

GENERACIÓN E INCORPORACIÓN DE PRODUCTOS DE VALOR AGREGADO A UN SERVIDOR DE MAPAS PARA EL MANEJO EPIDEMIOLÓGICO DE CHAGAS

por Diana Marcela Brito Hoyos

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

UNIVERSIDAD NACIONAL DE CÓRDOBA

Agosto, 2015

©IG - CONAE/UNC 2015

©FaMAF - UNC 2015

Director: Dr. Gabriel Jaime Parra Henao

Co-Directora: Mg. Estefania Aylén de Elia



Generación e incorporación de productos de valor agregado a un servidor de mapas para el manejo epidemiológico de Chagas por Diana Marcela Brito Hoyos distribuye bajo la Licencia *Creative Commons Atribución-NoComercial-CompartirIgual 2.5 Argentina*.

<http://creativecommons.org/licenses/by-nc-sa/2.5/ar/>.

Resumen

El presente trabajo de tesis describe la cadena de obtención de datos y procedimientos implementados para la generación de productos de valor agregado, en formato raster y vectorial, y su incorporación al sistema Web-GIS para el manejo epidemiológico de Chagas. Se describe detalladamente los aspectos epidemiológicos de la transmisión de la enfermedad de Chagas en Colombia y Argentina; así como las funcionalidades, arquitectura y tecnologías con las que fue construido el Web-GIS. Este documento describe paso a paso las fuentes de variables ambientales provenientes de datos del sensado remoto, el procedimiento, software, y scripts generados para obtener los 10 productos en formato raster, y las herramientas de transferencia de datos para la incorporación y publicación de las capas en la plataforma. Adicionalmente, se describe el procedimiento implementado para generar las capas vectoriales a partir de los registros de la planilla Chagas 6, su incorporación y publicación, haciendo principal énfasis en la generación de consultas SQL. 372 capas en formato raster y 11 en formato vectorial fueron generadas e incluidas al sistema, con lo que se pretende aportar con materia prima para la generación de modelos espaciales de la distribución de la enfermedad en Argentina y contribuir al soporte y mejoramiento de planes estratégicos para el control, prevención y erradicación del mal de Chagas en el país.

Palabras Clave: Web-SIG, Enfermedad de chagas, Geoserver, Raster, Vector

Abstract

This thesis describes the data collection chain and procedures implemented to the value-added products generation, in raster and vector format, and its incorporation into the Chagas disease epidemiological management Web-GIS system. Epidemiological aspects of Chagas disease transmission in Colombia and Argentina are described in detail; as well as the functionality, architecture and technologies with which the Web-GIS was built. This document describes step by step the environmental variables sources from remote sensing data, the procedures, software, and scripts to generate the 10 products in raster format, and the data transfer tools for the integration and publication of the layers into the platform. Furthermore, the procedure implemented to generate vector layers from the Chagas 6 records, its incorporation and publication are described, by making primary emphasis on SQL queries. 372 layers in raster format and 11 in vector format were generated and included into the system, this aims to provide raw material for generating spatial models of disease distribution in Argentina and to contribute with the support and improvement of strategic plans for the control, prevention and eradication of Chagas disease in the country.

Key words: Web-GIS, Chagas disease, Geoserver, Raster, Vector

Agradecimientos

- Agradezco a **Dios** por este nuevo paso en mi camino..
- Agradezco y dedico este trabajo a mi **familia** por el incansable apoyo para que pudiera estudiar, terminar mi carrera, salir de mi país y seguir. Por cada enseñanza que me dieron y me siguen dando, por cada grano que pusieron para construir la persona que soy hoy.
- Agradezco a **mi país** por mis raíces, porque orgullosamente soy Colombiana, conozco el valor de la vida y de las oportunidades únicas como esta maestría.
- Agradezco a **Argentina y a la educación pública** por permitirme estudiar en una Universidad de alta calidad como la UNC, por los recursos económicos que me fueron dedicados, por su hermosa cultura y gente que día a día hicieron de mi permanencia aquí una experiencia inolvidable.
- A la **CONAE**, y especialmente al **Instituto Mario Gulich**, por construir y mantener esta maestría, y por brindarme la oportunidad de la beca.
- Agradezco muy especialmente al **Doctor Marcelo Scavuzzo**, por sus consejos siempre oportunos, por el conocimiento, tiempo, buena disposición, amistad y por enseñarme un claro ejemplo de un profesional integral y una persona dispuesta a la labor más noble, la enseñanza.
- Agradezco a mi director **Gabriel Parra Henao** y a mi codirectora **Estefania De Elia** por su confianza y paciencia durante el desarrollo de este trabajo.
- Agradezco a los miembros del **jurado** por sus aportes y sugerencias.
- Agradezco la invaluable ayuda de **Pablo Zader**, por la paciencia, por cada consejo, por el conocimiento, por su tiempo. Su apoyo en todo momento fue trascendental en el desarrollo de esta tesis y de mi estadía en Argentina.
- Agradezco a **Veronica Andreo** por la ayuda con GRASS y por lidiar con mi poca afinidad con los manuales.
- Agradezco a cada uno de mis compañeros, **leones y leonas** de la vida ustedes fueron mis maestros. Más que mis compañeros son mis amigos, a ustedes gracias por ser mi familia en este lugar.

Contenidos

Agradecimientos	iii
Contenidos	iv
Lista de Figuras	vii
Lista de Tablas	x
Acrónimos	xi
1 Introducción General	1
1.1 Motivación	2
1.2 Objetivos	3
1.2.1 Objetivo general	3
1.2.2 Objetivos específicos	3
1.3 Estructura de la tesis	4
2 Introducción a la enfermedad de Chagas en Argentina y Colombia	5
2.1 Historia y distribución geográfica	5
2.2 Chagas en Argentina	6
2.3 Chagas en Colombia	9
2.4 Epidemiología	11
2.5 Ciclo de vida	12
2.6 Hospederos, reservorios y ciclos de transmisión vectorial	14
2.6.1 Ciclo de transmisión doméstica	14
2.6.2 Ciclo de transmisión selvática	15
2.6.3 Ciclo de transmisión peridoméstica	15
2.7 Vectores	17
2.8 Control	19
2.9 Patogénesis	22
2.9.1 Fase Aguda	22
2.9.2 Fase indeterminada	23
2.9.3 Fase Crónica	24
2.10 Tratamiento	24
2.11 Modelamiento y aplicaciones GIS	25
3 Sistema Argentino para el control y estratificación de riesgo de transmisión vectorial del <i>Trypanosoma cruzi</i>	27

3.1	Introducción	27
3.2	Funcionalidades	28
3.3	Tecnologías	30
3.3.1	Base de datos	30
3.3.2	Servidor de aplicaciones	31
3.3.3	Interface de Usuarios	32
3.4	Arquitectura	33
4	Cadena de procesamiento de productos raster e incorporación en la plataforma Chagas	36
4.1	Área de estudio	36
4.2	GRASS GIS	37
4.2.1	Estructura interna del <i>DATASET</i> de GRASS	38
4.3	Serie temporal de índices de vegetación (NDVI y EVI)	39
4.3.1	Fuente de Datos	39
4.3.2	Adquisición y preprocesamiento de datos	40
4.3.3	NDVI y EVI	41
4.3.4	Incorporación de índices verdes al GRASS GIS	42
4.3.5	Producto 1: Descriptivos anuales de la serie temporal de índices verdes	44
4.3.6	Producto 2: Anomalía anual de NDVI y EVI	48
4.3.7	Producto 3: Extracción de ciclos temporales mediante análisis de series de Fourier	51
4.3.7.1	Algoritmo r.hants	53
4.4	Serie temporal de temperatura de superficie (LST)	56
4.4.1	Fuente de datos	56
4.4.2	Adquisición y preprocesamiento de datos	57
4.4.3	Incorporación de la serie histórica de LST al GRASS GIS	58
4.4.4	Producto 4: Descriptivos anuales de la serie temporal de LST	59
4.4.5	Producto 5: Extracción de ciclos temporales de LST mediante análisis de series de Fourier	63
4.4.6	Producto 6: Grados día	66
4.5	Producto 7: Mapas de Cobertura de la tierra	70
4.6	Producto 8: Modelo de Elevación Digital (DEM)	74
4.7	Producto 9: Superficies climáticas con datos de terreno (WorldClim)	76
4.8	Producto 10: Índice topográfico de humedad (TWI).	81
4.9	Incorporación de productos a la plataforma Argentina <i>Web-GIS</i> de Chagas	83
5	Cadena de procesamiento de productos Vectoriales e incorporación en la plataforma Chagas	88
5.1	Planilla Chagas 6	88
5.2	PgAdmin y consultas SQL	89
5.2.1	Generación de consultas	91
5.2.1.1	Producto 1: Viviendas registradas	92
5.2.1.2	Producto 2: Viviendas evaluadas	95
5.2.1.3	Producto 3: Localidades evaluadas	97
5.2.1.4	Producto 4: Localidades sin infestación	99
5.2.1.5	Producto 5: Localidades con infestación	101

5.2.1.6	Producto 6: Animales intradomicilio	102
5.2.1.7	Producto 7: Animales peridomicilio	105
5.2.1.8	Producto 8: Índice de infestación intradomiciliaria	107
5.2.1.9	Producto 9: Índice de infestación peridomiciliaria	109
5.3	Producto 10: Distancia localidad infestada-No infestada	111
5.4	Producto 11: Combinación de Vector y Raster	114
6	Consideraciones finales y líneas de abordaje a futuro	120
6.1	Discusión	120
6.2	Líneas de abordaje a futuro	123
6.3	Conclusiones	124
A	Planilla Chagas 6	126
	Referencias	128

Índice de Figuras

2.1	Distribución geográfica de la enfermedad de Chagas (Schmunis, 2007)	6
2.2	Distribución geográfica de las principales especies de vectores en Argentina (Ministerio de Salud Argentina, 2012)	8
2.3	Riesgo de transmisión de la enfermedad de Chagas en Colombia. Rojo denota zonas de alto riesgo ; Amarillo indica zonas de riesgo medio ; Verde denota zonas de bajo riesgo ; Gris denota zonas no investigadas (Guhl et al., 2005) . .	10
2.4	Ciclo de vida y transmisión del <i>Trypanosoma cruzi</i> (Ministerio de Salud Argentina, 2012)	13
2.5	Intercambios entre Ciclos de transmisión vectorial del <i>Trypanosoma cruzi</i> (Coura y Dias, 2009)	16
2.6	Principales triatomíneos vectores del <i>Trypanosoma cruzi</i> en Colombia y Argentina Jurberg et al. (2004)	17
2.7	Iniciativas continentales para el control de los principales vectores de la enfermedad de Chagas asociados al hábitat humano (Rosas et al., 2007)	20
2.8	Complejo oftalmoganglionar o signo de Romana (Coura y Dias, 2009)	22
3.1	Modelo de arquitectura cliente servidor con tecnologías OpenGeo suite y su interacción	33
3.2	Arquitectura de la plataforma Chagas Argentina, interacción entre los cinco subsistemas	35
4.1	Área de estudio	37
4.2	Estructura del directorio GRASSDATA (Neteler y Mitasova, 2008).	39
4.3	NDVI mínimo anual para el año 2004	45
4.4	NDVI máximo anual para el año 2004	46
4.5	NDVI medio anual para el año 2004	46
4.6	EVI mínimo anual para el año 2004	46
4.7	EVI máximo anual para el año 2004	47
4.8	EVI medio anual para el año 2004	47
4.9	Flujo de trabajo para la obtención del producto descriptivos anuales de la serie temporal de NDVI y EVI	48
4.10	Anomalía de NDVI del año 2008 con respecto a la media del periodo 2004-2014	49
4.11	Anomalía de EVI del año 2008 con respecto a la media del periodo 2004-2014	50
4.12	Flujo de trabajo para la obtención del producto anomalía media anual de la serie temporal de NDVI y EVI	51
4.13	(a) Curva del coseno que representa el primer armónico; (b) curvas de los armónicos 1, 2, y 3; (c) curva generada por la suma de los armónicos de (b). (Jakubauskas et al., 2002)	52

4.14	Mapa de HANTS con los componentes de Fourier de NDVI del 2004: Rojo, ciclos anuales. Verde, ciclos semestrales. Azul, ciclos trimestrales	55
4.15	Flujo de trabajo para la obtención del producto extracción de ciclos temporales de NDVI y EVI	56
4.16	Temperatura de superficie media del año 2004	61
4.17	Temperatura de superficie mínima del año 2004	61
4.18	Temperatura de superficie máxima del año 2004	61
4.19	Desvío estándar de la temperatura de superficie del año 2004	62
4.20	Flujo de trabajo para la obtención del producto descriptivos anuales de la serie temporal de LST	63
4.21	Mapa de HANTS con los componentes de Fourier de temperatura (LST) 2004: Rojo, ciclos anuales. Verde, ciclos semestrales. Azul, ciclos trimestrales	65
4.22	Flujo de trabajo para la obtención del producto ciclos temporales de LST	66
4.23	Número de días del año 2004 en los que la temperatura de superficie fluctuó entre 0°C y 30°C	68
4.24	Flujo de trabajo para la obtención del producto Grados Día	69
4.25	Mapa de cobertura de la tierra del área de estudio en el 2012, según la clasificación del <i>International Geosphere Biosphere Programme</i> (IGBP) (United States Geological Survey, 2013).	73
4.26	Flujo de trabajo para la obtención del producto de Cobertura de la tierra	73
4.27	Modelo de Elevación Digital a partir de datos SRTM del área de estudio.	75
4.28	Flujo de trabajo para la obtención del producto Modelo de Elevación Digital a partir de datos SRTM	76
4.29	Coeficientes de correlación de Pearson para las 19 variables del data set BIOCLIM	79
4.30	Variables bio-climáticas a partir de datos de terreno <i>WorldClim</i>	80
4.31	Flujo de trabajo para la obtención del producto Superficies climáticas con datos de terreno (WorldClim)	81
4.32	Índice topográfico de humedad (TWI) del área de estudio	83
4.33	Flujo de trabajo para la obtención del producto Índice de humedad de terreno TWI a partir de un DEM SRTM	83
4.34	Menú de entrada a la interfaz gráfica del GeoServer	85
4.35	Ventana de creación de espacios de trabajo en el GeoServer	85
4.36	Editor de texto de la pantalla de creación de estilos del GeoServer	86
5.1	Interfaz gráfica de PgAdmin III con asistente para consultas SQL manuales	91
5.2	Salida del PgAdmin donde se detalla el resultado de la consulta SQL del total de viviendas agregadas a la base de datos	93
5.3	Ejemplo de la generación de nuevo origen de datos vectoriales en GeoServer	94
5.4	Ejemplo de la ventana de incorporación y publicación de nuevas capas del GeoServer	95
5.5	Salida del PgAdmin resultado de la consulta SQL del total de viviendas evaluadas	96
5.6	Salida del PgAdmin resultado de la consulta SQL con la suma de viviendas evaluadas por localidad	98
5.7	Representación gráfica de la capa publicada utilizando la herramienta <i>OpenLayers</i>	99
5.8	Salida del PgAdmin resultado de la consulta SQL con la suma de viviendas sin infestación por localidad	100
5.9	Salida del PgAdmin resultado de la consulta SQL con la suma de viviendas con infestación por localidad	102

5.10 Salida del PgAdmin resultado de la consulta SQL con el promedio por localidad de animales intradomiciliarios	104
5.11 Ejemplo gráfico del producto 6 utilizando la herramienta <i>OpenLayers</i>	105
5.12 Salida del PgAdmin resultado de la consulta SQL con el promedio por localidad de animales peridomiciliarios	106
5.13 Ejemplo gráfico del producto 7 utilizando la herramienta <i>OpenLayers</i>	107
5.14 Salida del PgAdmin de la consulta SQL con el IIID por localidad para cada año	108
5.15 Salida del PgAdmin resultado de la consulta SQL con el IIPD por localidad para cada año	110
5.16 Flujo de trabajo para la generación e incorporación de productos vectoriales al sistema Web-GIS	111
5.17 Ejemplo de la ventana de edición y asignación de estilo del GeoServer	114
5.18 Ejemplo de la ventana de edición y asignación de estilo del GeoServer de la capa de Degradación	119
6.1 Visualizador Web-GIS ISAGRO del proyecto PREISPA	123

Índice de Tablas

4.1	Capas que conforman el archivo .HDF (dataset) del producto MOD13Q1 (United States Geological Survey, 2011).	40
4.2	Capas que conforman el archivo .HDF (dataset) del producto MOD11A1 (United States Geological Survey, 2006).	57
4.3	Capas que conforman el archivo .HDF (dataset) del producto MCD12Q1 (United States Geological Survey, 2013).	70
4.4	Leyenda de cobertura de la tierra del <i>International Geosphere Biosphere Programme</i> (IGBP) (United States Geological Survey, 2013).	71
5.1	Opciones para introducir datos en PostGIS con la función <i>pgsql2shp</i>	93

Acrónimos

AEARTE	Aplicaciones Espaciales de Alerta y Respuesta Temprana a Emergencias
CONAE	Comisión Nacional de Actividades Espaciales
DEM	Modelo de Elevación Digital
EVI	Índice Mejorado de Vegetación
FaMAF	Facultad de Astronomía, Matemática y Física
FFT	Transformada Rápida de Fourier
FTP	<i>File Transfer Protocol</i>
GPL	Licencia Pública General
GRASS	<i>Geographic Resources Analysis Support System</i>
HANTS	<i>Harmonic Analysis of Time Series Algorithm</i>
HAP	<i>Health Appplication Project</i>
HDF	<i>Hierarchical Data Format</i>
IDL	<i>Interactive Data Language</i>
IG	Instituto de Altos Estudios Espaciales Mario Gulich
IGPB	<i>International Geosphere–Biosphere Programme</i>
IID	Índice de Infestación Intra-domiciliaria
IIPD	Índice de Infestación Peri-domiciliaria
LST	Temperatura de la Superficie de la Tierra
MODIS	<i>Moderate Resolution Imaging Spectroradiometer</i>
MSAL	Ministerio de Salud de la Nación Argentina
NASA	<i>National Aeronautics and Space Administration</i>
NDVI	Índice de Vegetación de Diferencia Normalizada
OMS	Organización Mundial de la Salud
RGB	<i>Red-Green-Blue</i>
SAR	Radar de Apertura Sintética
SIG	Sistemas de información Geográfica
SLD	<i>Styled Layer Descriptor</i>
SQL	<i>Structured Query Language</i>
SRTM	<i>Shuttle Radar Topography Mission</i>
TWI	Índice de humedad de Terreno
UNC	Universidad Nacional de Córdoba
USGS	<i>United States Geological Survey</i>
WCS	<i>Web Coverage Service</i>
WFS	<i>Web Feature Service</i>
WMS	<i>Web Map Service</i>

"No estalla como las bombas, ni suena como los tiros. Como el hambre, mata callando. Como el hambre, mata a los callados: a los que viven condenados al silencio y mueren condenados al olvido. Tragedia que no suena, enfermos que no pagan, enfermedad que no vende. El mal de Chagas no es negocio que atraiga a la industria farmacéutica, ni es tema que interese a políticos ni a los periodistas"....Eduardo Galeano

Capítulo 1

Introducción General

La enfermedad de Chagas descrita en 1909, es una infección crónica, sistémica y parasitaria causada por el protozoo *Trypanosoma cruzi* (Rassi et al., 2010); la principal forma de transmisión es vectorial, a través de deyecciones infectantes de insectos de la familia Reduviidae, Subfamilia Triatominae, conocidos de acuerdo a la zona geográfica como vinchucas, chupos, barbeiros, pitos, etc. De igual forma por transfusiones de sangre se puede transmitir la infección, por transmisión vertical (transplacentaria) o por alimentos contaminados, entre otras vías (World Health Organization, 2012).

La existencia de esta enfermedad en humanos fue un hecho generado en occidente, que ocurrió como causa de la entrada del hombre a ambientes naturales del vector, provocando desequilibrios ecológicos, y forzando a los triatomíneos infectados a ocupar viviendas humanas, llevándose a cabo el proceso de domiciliación, debido a que no solamente encuentran allí refugio sino también suficiente alimento en la sangre humana y de animales domésticos (Gulh et al., 2005). De esta manera entra el hombre a formar parte activa de la cadena de transmisión epidemiológica de la enfermedad de Chagas.

Esta enfermedad es uno de los problemas de salud pública más importantes en América Latina que actualmente afecta a un estimado de 8 millones de personas en 21 países, adicional al incremento de la dispersión hacia áreas no endémicas de la enfermedad debida a la migración humana (World Health Organization, 2012). Se estima que del total de las personas infectadas el 30-40 % padece o puede desarrollar cardiomiopatía, síndrome digestivo, o ambos (Rassi et al., 2010).

Triatoma infestans, *Rhodnius prolixus* y *Triatoma dimidiata* son las tres especies de vectores más importantes en la transmisión de *T. cruzi* al hombre (Rassi et al., 2010), siendo de gran relevancia para el seguimiento epidemiológico de la enfermedad. Históricamente, *T. infestans* ha sido, con mucho, el vector más importante en regiones endémicas sub-amazónicas (sur de Sudamérica); por su parte *R. prolixus* se reporta normalmente en el norte de América del Sur y América Central, y *T. dimidiata* ocupa una superficie similar, pero también se extiende más al norte, en México (World Health Organization, 2012).

Para el caso particular de Colombia, las especies de triatominos de mayor importancia debido a su adaptación a los ambientes domiciliarios, amplia distribución y asociación con la incidencia de la enfermedad de Chagas son *R. prolixus* y *T. dimidiata* ([World Health Organization, 2002](#)). *T. dimidiata* ha sido reportado habitando domicilios y peridomicilios principalmente en la región centro oriental de Colombia ([Gulh et al., 2007](#)), al igual que *R. prolixus*. Estas especies poseen un ciclo epidemiológico muy complejo; las poblaciones no domiciliadas causan dificultad en el control, debido a que pueden ser fuentes de infestación o de reinfestación de las viviendas ya tratadas y por lo tanto existe la posibilidad de reiniciar el ciclo de transmisión a los humanos ([Parra-Henao et al., 2015](#)).

En Colombia uno de los factores que determina la distribución de estos triatominos es la variación altitudinal propia de la geografía del país, encontrándose registros de vinchucas hasta los 2000 m.s.n.m así como variables ambientales que representan la vegetación y temperatura ([Parra-Henao et al., 2015](#)). Adicional a esto, existe una gran cantidad de ecosistemas diferentes, que pueden o no ser propicios para la presencia de una o varias de las especies que transmiten el parásito, hace que en el territorio colombiano hayan diferentes escenarios de transmisión.

Un gran número de variables ambientales que pueden ser obtenidas por satélites, como características de la vegetación, temperatura, régimen de lluvias y humedad ([Parra-Henao, 2012](#)), así como características de los domicilios son línea de partida para el análisis e implementación de modelos para entender dichos escenarios y de esta forma delimitar zonas de riesgo de infección ([Parra-Henao et al., 2015](#)), que contribuyan a ajustar los programas de control y prevención de la enfermedad en el país.

1.1 Motivación

Dentro del plan nacional para el manejo y control de Chagas en Argentina existe una herramienta que permite la estandarización de las fuentes de información y la sistematización de la recolección de los registros epidemiológicos de la enfermedad y sus vectores, incorporándolos a una plataforma SIG (Sistema de Información Geográfica) que asegura la calidad de la información y que permite la utilización de la información por parte de las entidades responsables de la prevención de la enfermedad en el país. Esta plataforma web permite la re-utilización de la información, debido a la disponibilidad de datos constante por parte de las entidades.

La arquitectura con la que fue construida la plataforma para Chagas es semejante a la implementada para el control de Dengue en Argentina, cuyo uso e implementación de dicho sistema se encuentra descrito por Porcasi y colaboradores (2012). Consiste en una infraestructura de carga de datos y visualización en un servidor de mapas on-line; que en primera instancia, permite que los registros obtenidos por los usuarios sean utilizados como base para el desarrollo de modelos y análisis con el objetivo de estratificar zonas de riesgo de transmisión vectorial en el territorio argentino.

Actualmente la plataforma Chagas recibe información de las planillas Chagas 6 (Anexo [A](#)) colectadas durante las acciones regulares de los programas nacionales y/o provinciales en la evaluación y rociado (tratamiento) de viviendas urbanas o rurales, o de la acción de los agentes de salud que realicen evaluaciones periódicas en viviendas situadas en regiones endémicas de

la enfermedad.

Por su parte, el programa nacional de investigación para la prevención, control y tratamiento integral de la enfermedad de Chagas en Colombia (Red Chagas) pretende contar con una plataforma SIG (Sistema de información Geográfica) que permita el flujo de datos entre las instituciones efectoras del control de la enfermedad como lo son las Secretarías de salud departamentales, el Ministerio de Salud y grupos de investigación; de una manera similar al manejo que se le ha dado en Argentina.

Con la ejecución del proyecto, la Red Chagas Colombia tiene como objetivo generar conocimiento que permita la interacción entre grupos de diversa naturaleza (estado, sector privado y sociedad civil) y de esta forma poder generar la integración del conocimiento obtenido en las estrategias de salud pública para contrarrestar la enfermedad. Desde el punto de vista epidemiológico se espera contar con una plataforma SIG que permita desarrollar modelos predictivos que mejoren la vigilancia y orienten las acciones en Salud Pública.

En este punto radica la importancia de este estudio que plantea la incorporación de capas ambientales con significado biológico para los vectores del *T. cruzi* que aporten la base para posteriores análisis de distribución de la enfermedad de Chagas, dichas capas son consideradas como productos de valor agregado. Se plantea automatizar la descarga de datos provenientes del sensado remoto y a su vez, de forma automática generar productos de valor agregado e incorporarlos a la plataforma Argentina. Dicho procedimiento enriquecerá la experiencia lograda con el sistema ya en uso en el país, a partir del cual se desarrollará la plataforma SIG en Colombia.

1.2 Objetivos

1.2.1 Objetivo general

Como objetivo general de la tesis se plantea (1) generar de forma automática productos de valor agregado con significado biológico para la enfermedad de Chagas, a partir de datos ambientales monitoreados por sensores remotos. (2) Incorporar de forma automática los productos finales generados y otras capas de información adicional a un servidor de mapas web.

1.2.2 Objetivos específicos

- Evaluar la disponibilidad web de datos satelitales ambientales.
- Automatizar la descarga de información satelital (raster y vectores) desde distintos servidores web y/o servidores FTP (*File Transfer Protocol*).
- Analizar diferentes algoritmos para la generación de productos de valor agregado.
- Implementar los algoritmos analizados (utilizando software libre) para la generación automática de los productos.

- Utilizar herramientas de transferencia de datos para automatizar la subida de los productos generados y otras capas ambientales a un servidor de mapas web, sin la interacción de los usuarios.

1.3 Estructura de la tesis

En el capítulo 2 se hace una revisión bibliográfica de conceptos básicos y del estado epidemiológico actual de la enfermedad de Chagas en Argentina y Colombia .

El capítulo 3 describe el sistema argentino para el control y estratificación de riesgo de transmisión vectorial del *Trypanosoma cruzi*, plataforma que actualmente funciona en Argentina. Se hace una descripción detallada de las funcionalidades y la arquitectura del sistema.

El capítulo 4 describe, a manera de tutorial, la cadena de procesamiento que se realizó para la obtención de todos los productos en formato raster generados. Desde la descarga de los datos de sensado remoto, seguido por los algoritmos implementados para la generación del producto y finalmente la subida y publicación de los datos en el servidor de Chagas.

Por su parte en el capítulo 5 se describe, de igual forma que en el capítulo 4, a manera de tutorial, la generación de capas en formato vectorial, que se obtienen de los registros pertenecientes a la planilla de Chagas almacenadas en el servidor. Se describen las consultas SQL, a partir de las cuales se obtienen los datos de interés, los algoritmos para la generación de las capas y su posterior integración al servidor.

Finalmente, en el capítulo 6 se discute la utilidad del trabajo y se plantean las consideraciones finales y líneas de abordaje a futuro.

Capítulo 2

Introducción a la enfermedad de Chagas en Argentina y Colombia

2.1 Historia y distribución geográfica

El antepasado del *Trypanosoma cruzi* fue introducido probablemente en América del Sur a través de los murciélagos hace aproximadamente 7-10 millones de años ([Steeverding, 2014](#)). Cuando los primeros seres humanos llegaron al Nuevo Mundo, el ciclo selvático de la enfermedad de Chagas fue establecido.

Datos paleoparasitológicos sugieren que la tripanosomiasis americana humana se originó en la zona andina, cuando se establecieron los primeros asentamientos en la región costera del desierto de Atacama ([Steeverding, 2014](#)).

La identificación de *T. cruzi* como el agente etiológico y los triatominos como vectores de transmisión de la enfermedad de Chagas se produjo al inicio del siglo 20; la enfermedad fue descubierta en 1909 por el médico brasileño Carlos Chagas (1879-1934), por lo que recibe comúnmente el nombre mal de Chagas.

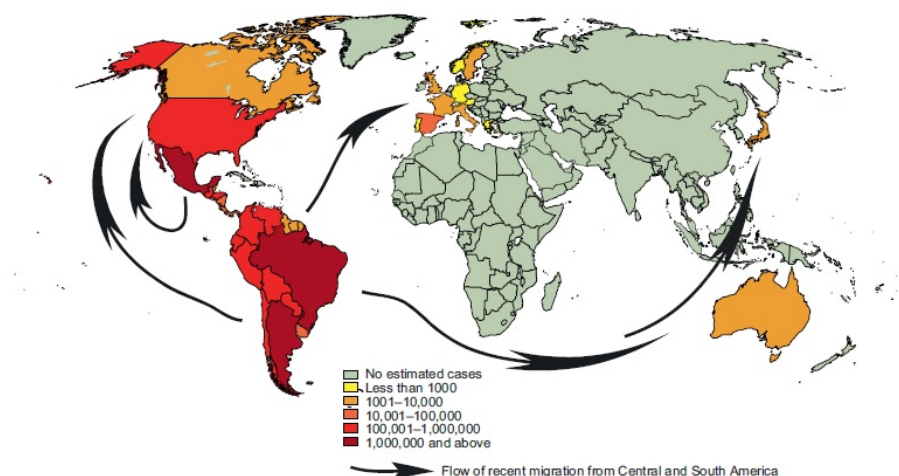


Figura 2.1: Distribución geográfica de la enfermedad de Chagas ([Schmunis, 2007](#))

El informe generado por el doctor Chagas es único en la historia de la medicina debido a que un solo científico describió con gran detalle tanto el ciclo de transmisión (vector, hospederos, y un nuevo organismo infeccioso), así como las manifestaciones clínicas con el análisis del primer caso humano ([Rassi et al., 2010](#)).

La infección por *T. cruzi* existe naturalmente solo en el continente americano, en un área que comprende desde la latitud 42° N en Estados Unidos (desde el estado de California al de Maryland), hasta alrededor del paralelo 34° S en Chile y el 42° S en la Argentina (Figura 2.1). También se han encontrado vectores y reservorios selváticos en la mayor parte de la región del Caribe, que antes se consideraba libre de la infección, y los casos fuera del continente americano no se han reportado como autóctonos ([Rassi et al., 2010](#)).

La historia nos enseña que la actividad humana que conduce a los cambios ambientales, en particular la deforestación, es la causa principal de la dispersión de la enfermedad. Recientemente, la migración de los pacientes infectados por *T. cruzi* desde áreas endémicas a lugares libres de la infección ha llevado a que la enfermedad se convierta en un problema global, llegando a países de América del Norte, Europa y la región del Pacífico occidental. El número total estimado de pacientes de Chagas fuera de América Latina es de más de 400.000 siendo los EE.UU. el país más afectado que representa tres cuartas partes de todos los casos dispersados ([Steverding, 2014](#)).

2.2 Chagas en Argentina

Actualmente en el mundo existen alrededor de 142 especies de triatomíneos distribuidos en 18 géneros y cinco tribus ([Galvao et al., 2003](#)), más de la mitad han sido natural y/o experimentalmente infectados con *T. cruzi*, pero son menos de 10 las especies que tienen importancia epidemiológica para el hombre, dado que son capaces de colonizar la vivienda

humana ([Ministerio de Salud Argentina, 2012](#)).

En Argentina, actualmente, se considera que existen 16 especies de triatomíneos, distribuidos en 3 géneros: *Psamolestes*, *Panstrongylus* y *Triatoma*.

Las especies de estos géneros vinculadas con la enfermedad son:

- *Psamolestes*: *coreodes*.
- *Panstrongylus*: *megistus*, *guentheri*, *geniculatus* y *rufotuberculatus*.
- *Triatoma*: *breyeri*, *delpontei*, *garciabesi*, *limai*, *platensis*, *eratyrsiformis*, *rubrovaria*, *patagonica*, *guasayana*, *sordida* e *infestans*.

Estas especies difieren en importancia epidemiológica según su hábitat, densidad poblacional y distribución geográfica. *T. guasayana*, *T.sordida*, *T. eratyrsiformis* y *T. patagonica*, si bien son especies silvestres y peridomiciliadas, pueden hallarse infectadas por *T. cruzi* y frecuentemente invadir y en algunos casos colonizar la vivienda con ataques a humanos y animales domésticos ([Ministerio de Salud Argentina, 2012](#)).

El resto de las especies son silvestres, siendo excepcionalmente atraídas por la luz hacia el interior de domicilio o peridomicilio, pero sin colonizar las viviendas, por lo que carecen de interés para la salud pública, sin embargo estos triatomíneos contribuyen al mantenimiento del ciclo de transmisión silvestre de *T. cruzi* ([Ministerio de Salud Argentina, 2012](#)).

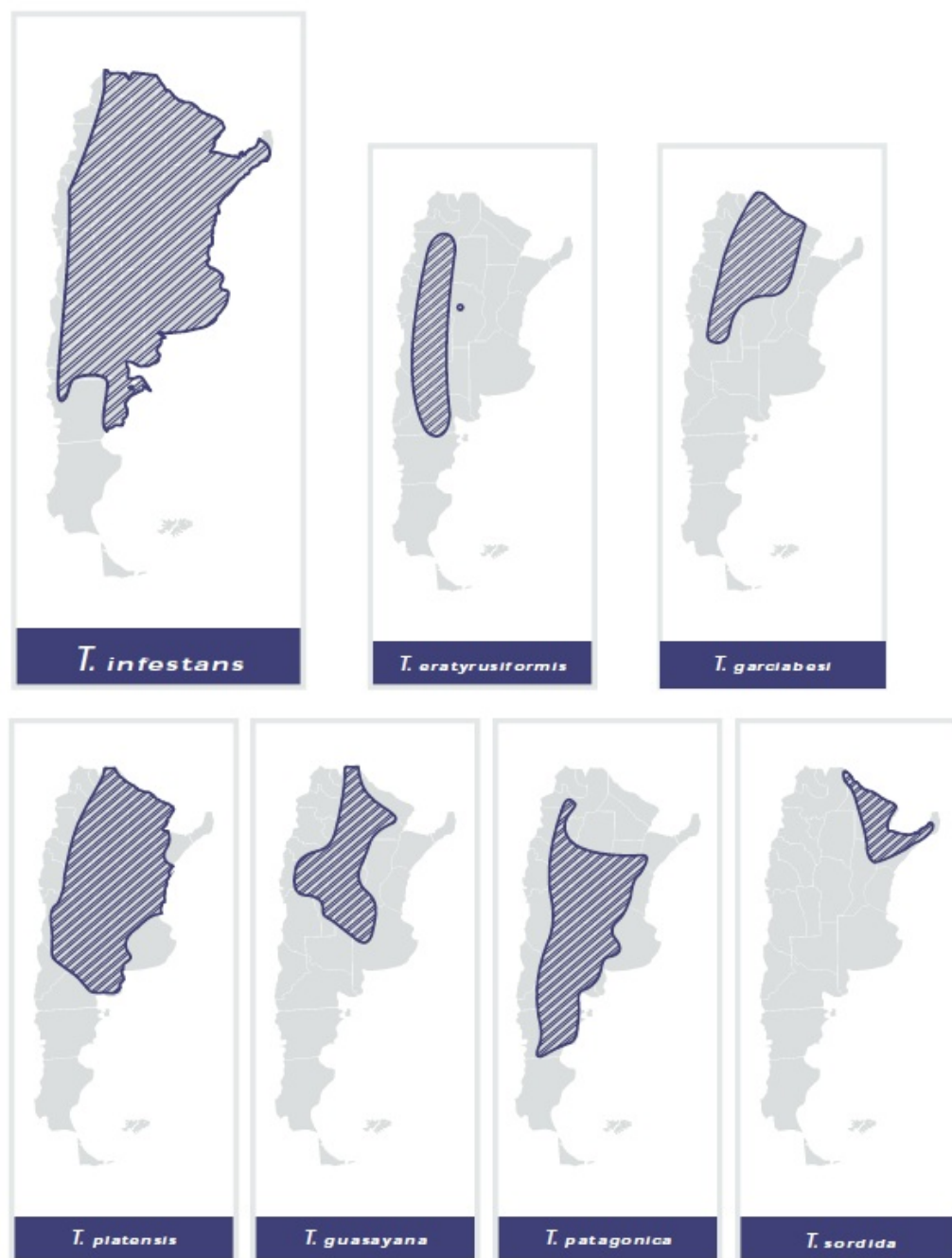


Figura 2.2: Distribución geográfica de las principales especies de vectores en Argentina
(Ministerio de Salud Argentina, 2012)

La distribución de los triatominos está relacionada principalmente con factores ambientales, climáticos y sociales; sin embargo los programas para el control vectorial son a nivel país y provincia, por lo que con fines didácticos se muestra su distribución por regiones en Argentina (Figura 2.2).

En el país al igual que en todo el Cono Sur *Triatoma infestans* es la principal especie vector del *T. cruzi*. Esta especie es endémica en diferentes regiones de América Latina; los programas de control lograron disminuir la densidad de las poblaciones de los vectores primarios a valores inferiores al del nivel crítico de transmisión vectorial, gracias a estrategias tales como la Iniciativa del Cono Sur (Moncayo y Ortiz, 2006).

En la Argentina el área de transmisión vectorial cubre el 60% del territorio al norte del paralelo 44° (Moncayo y Silveira, 2009). En 1980, la tasa media de infestación intradomiciliaria para el país fue de 30% ; posteriormente para 1998 fue de 1,2%, y ya en 2002, se redujo a 1% (Moncayo y Silveira, 2009), lo que equivale a una reducción del 98% en la actualidad.

A su vez la tasa de seroprevalencia para todo el país en el grupo de edad de 0-4 años es de 0,9% (Moncayo y Silveira, 2009), lo que confirma el bajo número de casos agudos en los niños de este grupo de edad. En el grupo de edad de 0-14 años, la tasa es del 1,9 %. Entre los hombres de 18 años de edad, la tasa de seroprevalencia disminuyó de 5,8% en 1981 al 1% en 1993 y 0,5% en 2002. Gracias a la iniciativa para el control y prevención de la enfermedad en el país se logró la interrupción de la transmisión vectorial en 10 de las 13 provincias endémicas de Argentina (Moncayo y Ortiz, 2006). Para controlar la transmisión por transfusiones de sangre, en el país se hace un análisis de las donaciones, en el sector público hay un 100% de cobertura de este estudio en los bancos de sangre, y en el sector privado un 80% (World Health Organization, 2012).

A pesar de los significativos logros de las iniciativas nacionales para controlar la enfermedad, no se ha logrado la sostenibilidad y la efectividad necesarias para interrumpir la transmisión del *T. cruzi*. La interrupción progresiva de la transmisión vectorial muestra un patrón espacial claro, dejando un gran territorio dentro del Gran Chaco, ecoregión que comprende Argentina, Bolivia y Paraguay, donde la enfermedad se sigue transmitiendo de forma activa por *T. infestans* (Bar et al., 2010), y registra los mayores niveles históricos de infección humana por *T. cruzi*, de miocardiopatías asociadas y de transmisión vertical (Porcasi et al., 2006).

El valor de los indicadores entomológicos actuales de la provincia de la Rioja, en particular, de los departamentos de Los Llanos es similar a la situación observada durante la década de 1960, anterior a la iniciación de las actividades de control; esta situación es el resultado de la desactivación virtual del programa de control de vectores de la provincia desde la segunda mitad de 1990 hasta 2003, y se asocia con la pobreza, bajos niveles de educación, y la cría de animales domésticos en el peridomicilio, que constituyen buenos refugios para los vectores (Porcasi et al., 2007).

Este escenario hace necesario que se reciba atención urgente por parte de las autoridades de salud en el Gran Chaco Argentino (Porcasi et al., 2006), tanto para el control vectorial como para la vigilancia de infecciones humanas, o tratamiento parasitológico en caso necesario.

2.3 Chagas en Colombia

En Colombia, en 1929, el Dr. César Uribe Piedrahita, profesor de Parasitología y Medicina Tropical de la Universidad Nacional, informó por primera vez el hallazgo del *T. cruzi* y del *T. rangeli*, encontrados en 15 ejemplares de *Rhodnius prolixus* (comunmente llamado en el país

como "Pito") en la población de Prado (Tolima) (Rosas et al., 2007).

La estadística más actualizada en Colombia muestra que 436.000 personas están infectadas por *T. cruzi* (tasa de prevalencia=0.96%), de las cuales 131.474 personas padecen cardiomiopatías. Cada año 5.250 casos son causados por transmisión vectorial y hay 1.000 nuevos casos de transmisión congénita de la enfermedad (Parra-Henao et al., 2015).

Colombia comenzó oficialmente el programa de prevención y control de la enfermedad de Chagas en 1996, teniendo en cuenta que en años anteriores se adelantaron medidas importantes de control, como por ejemplo el tamizaje obligatorio a todas las unidades transfusionales en bancos de sangre a nivel nacional, decreto que se promulgó en 1995 (Rosas et al., 2007), hoy en día el tamizaje es del 100% a nivel nacional y el estimativo de prevalencia en donantes a es del 2,1%.

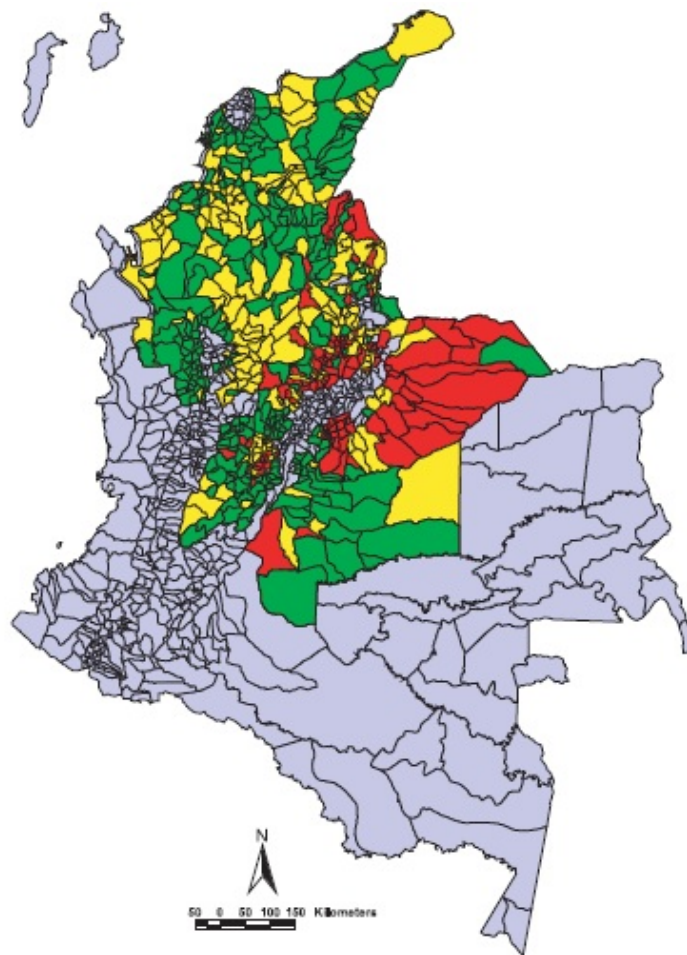


Figura 2.3: Riesgo de transmisión de la enfermedad de Chagas en Colombia. Rojo denota zonas de alto riesgo ; Amarillo indica zonas de riesgo medio ; Verde denota zonas de bajo riesgo ; Gris denota zonas no investigadas (Guhl et al., 2005)

En cuanto a los vectores, *Rhodnius prolixus* es responsable de la transmisión de Chagas en el norte de Sur América (Colombia y Venezuela y en Centro America es un importante vector en Honduras, El Salvador y Nicaragua); por su parte *Triatoma dimidiata* es también de suma

importancia epidemiológica (Parra-Henao, 2012).

Las áreas geográficas más comprometidas en Colombia (Figura 2.3) son Arauca, Boyacá, Cundinamarca, Santander, Norte de Santander, Casanare y Meta (Rosas et al., 2007).

En el país se han registrado 26 especies de triatomos, de las cuales 15 han sido relacionadas con la transmisión del *T. cruzi* (Gulh et al., 2007). Entre estas, *R. prolixus*, *T. dimidiata*, *T. maculata* y *T. venosa* han sido halladas infestando domicilios y peridomicilios en diferentes zonas del país (Gulh et al., 2007), con lo anterior las especies de triatomos de mayor importancia epidemiológica en Colombia, debido a su adaptación a los ambientes domiciliaros, amplia distribución y asociación con la incidencia de la enfermedad de Chagas son *R. prolixus* y *T. dimidiata* (Parra-Henao, 2012; World Health Organization, 2002).

2.4 Epidemiología

La mayoría de los casos de la enfermedad están relacionados con el contacto directo con el triatomo vector. El insecto se alimenta de la sangre de la persona, por lo general mientras duerme. La picadura se realiza en la cara u otras áreas expuestas del cuerpo como los brazos. La infección hacia el humano ocurre debido a que mientras el insecto se alimenta, defeca sobre la persona, quien luego por medio de heridas abiertas o rasguños sobre la picadura, ingresan al parásito al torrente sanguíneo. El ingreso del parásito también se puede dar por los ojos, cuando la persona se frota con las manos infectadas con la feca. Cuando el parásito se introduce en el ojo, se puede producir un edema unilateral del párpado llamado signo de Romana.

La transfusión de sangre, trasplante de órganos y transmisión congénita son métodos adicionales de transmisión de persona a persona; especialmente importantes en los países endémicos. El riesgo de adquirir la infección por transfusión aumenta con la cantidad de sangre recibida y la enfermedad en estos casos es típicamente asintomática o sus manifestaciones clínicas aparecen días a varias semanas después de la transfusión (World Health Organization, 2002). Los signos clínicos más frecuentes son fiebre, esplenomegalia y poliadenopatías (World Health Organization, 2002).

La transmisión a través de trasplante de órganos conlleva un efecto inmediato y más grave en la persona, debido a la inmunosupresión del paciente; esto conduce a casos graves de miocarditis, meningoencefalitis y finalmente la muerte si no se diagnostica rápidamente (World Health Organization, 2002). Este tipo de transmisión se ha convertido en una preocupación cada vez mayor en áreas no endémicas de la enfermedad debido a la inmigración de las personas infectadas con Chagas (Diaz, 2007)

Potencialmente se puede adquirir el mal de Chagas, a través de la ingestión de la leche de la madre infectada (Ferreira et al., 2001), por ingestión de alimentos o bebidas contaminadas con triatomos infectados y/o sus heces, y a través del consumo de alimentos o bebidas contaminados con heces o secreciones de las glándulas de algunos marsupiales (Jansen et al., 1999).

La transmisión oral representa un escenario epidemiológico emergente donde diferentes brotes asociados a la alimentación o consumo de bebidas han sido reportados en Argentina , Bolivia , Brasil , Ecuador y Venezuela . En Colombia , seis brotes orales humanos han sido reportados, lo que corrobora la importancia de esta ruta de transmisión en regiones donde los triatomíneos vectores son típicamente selváticos (Ramírez et al., 2013). Se estima que hay un aumento en la patogenicidad asociada con este modo de infección debido al efecto de los jugos gástricos.

En relación al riesgo de transmisión vectorial, la enfermedad de Chagas depende de un gran número de factores. Los estudios clásicos muestran que la probabilidad de transmisión vectorial de *T. cruzi* está relacionada con características referidas al tipo de construcción del domicilio y peridomicilio, el nivel socio-económico y cultural de los pobladores, la abundancia e infección de los vectores, la calidad de las intervenciones de los programas de control, etc.

2.5 Ciclo de vida

En condiciones naturales un triatómino infectado se alimenta de sangre y defeca simultáneamente, depositando con las heces formas infectantes (tripomastigotes metacíclicos) del parásito. Las heces contaminadas pueden ser llevadas a la conjuntiva ocasionando la principal manifestación de puerta de entrada, el signo de Romana, que se presenta como una prominente inflamación usualmente en la región de los párpados.

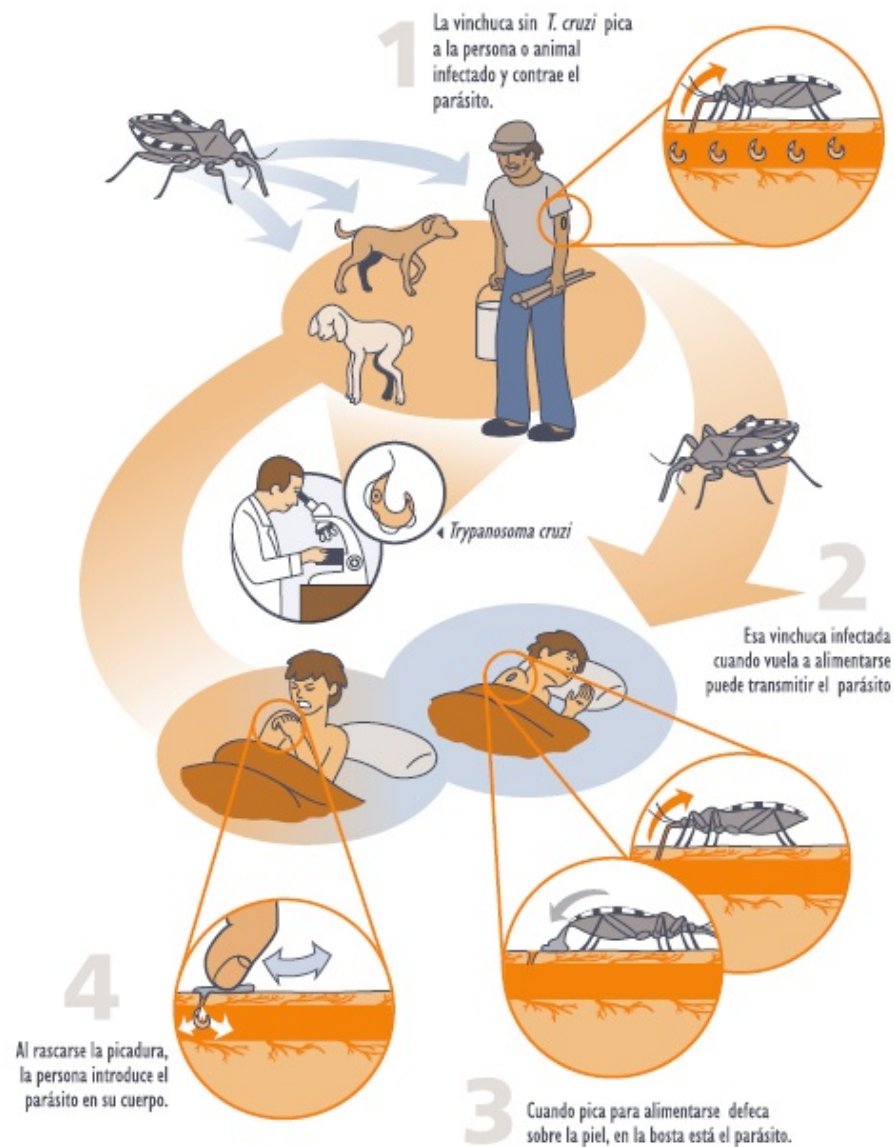


Figura 2.4: Ciclo de vida y transmisión del *Trypanosoma cruzi* (Ministerio de Salud Argentina, 2012)

Las formas infectantes también pueden penetrar al torrente circulatorio por otras vías, como las heridas en la piel, o por vía oral. El parásito presenta un especial tropismo por el tejido miocárdico, penetrando dentro de sus células para formar los típicos nidos de amastigotes. Cuando los amastigotes intracelulares se liberan y pasan al torrente sanguíneo de hospederos o reservorios, se convierten en tripomastigotes infectantes que circulan en la sangre y pueden ser ingeridos por un insecto vector no infectado, en donde se diferencian y se multiplican en su tracto digestivo hasta dar origen a las formas infectantes, completando así el ciclo (Figura 2.4).

El vector se vuelve infectante de los 30 a 40 días después de haber ingerido sangre infectada. La infección persiste en el intestino del triatomino durante toda su vida.

2.6 Hospederos, reservorios y ciclos de transmisión vectorial

El *T. cruzi* se ha encontrado en más de 180 especies de mamíferos, tanto domésticos como salvajes, pertenecientes a siete órdenes y 25 familias ([World Health Organization, 2002](#)). Reservorios comunes incluyen armadillos, marsupiales, osos perezosos, carpinchos, monos, puercoespines y zarigüeyas; estas últimas son de particular importancia para el ciclo de la enfermedad ya que llevan la etapa tripomastigote metacíclica infecciosa dentro de sus glándulas anales ([Jansen et al., 1999](#)).

Muchos mamíferos adquieren la infección a través de la ingestión de triatomíneos vectores infectados. Los cuises en Bolivia y Perú son frecuentemente encontrados con altas tasas de infección y se crían con frecuencia dentro de las viviendas o cerca de las viviendas ([World Health Organization, 2002](#)). Caninos y gatos domésticos son de importancia epidemiológica de la enfermedad, ya que sirven como centinelas para la transmisión mediada por vectores y se ha encontrado que tienen altas tasas de seroprevalencia en algunas zonas endémicas ([Gurtler et al., 2006](#)). La infección también se ha demostrado en una amplia variedad de otros animales domésticos, incluyendo cabras, ovejas, alpacas y cerdos, a pesar de que no son tan importantes para el ciclo de transmisión doméstico ([World Health Organization, 2002](#)).

Estudios indican que en la naturaleza se pueden encontrar al menos dos grupos principales de poblaciones de *T. cruzi*. El primero está estrechamente vinculado con el ciclo doméstico y produce infecciones y alta morbilidad en los seres humanos. El segundo está asociado con el ciclo silvestre y provoca infecciones más leves y menor morbilidad en los seres humanos. Existen algunas pruebas de que la distribución de estas poblaciones de parásitos, se relaciona con la distribución y otras características de las especies vector, y de que esto tiene importantes consecuencias epidemiológicas en la enfermedad de Chagas humana ([Rosas et al., 2007](#)).

En la naturaleza, *T. cruzi* mantiene tres tipos diferentes de transmisión vectorial, silvestre, peridoméstica y doméstica, a continuación una breve descripción de cada uno de estos ciclos.

2.6.1 Ciclo de transmisión doméstica

Este ciclo se mantiene por medio de triatomíneos domesticados que transmiten la infección ya sea de animales domésticos a los seres humanos, o entre seres humanos ([Coura y Dias, 2009](#)).

Se presenta en viviendas rurales o periurbanas usualmente en condiciones de mala calidad, con paredes de bahareque o adobe y techos de material vegetal ([Rosas et al., 2007](#)). Los principales reservorios del parásito son los seres humanos, los perros y una enorme gama de animales peridomésticos, especialmente el *Didelphis marsupialis*, o zarigüeya, que juega un papel epidemiológico muy importante en la transmisión del parásito ([Rosas et al., 2007](#)).

Los insectos vectores domiciliados viven y se multiplican en grietas de las paredes, agujeros del techo, debajo y detrás de los muebles o de los cuadros y en los anexos peridomiciliarios tales como gallineros, pilas de leña y arrumes de piedras o ladrillos. Es el caso de *Triatoma infestans*, principal vector domiciliado en los países de Cono Sur del continente, de *Rhodnius prolixus* y de *Triatoma dimidiata* en los países andinos y centroamericanos, y de *Triatoma barberi* en México.

2.6.2 Ciclo de transmisión selvática

El ciclo silvestre es enzoótico, presentandose relativamente pocos casos humanos de la enfermedad, y es mantenido por triatominos y animales salvajes (Coura y Dias, 2009).

Como se mencionó con anterioridad, a lo largo de la mayor parte del continente latinoamericano se han descubierto muchas especies y subespecies de pequeños mamíferos silvestres, terrestres o arbóreos, que son infectados de forma natural por el *T. cruzi*. En este ciclo intervienen triatominos selváticos que se infectan y a su vez infectan a roedores, marsupiales y otros animales salvajes.

Varias especies de triatominos conforman el ciclo silvestre de transmisión del *T. cruzi*, tales como *Panstrongylus geniculatus*, *Rhodnius colombiensis*, *Rhodnius brethesi*, *Rhodnius robustus* y *Rhodnius pallescens*, para mencionar algunos (Rosas et al., 2007).

2.6.3 Ciclo de transmisión peridoméstica

El ciclo peridoméstico se origina a partir del ciclo silvestre y es responsable de mantener la infección entre los animales domésticos en las zonas circundantes a las viviendas humanas (Coura y Dias, 2009).

En el ciclo peridoméstico interviene una gran variedad de mamíferos como roedores, marsupiales y perros, que entran y salen libremente de las viviendas, y triatomas selváticos atraídos a las casas por la luz y el alimento (Rosas et al., 2007).

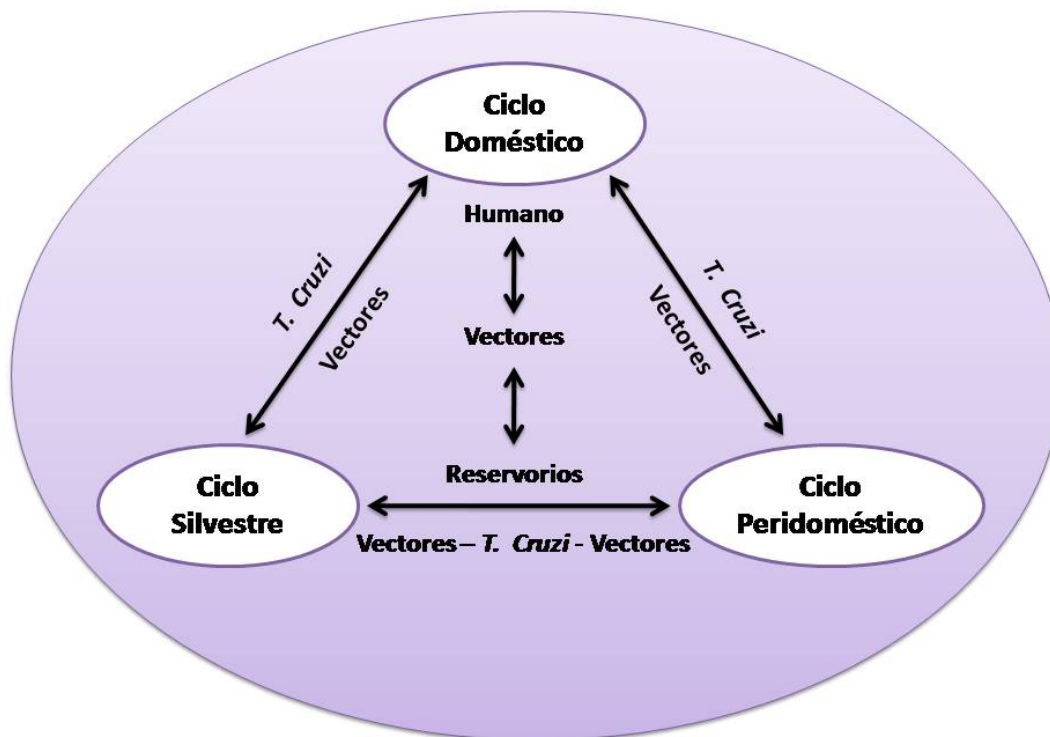


Figura 2.5: Intercambios entre Ciclos de transmisión vectorial del *Trypanosoma cruzi* (Coura y Dias, 2009)

Este ciclo sirve de nexo entre los ciclos doméstico y selvático (Figura 2.5) mediante la acción de los triatomíneos peridomésticos o domésticos, y/o animales del domicilio o peridomicilio; por ejemplo, los perros y gatos cazan animales salvajes que pueden penetrar al domicilio o áreas circundantes, y los animales salvajes, como ratas y zarigüeyas, actúan como reservorios del ciclo silvestre; de esta forma entra al domicilio la infección silvestre (Coura y Dias, 2009). Adicional a esto, se ha observado una gran capacidad de movilidad de especies domiciliadas como *T. dimidiata*, lo que hace que la vinchuca salga del domicilio e interactue con otras especies, vinculando los ciclos de transmisión (Coura y Dias, 2009).

En este punto es destacable, que la transmisión de este parásito se mantuvo evolutivamente de forma natural en el ciclo de transmisión silvestre, convirtiéndose en un problema de la humanidad en el momento en que el hombre empieza a utilizar espacios selváticos, modificando el entorno, lo que llevó a generar escenarios en los que las especies interactúan y se logra generar el intercambio entre estos ciclos de transmisión.

2.7 Vectores

Esta enfermedad es transmitida por unas 120 especies de Triatominos, los cuales son hemípteros, estrictamente hematófagos de la familia Reduviidae, subfamilia triatominae y géneros *rhodnius*, *triatoma* y *panstrongylus*, conocidos con nombres diferentes según el país: pitos, chinches besadores o besucones, chiribicos en Colombia; barbeiros, en Brasil; vinchucas, en Chile y Argentina; chipos, en Venezuela entre otros (Rosas et al., 2007).

Las vinchucas viven, tanto en el ambiente extra, como intradomiciliario del hombre; los más comunes son: *Triatoma infestans*, principal vector desde la línea ecuatorial hacia el Sur, y *Rhodnius prolixus* y *Triatoma dimidiata* al norte de la línea ecuatorial.

La cabeza es alargada y termina en una probóscide recta que durante el reposo se dobla en ángulo agudo contra la parte ventral del cuerpo, que se extiende en el momento de la picadura. Posee un par de ojos prominentes, por delante de los cuales emergen un par de antenas, cuyo punto de implantación sirve para la diferenciación de los géneros. El tórax es quitinoso y su segmento anterior o pronoto tiene forma de escudo. Las alas son dobles y se mantienen dobladas sobre el dorso y pocas especies no tienen alas (Rosas et al., 2007); en general los triatominos son más caminadores que voladores.

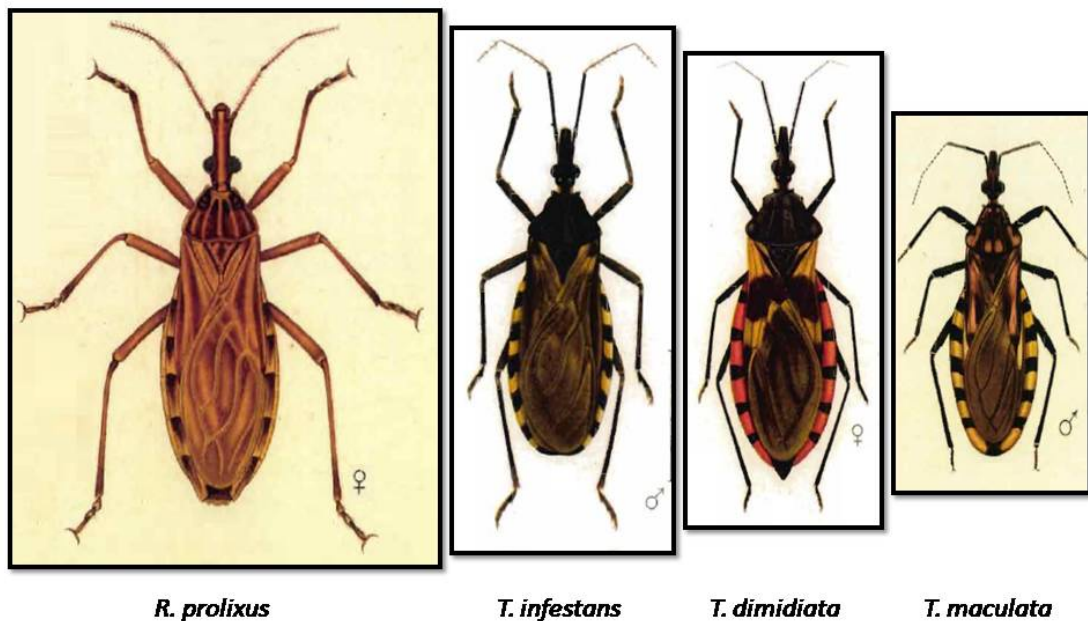


Figura 2.6: Principales triatominos vectores del *Trypanosoma cruzi* en Colombia y Argentina
Jurberg et al. (2004)

Se reproducen mediante huevos y hacen una metamorfosis incompleta pasando por cinco estados antes de llegar a adulto. Cada paso de un estado a otro se hace mudando el exoesqueleto, lo cual ocurre después de una comida completa de sangre. Para que haya un buen desarrollo de los huevos y posteriormente crecimiento de las ninfas, es necesario que exista alimentación, temperatura y humedad adecuadas. El ciclo de desarrollo completo varía con las especies y por lo general dura entre 84 y 134 días. La longevidad del insecto cambia también con la especie y se relaciona directamente con la capacidad de ayuno, la cual para los adultos y las ninfas puede ser de varios meses. Si se cuenta a partir del huevo, la vida oscila entre 300 a 350 días (Rosas et al., 2007).

El macho de *T. infestans* mide de 21 a 26 mm (Figura 2.6), mientras que la hembra puede alcanzar los 30 mm. Presenta manchas amarillas en el corio, abdomen y patas, de textura granulosa y pilosidades cortas y esparcidas. Esta especie está altamente adaptada al domicilio o peridomicilio y es el principal vector de Chagas en los países del cono sur (Argentina, Brasil, Chile, Bolivia, Paraguay, Perú y Uruguay). *T. infestans* es objeto de control directo por parte de las entidades en salud de los países del cono sur. Se considera domiciliada, pero existen poblaciones silvestres arborícolas y rupestres en Bolivia (Jurberg et al., 2004). Este insecto tolera temperaturas extremas y puede ser encontrado hasta las 4.000 m de altura. Su ciclo de vida lleva de 100 a 140 días y los adultos pueden vivir hasta 500 días (Jurberg et al., 2004).

El tamaño de *R. prolixus* varía de los 17.5 a 20 mm en los machos, y las hembras un poco más grandes pueden medir de 19,5 a 21,5 mm (Figura 2.6). Son de color castaño con manchas oscuras en varias regiones del cuerpo. Esta especie actualmente es exclusivamente domiciliada, siendo el vector más importante de la enfermedad de Chagas en su área de distribución que abarca del sur de México, toda América central y parte del norte suramericano (Colombia, Ecuador, Venezuela y Guyana) (Jurberg et al., 2004).

El tamaño de *T. dimidiata* es sumamente variado, de 24 a 35 mm para la hembra (Figura 2.6). Su coloración también tiene grandes variaciones, desde el negro al castaño, presenta manchas amarillas o naranjadas; estas variaciones tanto de tamaño como de color sugieren la existencia de un complejo de especies (Jurberg et al., 2004). Esta especie es considerada un importante vector de la enfermedad de Chagas en México, centroamérica y los países andinos hasta el norte de Perú. Según el área se adapta al domicilio, peridomicilio y ambientes silvestres, tanto terrestres como arborícolas, en donde se alimenta de marsupiales, murciélagos y edentados. Su ciclo de vida lleva 200 días y una hembra adulta puede llegar a poner hasta 680 huevos (Jurberg et al., 2004).

Por su parte el macho de *T. maculata* mide entre 16 y 21 mm, y la hembra entre 17,5 y 22 mm (Figura 2.6); presentan coloración variada de marrón oscuro a negro, con manchas amarillas, naranjas o rojo claro. Esta especie está en Colombia, Venezuela, Guyana y Brasil. Es preferiblemente arborícola en ambientes silvestres (palmas y nidos de aves, invade el peridomicilio y domicilio. Es de gran interés en salud pública porque en Venezuela se considera el principal vector del mal de Chagas (Jurberg et al., 2004).

2.8 Control

La experiencia histórica del control de la enfermedad de Chagas recomienda dos puntos principales para interrumpir la transmisión del *T. cruzi*: (1) la implementación de actividades de control de vectores en las casas con el fin de reducir y luego eliminar la transmisión vectorial y (2) el fortalecimiento de la capacidad de los bancos de sangre para prevenir la transmisión de la enfermedad de Chagas a través de transfusiones de sangre; lo anterior a través del desarrollo e implementación de políticas de estado que aseguren la permanencia en el tiempo del seguimiento a las áreas con mayor riesgo ([Moncayo y Silveira, 2009](#))

Los programas de control en países de América Latina se han centrado en la aplicación de insecticidas en las casas y en los anexos de hogares y edificios. La dieldrina, organoclorados y gamma-HCH (Hexaclorhidrato de benceno) fueron los primeros insecticidas sintéticos utilizados para el control de las poblaciones de vectores en los domicilios; posteriormente se empezaron a usar en mayor medida piretroides sintéticos ([Dias et al., 2002](#)).

Doce países de América tienen programas de control de activos que combinan la aplicación de insecticidas con educación para la salud. El patrón común de los programas de control centralizados, sigue varios pasos operativos o fases ([Moncayo y Silveira, 2009](#)):

- Una fase preparatoria para el mapeo y la programación general de actividades y la estimación de los recursos.
- Una fase de ataque durante la cual se realiza una primera fumigación masiva con insecticida a las casas y es seguido por una segunda pulverización de 6 a 20 meses más tarde, con evaluaciones adicionales para revisar la reinfestación.
- Una fase de vigilancia para la detección de focos residuales de triatomíneos. En esta última fase, la participación de la comunidad y la descentralización de las actividades de control residuales son esenciales.

En la actualidad uno de los mayores desafíos que se afrontan es controlar la reinfestación de las viviendas mediante el control en el peridomicilio.

El objetivo de la eliminación de la transmisión vectorial es más factible en áreas con vectores domiciliados, como *T. infestans* y *R. prolixus*.; sin embargo la enfermedad de Chagas ha estado ligada a la pobreza y a las condiciones de vivienda inadecuada que proporciona un hábitat adecuado para el insecto, tales como techos de paja y material de las paredes de adobe ([World Health Organization, 2002](#)); por lo que la mejora del material de las casas, se ha sugerido y ha tenido éxito en algunos estudios, pero los costos son altos y en las zonas rurales es difícil apuntar a los hogares que se necesitan mejorar, además existen diferencias culturales en los pueblos originarios como hábitos nómadas que dificultan este tipo de intervención.



Figura 2.7: Iniciativas continentales para el control de los principales vectores de la enfermedad de Chagas asociados al hábitat humano (Rosas et al., 2007)

Las iniciativas en Latino América han ido evolucionando con el tiempo; en 1990, la enfermedad de Chagas tomó mas importancia para los programas nacionales de Salud en todo América del Sur. Esto dió lugar a la formación de la Iniciativa del Cono Sur, en 1991, un acuerdo entre los gobiernos de Argentina, Bolivia, Brasil, Chile, Paraguay, Uruguay, y Perú para controlar la enfermedad a través de la eliminación selectiva de *T. infestans* (Dias et al., 2002). Esta iniciativa conllevó a un marcado descenso en las tasas de infección de la infancia en los nacidos después de la iniciativa, así como una reducción a gran escala en el rango de distribución de *T. infestans*; Uruguay y Chile fueron declarados libres de transmisión de Chagas en 1997 y 1999, respectivamente (Dias et al., 2002).

La Figura 2.7 muestra el mosaico epidemiológico en el cual se basaron las diferentes iniciativas continentales de control vectorial. Además, se muestran las áreas geográficas correspondientes a la distribución de los diferentes insectos vectores, así como los países que conforman las iniciativas y sus respectivas fechas de creación.

El modelo implementado en el Cono Sur se adaptó a las iniciativas de los países andinos en 1996 y en América Central en 1997 (World Health Organization, 2002), y más recientemente a la Iniciativa Amazónica, en 2004. Los avances en el control de la enfermedad de Chagas logrados en el período de 1991 a 2006 cambiaron el modelo epidemiológico de la enfermedad.

En Colombia, Ecuador, Perú y Venezuela, la eliminación de la transmisión vectorial se puso en marcha en una reunión intergubernamental celebrada en Bogotá, en febrero de 1997, durante el cual los planes detallados, país por país de acción, incluyendo metas anuales, necesidades presupuestarias, mecanismos de evaluación y necesidades de investigación fueron preparadas ([Moncayo y Silveira, 2009](#)). Los avances realizados en estos países, tanto en el desarrollo de nuevos métodos de evaluación y en las actividades de control, fueron los siguientes:

- El desarrollo de metodologías para la estratificación del riesgo
- La estratificación de riesgo de transmisión vectorial y programas de control en Colombia
- El establecimiento de un programa de control nacional en Ecuador y ejecución de las actividades de estratificación del riesgo
- La implementación de actividades de control regulares en la región sur de Perú
- El restablecimiento de las actividades de control de vectores y la vigilancia entomológica en Venezuela
- Cerca del 100% de la vigilancia en transfusiones de sangre

En Colombia puntualmente, se estimó que un 5% de la población que vive en las zonas endémicas se infecta, lo que equivale a aproximadamente 700.000 personas. Los departamentos con las mayores tasas de infección son Aráuca (21,1%), Casanare (10%), Santander (6,3%), Norte de Santander (5,2%), Boyacá (3,7%), Cundinamarca (1,9%) y Meta (1,7%) ([Moncayo y Silveira, 2009](#)). El tamizaje de bancos de sangre es obligatorio desde 1995 y hay 100% de cobertura en todo el país. Actualmente Colombia cuenta con el Programa Nacional de Investigación para la Prevención, Control y Tratamiento Integral de la enfermedad de Chagas en Colombia 2012-2020 (Red Chagas Colombia), que tiene por objetivo la generación de estrategias científicas para la promoción, prevención, diagnóstico, atención integral y rehabilitación, para la reducción de la carga de enfermedad de Chagas en Colombia ([Red Chagas Colombia, 2014](#)).

Administrativamente está constituida como una Unión Temporal, conformada por 5 instituciones: Instituto Nacional de Salud, la Universidad Industrial de Santander, la Universidad de los Andes, la Fundación Cardioinfantil y la Fundación Cardiovascular de Colombia. Adicionalmente, al programa se integran más de 30 Instituciones, entre universidades, hospitales, centros de investigación, asociaciones, que hacen parte de la red de conocimiento. Dentro de sus principales objetivos se listan ([Red Chagas Colombia, 2014](#))

- Generar una línea de base del conocimiento acumulado en epidemiología, vigilancia, diagnóstico, patogénesis, atención integral y participación comunitaria para la prevención y el control de la enfermedad de Chagas en Colombia, con miras a medir la carga de enfermedad.
- Diseñar, implementar y evaluar el impacto de estrategias integradas de promoción y prevención para el control de la transmisión de la enfermedad de Chagas en Colombia en los diferentes escenarios (vectorial, transfusional, oral y congénito).

- Desarrollar, implementar y evaluar estrategias para mejorar el acceso y oportunidad al tamizaje, diagnóstico y tratamiento etiológico tempranos de la enfermedad de Chagas.
- Desarrollar, implementar y evaluar estrategias para mejorar el pronóstico, rehabilitación y paliación por cardiopatía Chagásica.
- Integrar el conocimiento científico desarrollado a los programas de salud pública.
- Incrementar la capacitación y formación científica en los diferentes componentes del Programa.
- Generar innovación y desarrollo para la promoción, prevención, diagnóstico, tratamiento y rehabilitación de la enfermedad de Chagas.

2.9 Patogénesis

La infección por *T. cruzi* tiene tres etapas, aguda, indeterminada y crónica.

2.9.1 Fase Aguda



Figura 2.8: Complejo oftalmoganglionar o signo de Romana (Coura y Dias, 2009)

La fase aguda es asintomática en la mayoría de las personas infectadas, pero se caracteriza por la inflamación del lugar de entrada del parásito en el cuerpo (Rassi et al., 2010). Los individuos de *T. cruzi* se reproducen dentro de las células y las destruyen. Los parásitos libres invaden otras células que también se rompen y causan una reacción inflamatoria con infiltrado de diferentes tipos de leucocitos. La lesión inflamatoria localizada en la puerta de entrada, es visible como un chancro de inoculación y se conoce con el nombre de chagoma. La inflamación se extiende a los ganglios regionales, se bloquean los canales linfáticos y se produce edema local (Rosas et al., 2007).

Cuando compromete el párpado se conoce como el signo de Romaña (Figura 2.8) es el síntoma más reconocido de la enfermedad y se produce cuando la materia fecal del vector contaminada con el parásito es inadvertidamente frotada en los ojos por la persona (World Health Organization, 2002). Posteriormente, se encuentran parásitos intracelulares en otros ganglios linfáticos y órganos como bazo, médula ósea, corazón, tubo digestivo, suprarrenales, cerebro, ocasionalmente ovarios, testículos y tiroides (Rosas et al., 2007). Las células musculares, células adiposas, células de glía y en general las células del sistema reticuloendotelial sufren destrucción debido al crecimiento y multiplicación de los parásitos. A pesar de esto el índice de mortalidad en la fase aguda es bajo, cerca del 10% (Rosas et al., 2007).

Esta fase dura entre 4 y 8 semanas después de la infección; los síntomas en esta fase se reconocen en el 1-2% de todos los casos (World Health Organization, 2002). Otros síntomas generalizados de esta fase incluyen fatiga, fiebre, esplenomegalia, hepatomegalia, adenopatías, erupciones, anorexia, diarrea y vómitos. En los casos donde se llega a la muerte, es producida principalmente por miocarditis, meningoencefalitis y otras complicaciones como bronconeumonía. Las infecciones en individuos inmunodeprimidos son una excepción notable, y típicamente se presentan con miocardiopatía chagásica aguda o meningoencefalitis (Rosas et al., 2007).

2.9.2 Fase indeterminada

Después de la fase aguda ocurre una respuesta inmune que provoca disminución de la parasitemia y mantiene la infección en algunos focos selectivos. Este período va desde el final de la fase aguda hasta la aparición de los primeros síntomas de la fase crónica.

La etapa indeterminada es completamente asintomática y cualquier síntoma de fase aguda desaparecen por completo (World Health Organization, 2002), a pesar de las alteraciones que se presentan en los plexos parasimpáticos del corazón y del tubo digestivo. Hay poca parasitemia en tejidos (Rosas et al., 2007).

Esta etapa latente tiene una duración media de 10 años; la mayoría de las personas infectadas con la enfermedad permanecen en esta etapa, sin mostrar signos externos de infección. Estos individuos permanecen en el ciclo de la enfermedad actuando como reservorios naturales donde

la transmisión vectorial se sigue produciendo y también puede contribuir a la transmisión de la enfermedad congénita ([World Health Organization, 2002](#)).

2.9.3 Fase Crónica

Diez o más años después de la infección, aproximadamente el 20-30% de las personas infectadas desarrollará los síntomas relacionados con la fase crónica de la enfermedad ([Rassi et al., 2010](#)), se caracteriza por una reducida parasitemia.

La mayoría de las personas que presentan síntomas de esta fase sufren de cardiomiopatía Chagásica, que ofrece una gama de manifestaciones clínicas ([Rassi et al., 2010](#)).

Miocardopatía chagásica crónica es la causa principal de la enfermedad cardíaca y la muerte relacionada con el corazón en las poblaciones pobres y rurales de América Latina ([Rassi et al., 2010](#)). Las primeras manifestaciones de la cardiopatía chagásica implican bloqueo de la rama derecha y alteraciones de la contractilidad ventricular izquierda. Esto es seguido por la taquicardia ventricular, disfunción del nódulo sinusal, bloqueo cardíaco de alto grado, fenómenos tromboembólicos pulmonares y sistémicas y miocardopatía dilatada progresiva e insuficiencia cardíaca congestiva. Lecturas de ECG (Electrocardiogramas) anormales son útiles en la detección temprana de Chagas crónica y miocardopatías en pacientes sospechosos ([Rassi et al., 2010](#)).

En algunas partes de América del Sur, los individuos afectados sufren de mega-síndromes que incluyen la ampliación del esófago y colon, además de la participación de corazón ([Rassi et al., 2010](#)).

2.10 Tratamiento

El tratamiento antiparasitario siempre se indica en la fase aguda de la enfermedad de Chagas, aún en los asintomáticos. Se ha demostrado la eficacia de los compuestos nifurtimox y benznidazol para eliminar tanto las formas circulantes del parásito como las formas tisulares en la fase aguda de la enfermedad. La dosis recomendada de benznidazol es de 5 a 10 mg/kg de peso/día durante 60 días, mientras que la de nifurtimox es de 8 a 10 mg/kg/día durante 60 días ([Rosas et al., 2007](#)).

La alta eficacia del tratamiento antiparasitario en esta fase, resalta la importancia de efectuar una búsqueda activa de casos infectados asintomáticos entre las personas que conviven con los casos sintomáticos confirmados. Sin embargo estos medicamentos son muy discutibles debido a que no cumplen con los criterios de la Organización Mundial de la Salud (OMS) para ser un buen fármaco, no son efectivos en la fase crónica, son poco accesibles, no hay certeza sobre

efectos secundarios o efectos teratogénicos, y conllevan la necesidad de hospitalización para el tratamiento, etc (Coura y Dias, 2009).

2.11 Modelamiento y aplicaciones GIS

El componente espacial tiene un peso de importancia en la estimación del riesgo de transmisión vectorial de la enfermedad de Chagas; por ello, el uso de datos provistos por sensores remotos, los métodos de análisis espacial y el uso de navegadores satelitales en el contexto de sistemas de información geográfica tienen especial interés, ya que ofrecen la posibilidad de realizar detallados análisis con buena cobertura geográfica (Gorla et al., 2005). De igual forma, los datos del sensado remoto permiten realizar modelos predictivos de la distribución geográfica potencial de las especies de vectores con relevancia epidemiológica. Esta aplicación es crucial para la comprensión de las dimensiones geográficas de la transmisión y consecuente riesgo de la enfermedad.

Para el caso particular de la enfermedad de Chagas, existen estudios de cartografía y modelización relacionadas con los vectores en Argentina, Colombia, México y otros países (Parra-Henao et al., 2015; Parra-Henao, 2012; Gorla, 2002; Gorla et al., 2005; Gurgel-Goncalves et al., 2012; Rogers, 2000; Vazquez-Prokopec et al., 2012); sin embargo hace falta incrementar esfuerzos en el tema para lograr un entendimiento más certero del complejo ciclo de transmisión del parásito; que involucre sus vectores y reservorios silvestres y tenga particular énfasis en el proceso de domiciliación.

Los SIG han contribuido a la eficacia de las intervenciones a través del análisis de los datos espaciales sobre la salud, debido a que el uso de estas herramientas permite identificar patrones de distribución espacial de la incidencia o la mortalidad en regiones geográficas definidas (Martins-Melo et al., 2012). Adicional a esto los eventos pueden ser agrupados en el tiempo y de esta forma determinarse posibles causas de los procesos, asistir en el monitoreo, formulación de planes de prevención y toma de decisiones, mediante la construcción y prueba de modelos predictivos basados en inferencia estadística (Graham et al., 2004).

La incorporación de estas herramientas en los programas rutinarios de control y prevención, puede redundar en una mayor y más eficaz focalización de las áreas de riesgo y en una más eficiente identificación de las áreas en las que una intervención de las entidades públicas es de carácter prioritario, y por ende en acciones de control costo-efectivas que redundarán en el bienestar de los pobladores de las zonas endémicas (Parra-Henao, 2010).

En este contexto, el modelado de nicho ecológico es una herramienta que permite la exploración de los fenómenos geográficos y ecológicos sobre la base de ocurrencias conocidas de las especies y se ha aplicado ampliamente para conocer aspectos de la transmisión de la enfermedad de Chagas en la última década, incluyendo la caracterización de los nichos de las especies de vectores, y el análisis de las relaciones entre la distribución de vectores y

reservorios ([Gurgel-Goncalves et al., 2012](#)).

Existe un número de variables ambientales que tienen influencia directa o indirecta sobre la dinámica poblacional de los triatominos vectores de Chagas. Muchas de ellas pueden estimarse a partir de los datos registrados por sensores a bordo de plataformas en órbita espacial. Entre tales variables pueden mencionarse temperatura del aire, temperatura de superficie, índice de vegetación de diferencia normalizada (NDVI), radiación infrarroja media, déficit de saturación de vapor y altura de terreno ([Gorla et al., 2005](#)).

Capítulo 3

Sistema Argentino para el control y estratificación de riesgo de transmisión vectorial del *Trypanosoma cruzi*

3.1 Introducción

Gracias a la interdisciplinariedad de la epidemiología panorámica y su aplicación en el análisis de las enfermedades transmitidas por vectores, se han logrado grandes avances en el control de enfermedades como Chagas. Se plantea como principal objetivo de esta disciplina generar modelos de predicción del riesgo basados en elementos del paisaje y registros tomados en campo ([Porcasi et al., 2012](#)). Es importante y adecuado el aplicar las ideas de la epidemiología en proyectos operativos, que sean de utilidad tanto para las autoridades de vigilancia en salud, como para los investigadores externos.

Bajo este objetivo, en Argentina, el instituto de altos estudios espaciales “Mario Gulich” (CONAE – Universidad Nacional de Córdoba), hace énfasis en la generación de sistemas de alerta temprana en emergencias ambientales usando información espacial proveniente del sensado remoto.

El instituto considera que las enfermedades que afectan al hombre son una emergencia de interés primordial, por lo que se creó un convenio entre el Ministerio de Salud de la Nación (MSAL) y la Comisión Nacional de Actividades Espaciales (CONAE) para crear el sistema *Health Application Project* (HAP). Este programa tiene como principal objetivo desarrollar sistemas operativos que apoyen el plan estratégico para el control vectorial en Argentina, utilizando tecnologías geoespaciales ([Scavuzzo y Torrusio, 2010](#)). En el marco de este proyecto la primera infraestructura de datos que se creó, fue para el manejo del Dengue ([Porcasi et al., 2012](#)).

En particular, este capítulo se centrará en describir el sistema creado para el control y estratificación de riesgo de transmisión vectorial del *T. cruzi*, que actualmente funciona para todo Argentina, con el fin de apoyar las estrategias de vigilancia del Ministerio de Salud en la nación.

El amplio desarrollo de los GIS ha permitido que los usuarios tengan una visualización de la información espacial en un formato adecuado y de fácil manejo. Como resultado, la interpretación de la información se hace cada vez más simple. Sin embargo, no todo el mundo tiene acceso a los GIS, ni puede dedicar el tiempo necesario para utilizarlos de manera eficiente; en respuesta a esta problemática aparecen los Web-GIS.

Los Web-GIS son una forma económica y fácil de difusión de datos geoespaciales y de herramientas de procesamiento. Este servicio satisface la necesidad de muchas organizaciones interesadas en distribuir mapas y herramientas, sin restricción de tiempo ni ubicación de los usuarios.

Con la creación de un servidor web se puede interactuar visualmente con los datos, los clientes pueden producir mapas y gracias a que están en internet, estos nuevos productos pueden ser visualizados por otros clientes, desde cualquier lugar con acceso a la red ([Alesheikh y Helali, 2001](#)). La plataforma web creada en el instituto Gulich cuenta con grandes ventajas a la hora de administrar, compartir y visualizar los datos epidemiológicos, propias de los Web-GIS.

La experiencia del Instituto Gulich en el desarrollo de proyectos con aplicaciones en salud, permitió crear plataformas como la de Chagas que se desarrolló con las mismas características del sistema para la estratificación del riesgo de Dengue en Argentina ([Porcasi et al., 2012](#)). Actualmente se encuentran ambos sistemas operativos y son actualizados en el tiempo con nuevas versiones de las herramientas con las que fueron construidos. Adicionalmente, estos sistemas permanecen en constante implementación de nuevos algoritmos que hacen más robusta la capacidad de análisis de los datos.

3.2 Funcionalidades

Las funcionalidades del sistema responden a los requerimientos generados por un conjunto de usuarios finales del Ministerio de Salud de la nación Argentina, partiendo de la base de que el sistema está en un entorno web amigable, flexible e intuitivo que hace que la visualización y el manejo de la información sea simple para todo tipo de usuario final.

De manera abreviada los requerimientos ya implementados y funcionales en la plataforma de Chagas se presentan a continuación, de acuerdo con lo reportado por CONAE ([Porcasi et al., 2013](#)).

- Las capas de información se cargan de manera automática a diferentes niveles, nación, provincia, departamento, localidad, vivienda.
- Permite que las planillas de Chagas colectadas por los usuarios y las capas que son ingresadas y exportadas sean de diferentes formatos: vector, raster, planillas Excel, csv (valores separados por coma).
- Permite mantener la información de las planillas guardada en una base de datos para generar registros históricos.
- Permite individualizar las viviendas registradas con un ID (Identificador) único.
- Permite gestionar usuarios y permisos (nacionales o provinciales). Un usuario nacional puede visualizar los resultados de todo el país, mientras que los provinciales pueden cargar y visualizar los datos de su provincia; adicional a esto los usuarios solo pueden modificar sus propios datos (antes de 60 días después de cargados).
- El administrador del sistema puede dar de alta a nuevos usuarios en cualquier momento que se lo solicite.
- Las capas de información están georreferenciadas (actualmente el sistema usa el sistema de coordenadas geográficas WGS84).
- Se pueden agregar nuevas localidades (que no se encuentren en la base de datos inicial, creada a partir de los datos obtenidos del Instituto Nacional de Estadística y Censo - INDEC- y el Instituto Geográfico Nacional -IGN-).
- El sistema permite agregar coordenadas geográficas a localidades que carezcan de georreferencia.
- El sistema cuenta con capas de contexto como límites provinciales, ríos, rutas provinciales y nacionales.
- Permite una fácil y rápida visualización del estado de los indicadores. Los mapas de riesgo son generados de forma dinámica según los datos cargados en el sistema, con georeferencia.
- Permite descargar la base de datos según el nivel de permiso del usuario (el usuario nacional puede descargar datos de todas las provincias y el usuario provincial solo los datos de su provincia, los haya, o no, cargado el mismo usuario).
- El sistema permite imprimir los mapas generados en formato pdf o jpg.

Existen requerimientos que en la actualidad no son funcionales en la plataforma, los cuales son propósito de esta tesis para satisfacer las necesidades informadas por los usuarios finales; dentro de las funcionalidades que se adicionarán a la plataforma están:

- Incorporar datos ambientales en formato raster o vectorial a escala provincial
- Realizar búsquedas/consultas SQL (*Structured Query Language*) predefinidas a la base de datos.
- Incorporar capas generadas a partir del cruce de información ambiental y la información en la base de datos.

Adicional a lo anteriormente listado, quedan algunas funcionalidades a implementar que son objeto de próximas actualizaciones de la plataforma propias de la constante evolución en la que permanece en el Instituto Gulich.

- El sistema debe ser capaz de realizar análisis espaciales de cercanía, conectividad de las viviendas a escala rural (aprox. 1:10.000 a 1:100.000).
- Implementar un módulo flexible de consultas de datos accesible a través de internet, que permita realizar análisis epidemiológicos para diferentes escalas espaciales y temporales con los datos recolectados en terreno e ingresados al sistema.
- Exportar los datos obtenidos a través del módulo de consulta en distintos formatos (Excel y/o csv).
- Realizar búsquedas/consultas (SQL) configurables.
- Producir reportes gráficos, cruzando distintas variables (a definir). Ejemplo: gráficos de torta, gráfico de barras, histogramas.

3.3 Tecnologías

La arquitectura básica de un Web-GIS requiere inicialmente de una base de datos o sistema de almacenamiento de información, seguido de diferentes aplicaciones que permiten el intercambio de esta información con el visualizador en la web. Los usuarios finales interactúan con el sistema a través de un navegador web utilizado como cliente.

Existe una aplicación distribuida por la organización OpenGeo, llamada OpenGeo suite (*OpenSource*), que permite la incorporación de diferentes funcionalidades y herramientas para el usuario a la plataforma web. La OpenGeo suite está conformada por cinco componentes que corresponden a las principales tecnologías usadas para implementar el sistema de Chagas. A continuación se describen, siguiendo una secuencia lógica, las herramientas utilizadas desde el almacenamiento de datos hasta el desarrollo de la interfaz web.

3.3.1 Base de datos

PostGIS es un extensor de base de datos espaciales para la base de datos objeto-relacional PostgreSQL. Añade soporte de objetos geográficos para su utilización en SIG.

Para el sistema de Chagas se utilizó PostGIS 8.4.14 sobre PostgreSQL 8.4. El lenguaje que se utiliza para acceder a la información es SQL (Lenguaje de consulta estructurado).

PostGIS no almacena la información espacial en capas SHP (formato de archivo vectorial), como lo hacen los sistemas de archivos convencionales porque esto acarrea desventajas en cuanto a la capacidad de los discos de almacenamiento, tiempo de acceso y descarga de los datos. Por su parte la aplicación almacena la información como objetos dentro de una base de

datos espacial, estos objetos se relacionan entre sí por medio de funciones SQL, que requieren de una especificación de tipo de dato espacial.

De esta forma, en PostGIS la información se guarda como un tipo de dato espacial, por ejemplo, puntos, curvas, polígonos y superficies que tienen ciertas características; existe una jerarquía en el tipo de dato y a medida que se avanza en profundidad dentro de la jerarquía, el tipo de dato hijo va heredando todas las características del tipo de dato parental, de esta forma la información se puede guardar en forma desagregada y permite un manejo óptimo.

De manera similar sucede cuando la información espacial es multidimensional, la información se indexa desde la mínima unidad medible y hereda toda la información de los datos en jerarquías más altas. Por último PostGIS permite generar funciones espaciales para analizar, manipular y establecer relaciones entre las geometrías de los componentes que forman la base de datos espacial (tipo de dato espacial) ([OpenGeo Organization, 2011](#)).

3.3.2 Servidor de aplicaciones

La capa de servidor de aplicaciones es responsable de mediar entre la capa de datos y la capa de interfaz de usuario. Actúa como una puerta de enlace de protocolo, convirtiendo las peticiones que realicen en la interfaz web de usuario, al lenguaje específico para comunicarse con las bases de datos que descansa en el PostGIS. También mantiene las rutinas lógicas, servicios y procesamiento que están disponibles en el Web-GIS ([OpenGeo Organization, 2011](#)).

En el sistema de Chagas se utilizó el servidor de aplicaciones espaciales GeoServer (v.2.1.2) como bloque de construcción para los servicios web espaciales.

Geoserver fue integrado con el componente de cacheado GeoWebCache como un acelerador para los servicios de mapas. GeoWebCache es usado para reducir el tiempo de acceso a datos ubicados en la memoria principal que se utilizan con más frecuencia. Es decir, cuando se accede por primera vez a un dato se hace una copia de este en la memoria caché; luego cuando el microprocesador necesita leer o escribir datos, primero verifica si hay una copia en la memoria caché; si es así, el microprocesador entonces lee o escribe en el caché, que es mucho más rápido que hacerlo sobre la memoria principal.

GeoServer puede leer varios orígenes de datos, generar múltiples formatos de salida, y comunicarse usando múltiples protocolos estándar. Se integra fácilmente en la infraestructura y proporciona una vía de comunicación entre los componentes de softwares nuevos y antiguos.

Por último Apache Tomcat (v.6.0.24) fue utilizado como contenedor web.

Al igual que el servidor web Apache Tomcat, que proporciona un método de acceso HTTP (protocolo de transferencia de hipertexto) para archivos y servicios, GeoServer proporciona un método de acceso HTTP para objetos geoespaciales y consultas sobre esos objetos ([OpenGeo Organization, 2011](#)). GeoServer presenta datos espaciales (tablas en una base de datos PostGIS o archivos en un disco duro) como colecciones de atributos, y permite a los clientes HTTP realizar operaciones sobre los objetos que hacen parte de esas colecciones de datos.

El GeoServer emplea diferentes protocolos estandarizados por el Open Geospatial Consortium (OGC), diseñados para la transferencia de información geográfica desde y hacia el servidor, dependiendo del tipo de protocolo que utilice se pueden acceder a los datos originales (Raw data), atributos geográficos, o imágenes de mapa. Los protocolos usados en la plataforma de Chagas son ([OpenGeo Organization, 2011](#)):

- OGC Web Map Server (WMS) para la recuperación de las imágenes de mapas
- OGC Web Feature Server (WFS) para consultar y recuperar vectores con sus atributos
- OGC Web Coverage Server (WCS) para acceder a las coberturas de imágenes en formato raster. Se puede acceder a los datos de mapa de bits reales almacenados en el servidor.

De igual forma que el GeoServer, GeoWebCache es una puerta de enlace entre el visualizador web y la base de datos, pero funciona para mapas compuestos por téslas (piezas con las que se forma un mosaico) del tipo que están disponibles en Google maps, Bing maps y Open layers. Para acceder a este tipo de mapas se hacen solicitudes en paralelo que generan los mosaicos que son visualizados; como las téslas siempre tienen los mismos límites, y existen algunas que son consultadas con mayor frecuencia, son candidatas para estar guardadas en caché, de esto último se encarga GeoWebCache.

La aplicación recibe una solicitud de diferentes téslas y comprueba si están almacenadas en caché, si tiene una copia almacenada devuelve la respuesta y de esta forma se genera un acceso más ágil a la información; en caso de no estar disponible en el cache, traduce la solicitud mediante un protocolo WMS, WFS o WCS y el GeoServer devuelve la respuesta solicitada ([OpenGeo Organization, 2011](#)).

3.3.3 Interface de Usuarios

Finalmente para implementar las interfaces de usuarios se usaron librerías JavaScript: ExtJs, GeoExt y OpenLayers. A su vez, GeoExplorer fue el visualizador web elegido para la composición y publicación de mapas.

OpenLayers es una biblioteca de código abierto utilizada para mostrar mapas interactivos en los navegadores web. Ofrece un API (*Application Programming Interface*) para acceder a diferentes fuentes de información cartográfica: Web Map Services (WMS), Mapas comerciales (tipo Google Maps, Bing, Yahoo), Web Features Services (WFS), distintos formatos vectoriales, mapas de OpenStreetMap, capas de base personalizadas generadas por motores como GeoServer, MapGuide, Mapserver o ArcServer, etc y proporciona herramientas para la edición y captura de datos ([OpenGeo Organization, 2011](#)).

ExtJS es una librería con gran capacidad para la construcción de interfaces de usuario; permite incluir en el navegador utilidades de escritorio como visualizador, árbol de capas, barra de

herramientas, listas, paneles, tablas, cuadros de diálogo, etc. Por su parte, el set de componentes de GeoExt añade extensiones a ExtJS para vincularlo con los atributos de las capas que proporciona OpenLayers ([OpenGeo Organization, 2011](#)). La principal ventaja de usar GeoExt es que permite personalizar las aplicaciones que los usuarios requieren en el navegador del Web-GIS.

3.4 Arquitectura

En los últimos años, las iniciativas de las organizaciones e investigadores, tanto en epidemiología como en otras áreas muestran gran interés en mantener el flujo de información en la web, debido a que proporciona la ventaja de permitir la interacción de un gran número de usuarios. Una de las formas de implementar un Web-GIS es siguiendo el modelo de coordinación cliente/servidor que ofrece simplicidad en el flujo de datos.

La implementación de este modelo promueve diseños modulares que son flexibles y extensibles; de esta forma cualquier operación de recursos de datos puede ser organizada, integrada y accesible a manera de servicio ([Adler, 1995](#)). Con la capacidad de los Web-GIS de manejar ambientes multi-usuarios y multi-tareas y sumado al crecimiento de internet, la arquitectura comenzó mayormente a ser cliente-servidor. En esta arquitectura las aplicaciones se reparten entre, los proveedores de recursos o servicios (servidores) y los usuarios (clientes). Se logra que cada una de las partes que constituyan una aplicación puedan encontrarse físicamente en distintos lugares (computadores), aumentando no solo la capacidad sino también la necesidad de cómputo, permitiendo la evolución hacia las aplicaciones distribuidas de manera abierta en red.

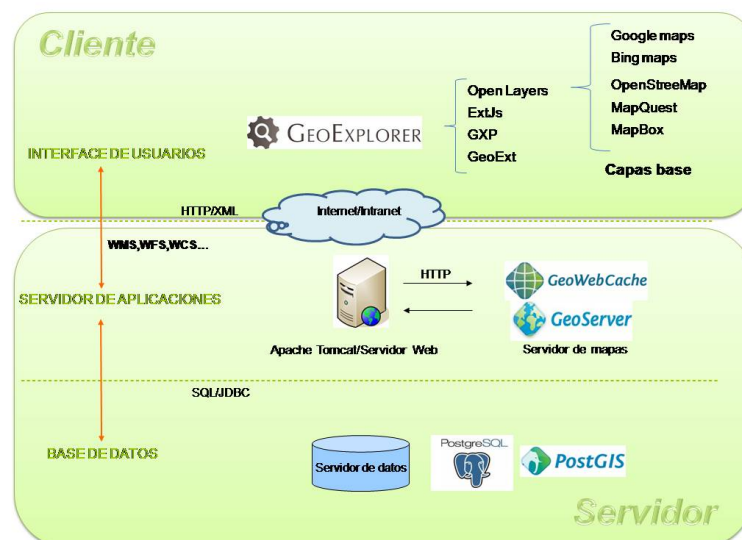


Figura 3.1: Modelo de arquitectura cliente servidor con tecnologías OpenGeo suite y su interacción

Siguiendo el diagrama usual de arquitectura cliente-servidor (Figura 3.1), en la capa interna ubicamos el motor de base de datos PostGIS, en la capa intermedia los servidores de aplicaciones GeoServer y GeoWebCache y en la capa externa las librerías OpenLayers y GeoExt, bases requeridas para el visualizador GeoExplorer.

La comunicación entre los componentes empieza desde el cliente GeoExplorer, en donde los usuarios realizan una solicitud. De inmediato se realiza una comunicación con el servidor de aplicaciones mediante el protocolo HTTP, si la solicitud de la información es alguna capa de información geográfica, entonces la comunicación se realiza mediante uno de los protocolos definidos por el OGC, sea WMS, WFS, o WCS. Por último el servidor de mapas obtiene las capas solicitadas consultando su sistema de archivos de disco o la base de datos (PostGIS/PostgreSQL). Las solicitudes al motor de base de datos se hace utilizando el lenguaje SQL.

Para el desarrollo del lado del cliente se utilizaron principalmente las librerías javascript, ExtJs y GeoExt contentándose con el servidor a través de peticiones Ajax.

Por su parte, para el desarrollo del lado del servidor se utilizaron los lenguajes de programación Java y PHP e internamente consiste en la suma de subsistemas y unidades independientes que se relacionan entre sí.

Los cuatro subsistemas que componen la plataforma del lado del servidor son los siguientes: Traducción de datos (TD), Visualizador (VZ), Base de Datos (BD) y Ejecutor de algoritmos (EA):

- Traducción de datos (TD): Este subsistema es el responsable de la carga de datos, ingresados por el usuario, en el servidor. TD recibe las planillas excel con los datos de campo (.xls o .gxp) y las planillas con las coordenadas geográficas de las viviendas; las procesa, valida los registros y almacena la información en la base de datos. Es decir, mediante este subsistema se logra la automatización de la carga de datos, su unificación y el control de fallos en la carga de los mismos.
- Visualizador (VZ): Este subsistema es el encargado de interactuar directamente con el usuario a través de la implementación de interfaces. Podemos dividirlo en dos partes: una vinculada a las interfaces web relacionadas con los datos planos (planillas de campo) y otra parte vinculada a los datos geoespaciales (mapas). Es decir que, en su totalidad, este subsistema es el encargado de brindar herramientas web para que el usuario pueda acceder, editar y obtener la información almacenada en la base de datos, a través de tablas, gráficos, informes y mapas. En este subsistema se implementa la navegación y consulta de capas de información disponibles y la posibilidad de crear nuevos índices. La funcionalidad de impresión de mapas también está incluida en este subsistema.
- Base de Datos (BD): Este subsistema, es el encargado de almacenar tanto la salida del subsistema TD como la salida del subsistema EA de manera óptima para ser luego usados por el subsistema VZ. En este subsistema se encuentra la base de datos (PostGIS/PostgreSQL). Es el único medio para acceder a la base de datos del subsistema, es capaz de importar y exportar cualquiera de los archivos que se intercambian entre las unidades del subsistema.

- Ejecutor de algoritmos (EA): En este subsistema se definen los módulos de procesamiento de datos. Aquí se implementan los algoritmos que vinculan datos espaciales con las planillas cargadas por los usuarios para generar los índices y las capas base de los modelos predictivos de riesgo. Además este subsistema satisface la necesidad de integrar la información ambiental, es decir que es aquí donde los algoritmos generados en esta tesis para obtener nuevos productos deben ser implementados como nuevas unidades al subsistema.

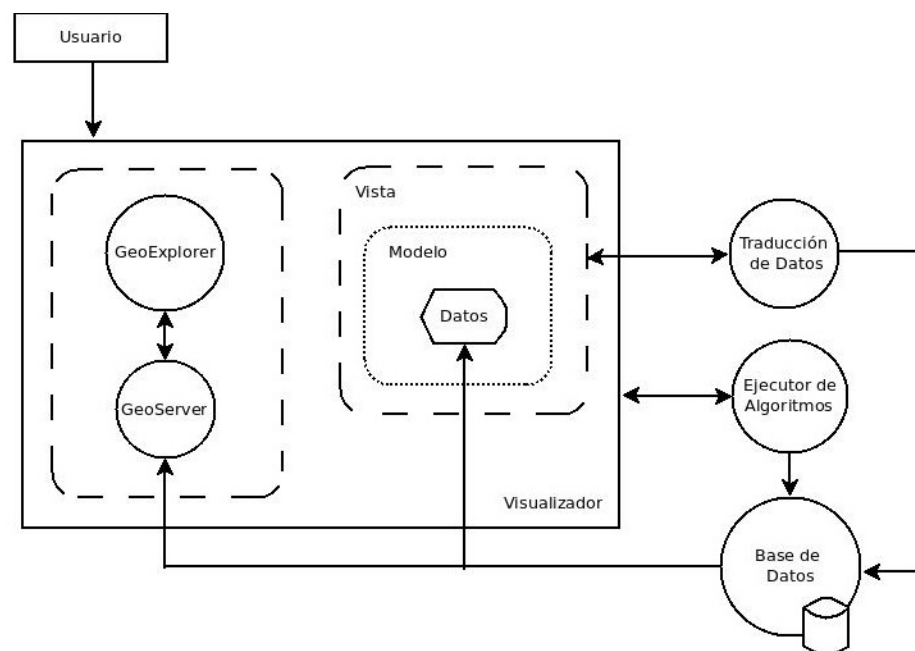


Figura 3.2: Arquitectura de la plataforma Chagas Argentina, interacción entre los cinco subsistemas

Mediante el subsistema TD se logra la automatización de la carga de datos, el control de fallos en la carga de los mismos y en menor medida proporciona flexibilidad en la ejecución de los algoritmos ya que mediante un cambio en la planilla, un algoritmo en particular puede ser modificado sin necesidad de un cambio en su implementación (Peralta et al., 2011).

La figura 3.2) muestra la interacción de los cuatro subsistemas, sus módulos y aplicaciones usadas para construir la plataforma de Chagas bajo la arquitectura cliente/servidor.

Capítulo 4

Cadena de procesamiento de productos raster e incorporación en la plataforma Chagas

4.1 Área de estudio

El área de estudio se sitúa en el noroeste y centro del territorio Argentino, en las provincias de Jujuy, Salta, Catamarca, Tucumán, Santiago del Estero, Chaco y Formosa. Estas provincias forman parte de la bioregión Gran Chaco, que a lo largo de la historia ha experimentado un acelerado deterioro de los ecosistemas y tiene áreas con altos niveles de pobreza. En la región son típicas las viviendas rurales hechas de adobe y paja, e históricamente es una zona hiperendémica de Chagas y otras enfermedades.

Por las características ambientales, sociales y culturales erradicar el vector *T. infestans* de las viviendas y por tanto interrumpir la transmisión vectorial de la enfermedad de Chagas en las poblaciones de la región, se ha convertido en un problema de difícil resolución.

Con fines prácticos y por la importancia de esta región se decidió tomar como zona de estudio el área comprendida entre los 20° y 30° de Latitud Sur y entre los 53° y 64° de Longitud Oeste en la zona norte del área de estudio, y los 57° y 69° Oeste, en el sur de la misma. (Figura 4.1)

El Gran Chaco presenta marcados gradientes climáticos, con temperaturas medias anuales entre 20°C y 28°C; evapotranspiración potencial entre 900 mm en el Sur y 1.600 mm en la frontera entre Paraguay y Bolivia. Este gran ecosistema es clasificado según su clima en tres sub-zonas, siguiendo un eje Noreste-Suroeste: el Chaco Sub-húmedo con lluvias de 1200 a 700 mm; el Chaco Semiárido con precipitaciones entre 750 y 500 mm; y el Chaco Árido con lluvias entre 500 y 300 mm por año en el extremo occidental (Cabrerá, 1976).

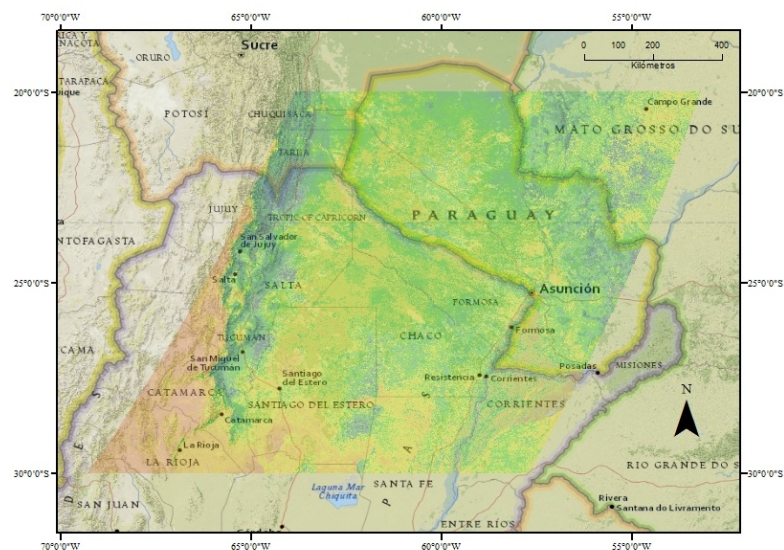


Figura 4.1: Área de estudio

La región presenta gran diversidad de ambientes, con extensas llanuras, sierras, grandes ríos que la atraviesan, sabanas secas e inundables, esteros, bañados, salitrales y una gran extensión y diversidad de bosques muy deteriorados por las actividades humanas. Esta amplia gama de ecosistemas contienen en forma colectiva una diversidad rica en especies y una tasa relativamente alta de endemismo en comparación con otros ambientes áridos, semi-áridos, y sub-húmedos secos (Cabrera, 1976).

Diferentes fuentes de datos provenientes del sensado remoto fueron utilizados para obtener productos que describan el comportamiento de las variables ambientales y de las coberturas presentes en el territorio que cubre el área de estudio. A continuación se describen los productos generados, su fuente de datos y procedimiento llevado a cabo hasta incorporar las capas de información a la plataforma Chagas del Ministerio de Salud Argentina.

4.2 GRASS GIS

GRASS GIS (Geographical Resources Analysis Support System) es uno de los proyectos de software libre aplicados a los sistemas de información geográfica más grandes que hay en la actualidad. Es liberado bajo la Licencia Pública General de GNU (GPL), que garantiza la libre disponibilidad del software. GRASS GIS combina potentes desarrollos integrado en un solo paquete de software para el procesamiento raster, vectorial, y de bases de datos. Incluye mas de 400 módulos con herramientas para el análisis espacial, el modelado, el procesamiento de imágenes y la visualización de las mismas (Neteler et al., 2012).

La licencia GPL que el equipo de desarrollo de GRASS ha adoptado desde 1999 protege a los desarrolladores contra el mal uso de su contribución dentro de proyectos propios que no permitan el libre acceso a su código fuente. La licencia GPL garantiza que todo el código desarrollado debe publicarse de nuevo bajo GPL. Los beneficios que aportan el uso de códigos de otros desarrolladores aumenta la motivación para participar y aportar a este proyecto. La licencia GNU GPL ofrece varias ventajas a los usuarios, no sólo el acceso al código fuente y el bajo costo, sino sobre todo el acceso a las nuevas características y capacidades desarrolladas permanentemente (Neteler et al., 2012).

GRASS proporciona un acceso completo a su estructura interna y algoritmos, por tanto, los usuarios avanzados que quieren escribir sus propios módulos pueden aprender y guiarse de los procedimientos existentes. Los módulos o funciones nuevas pueden ser integradas al GRASS y los algoritmos para automatizar un flujo de trabajo nuevo se escriben con *scripts* de Shell o Python, o en lenguajes de programación de más altos niveles (Neteler y Mitasova, 2008).

El modelo de desarrollo de GRASS es similar a otros proyectos *Open Source*. La columna vertebral del proyecto es internet, que apoya la distribución del software y soporte a usuarios en varias listas de correo y *blogs* de ayuda. El equipo de GRASS incluye desarrolladores de todo el mundo que continuamente mejoran y amplían las capacidades del software (Neteler y Mitasova, 2008).

Existe además la posibilidad de agregar al GRASS extensiones que se denominan *Add-ons* que aún no hacen parte de la versión estándar distribuida por GRASS, pero que son nuevas funcionalidades que incrementan las capacidades y el potencial del software. Cualquier persona puede desarrollar su propio *add-on*; un listado completo de las extensiones actuales con su manual de uso y link de descarga se encuentra disponible en http://grasswiki.osgeo.org/wiki/GRASS__AddOns

4.2.1 Estructura interna del *DATASET* de GRASS

Los datos (raster, vectores, volúmenes, ect) son almacenados en GRASS en un directorio GISDBASE. Este directorio comúnmente llamado GRASSDATA/ debe ser creado antes de empezar a trabajar con GRASS. Dentro del directorio, GRASS crea una estructura interna en la que los datos están organizados por proyectos almacenados en subdirectorios denominados LOCATIONS (Figura 4.2)

Cada LOCATION es definido por su sistema de coordenadas, proyección, y límites geográficos. Los subdirectorios y archivos que definen el location son creados automáticamente cuando GRASS es iniciado por primera vez con un nuevo LOCATION.

Cada LOCATION puede tener varios MAPSETs (subdirectorios de un LOCATION, ver figura 4.2) que son usados para subdividir el proyecto en diferentes temas, subregiones, o espacios de trabajo para diferentes usuarios. Es de resaltar que todos los datos contenidos en un LOCATION deben coincidir en la misma proyección, de otra forma la incorporación será errónea y el software no podrá incorporar los datos a su sistema de archivos. Cuando un usuario accede a su propio MAPSET, puede también ver mapas de otros MAPSET creados por

otros usuarios en el mismo LOCATION, pero únicamente puede modificar a remover datos de su propio MAPSET.

Todos los MAPSETs incluyen un archivo WIND (Figura 4.2) que almacena las coordenadas límite y la resolución del raster o vector en el que se esté trabajando.

Debido a la estructura de archivos que tiene GRASS, cualquier operación como incorporar capas, renombrarlas o eliminarlas, involucra un gran número de archivos internos y siempre debe ser realizado a través de comandos de GRASS, nunca modificando directamente datos sobre la carpeta GRASSDATA.

Los módulos de GRASS están organizados por nombre, de acuerdo a la clase de la función (display, general, imágenes, raster, vector o base de datos). La primera letra hace referencia a la clase de la función, seguida por un punto y una o dos palabras más, que describen los alcances del modulo. Por ejemplo v.in.ogr es el comando para importar los vectores a la estructura del GRASS.

La mayoría de los productos generados en esta tesis fueron desarrollados en GRASS GIS, generando *scripts* en bash con el flujo de trabajo, con la ventaja que pueden ser replicados con gran facilidad por futuras personas.

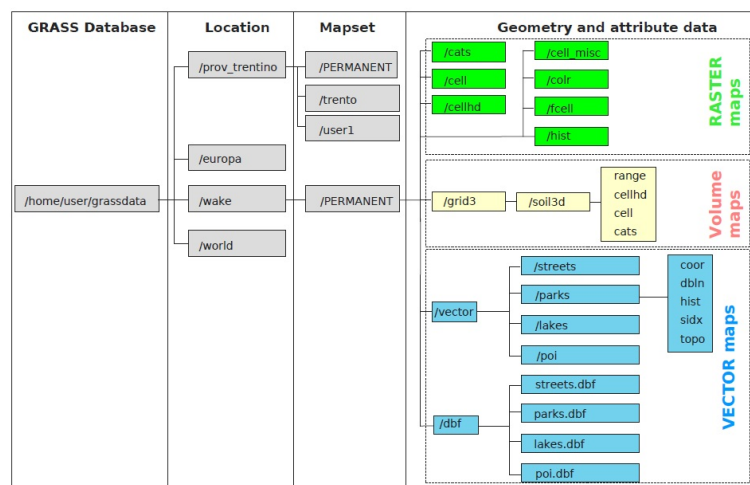


Figura 4.2: Estructura del directorio GRASSDATA (Neteler y Mitasova, 2008).

4.3 Serie temporal de índices de vegetación (NDVI y EVI)

4.3.1 Fuente de Datos

Se entiende por serie temporal a una secuencia de datos u observaciones ordenados cronológicamente y, normalmente, espaciados entre sí de manera uniforme. Considerando el tamaño del área de estudio y la necesidad de contar con información histórica y actual, para el análisis de índices de vegetación se utilizaron imágenes provenientes del sensor MODIS a

bordo de los satélites TERRA y AQUA, que cubren toda la superficie de la tierra en 1 a 2 días. La continuidad de los datos MODIS es vital para relacionar las variables climáticas con los ciclos temporales de las enfermedades y de esta manera estudiar a fondo la dinámica de dichas relaciones. Adicionalmente la resolución espacial de estos datos facilita avances en la obtención de modelos de la distribución del vector (Tatem et al., 2004).

El producto MODIS MOD13Q1 proporciona la posibilidad de comparar consistentemente en espacio y tiempo, las condiciones globales de la vegetación. Este producto puede ser utilizado para monitorear la actividad fotosintética sobre la superficie de la tierra como soporte para análisis fenológicos, detección de cambios e interpretaciones biofísicas. El producto proporciona un *dataset* de información en una grilla mundial que se genera cada 16 días con el mejor de los datos tomados durante ese periodo de tiempo; y permite hacer un preciso seguimiento estacional e interanual de las coberturas de la tierra (United States Geological Survey, 2011).

El formato de archivo en el que se suministra este producto es un HDF4 (*Hierarchical Data Format*) y la proyección geográfica en la que están los datos es sinusoidal.

Tabla 4.1: Capas que conforman el archivo .HDF (dataset) del producto MOD13Q1 (United States Geological Survey, 2011).

Capas del dataset (HDF)	Unidades	No-dato	Rango Válido	Factor de Escala
NDVI	Valor de NDVI	-3000	-2000, 10000	0.0001
EVI	Valor de EVI	-3000	-2000, 10000	0.0001
Calidad de índices (VI-QA)	Bits	65535	0, 65534	No aplica
Reflectancia Rojo	Reflectancia	-1000	0, 10000	0.0001
Reflectancia NIR	Reflectancia	-1000	0, 10000	0.0001
Reflectancia Azul	Reflectancia	-1000	0, 10000	0.0001
Reflectancia MIR	Reflectancia	-1000	0, 10000	0.0001
Ángulo de visión respecto al Zenith	Grados	-10000	-9000, 9000	0.01
Ángulo solar respecto al zenith	Grados	-10000	-9000, 9000	0.01
Azimuth	Grados	-4000	-3600, 3600	0.1
Fecha de toma del dato	Día Juliano	-1	1, 366	No aplica
Fiabilidad del píxel (QA)	Rango	-1	0, 3	No aplica

El dataset que compone al producto MOD13Q1 (Tabla 4.1) es generado con una resolución espacial de 250 m y tiene como primer producto al índice de vegetación de diferencia normalizada (NDVI). Asimismo, proporciona el índice de vegetación mejorado (EVI), el cual tiene mejor sensibilidad sobre regiones con gran biomasa respecto al NDVI. El producto MODIS también proporciona datos de reflectancia en las bandas del azul, rojo, infrarrojo cercano y medio; y datos de calidad del producto.

4.3.2 Adquisición y preprocesamiento de datos

Los productos MODIS son provistos por el Servicio Geológico de los Estados Unidos (USGS) y fueron descargados via http desde el sitio

<http://e4ftl01u.ocs.nasa.gov/MOLT/MOD13Q1.005/>.

Para cubrir toda el área de estudio, se seleccionó el *tile* H12V11 de la grilla de MODIS. El producto MOD13Q1 proporciona una imagen que contiene información de índices de vegetación cada 16 días. (23 imágenes por año). La serie histórica que se descargó son casi once años comprendidos entre el primero de enero de 2004 hasta el 31 de marzo de 2014, un total de 236 productos fueron adquiridos para el análisis temporal.

En congruencia con la finalidad de este trabajo, para cada uno de los procedimientos desde la descarga de la fuente datos hasta obtener el producto y subirlo a la plataforma Web-GIS, fueron utilizadas librerías y programas que permiten implementar cada uno de los pasos. Esto con la finalidad de manejar grandes cantidades de datos, resumir la información y hacer que estos procedimientos sean replicables en el tiempo de una manera más sencilla.

La descarga de los 236 productos que conforman la serie temporal se realizó via http de manera automática utilizando el software "PyMODIS" (Disponible en: <http://pymodis.fem-environment.eu/>). Esta librería de libre acceso (Delucchi, 2014) fue desarrollada como herramienta para descargar en cadena grandes cantidades de imágenes MODIS, conectándose de manera automática a la fuente de datos donde reposan estos productos (<http://e4ftl01u.ocs.nasa.gov/MOLT/MOD13Q1.005/>). Además permite crear mosaicos de varios *tiles* de la grilla MODIS; convertir un archivo HDF4 a GeoTiff y reprojectar datos utilizando internamente la herramienta *MODIS reprojection tools* (Ver: http://lpdaac.usgs.gov/tools/modis_reprojection_tool).

Al descargar la librería PyMODIS se adquieren 7 scripts en python que permiten manipular los datos MODIS, para este análisis solo se utilizó la función `modis_download.py`. A continuación se muestra la forma en la que se ejecutó esta función por línea de comando desde una terminal de Linux.

```
modis_download.py -s MOLA -t h12v11 -f 2014-04-01 -e 2004-01-01 -p MOD13Q1.005 verdes
```

Los parámetros de entrada se refieren al nombre del producto tal cual aparece en la http de MODIS (-p), la carpeta en donde esta el producto (-s), el *tile* a descargar (-t), la fecha de inicio (-f), finalización de la descarga (-e) y la carpeta en donde son guardados los archivos descargados. Una vez termina de correr este script, se crea un archivo.txt que informa todos los productos descargados y un archivo.log que reporta cualquier error que se haya producido. Finalmente se obtienen todos los archivos en formato .HDF y sus respectivos .XML que contienen los metadatos de cada producto.

4.3.3 NDVI y EVI

Un parámetro fundamental para el seguimiento y caracterización de la cobertura de la tierra son los índices de vegetación. Diversos estudios han demostrado la correlación existente entre estos parámetros y el grado de cobertura vegetal del terreno. Los índices de vegetación se calculan

mediante el cociente, píxel a píxel, entre los valores radiométricos contenidos en dos o más bandas de una misma imagen (Chuvieco, 2006).

Se utilizan ampliamente en dos situaciones: 1) para mejorar la discriminación entre dos coberturas con comportamiento reflectivo muy diferente, y 2) para reducir el efecto del relieve en la caracterización espectral de las coberturas (Chuvieco, 2006).

El empleo de los cocientes para discriminar masas vegetales se deriva del comportamiento radiométrico de la vegetación, que muestra un claro contraste en la banda roja del visible (0.6 a $0.7\mu\text{m}$) y el infrarrojo (0.7 a $1.1\mu\text{m}$), permitiendo separar, con relativa claridad, la vegetación sana de otras coberturas (Chuvieco, 2006). Estos índices se han venido utilizando desde los comienzos de la teledetección por satélite y se han propuesto multitud de ellos, debido a que se obtienen fácilmente a partir de los datos adquiridos por los sensores espaciales. En este estudio nos centramos en el uso de los índices NDVI y EVI.

A continuación se muestran las expresiones que permiten calcular los índices de vegetación NDVI y EVI, en ellas R hace referencia al valor de reflectividad registrado por el sensor para cada banda espectral.

Índice de Vegetación de Diferencia Normalizada (NDVI)

$$NDVI = \frac{R_{Infrarojo} - R_{rojo}}{R_{Infrarojo} + R_{rojo}} \quad (4.1)$$

Índice Mejorado de Vegetación (EVI)

$$EVI = G \times \frac{R_{Infrarojo} - R_{rojo}}{R_{Infrarojo} + C_1 \times R_{rojo} - C_2 \times R_{azul} + L} \quad (4.2)$$

Donde, C_1 y C_2 corresponden a coeficientes de resistencia atmosférica y L a un factor de corrección.

El EVI fue desarrollado para optimizar la señal de la vegetación con sensibilidad mejorada para altas densidades de biomasa, teniendo en cuenta el efecto atmosférico. Los valores de estos índices fluctúan entre -1 y 1, lo que facilita notablemente su interpretación. En términos generales cuanto mayor sea el resultado del cociente en el píxel, mayor vigor de la vegetación presenta la zona observada. Se puede señalar como umbral para identificar coberturas vegetales un valor de NDVI en torno a 0.1 y para la vegetación densa entre 0.5 y 0.7 (Chuvieco, 2006).

4.3.4 Incorporación de índices verdes al GRASS GIS

Para cada una de las fechas descargadas con los productos MOD13Q1, las escenas fueron convertidas desde la proyección sinusoidal original al sistema de coordenadas geográficas

lat/long WGS84; y posteriormente se incorporaron al GRASS.

Usualmente el comando con el que se importa al GRASS cualquier capa de datos raster es `r.in.gdal`, sin embargo, como se explicó con anterioridad, los productos MODIS son adquiridos en formato HDF, y el comando `r.in.gdal` tienen dificultades para extraer las capas del dataset MODIS. Adicionalmente para poder incorporar los datos en proyección sinusoidal al GRASS era necesaria la creación de un LOCATION con los datos de la proyección y no viene incorporada por defecto en el software. Para evitar estas dificultades los desarrolladores de GRASS crearon un módulo *Add-on* llamado `r.modis`.

`r.modis` es un set de herramientas para importar datos provenientes del sensor MODIS, usa la librería `PyMODIS` y *MODIS Reprojection Tool* para convertir y procesar los archivos HDF. Usando este modulo, la función `r.modis.import`, en un solo paso, extrae las capas de los archivos HDF, las transforma a TIFF, las reproyecta y finalmente incorpora las imágenes al GRASS. A continuación se muestra la manera en la que se usó la función para incorporar todos los índices de vegetación al GRASS.

```
#!/bin/bash
#$1: path a los hdf
#$2: path al MODIS reprojection tool

for i in `ls $1/*.hdf`
do
nome=`basename $i .hdf`
r.modis.import dns=$nome mrtpath=$2 spectral="(1,1,1,0,0,0,0,0,0,0,0,0)"
done
```

La función permite hacer una selección de las capas del dataset MODIS que se quieren incorporar, en este caso fueron incorporadas las capas 1, 2 y 3 que corresponden al NDVI, EVI y calidad del índice (Tabla 4.1).

Una vez terminado el comando, usando una sentencia general de GRASS, se pueden listar las capas que ya están dentro del MAPSET, con las cuales se puede empezar a trabajar. A continuación se observa el comando `g.list` para listar, y unas de las capas de índices verdes incorporadas.

```
g.list rast pat=MOD13*
...
MOD13A2.A2004001.h12v11.005.2007235190930.1_km_16_days-EVI
MOD13A2.A2004001.h12v11.005.2007235190930.1_km_16_days-NDVI
MOD13A2.A2004001.h12v11.005.2007235190930.1_km_16_days-VI-Quality
MOD13A2.A2004017.h12v11.005.2008224122331.1_km_16_days-EVI
MOD13A2.A2004017.h12v11.005.2008224122331.1_km_16_days-NDVI
MOD13A2.A2004017.h12v11.005.2008224122331.1_km_16_days-VI-Quality
...
...
...
```

4.3.5 Producto 1: Descriptivos anuales de la serie temporal de índices verdes

Los primeros productos que se generaron fueron los estadísticos descriptivos de toda la serie de tiempo, tanto para el NDVI como para el EVI. Se calculó el NDVI/EVI mínimo, máximo, y una media de la serie temporal. A continuación se describe la implementación en GRASS de cada uno de estos productos.

Para evitar errores y generar información que corresponde a los valores nulos de los productos, GRASS permite realizar una máscara que queda fija, una vez que se fabrica la capa denominada MASK, queda presente para todos los cálculos posteriores que se realicen en el mismo MAPSET. A continuación se muestran los comandos necesarios para identificar los datos nulos que en el producto MOD13Q1 son -3000 y posteriormente la fabricación de la máscara. Es de resaltar que la generación de la máscara con los datos nulos es de suma importancia, debido a que si se omite este paso los valores nulos que contienen -3000 serían incluidos en los cálculos de la media, el mínimo, y máximo, lo cual llevaría a resultados equívocos.

```
#!/bin/bash
#Null
r.null MOD13A2.A2014081.h12v11.005.2014098023307.1_km.16_days_NDVI setnull=-3000

#Mask
r.mask MOD13A2.A2014081.h12v11.005.2014098023307.1_km.16_days_NDVI
```

El siguiente paso consiste en un escalado de los índices. El rango de valores de las capas de NDVI y EVI del producto MOD13Q1 fluctúa entre -2000 y 10000 (Ver tabla 4.1). Debido a que normalmente el rango de valores para el índice NDVI fluctúa entre (-1,1) y en el producto MODIS dichos valores están re-escalados en 1 por 10000, se multiplicó el índice por el factor de escala 0.0001, mediante el uso del *mapcalculator* de GRASS.

```
#!/bin/bash
#escalado
for i in `g.mlist rast pattern=MOD13*NDVI`
do
r.mapcalc --o expression="$i=$i/10000.0"
done
```

Una vez escalados los datos, la serie temporal se encuentra preparada para estimar los descriptivos anuales. La importancia de estos productos es que de una manera sencilla se puede obtener información de la variabilidad inter-anual e intra-anual de los índices, y se resume la información. De esta forma, una serie temporal de 11 años con imágenes cada 16 días, puede ser interpretada con 3 productos por año (Mínimo, Máximo y Media).

Para calcular los descriptivos se usó la función *r.series* de GRASS, que permite aplicar diferentes cálculos que están definidos internamente en la función a una lista de imágenes. A continuación se muestra el uso de la función *r.series* para obtener los descriptivos.

```
#!/bin/bash
#Descriptivos anuales
```

```

for p in `seq 2004 2014`
do

r.series input='g.mlist rast pattern=MOD13*ASp*NDVI sep=', ' /
output=ndvi_mean_$p,ndvi_min_$p,ndvi_max_$p /
method=average ,minimum,maximum,stddev

echo "done $p"
done

```

La función `r.series` recibe como parámetro de entrada un *input* que para este caso son todas las imágenes de un año; además la función tiene varios métodos implementados, en este caso se usó la media, el mínimo y el máximo. Este procedimiento está en un ciclo *for* que permite que se replique el proceso para los 11 años de la serie. Finalmente todo el ciclo realizado con los productos de NDVI se repite para los de EVI.

Una vez obtenidos los raster con los descriptivos anuales se procedió a exportar del GRASS las imágenes generadas. Este paso es importante debido a que los productos deben estar siempre en formato GeoTiff y con el sistema de coordenadas geográficas Lat/Long WGS84, siendo estas las especificaciones para la subida de los productos al GeoServer. Todos los archivos son llevados a una carpeta de archivos de la cual serán tomados los raster para la posterior incorporación y publicación en la plataforma Web-GIS de Chagas que será explicada posteriormente.

Los raster fueron exportados utilizando la función básica de GRASS para crear archivos en formato GeoTiff. A continuación se muestra el uso de la función `r.out.gdal` para exportar del GRASS todos los raster creados con los descriptivos anuales de la serie temporal.

```

#!/bin/bash
#exportar a tiff
for map in `g.mlist rast pat=ndvi*`
do r.out.gdal input=$map output=/home/cursos/DBH/productos_chagas/$map format=GTiff
done

r.mask -r

```

Las figuras 4.3, 4.4 y 4.5 muestran los descriptivos de la serie temporal del NDVI, raster generados y exportados para la subida a la plataforma Web-GIS de Chagas. Por su parte las figuras 4.6, 4.7 y 4.8 muestran los raster generados con los descriptivos de la serie temporal de EVI.

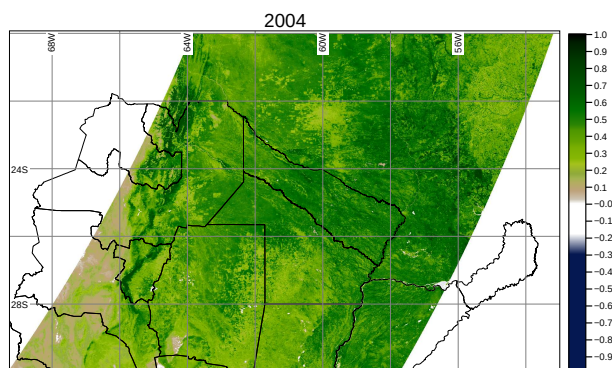
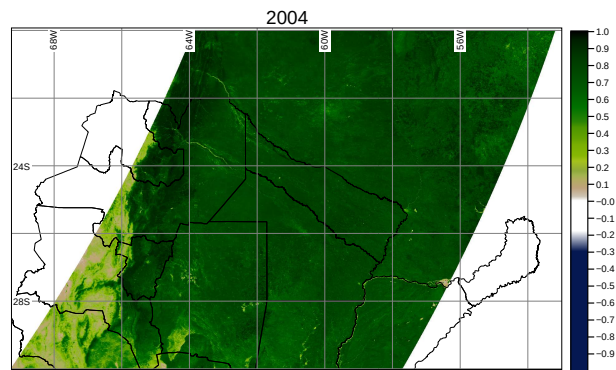
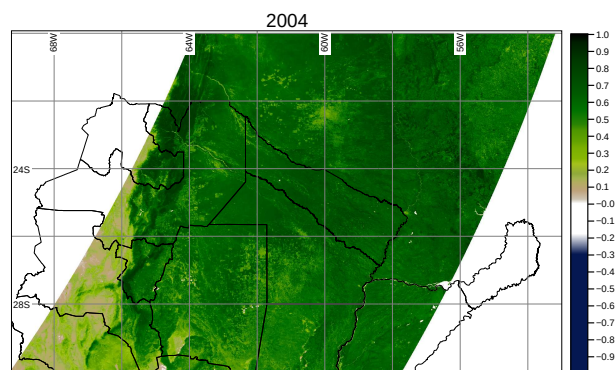
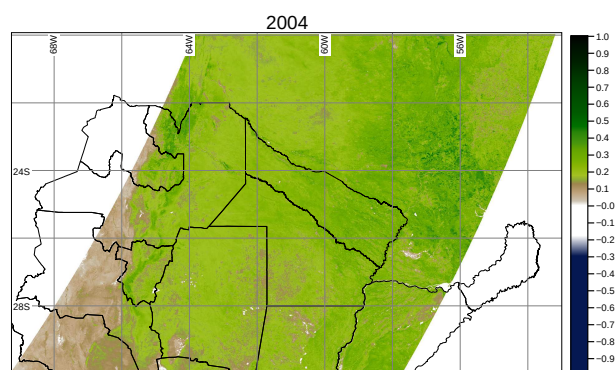
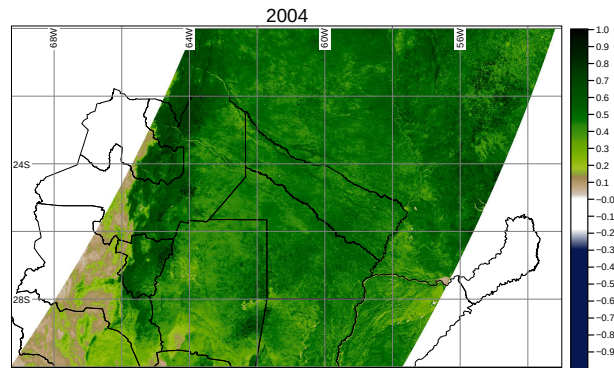
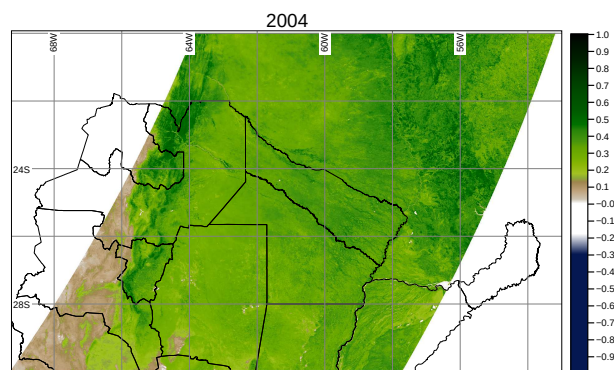


Figura 4.3: NDVI mínimo anual para el año 2004

**Figura 4.4:** NDVI máximo anual para el año 2004**Figura 4.5:** NDVI medio anual para el año 2004**Figura 4.6:** EVI mínimo anual para el año 2004

**Figura 4.7:** EVI máximo anual para el año 2004**Figura 4.8:** EVI medio anual para el año 2004

La figura 4.9 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de los descriptivos anuales de la serie temporal de NDVI y EVI. En la que se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron en GRASS hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

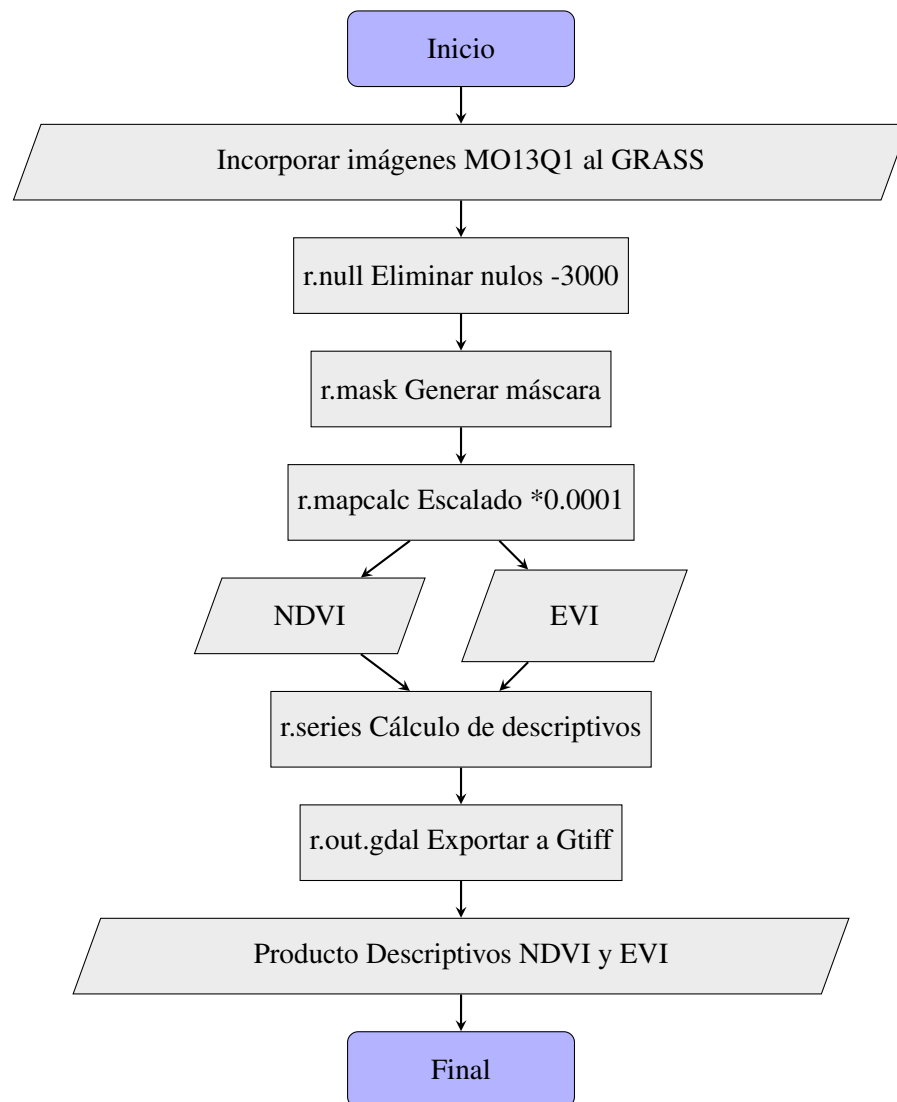


Figura 4.9: Flujo de trabajo para la obtención del producto descriptivos anuales de la serie temporal de NDVI y EVI

4.3.6 Producto 2: Anomalía anual de NDVI y EVI

El segundo producto a partir de la serie temporal de índices verdes consiste en el cálculo de la diferencia (anomalía) anual con respecto a la media de toda la serie.

Este producto permite identificar las zonas y los años en los que los índices se comportaron diferente a lo "usual" en los 11 años incluidos en este estudio, generando dos imágenes por año, una del índice NDVI y otra para el índice EVI.

Para obtener una capa anual con la anomalía de los dos índices, se utilizó la función `r.series` de GRASS, anteriormente explicada y utilizada. A continuación se muestra el uso de esta función, que para este caso recibe toda la lista de los productos de NDVI/EVI de los 11 años y calcula un promedio global; una vez calculada la media global, mediante la implementación de un ciclo *for* se toman las medias anuales que se generaron en el producto 1, y mediante la función

r.mapcalc se resta la media global a cada una de esas capas de la siguiente forma.

```
r.series input='g.mlist rast pattern=MOD13*ndvi* sep=,' /
output=ndvi_global method=average

for i in 'g.mlist rast pattern=ndvi_mean*'
do

r.mapcalc expression="delta_$i= $i-ndvi_global"

done
```

En la anterior implementación es notable la capacidad de GRASS para agilizar el trabajo del usuario, con tan solo un par de líneas y el uso de funciones ya establecidas se puede resumir una cantidad muy importante de datos y obtener productos que expresen la variabilidad de los mismos.

Una vez generados los productos se procede con la exportación del GRASS y generación de los archivos .tiff para su posterior incorporación al GeoServer. Al igual que en los productos anteriores, se usó la función r.out.gdal para archivar los 22 productos (11 años por índice) con las anomalías. De la siguiente forma se realizó la exportación de las capas.

```
for map in 'g.mlist rast pat=delta_ndvi*'
#for map in 'g.mlist rast pat=delta_evi*'
do r.out.gdal input=$map output=/home/curso/DBH/productos_chagas/$map format=GTiff
done
```

Las figuras 4.10 y 4.11 muestran la anomalía para los índices de vegetación NDVI y EVI respectivamente, se observa las zonas en el área de estudio en donde el año 2008 tuvo valores inusuales (bajos y altos).

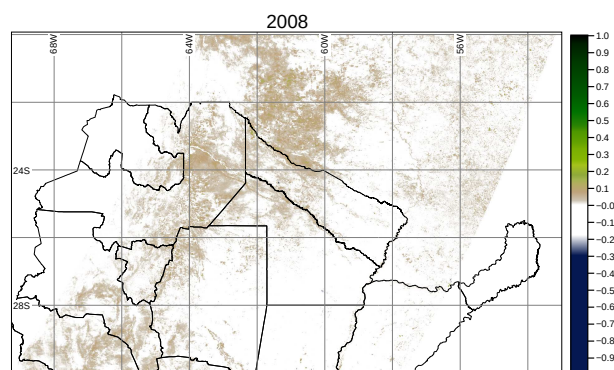


Figura 4.10: Anomalía de NDVI del año 2008 con respecto a la media del periodo 2004-2014

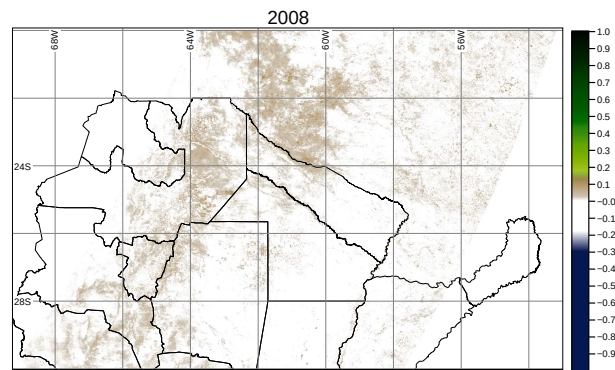


Figura 4.11: Anomalía de EVI del año 2008 con respecto a la media del periodo 2004-2014

La figura 4.12 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de anomalía anual de la serie temporal de NDVI y EVI. Se describen los pasos desde la incorporación de las imágenes al GRASS GIS, a pesar de estar omitidos en la descripción del producto, debido que una vez incorporados los productos en el GRASS, se generaron todos de manera consecutiva. Se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron en GRASS hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

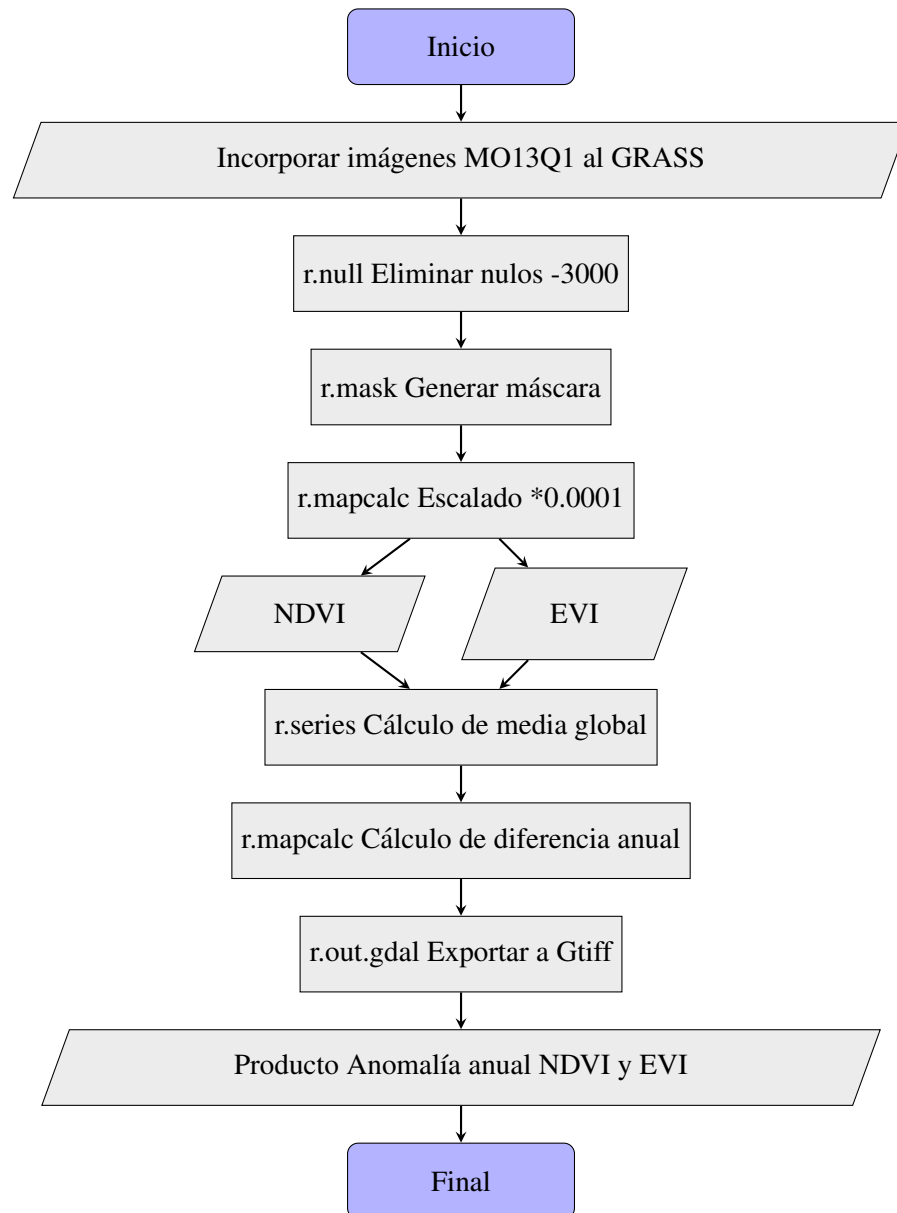


Figura 4.12: Flujo de trabajo para la obtención del producto anomalía media anual de la serie temporal de NDVI y EVI

4.3.7 Producto 3: Extracción de ciclos temporales mediante análisis de series de Fourier

El análisis de Fourier ha sido ampliamente utilizado en el procesamiento digital de imágenes para mejorar y analizar imágenes de una sola fecha, aplicando el método en dos dimensiones para mejorar la calidad de la misma o para analizar variaciones espaciales; más recientemente se incorporó este método para analizar conjuntos multitemporales de imágenes igualmente espaciadas en el tiempo (Jakubauskas et al., 2001).

Las series temporales vistas píxel a píxel a lo largo de la historia se comportan como una señal; el conjunto de datos temporal es una matriz de tres dimensiones, en los ejes X y Y tenemos el espacio y en el eje Z el tiempo, por lo cual cada coordenada X,Y a lo largo de Z es una señal

diferente. Para el caso de los índices de vegetación la variabilidad de esta señal dependerá de la cobertura. Es posible identificar y analizar ciclos a los largo del tiempo en cada una de las coberturas mediante el análisis de series armónicas de Fourier.

El análisis de series de Fourier descompone una señal en una serie infinita de componentes armónicos (ondas); cada uno de esos componentes inicialmente son ondas que siguen la función seno y, a su vez una onda coseno con igual frecuencia. Ambas ondas se combinan en una sola onda coseno, que tiene como característica intrínseca su propia amplitud (tamaño de la onda) y un ángulo de fase (Jakubauskas et al., 2001). El valor de la amplitud es la mitad de la altura de la onda, y el ángulo de fase (o simplemente fase) define el desplazamiento entre el origen y el pico de la onda sobre el rango de $0-2\pi$ (Figura 4.13). Cada término designa el número de ciclos completos que tiene la onda en un intervalo definido; por ejemplo en la parte b de la figura 4.13, el segundo término completa dos ciclos; al sumar los términos que son simplificados por la descomposición de Fourier se obtiene una curva compleja, como se muestra en la parte c de la figura 4.13, donde cada componente de la curva (término), representa un porcentaje del total de la variación de la señal original (Jakubauskas et al., 2002)

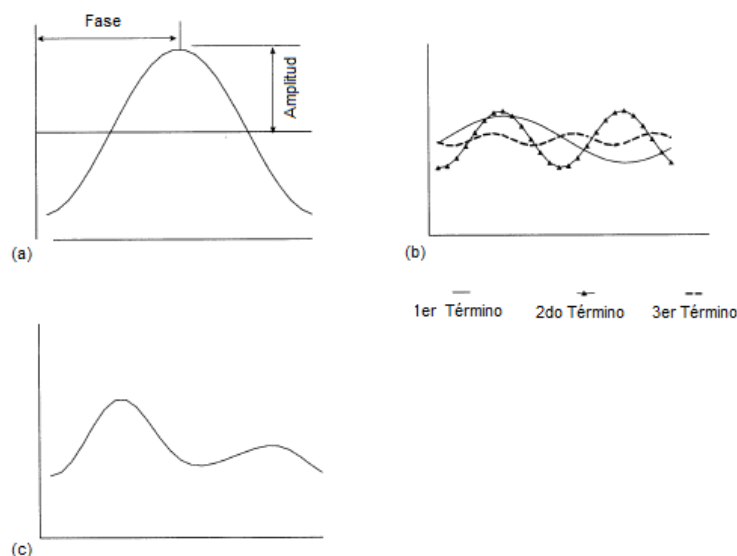


Figura 4.13: (a) Curva del coseno que representa el primer armónico; (b) curvas de los armónicos 1, 2, y 3; (c) curva generada por la suma de los armónicos de (b). (Jakubauskas et al., 2002)

Esta descomposición es de suma utilidad para identificar ciclos estacionales e intraestacionales; para series temporales de vegetación, el análisis de Fourier ofrece grandes promesas para identificar cambios de cobertura a lo largo del tiempo en un análisis en tres dimensiones, o cambios de uso del suelo distribuidos espacialmente en un análisis en dos dimensiones. En el contexto de esta tesis, y de diferentes trabajos relacionados con biodiversidad, su utilidad es notable, debido a que identificar estos ciclos en las coberturas puede relacionarse directamente con la presencia o ausencia de diferentes especies de interés.

Adicional a la identificación de ciclos temporales, este análisis permite, al igual que un análisis de componentes principales, resumir grandes cantidades de información; pero con la ventaja de no verse afectado por valores extremos en los datos, y la información generada puede ser

relacionada estrechamente con datos biológicos debido a que se pueden identificar ciclos incluso mensuales, facilitando la interpretación en términos biológicos (Rogers, 2000).

4.3.7.1 Algoritmo r.hants

Se realizó un análisis de serie de Fourier a la serie completa de NDVI y EVI con la doble motivación de obtener la variación temporal de las coberturas presentes en el área de estudio, y de resumir la serie de tiempo para ofrecer un producto con significado biológico que no represente una gran cantidad de datos para el usuario.

Se utilizó la implementación del algoritmo "Análisis armónico de series temporales (HANTS)", realizada por Markus Metz, incluida como add-on en la version 7 de GRASS GIS <http://grass.osgeo.org/grass70/manuals/addons/r.hants.html>.

HANTS fue desarrollado con la finalidad de manejar series temporales de observaciones irregularmente espaciadas (Roerink et al., 2000). Las principales capacidades del algoritmo son identificar y remover observaciones contaminadas por nubes, y hacer una interpolación temporal de las observaciones para reconstruir imágenes completas a intervalos preestablecidos de tiempo.

El algoritmo calcula una serie de Fourier para modelar las series de tiempo en cada píxel que compone la imagen en Z (a través del tiempo), e identifica valores extremos (*outliers*). Elimina esos *outliers* y los reemplaza con el valor (suavizado) dado por la serie de Fourier. Además, HANTS considera sólo las frecuencias más significativas esperadas en el perfil temporal y aplica un ajuste por cuadrados mínimos basado en componentes armónicos (senos y cosenos). Para cada frecuencia, de manera iterativa, se determina la amplitud y la fase de cada componente (Roerink et al., 2000).

Los datos entrantes con desviaciones positivas o negativas muy grandes respecto a la curva son removidos, asignándoles un valor de cero. Después de recalcular los coeficientes sobre la base de los puntos que quedan, el procedimiento se repite hasta que el error máximo sea aceptable o el número de puntos restantes sea muy pequeño (Roerink et al., 2000).

El ajuste de la curva es controlado por los siguientes parámetros de entrada que deben ser dados a la función r.hants (Roerink et al., 2000):

- Número de frecuencias (NOF): El usuario define cuantas frecuencias son usadas y qué tan largo es su período en unidades de tiempo muestrales. Esto resulta en una curva descrita por $2 \times \text{NOF} - 1$ componentes (fase y amplitud para cada frecuencia, menos la frecuencia cero (media) que no tiene fase)
- Indicador de supresión de valores altos y bajos: este parámetro indica los valores umbrales (máximo y mínimo) a partir de los cuales serán considerados *outliers* y eliminados durante el ajuste.

- Umbral de rechazo de datos inválidos: valores de la imagen que se consideran datos no válidos.
- Tolerancia de error en el ajuste (FET): Representa la diferencia absoluta entre los puntos y la curva ajustada. Esta diferencia se estima en cada iteración y el proceso termina cuando la diferencia es menor a la FET.
- Grado de sobre-determinación (DOD): Número de observaciones válidas, debe ser siempre mayor o igual al número de parámetros que describen la curva. Para obtener un ajuste más confiable el usuario puede decidir usar más puntos que el mínimo necesario.

Este algoritmo ofrece mayor flexibilidad en la elección de las frecuencias y la longitud de la serie de tiempo a un costo computacional en tiempo y memoria no tan elevado, en comparación con la transformada rápida de Fourier (FFT) implementada en otros GIS o softwares de procesamiento de imágenes, e incluso en el GRASS. Sin embargo los parámetros de entrada anteriormente listados, no se encuentran objetivamente descritos y la elección correcta de los mismos va ligado a experiencias previas o a diferentes corridas del algoritmo para comparar diferentes salidas.

Para la extracción de los ciclos temporales de la serie de índices de vegetación, fue necesario experimentar con varias combinaciones de los parámetros, para finalmente aplicar la función de la siguiente forma para los dos índices:

```
for p in `seq 2004 2013`
do
    for i in `g.mlist rast pattern=MOD13*ASP*NDVI sep=,`
    do
        r.hants input=$i nf=5 fet=0.05 idrt=+0.7 amplitude=amp_ndvi_$p \
        phase=phase_ndvi_$p
        echo "done $p"
    done
done
```

Como se observa en el código anterior, la función `r.hants` recibió como entrada todas las imágenes de NDVI/EVI de un año, se calcularon 5 frecuencias ($nf=5$), por lo cual el resultado son 5 amplitudes y 4 fases por año; el error absoluto de los residuos debió ser menor a 0.05 unidades de NDVI ($FET=0.05$) respecto a la curva obtenida; y los valores de NDVI mayores de 0.7 fueron rechazados bajo el supuesto de ser datos afectados por cobertura nubosa o sobre-estimados por el sensor ($idrt=+0.7$).

Una vez obtenidos los componentes armónicos de cada uno de los índices de vegetación se procedió con la exportación de los mismos del GRASS, generando productos en formato GeoTiff, para su posterior subida al servidor. De las cinco frecuencias calculadas, se exportaron únicamente las frecuencias 1, 2 y 4, las cuales se corresponden con los ciclos anual, semestral y trimestral. La exportación se realizó de la siguiente forma:

```
#exportar a tiff las frecuencias anual(1), semestral(2), y trimestral(4)

for map in `g.mlist rast pat=phase*ndvi*[1,2,4]`
do
    r.out.gdal input=$map output=/home/curso/DBH/productos_chagas/$map format=GTiff
done

for map in `g.mlist rast pat=amp*evi*[1,2,4]`
```



```
do r.out.gdal input=$map output=/home/curso/DBH/productos_chagas/$map format=GTiff
done
```

La figura 4.14 representa los componentes de amplitud del armónico de Fourier; es una composición RGB (*Red-Green-Blue*). El canal rojo representa las zonas donde las coberturas tienen ciclos anuales predominantemente; en verde las coberturas con ciclos semestrales, y en azul las zonas que se rigen por ciclos trimestrales; dentro del área se pueden encontrar diferentes combinaciones de estos componentes (amarillo, cian y magenta), debido a las diferencias en el uso de la tierra y condiciones climáticas. Se observan pequeñas áreas de color negro, en donde el suelo no está cubierto por vegetación, por lo tanto no hay un cambio en el ndvi a lo largo del tiempo.

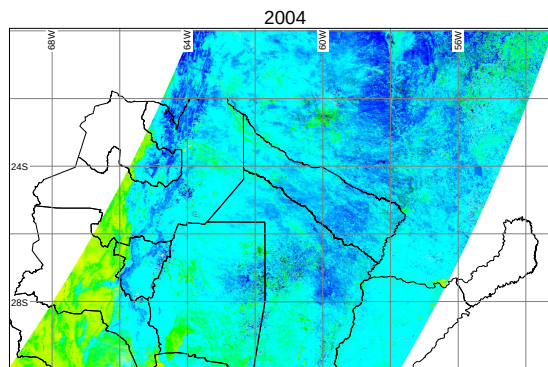


Figura 4.14: Mapa de HANTS con los componentes de Fourier de NDVI del 2004: Rojo, ciclos anuales. Verde, ciclos semestrales. Azul, ciclos trimestrales

La figura 4.15 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de ciclos temporales mediante el análisis armónico de Fourier, se describen los pasos desde la incorporación de las imágenes al GRASS GIS, a pesar de estar omitidos en la descripción del producto, debido que una vez incorporados los productos en el GRASS, se generaron todos de manera consecutiva. Se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron en GRASS hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

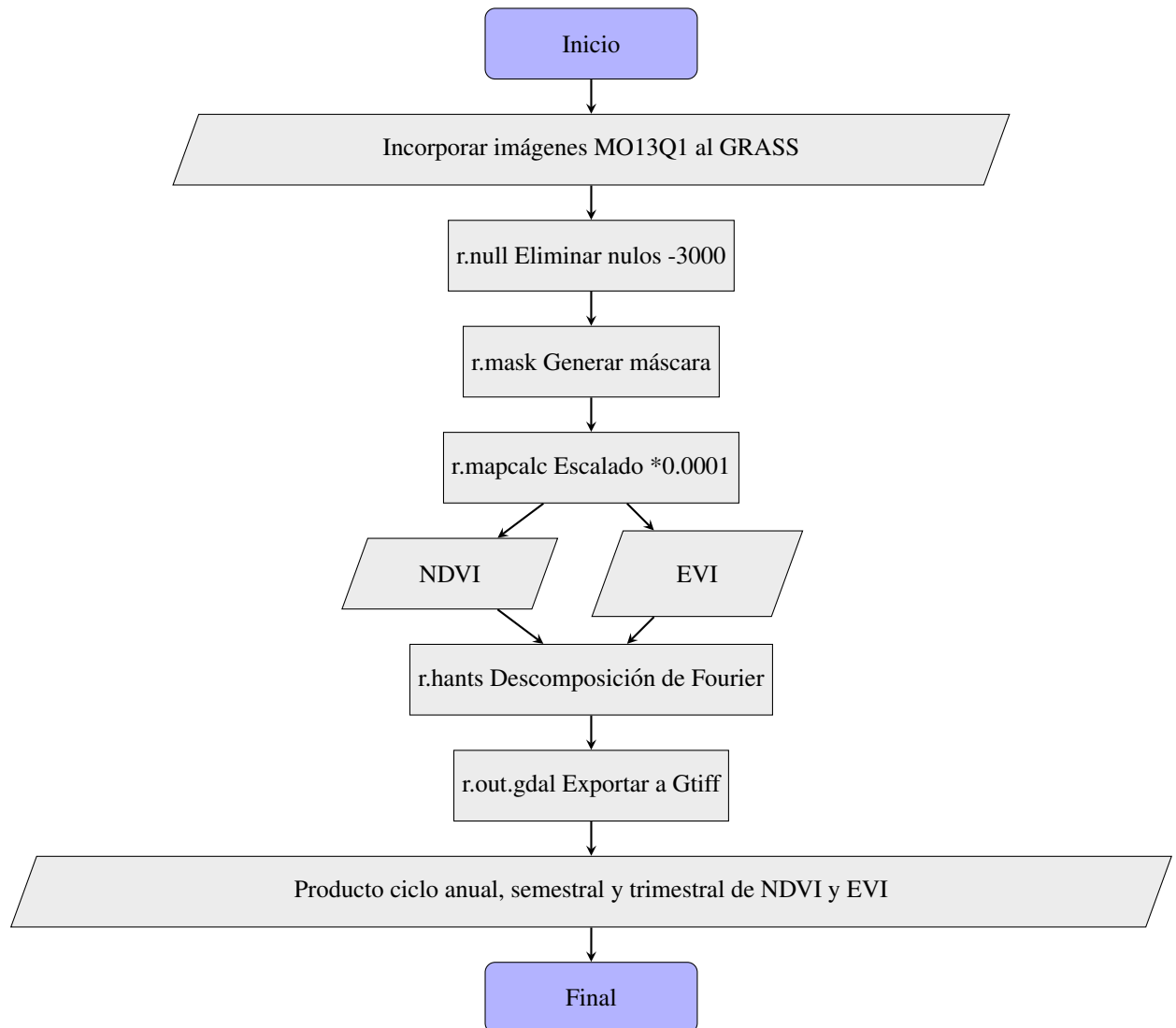


Figura 4.15: Flujo de trabajo para la obtención del producto extracción de ciclos temporales de NDVI y EVI

4.4 Serie temporal de temperatura de superficie (LST)

4.4.1 Fuente de datos

Para la elaboración de productos relacionados con la temperatura superficial de suelo (LST) se utilizaron una vez más imágenes del sensor MODIS (TERRA/AQUA), propiedad de la NASA. En general, los productos de temperatura de superficie están basados en el principio de que la energía electromagnética emitida por los objetos está relacionada con su temperatura, longitud de onda y emisividad.

El sensor multiespectral MODIS obtiene el valor de temperatura de la tierra basándose en un algoritmo de cálculo sobre las bandas de emisividad 31 y 32, 10.8 a 11.3 μm y 11.8 a 12.3 μm respectivamente (infra-rojo térmico) (United States Geological Survey, 2006). Históricamente este producto ha sido utilizado por su asociación con las condiciones hidrológicas del ambiente.

El producto MOD11A1 de temperatura LST diaria tiene una resolución espacial de 1km y es generado en una grilla con proyección sinusoidal para todo el planeta. Un *tile* de este producto contiene un raster de 1200 filas por 1200 columnas y exactamente cada píxel tiene 0.928km de lado (United States Geological Survey, 2006). El formato de archivo en el que se suministra este producto es HDF4 (*Hierarchical Data Format*) al igual que los índices de vegetación anteriormente descritos.

La tabla 4.2 resume el *dataset* de capas que se obtienen al descargar un archivo .HDF del producto MOD11A1, como se observa, el producto proporciona datos de temperatura tanto diurna como nocturna por lo cual para cada día del año se obtienen dos datos tomados por el sensor MODIS, esto resulta de la combinación de las pasadas del satélite aqua (noche) y TERRA (día) (United States Geological Survey, 2006). También se obtienen las capas de calidad de dato para cada pasada y de la cobertura nubosa a la hora de la toma de la emisividad; así mismo el *dataset* proporciona las bandas de emisividad 31 y 32 a partir de las cuales se obtiene el dato de LST.

Tabla 4.2: Capas que conforman el archivo .HDF (dataset) del producto MOD11A1 (United States Geological Survey, 2006).

Capas del dataset (HDF)	Unidades	No-dato	Rango Válido	Factor de Escala
LST_day	Kelvin	0	7500, 65535	0.02
Calidad de dato QC_day	No aplica	0	0, 255	No aplica
Hora de observación_day	Hora	0	0, 240	0.1
Ángulo de visión respecto al Zenith_day	Grados	255	0, 130	1.0
LST_night	Kelvin	0	7500, 65535	0.02
Calidad de dato QC_night	No aplica	0	0, 255	No aplica
Hora de observación_night	Hora	0	0, 240	0.1
Ángulo de visión respecto al Zenith_night	Grados	255	0, 130	1.0
Emisividad banda 31	No aplica	0	1, 255	0.002
Emisividad banda 32	No aplica	0	1, 255	0.002
Cobertura de nube_day	No aplica	0	0, 65535	0.0005
Cobertura de nube_night	No aplica	0	0, 65535	0.0005

4.4.2 Adquisición y preprocesamiento de datos

Al igual que los índices verdes, los productos MODIS fueron provistos por el Servicio Geológico de los Estados Unidos (USGS) y descargados via [http](http://e4ftl01u.eds.nasa.gov/MOLT/MOD11A1.005/) desde el sitio <http://e4ftl01u.eds.nasa.gov/MOLT/MOD11A1.005/>.

Para cubrir toda el área de estudio, se seleccionó el *tile* H12V11 de la grilla de MODIS. El producto MOD11A1 proporciona una imagen que contiene información de LST diario y nocturno. (365 imágenes por año). La serie histórica que se descargó son casi once años comprendidos entre el primero de enero de 2004 hasta el 31 de marzo de 2014, para un total de 3732 productos adquiridos para el análisis temporal.

De igual manera que para la serie temporal de NDVI y EVI, la descarga de los 3732 productos que conforman la serie temporal de LST se realizó via http de manera automática utilizando el software "PyMODIS" (Disponible en: <http://pymodis.fem-environment.eu/>).

Se utilizó la función `modis_download.py`. A continuación se muestra la forma en la que se ejecutó esta función por línea de comando desde una terminal de Linux.

```
modis_download.py -s MOLA -t h12v11 -f 2014-04-01 -e 2004-01-01 -p MOD11A1.005 lst
```

Los parámetros de entrada se refieren al nombre del producto tal cual aparece en la http de MODIS (-p), la carpeta en donde esta el producto (-s), el *tile* a descargar (-t), la fecha de inicio (-f), finalización de la descarga (-e) y la carpeta en donde son guardados los archivos descargados. Una vez termina de correr este script, se crea un archivo.txt que informa todos los productos descargados y un archivo.log que reporta cualquier error que se haya producido. Finalmente se obtienen todos los archivos en formato .HDF y sus respectivos .XML que contienen los metadatos de cada producto.

4.4.3 Incorporación de la serie histórica de LST al GRASS GIS

Para cada una de las fechas descargadas con los productos MOD11Q1, las escenas fueron convertidas desde la proyección sinusoidal original al sistema de coordenadas geográficas lat/long WGS84; y posteriormente se incorporaron al GRASS.

Se creó en primera instancia un *LOCATION* en GRASS con el sistema de coordenadas geográficas lat-lon wgs84; debido a que los datos del producto MOD11A1 están en proyección sinusoidal, la importación al GRASS se realizó con el comando `r.modis.import` que permite además de importar los datos, hacer la re-proyección desde el sistema de referencia original al lat-lon para coincidir con la referencia del *LOCATION*.

A continuación se muestra la manera en la que se usó la función para incorporar todas las imágenes de LST al GRASS.

```
#!/bin/bash
#$1: path a los hdf
#$2: path al MODIS reprojection tool

for i in `ls $1/*.hdf`
do
nome=`basename $i .hdf`
r.modis.import dns=$nome mrtpath=$2 spectral="(1,0,0,0,1,0,0,0,0,0,0,0)"
done
```

La función permite hacer una selección de las capas del *dataset* MODIS que se quieren incorporar, en este caso fueron incorporadas las capas 1 y 5 que corresponden a la temperatura

en el día y la noche (Tabla 4.2).

Una vez terminado el comando, usando usando la sentencia general de GRASS se puede listar las capas que ya están dentro del *MAPSET*, con las cuales se puede empezar a trabajar. A continuación se observa el comando *g.list* para listar, y algunas de las capas de LST incorporadas.

```
g.list rast pat=MOD11*
...
MOD11A1.A2004001.h12v11.005.2008163211519.LST_Day_1km
MOD11A1.A2004001.h12v11.005.2008163211519.LST_Night_1km
MOD11A1.A2004002.h12v11.005.2008165140219.LST_Day_1km
MOD11A1.A2004002.h12v11.005.2008165140219.LST_Night_1km
MOD11A1.A2004003.h12v11.005.2008165153703.LST_Day_1km
MOD11A1.A2004003.h12v11.005.2008165153703.LST_Night_1km
...
...
...
```

4.4.4 Producto 4: Descriptivos anuales de la serie temporal de LST

El primero de los productos generados a partir de la serie temporal de LST fue el de descriptivos anuales, siendo esta capa muy informativa para conocer la dinámica intra-anual de los datos de temperatura en el área de estudio.

A continuación se describe la implementación en GRASS GIS de todo el procedimiento para obtener los descriptivos, una vez importadas todas las imágenes.

Los píxeles del producto MOD11A1 con valor igual a cero corresponden a cobertura nubosa (ver tabla 4.2) y por lo tanto son datos nulos que necesitaremos recalcular para la creación de los nuevos productos.

Para realizar lo anterior, se utilizó la función *r.null*, en un ciclo *FOR*, que cambia el valor cero por *NULL* en cada pixel identificado, de la siguiente manera:

```
#Null
for map in `g.mlist rast pat="MOD11A1*LST*" `
do
r.null $map setnull=0
done
```

Una vez identificados los nulos de cada fecha, se procedió con una interpolación píxel a píxel a lo largo del espacio para estimar el dato perdido. La función *r.fillnulls* de GRASS permite "llenar" los píxeles nulos creando una función lineal sobre los píxeles vecinos, de esta forma estima el dato perdido y se reconstruye la imagen completa de cada fecha. El manual completo de esta función esta disponible en línea en <http://grass.osgeo.org/grass64/manuals/r.fillnulls.html>.

A continuación se muestra el uso de la función de GRASS `r.fillnulls`. Es notable que en esta oportunidad la cantidad de imágenes importadas en el GRASS es muy grande (3732), por lo cual hay que cambiar el número de archivos abiertos simultáneos, que por defecto está indicado en el sistema. En la implementación se observa el comando *ulimit*, que modifica este parámetro. Este cambio se hace por única vez para todo el procesamiento de los datos.

```
#Interpolar nulos
ulimit -n 4096

for map in `g.mlist rast pat="MOD11A1*LST*" `
do
r.fillnulls input=$map output=$map --o method=bilinear
done
```

La salida de la función `r.fillnulls`, es la imagen completa (sin nulos) de todas las fechas de la serie de tiempo, por lo que cualquiera de las imágenes puede ser utilizada como máscara para que el resto del procedimiento sea calculado a partir del área de interés. La máscara se construyó de la siguiente manera con la función de GRASS `r.mask`:

```
#Crear mascara
r.mask MOD11A1.A2004001.h12v11.005.2008163211519.LST.Day.1km --o
```

Los datos de los productos MOD11A1 vienen por defecto en grados Kelvin, por lo que se procedió con la conversión de los valores a grados Celsius, implementando la siguiente fórmula:

$$Celsius = (Kelvin * 0.02) - 273.15 \quad (4.3)$$

```
#Celsius
for map in `g.mlist rast pat="MOD11A1*LST*" `
do
r.mapcalc --o expression="$map=({map}*0.02)-273.15"
done
```

Una vez terminado todo el preprocesamiento de los datos, la serie temporal se encuentra preparada para obtener los productos que la sintetizan. Con la función `r.series` explicada en la serie temporal de índices verdes, se generaron productos para cada año de la serie temporal de la media, el mínimo, máximo y su desvío estándar, de la siguiente manera.

```
#Descriptivos
for p in `seq 2004 2014`
do

r.series input=`g.mlist rast pattern=MOD11A1*$p*$p sep=,` \
output=lst_mean_$p, lst_min_$p, lst_max_$p, lst_sd_$p \
method=average, minimum, maximum, stddev

echo "done $p"
done
```

Una vez generados los cuatro productos por año, se procedió con su exportación fuera del GRASS en formato GeoTiff, para su posterior subida al *GeoServer* del sistema *Web-GIS* de

Chagas. La exportación se realizó con la función `r.out.gdal` de la siguiente manera:

```
#exportar a tiff los descriptivos anuales
for map in `g.mlist rast pat=lst*`
do r.out.gdal input=$map output=/home/curso/DBH/productos_chagas/$map format=GTiff
done
```

Las figuras 4.16, 4.17, 4.18 y 4.19 representan la media, mínimo, máximo y desvío estándar para el año 2004, respectivamente.

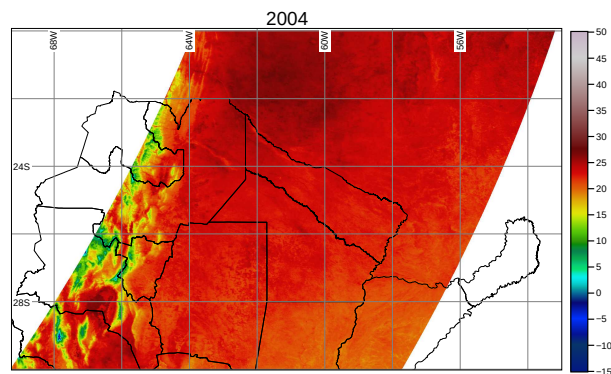


Figura 4.16: Temperatura de superficie media del año 2004

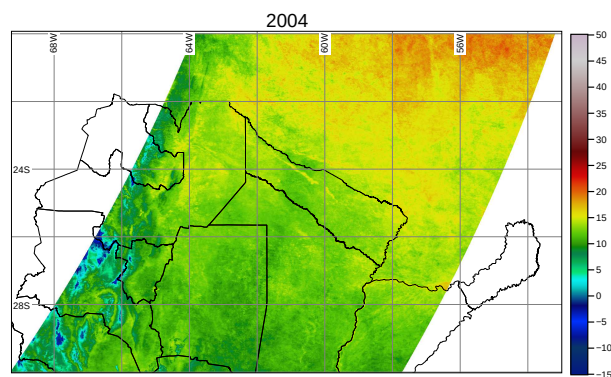


Figura 4.17: Temperatura de superficie mínima del año 2004

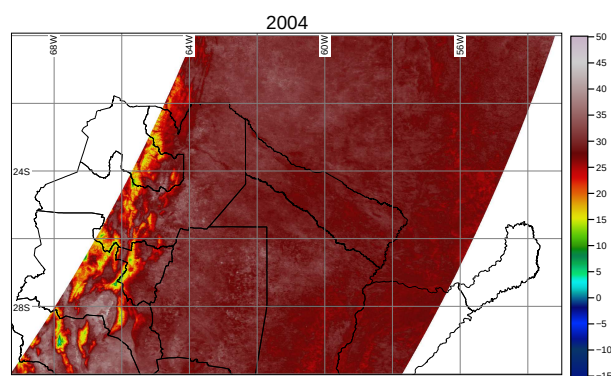


Figura 4.18: Temperatura de superficie máxima del año 2004

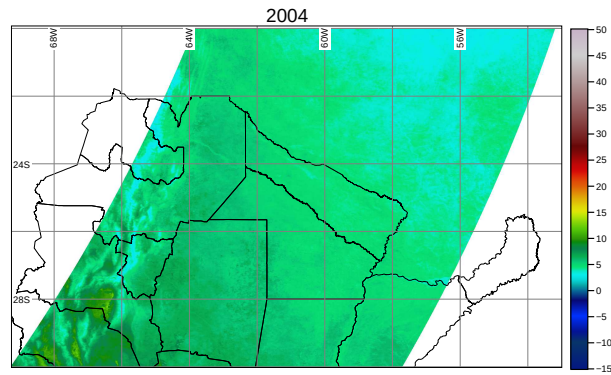


Figura 4.19: Desvío estándar de la temperatura de superficie del año 2004

La figura 4.20 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de los descriptivos anuales de la serie temporal de temperatura de superficie, se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron en GRASS hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

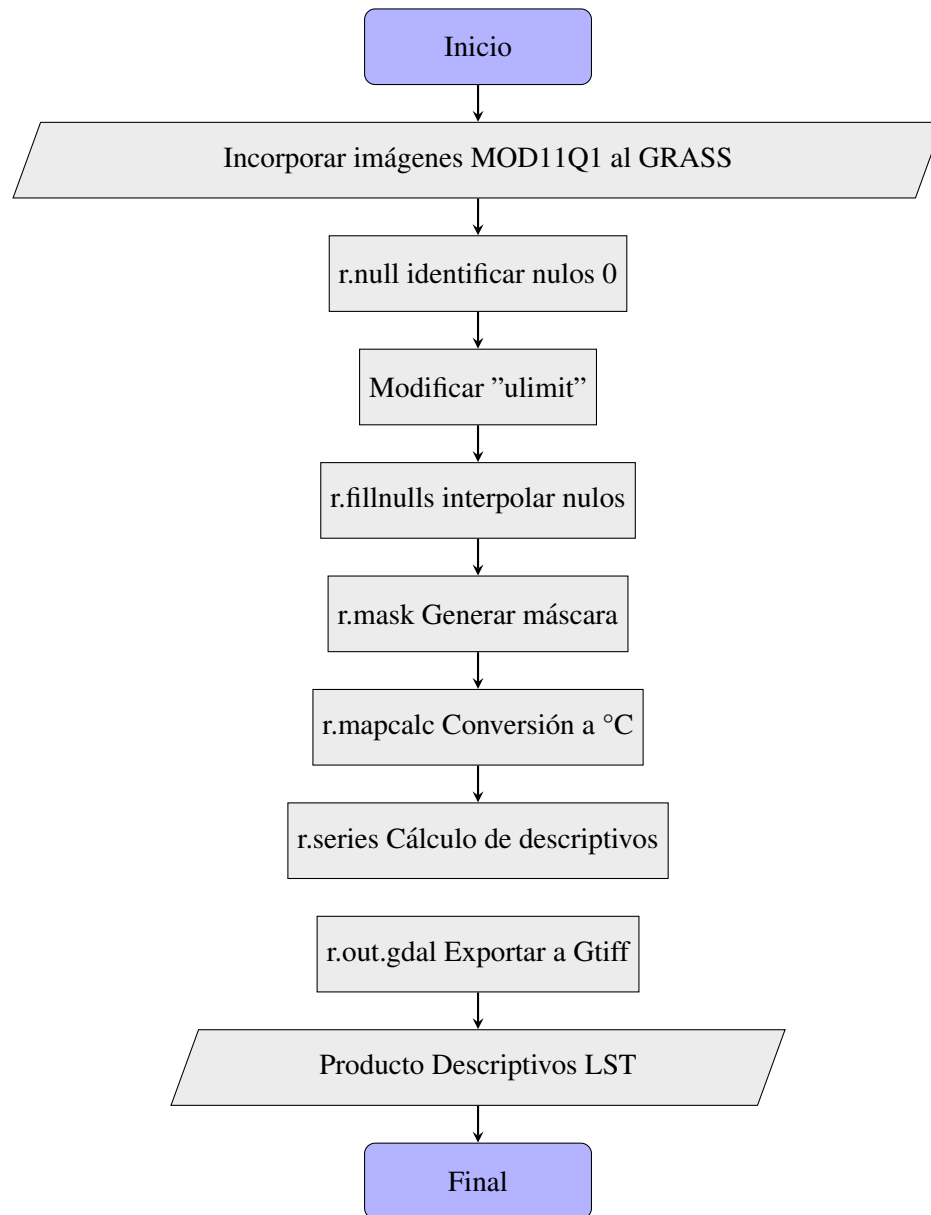


Figura 4.20: Flujo de trabajo para la obtención del producto descriptivos anuales de la serie temporal de LST

4.4.5 Producto 5: Extracción de ciclos temporales de LST mediante análisis de series de Fourier

Al igual que para la serie temporal de índices verdes, para la serie de LST se realizó un análisis de serie de Fourier con la motivación de obtener la variación temporal de la temperatura en el área de estudio, y de resumir la serie de tiempo para ofrecer un producto con significado biológico que no represente una gran cantidad de datos para el usuario.

Se utilizó la implementación del algoritmo "Análisis armónico de series temporales (HANTS)", realizada por Markus Metz, descrita con anterioridad.

Para la extracción de los ciclos temporales de la serie de LST, fue necesario experimentar con varias combinaciones de los parámetros, para finalmente aplicar la función de la siguiente forma:

```
for p in `seq 2004 2014`
do
    for i in `g.mlist rast pattern=MOD11A1.A$p*1km sep=,`
    do
        r.hants input=$i suffix=hants_$p nf=5 fet=0.05 idrt=+50 \
        amplitude=amp_$p phase=phase_$p
        echo "done $i"
    done
done
```

Como se observa en el código anterior, la función `r.hants` recibió como entrada todas las imágenes de temperatura diurna y nocturna de un año, se calcularon 5 frecuencias ($nf=5$), por lo cual el resultado son 5 amplitudes y 4 fases por año; el error absoluto de los residuos debió ser menor a 0.05 grados Celsius ($FET=0.05$) respecto a la curva obtenida; y los valores de temperatura mayores de 50 fueron rechazados bajo el supuesto de ser datos afectados por cobertura nubosa, sobre-estimados por el sensor o por la función `r.fillnulls` con la que se obtuvieron los datos faltantes en cada una de las fechas ($idrt=+50$).

Una vez obtenidos los componentes armónicos de la serie se procedió con la exportación de los mismos del GRASS, generando productos en formato GeoTiff, para su posterior subida al servidor. De las cinco frecuencias calculadas, se exportaron únicamente las frecuencias 1, 2 y 4, las cuales se corresponden con los ciclos anual, semestral y trimestral. La exportación se realizó de la siguiente forma:

```
#exportar a tiff las frecuencias anual(1), semestral(2), y trimestral(4)

for map in `g.mlist rast pat=phase*[1,2,4]`
do
    r.out.gdal input=$map output=/home/curso/DBH/productos_chagas/temperatura_$map \
    format=GTiff
done

for map in `g.mlist rast pat=amp*[1,2,4]`
do
    r.out.gdal input=$map output=/home/curso/DBH/productos_chagas/temperatura_$map \
    format=GTiff
done
```

La figura 4.21 representa los componentes de amplitud del armónico de Fourier; es una composición RGB. El canal rojo representa las zonas donde la temperatura sigue principalmente ciclos anuales; en verde las coberturas con ciclos de temperatura semestrales, y en azul las zonas que se rigen por ciclos trimestrales; dentro del área se pueden encontrar diferentes combinaciones de estos componentes (amarillo, cian y magenta), debido a las condiciones climáticas y de terreno que afectan la temperatura.

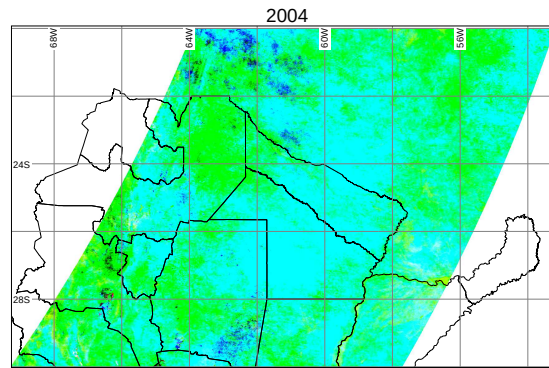


Figura 4.21: Mapa de HANTS con los componentes de Fourier de temperatura (LST) 2004:
Rojo, ciclos anuales. Verde, ciclos semestrales. Azul, ciclos trimestrales

La figura 4.22 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de ciclos temporales de LST mediante el análisis armónico de Fourier, se describen los pasos desde la incorporación de las imágenes al GRASS GIS, a pesar de estar omitidos en la descripción del producto, debido que una vez incorporados los productos en el GRASS, se generaron todos de manera consecutiva. Se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron en GRASS hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

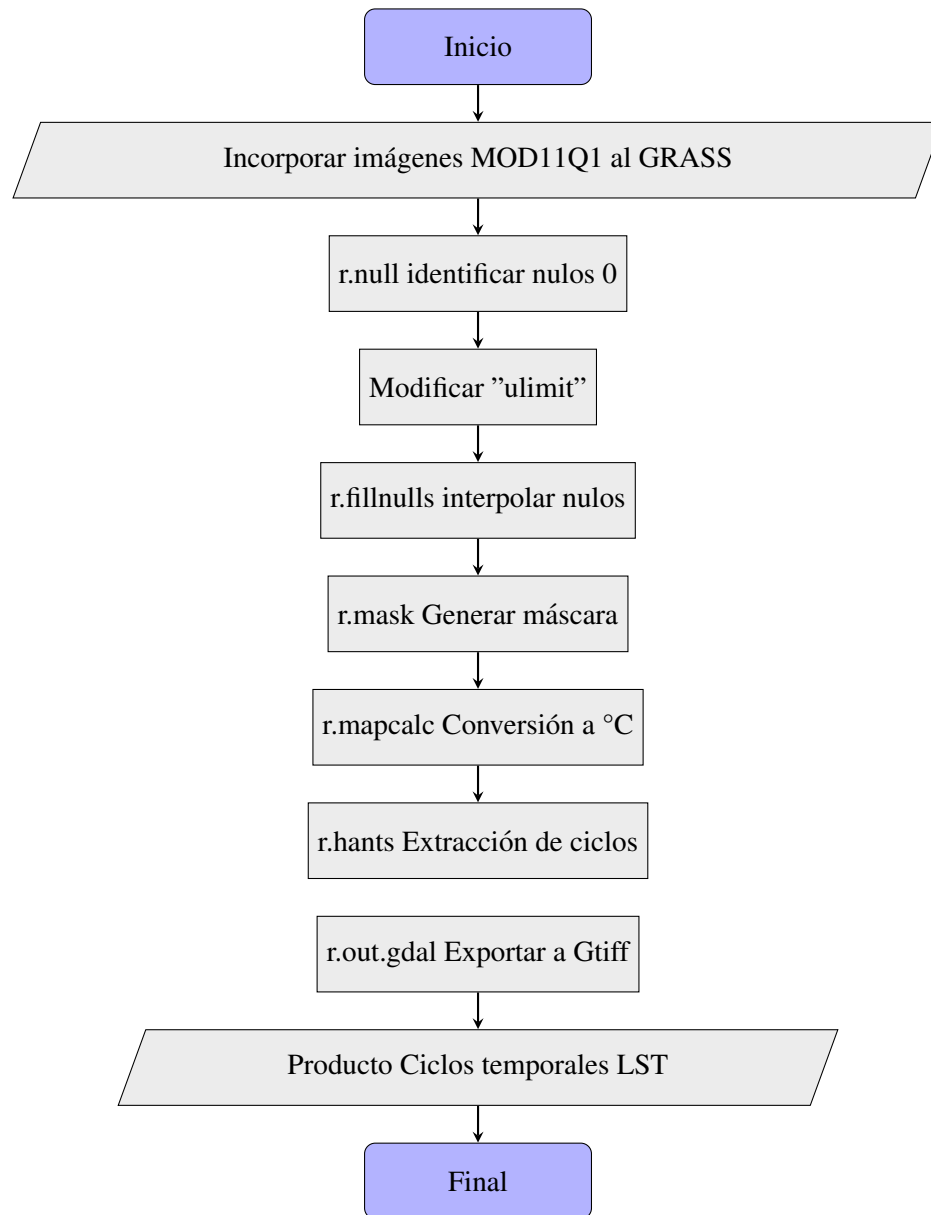


Figura 4.22: Flujo de trabajo para la obtención del producto ciclos temporales de LST

4.4.6 Producto 6: Grados día

La temperatura y humedad son determinantes fundamentales de la distribución y abundancia de la mayoría de las especies, y son variables esenciales para evaluar factores de riesgo de la transmisión de una enfermedad o la presencia de sus vectores (Nieto et al., 2006). Cada especie tiene sus propios requerimientos termo-hidrológicos que determinan su biología, preferencia de habitat, o los límites de tolerancia y óptimas condiciones para su establecimiento y desarrollo (Nieto et al., 2006).

El cálculo de los grados-día de un área puede ser utilizado para definir el gradiente de temperatura idóneo para la propagación y transmisión de las enfermedades transmitidas por vectores, razón por la que este producto se incluye en este estudio. En otras enfermedades, por ejemplo dengue, los grados-día han sido muy estudiados e incluidos en modelos de distribución de la enfermedad, debido a que se conoce que para el desarrollo de las ninfas del mosquito se necesita acumular un umbral de temperatura (Porcasi et al., 2012).

Para generar el producto grados día, se tomaron dos premisas iniciales, para que este producto tenga una fácil interpretación y así mismo pueda encontrarse su relación con la distribución de los vectores de Chagas en estudios futuros. Como premisa uno, fueron acumuladas las temperaturas que no fueron inferiores a 0°C, durante el día ni la noche. Como segunda premisa, fueron acumuladas las temperaturas que no superaron los 30°C en cada fecha. Estos puntos de corte se tomaron con base en estudios realizados en ninfas y adultos de *Triatoma infestans*, que sugieren que las ninfas se encuentran preferentemente en áreas con temperaturas de 26°C, mientras que los adultos con acceso a alimentación fácil (viviendas) permanecen en lugares con temperaturas entre los 28-29°C, y los adultos que no se alimentan con facilidad, permanecen en áreas con 25-26°C (Lazzari, 1991)

De esta forma el producto consiste en el total de días del año que en cada píxel la temperatura fluctuó en el rango 0°C-30°C. Las siguientes líneas fueron generadas en GRASS para obtener el producto:

```
for p in `seq 2004 2014`
do
    #definir arreglo
    declare -a day=(`g.mlist rast pat="MOD11A1*$p*LST*Day*1km"`)
    declare -a night=(`g.mlist rast pat="MOD11A1*$p*LST*Night*1km"`)

    #tomar longitud del arreglo
    len=${#day[@]}

    #crear mapa para acumular
    r.mapcalc --o expression="acum_$p=${day[0]}*0"

    #calcula gdd
    for ((i=0;i<${len};i++))
    do

        r.mapcalc --o expression="acum_$p=if (${day[$i]} >=0 && ${day[$i]} <=30 && \
        ${night[$i]} >=0 && \
        ${night[$i]} <=30 , acum_$p +1, acum_$p)"

    done
done
```

El script anterior define dos arreglos de imágenes, el primero de ellos consiste en el listado completo de las imágenes con temperatura diurna, y el segundo, el listado con las temperaturas nocturnas. Cada una de las imágenes (fechas) de la lista entra en cada iteración del ciclo *for*, en donde la función *r.mapcalc* revisa las dos condiciones anteriormente mencionadas, tanto para las imágenes diurnas y nocturnas, si las condiciones se cumplen, el raster con nombre "acum_año" va sumando una unidad. De esta forma, los píxeles en los que los 365 del año presentaron temperaturas diurnas y nocturnas entre 0°C y 30° tendrán en el raster "acum_año"

el valor de 365.

Una vez terminado un año, el procedimiento se repite para el año siguiente, hasta analizar toda la serie temporal.

Después de haber obtenido los acumulados para los once años de la serie, se procedió con la exportación de los mismos del GRASS, generando productos en formato GeoTiff, para su posterior subida al servidor. La exportación se realizó de la siguiente forma:

```
#exportar a tiff acumulados

for map in `g.mlist rast pat=acum*`
do
r.out.gdal input=$map output=/home/curso/DBH/productos_chagas/$map \
format=GTiff
done
```

A modo de ejemplo, la figura 4.23 representa para el año 2004 el número de días por píxel en los que la temperatura de superficie fluctuó entre 0°C y 30°C, las áreas en rojo son los lugares en donde la mayor parte del año se presentan estas temperaturas.

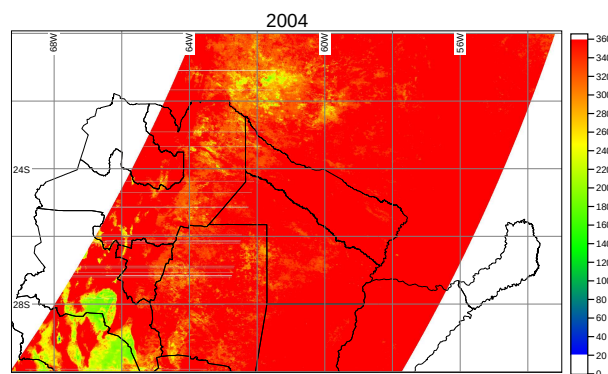
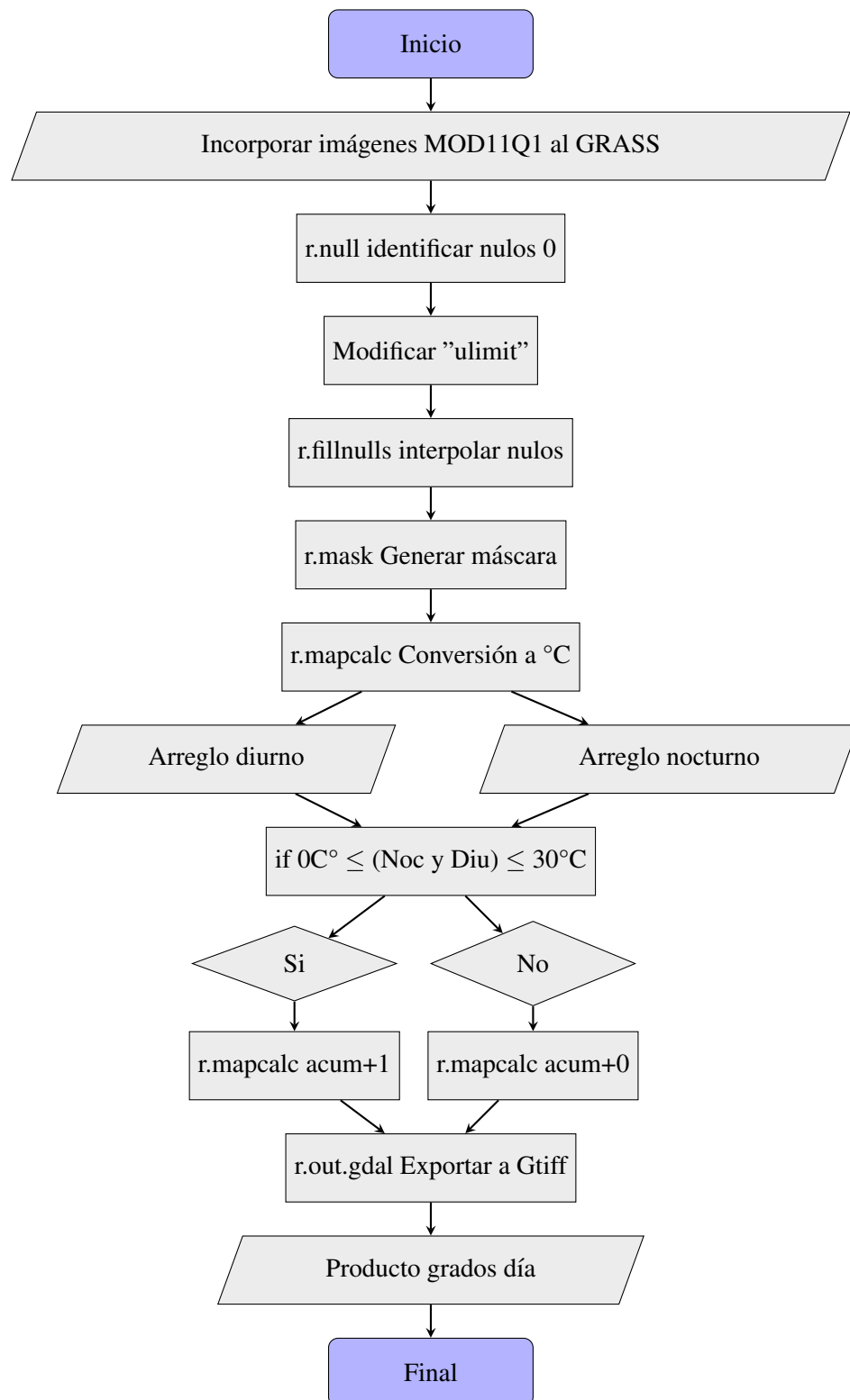


Figura 4.23: Número de días del año 2004 en los que la temperatura de superficie fluctuó entre 0°C y 30°C

La figura 4.24 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de grados día, se describen los pasos desde la incorporación de las imágenes al GRASS GIS, a pesar de estar omitidos en la descripción del producto, debido que una vez incorporados los productos en el GRASS, se generaron todos de manera consecutiva. Se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron en GRASS hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

**Figura 4.24:** Flujo de trabajo para la obtención del producto Grados Día

4.5 Producto 7: Mapas de Cobertura de la tierra

La cobertura de la superficie juega un papel importante en el clima y la biogeoquímica de la Tierra. Un uso importante de los datos de cobertura de la tierra, es la inferencia de los parámetros que influyen en los intercambios biogeoquímicos y de energía entre la atmósfera y la superficie de la tierra, para su uso en modelos y otras aplicaciones ([United States Geological Survey, 1999](#)).

Ejemplos de tales parámetros incluyen índice de área foliar, rugosidad, resistencia de la superficie de la evapotranspiración, fracción de verde de dosel, densidad de la vegetación, y fracción de la radiación absorbida fotosintéticamente activa ([United States Geological Survey, 2013](#)).

El producto *Land Cover Type* (MCD12Q1) ofrece un conjunto de tipos de cobertura del suelo que apoyan el análisis del cambio global, usando información espectral y temporal derivada de las imágenes MODIS ([United States Geological Survey, 2013](#)).

El producto proporciona información global del uso del suelo, con una resolución espacial de 500 m y una periodicidad de actualización anual entre los años 2001 y 2012 ([United States Geological Survey, 1999](#)). Desde el último año, no se encuentra disponible ninguna nueva versión.

Las clases de cobertura del producto MCD12Q1 se producen mediante el procesamiento de imágenes de 32 días, analizadas mediante árboles de decisión y clasificadas mediante redes neuronales. El producto está preparado para reconocer y cuantificar cambios sutiles y transformaciones progresivas del uso de la tierra.

Tabla 4.3: Capas que conforman el archivo .HDF (dataset) del producto MCD12Q1 ([United States Geological Survey, 2013](#)).

Capas del dataset (HDF)	Unidades	No-dato	Rango Válido
Cobertura tipo 1 (IGBP)	# de clase	255	0, 254
Cobertura tipo 2 (UMD)	# de clase	255	0, 254
Cobertura tipo 3 (LAI/fPAR)	# de clase	255	0, 254
Cobertura tipo 4 (NPP/BGC)	# de clase	255	0, 254
Cobertura tipo 5 (PFT)	# de clase	255	0, 254
Evaluación Cob tipo 1	Porcentaje	255	0, 254
Evaluación Cob tipo 2		255	0, 254
Evaluación Cob tipo 3		255	0, 254
Evaluación Cob tipo 4		255	0, 254
Evaluación Cob tipo 5		255	0, 254
Calidad (QC) de cobertura	No aplica	255	0, 254
Cob. secundaria tipo 1		255	0, 254
% Cob. secundaria tipo 1		255	0, 254
Propiedad 1 de Cobertura		255	0, 254
Propiedad 2 de Cobertura		255	0, 254
Propiedad 3 de Cobertura		255	0, 254

El formato de archivo del producto, al igual que para los índices de vegetación y temperatura de superficie, es un HDF, que contiene un *set* de capas (Tabla 4.3). El mismo proporciona

información sobre la cobertura del suelo agrupada según cinco leyendas preestablecidas.

- Leyenda del *International Geosphere Biosphere Programme* (IGBP).
- Leyenda de la universidad de Maryland (UMD).
- Leyenda propia derivada de datos LAI/fPAR (Índice de Longitud de Área foliar, Fracción de la radiación absorbida fotosintéticamente activa)
- Leyenda propia derivada de valores de Productividad Neta Primaria (NPP)
- Leyenda basada en la tipología funcional de las plantas (PFT, *Plant functional type*)

De las clasificaciones disponibles, para este estudio se utilizó únicamente la leyenda del IGBP. Esta clasificación reconoce 17 categorías de la cobertura de la tierra (Tabla 4.4), e incluye once categorías de vegetación natural, desglosadas por forma de vida; tres clases de tierras cultivadas y mosaico; y tres clases de tierras sin vegetación.

Tabla 4.4: Leyenda de cobertura de la tierra del *International Geosphere Biosphere Programme* (IGBP) (United States Geological Survey, 2013).

Clase	Cobertura
0	Agua
1	Perennifolias de hoja estrecha
2	Perennifolias de hoja ancha
3	Caducifolia de hoja estrecha
4	Caducifolia de hoja ancha
5	Bosques mixtos
6	Matorral denso
7	Matorral disperso
8	Sabana boscosa
9	Sabana
10	Praderas y pastizales
11	Zonas pantanosas
12	Cultivos
13	Superficies artificiales
14	Mosaico de Cultivos/Vegetación natural
15	Hielo y Nieve
16	Suelo desnudo o con poca vegetación
254	Sin clasificar
255	No dato

El producto de cobertura de la tierra se generó debido a la importancia de los datos que ofrece; ya sea para un usuario GIS avanzado o no, es una capa que brinda información base de importancia y puede ser incluida en cualquier modelo de distribución de especies, bajo la premisa que el uso del suelo tiene influencia o relación con la presencia, estabilidad y éxito reproductivo de las especies.

En congruencia, fueron adquiridos los productos MDC12Q1 disponibles para el área de estudio, con la finalidad de mantenerlos disponibles en la plataforma Web-GIS de Chagas.

A partir de la dirección ftp de MODIS disponible para los productos de cobertura de la tierra (<http://e4ftl01.cr.usgs.gov/MOTA/MCD12Q1.051/>), fueron descargados los productos MCD12Q1 de los años 2001 hasta 2012. Al igual que los productos de índices verdes y temperatura, estos productos se distribuyen por *tiles* en un grilla sinusoidal, y el área de estudio se ubica en el *tile* H12V11.

La descarga de los 12 productos (uno por año) que conforman la serie de cobertura de la tierra, se realizó via http de manera automática utilizando el software "PyMODIS" (Disponible en: <http://pymodis.fem-environment.eu/>).

Se utilizó la función `modis_download.py`. A continuación se muestra la forma en la que se ejecutó esta función por línea de comando desde una terminal de Linux.

```
modis_download.py -s MOTA -t h12v11 -f 2012-12-31 -e 2001-01-01 -p MCD12Q1.051 cover
```

Los parámetros de entrada se refieren al nombre del producto tal cual aparece en la http de MODIS (-p), la carpeta en donde esta el producto (-s), el *tile* a descargar (-t), la fecha de inicio (-f), finalización de la descarga (-e) y la carpeta en donde son guardados los archivos descargados. Finalmente se obtienen todos los archivos en formato .HDF y sus respectivos .XML que contienen los metadatos de cada producto.

El archivo .HDF contiene una gran cantidad de información de la cobertura de la tierra (tabla 4.3), sin embargo, es necesario extraer únicamente la capa con la clasificación IGBP, para lo cual se utilizó otra función de la librería PyMODIS. El comando `modis_convert.py` permite cambiar la proyección sinusoidal en la que vienen los datos por cualquier otra, en este caso las capas se re proyectaron al sistema de coordenadas geográficas lat-lon con datum WGS84. La función se utilizó de la siguiente forma:

```
for i in `ls $path*.hdf`
do

modis_convert.py -s "(1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)" -m $mrt -o $i.tif $i

done
```

El parámetro -s permite realizar un subset espectral, de esta forma se seleccionó, dentro de las 12 capas que están contenidas en el HDF, la clasificación del IGBP.

Una vez re-proyectada la capa queda un archivo .tiff por año, con la proyección adecuada para ser posteriormente incorporado al sistema Web-GIS de Chagas.

A modo de ejemplo, la figura 4.25 muestra el producto de cobertura de la tierra para el año 2012, con la leyenda IGBP y las 17 categorías listadas en la tabla 4.4.

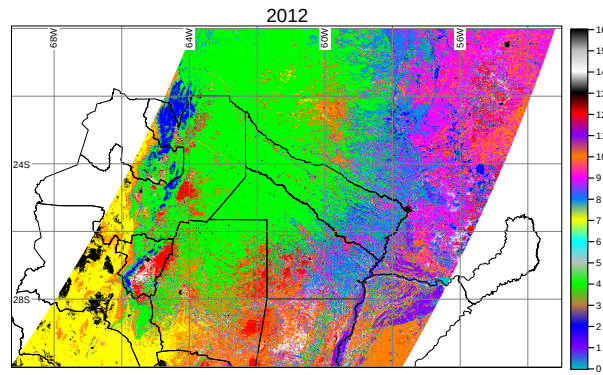


Figura 4.25: Mapa de cobertura de la tierra del área de estudio en el 2012, según la clasificación del *International Geosphere Biosphere Programme* (IGBP) ([United States Geological Survey, 2013](#)).

La figura 4.26 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de cobertura de la tierra, se describen los pasos desde la descarga de las imágenes, y cada uno de los sub-procesos que se realizaron usando la librería PyMODIS, hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

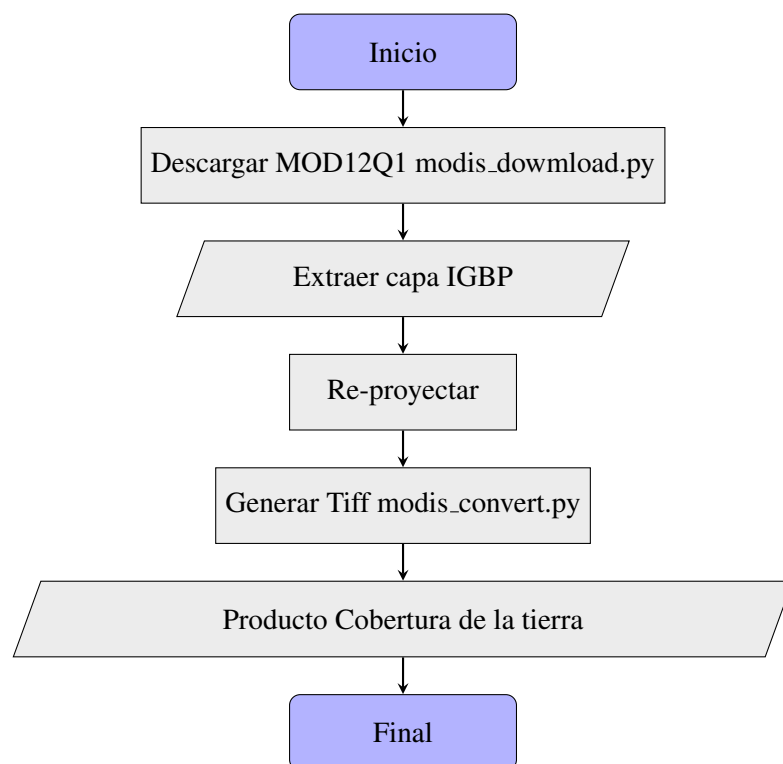


Figura 4.26: Flujo de trabajo para la obtención del producto de Cobertura de la tierra

4.6 Producto 8: Modelo de Elevación Digital (DEM)

En febrero del año 2000, por parte de una colaboración conjunta entre la NASA, la Agencia Nacional de Mapeo e Imágenes de los Estados Unidos (NIMA), y las agencias espaciales de Alemania e Italia, se llevó a cabo la misión espacial *Shuttle Radar Topography Mission* (SRTM), con la finalidad de mapear la superficie de la tierra a un nivel de detalle que para su época marco la diferencia por su resolución y acceso (Bowman, 2005). El modelo de elevación digital generado por la misión SRTM proporciona un salto muy notable en la disponibilidad de la información topográfica para muchas partes remotas del mundo (Neteler, 2005).

La misión SRTM captura los datos de topografía mediante interferometría en una sola pasada, para lo cual fueron montadas dos antenas de radar en el transbordador espacial para tomar simultáneamente dos imágenes con una ligera diferencia espacial entre ellas. Una antena esta sobre la superficie de la plataforma satelital y la otra en el extremo de un mástil que mide 60 metros de longitud (Neteler, 2005). La señal de radar enviada se refleja y es recibida por las dos antenas, de esta manera con la diferencia entre las dos señales es calculada la altura del terreno.

Los datos SRTM se distribuyen en diferentes resoluciones, dependiendo de la parte del mundo de interés. Para zonas fuera de Estados Unidos se distribuyen en 90 metros de resolución. La precisión vertical se estima como alrededor de 10 metros.

Existen diferentes formas de adquirir los datos SRTM, la más común es ingresando al ftp de la NASA (Ver: <ftp://e0mss21u.ecs.nasa.gov/srtm/>), en donde se encuentran disponibles para la descarga los *tiles* de todo el mundo de las imágenes SRTM con 90m de resolución espacial. Los archivos en esta ftp están en formato HGT (*Height format*), comprimidos en un archivo zip.

GRASS GIS tiene implementado el manejo de modelos de elevación digital (DEM) a partir de datos SRTM, desde la descarga de los mismos, incorporación y manejo dentro del GIS.

Como primera instancia, se debe generar un *LOCATION* en GRASS con coordenadas geográficas lat-lon y datum WGS84. La generación de la capa de DEM a partir de los datos SRTM, se realizó para todo el área de estudio que cubre los productos de índices verdes, temperatura y cobertura de la tierra (*tile* de MODIS h12v11). Utilizando el *LOCATION* y la región de trabajo de los anteriores productos, se procedió a obtener los datos SRTM de la misma área de interés. La función `r.in.srtm.region` de GRASS, permite descargar todos los *tiles* con los datos SRTM necesarios para cubrir la región de trabajo; dicha función se usó de la siguiente forma:

```
r.in.srtm.region --overwrite output=dem
```

Con tan solo esta línea de comando, GRASS GIS se conecta con el ftp de NASA, descarga todos los *tiles* en formato .hgt.zip, los lee, los incorpora al *LOCATION* en uso, y genera un mosaico de los mismos. De esta forma se crea la capa denominada "dem" que contiene el

modelo de elevación digital para todo el área de estudio.

Parte del post-procesamiento de los datos SRTM es la corrección de los huecos (datos nulos). Estos errores son particularmente frecuentes en los *tiles* que abarcan superficies con mucha pendiente o rugosidad, y son debidos a la técnica de adquisición de datos de Radar de Apertura Sintética (SAR). Así mismo, zonas con montañas muy altas que generan sombras, también llevan a errores de cálculo con interferometría, por lo cual generan datos nulos.

La función `r.fillnulls` normalmente hace un trabajo aceptable al llenar estos huecos de los datos SRTM, extrae una ventana de los píxeles vecinos al dato nulo y calcula un promedio o alguna función de ajuste para estimar el valor del píxel faltante. Se utilizó esta función de GRASS de la siguiente forma:

```
r.fillnulls input=dem output=dem1
```

Finalmente, se obtuvo el modelo de elevación digital "dem1" corregido en sus valores nulos. Se procedió con la exportación de la capa, generando el producto en formato GeoTiff, para su posterior subida al servidor. La exportación del GRASS GIS se realizó de la siguiente forma:

```
r.out.gdal -o input=dem1 output=/home/curso/DBH/productos_chagas/dem.h12v11 \
format=GTiff
```

La figura 4.27 representa el modelo de elevación digital del área de interés, se observan hacia el costado izquierdo las zonas en blanco con mayor altura que corresponden a la cordillera y pre-cordillera. Por su parte en tonos verdes, se observan las zonas bajas que comprenden gran porcentaje del área, y en donde los límites biológicos del vector pueden llevar a que sean los sitios con mayor interés a la hora de estimar modelos de distribución de *T. infestans*.

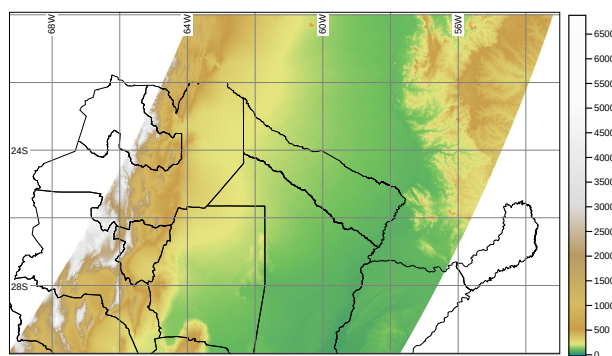


Figura 4.27: Modelo de Elevación Digital a partir de datos SRTM del área de estudio.

La figura 4.28 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto Modelo de Elevación Digital a partir de datos SRTM del área de estudio, se describen los pasos desde la descarga e incorporación de las imágenes al GRASS GIS, se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron en GRASS hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

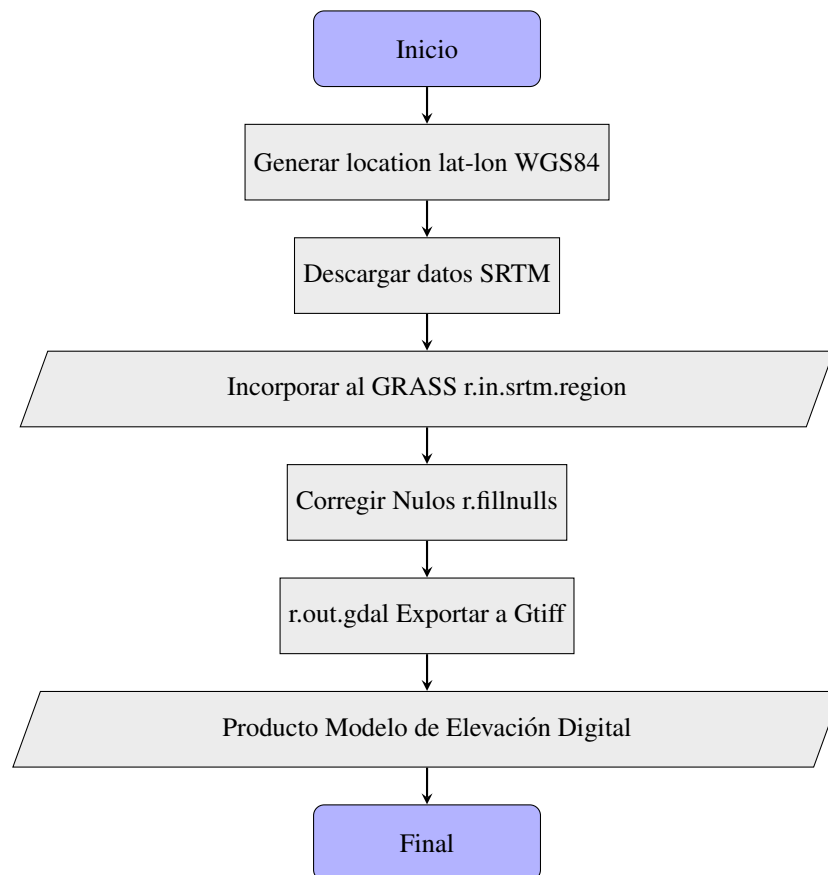


Figura 4.28: Flujo de trabajo para la obtención del producto Modelo de Elevación Digital a partir de datos SRTM

4.7 Producto 9: Superficies climáticas con datos de terreno (WorldClim)

Los mapas de superficies climáticas a partir de la interpolación de datos de terreno son usados en muchas aplicaciones, particularmente en ciencias ambientales, agrícolas y biológicas, por esta razón se consideró importante incluir variables de este tipo (Hijmans et al., 2005).

La resolución espacial de estas interpolaciones depende de las necesidades de esa aplicación y de la cantidad de datos disponibles; en muchas aplicaciones de fenómenos locales son necesarias resoluciones menores a 1 km, para capturar la variabilidad ambiental, sobretudo en zonas montañosas que presentan gradientes climáticos con cambios abruptos. Sin embargo, tener datos con esta resolución es difícil, por la poca cobertura de estaciones de medición terrena en la mayoría de la superficie de la tierra.

Usualmente se usan interpolaciones con una resolución espacial de 18 a 55 km, sin embargo existe un set de datos denominado *WorldClim* desarrollado por Robert J. Hijmans, Susan Cameron, and Juan Parra en el Museo de zoología de vertebrados de la Universidad de California (Hijmans et al., 2005).

WorldClim es un set de capas climáticas globales en formato raster, son generadas con una

resolución espacial de aproximadamente 0.8 km dependiendo de la latitud ($\sim 1\text{km}$), y su dataset es distribuido de manera libre con propósitos académicos desde la dirección <http://www.worldclim.org>. Las capas climáticas son generadas mediante la recopilación de promedios mensuales de datos climáticos, a partir de estaciones meteorológicas globales, regionales, nacionales y locales (Hijmans et al., 2005). Los datos son en su mayoría para el periodo 1950-2005, la información de terreno se interpola y de esta forma se obtienen mapas de superficies climáticas para la precipitación y temperatura mensual mínima, media y máxima (Hijmans et al., 2005).

WorldClim ofrece también variables bio-climáticas (BIOCLIM) derivadas de los promedios mensuales de lluvia y temperatura. Estas capas fueron generadas para ofrecer a los usuarios variables con un mayor sentido biológico y son usualmente incluidas en modelos de distribución de especies o nicho ecológico (Hijmans et al., 2005). Las capas del BIOCLIM representan tendencias y datos extremos de factores ambientales, son un conjunto de 19 variables listadas a continuación:

- BIO1 = Temperatura media anual
- BIO2 = Rango medio de temperatura diurna (Media mensual(max temp - min temp))
- BIO3 = Isotermalidad (BIO2/BIO7) (* 100)
- BIO4 = Estacionalidad de la temperatura (Desvío estándar *100)
- BIO5 = Temperatura máxima del mes más caliente
- BIO6 = Temperatura mínima del mes más frío
- BIO7 = Rango anual de temperatura (BIO5-BIO6)
- BIO8 = Temperatura media de trimestre más húmedo
- BIO9 = Temperatura media de trimestre más seco
- BIO10 = Temperatura media de trimestre más caliente
- BIO11 = Temperatura media de trimestre más frío
- BIO12 = precipitación anual
- BIO13 = Precipitación del mes más húmedo
- BIO14 = Precipitación del mes más seco
- BIO15 = Estacionalidad de la precipitación (Coeficiente de variación)
- BIO16 = Precipitación del trimestre más húmedo
- BIO17 = Precipitación del trimestre más seco
- BIO18 = Precipitación del trimestre más caliente
- BIO19 = Precipitación del trimestre más frío

Los datos WorldClim pueden ser descargados y manejados con facilidad con el *software* para análisis estadístico R (R Core Team, 2014), que tiene disponible la librería "Raster" con una gran cantidad de funciones implementadas para generar procedimientos que van desde la descarga y procesamiento de variables de diferentes fuentes de datos espaciales (Hijmans y Van Etten, 2012).

Una de las funciones disponibles dentro de la librería *Raster* es *GetData*, permite obtener datos geográficos de cualquier lugar del mundo. De forma automática la función lee los datos después de ser descargados desde la dirección donde se encuentra ubicado el repositorio *online* (Hijmans y Van Etten, 2012). Usando esta función es posible descargar información a partir de diferentes fuentes; para este caso se usó únicamente para obtener las variables del BIOCLIM que ofrece el set de datos de *WorldClim*. La función *GetData* se usó de la siguiente forma:

```
require(raster)
w = getData('worldclim', var='bio', res=0.5, lon=-67, lat=-30)
#tile 43
z = getData('worldclim', var='bio', res=0.5, lon=-66, lat=-25)
#tile 33
```

Cuando la función *GetData* recibe como parámetro la descarga de datos *WorldClim*, es necesario indicarle a su vez, la variable a descargar. Para el caso de las capas BIOCLIM, de manera automática se descargan las 19 variables incluidas en el set de datos. Adicionalmente es necesario indicar la resolución de los datos a descargar. La mejor resolución disponible es $\sim 1\text{km}$, y estos datos son distribuidos en *tiles*, por lo que la función también requiere que se especifique al menos una coordenada para ubicar el *tile* a ser descargado.

El área de interés se encuentra en el límite entre dos *tiles* (43 y 33), por lo que se descargaron de manera consecutiva los *tiles* y posteriormente, mediante el uso de la función "mosaic" se unieron los dos *tiles*. La función une dos o mas objetos, generando uno solo con una extensión espacial mayor; en particular esta función permite decidir que procedimiento seguir con los píxeles que se sobrelapan en los límites de los *tiles*, lo que la hace diferente a un simple *merge* (Hijmans y Van Etten, 2012). El mosaico de las 19 variables del BIOCLIM se realizó de la siguiente forma, en donde w y z son los *tiles* descargados con anterioridad:

```
m=mosaic(w,z,fun=min)
```

Con la finalidad de seleccionar dentro de las 19 variables del BIOCLIM, las capas que representen información poco redundante, se realizó un análisis de correlación entre las 19 capas; la función "LayerStats" de la librería *Raster* permite computar análisis de correlación y covarianza para objetos multi-capas (Hijmans y Van Etten, 2012). Se utilizó de la siguiente forma la función para obtener el coeficiente de correlación de Pearson para cada par de capas:

```
cor<-layerStats(m,'pearson',na.rm=TRUE)
```

La figura 4.29 representa los coeficientes de correlación de Pearson para cada par de capas incluidas en el data set de 19 variables del BIOCLIM; las variables con correlaciones positivas

son representadas en azul, y las correlaciones negativas en rojo. Por su parte el tamaño del círculo e intensidad del tono representan que tan fuerte es la correlación entre las variables; al lado izquierdo de la diagonal de la matriz de correlación se muestran los valores puntuales para cada una de las comparaciones pareadas.

Mediante la interpretación de este gráfico se seleccionaron 6 capas de las 19 variables, que se consideraron aportan información poco redundante entre ellas (poco correlacionadas entre sí); las variables seleccionadas fueron BIO2 (Rango medio de temperatura diurna), BIO3 (Isotermalidad), BIO7 (Rango anual de temperatura), BIO14 (Precipitación del mes más seco), BIO18 (Precipitación del trimestre más caliente) y BIO19 (Precipitación del trimestre más frío).

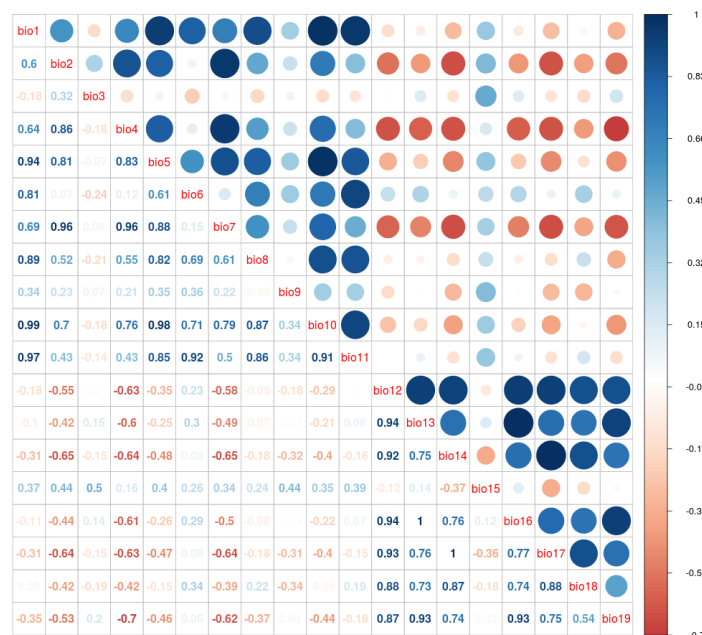


Figura 4.29: Coeficientes de correlación de Pearson para las 19 variables del data set BIOCLIM

Una vez seleccionadas las capas, mediante la función de R *"dropLayer"*, que permite eliminar capas de un *stack* de variables, la función recibe como parámetro de entrada el objeto *"m"* con las 19 variables y un arreglo con el número identificador de las capas a ser eliminadas; fueron descartadas las capas que están altamente correlacionadas (figura 4.29). La función se utilizó de la siguiente forma:

```
m<-dropLayer(m, c(1,4,5,6,8,9,10,11,12,13,15,16,17))
```

Una vez obtenido un *stack* con las seis capas seleccionadas, se procedió con la exportación de las mismas, generando productos en formato GeoTiff, para su posterior subida al servidor. La exportación se realizó de la siguiente forma, usando la función de R *"writeRaster"*, que recibe como parámetro de entrada el objeto con las capas, y la asignación *"bylayer=TRUE"* genera un tiff individual para cada capa dentro del *stack* (imágenes apiladas):

```
M<-writeRaster(m, filename="/home/curso/DBH/productos_chagas/mosaico_worldclim.tif", \
format="GTiff", bylayer=TRUE, overwrite=TRUE)
```

La figura 4.30 muestra las seis variables del BIOCLIM, que fueron seleccionadas, y exportadas en formato GeoTiff, para su posterior subida al sistema Web-GIS de Chagas.

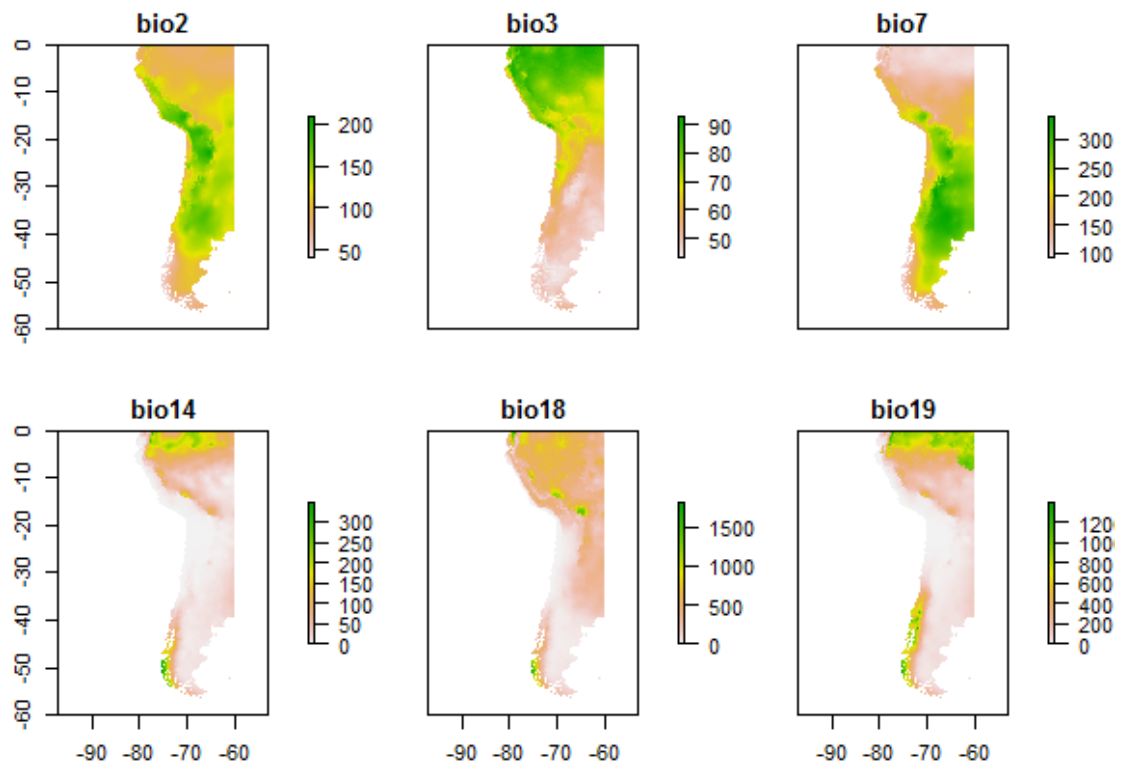


Figura 4.30: Variables bio-climáticas a partir de datos de terreno *WorldClim*

La figura 4.31 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de Superficies climáticas con datos de terreno (WorldClim), se describen los pasos desde la descarga, incorporación de las imágenes y procesamiento dentro del software R, así como las funciones utilizadas en cada paso, se muestran las imágenes de entrada, y cada uno de los sub-procesos que se implementaron hasta obtener los raster en formato GeoTiff preparados para la incorporación al sistema Chagas.

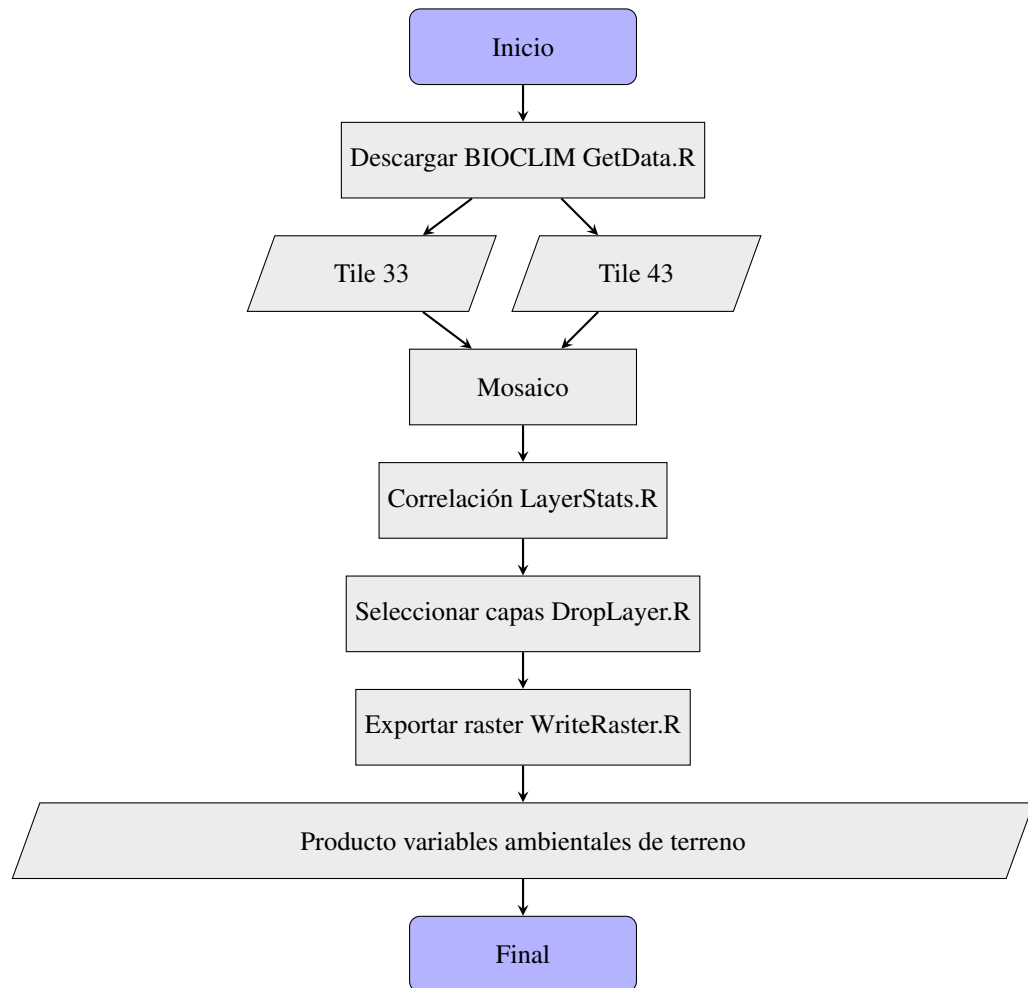


Figura 4.31: Flujo de trabajo para la obtención del producto Superficies climáticas con datos de terreno (WorldClim)

4.8 Producto 10: Índice topográfico de humedad (TWI).

El índice topográfico de humedad, *Topographic Wetness Index* (TWI), es una estimación de la acumulación de humedad (o agua) en una determinada unidad espacial y relaciona el 'área de acumulación' (medida del agua que fluye hacia un cierto punto) con la pendiente local (medida de la transmisividad lateral subsuperficial) (Beven y Kirkby, 1979). Esta función determina que las áreas con valores más altos de TWI se saturarán durante una lluvia o el acontecimiento de deshielo con mayor probabilidad (Lanfri, 2011).

El TWI se define como (Beven y Kirkby, 1979):

$$TWI = \ln \frac{a}{\tan \beta} \quad (4.4)$$

Donde β es el ángulo de pendiente local y a es la acumulación de flujo por unidad de superficie. La acumulación de flujo a , como se especificó anteriormente, es el área que contribuye a cada píxel y puede ser estimada como el producto del número de píxeles que drenan multiplicado por el área de cada píxel (Beven y Kirkby, 1979). El índice topográfico de humedad es adimensional. Se consideró importante incluir este producto al sistema Chagas debido a que el TWI podría proveer una descripción simple y biológicamente significativa acerca de cómo la topografía afecta el riesgo de transmisión de Chagas, a través de la facilitación de sitios potenciales para el desarrollo de los vectores. En malaria por ejemplo, ha sido demostrado que la incidencia local de la enfermedad en una región de Kenya occidental propensa a esta epidemia, tiene una fuerte asociación con índices de humedad topográficamente determinados (Cohen et al., 2008). A su vez, los índices topográficos de humedad locales han predicho hábitats propicios para la presencia de malaria, en algunos casos mejor que variables como uso y cobertura terrestre (Cohen et al., 2008).

Para la enfermedad de Chagas se ha demostrado la relación entre la distribución geográfica de algunas especies de Triatomos y variables como la temperatura, humedad, precipitación y altura sobre el terreno (Gorla, 2002). Por lo anterior, se consideró que el TWI tiene un potencial importante para ser vinculado con modelos de distribución de los vectores de la enfermedad de Chagas.

El TWI se calculó a partir de la metodología propuesta por Lanfri (2011), y utilizando una de las aproximaciones generadas por la autora; en su trabajo lleva a cabo la implementación en el lenguaje *Interactive Data Language* (IDL), de diferentes aproximaciones para el cálculo de un índice topográfico de humedad. El código desarrollado en su trabajo es de dominio público, en el sentido de que es permitido su uso, copia, modificación o redistribución con o sin fines de lucro, y su aproximación captura las características espaciales topográficas relativas a flujo y acumulación de agua de un modo más eficiente que otros modelos dinámicos de simulación hidrológica disponibles en otros softwares.

La metodología propuesta por Lanfri (2011) consiste en tomar un DEM del área de interés, a partir del cual, mediante el código de IDL desarrollado por la autora, se calcula 1) la pendiente del terreno, 2) dirección y 3) acumulación de flujo; una vez obtenidas estas capas, son ingresadas como *inputs* para finalmente calcular el TWI.

Para obtener el producto del área de interés de este estudio, se utilizó como entrada el DEM generado a partir de los datos SRTM explicado con anterioridad.

La figura 4.33 muestra el índice topográfico de humedad (TWI) del área de estudio generado para el servidor de Chagas de Argentina; los valores bajos de TWI representan las áreas que tienen menor capacidad de acumular agua, se corresponden con las zonas con mayor pendiente; a partir de la cual el agua tiende a ser escurrida hacia las zonas con menor pendiente, y por lo tanto mayor capacidad de acumular agua y humedad (valores mayores de TWI).

La salida del algoritmo desarrollado por Lanfri (2011) es un archivo GeoTiff, que se encuentra preparado para ser posteriormente publicado en el servidor de Chagas. La figura 4.33 resume la cadena de procesamiento anteriormente descrita, con la que se generó el producto de índice de humedad de terreno TWI del área de estudio. Se describen los pasos desde la incorporación del DEM al algoritmo de IDL desarrollado por Lanfri (2011).

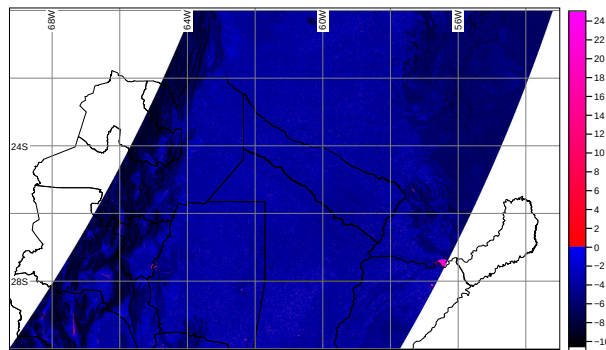


Figura 4.32: Índice topográfico de humedad (TWI) del área de estudio

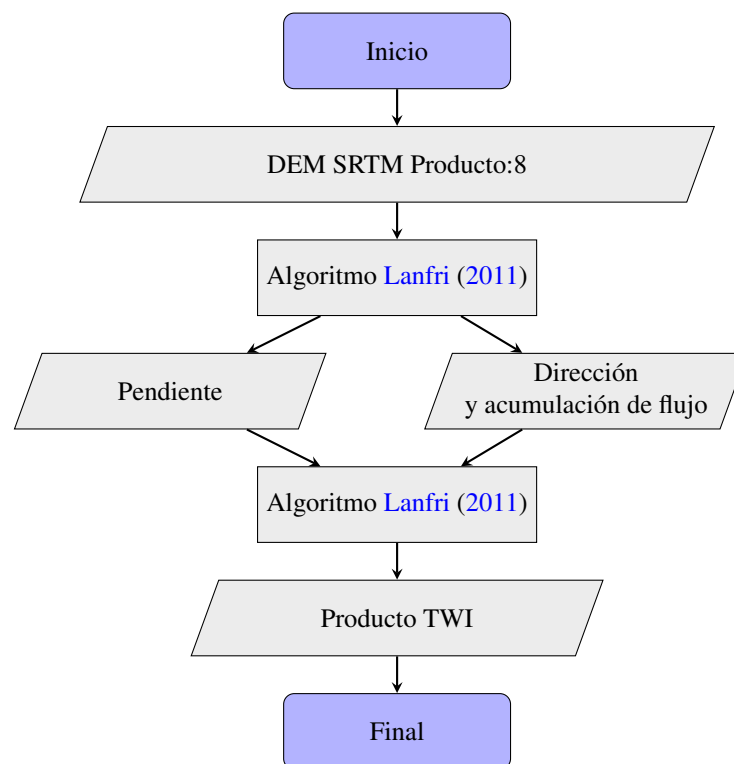


Figura 4.33: Flujo de trabajo para la obtención del producto Índice de humedad de terreno TWI a partir de un DEM SRTM

4.9 Incorporación de productos a la plataforma Argentina Web-GIS de Chagas

Como se explicó en el capítulo 3, dentro de la arquitectura cliente-servidor del Web-GIS de Chagas, el GeoServer se encuentra en la segunda capa del sistema y funciona como un puente entre la información de la base de datos y las aplicaciones con las que interactúan los usuarios (GeoExplorer). Con esta herramienta es posible configurar los estilos de las capas incluidas en el servidor, modificar tamaños, íconos, espacios de trabajo, fuentes de datos y un sinnúmero de posibilidades para moldear de la manera requerida los datos almacenados.

Así mismo, se puede conectar a diferentes tipos de bases de datos y unificar su información, entregar metadatos así como compartir los mapas por medios web utilizando principalmente los protocolos WMS, WCS y WFS.

Cabe destacar que para el proceso de subida y publicación de las capas en un sistema construido con las aplicaciones de OpenGeo-Suite es posible no utilizar el servidor de mapas (GeoServer), en dicho caso podríamos conectarnos directamente a la base de datos lo cual es válido y permitido por la mayoría de los Web-GIS, sin embargo las facilidades de Geoserver son muy prácticas e intuitivas para la gestión de los datos.

La herramienta web de administración de GeoServer ofrece una interface gráfica amigable y de fácil uso, no siendo necesario modificar directamente los ficheros de configuración y pudiendo probar nuevas configuraciones antes de salvarlas, lo cual permite la administración y mejoramiento de los datos geográficos a publicar.

Para incorporar capas, inicialmente se debe crear un espacio de trabajo, indicando que tipo de información se va a publicar. El orden lógico de la incorporación de capas en el servidor es: primero subir las capas al sistema de archivos del servidor y luego publicar las capas desde Geoserver, es decir hacerlas visibles para el usuario.

Estos son procedimientos fáciles de hacer, sin embargo cuando se tiene gran cantidad de capas de información para ser incorporadas al servidor de mapas, realizar esta tarea de forma manual (por cada archivo) se hace un proceso pesado y repetitivo. Esto puede ser fácilmente automatizado mediante el uso de herramientas de software para transferir archivos con sintaxis URL mediante intérprete de comandos, como cURL. El principal propósito de cURL es automatizar transferencias de archivos o secuencias de operaciones no supervisadas, por ejemplo para simular las acciones de usuarios, como la subida de capas a Geoserver.

El total de capas generadas con los 10 productos en formato raster anteriormente descritos fue 372, por lo que la incorporación a la plataforma Web-GIS de Chagas se hizo de forma automática; sin embargo, antes de incorporar las capas, se gestionó todo el entorno en el GeoServer para preparar la publicación de las mismas.

El primer paso de la gestión del entorno fue crear un espacio de trabajo en el cual se van a ubicar las capas, y posteriormente se generaron los estilos para los 10 productos. El procedimiento de la gestión del entorno para la publicación se describe a continuación:

Para entrar en la herramienta web (interfaz gráfica) del GeoServer será necesario introducir en nuestro navegador: <http://localhost:8080/geoserver/>. En donde "localhost" es la dirección ip de la máquina de escritorio en la que reposa instalado el servidor de mapas. Al ingresar a este sitio web, el GeoServer requiere de un nombre de usuario y contraseña para conectarse con la máquina de escritorio, la figura 4.34 muestra el menú de entrada a la interfaz gráfica del GeoServer; en la esquina superior derecha se encuentran los campos requeridos para ingresar al sistema.

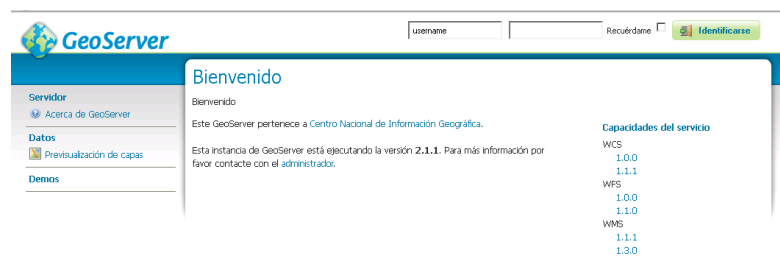


Figura 4.34: Menú de entrada a la interfaz gráfica del GeoServer

Una vez conectados al GeoServer, se procede con la creación de un espacio de trabajo. Los espacios de trabajo son usados para proporcionar elementos y atributos con nombre único a un grupo de capas, la definición de un espacio de nombres es esencial si el servidor va a ser público. El GeoServer usa los espacios de trabajo para evitar ambigüedades y para organizar y agrupar capas, de acuerdo con algún criterio, permitiendo la existencia de dos capas con el mismo nombre pertenecientes a diferentes espacios de trabajo. De este modo, las capas se identifican de la forma `espacio_de_trabajo:nombre_capa`.

Para crear el espacio de trabajo con nombre "ambientales" en donde se incorporan las 372 capas generadas se debe elegir la opción "Espacios de Trabajo" y hacer click sobre esta. Abierta la ventana 4.35 se introduce el nombre y la URI y se pulsa a enviar.

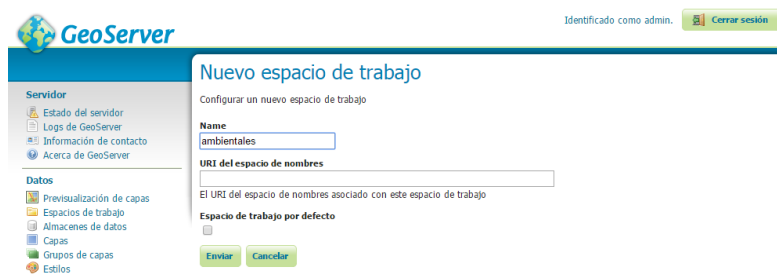


Figura 4.35: Ventana de creación de espacios de trabajo en el GeoServer

Posterior a la creación del espacio de trabajo se procede con la creación de los 10 estilos para cada uno de los productos generados; los estilos definen cómo se van a representar los datos. GeoServer trabaja con archivos de estilo basados en la especificación OGC SLD (*Styled Layer Descriptor*). Los archivos SLD son archivos XML que describen cómo va a ser visualizada una capa.

Al pulsar la opción "Estilos" aparece una lista de estilos por defecto, así como la opción de agregar un nuevo estilo o eliminar los estilos existentes.

Se puede crear un estilo nuevo, al que habrá que dar un nombre, utilizando la opción "Copiar" un estilo existente, o bien cargando un archivo SLD externo, para lo cual habrá que pulsar "Examinar", seleccionar el archivo SLD y posteriormente hacer click en Subir.

En cualquiera de los casos el XML completo se mostrará en la ventana del editor de texto 4.36, donde se podrá añadir, modificar o eliminar aquellos elementos del SLD que se considere oportuno.

Es importante resaltar que la generación del archivo SLD puede hacerse de forma manual

directamente en la ventana del editor de texto del GeoServer, o bien utilizando, de forma externa en algún GIS de escritorio que permita exportar archivos de estilo, por ejemplo, QGis.

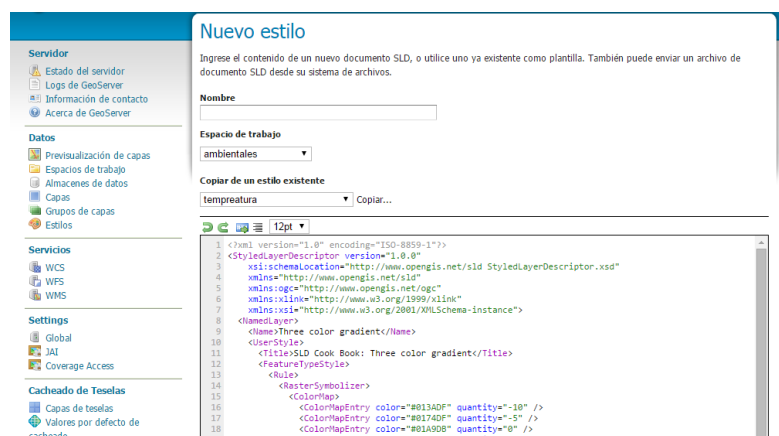


Figura 4.36: Editor de texto de la pantalla de creación de estilos del GeoServer

Una vez editado el estilo, se selecciona la opción "Validar" para que GeoServer compare el XML con el esquema de la especificación SLD, de modo que si no es compatible con éste, se impedirá guardar los cambios efectuados. En la parte superior de la pantalla se indica, por lo general, el error detectado por el GeoServer.

Si el GeoServer no detecta ningún error, aparecerá el mensaje correspondiente, y se procede a seleccionar la opción "Enviar" y el estilo creado es añadido a la lista de estilos disponibles.

Este procedimiento se realizó para los 10 productos raster generados y explicados con anterioridad, una vez terminada la generación de estilos, se termina la gestión del entorno en el GeoServer y se procede con la incorporación de las capas al sistema de archivos y su publicación.

Para realizar una incorporación masiva de datos al sistema Web-GIS es recomendada la herramienta de Linux cURL, que permite mediante línea de comando incorporar datos, gestionar espacios de trabajo y estilos, junto con varias funcionalidades relacionadas con la administración de la base de datos del sistema.

La herramienta cURL se incluyó en el siguiente script *bash* que primero sube cada capa y luego le asigna el estilo:


```

pathdatos=""
ipgeo=""
port=80
workspace=ambientales
user=""
pass=""

cd $pathdatos

vars=(acumulado_gdd delta dem Fourier_amplitud_ndvi_4 Fourier_amplitud_temperatura \
Fourier_amplitud_verdes Fourier_phase land_cover lst_sd ndvi tempreatura TWI \
wc_bio3 wc_bio7 wc_bio14 wc_bio18 wc_bio19 world_clim)

for var in ${vars[*]}
do
    list='ls $pathdatos/$var | grep .tif '

    for i in $list; do
        layer=$(echo $i | awk -F'.' '{print $1}')

        #subo la capa
        curl -u $user:$pass -v -XPUT -H "Content-type: image/tiff" \
        --data-binary @$pathdatos/$var/$i \ http://$ipgeo:$port/geoserver/rest/ \
        workspaces/$workspace/coveragestores/$layer/file.geotiff

        #seteo estilo
        curl -u $user:$pass -XPUT -H 'Content-type: text/xml' -d \
        '<layer><defaultStyle><name>$var</name><workspace>\
        $workspace</workspace></defaultStyle></layer>' \
        http://$ipgeo:$port/geoserver/rest/layers/$workspace:$layer

    done
done

exit 0

```

Como parámetros de entrada, el script recibe la información para conectarse al servidor de mapas, "ipgeo" corresponde a la ip de la maquina de escritorio en donde reposa la plataforma de Chagas, "port" es el puerto por el cual se genera el enlace, "user" y "pass" son los datos de seguridad para identificarse y generar la conexión. Por su parte "pathdatos" es la ruta a las carpetas en donde están ubicadas las capas en formato .tiff a ser subidas al servidor. El parámetro "workspace" se corresponde con el espacio de trabajo "ambientales" que se generó con anterioridad en el GeoServer.

La variable "vars" indica los estilos que deben ser asignados a las capas. El comando cURL se incorporó en un ciclo *For* que itera el procedimiento indicado para cada una de los 372 archivos .tiff. En el script se observa que el comando cURL aparece dos veces, la primera implementación es la que se usó para subir las capas al espacio de trabajo "ambientales" utilizando los parámetros de conexión, una vez incorporada la capa, la segunda implementación del cURL le asigna el estilo.

Finalmente, en cuanto terminó de ejecutarse este script, las capas quedan disponibles para la visualización por parte de los usuarios en el sistema Chagas (GeoExplorer), con este paso finaliza la cadena de procesamiento e incorporación a la plataforma Web-GIS de los productos raster generados.

Capítulo 5

Cadena de procesamiento de productos Vectoriales e incorporación en la plataforma Chagas

5.1 Planilla Chagas 6

El flujo de datos para el control vectorial y estratificación de riesgo de transmisión del *T. cruzi* en Argentina se inicia con la captura de datos, a través de las campañas de recolección que se realizan en terreno, hasta su almacenamiento y post-procesamiento en un sistema automático avanzado, mediante la planilla Chagas 6 (Apéndice A), generada por el ministerio de salud en el marco del Programa Nacional de Chagas (MSAL y CONAE, 2013).

La plataforma cliente-servidor al que se accede a través de la web mediante una interfaz intuitiva que permite que el usuario gestione la información integrada al mismo y proveniente de distintas fuentes con distintos niveles (localidad, departamento y provincia). El sistema posee un modulo de gestión de datos que cuando el usuario ingresa las planillas excel al sistema, estas son procesadas y almacenadas en una base de datos objeto-relacional (Base de datos espacial PostGIS ver 3). Luego, estos datos deben ser extraídos y visualizados correctamente para permitirle al usuario del sistema la visualización y manipulación de los mismos. El modulo de Gestión de datos que implementa esta solución fue diseñado e implementado por la CONAE de acuerdo a los requerimientos descritos a continuación, teniendo en cuenta que existen usuarios Nacionales y Provinciales, ambos con distintos privilegios de acceso a la información:

- Realizar búsquedas/consultas SQL (*Structured Query Language*) a la base de datos
- Permitir editar los datos cargados previamente (esta funcionalidad solo está disponible para algunos niveles de usuarios).
- Tiempo de edición de la base de datos: Pasado los 30 días de la carga de la planilla, los datos NO pueden ser modificados.

- Permitir eliminar los datos cargados previamente (esta funcionalidad solo está disponible para algunos niveles de usuarios). Las modificaciones solo pueden realizarse por quien ingresa los datos.
- Permitir la descarga de los resultados de las búsquedas SQL.
- La descarga de la base de datos y los resultados de las búsquedas deberán ser guardadas en archivos con formato compatible con sistemas de administración de bases de datos de tipo CSV (*Comma Separated Values*).

Los datos que se incorporen a este sistema Web-GIS se colectarán durante las acciones regulares de los programas nacionales y/o provinciales en la evaluación y rociado (tratamiento) de viviendas urbanas/rurales de zona endémica y/o la acción de los agentes de salud que realicen evaluaciones periódicas en viviendas en regiones endémicas ([MSAL y CONAE, 2013](#)). Una vez que se incorpore esta nueva planilla de Chagas 6 al sistema (carga virtual), se deberá eliminar las antiguas Planillas de Chagas 6 (Para efectores) y Chagas 7 (Para supervisores).

La planilla Chagas 6 es distribuida en formato *Excel*, es de un llenado intuitivo con la información que se pide, es importante no modificar el formato pre-existente de la misma. La planilla Chagas 6 digital, colecta de datos de las viviendas para actividades de control vectorial. Deberá cargarse una planilla por cada localidad que se haya visitado y de las cuales se posean datos nuevos.

La implementación de la planilla en las evaluaciones de terreno permite que cada vivienda sea ubicada espacialmente, mediante la utilización de un GPS (*Global Positioning System*), de esta forma en la plataforma virtual se crea una base de datos espacial que contiene la información de las localidades en donde se hizo el control vectorial, que proporciona datos a nivel vivienda de presencia de triatomíneos, si se realizó control químico, habitantes de la vivienda por grupos de edades, condiciones de la infraestructura de la vivienda, e índices de infestación e infección intra/peridomiciliaria.

Es importante resaltar que la CONAE y MSAL han sumado esfuerzos para lograr que las entidades a nivel local, provincial y nacional proporcionen la información de la planilla y realicen la carga virtual a la plataforma Web-GIS, desde su interfaz de usuario. El éxito y utilidad del sistema esta directamente relacionado con los esfuerzos de las instituciones por compartir la información recolectada en campo.

5.2 PgAdmin y consultas SQL

Para generar productos en formato vectorial con valor agregado para el estudio de Chagas en Argentina, la fuente principal de datos es la base incorporada en el Web-GIS a partir de la planilla Chagas 6. Para interactuar con la base de datos que reposa en la parte interna de un

sistema construido bajo la arquitectura cliente-servidor (explicada en el capítulo 3) se han desarrollado varias herramientas que combinadas con sentencias de línea de comando de Linux, logran realizar de una forma adecuada la gestión de los datos y posteriormente, la creación de capas vectoriales (shapes) a partir de ellos.

Al ser suministrada nueva información de la planilla por parte de los usuarios al sistema, esta se incorpora a la estructura de base de datos espacial, la cual consiste en una tabla relacional.

Una base de datos relacional está compuesta por tablas o relaciones donde una tabla tiene una serie de atributos o columnas y esta formada por filas que tienen un valor para cada uno de los atributos. Una base de datos PostGIS es considerada como una base de datos orientada a objetos por sus principales características, así como también por los tipos de datos que manipula (relaciona atributos con clases geométricas en el espacio).

Cada atributo está unido correlacionalmente con otras tablas y todos responden a una columna de tipo de datos de geometría (punto, línea, polilíneas, etc), de esta forma se le asigna un lugar en el espacio.

Las búsquedas SQL (*Structured Query Language*) son el método para acceder a los datos que reposan en la base de datos correlacional del PostGIS.

SQL es el lenguaje estándar de definición, manipulación y control de bases de datos relacionales. Es un lenguaje declarativo, y mediante su uso sólo hay que indicar qué se quiere hacer o extraer de las tablas, es muy parecido al lenguaje natural; por estas razones, y como lenguaje estándar, el SQL es un lenguaje con el que se puede acceder a todos los sistemas relacionales comerciales actuales. Una consulta o búsqueda SQL consiste en listar toda la información disponible en la base de datos para ciertos atributos que se soliciten, y adicionalmente se pueden modificar, borrar, mostrar y agregar datos a la tablas.

GeoServer permite la creación de capas a partir de consultas SQL. Una consulta SQL genera una nueva tabla a partir de las tablas existentes, con los datos que sean solicitados. Estas tablas son temporales, y mediante diferentes comandos pueden ser incorporadas de forma permanente en la base de datos espacial; un ejemplo de esto es la creación de una capa (*shape*) a partir de la consulta.

El proceso para poder generar un archivo .shp (shape) a partir de la base de datos se hace mediante una función de PostgreSQL desde línea de comando. La función `pgsql2shp` se conecta directamente con la base de datos, requiriendo datos de acceso (ip del servidor, usuario y clave) y convierte una tabla temporaria, que es generada a partir de la consulta SQL, en un archivo de figuras .shp.

El procedimiento de la generación del archivo shape es sumamente simple mediante el comando `pgsql2shp`, sin embargo requiere de tener la consulta SQL preparada para que solicite los datos pertinentes y genere la tabla de atributos de la forma deseada; por lo anterior es recomendable construir las consultas de manera preliminar y revisar las tablas (atributos) que se generan a partir de ellas.

Existe un sistema de gestión de bases de datos llamado PgAdmin III. Es una herramienta de

código abierto para la administración de bases de datos PostgreSQL y derivados (PostGIS), que incluye una interfaz administrativa gráfica, herramienta de consulta SQL, editor de códigos y varias funcionalidades avanzadas. PgAdmin fue diseñado para responder a las necesidades de la mayoría de los usuarios, desde escribir simples consultas SQL hasta desarrollar bases de datos complejas. La interfaz gráfica soporta todas las características de PostgreSQL y hace simple la administración. Está disponible en más de una docena de lenguajes y para varios sistemas operativos, incluyendo Microsoft Windows y Linux (PgAdmin, 2015).

Previo a la generación de los archivos .shp, se usó la aplicación gráfica de PgAdmin para generar las consultas deseadas que posteriormente son usadas para obtener los productos vectoriales de valor agregado que serán explicados a continuación. Con lo anterior, fueron verificadas de forma manual las tablas generadas a partir de las consultas SQL. La figura 5.1 muestra la ventana de conexión del PgAdmin III con la base de datos, en donde se encuentra la herramienta para generación de consultas SQL que fue utilizada.

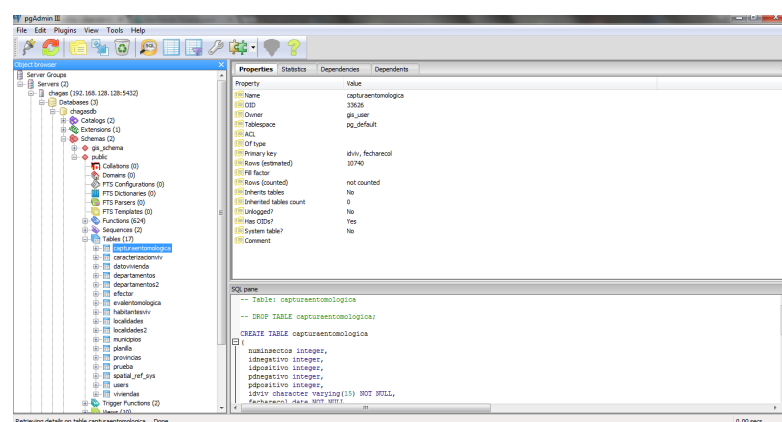


Figura 5.1: Interfaz gráfica de PgAdmin III con asistente para consultas SQL manuales

5.2.1 Generación de consultas

Se generaron nueve consultas SQL, para generar las capas vectoriales en formato .shp, con información de la planilla Chagas 6.

Generar capas geográficas a partir de los datos recolectados en las viviendas, tiene un potencial importante a la hora de relacionar la distribución espacial de *T. infestans* u otro vector de Chagas a nivel localidad, con variables demográficas, y generar modelos que vinculen los productos de variables ambientales explicados en el capítulo 4, para ser utilizados en la predicción de áreas con alto riesgo de infestación.

Por lo anterior, se buscó generar capas que expresen de manera sintética información relevante para la epidemiología de la enfermedad y la infestación vectorial, así como información del estado de vigilancia actual del mal de Chagas en el país. Las capas generadas fueron 9: 1. viviendas registradas en la base de datos y 2. viviendas evaluadas. Estas dos capas dan una idea del estado actual de la revision epidemiológica del país, y de la cantidad de datos que han sido ingresados al sistema por parte de las instituciones encargadas de la vigilancia domiciliaria.

3. Viviendas evaluadas por localidad, 4. viviendas evaluadas sin infestación y 5. viviendas evaluadas con infestación. Los productos 3, 4 y 5 resumen el estado de la infestación domiciliaria en las localidades.

Las capas 6 y 7 indican el promedio de animales intra-domiciliares y presentes en el peri-domicilio, respectivamente, estas capas son de importancia para ser incluidos en modelos de estimación de riesgo de transmisión por el papel que cumplen las mascotas o animales de corral en el ciclo de transmisión de la enfermedad (ver capítulo 2).

Por último las capas 8 y 9 corresponden al índice de infestación intra-domiciliaria (IIID) y peri-domiciliaria (IIPD) respectivamente.

A continuación se describe el procedimiento para obtener las 9 capas anteriormente mencionadas, que consiste en generar y probar la consulta SQL manualmente en el PgAdmin, utilizar el comando "pgsql2shp", generar el archivo de estilo de GeoServer y publicación de la capa en el sistema Web-GIS.

5.2.1.1 Producto 1: Viviendas registradas

Para generar la capa vectorial con el total de viviendas incluidas en la base de datos (con o sin evaluación entomológica) se construyó la siguiente consulta SQL, utilizando la herramienta gráfica del PgAdmin III.

```
SELECT viv.idviv , viv.the_geom, loc.codpcia , loc.coddepto ,
loc.codloc , loc.nombreloc
FROM localidades as loc ,

(SELECT viviendas.the_geom, viviendas.codloc , viviendas.coddepto ,
viviendas.codpcia , viviendas.idviv from viviendas) viv
WHERE viv.coddepto=loc.coddepto and viv.codloc=loc.codloc and
viv.codpcia=loc.codpcia

ORDER BY loc.coddepto ,loc.codpcia , loc.codloc
```

La consulta selecciona de toda la base de datos, únicamente los atributos "idviv", que representa el identificador único para cada una de las casas registradas, "the_geom", que corresponde al código geográfico que PostGIS le asigna a cada elemento para ubicarlo en el espacio; seleccionar este atributo es de suma importancia porque es a partir del cual se creará el shape, si no llegara a estar presente en la consulta, sería imposible generar una capa con puntos geográficos. Los atributos "Codpcia", "Coddepto" y "Codloc" son los códigos con los que se identifican en la base de datos las provincias, departamentos y localidades, respectivamente. Por último el atributo "Nombreloc" representa el nombre de la localidad en donde se ubica la vivienda registrada.

La figura 5.2 muestra la tabla que resulta de la consulta anteriormente descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar, siendo esta la principal ventaja y ayuda que proporciona el PgAdmin al proceso de creación de capas vectoriales a partir de la base de datos PostGIS.

Query - chagadb on gi_user@192.168.128.128:5432

SQL Editor

```
(SELECT vivienda.the_geom, viviendas.codloc, viviendas.coddepto,
viviendas.codpcia, viviendas.idvivi from viviendas) vivi
WHERE vivi.coddepto=loc.coddepto and vivi.codloc=loc.codloc and
vivi.codpcia=loc.codpcia
ORDER BY loc.coddepto, loc.codpcia, loc.codloc
```

Output pane

idvivi	the_geom	codpcia	coddepto	codloc	nombrelot
895	300087008	010100002 30	008	7	COLONIA HUGHES
896	300087006	010100002 30	008	7	COLONIA HUGHES
897	300087007	010100002 30	008	7	COLONIA HUGHES
898	300087008	010100002 30	008	7	COLONIA HUGHES
899	300087009	010100002 30	008	7	COLONIA HUGHES
900	300087010	010100002 30	008	7	COLONIA HUGHES
901	300087011	010100002 30	008	7	COLONIA HUGHES
902	300087012	010100002 30	008	7	COLONIA HUGHES
903	300087013	010100002 30	008	7	COLONIA HUGHES
904	300087014	010100002 30	008	7	COLONIA HUGHES
905	300087015	010100002 30	008	7	COLONIA HUGHES
906	300087016	010100002 30	008	7	COLONIA HUGHES
907	300087017	010100002 30	008	7	COLONIA HUGHES
908	300087018	010100002 30	008	7	COLONIA HUGHES
909	300087019	010100002 30	008	7	COLONIA HUGHES
910	58014103	010100002 58	014	2	AGUADA DE SAN ROQUE
911	58014118	010100002 58	014	2	AGUADA DE SAN ROQUE
912	58014120	010100002 58	014	2	AGUADA DE SAN ROQUE
913	58014111	010100002 58	014	2	AGUADA DE SAN ROQUE
914	58014101	010100002 58	014	2	AGUADA DE SAN ROQUE

OK

Unix Ln 10, Col 46, Ch 365

10257 rows. 2685 ms

Figura 5.2: Salida del PgAdmin donde se detalla el resultado de la consulta SQL del total de viviendas agregadas a la base de datos

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp`. Este comando conecta directamente con la base de datos y convierte una tabla en un archivo de figuras (vector.shp) y la sintaxis que se usa es:

pgsql2shp [opciones] [base de datos] [consulta]

Las opciones que permite definir la función `pgsql2shp` (Tabla 5.1) corresponden a los datos de conexión con la base de datos, como lo son , el host donde se encuentra físicamente la base de datos, el usuario y clave de acceso al sistema.

Tabla 5.1: Opciones para introducir datos en PostGIS con la función *pgsql2shp*

Opciones	Significado
-f	<i>Filename:</i> Usar esta función para especificar el nombre del archivo .shp a crear.
-h	<i>Host:</i> Permite definir la conexión a otra maquina diferente a la local.
-p	<i>Port:</i> Permite definir el puerto para realizar la conexión.
-u	<i>User:</i> Conectar con la base de datos con un usuario específico.
-P	<i>Password:</i> Contraseña de acceso.
-g	<i>Geometry_column:</i> Permite especificar la columna con datos geométricos.

La función `pgsql2shp` se usó de la siguiente forma para generar la capa `viviendas.total.shp`, recibiendo como entrada la consulta SQL descrita con anterioridad para generar la tabla de atributos:

```
pgsql2shp -f viviendas_total.shp -h localhost -u use_gis -P **** chagadb viv_total.sql
```

Una vez creado el shape, se procede con la incorporación del mismo al GeoServer, y su posterior publicación al GeoExplorer. A diferencia de los productos raster, en los que la incorporación y publicación de los datos al sistema Web-GIS se hizo de forma automática con el comando cURL, para las capas vectoriales, debido a que son nueve productos, se realizó de forma manual con las opciones de la interfaz gráfica del GeoServer; con lo anterior, fueron utilizados dos métodos diferentes de incorporación de información al servidor (automático y manual).

Para incorporar la capa al GeoServer se debe crear un nuevo origen de datos vectoriales, para esto, sobre la ventana principal del GeoServer, se accede a la opción "Almacenes de datos", seguido de la opción "Agregar nuevo almacén" que cambia la ventana por un listado de fuentes de datos de diferentes tipo (raster, shape, wms, wfs entre otros). Se selecciona la opción "Shapefile - ESRI(tm) Shapefiles (*.shp)" y aparece la ventana de configuración del nuevo origen de datos vectoriales (figura 5.3).

Para crear el nuevo origen de datos es necesario indicar: 1) El espacio de trabajo en donde serán visualizadas las capas "GRAL", 2) El nombre del origen de datos "vectores_planilla" y 3) La ruta con la ubicación de la capa vectorial o carpeta de archivos dentro del sistema PostGIS, en donde reposan los shapes generados con anterioridad.



Figura 5.3: Ejemplo de la generación de nuevo origen de datos vectoriales en GeoServer

Con el nuevo origen de datos creado, se encuentra preparado el entorno para ser incorporada la capa al GeoServer. En la ventana principal se accede a la opción "Capas", que despliega un listado con todos los orígenes de datos existentes; se selecciona la opción "vectores_planilla" recientemente creada y se abre una nueva ventana (figura 5.4) con el listado de las shapes presentes en la carpeta del PostGIS con la que se conecta el GeoServer.



Figura 5.4: Ejemplo de la ventana de incorporación y publicación de nuevas capas del GeoServer

Se selecciona la capa "viviendas_total" y se accede a la opción "publicar". Con esto, se despliega la ventana de edición de la capa, que corresponde a todas las opciones de visualización de la misma. GeoServer permite configurar y habilitar opciones de: 1) Nombre, 2) Resumen o breve descripción del producto, 3) Sistema de referencia geográfico, 4) Encuadre o límites geográficos de extensión de la capa, y 5) detalles de la tabla de atributos. Para la publicación del producto "viviendas_total" se usaron los parámetros por defecto que se asignan en todas las opciones y el encuadre geográfico se calculó a partir de los límites de la propia capa.

Con lo anterior, solo resta asignar el estilo con el que serán visualizados los datos en la plataforma Web-GIS. Debido a que la capa "viviendas_total" representa con un punto la ubicación espacial de las viviendas incluidas en la base de datos, no fue necesario crear un estilo particular para la capa, por lo que se usó el archivo point.sld que esta por defecto en el GeoServer.

Una vez asignado el estilo se selecciona la opción "Guardar" y finalmente la capa "viviendas_total" queda incorporada y publicada en la plataforma Web-GIS del sistema Chagas.

5.2.1.2 Producto 2: Viviendas evaluadas

El segundo producto vectorial que se generó fue el de viviendas evaluadas, que representa los puntos geográficos en donde se ubican cada una de las viviendas registradas en la base de datos que fueron evaluadas por alguno de los efectores del control vectorial domiciliario.

La siguiente consulta SQL se realizó en el PgAdmin de forma manual para crear una tabla con los atributos deseados para este producto:

```
SELECT nombreloc , t1.*
FROM localidades ,
```

```
(SELECT viv2.the_geom, viv2.codpcia as codpcia2, viv2.coddepto as coddepto2,
viv2.codloc as codloc2, eval2.estadoviv as evaluadas, eval2.numidad
as numidad2, eval2.numidn as numidn2, eval2.presidad as
presidad2, eval2.presidn as presidn2
FROM evalentomologica as eval2, viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%'
ORDER BY codloc, coddepto, codpcia, eval2.estadoviv) t1

WHERE t1.codpcia2=localidades.codpcia and t1.coddepto2=localidades.coddepto
and t1.codloc2=localidades.codloc
```

La consulta selecciona de toda la base de datos, únicamente los atributos "the_geom", con las coordenadas geográficas de las viviendas, "nombreloc" que contiene el nombre de la localidad en donde se ubica la vivienda registrada, "Codpcia", "Coddepto" y "Codloc" con los códigos que identifican a las provincias, departamentos y localidades en la base de datos respectivamente, y "evaluadas" que corresponde a la columna del estado de la vivienda, para verificar que todas las entradas de la tabla cumplan con el requisito de haber sido visitada por el agente efector.

La figura 5.5 muestra la tabla que resulta de la consulta anteriormente descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

```

Query - chagabos on gis_user@192.168.128.128:5432 *
File Edit Query Favourites Macros View Help
[Icons] chagabos on gis_user@192.168.128.128:5432
SQL Editor Graphical Query Builder
Previous queries [v] [Delete] [Delete All]
[SELECT vivi2.the_geom, vivi2.codpcia as codpcia2, vivi2.coddepto as coddepto2,
vivi2.codloc as codloc2, eval2.estadoviv as evaluadas, eval2.munind as
as munid2, eval2.munind as munid2, eval2.presid as
presid2, eval2.presid as presid2
FROM evalentomologica as eval2, viviendas as vivi2
WHERE vivi2.idviv=eval2.idviv and eval2.estadoviv
]
Output pane
Data Output Explain Messages History
the_geom nombreloc codpcia2 coddepto2 codloc2 evaluadas
geometry character varying(50) character varying(2) character varying(3) character varying(4) character varying(50)
vivi2 EVALUADA
674 0101000020E61000004E30 BASUALDO 30 042 11 Evaluada
675 0101000020E61000004E30 BASUALDO 30 042 11 Evaluada
676 0101000020E610000013F2 BASUALDO 30 042 11 Evaluada
677 0101000020E610000035EB BASUALDO 30 042 11 Evaluada
678 0101000020E610000035EB BASUALDO 30 042 11 Evaluada
679 0101000020E610000035EB BASUALDO 30 042 11 Evaluada
680 0101000020E610000035EB BASUALDO 30 042 11 Evaluada
681 0101000020E610000035EB BASUALDO 30 042 11 Evaluada
682 0101000020E610000035EB BASUALDO 30 042 11 Evaluada
683 0101000020E610000035EB BASUALDO 30 042 11 Evaluada
684 0101000020E6100000DCD4 BASUALDO 30 042 11 Evaluada
685 0101000020E6100000DCD4 BASUALDO 30 042 11 Evaluada
686 0101000020E6100000DCD4 BASUALDO 30 042 11 Evaluada
687 0101000020E6100000DCD4 BASUALDO 30 042 11 Evaluada
688 0101000020E6100000A0AD BASUALDO 30 042 11 Evaluada
689 0101000020E6100000A0AD BASUALDO 30 042 11 Evaluada
690 0101000020E6100000E0E30 BASUALDO 30 042 11 Evaluada
691 0101000020E6100000E0E30 BASUALDO 30 042 11 Evaluada
692 0101000020E6100000B474 BASUALDO 30 042 11 Evaluada
693 0101000020E6100000B474 BASUALDO 30 042 11 Evaluada

```

Figura 5.5: Salida del PgAdmin resultado de la consulta SQL del total de viviendas evaluadas

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp` de la siguiente forma para generar la capa `viviendas_evaluadas.shp`, recibiendo como entrada la consulta SQL descrita con anterioridad para generar la tabla de atributos:

```
pgsql2shp -f viviendas_evaluadas.shp -h localhost -u user -P *** chagasdb viv_eval.sql
```

Con el shape "viviendas_evaluadas.shp" creado, se procede con la incorporación al GeoServer y su posterior publicación.

Como el almacén de datos "vectores_planilla" fue creado con anterioridad, el GeoServer tiene

el vínculo con la capa del PostGIS, y para publicar la capa solo se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa "viviendas_evaluadas".

Se usaron las opciones de edición de la capa por defecto, y se calcularon los límites geográficos a partir de la capa.

Por su parte el estilo "point" que también está por defecto incorporado al GeoServer, fue usado para representar con un punto cada una de las viviendas presentes en la capa.

Con lo anterior el producto "viviendas_evaluadas.shp" finalmente quedó incorporado y publicado en la plataforma Web-GIS de Chagas.

5.2.1.3 Producto 3: Localidades evaluadas

El tercer producto vectorial que se generó corresponde al total (suma) de viviendas que fueron evaluadas por localidad, esta capa es importante debido a que proporciona información real del estado de seguimiento epidemiológico por parte de los efectores. Para crear la capa fue necesario realizar la siguiente consulta SQL, que dentro de su código recibe función "count" que es la que se encarga de sumar el número de registros por localidad y lo ubica en una columna atributo denominada "Evaluadas":

```
SELECT the_geom, nombreloc, t1.*
FROM localidades ,

(SELECT viv2.codpcia as codpcia2, viv2.coddepto as coddepto2, viv2.codloc
as codloc2, count(eval2.estadoviv) as evaluadas
FROM evalentomologica as eval2, viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%'
GROUP BY codloc, coddepto, codpcia) t1

WHERE t1.codpcia2=localidades.codpcia and t1.coddepto2=localidades.coddepto
and t1.codloc2=localidades.codloc

ORDER BY codpcia, coddepto, codloc
```

La búsqueda SQL selecciona de toda la base de datos, únicamente los atributos "the_geom", con las coordenadas geográficas de las viviendas, "nombreloc" con el nombre de la localidad en donde se ubica la vivienda registrada, "Codpcia", "Coddepto" y "Codloc" con los códigos que identifican a las provincias, departamentos y localidades en la base de datos respectivamente, y "evaluadas" que como se mencionó con anterioridad, corresponde al total de las viviendas evaluadas en dicha localidad.

La figura 5.6 muestra la tabla que resulta de la consulta anteriormente descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

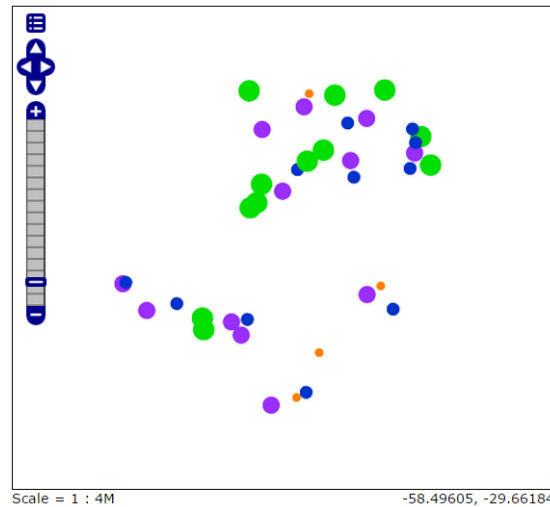


Figura 5.7: Representación gráfica de la capa publicada utilizando la herramienta *OpenLayers*

5.2.1.4 Producto 4: Localidades sin infestación

La capa total de viviendas evaluadas fue separada en dos productos que por sí mismos generan información útil para ser incorporada en modelos de predicción de los factores que se asocian con la presencia de vinchucas en los hogares. Los productos son el número de viviendas sin y con infestación por localidad.

El cuarto producto es viviendas sin infestación y fue generado con la siguiente consulta SQL de la base de datos, que una vez más usa la función "count" para totalizar el número de registros que cumplan con los parámetros de haber sido evaluada la vivienda y no presentar infestación al interior del domicilio ni en el peridomicilio:

```
SELECT the_geom, t1.*, nombreloc
FROM localidades ,

(SELECT viv2.codpcia as codpcia2, viv2.coddepto as coddepto2, viv2.codloc
as codloc2, count(eval2.estadoviv)
FROM evalentomologica as eval2, viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%' and
(((eval2.numidad is NULL or eval2.numidad = 0 ) and
(eval2.numidn is NULL or eval2.numidn = 0) ) and
((eval2.presidad='No' or eval2.presidad is NULL ) and
(eval2.presidn='No' or eval2.presidn is NULL)))

GROUP BY codloc ,coddepto , codpcia) t1

WHERE t1.codpcia2=localidades.codpcia and
t1.coddepto2=localidades.coddepto and
t1.codloc2=localidades.codloc

ORDER BY codloc ,coddepto , codpcia
```

La búsqueda SQL selecciona de toda la base de datos, únicamente los atributos "the_geom", con las coordenadas geográficas de las viviendas, "nombreloc" con el nombre de la localidad en donde se ubica la vivienda registrada, "Codpcia", "Coddepto" y "Codloc" con los códigos que identifican a las provincias, departamentos y localidades en la base de datos

respectivamente, y "count" que indica el número de viviendas sin infestación.

La figura 5.8 muestra la tabla que resulta de la consulta anteriormente descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

	the_geom	the_geom geometry	codpcia2	coddepto2	codloc2	count	nombreloc
			character varying(2)	character varying(3)	character varying(4)	bigint	character varying(50)
146	0101000020E610C58			035	7	1	EL CHOCÓN
147	0101000020E610C30			077	7	29	COLONIA LA LLAVE
148	0101000020E610C30			091	7	3	GOBERNADOR MANZILLA
149	0101000020E610C58			091	7	12	BUFA RAMOQUE
150			46	119	7	65	LAS TALAS
151	0101000020E610C46			021	8	64	PINCHAS
152	0101000020E610C30			035	8	8	COLONIA SAN ISIDRO
153	0101000020E610C30			042	8	202	SAN JOSE DE FELICIANO
154	0101000020E610C58			063	8	1	CHORRILACA
155	0101000020E610C30			091	8	17	GOBERNADOR SOLA
156	0101000020E610C42			112	8	24	FUELEN
157	0101000020E610C46			119	8	154	SALICAS
158	0101000020E610C46			007	9	7	UDPINANGO
159	0101000020E610C58			014	9	29	LOS CHIRIPIDOS
160	0101000020E610C46			028	9	414	SUNUNCOO
161			30	035	9	10	COLONIA SAN LORENZO
162	0101000020E610C30			042	9	27	SAN VICTOR
163	0101000020E610C14			077	9	75	CHANCANI
164	0101000020E610C30			091	9	16	GUARDAMONTE
165	0101000020E610C58			112	9	23	RAMÓN M. CASTRO
166	0101000020E610C46			119	9	86	SAN BLAS

Figura 5.8: Salida del PgAdmin resultado de la consulta SQL con la suma de viviendas sin infestación por localidad

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp` de la siguiente forma para generar la capa `loc_sin_infestacion.shp`, recibiendo como entrada la consulta SQL descrita con anterioridad para generar la tabla de atributos:

```
pgsql2shp -f loc_sin_infestacion.shp -h host -u user -P *** chagasdb sin_inf.sql
```

Con el shape "loc_sin_infestacion.shp" creado, se procede con la incorporación al GeoServer y su posterior publicación.

Para publicar la capa solo se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa "loc_sin_infestacion.shp".

Como cada entrada de esta capa representa un número con la suma de viviendas, el estilo creado para el producto 3, fue usado también para esta capa, de esta forma conforme aumenta el número de viviendas sin infestación dentro de la localidad, el punto que representa la ubicación geográfica de la localidad será más grande y de un color diferente. Con lo anterior el producto "loc_sin_infestacion.shp" finalmente quedó incorporado y publicado en la plataforma Web-GIS de Chagas.

5.2.1.5 Producto 5: Localidades con infestación

El quinto producto vectorial generado fue la capa de localidades evaluadas con infestación, que representa el total por localidad de las viviendas que tuvieron visita domiciliaria por parte de los efectores de control, en donde fueron encontradas vinchucas, dentro de la vivienda o en su peridomicilio.

El producto fue generado con la siguiente consulta SQL de la base de datos, con la función "count" que totaliza el número de registros que cumplan con los parámetros de haber sido evaluada la vivienda y tener infestación al interior del domicilio o en el peridomicilio:

```
SELECT the_geom, t1.*, nombreloc
FROM localidades ,

(SELECT viv2.codpcia as codpcia2, viv2.coddepto as coddepto2,
viv2.codloc as codloc2, count(eval2.estadoviv)
FROM evalentomologica as eval2, viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%' and
(eval2.numidad > 1 or eval2.numidn > 1 or eval2.presidad = 'Si'
or eval2.presidn='Si')

GROUP BY codloc ,coddepto , codpcia) t1

WHERE t1.codpcia2=localidades.codpcia and t1.coddepto2=localidades.coddepto
and t1.codloc2=localidades.codloc

ORDER BY codloc ,coddepto , codpcia
```

La búsqueda SQL selecciona de toda la base de datos, los atributos descritos con anterioridad, "the_geom", con las coordenadas geográficas de las viviendas, "nombreloc" con el nombre de la localidad en donde se ubica la vivienda registrada, "Codpcia", "Coddepto" y "Codloc" con los códigos que identifican a las provincias, departamentos y localidades en la base de datos respectivamente, y "count" que indica el número de viviendas con infestación.

La figura 5.9 muestra la tabla que resulta de la consulta descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

Query - chagesdb on gis_user@192.168.128.128:5432 *

File Edit Query Favourites Macros View Help

chagesdb on gis_user@192.168.128.128:5432

SQL Editor Graphical Query Builder

Previous queries

Delete Delete All

```
SELECT the_geom, t1.*, nombreloc
FROM localidades,
(
  utm3_cedente, utm3_cedente2, utm3_cedente3, utm3_cedente4
)
```

Output pane

	the_geom	codpca2	coddepto2	codloc2	count	nombreloc
	geometry	character varying(2)	character varying(3)	character varying(4)	bigint	character varying(50)
1	0101000020E610C46	007	10	1	VILLA MAZAN	
2	0101000020E610C46	119	10	1	SCHAQUI	
3	0101000020E610C46	021	11	1	SANTA CRUZ	
4	0101000020E610C46	119	13	2	ALPASINCHE	
5	0101000020E610C46	119	14	1	LOS ROBLES	
6	0101000020E610C46	105	20	4	FIAMBALA	
7	0101000020E610C10	084	21	7	SIJANI	
8	0101000020E610C46	007	3	5	BABADOS DEL PANTANO	
9	0101000020E610C46	021	3	3	ANTILLACO	
10	0101000020E610C10	042	3	3	CAPAYAN	
11	0101000020E610C10	070	39	3	QUIROS	
12	0101000020E610C10	042	4	2	CHUMBUCHA	
13	0101000020E610C46	119	4	1	CHAUPICHUASI	
14	0101000020E610C10	070	40	2	RECREO	
15	0101000020E610C10	105	45	12	TINGASTA	
16	0101000020E610C46	119	5	2	CUTIPAN	
17	0101000020E610C10	070	51	5	Las Esquilas	
18	0101000020E610C46	007	6	1	MACHIGASTA	
19	0101000020E610C46	119	7	1	LAS TALAS	
20	0101000020E610C46	119	8	1	SALICAS	
21	0101000020E610C46	028	9	16	GUANDACOL	

OK

1 row 16 rows 63 ms

Figura 5.9: Salida del PgAdmin resultado de la consulta SQL con la suma de viviendas con infestación por localidad

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp` de la siguiente forma para generar la capa `loc_con_infestacion.shp`, recibiendo como entrada la consulta SQL para generar la tabla de atributos:

```
pgsql2shp -f loc_con_infestacion.shp -h host -u user -P *** chagasdb con_inf.sql
```

Con el shape "loc_con_infestacion.shp" creado, se procede con la incorporación al GeoServer y su posterior publicación.

Para publicar la capa solo se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa "loc_con_infestacion.shp".

Como cada entrada de esta capa representa un numero con la suma de viviendas, el estilo creado para el producto 3, fue usado también para esta capa, de esta forma conforme aumenta el numero de viviendas sin infestación dentro de la localidad, el punto que representa la ubicación geográfica de la localidad sera más grande y de un color diferente. Con lo anterior el producto "loc_con_infestacion.shp" finalmente quedó incorporado y publicado en la plataforma Web-GIS de Chagas.

5.2.1.6 Producto 6: Animales intradomicilio

La presencia de animales en las viviendas o en su peridomicilio ha sido asociada con el ciclo de transmisión del *T. cruzi* (Parra-Henao et al., 2015), debido a que constituyen alimento para los vectores, así mismo, también ha sido discutida la actividad de los animales como escudo para la transmisión, debido a que las vinchucas se alimentan de la sangre de los animales, dejando a un lado el recurso que para ellas suministra la sangre de las personas en las viviendas (Ministerio de Salud Argentina, 2012).

En congruencia se consideró oportuna la generación de un producto que represente el número de animales presentes en el domicilio y que conviven con los humanos. El sexto producto vectorial que se generó corresponde al promedio por localidad de animales en el intradomicilio. Los agentes efectores en las visitas de control en las viviendas toman el dato del número de perros, gatos, gallinas y cabras, considerando como una generalidad que los animales que permanecen la mayor parte del tiempo dentro del domicilio (pernoctar dentro de la vivienda) son los perros y gatos, estos fueron los que se incluyeron en el cálculo del promedio.

Para crear la capa fue necesario realizar la siguiente consulta SQL, que dentro de su código recibe función "sum" que es la que se encarga de sumar el número de gatos y perros en las viviendas, que posteriormente se divide en el número de viviendas evaluadas de cada localidad presente en la base de datos:

```
SELECT codloc , coddepto , codpcia , the_geom, nombreloc , aintra
FROM localidades ,

(SELECT codloc1 ,coddepto1 , codpcia1 , (animalesintra::float8/evaluadas) as aintra

FROM (SELECT viv1.codloc as codloc1 , viv1.coddepto as coddepto1 , viv1.codpcia
as codpcia1 , (COALESCE(SUM(carac.numperro::float8),0)+COALESCE(SUM(carac.numgato),0))
as animalesintra
FROM caracterizacionviv as carac , viviendas as viv1
WHERE viv1.idviv=carac.idviv and (carac.numperro > 0 OR carac.numgato > 0 )
GROUP BY codloc ,coddepto , codpcia) t1 JOIN

(SELECT viv2.codloc as codloc2 ,viv2.coddepto as coddepto2 , viv2.codpcia as codpcia2 ,
COUNT(eval2.estadoviv) as evaluadas
FROM evalentomologica as eval2 , viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%'
GROUP BY codloc ,coddepto , codpcia) t2 ON t1.codloc1=t2.codloc2
and t1.coddepto1=t2.coddepto2 and t1.codpcia1=t2.codpcia2) t3

WHERE t3.codloc1=codloc and t3.coddepto1=coddepto and t3.codpcia1=codpcia ;
```

La búsqueda SQL selecciona de toda la base de datos, los atributos "the_geom", "nombreloc", "Codpcia" y "Coddepto", con los datos de ubicación e identificación de cada localidad y genera una columna con el atributo "Aintra" que corresponde al promedio de animales calculado como se mencionó con anterioridad.

La figura 5.10 muestra la tabla que resulta de la consulta anteriormente descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

Query - chagadb on gi_user@192.168.128.128:5432

SQL Editor

```
SELECT codloc, coddepto, codpcia, the_geom, nombreloc, aintra
FROM localidades,
(SELECT codloc, coddepto1, codpcia1, (animalesintra::float8/evaluadas) as aintra

```

Output pane

	codloc character varying(8)	coddepto character varying(3)	codpcia character varying(2)	the_geom geometry	nombreloc character varying(50)	aintra double precision
1	1	007	46	0101000020E61000000	AIMOGASTA	2.2
2	1	021	46		AGUA BLANCA	1.875
3	1	028	30	0101000020E61000000	CHAJARI	2.17857142857143
4	1	035	30	0101000020E61000000	ARRIO DEL MEDIO	3.46153846153846
5	1	042	30	0101000020E61000000	ATENCIO	3.7
6	1	077	30	0101000020E61000000	20 DE SEPTIEMBRE	3.82142857142857
7	1	084	42	0101000020E61000000	CUCHILLO CO	2.35714285714286
8	1	091	30	0101000020E61000000	ALTAMIRANO NORTE	4.33333333333333
9	1	119	46	0101000020E61000000	AMUSCHITA	1
10	10	007	46	0101000020E61000000	VILLA MAZAN	0.495
11	10	021	46		SAN PEDRO	1.74285714285714
12	10	035	58	0101000020E61000000	NEUQUEN	5.74468085106383
13	10	049	30	0101000020E61000000	GUAYLEGUAY	3.21739130434783
14	10	091	30	0101000020E61000000	LAS GUACHAS	2.75
15	10	119	46	0101000020E61000000	SCHAQUI	1.24418804651163
16	11	021	46		SANTA CRUZ	3.16666666666667
17	11	042	30		BASUALDO	4.65
18	11	042	58		EL ALAMITO	6.5
19	11	091	30	0101000020E61000000	MACIA	3.8125
20	11	119	46	0101000020E61000000	SURIVACO	1.46153846153846
21	12	035	58	0101000020E61000000	PIOTTER	5.15384615384615
22	12	077	14	0101000020E61000000	EL CADILLO	4.11111111111111

OK. Unix Ln 17, Col 18, Ch 754 158 rows. 796 ms

Figura 5.10: Salida del PgAdmin resultado de la consulta SQL con el promedio por localidad de animales intradomiciliarios

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp` de la siguiente forma para generar la capa "animales_intradomicilio.shp", recibiendo como entrada la consulta SQL descrita con anterioridad para generar la tabla de atributos:

```
pgsql2shp -f animales_intradomicilio.shp -h host -u user -P *** chagadb ani_int.sql
```

Con el shape "animales_intradomicilio.shp" creado, se procede con la incorporación al GeoServer y su posterior publicación.

Para publicar la capa solo se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa "animales_intradomicilio.shp".

Se usaron las opciones de edición de la capa por defecto, y se calcularon los límites geográficos a partir de la capa.

Se creó un estilo particular para esta capa con el objetivo que visualmente se pueda interpretar el promedio de los animales de la localidad, por lo que se creó un estilo que agrupe en 4 categorías las localidades, de esta forma: 1) Menos de 2 animales, 2) entre 2 y 4, 3) entre 4 y 6, 4) más de 6 perros o gatos, conforme aumenta el número de animales en el intradomicilio, el punto que representa la localidad aumenta el tamaño y cambia su color.

Este estilo se creó en el editor de archivos SLD del GeoServer como se explicó previamente. La figura 5.11 muestra la previsualización de la capa publicada con el estilo generado. Con lo anterior el producto "animales_intradomicilio.shp" finalmente quedó incorporado y publicado en la plataforma Web-GIS de Chagas.

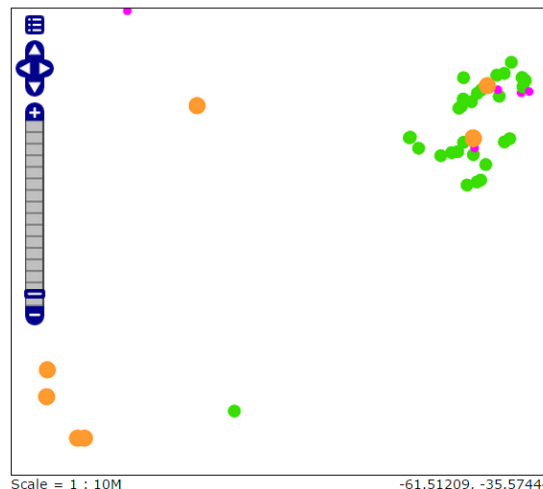


Figura 5.11: Ejemplo gráfico del producto 6 utilizando la herramienta *OpenLayers*

5.2.1.7 Producto 7: Animales peridomicilio

El séptimo producto vectorial generado fue la capa que representa el promedio por localidad de animales presentes en el peridomicilio; los animales que están incluidos en el cálculo del promedio son cabras y gallinas.

Para crear la capa fue necesario realizar la siguiente consulta SQL, que dentro de su código recibe función "sum" que es la que se encarga de sumar el número de cabras y gallinas en el peridomicilio de las viviendas (galpones, gallineros, corrales, etc.) que posteriormente se divide en el número de viviendas evaluadas de cada localidad presente en la base de datos:

```
SELECT codloc , coddepto , codpcia , the_geom , nombreloc , aperi
FROM localidades ,

(SELECT codloc1 ,coddepto1 , codpcia1 , (animalesperi::float8/evaluadas) as aperi
FROM (SELECT viv1.codloc as codloc1 , viv1.coddepto as coddepto1 , viv1.codpcia
as codpcia1 , (COALESCE(SUM(carac.numgallinas::float8),0)
+COALESCE(SUM(carac.numcabras),0)) as animalesperi
FROM caracterizacionviv as carac , viviendas as viv1
WHERE viv1.idviv=carac.idviv and (carac.numgallinas > 0 OR carac.numcabras > 0 )
GROUP BY codloc ,coddepto , codpcia) t1 JOIN

(SELECT viv2.codloc as codloc2 ,viv2.coddepto as coddepto2 , viv2.codpcia as codpcia2 ,
COUNT(eval2.estadoviv) as evaluadas
FROM evalentomologica as eval2 , viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%'
GROUP BY codloc ,coddepto , codpcia) t2 ON t1.codloc1=t2.codloc2
and t1.coddepto1=t2.coddepto2 and t1.codpcia1=t2.codpcia2) t3

WHERE t3.codloc1=codloc and t3.coddepto1=coddepto and t3.codpcia1=codpcia ;
```

La búsqueda SQL selecciona de toda la base de datos, los atributos "the_geom", "nombreloc", "Codpcia" y "Coddepto", con los datos de ubicación e identificación de cada localidad y genera una columna con el atributo "Aperi" que corresponde al promedio de animales calculado como se mencionó con anterioridad.

La figura 5.12 muestra la tabla que resulta de la consulta anteriormente descrita, que fue

realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

	codloc character varying(8)	coddepto character varying(3)	codpcia character varying(2)	the_geom geometry	nombreloc character varying(50)	apert double precision
48	2	014	58	0101000020E61000000	AGUADA DE SAN ROQUE	106.538461538462
49	2	021	46	0101000020E61000000	AMINBA	4.125
50	2	028	30	0101000020E61000000	COLONIA ALENAMA	12.4285714285714
51	2	070	30	0101000020E61000000	BOVRIL	10.82
52	2	091	30	0101000020E61000000	ALTAMIRANO SUR	13.75
53	2	126	46	0101000020E61000000	EL HUACO DE ARRIBA	1.25
54	20	021	30		PUERTO LAS CUEVAS	6.70588235294118
55	20	028	30		COLONIA PEÑA	16.3714285714286
56	20	035	30	0101000020E61000000	NUEVA VIZCAYA	9.85652173913044
57	20	084	58	0101000020E61000000	TAQUIMILAN	122.833333333333
58	20	105	10		FIAMBALA	3.01307189542484
59	21	021	30	0101000020E61000000	MOLINO DOLL	15.2758620689655
60	21	070	30	0101000020E61000000	SAN GUSTAVO	13.45
61	21	084	10	0101000020E61000000	SIJAN	6.31714285714286
62	22	008	30	0101000020E61000000	COLONIA HOCKER	5.66666666666667
63	22	028	46		SANTA CLARA	3.73529411764706
64	22	035	30	0101000020E61000000	SAUCE DE LUNA	8.28
65	22	049	30		DISTRITO SEXTO COS	16.0666666666667
66	22	070	30		SANTA ELENA	11.875
67	23	021	30	0101000020E61000000	STROBEL	3.33333333333333
68	26	008	30		SAN JOSE	9.86666666666667
69	26	028	30	0101000020E61000000	FEDERACION	4.92
70	26	035	30		BAUDERAS	16.5909090909091

Figura 5.12: Salida del PgAdmin resultado de la consulta SQL con el promedio por localidad de animales peridomiciliarios

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp` de la siguiente forma para generar la capa "animales_peridomicilio.shp", recibiendo como entrada la consulta SQL descrita con anterioridad para generar la tabla de atributos:

```
pgsql2shp -f animales_peridomicilio.shp -h host -u user -P *** chagadb ani_per.sql
```

Con el shape "animales_peridomicilio.shp" creado, se procede con la incorporación al GeoServer y su posterior publicación.

Para publicar la capa solo se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa "animales_peridomicilio.shp".

Se usaron las opciones de edición de la capa por defecto, y se calcularon los límites geográficos a partir de la capa.

Se creó un estilo particular para esta capa con el objetivo que visualmente se pueda interpretar el promedio de los animales de la localidad, por lo que se creó un estilo que agrupe en 4 categorías las localidades, de esta forma: 1) Menos de 10 animales, 2) entre 10 y 20) entre 20 y 30, 4) más de 30 cabras o gallinas, conforme aumenta el número de animales en el peridomicilio, el punto que representa la localidad aumenta el tamaño y cambia su color.

Este estilo se creó en el editor de archivos SLD del GeoServer como se explicó previamente. La figura 5.13 muestra la previsualización de la capa publicada con el estilo generado. Con lo anterior el producto "animales_peridomicilio.shp" finalmente quedó incorporado y publicado en la plataforma Web-GIS de Chagas.

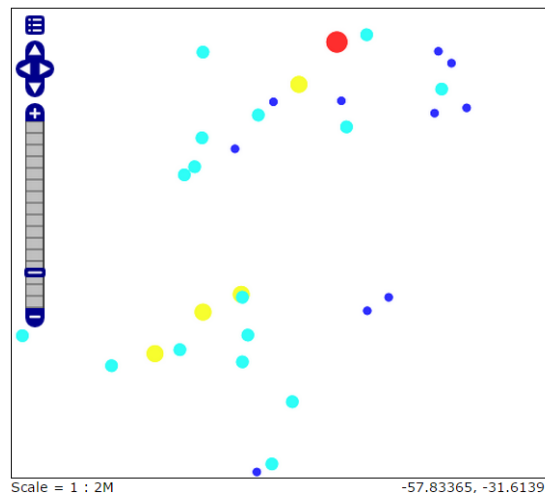


Figura 5.13: Ejemplo gráfico del producto 7 utilizando la herramienta *OpenLayers*

5.2.1.8 Producto 8: Índice de infestación intradomiciliaria

El octavo producto vectorial que se generó representa el índice de infestación intradomiciliaria (IIID) por localidad, que se calcula mediante la ecuación 5.1. Este producto es de suma importancia debido a que estandariza la información epidemiológica de las localidades, permitiendo análisis comparativos entre localidades.

$$IIID = \frac{\text{Número de viviendas infestadas ID}}{\text{Número de viviendas evaluadas}} \times 100 \quad (5.1)$$

La siguiente consulta SQL fue generada para totalizar el número de viviendas evaluadas de cada localidad, mediante la función "count", después sumar el número de viviendas con infestación y luego generar el cálculo de la ecuación 5.1:

```
SELECT the_geom, codloc, coddepto, codpcia, nombreloc, iiid, fecha
FROM localidades,

(SELECT codloc1, coddepto1, codpcia1, (ipositivas::float8/evaluadas)*100 as iiid, fecha
FROM (SELECT viv1.codloc as codloc1, viv1.coddepto as coddepto1, viv1.codpcia
as codpcia1, COUNT(viv1.idviv) as ipositivas, eval1.fecharecol as fecha
FROM evalentomologica as eval1, viviendas as viv1

WHERE viv1.idviv=eval1.idviv and (eval1.numidad > 0 OR eval1.numidn > 0
OR eval1.presidad='Si' OR eval1.presidn='Si') and (eval1.fecharecol >= '2011-01-01'
and eval1.fecharecol <='2015-12-31')

GROUP BY codloc, coddepto, codpcia, eval1.fecharecol) t1 JOIN

(SELECT viv2.codloc as codloc2, viv2.coddepto as coddepto2, viv2.codpcia as codpcia2,
COUNT(eval2.estadoviv) as evaluadas
FROM evalentomologica as eval2, viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%' and eval2.fecharecol >= '2011-01-01'
and eval2.fecharecol <='2015-12-31'
```

```

GROUP BY codloc , coddepto , codpcia ) t2

ON t1.codloc1=t2.codloc2 and t1.coddepto1=t2.coddepto2 and
t1.codpcia1=t2.codpcia2) t3

WHERE t3.codloc1=codloc and t3.coddepto1=coddepto and t3.codpcia1=codpcia

ORDER BY fecha , codpcia , coddepto , codloc ;

```

La búsqueda SQL selecciona de toda la base de datos, los atributos "the_geom", "nombreloc", "Codpcia" y "Coddepto", con los datos de ubicación e identificación de cada localidad, también genera una columna con el atributo "iid" que corresponde al índice calculado; adicionalmente la consulta genera el atributo "fecha" debido a que este producto fue pensado para que genere un índice IIID para cada año en las localidades. Es importante resaltar que en la actualidad existen registros de la planilla Chagas 6 incorporados al sistema únicamente de los años 2013, 2014 y 2015.

La figura 5.14 muestra la tabla que resulta de la consulta anteriormente descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

Query - chagasdb on gis_user@192.168.128.128:5432*

File Edit Query Favourites Macros View Help

</

Figura 5.14: Salida del PgAdmin de la consulta SQL con el IIID por localidad para cada año

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp` de la siguiente forma para generar la capa "iiid.año.shp", recibiendo como entrada la consulta SQL descrita con anterioridad para generar la tabla de atributos:

```
pgsql2shp -f iiid_año.shp -h host -u user -P *** chagasdb iiid_año.sql
```

Con el shape "iiid.año.shp" creado, se procede con la incorporación al GeoServer y su posterior publicación.

Para publicar la capa solo se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa

”iiid_año.shp”.

Se usaron las opciones de edición de la capa por defecto, y se calcularon los límites geográficos a partir de la capa.

Se creó un estilo particular para esta capa con el objetivo de interpretar visualmente la diferencia de los años en el IIID, por ello los puntos que corresponden a las localidades se visualizan de diferente color según el año que represente el valor del índice. Este estilo se creó en el editor de archivos SLD del GeoServer como se explicó previamente. Con lo anterior el producto ”iiid_año.shp” finalmente quedó incorporado y publicado en la plataforma Web-GIS de Chagas.

5.2.1.9 Producto 9: Índice de infestación peridomiciliaria

El último producto vectorial generado a partir de los datos de la planilla representa el índice de infestación peridomiciliaria (IIPD) por localidad, que se calcula mediante la ecuación 5.2.

$$IIPD = \frac{\text{Número de viviendas infestadas PD}}{\text{Número de viviendas evaluadas}} \times 100 \quad (5.2)$$

La siguiente consulta SQL fue generada para totalizar el número de viviendas evaluadas de cada localidad, mediante la función ”count”, después sumar el número de viviendas con infestación en el peridomicilio y luego generar el cálculo de la ecuación 5.2:

```
SELECT the_geom, codloc, coddepto, codpcia, nombreloc, iiud, fecha
FROM localidades, (SELECT codloc1, coddepto1, codpcia1,
(ipositivas::float8/evaluadas)*100 as iiud, fecha

FROM (SELECT viv1.codloc as codloc1, viv1.coddepto as coddepto1, viv1.codpcia
as codpcia1, COUNT(viv1.idviv) as ipositivas, eval1.fecharecol as fecha

FROM evalentomologica as eval1, viviendas as viv1
WHERE viv1.idviv=eval1.idviv AND (eval1.numpda > 0 OR eval1.numpdn > 0
OR eval1.prespda='Si' OR eval1.prespdn='Si' ) and eval1.fecharecol >= '2012-01-01'
and eval1.fecharecol <='2014-12-31'
GROUP BY codloc, coddepto, codpcia, eval1.fecharecol) t1 JOIN

(SELECT viv2.codloc as codloc2, viv2.coddepto as coddepto2, viv2.codpcia as codpcia2,
COUNT(eval2.estadoviv) as evaluadas
FROM evalentomologica as eval2, viviendas as viv2
WHERE viv2.idviv=eval2.idviv and eval2.estadoviv
LIKE '%Evaluada%' and eval2.fecharecol >= '2012-01-01'
and eval2.fecharecol <='2014-12-31'
GROUP BY codloc, coddepto, codpcia) t2

ON t1.codloc1=t2.codloc2 and t1.coddepto1=t2.coddepto2 and t1.codpcia1=t2.codpcia2) t3

WHERE t3.codloc1=codloc and t3.coddepto1=coddepto and t3.codpcia1=codpcia

ORDER BY fecha, codpcia, coddepto, codloc;
```

La búsqueda SQL selecciona de toda la base de datos, los atributos ”the_geom”, ”nombreloc”, ”Codpcia” y ”Coddepto”, con los datos de ubicación e identificación de cada localidad,

también genera una columna con el atributo "iuid" que corresponde al índice calculado; adicionalmente la consulta genera el atributo "fecha" debido a que de igual forma que el producto de IIID, este producto fue pensado para que genere un índice IIPD para cada año en las localidades.

La figura 5.15 muestra la tabla que resulta de la consulta anteriormente descrita, que fue realizada de forma manual en el PgAdmin en donde, como se observa, es posible visualizar si la salida de la consulta SQL es realmente la tabla de atributos de la capa que se quiere generar.

Query - chagasdb on gis_user@192.168.128.128:5432

```

as codpcia1, COUNT(vivi1.idvivi) as ipositivas, eval1.fechaarecol as fecha
FROM evalentomologica as eval1, viviendas as vivi1
WHERE vivi1.idvivi=eval1.idvivi AND (eval1.numpda > 0 OR eval1.numpdn > 0
OR eval1.prespda='SI' OR eval1.prespdn='SI') and eval1.fechaarecol >= '2012-01-01'
and eval1.fechaarecol <= '2012-12-31'

```

	the_geom	codloc	coddepto	codpcia	nombrelot	aud	fecha
	geometry	character varying(8)	character varying(3)	character varying(2)	character varying(50)	double precision	date
1	0101000020E61000008B7013	119	46	ALPASINCH	2.08333333333333	2013-04-23	
2	0101000020E6100000000004	119	46	CHAUPHUASI	4.10958904109589	2013-04-23	
3		6	119	46	EL RETIRO	100	2013-04-23
4	0101000020E61000008B7013	119	46	ALPASINCH	5.20833333333333	2013-04-24	
5	0101000020E6100000000004	119	46	CHAUPHUASI	2.73972602739726	2013-04-24	
6	0101000020E61000008B7013	119	46	ALPASINCH	5.20833333333333	2013-04-25	
7	0101000020E6100000000004	119	46	CHAUPHUASI	1.36986301369863	2013-04-25	
8	0101000020E6100000000005	119	46	CUIPAN	1.16279069767442	2013-04-25	
9	0101000020E6100000000005	119	46	CUIPAN	1.16279069767442	2013-04-26	
10	0101000020E6100000000005	119	46	SAN BLAS	1.16279069767442	2013-04-26	
11	0101000020E61000003CDE14	119	46	LOS ROBLES	0.909090909090909	2013-04-29	
12	0101000020E6100000000008	119	46	SALICAS	0.641025641025641	2013-04-29	
13	0101000020E6100000000009	119	46	SAN BLAS	2.32558139534884	2013-04-29	
14	0101000020E61000003CDE14	119	46	LOS ROBLES	0.909090909090909	2013-04-30	
15	0101000020E6100000000008	119	46	SALICAS	2.56410256410256	2013-04-30	
16	0101000020E61000003CDE14	119	46	LOS ROBLES	1.81818181818182	2013-05-02	
17	0101000020E6100000000008	119	46	SALICAS	0.641025641025641	2013-05-02	
18	0101000020E6100000000009	119	46	SAN BLAS	1.16279069767442	2013-05-02	
19	0101000020E61000003CDE14	119	46	LOS ROBLES	2.72727272727273	2013-05-03	
20	0101000020E6100000000005	119	46	CUIPAN	1.16279069767442	2013-05-03	
21	0101000020E6100000000009	119	46	SAN BLAS	1.16279069767442	2013-05-03	
22	0101000020E61000003CDE14	119	46	LOS ROBLES	0.909090909090909	2013-05-03	

Figura 5.15: Salida del PgAdmin resultado de la consulta SQL con el IIPD por localidad para cada año

Una vez lista la consulta, se procede a utilizar la función `pgsql2shp` de la siguiente forma para generar la capa "iipd_año.shp", recibiendo como entrada la consulta SQL descrita con anterioridad para generar la tabla de atributos:

```
pgsql2shp -f iipd_anio.shp -h host -u user -P *** chagasdb iipd_anio.sql
```

Con el shape "iipd_año.shp" creado, se procede con la incorporación al GeoServer y su posterior publicación.

Para publicar la capa solo se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa "iipd_año.shp".

Se usaron las opciones de edición de la capa por defecto, y se calcularon los límites geográficos a partir de la capa.

Para la visualización de la capa en el sistema Chagas, fue asignado el mismo estilo que en la capa de "iiid_año.shp", los puntos que corresponden a las localidades se visualizan de diferente color según el año que represente el valor del índice. Con lo anterior el producto "iipd_año.shp" finalmente quedó incorporado y publicado en la plataforma Web-GIS de Chagas.

Con la incorporación al sistema Chagas de esta última capa vectorial, se concluyó con la generación de los productos a partir de la base de datos PostGIS, la figura 5.16 representa el

ciclo de trabajo que resume el procedimiento que se explicó para las nueve capas, siendo estos los pasos a seguir para la generación de un shape y su incorporación al GeoServer.

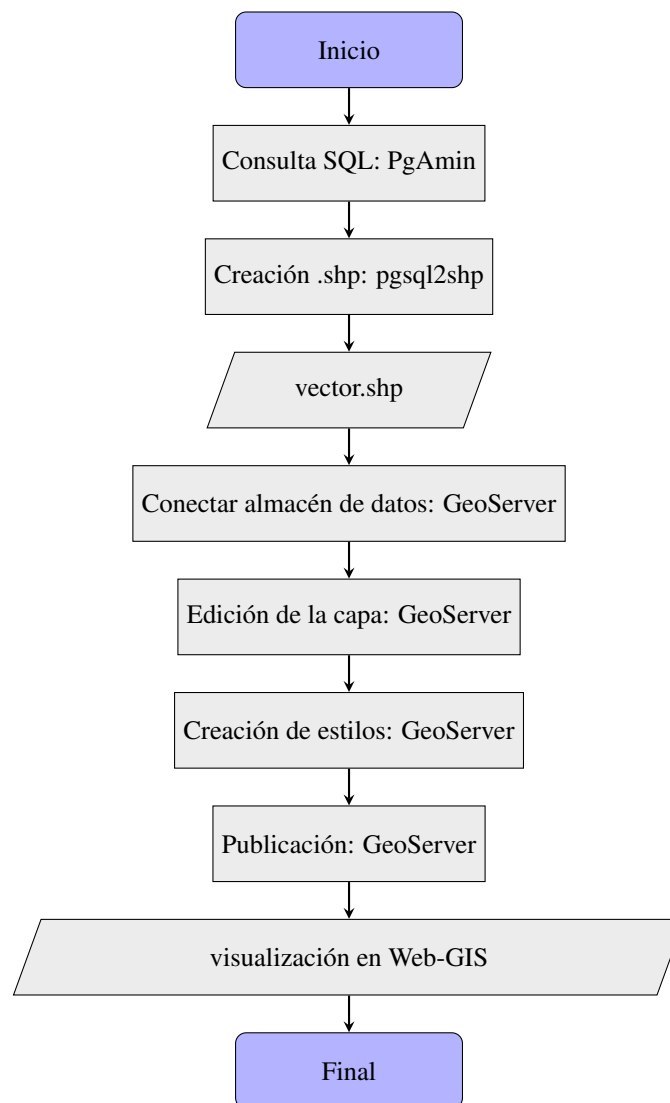


Figura 5.16: Flujo de trabajo para la generación e incorporación de productos vectoriales al sistema Web-GIS

5.3 Producto 10: Distancia localidad infestada-No infestada

Una de las principales utilidades de los productos que fueron generados a partir de la información de las bases de datos con los registros de la Planilla Chagas 6, es que constituyen en sí mismos información base de fácil interpretación y uso por parte de los usuarios, investigadores o efectores del control epidemiológico de Chagas, y que a partir de la información de las capas se puede generar otro tipo de datos que aportan al análisis de la enfermedad.

En congruencia con esto se buscó crear un producto a partir de las capas ya generadas e incorporadas en el sistema; [Vazquez-Prokopec et al. \(2012\)](#) realizaron un importante análisis de diferentes variables, en busca de asociar algunas de ellas con la enfermedad y obtener mapas de riesgo de transmisión de Chagas; una de las capas incluidas en los modelos que respondieron significativamente al análisis realizado por los autores, fue la distancia en metros que existe desde las localidades sin infestación hasta la localidad con viviendas infestadas más cercana. Lo cual es posible calcular de manera muy sencilla a partir de las capas "loc_sin_infestación" y "loc_con_infestación" explicadas con anterioridad (Productos vectoriales 4 y 5 respectivamente).

Para generar la capa vectorial que se denomina "distancia.shp", hace falta incorporar los productos "loc_sin_infestación" y "loc_con_infestación" a un GIS de escritorio. La capa se construyó en el GRASS GIS, utilizando el *LOCATION* y *MAPSET* creado para la generación de productos raster, con el área de interés descrita en el capítulo 4, por lo que únicamente las localidades registradas en la base de datos PostGIS con los registros de la planilla Chagas 6 que estén ubicadas geográficamente dentro del área de estudio son analizadas.

Con la finalidad de incorporar los shapes al GRASS GIS se utilizó la función `v.in.ogr` que recibe como parámetro de entrada el shape y su carpeta de ubicación, y un nombre "output" con el cual será reconocida la capa dentro del entorno GRASS GIS; con las siguientes líneas de comando se incorporaron las dos capas:

```
v.in.ogr input=/loc_sin_infestacion.shp output=localidades_sin_infestacion
v.in.ogr input=/loc_con_infestacion.shp output=localidades_con_infestacion
```

Una vez están los shapes en el GRASS, se identifica que la distancia a calcular parte desde las localidades sin infestación, hacia las localidades con infestación, por lo que se requiere agregar un atributo (columna) a la tabla asociada al shape de las localidades sin infestación.

La función `v.db.addcolumn` añade uno o más columnas a la tabla de atributos conectada a un vector dado; por lo que satisface lo que se quiere lograr con la capa "localidades_sin_infestación"; la función se utilizó de la siguiente forma, y requiere únicamente el nombre de la capa a modificar, el nombre de la columna y tipo de dato que será añadido a la tabla:

```
v.db.addcolumn map=localidades_sin_infestacion column="nrst double precision"
```

Con lo anterior se creó el espacio para que el cálculo de la distancia sea ubicado en la columna "nrst" recientemente agregada a la tabla de atributos.

La función de GRASS `v.distance` encuentra el elemento más cercano de una capa vectorial "TO", desde los elementos de otra capa vectorial "FROM". La distancia que calcula es en metros y es posible delimitar un radio máximo de búsqueda. La distancia se calculó con la siguiente línea de comando usando la función `v.distance` que recibe como parámetros de entrada las capas "TO" y "FROM", y el nombre de la columna a actualizar con el dato de distancia:

```
v.distance from=localidades_sin_infestacion to=localidades_con_infestacion \
upload=dist column=nrst
```

Una vez obtenida la capa con la distancia entre las localidades se procedió con su exportación del GRASS, generando el producto en formato Shape, para su posterior subida al servidor. La exportación se realizó de la siguiente forma:

```
v.out.ogr input=localidades_sin_infestacion output=/distancia.shp format=ESRI-Shapefile
```

Con el producto `distancia.shp` creado se procedió con la incorporación del mismo a la plataforma Chagas; esta vez la incorporación de la capa se realizó de forma manual mediante una conexión por el protocolo SSH en el *FileZilla*; para lo cual se requieren los parámetros de acceso a la base de datos como son, dirección ip donde se encuentra el sistema Chagas, usuario y contraseña de acceso y puerto de conexión. Una vez suministrados estos datos se procedió a transferir el archivo "distancia.shp" a la misma ubicación (*path*) donde se crearon los 9 productos vectoriales a partir de la base de datos PostGIS del sistema.

Para incorporar la capa al GeoServer se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a la capa "distancia.shp". No es necesario crear una nueva fuente de datos vectoriales ni un almacén de datos, debido a que el producto se transfirió a la misma ubicación de las otras capas de la planilla, y el GeoServer ya tiene generado el vínculo con esa fuente de datos.

Se usaron las opciones de edición de la capa por defecto, y se calcularon los límites geográficos a partir de la capa.

Se creó un estilo particular para esta capa con el objetivo que visualmente se pueda interpretar la distancia entre las mismas, la figura 5.17 muestra la ventana de edición de la capa en donde se le asignó el estilo que representa con colores diferentes las localidades sin infestación que tienen una localidad infestada a una distancia clasificada en cuatro categorías: 1) Menos de 50 km, 2) Entre 50 y 70 km, 3) Entre 70 y 100 km, 4) Mayor a 100 km.

Con lo anterior el producto "distancia.shp" que resultó de la combinación de dos capas vectoriales generadas a partir de la planilla Chagas, finalmente quedó incorporado y publicado en la plataforma Web-GIS.

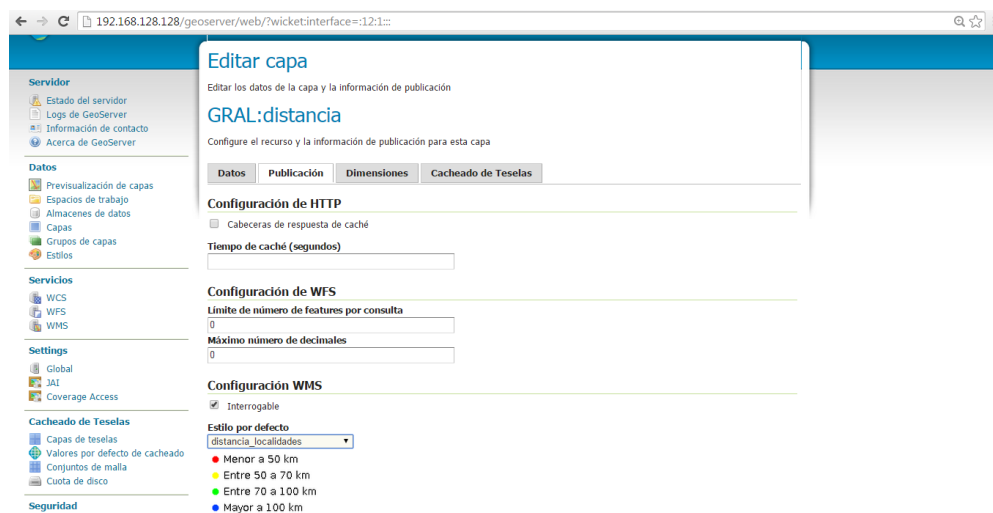


Figura 5.17: Ejemplo de la ventana de edición y asignación de estilo del GeoServer

5.4 Producto 11: Combinación de Vector y Raster

Como se mencionó previamente una de las principales utilidades de los productos generados es la posible combinación de los mismos para generar información con nuevo valor agregado para el análisis de la enfermedad y la biología de sus vectores. El último producto desarrollado como ejemplo de esta utilidad es una capa de degradación de suelo en la periferia de los centros poblados registrados en la base de datos.

En el análisis de diferentes variables realizado por [Vazquez-Prokopec et al. \(2012\)](#) encontraron significativa la relación entre las zonas donde se presentan *clusters* (agrupaciones) de comunidades rurales con alta infestación y el modelo que incluye a la variable que denominan como "Porcentaje alto de tierras degradadas o deforestadas", siendo la variable de alto aporte para el modelo que obtuvieron. Esta variable es representada como el porcentaje de tierra que presenta actividad agrícola o ganadera en un buffer de 2 km en la periferia de las comunidades. Por la importancia estadística de esta variable para estimar áreas con alta infestación de *T. infestans* se generó esta capa a partir de la combinación de la serie de tiempo de NDVI (Ver capítulo 4) y el producto vectorial con las localidades evaluadas explicado con anterioridad.

El producto fue creado en GRASS-GIS utilizando el *LOCATION* y *MAPSET* creado para la generación de productos de la serie temporal de índices verdes NDVI y EVI, descrito en el capítulo 4. De igual forma que en el producto 10, se utilizó una de las capas vectoriales generada a partir de la base de datos del servidor de Chagas, la capa "localidades_evaluadas.shp" fue incorporada con el comando `v.in.ogr` de la siguiente forma, teniendo en cuenta que la región de trabajo del *MAPSET* es únicamente el área de estudio descrita en el capítulo 4, por lo que únicamente las localidades incluidas en el área son importadas al GRASS:

```
v.in.ogr --overwrite input=localidades_evaluadas.shp output=localidades_evaluadas
```

Con la capa vectorial de las localidades incorporada al GRASS, se procede con la generación del producto raster que representa la degradación del suelo.

Eventos como incendios o inundaciones tienen efectos en el NDVI, reduciendo su valor considerablemente, lo que podría ser interpretado como una pérdida de la vegetación en determinada área de una fecha a otra, sin embargo, si estos eventos son accidentales o resultado de alguna emergencia ambiental, el valor del índice verde subirá progresivamente en fechas posteriores al evento hasta aproximarse a su valor original. Por el contrario, la pérdida de cobertura de vegetación boscosa, con la finalidad de utilizar las tierras para el agro o ganadería, representan una baja en el NDVI que se mantiene a lo largo del tiempo. O en caso contrario, cuando coberturas originales que contienen sabanas o suelos sin cobertura boscosa se transforman en cultivos, el NDVI tiene un aumento propio que varía según la fenología del cultivo.

Estos cambios en el NDVI se hacen permanentes en el tiempo, y constituye una degradación por prácticas agrícolas del suelo que se expresa en cambios (aditivos o sustractivos) del valor del índice. En congruencia el producto de degradación de suelo fue planteado como dos capas que se denominan "degradación_positiva", si el cambio en el NDVI fue aditivo, y "degradación_negativa" si el cambio fue sustractivo.

El desafío en la creación de este producto es identificar un umbral en los valores de NDVI a partir del cual se considere que hay cambio en la vegetación presente en cada píxel. A continuación se describe la metodología y las líneas de comando de GRASS implementadas en BASH para establecer dichos umbrales, y para revisar que los píxeles en donde se identificó cambio, permanezcan en el tiempo con esos valores, de esta forma evitar la confusión con eventos aleatorios (inundaciones, incendios, etc).

El umbral de corte a partir del cual se considera que el valor del índice cambió, se implementó con la ecuación 5.3 que calcula valores atípicos a partir de la media y el desvío estándar de la serie de tiempo. Con esto, los píxeles que superen este valor (límite positivo) serán identificados como píxeles con degradación positiva (aditiva), es decir, que sus valores normales de NDVI, aumentaron considerablemente en una fecha. Por su parte el límite inferior se establece con la ecuación 5.4 y los píxeles con valor de índice de NDVI inferior a este umbral, son identificados como degradación negativa (cambio sustractivo).

$$\text{Límite_superior} = \text{Media} + (2 \times \text{Desvío Estandar}) \quad (5.3)$$

$$\text{Límite_inferior} = \text{Media} - (2 \times \text{Desvío Estandar}) \quad (5.4)$$

Para aplicar estas formulas a las imágenes de la serie de tiempo se calcularon las capas de Desvío estándar y Media de toda la serie histórica de NDVI, mediante el uso de la función de GRASS `r.series`, que permite aplicar diferentes cálculos que están definidos internamente en la función a una lista de imágenes. A continuación se muestra el uso de la función `r.series` para obtener la media y el desvío:

```
r.series -o input='g.list rast pattern=MOD13*NDVI sep=', '\
```

```
output=ndvi_mean_TOTAL, ndvi_sd_TOTAL \
method=average , stddev
```

Cada píxel incluido en el área de estudio contiene un valor diferente del promedio y su desvío, el cual varía en el espacio debido a la cobertura e historia de cada píxel, por lo que el límite máximo y mínimo no es un valor fijo, sino una imagen con dichos valores; las formulas 5.3 y 5.4 se aplicaron para obtener los límites mediante la función `r.mapcalc`, tomando como entrada las capas de la media y el desvío histórico calculado con anterioridad, de la siguiente forma:

```
r.mapcalc --o expression="lim_max=ndvi_mean_TOTAL+(2*ndvi_sd_TOTAL)"
r.mapcalc --o expression="lim_min=ndvi_mean_TOTAL-(2*ndvi_sd_TOTAL)"
```

Cada una de las imágenes que componen la serie de tiempo es comparada con las capas que contienen los límites superior e inferior mediante la siguiente implementación de la función `r.mapcalc`:

```
for map in `g.list rast pattern=MOD13*NDVI`
do
r.mapcalc --o expression="outlier_max_$map=$map*0"
r.mapcalc --o expression="outlier_max_$map=if($map > lim_max, outlier_max_$map + 1 ,\
outlier_max_$map)"

r.mapcalc --o expression="outlier_min_$map=$map*0"
r.mapcalc --o expression="outlier_min_$map=if($map < lim_min, outlier_min_$map + 1 ,\
outlier_min_$map)"
done
```

La implementación agrupa cuatro líneas con la función `r.mapcalc` en un ciclo *FOR*, que se repite hasta que todas las imágenes hayan sido comparadas con los límites superior e inferior. Con cada iteración del ciclo se crean dos imágenes que contienen valores cero en donde la imagen de NDVI no fue superior o inferior a los límites establecidos, o valores uno en donde superó los umbrales. La salida de esta implementación son dos series de capas, una imagen para cada fecha contenida en la serie de tiempo, para los dos límites (positivo y negativo)

Una vez identificadas las fechas en donde hubo un cambio considerable en los valores del índice verde, se procedió con la implementación para verificar que ese cambio perdure en el tiempo, con la finalidad de distinguir entre un evento aleatorio y la degradación de suelo con fines agrícolas.

Para esto es necesario registrar las dos series de imágenes binarias recientemente construidas para que sean consideradas como una lista y permita una manipulación más simple.

El comando de *Bash* que se muestra a continuación permitió crear una lista con las imágenes comparadas con el límite máximo, y otra para el límite mínimo:

```
declare -a lista_max=('g.list rast pat=outlier_max*')
declare -a lista_min=('g.list rast pat=outlier_min*')
```

Con la siguiente implementación se automatizó el procedimiento para generar la capa de degradación positiva:

```
len_max=${#lista_max[@]} # get list length
r.mapcalc --o expression="degrad_positiva=${lista_max[0]}*0"
len1=${len_max}-3
for ((i=0;i<${len1};i++))

do
r.mapcalc --o expression="degrad_positiva=
if (($ {lista_max[$i]}+${lista_max[$i+1]}+${lista_max[$i+2]}+${lista_max[$i+3]}) >= 4,
degrad_positiva+1,
degrad_positiva)"
done
```

La variable "len_max" contiene el número de imágenes de la lista "lista_max" (longitud), es necesario identificar este dato debido a que el ciclo *FOR* implementado itera tantas veces como la longitud de la lista.

Se crea una imagen "degrad_positiva" que contiene únicamente valores cero, con la finalidad de usarla como espacio para que la salida de cada iteración del ciclo sume una unidad a los píxeles que cumplen con la condición especificada, o permanezca en cero si no la cumplen. Dentro de la función *r.mapcalc* de GRASS, se implementó la condición *IF* que revisa si la fecha que se esta revisando y tres fechas hacia adelante en la lista de la serie de tiempo de índices verdes, tienen valores de uno, es decir, en el paso anterior cuatro fechas seguidas superaron el límite superior, entonces en el píxel de la imagen "degrad_positiva" se suma una unidad. Lo anterior fue la lógica que se construyó para identificar cambios del NDVI que perduren en el tiempo, recordando que cuatro imágenes de índice verde corresponden a 2 meses (Un producto cada 16 días, Ver capítulo 4).

Cuando el ciclo termina las iteraciones, se obtiene el producto degradación positiva; esta misma implementación se utilizó para generar la imagen degradación negativa, utilizando como entrada la lista de imágenes con los límites inferiores.

Con las capas de degradación en formato raster creadas, se procede con la combinación de la capa vectorial de las localidades evaluadas que fue importada al GRASS con anterioridad.

La función *v.buffer* crea un área circular tomando como centro la localidad, se usó la función de la siguiente manera para crear los círculos (polígonos) con un radio de 10 km.

```
v.buffer -t --overwrite input=localidades_evaluadas \
out=buffer_deg_positiva type=point distance=0.1

v.buffer -t --overwrite input=localidades_evaluadas \
out=buffer_deg_negativa type=point distance=0.1
```

Se generaron dos capas con los polígonos, una que se cruza con la capa de degradación positiva (*buffer_deg_positiva*) y otra con la capa de degradación negativa (*buffer_deg_negativa*).

Con la función *v.rast.stats* se cruza la información de la capa raster con la capa vectorial, la función se usó de la siguiente manera para calcular el promedio de los píxeles degradados que

están presentes en el buffer de 10 km de radio a partir de la localidad:

```
v.rast.stats -c --overwrite map=buffer_deg_positiva raster=degrad_positiva \
column_prefix=b method=average

v.rast.stats -c --overwrite map=buffer_deg_negativa raster=degrad_negativa \
column_prefix=b method=average
```

Para expresar los valores en porcentaje, con la función `v.db.update`, se actualiza la capa vectorial en la columna que contiene el promedio calculado con la función `v.rast.stats` multiplicando por 100, de la siguiente forma:

```
v.db.update buffer_deg_positiva col=b_average qcol="b_average*100"
v.db.update buffer_deg_negativa col=b_average qcol="b_average*100"
```

Finalmente se obtuvieron los productos vectoriales de degradación positiva y negativa. Se procedió con la exportación de la capa, generando el producto en formato `.shp`, para su posterior subida al servidor. La exportación del GRASS GIS se realizó de la siguiente forma:

```
v.out.ogr input=buffer_deg_positiva output=buffer_deg_positiva format=ESRI_Shapefile
v.out.ogr input=buffer_deg_negativa output=buffer_deg_negativa format=ESRI_Shapefile
```

Se procedió con la incorporación de los productos a la plataforma Chagas; esta vez la incorporación de la capa se realizó de forma manual mediante una conexión por el protocolo SSH en el *FileZilla* de igual forma que para el producto vectorial 10 con la distancia de las localidades.

Para incorporar la capa al GeoServer se requiere ir a la opción "capas", seleccionar el origen de datos "vectores_planilla" y posteriormente seleccionar la opción "publicar" frente a las capas "buffer_deg_positiva.shp" y "buffer_deg_negativa.shp". No es necesario crear una nueva fuente de datos vectoriales ni un almacén de datos, debido a que el producto se transfirió a la misma ubicación de las otras capas de la planilla, y el GeoServer ya tiene generado el vínculo con esa fuente de datos.

Se usaron las opciones de edición de la capa por defecto, y se calcularon los límites geográficos a partir de la capa.

Se creó un estilo particular para esta capa con el objetivo que visualmente se pueda interpretar la distancia entre las mismas, la figura 5.18 muestra la ventana de edición de la capa en donde se le asignó el estilo que representa con colores diferentes el porcentaje de degradación presente en el polígono periférico a la localidad. Los porcentajes se clasificaron en cuatro categorías: 1) Menos de 25%, 2) Entre 25% y 50%, 3) Entre 50% y 75% y 4) Mayor a 75%

Con lo anterior los productos de porcentaje de degradación que resultaron de la combinación de capas raster y vectoriales generadas a partir de la planilla Chagas, finalmente quedaron incorporados y publicados en la plataforma Web-GIS.

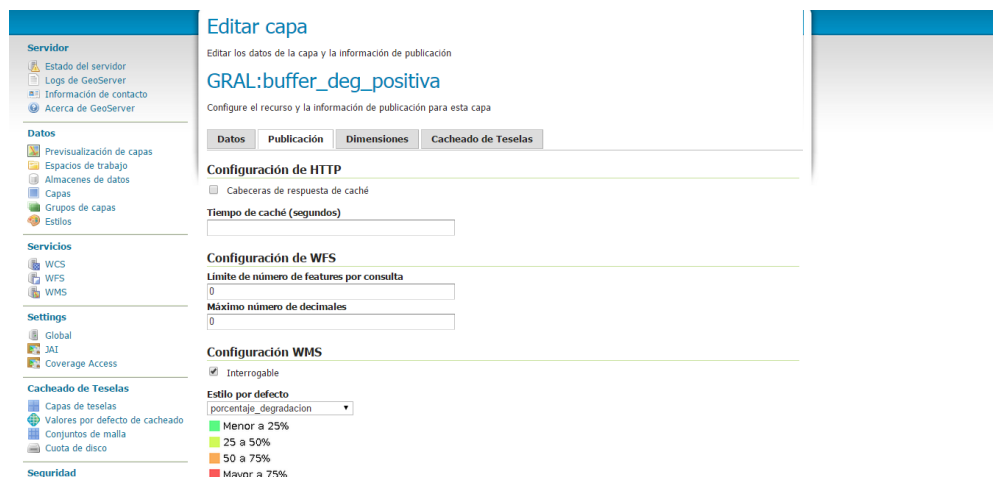


Figura 5.18: Ejemplo de la ventana de edición y asignación de estilo del GeoServer de la capa de Degradación

Capítulo 6

Consideraciones finales y líneas de abordaje a futuro

6.1 Discusión

Cada especie tiene requerimientos únicos que determinan sus preferencias biológicas o de hábitat, debido a esto, identificar los límites de tolerancia y las condiciones óptimas para el desarrollo de los vectores es de suma importancia para estimar los lugares con mayor probabilidad de encuentro entre los vectores, los patógenos y el hombre.

Variables ambientales como la temperatura, la precipitación, humedad del ambiente y otras, son factores determinantes y fundamentales para la distribución y abundancia de una especie y son una parte esencial en cualquier intento de evaluar los factores de riesgo ambientales de presencia de patógenos y sus vectores.

Cuando estas variables y sus límites se desconocen, es posible utilizar la información de registros en donde se tiene seguridad de la presencia de los vectores, y dentro de un SIG, superponer y analizar diferentes variables que se registran en las zonas con presencia; con esto se definen los requisitos de las especies y se pueden modelar zonas en donde no se conoce en terreno la presencia de individuos. Por ejemplo, modelos de nicho ecológico estiman la distribución de una especie basándose en la probabilidad de tener en zonas desconocidas, los mismo valores de las variables que tienen los lugares donde sí se conoce la presencia de individuos.

Lo anterior es posible gracias a las herramientas de análisis geo-espaciales disponibles en los SIG. La materia prima para este tipo de análisis que estiman el riesgo de presencia de enfermedades o sus vectores, son además de los registros de presencia en campo, capas que espacializan las variables ambientales, de esta forma se puede extrapolar el conocimiento local y llevarlo a zonas donde el esfuerzo en campo no se logra, y realizar estimaciones que sirvan como base para la toma de decisiones en la gestión del control de la enfermedad de interés.

La interacción entre la ubicación, el tiempo y la epidemiología está bien establecida en la comunidad médica, donde las enfermedades se pueden caracterizar por las exigencias

ecológicas de los organismos (insectos vectores) involucrados, los factores ambientales o estacionalidad; conociendo esto, se puede aprovechar esta información para encontrar soluciones de manejo. El papel de la geografía, como una disciplina que trasciende tanto la ubicación y el tiempo y que crea herramientas útiles, tales como la teledetección y SIG, es fundamental para la comunidad médica para hacer frente a estos problemas de salud pública. Las herramientas que se necesitan para combatir el objetivo siempre en movimiento de la salud humana, seguirán siendo aquellas que entran en el ámbito de las ciencias geográficas, que permitan analizar patrones a todas las escalas de estudio y manejo; y que además permitan la interdisciplinariedad de los investigadores y la interoperabilidad de las instituciones efectoras. Como respuesta a esto, considerables esfuerzos están en marcha para desarrollar funcionalidades SIG en la web. Al igual que la plataforma WebGIS de Chagas descrita en este trabajo, a lo largo del mundo existen esfuerzos de desarrollo de estas tecnologías que han ido mutando según los avances en aplicaciones como MapServer u OpenGeouite, pero siempre obteniendo resultados satisfactorios para el manejo de las enfermedades.

Como ejemplo, [Yang et al. \(2005\)](#) proponen el desarrollo de un WebGIS para el control de la esquistosomiasis en China, esfuerzo que fue exitoso y novedoso para la época, desarrollado bajo la tecnología MapSever. Por su parte, [Furlanello et al. \(2002\)](#) presentan el ULSS Belluno, un WebGIS para el mapeo de enfermedades transmitidas por garrapatas; en donde, al igual que el WebGIS de Chagas Argentina, integran información ambiental con datos epidemiológicos de picaduras de los vectores, y obtienen modelos de riesgo con una precisión de hasta el 82%. El desarrollo de [Furlanello et al. \(2002\)](#) fue un gran avance que permitió la obtención de modelos predictivos de riesgo, como base para tomar medidas de salud pública eficaces y sistémicas.

[Laosuwan \(2012\)](#) desarrollaron un WebGIS para el control de la leptospirosis, con tecnologías de código abierto. En esta investigación los autores desarrollaron una arquitectura orientada a servicios, lo que asegura, con el uso de tecnología más actual, la interoperabilidad de su plataforma.

Los ejemplos anteriormente mencionados, reflejan algunas experiencias históricas en diferentes lugares, en los cuales siempre fueron obtenidos resultados positivos para la vigilancia de cada una de las enfermedades de interés.

Dentro del plan nacional para el manejo y control de Chagas en Argentina existe la activa intención de estandarizar la información que todos los efectores del control epidemiológico de la enfermedad realizan en campo, de esta forma nace la plataforma Web-GIS descrita en este trabajo, como herramienta para compartir información que sirva como base para realizar análisis que mejoren el entendimiento de la enfermedad y en consecuencia las medidas de control y erradicación de la misma.

El aporte de este trabajo consistió en generar un gran número de variables ambientales debidamente especializadas en un territorio extenso, que tienen significado biológico para los vectores de Chagas, y que pueden ser incluidos en modelos que mejoren las estimaciones del riesgo de la enfermedad.

Adicional al potencial de los productos generados, también se explicó con suficiente nivel de detalle los métodos utilizados para su generación, por lo cual se pueden aprovechar para obtener las mismas capas en otras zonas de interés o actualizaciones a futuro.

Integrar las capas al sistema Web-GIS de Chagas aporta a la intención del Ministerio, debido a que los productos están disponibles para ser utilizadas por parte de los efectores o diferentes usuarios del sistema; brindando materia prima muy relevante para análisis posteriores, y que facilitan la generación de estudios, principalmente para usuarios que no tienen amplios conocimientos en SIG, procesamiento de imágenes, o manejo de series históricas con un gran

número de datos, como las que fueron resumidas con los productos de índices verdes y temperatura de superficie.

Los efectores, investigadores y usuarios de todo tipo, encuentran en la plataforma Web-GIS todos los productos debidamente preparados para su uso, en formatos comunes (GeoTiff y shape), estandarizados y de libre acceso, lo que evita la necesidad de herramientas sofisticadas para extraer información a partir de ellos.

El mundo de la información geográfica vive un momento de transformación, en donde el uso de herramientas libres (*OpenSource*) y tecnologías en la web están marcando tendencia por las grandes posibilidades de generar sistemas muy sofisticados con herramientas de fácil acceso. Todo el sistema de Chagas con el que se interactuó en este trabajo esta construido con este tipo de herramientas, y en sí mismo constituye una excelente iniciativa, novedosa y robusta que aporta en gran medida a mejorar la vigilancia epidemiológica de la enfermedad en Argentina.

La experiencia adquirida con la construcción de este tipo de plataformas dedicadas a problemas de salud pública ya ha tenido otros escenarios satisfactorios en Argentina (Dengue), y países como Colombia encuentran en esta experiencia un ejemplo base para proponer y construir sus propios sistemas.

Esto genera un gran impacto social por el hecho de que en Colombia no existen sistemas de esta envergadura, y la creación de una plataforma para la vigilancia epidemiológica del mal de Chagas es un proyecto pionero en dicho país, llevando las consecuencias positivas que ya fueron experimentadas en Argentina.

Uno de los propósitos que se lograron con el desarrollo de este trabajo fue entender como se estructura un Web-GIS, en todas sus capas de implementación, y como interactúan todas las partes de la plataforma para que el usuario pueda visualizar la información que reposa en el sistema. Toda la estructura del Web-GIS estudiada, los scripts desarrollados, los lenguajes aprendidos, el flujo de trabajo desde descargar datos, manejar series de tiempo y datos de gran tamaño, interactuar con diferentes formatos de información satelital, construir capas, integrarlas al sistema y publicar los productos fue debidamente documentado en este trabajo y representa un gran logro para la formación profesional; principalmente porque aporta en gran medida a la formación de un recurso humano interdisciplinario.

Con mi formación como bióloga fue un reto muy importante el abordar un trabajo que responde a habilidades computacionales, y me llevó a entender la importancia y el mundo de posibilidades que representa el conocer este tipo de tecnologías y herramientas, que aplicadas al área biológica generan aportes de mucho valor, y en lo personal me ha brindado en muy corto tiempo la posibilidad de aplicar lo aprendido en otros proyectos.

El conocimiento adquirido con el desarrollo de este trabajo es aplicable en cualquier proyecto que requiera de estandarizar, automatizar y compartir información geográfica en la Web, lo que me permitió participar en el programa regional de empleo de información satelital para la productividad agrícola (PREISPA).

El proyecto PREISPA fue financiado por el Banco Interamericano de Desarrollo (BID) y

desarrollado de forma colectiva entre instituciones de los países del cono sur, como lo son CONAE, la Agencia Chilena del Espacio (ACE), el Centro de Información de Recursos Naturales de Chile (CIREN), el Consejo Nacional de Ciencia y Tecnología de Paraguay (CONACYT), el Instituto Nacional de Investigación Agropecuaria (INIA) y el Ministerio de Ganadería, Agricultura y Pesca de Uruguay (MGAP).

Mi labor como consultora del proyecto fue utilizar todas las herramientas obtenidas en esta tesis para generar un gran número de productos ambientales que fueron incorporados al Web-GIS (Figura 6.1), denominado ISAGRO (Información Satelital para el Agro).

El aporte fue la construcción de capas con el pronóstico de heladas actual e histórico, NDVI actual, histórico y sus anomalías, índices de sequía y humedad, precipitaciones históricas y actuales, y temperatura en tiempo real. Teniendo en cuenta la gran envergadura del proyecto PREISPA, mi participación en este proyecto es un logro de suma importancia, que derivó de la realización de este trabajo de tesis.

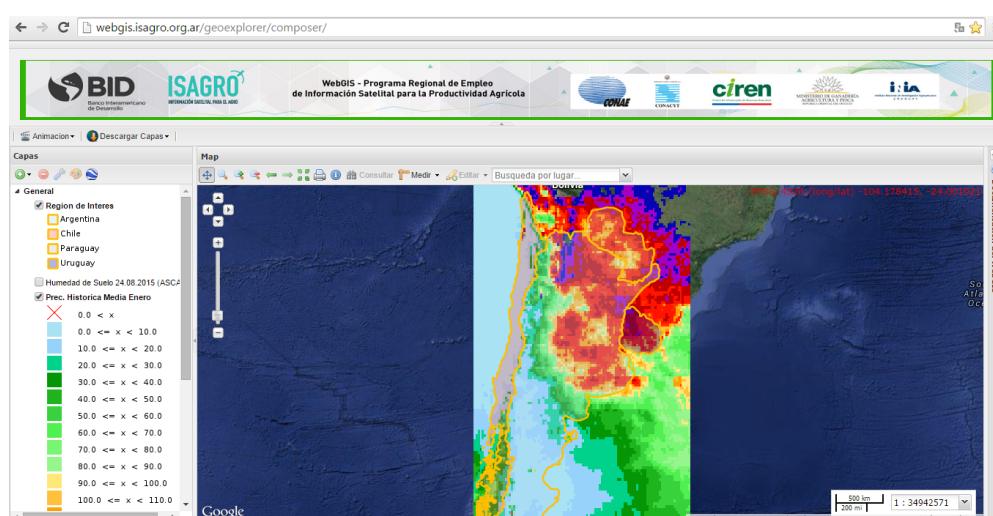


Figura 6.1: Visualizador Web-GIS ISAGRO del proyecto PREISPA

6.2 Líneas de abordaje a futuro

En lo que respecta a trabajo futuro queda un largo camino, a pesar de que con el desarrollo de este trabajo se logró que la plataforma de Chagas que antes únicamente contaba con los datos de planilla, ahora tenga un listado de productos ambientales preparados para ser analizados por los usuarios. Hay varios esfuerzos que requieren ser llevados a cabo para que esta herramienta sea cada vez más importante para las instituciones encargadas de la vigilancia de la enfermedad, a continuación son mencionadas las principales líneas de abordaje a futuro que se consideran de mayor relevancia:

- Es importante continuar y mejorar los esfuerzos para que las instituciones locales, provinciales, y nacionales aumenten la integración de sus registros en campo de la planilla Chagas 6 a la plataforma Web-GIS, de esta manera se logrará obtener un set de datos histórico y de mejor cobertura de las localidades, lo que conllevará a mejores análisis de la enfermedad en Argentina.

- En cuanto a mejoras del sistema, el desarrollo de un módulo de consultas SQL predefinidas, que le den al usuario la posibilidad de obtener datos de los registros históricos de la planilla, sería de gran utilidad. Posibilitando al usuario seleccionar de una lista de atributos el tipo de resumen que desee, esto hace que para los efectores sea transparente toda la generación de la de la búsqueda, debido a que las consultas SQL estarán implementadas e integradas internamente en el sistema.
- Un trabajo de implementación a futuro muy importante, es construir un módulo que integre los scripts generados en este trabajo para la generación de productos, ya sea a pedido, o para que funcione periódicamente actualizando la información de la series temporales. De esta forma la cadena de trabajo que incluye descarga de datos, generación del producto, incorporación al sistema y publicación de las capas se realiza de forma automática cada vez que sea requerido.
- Ampliar la cobertura espacial de las capas generadas que en su mayoría fueron únicamente para el Nor-orienté Argentino, con esto se busca que los productos generados en este trabajo sean extendido a todo el territorio argentino.
- Implementar en la interfaz gráfica herramientas estadísticas que permitan que los usuarios obtengan gráficos de los valores de las capas ambientales, o de los datos de planilla.
- Implementar geo-procesos que funcionen en la Web, mediante la interfaz gráfica, entre las capas vectoriales y raster que se encuentran incorporadas al sistema. Este punto es importante, debido a que entre más funcionalidades SIG estén implementadas en el Web-GIS, se hace más fácil la interpretación de los datos, y hace que todos los procesos sean transparentes para el usuario, lo que apunta a que la persona que haga uso del sistema requiera menos conocimientos en temáticas relacionadas con manejo de SIG, apuntando a un público más general.

6.3 Conclusiones

En términos generales con este trabajo de tesis fue posible generar el flujo de trabajo que automatiza un gran número de productos con significado biológico para la enfermedad de Chagas, a partir de datos ambientales monitoreados por sensores remotos y de la planilla Chagas 6. También se logró Incorporar al sistema Web-GIS los productos finales generados para el uso por parte de los usuarios efectores de la vigilancia epidemiológica de la enfermedad en Argentina.

Este desarrollo constituye una mejora al sistema Chagas que no contaba con datos de variables ambientales, y permite el acceso a la información entomológica y a los productos generados desde cualquier lugar en donde los usuarios efectores cuenten con acceso a la Web para ingresar y hacer uso del sistema y sus datos, logrando de esta manera cumplir con el objetivo general planteado.

La mejora de los servicios que ofrece el sistema Web-GIS de Chagas pretende contribuir a la vigilancia epidemiológica, proporcionando materia prima para la construcción de modelos de distribución de la enfermedad en el país que mejoren la estimación de las zonas con alto riesgo, que conlleva a que el Web-GIS de Chagas sirva como herramienta de apoyo para la toma de decisiones por parte de las instituciones efectoras del control de la enfermedad en Argentina.

Se cumplieron tanto los objetivos generales planteados como los objetivos específicos, destacando lo siguiente:

- Se evaluó la disponibilidad web de datos satelitales ambientales, seleccionando las fuentes de datos con fácil acceso y de mayor relevancia biológica para la enfermedad, y fueron debidamente documentados para la utilización futura de los mismos.
- Fue automatizada la descarga de información satelital de las variables ambientales en formato raster desde distintos servidores web y FTP; se implementaron diferentes métodos de acceso como lo fueron la conexión y descarga de los productos MODIS por medio de la librería PyModis, la descarga de los datos *WorldClim* usando la función *GetData* del Software R, y la obtención del modelo de elevación digital desde el GRASS-GIS. Por su parte para los productos vectoriales, para los cuales la fuente de datos es la información contenida en la base de datos PostGIS del sistema Chagas, que provienen de la planilla Chagas 6, se estudió e implementó la interacción de las consultas SQL y la función de PostGIS `pgsql2shp` para extraer los atributos deseados y generar los productos.
- Se analizaron diferentes algoritmos para generar los productos con valor agregado de interés, y fueron seleccionadas dentro de las opciones los que representaran un menor costo en relación al consumo de recursos computacionales y tiempo.
- Fueron implementados los algoritmos estudiados que generan 10 productos raster para un total de 372 capas de variables ambientales, 10 productos vectoriales con información epidemiológica, y un producto que deriva de la interacción de los dos tipos de datos; todos los algoritmos y scripts implementados fueron debidamente documentados a lo largo de este trabajo, y se generaron mediante el uso de software libre, principalmente en GRASS-GIS.
- Para concluir con la cadena de trabajo realizada, fueron utilizadas herramientas de transferencia de datos para la subida de los productos generados y otras capas ambientales a la plataforma Web-GIS sin requerir de la interacción del usuario, se describieron tres formas de incorporar información; la primera de forma automática mediante el comando `cURL`, utilizado principalmente para los productos raster; la segunda mediante la conexión con el servidor por SSH utilizando la función `pgsql2shp`, y la tercera de forma manual con FileZilla. Adicionalmente se documentó el procedimiento necesario para generar el entorno en el GeoServer, como lo son los estilos y almacenes de datos necesarios para la publicación de las capas en el sistema.

Planilla Chagas 6

CHAGAS 6



Ministerio de Salud de la Nación
Dirección de Enfermedades Transmisibles por Vectores

MINISTERIO DE SALUD DE LA NACION, ARGENTINA
DIRECCION DE ENFERMEDADES TRANSMISIBLES POR VECTORES
PROGRAMA NACIONAL DE CHAGAS
CONTROL DE LA VINCHUCA



HOJA NRO.: 2

[illegible]

CHAGAS 6

CONFECCIONADA POR EFECTOR

Referencias

- Adler, R. M. (1995). Emerging standards for component software. *IEEE Computer*, 28(3):68–77.
- Alesheikh, A. A. & Helali, H. (2001). Distributing national geospatial information. *Proceedings of digital Earth*.
- Bar, M. E., Oscherov, E. B., Damborsky, M. P., & Borda, M. (2010). Estudio transversal de la enfermedad de chagas en un área endémica de la provincia de corrientes, argentina. *Boletín de Malariología y salud ambiental*, 2:207–217.
- Beven, K. J. & Kirkby, M. J. (1979). A physically based, variable contributing area model of basin hydrology. *Hydrological Sciences Bulletin*, 24(1):43–69.
- Bowman, M. H. (2005). Mapping freely available high resolution global elevation and vector data in grass. *GRASS Newsletter*, 3:7–10.
- Cabrera, A. (1976). *Regiones Fitogeográficas Argentinas. Enciclopedia Argentina de Agricultura y Jardinería.*, (2 ed.).
- Chuvieco, E. (2006). *Teledetección ambiental: La observación de la Tierra desde el Espacio.*, (2 ed.).
- Cohen, J. M., Ernst, K. C., Lindblade, K. A., Vulule, J. M., John, C. C., & Wilson, M. L. (2008). Topography-derived wetness indices are associated with household-level malaria risk in two communities in the western kenyan highlands. *Malaria Journal*, 7(40):1–12.
- Coura, J. R. & Dias, J. C. (2009). Epidemiology, control and surveillance of chagas disease - 100 years after its discovery. *Memórias do Instituto Oswaldo Cruz*, 104(1):31–40.
- Delucchi, L. (2014). Pymodis documentation release 1.0.0. http://pymodis.fem-environment.eu/_static/pyModis.pdf.
- Dias, J., Silveira, A., & Schofield, C. (2002). The impact of chagas disease control in latin america – a review. *Memórias do Instituto Oswaldo Cruz*, 97:603–612.
- Diaz, J. H. (2007). Chagas disease in the united states: A cause for concern in louisiana? *Journal of the Louisiana State Medical Society*, 159:21–23.
- Ferreira, C. S., Marthinho, P. C., Amato-Neto, V., & Cruz, R. R. B. (2001). Pasteurization of human milk to prevent transmission of chagas disease. *Revista do Instituto de Medicina Tropical de São Paulo*, 43:161–162.

- Furlanello, C., Merler, S., Menegon, S., Mancuso, S., & G., B. (2002). New webgis technologies for geolocation of epidemiological data: an application for the surveillance of the risk of lyme borrelliosis disease. *Giornale Italiano di Aritmologia e Cardiostimolazione*, 5(1):241–245.
- Galvao, C., Carcavallo, R., Rocha, D., & Jurberg, J. (2003). A checklist of the current valid species of the subfamily triatominae jeannel, 1919 (hemiptera, reduviidae) and their geographical distribution, with nomenclatural and taxonomic notes. *Zootaxa*, 2002:1–36.
- Gorla, D. E. (2002). Variables ambientales registradas por sensores remotos como indicadores de la distribución geográfica de *Triatoma infestans* (heteroptera: Reduviidae). *Ecología Austral*, 12(2):117–127.
- Gorla, D. E., Porcasi, X., & Catalá, S. S. (2005). Sistemas de información geográfica y sensores remotos como herramientas en los programas de control vectorial de la enfermedad de chagas. *Proceedings de la VI Reunión de la Iniciativa Andina para el Control de la Enfermedad de Chagas*, pages 265–276.
- Graham, A. J., Atkinson, P. M., & Danson, F. M. (2004). Spatial analysis for epidemiology. *Acta Tropical*, 91:219–225.
- Guhl, F., Restrepo, M., Angulo, V. M., Antunes, C., Campbell-Lendrum, D., & Clive, R. D. (2005). Lessons from a national survey of chagas disease transmission risk in colombia. *Trends in Parasitology*, 21(6):259–262.
- Gulh, F., Aguilera, G., Pinto, N., & Vergara, D. (2007). Actualización de la distribución geográfica y ecoepidemiología de la fauna de triatominos (reduviidae: Triatominae) en colombia. *Biomédica*, 27(1):143–162.
- Gulh, F., Pinto, N., Aguilera, G., & Silveira, A. C. (2005). Distribución actual de las principales especies de triatominos asociadas al domicilio en colombia. *Memorias del primer taller internacional sobre control de la enfermedad de Chagas*, pages 49–54.
- Gurgel-Goncalves, R., Galvao, C., Costa, J., & Peterson, T. (2012). Geographic distribution of chagas disease vectors in brazil based on ecological niche modeling. *Journal of tropical medicine*, pages 1–15.
- Gurtler, R. E., Cecerre, M. C., Lauricella, M. A., Cardinal, M. V., Kitron, U., & Cohen, J. E. (2006). Domestic dogs and cats as sources of *Trypanosoma cruzi* infection in rural northwestern argentina. *Parasitology*, 134:69–82.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25(15):1965–1978.
- Hijmans, R. J. & Van Etten, J. (2012). *Geographic analysis and modeling with raster data. R package "Raster"*. R Foundation for Statistical Computing.
- Jakubauskas, M. E., Legates, D. R., & Kastens, J. H. (2001). Harmonic analysis of time-series avhrr ndvi data. *Photogrammetric Engineering and Remote Sensing*, 67(4):461–470.
- Jakubauskas, M. E., Legates, D. R., & Kastens, J. H. (2002). Crop identification using harmonic analysis of time-series avhrr ndvi data. *Computers and Electronics in Agriculture*, 37:127–139.

- Jansen, A. M., Pinho, A. P., Lisboa, C. V., Cupolillo, E., Mangia, R. H., & Fernandes, O. (1999). The sylvatic cycle of *trypanosoma cruzi*: a still unsolved puzzle. *Memórias do Instituto Oswaldo Cruz*, 94:203–206.
- Jurberg, J., Galvão, C., Noireau, F., Carcavallo, R. U., Da Silva Rocha, D., & Lent, H. (2004). Uma iconografia dos triatomíneos (hemiptera, reduviidae). *Entomología y Vectores*, 11(1):457–494.
- Lanfri, S. (2011). *Desarrollo de una metodología para la detección de cuerpos de agua mediante el análisis de imágenes SAR COSMO SkyMed y de DEMs*. PhD thesis, Universidad Nacional de Córdoba, Argentina.
- Laosuwan, T. (2012). Online web gis-based services for spatial data and sharing of leptospirosis epidemiology information; development of pilot project in mahasarakham province, thailand. *International journal of geomatics and geosciences*, 3(1):121–133.
- Lazzari, C. R. (1991). Temperature preference in *Triatoma infestans* (hemiptera: Reduviidae). *Bulletin of Entomological Research*, 81:273–276.
- Martins-Melo, F. R., Novaes, A. J., Alencar, C. H., & Lange, W. (2012). Mortality of chaga's disease in brazil: spatial patterns and definition of high-risk areas. *Tropical Medicine and International Health*, 17(9):1066–1075.
- Ministerio de Salud Argentina (2002). Informe chagas. Technical report, MSA.
- Ministerio de Salud Argentina (2012). Guía para el control vectorial de la enfermedad de chagas. *Programa Nacional de Chagas*, pages 1–65.
- Moncayo, A. & Ortiz, M. I. (2006). An update on chagas disease (human american trypanosomiasis). *Annals of Tropical Medicine and Parasitology*, 100(6):1–15.
- Moncayo, A. & Silveira, A. C. (2009). Current epidemiological trends for chagas disease in latin america and future challenges in epidemiology, surveillance and health policy. *Memórias do Instituto Oswaldo Cruz*, 104:17–30.
- MSAL & CONAE (2013). Instructivo general de captura, almacenamiento y procesamiento de datos. Technical report, Ministerio de Salud Argentina y Comisión Nacional de Actividades Espaciales.
- Neteler, M. (2005). Srtm and vmap0 data in ogr and grass. *GRASS Newsletter*, 3:2–6.
- Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). Grass gis: A multi-purpose open source gis. *Environmental Modelling and Software*, 31:124–130.
- Neteler, M. & Mitasova, H. (2008). *Open source GIS: A GRASS GIS approach*, (3 ed.).
- Nieto, P., Malone, J. B., & Bavia, M. E. (2006). Ecological niche modeling for visceral leishmaniasis in the state of bahia, brazil, using genetic algorithm for rule-set prediction and growing degree day-water budget analysis. *Geospatial Health*, 1:115–126.
- OpenGeo Organization (2011). The opengeo architecture. Technical report, OpenGeo.
- Parra-Henao, G., Segura-Cardona, A., Quirós-Gómez, O., Angulo, V., & Neal, A. (2015). House-level risk factors for *triatoma dimidiata* infestation in colombia. *American Journal of Tropical Medicine and Hygiene*, 92(1):193–200.

- Parra-Henao, G. J. (2010). Sistemas de información geográfica y sensores remotos. aplicaciones en enfermedades transmitidas por vectores. *Revista CES Medicina*, 24(2):75–90.
- Parra-Henao, G. J. (2012). *Triatoma dimidiata en Colombia. Distribución, Ecología, Análisis de Factores de Riesgo y Generación de Mapas predictivos de Riesgo*. PhD thesis, Universidad de Antioquia, Colombia.
- Peralta, G., De Elia, E. A., Lanfri, M., Porcasi, X., Lanfri, S., Frutos, N., Rotela, C., & Scavuzzo, C. M. (2011). Geomática aplicada a un sistema de alerta temprana. Technical report, CONAE.
- PgAdmin (2015). Introduction. <http://www.pgadmin.org>.
- Porcasi, X., Catala, S., Hrellac, H., Scavuzzo, C., & Gorla, D. (2006). Infestation of rural houses by triatoma infestans (hemiptera: Reduviidae) in southern area of gran chaco in argentina. *Journal of medical entomology*, 43(5):1060–1067.
- Porcasi, X., Catala, S., Moreno, M., Abraham, L., Hernandez, L., & Gorla, D. (2007). Infestation of rural houses by triatoma infestans in the region of los llanos (la rioja, argentina). *Memórias do Instituto Oswaldo Cruz*, 102(1):63–68.
- Porcasi, X., De Elia, E. A., & Scavuzzo, C. M. (2013). Documentación sistema de chagas. Technical report, CONAE.
- Porcasi, X., Rotela, C. H., Introini, M. V., Frutos, N., Lanfri, S., Peralta, G., De Elia, E. A., Lanfri, M., & Scavuzzo, C. M. (2012). An operative dengue risk stratification system in argentina based on geospatial technology. *Geospatial Health*, 6(3):531–542.
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- Ramírez, J. D., Montilla, M., Cucunubá, Z. M., Florez, A. C., & Zambrano, P. (2013). Molecular epidemiology of human oral chagas disease outbreaks in colombia. *Plos Neglected Tropical Diseases*, 7(2):1–7.
- Rassi, A. J., Rassi, A., & Marin-Neto, J. A. (2010). Chagas disease. *Lancet*, 375:1388–1402.
- Red Chagas Colombia (2014). Programa nacional de investigación para la prevención, control y tratamiento integral de la enfermedad de chagas en colombia 2012-2020. <http://www.redchagascolombia.org/>.
- Roerink, G. J., Menenti, M., & Verhoef, M. (2000). Reconstructing cloudfree ndvi composites using fourier analysis of time series. *International Journal of Remote Sensing*, 21(9):1911–1917.
- Rogers, D. (2000). *Satellites, space, time and the African trypanosomiasis*. En: *Remote sensing and Geographic Information Systems in epidemiology. Advances in Parasitology*, [Hay S. I., Randolph, S. E. and Rogers, D. J.], (1 ed.).
- Rosas, F., Vanegas, D. I., & Cabrales, M. F. (2007). *Enfermedad de chagas*, (1° ed.). Sociedad Colombiana de Cardiología y Cirugía vascular.
- Scavuzzo, C. M. & Torrusio, S. (2010). Project management plan for hap. Technical report, CONAE.
- Schmunis, G. A. (2007). Epidemiology of chagas disease in non-endemic countries: the role of international migration. *Memórias do Instituto Oswaldo Cruz*, 102:75–85.

- Steverding, D. (2014). The history of chagas disease. *Parasites and Vectors*, pages 7–317.
- Tatem, A. J., Goetz, S., & Hay, S. (2004). Terra and aqua: new data for epidemiology and public health. *International Journal of Applied Earth Observation and Geoinformation*, 6:36–41.
- United States Geological Survey (1999). Modis land cover product algorithm theoretical basis document (atbd). http://modis.gsfc.nasa.gov/data/atbd/atbd_mod12.pdf.
- United States Geological Survey (2006). Modis land surface temperature products user's guide. http://lpdaac.usgs.gov/products/modis_products_table/mod11a1/.
- United States Geological Survey (2011). Vegetation indices 16-day 13 global 250m. http://lpdaac.usgs.gov/products/modis_products_table/mod13q1/.
- United States Geological Survey (2013). User guide for the modis land cover type product (mcd12q1). https://lpdaac.usgs.gov/sites/default/files/public/modis/docs/MCD12Q1_UserGuide_07302013.pdf.
- Vazquez-Prokopec, G. M., Spillmann, C., Zaidenberg, M., Gürtler, R. E., & Kitron, U. (2012). Spatial heterogeneity and risk maps of community infestation by triatoma infestans in rural northwestern argentina. *Plos Neglected Tropical Diseases*, 6(8):e1788.
- World Health Organization (2002). Control of chagas disease. Technical Report 905, WHO.
- World Health Organization (2012). Research priorities for chagas disease, human african trypanosomiasis and leishmaniasis. Technical Report 975, WHO.
- Yang, G. J., Vounatsou, P. B., Xiao, Z., Utzinger, J. B., & Tanner, M. B. (2005). A review of geographic information system and remote sensing with applications to the epidemiology and control of schistosomiasis in china. *Acta Tropica*, 96:117–129.