

## Variables ambientales registradas por sensores remotos como indicadores de la distribución geográfica de *Triatoma infestans* (Heteroptera: Reduviidae)

DAVID E GORLA ✉

Centro Regional de Investig. Científ. y Transferencia Tecnológica (CRILAR), Anillaco, La Rioja, ARGENTINA

**RESUMEN.** El trabajo presenta un análisis de la distribución geográfica de *Triatoma infestans* basado en variables biofísicas calculadas a partir de datos obtenidos por satélites meteorológicos. Se consideran datos derivados de un Análisis Temporal de Fourier sobre imágenes mensuales que contienen un conjunto de 14 descriptores estadísticos del promedio, mínimo, máximo, variabilidad, amplitud y fase de la temperatura del aire, la temperatura de superficie, la radiación infrarroja media y el índice de vegetación calculados con datos colectados por el sensor AVHRR. La serie temporal considerada cubre el período 1982–2000 y el área abarca desde Panamá hasta Tierra del Fuego. El análisis de variables individuales mostró que la fase anual del índice de vegetación identifica correctamente el 80% de los sitios donde *Triatoma infestans* está presente y ausente. El análisis confirmó la importancia de la temperatura (especialmente la del aire) sobre la distribución de *Triatoma infestans*, ya que 7 descriptores estadísticos de la temperatura del aire (de un total de 14) identificaron correctamente >70% de los sitios donde *Triatoma infestans* está ausente y presente. Un Análisis Lineal Discriminante por Pasos mostró que una función con seis descriptores estadísticos (cuatro asociados con la temperatura del aire, uno con la radiación infrarroja media y uno con el índice de vegetación) identificaron correctamente el 90% de los sitios donde *Triatoma infestans* está presente y ausente. El análisis muestra la importante asociación entre las variables estudiadas a escala continental y la distribución de una especie doméstica y hematófaga. El enfoque de trabajo puede resultar de utilidad en estudios de escala geográfica amplia sobre la distribución de otras especies con una relación más directa y causal con la vegetación y el clima que *Triatoma infestans*. [Palabras clave: *Triatoma infestans*, enfermedad de Chagas, distribución geográfica, sensores remotos.]

**ABSTRACT. Remotely sensed environmental variables as indicators of *Triatoma infestans* (Heteroptera: Reduviidae) distribution:** The paper presents an analysis of the geographic distribution of *Triatoma infestans* based on biophysical variables recorded by meteorological satellites. The analysis includes data derived from a temporal Fourier Analysis of monthly images over a set of 14 derived statistical descriptors of the average, minimum, maximum, variability, amplitude and phase of the air temperature, land surface temperature, medium infrared radiation and a vegetation index remotely sensed by the AVHRR. Temporal series covered the period 1982–2000, and the area from Panamá to Tierra del Fuego. The analysis of individual variables showed that the annual phase of the normalised difference vegetation index identifies correctly 80% presence and absence sites of *Triatoma infestans*. The analysis confirms the importance of temperature (especially air temperature) for *Triatoma infestans* distribution, as 7 (out of 14) statistical descriptors associated with air temperature identified correctly >70% of presence and absence sites. A Stepwise Linear Discriminant Analysis showed that a function with six statistical descriptors (four of them associated with air temperature, one associated with medium infrared radiation, and one with the vegetation index) identified correctly 90% of presence and absence sites of *Triatoma infestans*. The analysis shows the important association between the environmental variables studied at the continental scale and the distribution of a strictly domestic and haematophagous species. The approach could be useful for large scale studies of the distribution of other species with more direct and causal relation with vegetation and climate than *Triatoma infestans*. [Keywords: *Triatoma infestans*, Chagas disease, geographic distribution, remote sensing.]

## INTRODUCCIÓN

La enfermedad de Chagas es la enfermedad parasítica más importante de las Américas. Su impacto social y económico supera el impacto combinado de otras enfermedades parasíticas tales como la malaria, la esquistosomiasis y la leishmaniasis. Existen actualmente entre 16 y 18 millones de personas infectadas y unas 100 millones consideradas bajo riesgo de contraer la enfermedad. La pérdida anual debido a la enfermedad de Chagas es de 2740000 DALYs ("disability-adjusted life years") y representa una pérdida económica en los países endémicos latinoamericanos de más de 6500 millones de dólares anuales (WHO 1997; Schofield & Dias 1999). Las iniciativas continentales (Cono Sur, Pacto Andino, Centro América y México) coordinadas por la Organización Mundial de la Salud para interrumpir la transmisión vectorial del Chagas en Latinoamérica mostraron grandes avances durante la última década, especialmente en el cono sur de América del Sur (Moncayo 1999). Sin embargo, aún en las regiones donde los programas de control son considerados exitosos (como en los países del cono sur), el objetivo de interrupción de la transmisión vectorial de la enfermedad aún está distante debido al progresivo desmantelamiento observado desde mediados de la década de 1990 en muchos de los servicios de control de vectores del Chagas, con la consecuente reinfestación de regiones en las que se había conseguido eliminar al principal insecto vector de la enfermedad, *Triatoma infestans*.

La distribución geográfica de los vectores de enfermedades humanas es un aspecto de reconocida importancia para la evaluación epidemiológica en los programas de control, siendo ampliamente conocida la influencia de factores ambientales sobre la distribución y abundancia de insectos. La temperatura mínima mensual promedio es un buen indicador de la distribución potencial y de la capacidad de crecimiento de *Triatoma infestans* en el territorio de la República Argentina (Gorla et al. 1997), analizadas sobre la base de un modelo de regulación de la abundancia poblacional de *Triatoma infestans* desarrollado por Gorla & Schofield (1989) y Gorla (1991, 1992). Gorla et al. (1997) analizaron la relación entre la temperatura y la tasa intrínseca de crecimiento natural ( $r$ ) de poblaciones de

*Triatoma infestans*, bajo el supuesto de que poblaciones de la especie con muy baja densidad según el valor de  $r$  sin estar afectadas por factores de regulación de la abundancia dependientes de la densidad. Aunque la temperatura no es la única variable climática que limita la capacidad de crecimiento de poblaciones de *Triatoma infestans* (y, por lo tanto, su distribución geográfica), la predicción del modelo de regresión coincide estrechamente con la distribución geográfica conocida de la especie.

La relación entre la distribución geográfica y la temperatura, humedad, precipitación y altura sobre el terreno fue mostrada para algunas especies de Triatominae (e.g., Carcavallo 1999; Ramsey et al. 2000). Los estudios de la relación entre la vegetación y la distribución de Triatominae no son numerosos y, en general, se refieren a relaciones cualitativas. Hay algunas relaciones obvias, como en el caso de la estrecha asociación entre algunas especies de los géneros *Rhodnius* y *Parabelminus* y ciertas especies de bromeliáceas o palmas (Lent & Wygodzinsky 1979). Excepto situaciones en que la especie de triatomo tiene un rango de distribución muy estrecho, la idea generalizada es que la vegetación tendría poca o ninguna asociación directa con insectos hematófagos, especialmente con aquellos que muestran una marcada domiciliación (sensu Schofield et al. 1999; por ejemplo, *Triatoma infestans*). Sin embargo, a menudo esa relación existe aunque más no sea por razones indirectas, como suele ocurrir con cierto tipo de vegetación que aloja a los hospedadores de sangre caliente de los cuales se alimentan los triatomos. Es el caso de *Triatoma guasayana*, que en Santiago del Estero tiene una alta preferencia por quimiles, una preferencia menor por chaguares y casi ninguna por huecos de árboles (Wisnivesky-Colli et al. 1997).

La vegetación es una variable que integra efectos de la temperatura, la precipitación y de propiedades edáficas; por ello, en algunas circunstancias puede resultar un indicador relacionado con variables que influyen directamente sobre procesos demográficos básicos (mortalidad, natalidad, dispersión). Lo que frecuentemente ocurre es que la distribución de una especie animal no está asociada con los clásicos patrones de distribución de comunidades vegetales, sino con indicadores de la

variabilidad espacial y temporal de la biomasa fotosintéticamente activa, como fuera mostrado para moscas tse-tsé (Rogers & Randolph 1993), para especies de Culicidae (Rogers et al. 2002) y para ácaros (Randolph & Rogers 1997). La vegetación puede caracterizarse sobre la base de sus patrones espaciales y temporales de cambio (Hay 2000). Investigaciones recientes mostraron el valor informativo contenido en series temporales del índice de vegetación calculado con datos de satélites meteorológicos para el desarrollo de sistemas de seguimiento de insectos plaga a escala regional (Gleiser et al. 1997; Grilli & Gorla 1997).

El objetivo de este trabajo es evaluar la capacidad que tienen algunas variables ambientales obtenidas con datos tomados por satélites meteorológicos para describir la distribución geográfica de *Triatoma infestans*, utilizando series temporales de imágenes satelitales en formato digital.

## MÉTODOS

### *Datos satelitales*

Se usaron imágenes producidas por el sensor AVHRR ("Advanced Very High Resolution Radiometer") que se encuentra a bordo de satélites meteorológicos de la serie NOAA ("National Oceanic and Atmospheric Administration"). El sensor AVHRR produce imágenes en cinco canales o bandas ("ventanas" del espectro electromagnético), que solas o combinadas permiten calcular diferentes propiedades biofísicas de la superficie terrestre. El satélite tiene una órbita polar que permite dos pasadas diarias (una diurna de norte a sur y otra vespertina de sur a norte) por un mismo punto de la superficie terrestre (resolución temporal). En su órbita, el sensor registra la energía que llega desde áreas de 1.1×1.1 km (resolución espacial en el zenit) de la superficie terrestre. La imagen es almacenada en formato digital como una matriz numérica, donde cada celda (pixel) contiene el valor de la energía que llega al sensor. Las imágenes utilizadas en este trabajo fueron obtenidas desde el servidor [ftp://daac.gsfc.nasa.gov/avhrr/global\\_8km/](ftp://daac.gsfc.nasa.gov/avhrr/global_8km/) y corresponden a productos mensuales (mosaicos compuestos a partir de imágenes diarias vespertinas), calculados con el método de máximo valor de composición (Holben 1986). Este método minimiza la contaminación de las

imágenes por nubes. La serie temporal de imágenes mensuales comprende el período 1982–2000. Debido al método de procesamiento de las imágenes originales del AVHRR, la resolución espacial de las imágenes disponibles es de 8×8 km.

Las variables biofísicas consideradas en este trabajo corresponden al índice de vegetación (*NDVI*, estrictamente un valor entre -1 y +1, aunque áreas con vegetación tienen generalmente un valor de *NDVI* > 0), temperatura de superficie (*LST*, en °C), temperatura del aire (*AT*, en °C) y radiación de infrarrojo medio (*MIR*, en equivalentes °C). El *NDVI* se calcula con datos de los canales 1 y 2, la *LST* con los canales 4 y 5, la *AT* con los canales 1, 2, 4 y 5, y la *MIR* es la lectura directa del canal 3 (Hay 2000), según las siguientes fórmulas:

$$NDVI = (Ch2 - Ch1) / (Ch1 + Ch2),$$

donde *Ch1* es el valor del canal 1 y *Ch2* el valor del canal 2 del AVHRR, y

$$LST = Ch4 + 3.33 (Ch5 - Ch4),$$

donde *Ch4* y *Ch5* son los valores de los canales 4 y 5 del AVHRR (Price 1984).

*AT* es una función de *LST* evaluada para valores de *NDVI* = 0.7, sobre la base del algoritmo TvX, cuyos detalles se pueden consultar en Goetz et al. (2000).

La ventana geográfica analizada (vértice superior izquierdo: 82°W, 15°N; vértice inferior derecho: 34°W, 57°S) abarca el continente sudamericano. Cada imagen tiene 960 columnas por 1440 filas. Contando el tamaño de cada imagen (una por mes), los cinco canales y 216 meses, el volumen de datos analizado es de 1.5 Gbyte. La gran cantidad de datos contenidos en series temporales largas de imágenes satelitales llevó a diferentes autores a desarrollar una variedad de métodos para reducir la cantidad de datos a analizar sin perder información. El Análisis de Componentes Principales estandarizado es uno de tales métodos, el cual permite calcular las contribuciones de la variación espacial y de la variación temporal en series largas de imágenes satelitales (Gorla 2001). Sin embargo, aún cuando este análisis resulta útil para reducir la cantidad de datos, usualmente resulta difícil extender sus resultados a otros tiempos y/o lugares debido a que la rotación de ejes de los componentes principales está determinado por el conjunto

original de datos observados (Rogers et al. 1996).

Un procedimiento alternativo para reducir datos, usando un Análisis de Fourier de series temporales, mostró resultados que fueron relacionados con procesos biológicos a escala regional y continental. El Análisis de Fourier divide la variabilidad de la serie temporal en componentes ortogonales con períodos iguales a 1, 1/2, 1/3, 1/n veces la duración de las  $n$  observaciones. Cada componente es esencialmente una función coseno, que queda definido por su amplitud y su fase (Bloomfield 2000). Si las observaciones son mensuales, el Análisis de Fourier divide a la serie en frecuencias equivalentes a períodos que van desde el de la serie completa (el más largo) hasta dos meses (el más corto). El conjunto total de componentes describe con exactitud la serie original, pero no todos los componentes tienen una misma ponderación. A semejanza de lo que ocurre con un Análisis de Componentes Principales, en el que los primeros componentes contienen la mayor parte de la variación, frecuentemente un conjunto de siete variables permite representar la serie sin pérdida significativa de información: el promedio de la serie y las amplitudes y fases de los tres primeros componentes que representan la variación en ciclos anuales, bianuales y trianuales (Rogers et al. 1996).

#### *Análisis de datos*

Para cada una de las cuatro variables biofísicas mencionadas arriba (*AT*, *LST*, *MIR* y *NDVI*) se usaron los datos promedio, mínimo y máximo de la variable para toda la serie, la amplitud ( $A_1$ ,  $A_2$ ,  $A_3$ ) y fase ( $P_1$ ,  $P_2$ ,  $P_3$ ) para cada uno de los tres primeros componentes de Fourier, la proporción de varianza explicada por cada uno de los primeros tres componentes de Fourier ( $D_1$ ,  $D_2$ ,  $D_3$ ), el porcentaje de varianza total capturada por los tres primeros componentes de Fourier (*DALL*) y la varianza de la serie completa (*VR*). El conjunto de datos analizados en este trabajo fue provisto por el TALA Group.

La hipótesis básica que guía este estudio es que a la escala del continente sudamericano, las características espaciales y temporales de las variables ambientales definen un perfil característico para cada sitio que permite iden-

tificar áreas donde la especie está presente y áreas donde la especie está ausente. Para el procedimiento que se analiza, el perfil ambiental definirá las condiciones de cada punto sobre el espacio (en este caso de tamaño  $8 \times 8$  km) representadas por un vector columna de 56 elementos.

La distribución geográfica de *Triatoma infestans* se conoce con bastante precisión, debido al esfuerzo acumulado de los trabajos de campo que los programas de control de la enfermedad de Chagas realizan en el Cono Sur sudamericano desde la década de 1940 (Schofield 1989). Los perfiles ambientales de los sitios donde la especie está presente o ausente se estimaron a partir de la selección aleatoria de puntos muestrales a partir de un mapa de la distribución geográfica de *Triatoma infestans* construido principalmente a partir de la información anteriormente mencionada (Schofield 1989). El número de puntos muestrales para áreas de presencia y ausencia fue proporcional al tamaño de cada una de estas áreas: 124 para áreas de presencia y 333 para áreas de ausencia.

El análisis se realizó en dos partes. En la primera se realizó un análisis univariado, con el objeto de evaluar la capacidad de las variables individuales para describir la distribución geográfica de *Triatoma infestans*. En la segunda parte se usó un Análisis Discriminante (lineal) por Pasos con la opción "backward", con el objeto de identificar las variables que mejor describen el conjunto de datos. La función derivada fue usada para clasificar áreas con ausencia y presencia de la especie dentro de América del Sur. La unidad de clasificación es un pixel de  $8 \times 8$  km y el número total de pixeles a clasificar sobre el continente sudamericano es de 627 616. La habilidad del procedimiento para describir la distribución observada fue medida como la proporción de clasificaciones correctas (número de presencias y ausencias identificadas correctamente / número total de pixeles), la proporción de falsos positivos y de falsos negativos (falsas clasificaciones sobre presencias y ausencias), la sensibilidad (proporción de clasificaciones sobre presencias correctas) y la especificidad (proporción de clasificaciones sobre ausencias correctas). Todas las operaciones que involucraron álgebra de imágenes digitales se realizaron con el programa Idrisi™ v. 2.0.

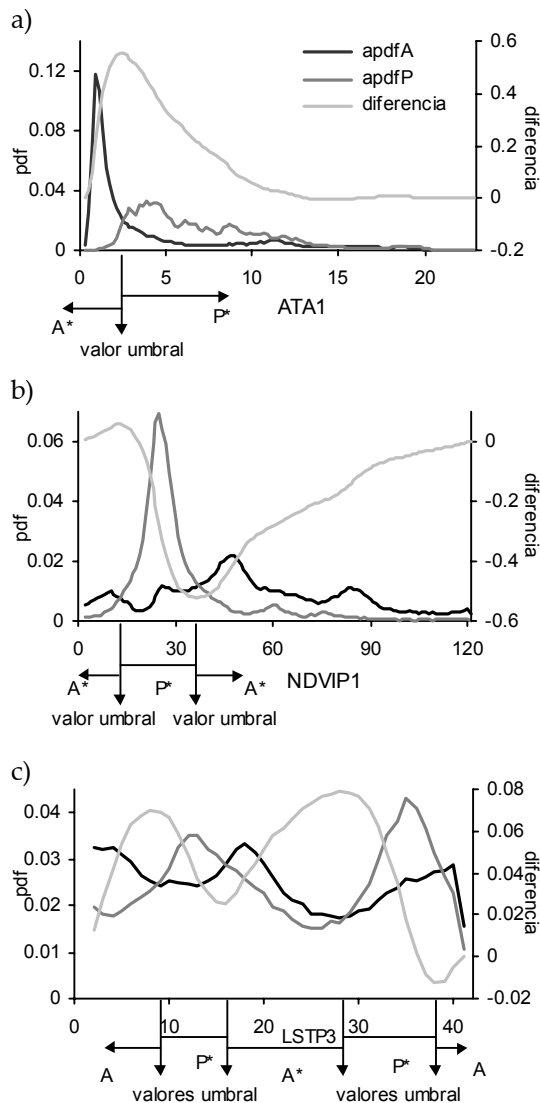
## RESULTADOS

### Análisis univariado

Para cada variable se estimó la función de densidad de probabilidad de valores asociados a la ausencia (pdfA) y presencia (pdfP) de la especie, sobre la base de la distribución en intervalos de clase de frecuencias relativas ( $h_i$ ) de los valores de las variables para áreas de ausencia y presencia. La frecuencia relativa se calculó como  $h_i = n_i/n$ , donde  $n_i$  es el número de píxeles en el  $i$ -ésimo intervalo de valores de la variable y  $n$  es el número total de píxeles dentro del área de ausencia o presencia. En general, pdfA y pdfP tienen rangos de valores de la variable superpuestos, pero en cada caso existe un valor umbral óptimo de la variable que separa valores de las distribuciones asociadas a áreas donde la especie está ausente y donde está presente. Una forma de encontrar el valor umbral óptimo es calcular el máximo de la diferencia de las funciones de densidad de probabilidad acumuladas para áreas de ausencia (apdfA) y presencia (apdfP) (derivado de Robinson et al. 1997).

Como ejemplos de los análisis realizados, se presentan en la Figura 1 los casos de las variables *ATA1* (amplitud anual de la temperatura del aire), fase del ciclo anual en el índice de vegetación (*NDVIP1*) y fase del ciclo trianual de la temperatura de superficie (*LSTP3*). Cada uno ilustra una situación particular de análisis, que se repite para las otras variables.

Los valores de *ATA1* en los sitios donde la especie está ausente son menores que aquellos donde la especie está presente (Figura 1a); la pdfA es leptokúrtica y con fuerte asimetría a la derecha, en tanto que la pdfP es platikúrtica con asimetría derecha. La diferencia entre las funciones de densidad de probabilidad acumuladas (apdfA - apdfP) tiene una distribución unimodal que muestra un máximo en 2.3, valor del umbral óptimo para separar ambas distribuciones. Valores de *ATA1* < 2.3 están asociados a sitios donde *Triatoma infestans* está ausente, valores de *ATA1* > 2.3 están asociados a sitios donde *Triatoma infestans* está presente. El valor 2.3 es un umbral óptimo de separación, pero no perfecto, ya que valores de *ATA1* < 2.3 están presentes en la pdfP y valores de *ATA1* > 2.3 están presentes en la pdfA. La identificación de estas diferentes situaciones permite calcular la pro-



**Figura 1.** Funciones de densidad de probabilidad (pdf) para áreas donde *Triatoma infestans* está presente (P\*) y ausente (A\*) para 3 de las 56 variables analizadas, junto con la función de densidad de probabilidad resultante de la diferencia entre las funciones acumuladas de ausencia (apdfA) y presencia (apdfP). Las variables mostradas son: a) la amplitud del ciclo anual de la temperatura del aire (*ATA1*), b) la fase del ciclo anual del índice de vegetación de diferencia normalizada (*NDVIP1*), y c) la fase del ciclo trianual de la temperatura de superficie (*LSTP3*).

**Figure 1.** Probability density functions (pdf) for areas where *Triatoma infestans* is present (P\*) or absent (A\*), for 3 of the 56 analysed variables, and the probability density function resulting from the difference between accumulated pdf of absence (apdfA) and presence (apdfP). Variables shown are: a) amplitude of the annual cycle of air temperature (*ATA1*), b) phase of the annual cycle of the normalised difference vegetation index (*NDVIP1*), and c) phase of the tri-annual cycle of surface temperature (*LSTP3*).

	AT				LST				MIR				NDVI				Correcto VD
	Presencia (rango)		Falso positivo	Falso negativo	Presencia (rango)		Falso positivo	Falso negativo	Presencia (rango)		Falso positivo	Falso negativo	Presencia (rango)		Falso positivo	Falso negativo	
	Correcto	Incorrecto			Correcto	Incorrecto			Correcto	Incorrecto			Correcto	Incorrecto			
Promedio	21.9–39.5	0.64	0.41	0.42	35.1–46.1	0.70	0.19	0.58	30.7–48.3	0.68	0.33	0.29	0.30–0.63	0.64	0.41	0.23	0.66
A1	>2.3	0.72	0.36	0.08	2.9–20.9	0.66	0.45	0.03	25–175	0.66	0.47	0.02	70–202	0.66	0.30	0.42	0.68
A2	10–25	0.62	0.42	0.26	12–30	0.66	0.35	0.30	11–26	0.62	0.38	0.36	25–53	0.56	0.45	0.39	0.62
A3	3–11	0.54	0.54	0.26	5–14	0.58	0.45	0.34	4–11	0.62	0.37	0.41	8–22	0.52	0.51	0.41	0.56
P1	107 < P1 < 5	0.71	0.34	0.14	107 < P1 < 3	0.73	0.28	0.27	107 < P1 < 2	0.72	0.24	0.38	13–37	0.80	0.19	0.23	0.74
P2	25–37	0.59	0.45	0.41	24–30	0.58	0.44	0.39	23–49	0.57	0.48	0.29	38–53	0.59	0.43	0.36	0.58
P3	20–37	0.59	0.38	0.50	8–16; 28–38	0.57	0.43	0.43	5–15; 30–41	0.57	0.46	0.34	16–36	0.52	0.59	0.20	0.56
D1	31–79	0.73	0.30	0.18	28–88	0.66	0.42	0.14	20–84	0.66	0.46	0.08	24–78	0.67	0.33	0.33	0.68
D2	32 < D2 < 5	0.53	0.54	0.27	40 < D2 < 5	0.58	0.42	0.42	41 < D2 < 4	0.56	0.48	0.34	4–12	0.60	0.36	0.50	0.57
D3	5 < D3 < 3	0.42	0.69	0.18	2–10	0.62	0.27	0.78	2–4	0.68	0.10	0.88	2–12	0.50	0.43	0.70	0.54
DALL	35–79	0.73	0.32	0.15	33–88	0.65	0.46	0.08	25–86	0.72	0.36	0.06	21–87	0.56	0.54	0.16	0.67
mín	10.6–29.5	0.71	0.31	0.23	13.1–28.5	0.74	0.19	0.43	13.3–26.3	0.74	0.19	0.52	0.21–0.55	0.64	0.41	0.22	0.71
máx	35.1–43.9	0.70	0.25	0.44	37.3–57.1	0.72	0.33	0.17	32.9–50.5	0.66	0.44	0.08	0.33–0.69	0.63	0.41	0.28	0.68
VR	13–142	0.76	0.28	0.14	19–274	0.68	0.43	0.05	17–193	0.65	0.46	0.05	97–307	0.61	0.38	0.52	0.67
Correcto VRS		0.65				0.66				0.65				0.61			

**Tabla 1.** Capacidad descriptiva de cada variable biofísica incluida en el análisis univariado (AT, LST, MIR y NDVI; ver descripción en el texto). Para cada una se incluyen los valores umbrales óptimos para separar las distribuciones de valores asociados a sitios de presencia y ausencia, la proporción de falsos positivos y negativos y la proporción de sitios correctamente identificados como de presencia o ausencia. Correcto VRS representa la proporción de sitios correctamente identificados por cada variable biofísica considerada. Correcto VD representa la proporción de sitios correctamente identificados (presencia o ausencia) por cada parámetro derivado del Análisis de Fourier.

**Table 1.** Descriptive capacity of each biophysical variable included in the univariate analysis (AT, LST, MIR y NDVI; see text for a description). For each variable, the optimum threshold values to discriminate the value distributions associated to presence and absence sites, the proportion of false positives and false negatives, and the proportion of sites correctly identified are included. Correcto VRS represents the proportion of sites correctly identified by each biophysical variable considered. Correcto VD represents the proportion of sites correctly identified (presence or absence) by each variable derived from the Fourier analysis.

babilidad de asociar correcta e incorrectamente valores de *ATA1* a áreas de presencias y ausencias.

Para el caso de *NDVIP1* (Figura 1b) se observa que valores bajos y altos están asociados a sitios donde *Triatoma infestans* está ausente, en tanto que valores intermedios están asociados a sitios donde la especie está presente. La diferencia entre las apdf muestra un máximo en 13 y un mínimo en 37, indicando que la especie está preferentemente ubicada dentro del rango  $13 < \text{NDVIP1} < 37$ .

Para el caso de *LSTP3* (Figura 1c) se observa que la diferencia entre las apdf tiene una distribución bimodal, existiendo 2 rangos de valores de la variable asociados a la presencia de *Triatoma infestans* (8–16 y 28–38).

Debido a que la diferencia entre las apdf se calcula como  $\text{apdfA} - \text{apdfP}$ , cuando la primera es menor que la segunda el resultado es negativo. Un valor máximo de esta diferencia representa el límite inferior y un valor mínimo representa el límite superior de valores de la variable asociados a sitios donde *Triatoma infestans* está presente. La diferencia entre el máximo y mínimo de  $\text{apdfA} - \text{apdfP}$  es una indicación del grado de discriminación esperable entre las funciones de densidad de probabilidad de los valores. La diferencia entre las apdf correspondientes a *NDVIP1* es la que muestra mayor rango (0.578), en tanto que la correspondiente a la variable *D3* muestra el menor rango ( $< 0.14$ ).

La Tabla 1 contiene las proporciones de áreas correctamente identificadas como ausencia o presencia, falsos positivos (áreas erróneamente identificadas como sitios donde la especie está presente) y falsos negativos (áreas erróneamente identificadas como sitios donde la especie está ausente). Ninguna de las cuatro variables derivadas de los datos provistos por el sensor AVHRR (*AT*, *LST*, *MIR*, *NDVI*) se destaca particularmente por su capacidad para describir correctamente la distribución de *Triatoma infestans*; en general, cada una de ellas clasifica correctamente la presencia y la ausencia en el 60–65% de los sitios. Sin embargo, los parámetros derivados del Análisis de Fourier permiten estimar correctamente presencia y ausencia hasta en el 80% de los casos (*NDVIP1*).

En la valoración de la capacidad descriptiva de cada variable debe considerarse que cada

punto del espacio tiene una probabilidad a priori de pertenecer a áreas con presencia o ausencia de *Triatoma infestans* igual a 0.5 (i.e., una variable que identifique correctamente el 70% de los sitios de presencia o ausencia representa un aumento de la capacidad predictiva del 40% por encima del mero azar). En la Tabla 1 se resaltan las variables derivadas que mostraron valores  $\geq 0.70$  de clasificaciones correctas. Hay siete variables derivadas de la temperatura del aire (*AT*) que mostraron clasificaciones correctas por encima del 70%, cuatro asociadas a *LST*, tres a *MIR* y una al *NDVI*. Considerando cada variable derivada (e.g., promedio, *A1*) sobre el conjunto de las variables registradas por el AVHRR, se observa que solo los valores correspondientes a la fase del ciclo anual (*P1*) y los mínimos consiguen identificar correctamente el 70% o más de los sitios.

En general, las variables registradas por sensores remotos sobre un mismo punto del espacio muestran una alta correlación (Rogers 2000). Analizando las 15 variables que mostraron capacidad para identificar correctamente  $\geq 70\%$  de sitios de ausencia o presencia de *Triatoma infestans*, se observa que el grado de correlación existente entre ellas es muy bajo. De las 105 correlaciones posibles, sólo 9 ( $< 9\%$ ) muestran un Coeficiente de Correlación Lineal  $\geq 0.70$ . La variable con mayor capacidad de identificar correctamente sitios de presencia y ausencia (*NDVIP1*) muestra una correlación muy baja ( $r < |0.26|$ ) con cada una de las restantes 14 variables.

#### *Análisis multivariado*

De las 56 variables introducidas en el Análisis Discriminante, 6 permiten clasificar los sitios muestrales en las categorías de ausencia y presencia. En la Tabla 2 se incluyen las variables, sus coeficientes estandarizados y no estandarizados y los valores de los coeficientes de las funciones de clasificación. De las seis variables, cuatro corresponden a temperatura del aire (*AT*) y las dos restantes a los máximos de la radiación infrarroja (*MIR*) y del *NDVI*. La función estimada clasifica correctamente 89.7% de los 457 puntos muestrales usados en el análisis, con una sensibilidad de 91.9% y una especificidad de 83.9%.

La inspección de los coeficientes estandarizados de la función (indicativos de las ponde-

**Tabla 2.** Coeficientes estandarizados y no estandarizados de las funciones discriminantes y coeficientes de las funciones de clasificación para sitios con ausencia y presencia de *Triatoma infestans*.

**Table 2.** Standardized and unstandardized coefficients for the discriminant functions and coefficients of the classification functions for absence and presence sites of *Triatoma infestans*.

Variable	Coeficientes estandarizados	Coeficientes no estandarizados	Función para ausencia	Función para presencia
ATDALL	-7.627	-0.078	1.90	2.09
Promedio AT	5.759	0.086	0.77	0.62
ATA1	-3.007	-0.069	-1.24	-1.16
ATD3	0.671	0.032	-0.67	-0.84
máx MIR	0.360	0.008	-0.38	-0.37
máx NDVI	-0.350	-0.014	-0.86	-0.89
Constante		-25.882	-1833.88	-1895.36

raciones de cada variable en el valor final de la función discriminante) muestra que las variables con mayor importancia se relacionan con la temperatura del aire, siendo la variabilidad conjunta de los ciclos anual, bianual y trianual de la temperatura del aire (ATDALL) la que más peso tiene en la discriminación de áreas de presencia/ausencia, seguida en importancia por la temperatura promedio de la serie (Promedio AT) y luego por la amplitud del ciclo anual (ATA1).

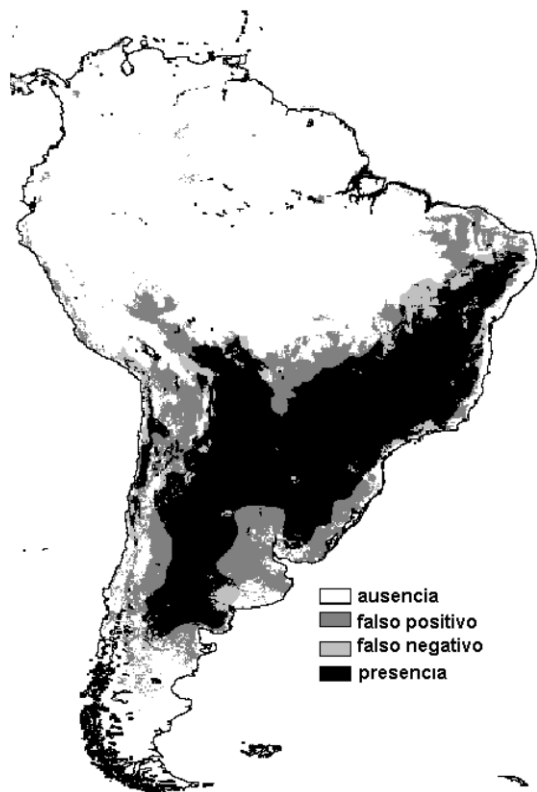
En este análisis, además de los valores muestrales contamos con lo que consideramos la "verdadera" distribución de *Triatoma infestans* (reportada por Schofield 1989). Usando las funciones de clasificación para cada una de las categorías (presencia o ausencia) resultantes del Análisis Discriminante (Tabla 2) y los datos de cada una de las seis variables identificadas como de mayor importancia para la discriminación, se calculó la categoría de pertenencia para cada uno de los píxeles para los que existen datos sobre América del Sur ( $n = 627\,616$ ). De los 176015 píxeles que ocupa la distribución geográfica de *Triatoma infestans*, el análisis clasificó correctamente 157536 (sensibilidad = 0.895) y de los 451649 en los que la especie está ausente clasificó correctamente 374963 (especificidad = 0.83). En la Figura 2 se muestra una síntesis de los resultados de la clasificación, con la identificación de las áreas con presencia y ausencia de la especie, así como de los falsos positivos y falsos negativos resultantes del análisis.

## DISCUSIÓN

*Triatoma infestans* es una especie estrechamente asociada a estructuras domésticas y peridomésticas. El éxito de la especie es función de su habilidad para explotar eficientemente los recursos disponibles en ambientes humanos. A pesar de ello, la capacidad de las variables ambientales para describir la distribución geográfica de la especie a escala regional mostrada por los resultados del análisis presentado, indican que, a esta escala, existe un perfil ambiental que resulta adecuado para la existencia de la especie, que no necesariamente está asociado con la disponibilidad de viviendas humanas. Si bien las condiciones microambientales en las estructuras domésticas y peridomésticas modifican las condiciones macroambientales percibidas por el insecto (i.e., amortiguando los extremos de temperatura), éstas tienen una influencia suficientemente fuerte como para determinar el conjunto de condiciones que definen dónde la especie puede estar y dónde no. Para el caso analizado, la importancia de los resultados no es la postdicción de la distribución geográfica de *Triatoma infestans* (conocida con cierto detalle desde mucho tiempo atrás, y fuertemente modificada debido a los programas de control de vectores de los países del cono sur), sino la evidencia de que un conjunto de variables ambientales registradas en formato digital por sensores remotos tienen la capacidad de describir la distribución geográfica de



una especie que mantiene relaciones tan estrechas con estructuras humanas. Es razonable pensar que, para el caso de otras especies más estrechamente asociadas a variables ambientales naturales (e.g., vegetación, clima), el enfoque produzca iguales o mejores resultados.



**Figura 2.** Áreas de presencia y ausencia de *Triatoma infestans* correctamente identificadas y áreas incorrectamente identificadas (falsos positivos y falsos negativos), derivadas de la clasificación basada en las 6 variables biofísicas de mayor peso en el Análisis Discriminante por Pasos de 56 estadísticos estimados luego de un Análisis Temporal de Fourier de variables calculadas con datos del AVHRR. El mapa base con el que se contrasta la distribución estimada es el reportado por Schofield (1989) a partir de datos de terreno.

**Figure 2.** Presence-absence areas of *Triatoma infestans* correctly identified, and areas incorrectly identified (false positive and false negative), derived from the classification based on the 6 biophysical variables with the highest weight in the Stepwise Discriminant Analysis of 56 statistics estimated after a Temporal Fourier Analysis of remotely sensed variables calculated from AVHRR data. The base map that is contrasted with the estimated distribution is the one reported by Schofield (1989) from field data.

El análisis univariado mostró que la fase del ciclo anual (P1), tanto para la temperatura del aire (AT), la temperatura de superficie (LST), la radiación infrarroja media (MIR) como para el índice de vegetación (NDVI), aparece como el parámetro derivado del Análisis Temporal de Fourier que mayor capacidad descriptiva tiene sobre la distribución de *Triatoma infestans* (> 70%), en tanto que 7 de las 14 variables derivadas del Análisis de Fourier con valores de la temperatura del aire (AT) muestran una capacidad de descripción superior al 70% para identificar correctamente sitios de presencia y ausencia de *Triatoma infestans*. El resultado del análisis remarca la importancia de la temperatura en la distribución de *Triatoma infestans*, pero llama la atención el peso que parece tener la fase del ciclo anual, característica que hasta ahora no había sido considerada en los análisis de la distribución de la especie.

La fase del ciclo anual del NDVI (NDVI<sub>P1</sub>) fue la variable que mostró la mayor capacidad para describir correctamente la distribución de *Triatoma infestans* (80% de aciertos). Llama la atención tal capacidad predictiva de las propiedades espacio-temporales de la vegetación para una especie de insecto hematófago y estrechamente asociado a estructuras domiciliarias. Aún cuando no existe una relación causal directa, la vegetación constituye una variable sintética que refleja cambios en la temperatura, la disponibilidad de agua y propiedades edáficas que afectan tanto tasas demográficas básicas de *Triatoma infestans* como la disponibilidad de refugios para sus hospedadores (viviendas humanas y estructuras peridomiciliarias). La relación estadística entre vegetación y distribución de *Triatoma infestans* es evidente en los resultados presentados, pero la identificación de las relaciones causales dentro de esta relación estadística requerirá de estudios experimentales que superen las limitaciones de estudios observacionales como el presente. Las variables biofísicas estimadas a partir de datos del AVHRR están fuertemente correlacionadas, pero el contenido informativo de la fase anual del NDVI muestra características diferenciales, a juzgar por su baja correlación con cualquiera de las otras variables, al menos cuando se considera la capacidad de variables individuales para describir la distribución de *Triatoma infestans*.

De las 56 variables analizadas para evaluar su capacidad de descripción de la distribución de *Triatoma infestans*, el 70% de los casos mostró que el error de falsos positivos (predicción de presencia errónea) es mayor que el de falsos negativos (predicción de ausencia errónea). Un falso negativo es una indicación unívoca de error en la predicción (ya que la especie se encontró allí donde el análisis indica que no debería estar). Un falso positivo no es una indicación unívoca de error, ya que la especie podría estar presente donde el análisis lo indica, pero no haber sido reportada, errores frecuentes en la construcción del tipo de mapa presentado por Schofield (1989); si esto fuera cierto, el nivel de acierto en las predicciones sobre la distribución sería mayor. Si bien la distribución geográfica de *Triatoma infestans* es bien conocida, es posible que hacia los márgenes de la distribución de la especie existan (o hayan existido) áreas donde la especie está (o estuvo) presente.

El análisis multivariado mostró que un conjunto de seis variables calculadas con los datos del AVHRR es capaz de identificar correctamente casi el 90% de los sitios de ausencia y presencia de *Triatoma infestans*. Entre las seis variables, cuatro corresponden a variables derivadas de la temperatura del aire (AT) en el Análisis de Fourier, remarcando la importancia que tiene la temperatura ambiente sobre la distribución de *Triatoma infestans*. Esto coincide con lo encontrado por otros autores que realizaron estudios experimentales tanto en laboratorio como en el campo (e.g., Gorla 1992). De acuerdo al análisis realizado, entre las variables derivadas de AT, la distribución de *Triatoma infestans* queda definida con mayor peso por la variabilidad de los ciclos anual, bianual o trianual, la temperatura promedio y la amplitud del ciclo anual.

La información producida por sensores remotos es una poderosa herramienta de apoyo que incrementa nuestra habilidad para investigar fenómenos a gran escala sobre la distribución de especies de insectos vectores de enfermedades humanas, imposibles de obtener con estudios de terreno debido a los costos operacionales que ello acarrearía. En muchos casos, el enfoque de trabajo permite sentar las bases para diseñar programas de seguimiento de plagas sobre grandes áreas, gracias a su

capacidad para producir caracterizaciones espaciales del ambiente y, especialmente, de las variaciones temporales. En esta línea de trabajo, los trabajos realizados en Argentina sobre *Aedes albifasciatus* (Gleiser et al. 1997) y *Delphacodes kuscheli* (Grilli & Gorla 1997) muestran la versatilidad del método.

La descripción estadística de la distribución geográfica de *Triatoma infestans* (o de cualquier otra especie) no es en modo alguno sustituto para la búsqueda de conocimiento biológico que causa el fenómeno observado. Tal búsqueda debe basarse en el estudio de procesos demográficos básicos; sin embargo, la información sobre tales procesos no abunda, dejando por el momento el enfoque estadístico como el único disponible para realizar afirmaciones a la escala de análisis presentada en este trabajo. Los resultados informados en este artículo muestran que aún con un conocimiento parcial, la distribución geográfica de la especie puede estimarse con razonable precisión si se cuenta con información sobre variables ambientales que cubra todo el rango de distribución de la especie. En este sentido, la información proporcionada por variables registradas por sensores remotos, evaluada con herramientas analíticas adecuadas, constituye una base de información de bajo costo insustituible para abordar estudios a escalas espaciales por encima de la escala del paisaje.

## AGRADECIMIENTOS

Este trabajo recibió apoyos parciales de la red Europeo Latinoamericana para el Estudio de la Genética y Control de Triatominae (ECLAT) y de CONICET. Se agradece al Dr. David Rogers, coordinador del Tala Research Group del Zoology Department (Oxford University, Reino Unido) por la provisión de las imágenes con los análisis temporales de Fourier para las variables ambientales usadas en los análisis. S. Catalá, M. Grilli, J. Rabinovich, J. Paruelo, W. Hasperué, P. Guerscham e I. Velasco ayudaron a mejorar versiones iniciales de este manuscrito.

## BIBLIOGRAFÍA

- BLOOMFIELD, P. 2000. *Fourier analysis of time series*. 2<sup>da</sup> edn. John Wiley & Sons. 261 pp.
- CARCAVALLO, R. 1999. Climatic factors related to Chagas disease transmission. *Mem. I. Oswaldo Cruz* 94:367-369.

- GLEISER, RM; DE GORLA & FF LUDUEÑA ALMEIDA. 1997. Monitoring *Aedes (Ochlerotatus) albifasciatus* (Macquart 1838) (Diptera: Culicidae) abundance in the southern region of Mar Chiquita Lake, with the aid of remote sensing. *Ann. Trop. Med. Parasit.* **91**:917–926.
- GOETZ, SJ; SD PRINCE & J SMALL. 2000. Advances in satellite remote sensing of environmental variables for epidemiological applications. *Adv. Parasit.* **47**:289–307.
- GORLA, DE. 1991. Recovery of *Triatoma infestans* populations after insecticide application: an experimental field study. *Med. Vet. Entomol.* **5**:311–324.
- GORLA, DE. 1992. Population dynamics and control of *Triatoma infestans*. *Med. Vet. Entomol.* **6**:91–97.
- GORLA, DE. 2001. Análisis de la distribución geográfica de Triatominae a escala continental en base a información de variables ambientales. *UNLAR Ciencia* **2**:2–8.
- GORLA, DE; SS CATALÁ & MP GRILLI. 1997. Efecto de la temperatura sobre la distribución de *Triatoma infestans* y el riesgo de transmisión vectorial de la enfermedad de Chagas en Argentina. *Acta Toxicológica Argentina* **5**:36–39.
- GORLA, DE & CJ SCHOFIELD. 1989. Population dynamics of *Triatoma infestans* under natural climatic conditions in the Argentine Chaco. *Med. Vet. Entomol.* **3**:179–194.
- GRILLI, MP & DE GORLA. 1997. The spatio-temporal pattern of *Delphacodes kuscheli* (Homoptera: Delphacidae) abundance in central Argentina. *B. Entomol. Res.* **87**:45–53.
- HAY, S. 2000. An overview of remote sensing and geodesy for epidemiology and public health applications. *Adv. Parasit.* **47**:1–35.
- HOLBEN, BN. 1986. Characteristics of the maximum-value composite images from temporal AVHRR data. *Int. J. Remote Sens.* **7**:1417–1434.
- LENT, H & P WYGODZINSKY. 1979. Revision of the Triatominae (Hemiptera, Reduviidae) and their significance as vectors of Chagas' disease. *B. Am. Mus. Nat. Hist.* **163**:123–520.
- MONCAYO, A. 1999. Progress towards the interruption of Chagas disease. *Mem. I. Oswaldo Cruz* **94**:401–404.
- PRICE, JC. 1984. Land surface temperature measurements from the split window channels of the NOAA 7 advanced very high resolution radiometer. *J. Geophys. Res.* **89**:7231–7237.
- RAMSEY, J; R ORDÓÑEZ; A CRUZ-CELIS; AL ALVERA; V CHAVEZ ET AL. 2000. Distribution of domestic Triatominae and stratification of disease transmission in Oaxaca, Mexico. *Med. Vet. Entomol.* **14**:19–30.
- RANDOLPH, SE & DJ ROGERS. 1997. A generic population model for the Africa tick *Rhipicephalus appendiculatus*. *Parasitology* **115**:265–279.
- ROBINSON, T; D ROGERS & B WILLIAMS. 1997. Univariate analysis of tse tse habitat in the common fly belt of Southern Africa using climate and remotely sensed vegetation data. *Med. Vet. Entomol.* **11**:223–234.
- ROGERS, DJ. 2000. Satellites, space, time and the African trypanosomiasis. *Adv. Parasit.* **47**:129–171.
- ROGERS, DJ; SI HAY & MJ PACKER. 1996. Predicting the distribution of tse tse flies in West Africa using temporal Fourier processed meteorological satellite data. *Ann. Trop. Med. Parasit.* **90**:225–241.
- ROGERS, DJ & SE RANDOLPH. 1993. Distribution of tse tse and ticks in Africa: past, present and future. *Parasitol. Today* **9**:266–271.
- ROGERS, DJ; SE RANDOLPH; RW SNOW & SI HAY. 2002. Satellite imagery in the study and forecast of malaria. *Nature* **415**:710–715.
- SCHOFIELD, CJ. 1989. The American trypanosomiasis. Pp. 81–92 in: WHO (ed). *Geographical distribution of arthropod-borne disease and their principal vectors*. WHO/VBC/89.967. 134 pp.
- SCHOFIELD, CJ & JC DIAS. 1999. The Southern Cone programme against Chagas disease. *Adv. Parasitol.* **42**:1–25.
- SCHOFIELD, CJ; L DIOTAIUTI & JP DUJARDIN. 1999. The process of domestication in Triatominae. *Mem. I. Oswaldo Cruz* **94**(Suppl. 1):375–378.
- WHO. 1997. Chagas disease. Interruption of transmission. *Weekly Epidemiological Record* **72**:1–5.
- WISNIVESKY-COLLI, C; NJ SCHWEIGMANN; S PIETROKOVSKY; V BOTTAZZI & JE RABINOVICH. 1997. Spatial distribution of *Triatoma guasayana* (Hemiptera: Reduviidae) in hardwood forest biotopes in Santiago del Estero, Argentina. *J. Med. Entomol.* **34**:102–109.