

# Desarrollo de un Modelo de Predicción de Riesgo de Fiebre del Dengue en León, Nicaragua

Informe técnico final

*Dr. Peter C. Esselman, Michigan State University, East Lansing, MI, USA;*

*pesselman@hotmail.com*

*10 de Agosto de 2009*

## Introducción

Uno de los objetivos principales de los Objetivos de Desarrollo del Milenio (ODMs) es el detener y revertir la incidencia de las principales enfermedades para el 2015. Una de estas enfermedades es la fiebre del dengue. El dengue es una enfermedad viral de gran importancia epidemiológica causada por la picada de un mosquito que infecta casi un millón de personas por año en América Latina ([www.paho.org](http://www.paho.org)). El problema del dengue es tan severo que en el 2001, la Organización Panamericana de la Salud (OPS) aprobó una declaración política (Resolución CD43.R4) para “reducir la morbilidad, mortalidad y la carga social y económica causada por las ocurrencias súbitas y epidemias de dengue”. El dengue es transmitido por el mosquito *Aedes aegypti*, el cual pica durante el día, se multiplica en charcos y otros sitios con agua estancada y prospera en ambientes semi-urbanos. La transmisión ocurre cuando un mosquito hembra se alimenta de un individuo afectado y pasa el virus a otros individuos en alimentaciones de sangre subsecuentes. Para detener y revertir la incidencia de la fiebre del dengue y lograr el objetivo de los ODMs, se necesitan intervenciones altamente enfocadas para controlar las poblaciones de mosquitos, minimizar la transmisión, y proteger a individuos susceptibles en forma eficaz.

Los modelos de predicción de enfermedades tienen un gran potencial de ayudar con el desarrollo de intervenciones para reducir los impactos del dengue. Los modelos de predicción utilizan datos sobre tasas de infección humana y los factores que afectan estas tasas para predecir los lugares y temporadas con altos niveles de riesgo de ocurrencias súbitas en el futuro. En el caso del dengue, estudios científicos previos han mostrado vínculos entre la enfermedad y condiciones climáticas, ecológicas, socioeconómicas y demográficas (Koopman et al. 1991). Estas fuerzas conductoras y correlaciones a las ocurrencias súbitas de dengue muestran variaciones en el tiempo (por ejemplo, condiciones climáticas) y en el espacio (por ejemplo, factores ecológicos y socioeconómicos), y por lo tanto requieren modelos que incorporen tanto patrones espaciales como temporales en sus predicciones de riesgo de la enfermedad.

Este informe presenta un prototipo de un modelo de predicción espacial-temporal que utiliza condiciones climáticas, ecológicas, demográficas y socioeconómicas para predecir el riesgo de dengue nueve semanas en el futuro de acuerdo a condiciones climáticas recientes dadas. En el centro de este sistema se encuentran modelos predictivos empíricos que fueron calibrados y probados con un conjunto de datos históricos de 10 años sobre los sitios donde ocurrieron casos de dengue en la ciudad de León, en Nicaragua. Más abajo detallo los conjuntos de datos, los métodos utilizados, y los resultados de este esfuerzo para desarrollar y validar un sistema de predicción prototipo.

## Métodos

### *Descripción del sitio*

La ciudad de León se localiza en la ladera del Pacífico de Nicaragua a unos 70 km al noroeste de Managua y a 17 km del Océano Pacífico (Figura 1). Esta histórica ciudad colonial española

con 175,000 residentes tiene unos 100 barrios que cubren un área de 15 km<sup>2</sup>. Las densidades de población en estos barrios van desde 800 personas/km<sup>2</sup> hasta más de 20,000 personas/km<sup>2</sup> en una variedad de condiciones de vivienda. Un promedio del 5% de las casas en cada barrio no cuentan con servicios de desechos higiénicos (escusado o letrina), 38% de las casas no cuentan con servicio de recolección de basura, 15% no cuentan con servicio de agua potable ni con aljibes, y el 26% de las casas existen en pobreza extrema (INIDE 2008). El clima de León es altamente cíclico con una estación seca fuerte entre enero y mayo y una estación de lluvias fuertes de enero a diciembre. Las temperaturas anuales varían entre 20 y 40°C (promedio 1998 – 2008 = 27.5°C) y la precipitación total anual dio un promedio de 1744 mm por año entre 1998 y 2008. La ciudad se encuentra rodeada por campos agrícolas, un terreno suavemente inclinado, y dos pequeños ríos que fluyen a través de la ciudad, descargando eventualmente en el Pacífico.

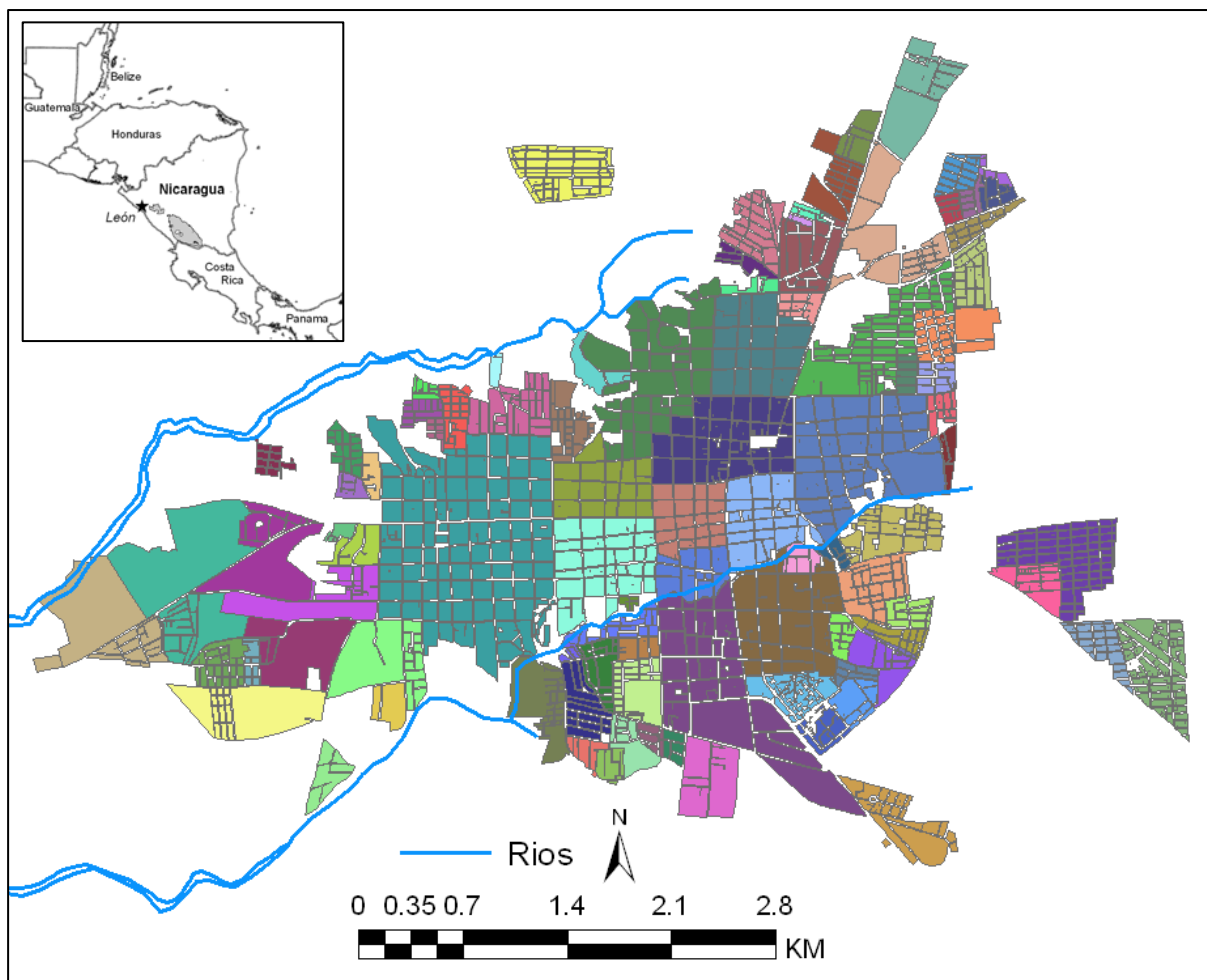


Figura 1. Mapa localizador mostrando donde se asienta León en la parte oriental de Nicaragua, y la disposición de la ciudad en relación a los ríos. Los colores diferentes representan los barrios de la ciudad.

### *Perspectiva del enfoque*

Se seleccionó un enfoque de modelo que utiliza datos sobre las fechas y sitios de casos de dengue confirmados junto con datos de ecosistemas, clima, y características poblacionales y socioeconómicas para predecir la probabilidad de que un caso de dengue ocurra en una localidad dada en una fecha dada. Una suposición central que nos permite predecir el riesgo de dengue en el futuro es que existe un *lapso de tiempo* constante entre el comienzo de las condiciones climáticas que son favorables para los mosquitos y la aparición de casos de dengue dentro de la población humana (Depradine y Lovell 2004). Mientras que las influencias climáticas cambian a través del tiempo, una variedad de otros factores que controlan los sitios de anidamiento de los mosquitos, las tasas de transmisión, y el contagio humano cambian en el espacio. Por ejemplo, la topografía influencia la tendencia del agua de acumularse en charcos y así puede ejercer influencia en la abundancia de sitios de anidamiento para mosquitos en una localidad. Al utilizar datos sobre los factores que pueden afectar la incidencia de dengue y correlaciones del dengue en el espacio y tiempo con las localidades de casos conocidos de dengue en un periodo de 10 años en León, esta investigación trata de hacer predicciones sobre cuando y donde es más probable que ocurran casos de dengue dentro de los límites de la ciudad.

Las factores que afectan la incidencia de dengue fueron identificadas a través de una revisión de la literatura científica y fueron comparadas con conjuntos de datos geoespaciales que pudieron ser mapeados y manipulados con programas de sistemas de información geográfica (SIG). Se acumularon diez años de información climática y se utilizaron en conjunto con datos de casos de dengue/semana para identificar el tiempo en semanas entre las condiciones climáticas apropiadas y una alta incidencia de casos de dengue. Durante este lapso de tiempo, se presume que el mosquito del dengue cumple su ciclo de vida, y pica a una persona infectada e ingiere sangre que contiene el virus del dengue. A continuación, el virus se replica durante un período de incubación dentro del mosquito, el mosquito pica a una persona susceptible y le transmite el virus también a ésta, y finalmente, el virus se replica en la segunda persona y produce síntomas (Figura 2). Este lapso de tiempo es utilizado en el proceso de desarrollo del modelo para asociar cada caso de dengue con las condiciones climáticas que se presume iniciaron la cadena de eventos que llevó a la ocurrencia de ese caso. Este enfoque de capacitar modelos a condiciones climáticas al principio del lapso de tiempo tiene una ventaja clave para la predicción en que, una vez que los modelos han sido capacitados sobre datos del pasado, el riesgo puede ser proyectado hacia el futuro al final del lapso de tiempo utilizando condiciones climáticas recientes.

Este enfoque específico de modelos de predicción utilizado se llama el método de entropía máxima, o Maxent. Maxent utiliza datos de ocurrencias (por ejemplo, puntos de dengue) y condiciones ambientales asociadas con esos datos para calcular la distribución más uniforme (con "entropía máxima") a través del área de estudio, dada la restricción de que el valor esperado de cada variable ambiental de predicción bajo esta distribución calculada sea igual a su promedio empírico para el conjunto de registros de presencia de la especie (Phillips et al. 2006). Como se mencionó anteriormente, los insumos para Maxent son mapas de condiciones sociales y ambientales que tienen influencia sobre la ocurrencia de dengue a través del área de estudio, así como información sobre las localidades en donde han ocurrido casos de dengue. Los resultados del modelo Maxent son una superficie continua de valores de probabilidad entre 0 y 1, en donde los valores más altos indican una probabilidad mayor de que un caso de dengue ocurrirá. Los lectores con un interés más profundo en los detalles matemáticos o de procedimiento sobre el proceso de predicción de Maxent pueden consultar varios artículos excelentes sobre estos temas (Phillips et al. 2006, Elith et al. 2006, Hernández et al. 2006).

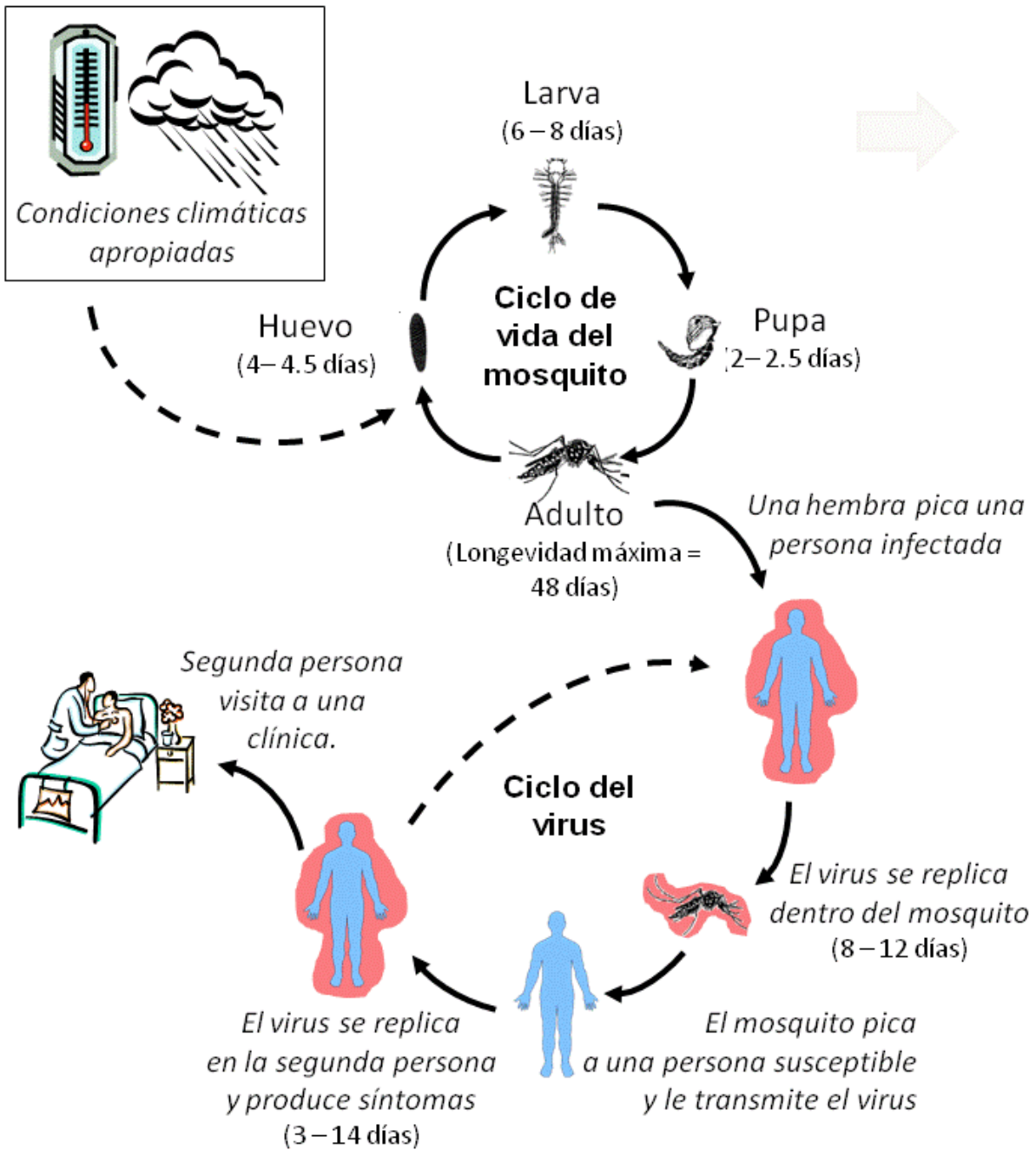


Figura 2. Los eventos que pasan durante el lapso de tiempo entre condiciones climáticas apropiadas y una visita de una persona con la fiebre del dengue a una clínica. Durante este lapso de tiempo, se presume que el mosquito del dengue cumple su ciclo de vida (12 a 15 días; Beserra y de Castro 2008), y pica a una persona infectada e ingiere sangre que contiene el virus del dengue. A continuación, el virus se replica durante un período de incubación dentro del mosquito (8 a 12 días), el mosquito pica a una persona susceptible y le transmite el virus también a ésta, y finalmente, el virus se replica en la segunda persona y produce síntomas (3 a 14 días; CDC 2002).

El proceso básico de creación de modelos de predicción involucra los siguientes pasos: (1) preparación de los datos; (2) identificación del lapso de tiempo; (3) creación de las carpetas de insumo para Maxent; (4) capacitación del modelo; (5) pruebas de la validez del modelo; y (6) proyección y mapeo de las predicciones del modelo. Cada uno de estos pasos será descrito abajo.

#### *Preparación de los datos*

Maxent requiere que los mapas de cada variable de insumo sean presentados como una matriz de celdas idénticas con un valor asociado con cada celda para cada variable de predicción. En este estudio, la matriz estaba compuesta de celdas de 30 m por 30 m que contenían datos solo en localidades que traslapaban con uno de los barrios de León. Una lista de variables candidato que tenían posibilidad de ser útiles para predecir dengue fue desarrollada de una revisión de la literatura científica sobre *Aedes aegypti* y la transmisión de dengue. Esta lista fue después comparada con los datos disponibles sobre la topografía de la ciudad (por ejemplo, elevación, inclinación, y la predisposición del agua a encharcarse), factores demográficos (por ejemplo, densidad de población, tasa de madres a niños, tasas de pobreza); y características vinculadas a factores socioeconómicos (por ejemplo, tasas de pobreza, materiales de las viviendas, servicios sanitarios). Los datos climatológicos para todo el periodo del estudio, el cual se extendió del 1 de enero de 1998 al 31 de agosto de 2008, fueron adquiridos del Instituto Nicaragüense de Estudios Territoriales (INETER). La lista final consistió de 17 capas disponibles como insumo para Maxent (Cuadro 1).

Se utilizó un programa SIG (ArcMap<sup>®</sup> v.9.1) para convertir los datos sin procesar a mapas con matrices. Por ejemplo, los datos demográficos y socioeconómicos estaban disponibles solamente en informes impresos. Estos datos tuvieron que ser ingresados en hojas de cálculo y después unidos a un mapa (shapefile) de los barrios de León. Estos datos fueron transformados a porcentajes y medidas de área-por-unidad de manera que los barrios con diferentes áreas geográficas y tamaños de población pudieran ser comparados directamente, y luego convertidos a una matriz individual para cada variable socioeconómica. En otros casos, los mapas estaban ya disponibles como matrices. Estos fueron simplemente re-proyectados para tener la misma medida de celda y tamaño igual al de las otras capas.

Además de las 17 variables de insumo, se obtuvieron registros de casos de dengue de bases de datos del Ministerio Nicaragüense de Salud (MINSAL). Cada uno de los 252 casos en la base de datos del MINSAL para León incluyó los siguientes detalles de cada paciente: la fecha en que acudió a la clínica, domicilio, género, ocupación, y si el paciente estaba viajando cuando contrajo el virus. Los pacientes que tenían posibilidad de haber contraído la enfermedad fuera de León, ya sea que fuere porque estaban viajando cuando contrajeron el virus, o porque tienen ocupaciones que frecuentemente los llevan fuera de la ciudad (por ejemplo, conductor de camiones de carga), fueron eliminados del conjunto de datos. Para cada uno de los registros de casos que permanecieron una vez que se limpiaron los datos, se ingresó también la latitud y longitud de su domicilio en la base de datos.

Cuadro 1. Variables de predicción utilizadas en el análisis, incluyendo su tipo de variable y una descripción breve. DC = modelo Día Calendario; NDC = modelo No Día Calendario; 3SE = modelo 3 Semanas Clima.

<b>Tipo</b>	<b>Variable (unidades)</b>	<b>Descripción</b>	<b>DC</b>	<b>NDC</b>	<b>3SE</b>
Temporal	Día calendario (numeral)	Día del año de 1 a 366.	X		
Clima	Temp. Max menos 9semanas (grados C)	Promedio semanal de temperatura máxima diaria en un periodo de 9 semanas anteriores	X	X	X
Clima	Temp. Max. menos 10semanas (grados C)	Promedio semanal de temperatura máxima diaria en un periodo de 10 semanas anteriores			X
Clima	Temp. Max. menos 11semanas (grados C)	Promedio semanal de temperatura máxima diaria en un periodo de 11 semanas anteriores			X
Clima	Precipitación menos 10semanas (mm)	Promedio semanal de precipitación con un periodo de 10 semanas anteriores	X	X	X
Clima	Precipitación menos 11semanas (mm)	Promedio semanal de precipitación con un periodo de 11 semanas anteriores			X
Clima	Precipitación menos 12semanas (mm)	Promedio semanal de precipitación con un periodo de 12 semanas anteriores			X
Topográfica	Elevación (masl)	Elevación del terreno sobre la superficie	X	X	X
Topográfica	Acumulacion de Flujo (pixeles)	Numero de pixels cuenca arriba de cualquier punto dado	X	X	X
Topográfica	Inclinación (grados)	Inclinación del terreno	X	X	X
Topográfica	Indice Topográfico compuesto	Una medida de la predisposición del agua a encharcarse en base al flujo de acumulación y a la inclinación.	X	X	X
Demográfica	Densidad poblacional (personas/km2)	Población de un barrio dividida por el área de ese barrio.	X	X	X
Socioeconómica	Densidad de vivienda	Número de viviendas en el barrio dividió por el área de ese barrio.			
Socioeconómica	Hogares en pobreza extrema (%)	Número de hogares con ingresos menores al nivel especificado para pobreza extrema.	X	X	X
Demográfica	RNM	Tasa de mujeres fértiles con relación a menores de 5 años	X	X	X
Demográfica	Población con menos de 15 años (%)	Porcentaje de la población menor de 15 años			
Socioeconómica	Hogares sin servicios higiénicos (%)	Porcentaje de hogares que no cuentan con inodoro o letrina para la eliminación de excretas.	X	X	X
Socioeconómica	Hogares sin servicios de recolección de basura (%)	Porcentaje de hogares que no tienen acceso al servicio comunal de recolección de basura.	X	X	X
Socioeconómica	Hogares sin agua potable (%)	Número de viviendas que no son abastecidas por tubería dentro o fuera de la vivienda	X	X	X
Socioeconómica	Viviendas inadecuadas (%)	Porcentaje de viviendas clasificadas como "inadecuadas" por su construcción y servicios.	X	X	X
Demográfica	RDE	La proporción de niños de menos de 15 años más adultos mayores de 65 años con relación al número de personas entre 15 y 65 años.	X	X	X

### Identificación del lapso de tiempo

Como se mencionó arriba, una suposición central del enfoque de este modelo de predicción es que los mosquitos vectores del virus del dengue son afectados por variables climáticas en un lapso de tiempo igual al tiempo que les toma a los mosquitos salir del huevo, madurar, alimentarse en un individuo infectado, alimentarse de nuevo en un individuo susceptible, y para que la enfermedad comience a mostrar síntomas en el anfitrión. Se utilizó un enfoque estadístico llamado *correlación cruzada de lapso de tiempo* para identificar el lapso de tiempo en el cual la asociación más fuerte entre las variables climáticas y un brote en casos de dengue tuvieron lugar. La correlación cruzada de lapso de tiempo examina la robustez de la correlación entre casos de dengue y variables climáticas a diferentes lapsos de tiempo en periodos de una semana. Las correlaciones más robustas se presentaron con una temperatura máxima con un lapso de 9 semanas (Figura 3a), y con precipitación en un lapso de 10 semanas (Figura 3b). La asociación entre la fecha del brote de dengue y las variables climáticas utilizadas para predecir este caso de dengue fueron ajustadas para estos lapsos.

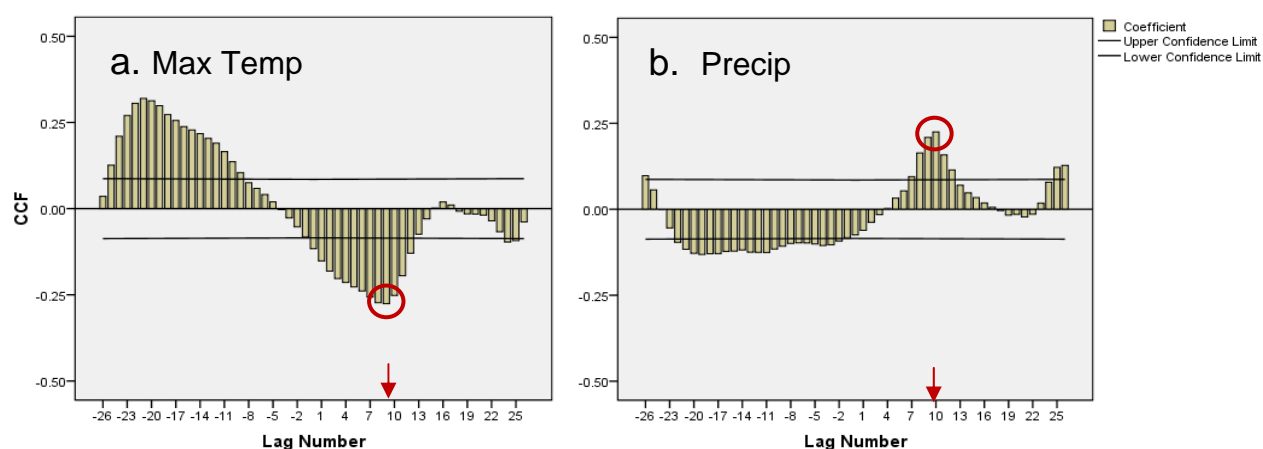


Figura 3. Diagramas de correlación cruzada de lapso de tiempo mostrando la robustez de la correlación entre casos de dengue y (a) temperatura máxima (coeficiente de correlación cruzada = -0.28), y (b) precipitación (coeficiente de correlación = 0.23). Estas dos correlaciones fueron ambas estadísticamente altamente significativas (nivel de confianza de 95% mostrado por la línea negra horizontal).

### Creación de carpetas de insumo de Maxent

El paso final anterior a la creación del modelo fue generar tres bases de datos de punto para ingresar en el programa. La primera base de datos de punto contenía las localidades de la vivienda de cada paciente de los registros de casos de dengue del MINSA. La segunda base de datos contenía 10,000 puntos generados al azar, y se utilizó en el proceso de desarrollo del modelo para establecer la gama completa de las condiciones de “antecedentes” ambientales presentes en el área de estudio entre las cuales el modelo debe distinguir las condiciones apropiadas para que aparezca un caso. La tercera base de datos consistió en 25,000 puntos seleccionados al azar en proporción directa a la densidad poblacional en cada barrio. Este conjunto de datos se utilizó para crear 100 modelos al azar, los cuales se utilizaron para probar la validez de los modelos de predicción de dengue generados de los datos verídicos de casos de dengue (que se explica en más detalle abajo). Una vez que estos conjuntos de datos fueron generados, fueron interceptados con las matrices de todas las variables de predicción excepto la temperatura máxima y la precipitación en un SIG. Para los datos de casos de dengue ocurridos, los campos de temperatura y precipitación en los cuadros fueron poblados con las condiciones climáticas reales presentes en la semana anterior a la fecha en que el paciente fue admitido a una clínica menos el lapso de tiempo (10 semanas para precipitación y 9 semanas para temperatura máxima). Para los conjuntos de datos de puntos al azar, los campos de

temperatura y precipitación fueron poblados al azar tomados de una distribución estadística similar a la distribución de datos climáticos verdadera del periodo de todo el estudio.

#### *Capacitación del modelo de predicción de dengue*

El primer paso en el proceso del modelo de predicción después de la preparación de los datos fue la reducción de los datos para eliminar variables redundantes. Se utilizó el análisis de correlación de Pearson para examinar la robustez de las relaciones entre variables. Se identificaron dos pares de variables redundantes: densidad poblacional con densidad de vivienda, y población menor de 15 años con la tasa de mujeres fértiles en relación a niños menores de 5 años (RNM). Para remover la redundancia entre capas de estos datos, no se utilizaron en la capacitación del modelo la densidad de vivienda y la población menor de 15 años porque se percibieron como menos útiles para la interpretación de los resultados del modelo.

Se utilizó el programa Modelo de Entropía Máxima de Distribución de Especies, Versión 3.2.19 (<http://www.cs.princeton.edu/~schapire/maxent/>) para capacitar a los modelos en la predicción de la probabilidad de que un caso de dengue ocurra en cada pixel del área de estudio en cada paso de tiempo de una semana. Con el propósito de hacer comparaciones, se capacitaron tres modelos diferentes con el conjunto de datos completo de 225 casos de dengue. El primer modelo, al cual de aquí en adelante nos referiremos como “Día Calendario” (DC), incluyó el día del año y las condiciones climáticas en la semana anterior al comienzo del lapso de tiempo. El segundo modelo, “No Día Calendario” (NDC), incluyó las condiciones climáticas de la semana anterior al comienzo del lapso de tiempo sin el día del año. El tercer modelo, “3 Semanas Clima” (3SE), utilizó las condiciones climáticas de la semana anterior al comienzo del lapso de tiempo y dos semanas adicionales y omitió el día calendario (ver Cuadro 1 para conocer las variables incluidas en cada modelo). Estas tres configuraciones fueron evaluadas una contra la otra para identificar el modelo más fuerte. Cada modelo fue corrido una vez con todos los datos de casos ocurridos para obtener la mejor predicción posible.

#### *Pruebas de la validez del modelo*

Se utilizaron dos enfoques para evaluar el comportamiento de cada modelo: (1) pruebas en contra de un modelo nulo; y (2) pruebas con un subconjunto del 25% de los datos. Cada uno se explica brevemente más abajo.

El enfoque del modelo nulo es una prueba estadística de la hipótesis que dice que “el modelo se comporta significativamente mejor que una predicción completamente al azar”. En esta prueba, el comportamiento de cada modelo de predicción de dengue es comparado con el límite más alto de un intervalo de confianza del 95% construido de 100 modelos de puntos obtenidos al azar del conjunto de datos de 25,000 puntos descrito más arriba. La medida del comportamiento del modelo utilizado para hacer esta comparación es “el área bajo la curva” (AUC por sus siglas en inglés) del diagrama característico del operador recibidor, el cual proporciona una sola medida independiente del umbral de la veracidad general (consultar Fielding y Bell 1997). Si la AUC del modelo de predicción de dengue queda fuera del intervalo de confianza del 95% establecido por la varianza entre el AUC del modelo al azar, entonces se confirma la hipótesis de que el modelo se comporta significativamente mejor que uno al azar (Raes y Steege 2007).

En el procedimiento de validación con el subconjunto del 25%, se llevan a cabo 100 corridas del modelo utilizando una porción fija de todos los datos de casos de dengue ocurridos, y los puntos restantes se utilizan para pruebas. En este caso el 75% de los puntos fueron utilizados para capacitar el modelo y los restantes 25% fueron utilizados para probar el comportamiento



del modelo. Este procedimiento de validación requiere la definición del umbral arriba del cual se presume que el dengue está presente. En este caso, un umbral fijo es igual a la probabilidad en la cual el 90% de las localidades utilizadas donde ocurrieron casos en la capacitación fueron clasificadas correctamente como presentes (por ejemplo, quedaron arriba del umbral). Esta proporción de puntos del 25% de los datos separados para hacer pruebas clasificados correctamente como presentes fue también utilizada como medida de comportamiento. Para obtener una estimación robusta del porcentaje de clasificación correcto, se llevaron a cabo 100 corridas con diferentes partes de 25% separadas y se calcularon los promedios del porcentaje correctos y la desviación estándar.

#### *Proyección del mapeo de las predicciones del modelo*

El programa Modelo de Entropía Máxima de Distribución de Especies genera un conjunto de coeficientes matemáticos que permite que el modelo sea utilizado con nuevos datos tanto para “predecir hacia el pasado” como para “predecir hacia el futuro” el riesgo de dengue sobre diferentes combinaciones de condiciones ambientales. Esta capacidad de proyectar el modelo sobre condiciones nuevas le da al modelo su utilidad para visualizar escenarios pasados y futuros. Para mapear los resultados de una predicción de riesgo de dengue desde Maxent, es necesario aplicar estos coeficientes a las condiciones ambientales presentes en el área de estudio. Después de seleccionar el modelo más robusto utilizando los métodos de validación descritos arriba, utilicé el modelo capacitado para predecir hacia el pasado el riesgo de dengue para cada semana de todo el periodo de tiempo cubierto por el estudio (544 semanas en total). Luego se mapearon las predicciones de cada semana para una visualización final de riesgo de dengue durante los 10 años del estudio.

### **Resultados y Discusión**

De las tres configuraciones del modelo que se corrieron, la del modelo 3SE se comportó mejor que las de los modelos DC o NDC (Cuadro 2). Una comparación de los resultados de comparación de las validaciones de los modelos mostró que los modelos 3SE y DC se comportaron significativamente mejor que uno al azar, como se indicó por el hecho que la AUC de estos modelos fue mayor que el umbral superior del intervalo de confianza del 95% formado de las AUCs de los 100 modelos al azar. Una comparación de las tasas de omisión de prueba del procedimiento del subconjunto de 25% mostró que bajo el modelo 3SE el 86% de los puntos de datos de prueba fueron clasificados correctamente como presentes, y que el modelo DC clasificó correctamente el 83% de los casos como presentes, y que el modelo NDC clasificó correctamente el 82% de los casos de dengue. Estos resultados indican que tanto el modelo 3SE y el DC hicieron una buena labor prediciendo casos de dengue con tasas de veracidad mayores al 80%. En resumen, los modelos desarrollados del conjunto de datos de 10 años de León predijeron correctamente una alta proporción de ocurrencias de dengue en la última década, y por lo tanto pueden ser utilizados razonablemente para predecir riesgo de dengue en el futuro.

Un modelo Maxent capacitado hace predicciones que son espacialmente explícitas y que pueden ser representadas en mapas. Una proyección típica de uno de estos modelos muestra una estimación del riesgo de dengue en cada celda de 30 m por 30 m en la ciudad, dados los datos de insumo presentados al modelo para su proyección (Figura 4). Cuando el riesgo de dengue es proyectado en una serie temporal (Figura 5), los patrones estacionales de nivel de riesgo son obvios. Un diagrama de la influencia de la variable día calendario sobre la probabilidad de que ocurra dengue muestra que los riesgos de dengue son mayores entre la mitad de octubre y la mitad de enero cada año, correspondiente con la estación de lluvias.

Cuadro 2. Resultados de los ejercicios de validación de las tres configuraciones de modelos de predicción probados en este estudio. La columna AUC del modelo muestra el valor para la AUC de cada corrida del modelo con todos los 225 puntos de datos. El siguiente valor muestra el límite superior del intervalo de confianza de 95% creado de los AUCs de las 100 corridas del modelo con puntos al azar. Una AUC de modelo que es mayor que el límite superior del intervalo de confianza del 95% se comporta significativamente mejor que uno al azar. La última columna muestra la proporción de puntos en la rutina del subconjunto de 25% que fueron clasificados correctamente como presentes por cada modelo.

<b>Modelo</b>	<b>AUC del Modelo</b>	<b>Límite superior del IC de 95% para un AUC Nulo</b>	<b>% Clasificado correctamente como presente</b>
Día Calendario	0.930	0.9272	83 +/- 5
No Día Calendario	0.913	0.9173	82 +/- 6
3 Semanas Clima	0.955	0.9439	86 +/- 5

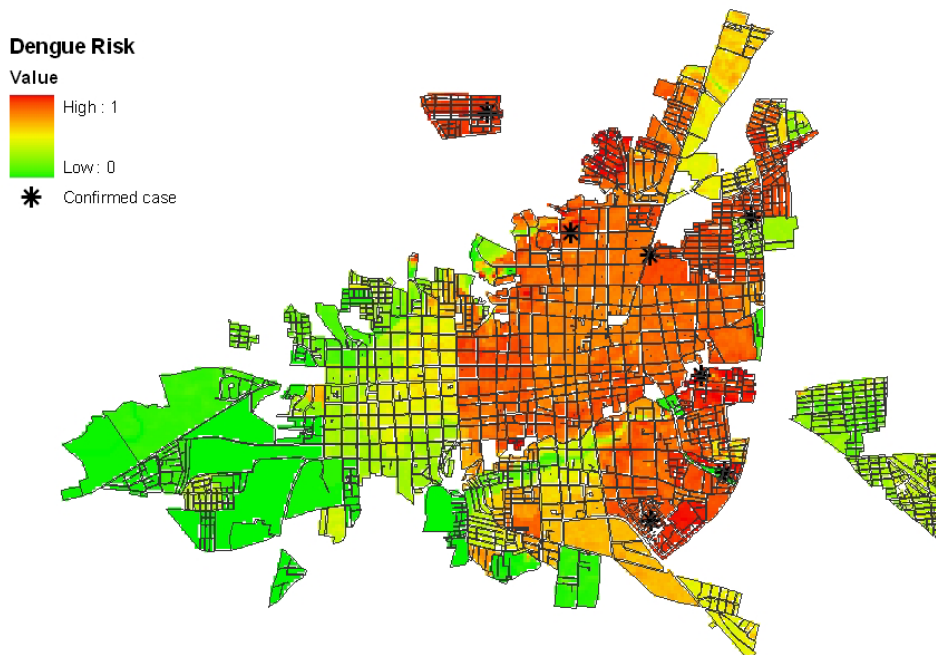


Figura 4. Una proyección del modelo Día Calendario sobre condiciones ambientales presentes durante principios de noviembre de 2001. El mapa muestra los barrios de León, con los colores más cálidos (rojos, naranjas) indicando un mayor riesgo de dengue. Los asteriscos negros representan localidades confirmadas de casos de fiebre de dengue.

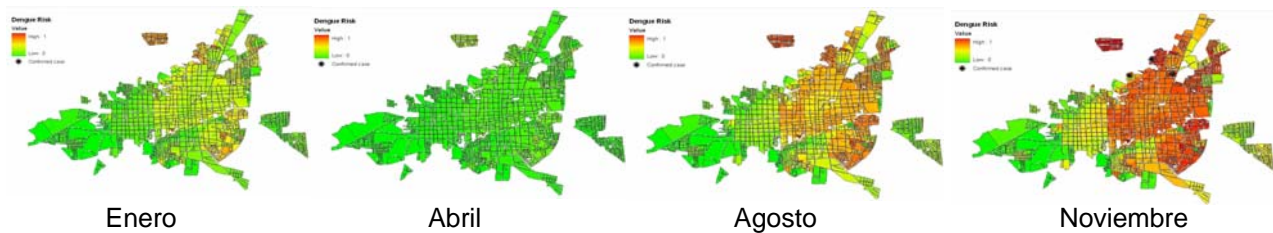


Figura 5. Serie temporal de predicciones del riesgo de dengue utilizando datos actuales de 2001.

Un examen de la influencia de diferentes variables sobre los modelos DC y 3SE muestra que las variables climáticas (incluyendo Día calendario si es tomada como un sucedáneo para las estaciones) hace la mayor contribución relativa a los resultados del modelo (Figura 6). En el modelo DC, los factores topográficos incorporaron la siguiente mayor influencia, seguidos por los factores demográficos (densidad poblacional y RNM). En el modelo 3SE, las condiciones climáticas fueron todavía más influyentes, seguidas por los factores topográficos, factores demográficos y después condiciones socioeconómicas

Dado que dos modelos validos fueron desarrollados en este proyecto, cuál de los dos tiene la mayor utilidad para predecir riesgo de dengue? El modelo 3SE tiene ciertas ventajas sobre el modelo DC que lo hacen ser el mejor de los dos modelos. Primero, es más veraz, habiendo clasificado verazmente el 86% de los casos conocidos de dengue (contra el 83% del modelo DC). La otra ventaja del modelo 3SE es que es más apropiado para una realidad en donde los patrones globales de clima son cada vez más inestables debido a cantidades elevadas de CO<sub>2</sub> en la atmosfera. Si la temporalidad de la lluvia y las temperaturas cambian substancialmente, u ocurren mas anomalías en la precipitación, un modelo que se apoya solamente en la temperatura y la precipitación es un modelo más robusto que uno que se apoya en el día calendario, el cual guarda diferencias históricas en el clima de las estaciones. Dado que la fiebre de dengue puede intensificarse con el calentamiento del clima global, es importante tener un modelo que puede manejar el cambio climático. Por estas razones el modelo 3SE se recomienda como el mejor modelo para ser utilizado en un sistema de toma de decisiones sobre el dengue.

Todos los modelos matemáticos están sujetos a tener problemas y a dar resultados no veraces, y por ello deben ser usados en la forma correcta para apoyar la toma de decisiones. Los modelos matemáticos simplifican la realidad en un intento de generalizar las relaciones entre variables. No solo están limitados estos modelos en su capacidad de representar la realidad, sino que dependen mucho de la calidad de los datos utilizados para desarrollar el modelo y de que tan completos estén estos datos. Debido a que los procesos que controlan a los *Aedes* y las poblaciones anfitrionas humanas son muy complejos, los datos utilizados en este modelo dejaron por fuera detalles importantes. Por ejemplo, debido a que no había datos disponibles, nuestro modelo no sigue las cuatro diferentes variables (serotipos A, B, C, y D) del virus del dengue que se conocen actualmente. Además, los modelos no incorporan el hecho de que una vez que un individuo se contagia con una variedad de dengue se vuelve resistente a esa variedad por el resto de su vida. No había datos disponibles sobre el tipo e intensidades de los esfuerzos para controlar mosquitos llevados a cabo anualmente en la ciudad, ni de los sitios en donde los pacientes pasan el día fuera de su casa. Como resultado de estas y otras fuentes potenciales de error en nuestros modelos, estos deben ser considerados como *un primer cálculo aproximado del riesgo de fiebre de dengue* que predice lo que puede suceder en el futuro en base a patrones observados en el pasado reciente.

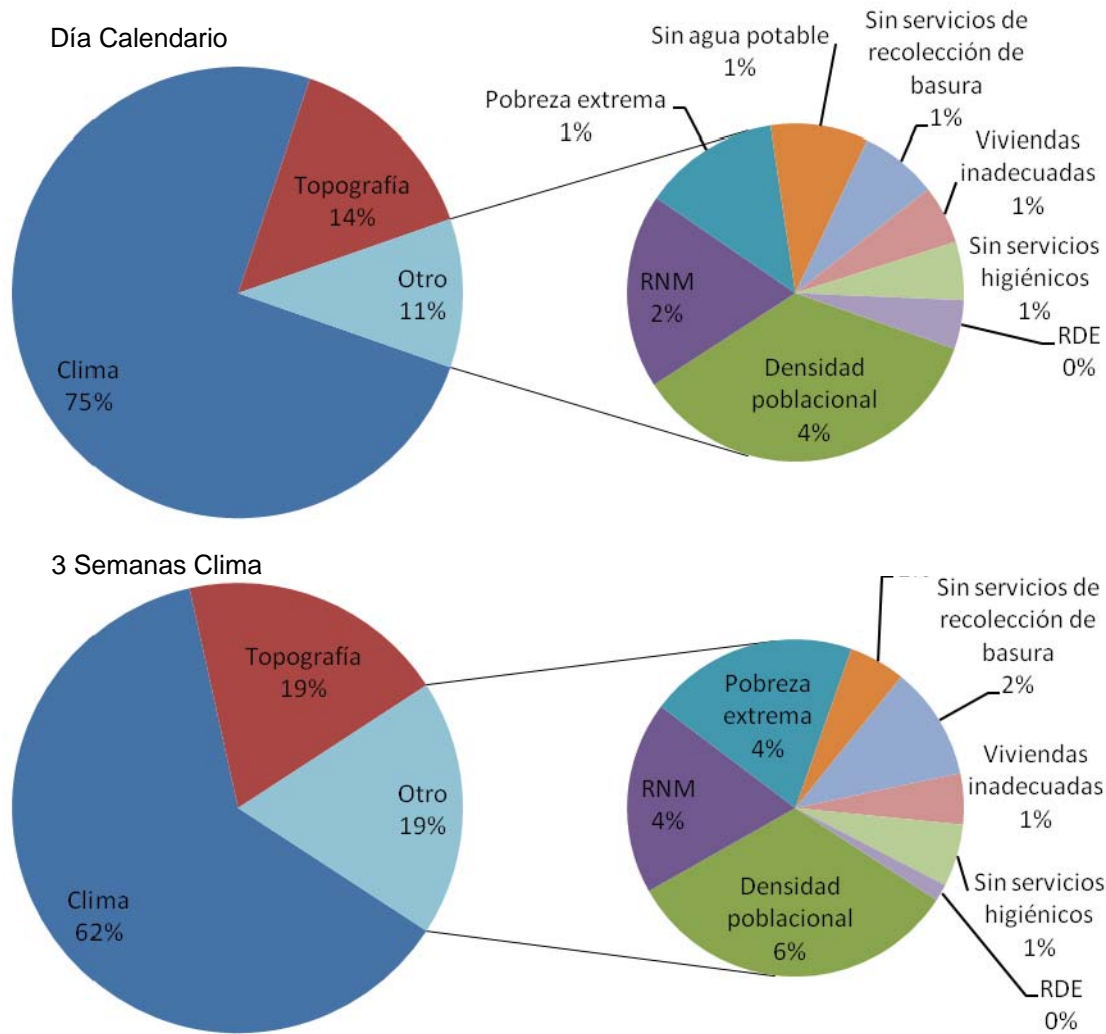


Figura 6. Contribuciones relativas de las clases de variables del modelo Día Calendario (parte de arriba) y del modelo 3 Semanas Clima (parte de abajo). La gráfica circular a la derecha de la principal gráfica circular muestra la división para las variables socioeconómicas, con Densidad poblacional haciendo la mayor contribución relativa en ambos casos.

A pesar de estas debilidades, los modelos presentados aquí muestran que es posible predecir el riesgo de que ocurran casos de fiebre de dengue con alguna veracidad y que se justifica más investigación y desarrollo de más modelos de predicción. Conjuntos de datos temporales y geográficos más extensos muy probablemente mejorarán la capacidad de los modelos para predecir verazmente el riesgo de dengue a través de una gama de contextos. Una consideración más explícita de la ecología de los mosquitos, de la dinámica de la enfermedad dentro de poblaciones humanas y la inclusión de diferencias de serotipos entre tipos de dengue pueden también dar más realismo y/o capacidad de predicción a los modelos. Los beneficios de tener un modelo de predicción funcional para anticipar emergencias de salud pública son lo suficientemente amplios para que se lleven a cabo mayores esfuerzos para expandir el área geográfica sobre la cual los modelos de dengue pueden ser aplicados al incluir datos de otros ambientes urbanos en América Latina.

## Literatura Citada

- Beserra, E.B., y F.P. de Castro, Jr. 2008. Biología comparada de poblaciones *Aedes (Stegomyia) aegypti* (L.) (Diptera: Culicidae) da Paraíba. *Neotropical Entomology* 37: 81 – 85.
- Centers for Disease Control and Prevention (CDC). 2002. Dengue Fever Slide Sets. Electronic resource: <http://www.cdc.gov/ncidod/dvbid/dengue/slideset/set1/i/slide05.htm>; (Accessed 28 August 2009).
- Deprandine, C., y E. Lovell. 2004. Climatological variables and the incidence of Dengue fever in Barbados. *International Journal of Environmental Health Research* 14: 429-441.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudik, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberon, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129-151.
- Fielding, A. H. y J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38-49.
- Hernandez, P. A., C. H. Graham, L. L. Master, y D. L. Albert. 2006. The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography* 29:773-785.
- Instituto Nacional de Información de Desarrollo (INIDE). 2008. *León en cifras*. Instituto Nacional de Información de Desarrollo, Managua, Nicaragua.
- Koopman, J.A., D.R. Prevots, M.A.V., H.G. Martin Dantos, M.L.Z. Aquino, I.M. Longini Jr. y J.S. Amor. 1991. Determinants and predictors of dengue infection in Mexico. *American Journal of Epidemiology* 133: 1168 - 1178.
- Phillips, S. J., R. P. Anderson, y R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190:231-259.
- Raes, N. y H. ter Steege. 2007. A null-model for significance testing of presence-only species distribution models. *Ecography* 30:727-736.