

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE CIENCIAS**

**MODELACIÓN MATEMÁTICA PARA LA PREDICTIBILIDAD DEL  
DENGUE EN ECUADOR**

**TRABAJO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERA  
MATEMÁTICA**

**PROYECTO DE INVESTIGACIÓN**

**DIANA PAULINA TAIPE HIDALGO**  
**diany\_paulina28@hotmail.com**

**Directora: DRA. SANDRA ELIZABETH GUTIÉRREZ POMBOSA**  
**sandra.gutierrez@epn.edu.ec**

**QUITO, ENERO, 2017**

## DECLARACIÓN

Yo DIANA PAULINA TAIPE HIDALGO, declaro bajo juramento que el trabajo aquí escrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual, correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normatividad institucional vigente.



Diana Paulina Taipe Hidalgo

## **CERTIFICACIÓN**

Certifico que el presente trabajo fue desarrollado por DIANA PAULINA TAIPE HIDALGO, bajo mi supervisión.



---

Dra. Sandra Elizabeth Gutiérrez Pombosa  
Director del Proyecto

## AGRADECIMIENTOS

Agradezco al Instituto Nacional de Investigación en Salud Pública (**INSPI**), especialmente al Centro de Investigación y Referencia de Vectores (**CIREV**) quienes en conjunto con la Dra. Varsovia Cevallos, el Lcdo. Diego Morales y su equipo de colaboradores me han brindado el tiempo y las herramientas necesarias para culminar este proyecto, siempre a la vanguardia de la información de campo, la elaboración de mapas y sobretodo en el ámbito de los conceptos biológicos que permitieron entender la dinámica de los vectores y la interacción con el hombre. De igual manera extiendo mi gratitud a la Dra. Emmanuelle Quentin por su aporte en la entrega de datos.

Mis sinceros agradecimientos a la Dra. Sandra Gutiérrez por su labor continua, su consejo a nivel académico y el tiempo infinito para ayudarme con paciencia en todas las adversidades.

A Rubén mi padre, por tomarme de su mano enseñándome desde muy pequeña el valor del esfuerzo y la paciencia.

A mis hermanos: Ricardo y Pablo, por ser los cómplices incondicionales que alegran mi vida.

A Berthita y Rita por estar junto a mí en cada paso, siempre con una sonrisa presa y con su cariño sincero.

A María José, Esteban y Daniel, mis grandes amigos, por regalarme la oportunidad de compartir junto a ellos este camino de grandes anécdotas. Sin su compañía, el diario vivir no habría tomado el humor y el optimismo para culminar nuestra carrera universitaria.

A mis compañeros Andrés y Saúl por animarme en las fases finales de este proyecto.

## **DEDICATORIA**

*A Elena, le dedico este trabajo por ser la más sublime inspiración de mi vida. Gracias madre por construirme alas preciosas para conocer la libertad y por cambiar el color de mis días con tu sonrisa infinita.*

*Mi agradecimiento eterno a mi más perfecta amiga.*

*A Delia y María Erlinda, mis abuelas, por dejar en mí una herencia infinita de cariño y recuerdos únicos. Gracias por las lecciones de vida y primordialmente por mostrarme que el esfuerzo en el presente es la recompensa del mañana, aún cuando el sol se demore en salir.*

# Índice general

<b>Resumen</b>	<b>xiv</b>
<b>Abstract</b>	<b>xv</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Antecedentes . . . . .	2
1.2. Área de Estudio . . . . .	3
1.3. Justificación . . . . .	5
1.4. Objetivos . . . . .	6
1.4.1. Objetivos Generales . . . . .	6
1.4.2. Objetivos Específicos . . . . .	6
1.5. Hipótesis . . . . .	7
1.6. Metodología . . . . .	7
1.7. Alcances . . . . .	8
1.8. Estructura del Trabajo . . . . .	8
<b>2. Marco Teórico</b>	<b>10</b>
2.1. Nociones Básicas . . . . .	12
2.1.1. Distribución a Priori . . . . .	16
2.2. Métodos Monte Carlo de Cadenas de Markov (MCMC) . . . . .	16
2.2.1. Algoritmo Metropolis Hastings . . . . .	18
2.2.2. Muestreador de Gibbs . . . . .	19
2.3. Modelo Condicional Autorregresivo . . . . .	20
2.4. Criterios para la comparación de Modelos . . . . .	21
2.4.1. Criterio de Información de Akaike (AIC) . . . . .	22

2.4.2. Criterio de Información Bayesiano (BIC) . . . . .	22
2.4.3. Criterio de Información de Devianza (DIC) . . . . .	23
2.5. Índice I de Morán . . . . .	24
2.6. Estudio de los Residuos . . . . .	25
<b>3. Construcción de la Base de Datos</b>	<b>27</b>
3.1. Clasificación de los casos de dengue en el período 2010-2015 . . . . .	27
3.2. Metodología para interpolar tamaños poblacionales. . . . .	30
3.3. Construcción de Variables . . . . .	30
3.3.1. Indicadores sociales . . . . .	31
3.3.2. Indicadores climáticos . . . . .	36
3.4. Uso de WinBUGS14 . . . . .	42
3.4.1. Análisis de Convergencia . . . . .	43
3.5. Construcción de Mapas por Circuitos . . . . .	45
<b>4. Modelo Bayesiano Espacio Temporal</b>	<b>47</b>
4.1. Distribuciones a priori . . . . .	50
4.2. Distribuciones a posteriori . . . . .	51
4.2.1. Distribución condicional para $\alpha_t$ . . . . .	52
4.2.2. Distribución condicional para $\beta_t^j$ . . . . .	52
4.2.3. Distribución condicional para $\gamma_s^k$ . . . . .	53
4.2.4. Distribución condicional para $\nu_{st}$ . . . . .	53
4.2.5. Distribución condicional para $\mu_{st}$ . . . . .	53
4.2.6. Distribución condicional para $\tau_\mu$ . . . . .	53
4.2.7. Distribución condicional para $\tau_\nu$ . . . . .	54
<b>5. Resultados</b>	<b>55</b>
5.1. Descripción de los modelos . . . . .	55
5.2. Modelos espacio-temporales monitoreados en WinBUGS14 . . . . .	56
5.3. Comparación de los Modelos . . . . .	57
5.4. Modelo elegido . . . . .	58
5.5. Monitoreo de la Convergencia para el Modelo Elegido . . . . .	59
5.5.1. Convergencia para el parámetro $\beta_t^1$ . . . . .	59

5.5.2. Convergencia para el parámetro $\beta_t^2$	65
5.5.3. Convergencia para el parámetro $\gamma_s^1$	72
5.5.4. Convergencia para $\tau_\mu$	77
5.6. Tablas y Mapas de Riesgo para el período 2010-2015	80
5.6.1. Riesgo Relativo para el año 2010	81
5.6.2. Riesgo Relativo para el año 2011	85
5.6.3. Riesgo Relativo para el año 2012	90
5.6.4. Riesgo Relativo para el año 2013	94
5.6.5. Riesgo Relativo para el año 2014	98
5.6.6. Riesgo Relativo para el año 2015	103
5.7. Análisis de Residuos	106
<b>6. Limitaciones, Conclusiones, Recomendaciones, y Trabajo Futuro</b>	<b>110</b>
6.1. Limitaciones	110
6.2. Conclusiones	110
6.3. Recomendaciones	112
6.4. Trabajo Futuro	113
<b>7. Anexos</b>	<b>114</b>
7.1. Definición de las Variables	114
7.2. Matriz de vecindad	120
7.3. Códigos	120
7.3.1. Datos	122
7.3.2. Datos iniciales para la cadena 1	130
7.3.3. Datos iniciales para la cadena 2	131
<b>Bibliografía</b>	<b>134</b>

# Índice de cuadros

1.1. División territorial por circuitos de Guayaquil, Fuente: SENPLADES, Elaboración: CIREV . . . . .	5
3.1. Ilustración de la primera semana epidemiológica del año 2016 . . . . .	28
3.2. Matriz de correlaciones para las variables agptbl, alcrt, anlf, bsr, dsmpl, vstd . . . . .	32
3.3. Importancia de los componentes principales, CP: Componente Prin- cipal . . . . .	33
3.4. Coeficientes para CP1 y CP2, de acuerdo a cada variable. . . . .	34
3.5. Tabla de códigos para los circuitos de la Ciudad de Guayaquil . . . . .	45
5.1. Ajuste bayesiano para los modelos monitoreados en WinBUGS14 . .	58
5.2. Estadísticos a posteriori para $\beta^1$ , año 2010, estaciones lluviosa y seca	64
5.3. Estadísticos a posteriori para $\beta^1$ , año 2011, estaciones lluviosa y seca	64
5.4. Estadísticos a posteriori para $\beta^1$ , año 2012, estaciones lluviosa y seca	64
5.5. Estadísticos a posteriori para $\beta^1$ , año 2013, estaciones lluviosa y seca	65
5.6. Estadísticos a posteriori para $\beta^1$ , año 2014, estaciones lluviosa y seca	65
5.7. Estadísticos a posteriori para $\beta^1$ , año 2015, estaciones lluviosa y seca	65
5.8. Estadísticos a posteriori para $\beta^2$ , año 2010, estaciones lluviosa y seca	71
5.9. Estadísticos a posteriori para $\beta^2$ , año 2011, estaciones lluviosa y seca	71
5.10. Estadísticos a posteriori para $\beta^2$ , año 2012, estaciones lluviosa y seca	71
5.11. Estadísticos a posteriori para $\beta^2$ , año 2013, estaciones lluviosa y seca	71
5.12. Estadísticos a posteriori para $\beta^2$ , año 2014, estaciones lluviosa y seca	72
5.13. Estadísticos a posteriori para $\beta^2$ , año 2015, estaciones lluviosa y seca	72
5.14. Estadísticos a posteriori para $\gamma_1$ para los circuitos: (1-24) . . . . .	76
5.15. Estadísticos a posteriori para $\gamma_1$ para los circuitos: 25-48 . . . . .	77

5.16. Estadísticos a posteriori para $\tau_\mu$ , año 2010 . . . . .	79
5.17. Estadísticos a posteriori para $\tau_\mu$ , año 2011 . . . . .	79
5.18. Estadísticos a posteriori para $\tau_\mu$ , año 2012 . . . . .	79
5.19. Estadísticos a posteriori para $\tau_\mu$ , año 2013 . . . . .	79
5.20. Estadísticos a posteriori para $\tau_\mu$ , año 2014 . . . . .	80
5.21. Estadísticos a posteriori para $\tau_\mu$ , año 2015 . . . . .	80
5.22. Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2010, RR=riesgo relativo, IC=Intervalo de confianza . . . . .	81
5.23. Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2011, RR=Riesgo Relativo, IC=Intervalo de confianza . . . . .	85
5.24. Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2012 . . . . .	90
5.25. Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2013 . . . . .	94
5.26. Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2014 . . . . .	99
5.27. Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2015 . . . . .	103
5.28. Residuos del modelo elegido, año 2010 . . . . .	106
5.29. Residuos del modelo elegido, año 2011 . . . . .	107
5.30. Residuos del modelo elegido, año 2012 . . . . .	107
5.31. Residuos del modelo elegido, año 2013 . . . . .	108
5.32. Residuos del modelo elegido, año 2014 . . . . .	108
5.33. Residuos del modelo elegido, año 2015 . . . . .	109
7.1. Matriz W de adyacencia para los 48 circuitos de la Ciudad de Guayaquil	120

# Índice de figuras

1.1. Mapa de la Ciudad de Guayaquil sectorizada en Circuitos, Fuente: SENPLADES . . . . .	4
2.1. Esquema intuitivo del comportamiento de las distribuciones a priori y posteriori . . . . .	14
3.1. Serie de casos de dengue por circuitos en la Ciudad de Guayaquil, en el período 2010-2015 . . . . .	29
3.2. Proporción de variabilidad explicada para cada componente asocia- dos los indicadores seleccionados . . . . .	33
3.3. Círculo de correlaciones para las variables: agptbl, alcstr, anlf, bsr, dsmpl, vstd . . . . .	34
3.4. Mapa de Precipitación, época lluviosa, Fuente:EPISIG-INSPI . . . . .	37
3.5. Mapa de Precipitación, época seca, Fuente:EPISIG-INSPI . . . . .	38
3.6. Mapa de Temperatura, época lluviosa, Fuente:EPISIG-INSPI . . . . .	39
3.7. Mapa de Temperatura, época seca, Fuente:EPISIG-INSPI . . . . .	39
3.8. Mapa del Índice de vegetación, época lluviosa, Fuente:EPISIG-INSPI . . . . .	40
3.9. Mapa del Índice de vegetación, época seca, Fuente:EPISIG-INSPI . . . . .	41
3.10. Mapa de Altitud, Fuente: EPISIG-INSPI . . . . .	41
3.11. Nomenclatura para definir el código de un circuito . . . . .	45
5.1. History plot para la variable $\beta_t^1$ , desde enero 2010 - agosto 2010 . . . . .	59
5.2. History plot para la variable $\beta_t^1$ , desde septiembre 2010 - abril 2011 . . . . .	60
5.3. History plot para la variable $\beta_t^1$ , desde mayo 2011 - diciembre 2011 . . . . .	60
5.4. History plot para la variable $\beta_t^1$ , desde enero 2012 - agosto 2012 . . . . .	60
5.5. History plot para la variable $\beta_t^1$ , desde septiembre 2012 - abril 2013 . . . . .	61
5.6. History plot para la variable $\beta_t^1$ , desde mayo 2013 - diciembre 2013 . . . . .	61

5.7. History plot para la variable $\beta_t^1$ , desde enero 2014 - agosto 2014 . . . . .	61
5.8. History plot para la variable $\beta_t^1$ , desde septiembre 2014 - abril 2015 . . . . .	62
5.9. History plot para la variable $\beta_t^1$ , desde mayo 2015 - diciembre 2015 . . . . .	62
5.10. Gráfico de la densidad para la variable $\beta_t^1$ , desde enero 2010- junio 2012 . . . . .	63
5.11. Gráfico de la densidad para la variable $\beta_t^1$ , desde julio 2012-diciembre 2014 . . . . .	63
5.12. Gráfico de la densidad para la variable $\beta_t^1$ , desde enero 2015- diciembre 2015 . . . . .	64
5.13. History plot para la variable $\beta_t^2$ , desde enero 2010 - agosto 2010 . . . . .	65
5.14. History plot para la variable $\beta_t^2$ , desde septiembre 2010 - abril 2011 . . . . .	66
5.15. History plot para la variable $\beta_t^2$ , desde mayo 2011 - diciembre 2011 . . . . .	66
5.16. History plot para la variable $\beta_t^2$ , desde enero 2012 - agosto 2012 . . . . .	67
5.17. History plot para la variable $\beta_t^2$ , desde septiembre 2012 - abril 2013 . . . . .	67
5.18. History plot para la variable $\beta_t^2$ , desde mayo 2013 - diciembre 2013 . . . . .	68
5.19. History plot para la variable $\beta_t^2$ , desde enero 2014 - agosto 2014 . . . . .	68
5.20. History plot para la variable $\beta_t^2$ , desde septiembre 2014 - abril 2015 . . . . .	69
5.21. History plot para la variable $\beta_t^2$ , desde mayo 2015 - diciembre 2015 . . . . .	69
5.22. Gráfico de la densidad para la variable $\beta_t^2$ , desde enero 2010- junio 2012 . . . . .	70
5.23. Gráfico de la densidad para la variable $\beta_t^2$ , desde julio 2012-diciembre 2014 . . . . .	70
5.24. Gráfico de la densidad para la variable $\beta_t^2$ , desde enero 2015- diciembre 2015 . . . . .	71
5.25. History plot para la variable $\gamma_s^1$ para los circuitos: (1-8) . . . . .	72
5.26. History plot para la variable $\gamma_s^1$ para los circuitos: (9-16) . . . . .	73
5.27. History plot para la variable $\gamma_s^1$ para los circuitos: (17-24) . . . . .	73
5.28. History plot para la variable $\gamma_s^1$ para los circuitos: (25-32) . . . . .	74
5.29. History plot para la variable $\gamma_s^1$ para los circuitos:(33-39) . . . . .	74
5.30. History plot para la variable $\gamma_s^1$ para los circuitos: (40-48) . . . . .	75
5.31. Gráfico de la densidad para la variable $\gamma_s^1$ , para los circuitos: (1-24) . . . . .	75
5.32. Gráfico de la densidad para la variable $\gamma_s^1$ , para los circuitos: (25-48) . . . . .	76
5.33. Gráfico de autocorrelación para la variable $\tau_\mu$ para los meses: 1 al 30 . . . . .	78

5.34. Gráfico de autocorrelación para la variable $\tau_\mu$ para los meses: 31 al 60	78
5.35. Gráfico de autocorrelación para la variable $\tau_\mu$ para los meses: 61 al 72 . . . . .	79
5.36. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2010, época lluviosa . . . . .	82
5.37. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2010, época seca . . . . .	83
5.38. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2011, época lluviosa . . . . .	87
5.39. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2011, época seca . . . . .	88
5.40. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2012, época lluviosa . . . . .	92
5.41. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2012, época seca . . . . .	93
5.42. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2013, época lluviosa . . . . .	96
5.43. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2013, época seca . . . . .	97
5.44. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2014, época lluviosa . . . . .	101
5.45. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2014, época seca . . . . .	102
5.46. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2015, época lluviosa . . . . .	104
5.47. Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2015, época seca . . . . .	105

# Resumen

La aparición de las enfermedades vectoriales, en especial las transmitidas por el mosquito *Aedes aegypti* se han convertido en un problema en el ámbito de la salud pública, especialmente por el continuo movimiento poblacional, el crecimiento en las zonas urbanas e incluso la variabilidad mensual en la precipitación y temperatura. En este proyecto se utiliza la teoría bayesiana y la generación de Cadenas de Markov, fusionando el conocimiento a priori del investigador y algunos criterios adquiridos en el estudio de vectores para aproximar la realidad hacia un conjunto de factores que expliquen la propagación, contagio y dispersión de la enfermedad, concentrándose a nivel de circuitos con temporalidad mensual. Apoyados con este conjunto de herramientas, el riesgo relativo para cada año en el período 2010 al 2015 nos da una idea de cómo cambia este indicador de acuerdo a los fluctuaciones climáticas y la inclusión de las condiciones de vivienda y salubridad, propias de cada localidad. En los resultados determinamos que el aumento en el número de casos de dengue está relacionado con las zonas que aún presentan deficiencias como limitación en la recolección de basura, el acceso al agua corriente, el inadecuado sistema de alcantarillado, que en combinación con el incremento de lluvia en los meses de febrero a mayo, julio y diciembre, son escenario idóneo para la producción de criaderos del mosquito transmisor y su reproducción.

# Abstract

The appearance of vector diseases, especially diseases transmitted by *Aedes aegypti* mosquito have become a problem in the field of public health, especially by the continuous movement population, growth in urban areas and even the monthly variability in precipitation and temperature. In this project we use the Bayesian theory and the generation of Markov chains, merging the a priori knowledge of the researcher and some criteria acquired in the study of vectors to approximate reality towards a set of factors that explain the spread of the disease, concentrating at the level of circuits with monthly temporality. Supported with this toolkit, the relative risk for each year in the period 2010 to 2015 gives us an idea how to change this indicator according to climatic fluctuations and the inclusion of housing conditions and health, specific to each locality. The results determined that the increase in the number of dengue cases is related to the areas still limitation deficient as garbage collection, water access current, inadequate sewage system, which in combination with the increase rain in the months of February to May, July and December are scenario suitable for the production of mosquito breeding sites and reproduction

# Capítulo 1

## Introducción

El dengue es una enfermedad vírica transmitida por mosquitos hembra de la especie *Aedes aegypti*. Una vez que una persona es contagiada por cualquiera de los cuatro serotipos, los síntomas aparecen entre 4 y 7 días (en promedio) posterior a la picadura. Se conoce que si un individuo padece de cualquiera de los 4 serotipos nunca volverá a padecer del mismo, pero es posible que pueda contraer cualquiera de los restantes. Los síntomas que acompañan esta enfermedad son: dolor de cabeza intenso, dolor detrás de los globos oculares, dolores en músculos y articulaciones, náuseas, vómito o sarpullido. [17]

Cuando el mosquito se alimenta de la sangre de un individuo enfermo con dengue es capaz de contagiar a otra persona, es decir, es la única manera de que la enfermedad sea transmitida, puesto que no lo es directamente de una persona a otra, ni a través de objetos. La reproducción del *Aedes aegypti* se lleva a cabo principalmente en recipientes artificiales que contienen agua estancada, además la mayoría de mosquitos hembras pueden pasar su ciclo vital en el interior de las casas en las que se han convertido en adultos o alrededor de ellas. Los representantes de esta especie suelen volar alrededor de 400 metros, lo que nos permite intuir que son las personas más que los mosquitos quienes se desplazan y se convierten en las principales propagadoras del virus en sus localidades de residencia. [17]

Actualmente, la dispersión del dengue afecta a la mayor parte de los países de Asia y América Latina, convirtiéndose en una de las principales causas de hospitalización y muerte, especialmente para los niños de dichas regiones [17].

## **1.1. Antecedentes**

En las últimas décadas, la incidencia del dengue en el mundo ha manifestado un crecimiento considerable, contemplando recientes estimaciones alrededor de 3.900 millones de personas y 128 países afectados, todos vulnerables al riesgo de infección a causa de este virus [7].

En el panorama mundial se puede decir que la enfermedad es endémica en las regiones de: África, América, Mediterráneo Oriental, Asia Sudoriental y el Pacífico Occidental. Según los reportes oficiales de la Organización Mundial de la Salud (OMS), para el período 2008-2013 se reportaron 1.2 millones de casos. Posteriormente se registraron 2.35 millones de enfermos sólo en las Américas y en la transición del año 2014 al 2015 los países insulares del Pacífico, Brasil y sus países vecinos tuvieron un incremento alarmante en el número de casos. Anualmente, a nivel mundial alrededor de 50.000 personas que padecen dengue grave necesitan hospitalización y un 2.5 % de ellas fallecen [17].

En Ecuador, el virus del dengue fue detectado en el año 1988 y hasta la actualidad la región litoral registra la mayor cantidad de casos de dengue clásico, seguida por la región Oriental. En tanto, en las Islas Galápagos el dengue clásico se detectó en el 2002 y hacia el 2006 aparecieron los casos de dengue hemorrágico.

En las gacetas epidemiológicas emitidas por el Ministerio de Salud Pública (MSP), en los meses de enero y diciembre del 2014 los casos de dengue fueron 15.446, mientras que en el 2015 la cifra ascendió a 27.400 enfermos, lo que claramente desencadenó un avance progresivo del dengue relacionado con la dispersión del vector, la circulación de serotipos virales observados en varios ciclos epidémicos y que coincidieron con el incremento de las precipitaciones.

En este entorno, en la interacción entre el vector, la transmisión del virus y el contacto con el hombre se encuentran las condiciones propicias para la explosiva reproducción del mosquito, el mismo que deposita sus huevos en el agua, en los alrededores o interiores de viviendas y especialmente en recipientes destinados al almacenamiento de agua u objetos considerados como contenedores (jarrones, tarros, floreros, tinas, neumáticos viejos, otros) [37]. El aporte en los avances en prevención, detección oportuna y la asistencia médica a tiempo demuestran una reducción del 1 % en las tasas de mortalidad, según la OMS [17].

En el desarrollo de este proyecto de investigación, el estudio estará direccionado a la ciudad de Guayaquil, con la finalidad de indagar aquellos factores asociados a la incidencia y contagio de esta enfermedad, dado que el número de casos reportados obedece a una serie de condiciones relacionadas con el aumento de la densidad poblacional, la movilización de las personas y el vector y la urbanización descontrolada vinculada a la calidad de vida de los habitantes. Además prestaremos atención a las condiciones climáticas propias de esta región las cuales favorecen idóneamente a la reproducción del vector [34].

Para alcanzar este objetivo empezaremos con el planteamiento de modelos espaciales y temporales mediante un enfoque bayesiano, luego trataremos de plasmar los resultados en mapas de riesgo que permitan la visualización de aquellos puntos focales que necesitan mayor atención y la aplicación de medidas activas para la generación de una alerta temprana.

En un futuro se desea incentivar este proceso hacia la cobertura de nuevas enfermedades epidemiológicas transmitidas por el mismo vector, haciendo hincapié en la difusión de información preventiva para formular una propuesta que a tiempo evite las pérdidas humanas y aquellas que pueden ser cuantificables en materia económica.

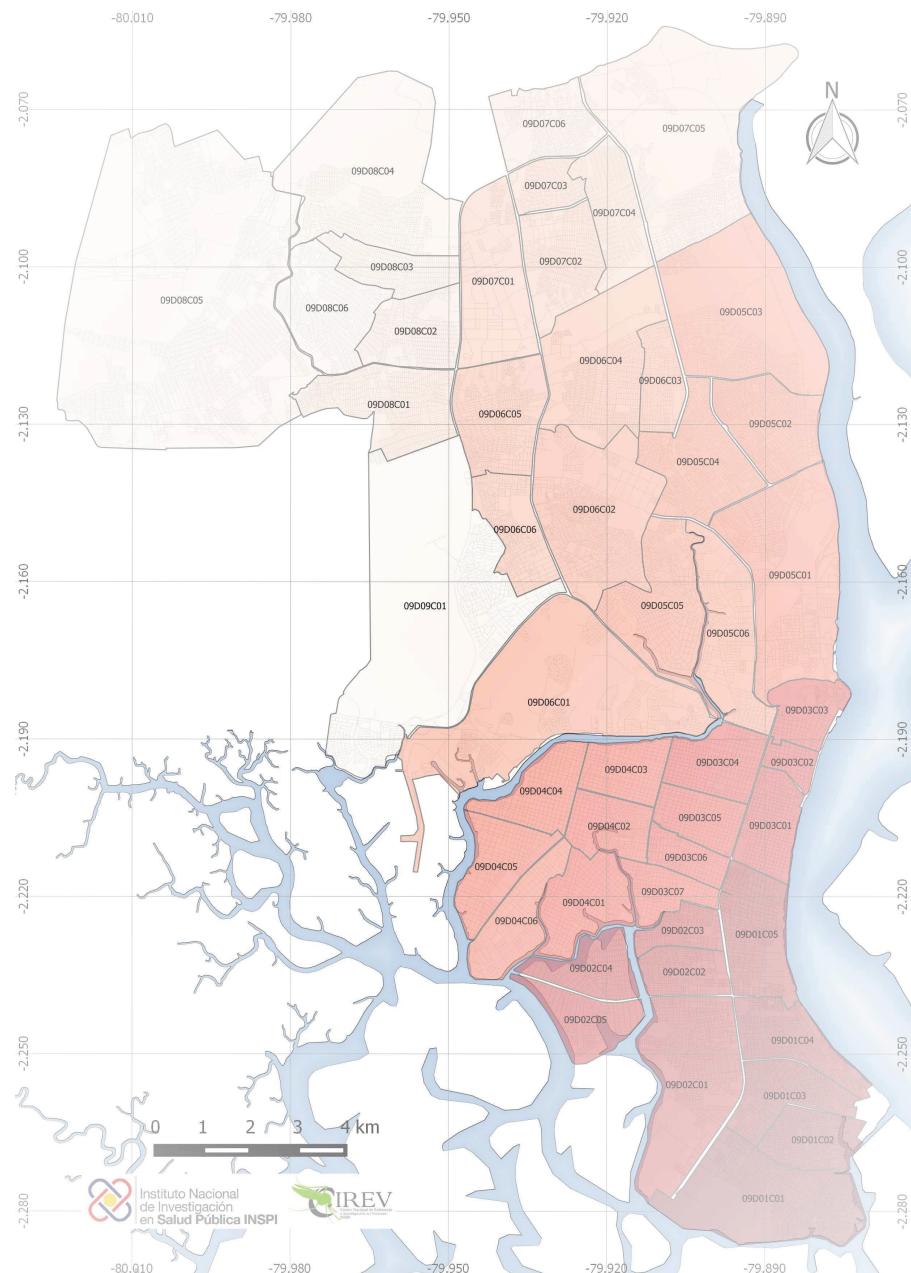
## 1.2. Área de Estudio

Guayaquil está localizada entre las coordenadas  $2^{\circ}11'0''$  S y  $79^{\circ}53'0''$  O, en el Sur Oeste de la provincia de Guayas y se constituye como la ciudad más poblada del Ecuador con 2'291.158 habitantes y con una superficie total de  $345 \text{ km}^2$  [27].

Posee una temperatura cálida que oscila entre los  $25^{\circ}\text{C}$  y  $28^{\circ}\text{C}$  durante casi todo el año, no obstante, su proximidad al Océano Pacífico hace que las corrientes de Humboldt (fría) y la de El Niño (cálida) marquen dos períodos climáticos: una temporada húmeda-lluviosa (período en el que ocurre el 97 % de la precipitación anual) extendida de diciembre a abril y la temporada seca que va de mayo a noviembre [47].

Respecto de la división territorial, la Secretaría Nacional de Planificación y Desarrollo (SENPLADES) ha delimitado las zonas de Guayaquil por distritos y circuitos, para lo cual usaremos la sectorización que se muestra en la figura (1.1), debido a que la información usada posteriormente está mejor concentrada por circuitos en el

sentido de la calidad de los datos.



**Figura 1.1:** Mapa de la Ciudad de Guayaquil sectorizada en Circuitos, Fuente: SEN-PLADES

Código	Círcito	Código	Círcito
09D01C01	GUASMO	09D01C02	UNIÓN DE BANANEROS
09D01C03	FLORESTA	09D01C04	SIETE LAGOS
09D01C05	CENTENARIO	09D02C01	ESTEROS
09D02C02	MALVINAS 1	09D02C03	MALVINAS 2
09D02C04	TRINITARIA 1	09D02C05	TRINITARIA 2
09D03C01	CHILE	09D03C02	VICTORIA
09D03C03	9 DE OCTUBRE	09D03C04	GARAY
09D03C05	VENEZUELA	09D03C06	ANTEPARA
09D03C07	CRISTO DEL CONSUELO	09D04C01	EL CISNE
09D04C02	CHALÁ	09D04C03	SALINAS
09D04C04	PORTETE	09D04C05	BATALLÓN DEL SUBURBIO
09D04C06	SUBURBIO	09D05C01	ATARAZANA
09D05C02	SAUCES	09D05C03	GUAYACANES
09D05C04	ALBORADA	09D05C05	URDESA
09D05C06	KENNEDY	09D06C01	BELLAVISTA
09D06C02	MARTHA ROLDOS	09D06C03	SAMANES
09D06C04	JUAN MONTALVO	09D06C05	FLORIDA
09D06C06	MAPASINGUE	09D07C01	MONTE BELLO
09D07C02	BASTIÓN POPULAR 2	09D07C03	BASTIÓN POPULAR 1
09D07C04	MUCHO LOTE	09D07C05	ORQUÍDEAS
09D07C06	PASCUALES	09D08C01	NUEVA PROSPERINA
09D08C02	FORTÍN	09D08C03	PARAÍSO DE LA FLOR
09D08C04	FLOR DE BASTIÓN	09D08C05	MONTE SINAÍ
09D08C06	NUEVO GUAYAQUIL	09D09C01	LOS CEIBOS

**Cuadro 1.1:** División territorial por circuitos de Guayaquil, Fuente: SENPLADES,  
Elaboración: CIREV

### 1.3. Justificación

En el país, la transmisión del virus del dengue se convirtió en un problema de salud pública por la dispersión del vector y la recurrencia de los ciclos epidémicos. Actualmente se ha registrado el desplazamiento del dengue a zonas donde no era común el hallazgo de casos de contagio [13], lo que motiva a la identificación de localidades vulnerables, teniendo en cuenta que la medida más recurrente es la prevención a través del control vectorial, la asistencia sanitaria, el mejoramiento de las condiciones de vida y la adecuada información para los habitantes de los lugares de mayor atención [17].

En adelante, usaremos algunos conceptos de la teoría bayesiana para encontrar un modelo espacio-temporal que nos permita construir mapas de riesgo con la ubi-

caci n de lugares susceptibles. En comparaci n con el an lisis espacial de la fiebre del dengue realizado en Guayaquil [10], trataremos de implementar estructuras de simulaci n a trav s de factores clim ticos que se convertir n en elementos de temporalidad variable, conjuntamente, con la inclusi n de indicadores sociales que provengan una idea visual pr ctica para la toma de decisiones.

Con este planteamiento buscamos aplicar estos modelos al seguimiento de las enfermedades vectoriales, pues en diversos lugares del pa s se reportan nuevos contagios relacionados a la aparici n de los virus como Chikungunya y Zika originados por *Aedes aegypti*.

En conjunto, estas enfermedades vectoriales causan impacto econ mico relacionado con los costos por prevenci n, control, vigilancia, morbilidad, tratamiento y hospitalizaci n, los mismos que se ven reflejados en toda la gesti n de implementaci n de decisiones de intervenci n para el monitoreo de criaderos, la divulgaci n de medidas que involucren activamente a la comunidad, la atenci n oportuna en las ´reas de salud o la dotaci n de recursos necesarios a los laboratorios con el fin de integrar una red de diagn stico veraz.

## 1.4. Objetivos

### 1.4.1. Objetivos Generales

- Ajustar un modelo jer rquico bayesiano espacio temporal que permita identificar las variables de mayor influencia en la incidencia y contagio del dengue en la ciudad de Guayaquil.

### 1.4.2. Objetivos Espec ficos

- Agrupar las series de datos en estudio para adecuarles a las temporalidades lluviosa y seca.
- Ajustar los diferentes modelos jer rquicos que expliquen la enfermedad, basados en el comportamiento espacio-temporal.
- Seleccionar el modelo que se ajuste a los criterios de validaci n propuestos posteriormente.

- Establecer mapas de riesgo específicos a partir de los diferentes escenarios encontrados.

## 1.5. Hipótesis

Es posible encontrar un modelo adecuado desagregado a nivel mensual y por circuitos, que revele tanto los factores asociados a la incidencia como al contagio del dengue, usando información relacionada con variables vinculadas a este hecho, que pueden ser tomadas de la bibliografía propuesta [34], [36], [10]. Posteriormente, a través de los ajustes que se sugieren en los capítulos próximos se puede elegir un modelo bayesiano que se apegue a nuestros objetivos para armar mapas de riesgo que señalen las zonas más importantes del área en estudio, las mismas que serán puntos focales y sobretodo los candidatos para implementar las medidas de control.

## 1.6. Metodología

En una visión global de este proyecto de investigación se tomará como región de estudio a la ciudad de Guayaquil, sectorizada en 48 circuitos durante los meses de enero a diciembre del período 2010-2015. A continuación se detallan las fases que se llevarán a cabo.

**Fase 1:** Inicia con un estudio exploratorio para seleccionar las variables contundentes que tengan un aporte significativo o hayan sido el resultado de investigaciones anteriores, como datos procedentes de información climática, socioeconómica y epidemiológica, proporcionadas por el Instituto Nacional de Investigación en Salud Pública. [21]

**Fase 2:** En la depuración y relleno de datos se propondrá un control de calidad asociado a todas las observaciones. Se aplicarán ciertas técnicas para los datos faltantes según el caso requiera, respetando las tendencias de cada punto de estudio para organizar la información por circuitos y a nivel mensual.

**Fase 3:** En la organización de los datos, el paso siguiente será sintetizarlos evitando la redundancia mediante el análisis de componentes principales en el caso de los indicadores sociales y económicos.

**Fase 4:** En la estructura del modelo se usará una distribución de Poisson para modelar el número de casos de dengue a través de la información conocida, los casos esperados y la población de cada circuito. Además se dotará al riesgo relativo en cada área de una estructura formada con las variables climáticas y las

nuevas variables explicativas procedentes del análisis de componentes principales, a través de la función enlace (logaritmo) que se detalla en el capítulo de construcción de la base de datos.

**Fase 5:** En el monitoreo de los modelos se usa el programa WinBUGS14 ayudados de los criterios de devianza DIC y pérdida predictiva a posteriori para seleccionar el que presente un mejor ajuste.

**Fase 6:** Finalmente, se presentarán mapas de riesgo para cada año en los que se pinten los circuitos de acuerdo a la clasificación del riesgo en alto, medio y bajo, desagregándolos por época lluviosa y seca en todo el período 2010 al 2015.

## 1.7. Alcances

A partir del modelo seleccionado se espera:

- Identificar las variables de mayor influencia en la incidencia y contagio del dengue.
- Reconocer la importancia del componente de heterogeneidad espacial y la existencia de algún patrón especial relacionado con la época lluviosa o seca.
- Clasificar cada uno de los circuitos de la ciudad de Guayaquil, de acuerdo a los niveles de riesgo establecidos para agrupar las zonas de mayor atención.

## 1.8. Estructura del Trabajo

En el capítulo II se describen las definiciones que apoyan la inclusión de la teoría bayesiana, la elección de las distribuciones a priori, los métodos para incluir el uso de las Cadenas de Markov, la importancia de añadir los componentes de heterogeneidad espacial y la influencia del modelo condicional autorregresivo. Exponemos también los criterios de comparación para el ajuste de los modelos seleccionados y para el estudio de los residuos del modelo final.

El capítulo III hace una descripción detallada de la construcción de la base de datos, la metodología para la interpolación de los tamaños poblacionales y la elección de los indicadores climáticos y sociales. Aquí también haremos mención del uso del software WinBUGS14 y los diferentes criterios para probar la convergencia de las Cadenas de Markov generadas para el modelo final ajustado y el proceso para la

elaboración de los mapas con los resultados tanto para la época lluviosa como para la seca, en el período 2010 al 2015.

En el capítulo IV exponemos la elección de las distribuciones a priori para cada uno de los parámetros del modelo y el cálculo de sus distribuciones condicionales a posteriori.

En el capítulo V se contempla la descripción de todos los modelos monitoreados en el programa WinBUGS14 y la selección y ajuste del modelo que se usará para los fines de este proyecto. Además se exhibe la convergencia para todos los parámetros, las tablas y mapas con el riesgo relativo para cada circuito y finalmente el análisis de residuos correspondiente.

# Capítulo 2

## Marco Teórico

En esta sección mencionamos los conceptos que envuelven el uso de la teoría bayesiana, la influencia de las distribuciones a priori reflejadas en los resultados obtenidos cuando calculemos las distribuciones a posteriori. Se describirán también los algoritmos de Metropolis-Hastings y Gibbs para la generación de las cadenas de Markov y el comportamiento aproximado de la convergencia de las mismas. A continuación se hace la comparación de modelos a través de los criterios de devianza y pérdida predictiva a posteriori para la elección de aquel que se ajuste a los objetivos del proyecto.

Los modelos estadísticos paramétricos generalmente se basan en las relaciones entre la variable de interés o de respuesta y ciertas variables explicativas que ayudan a comprender el comportamiento de un fenómeno de estudio, sin embargo, en ocasiones se pone de manifiesto la necesidad de implementar diseños experimentales complejos en los cuales las variables de respuesta pueden ser: recuentos, proporciones, variables de escala, entre otras, que claramente indican que los errores no se apegan a los supuestos de normalidad.

Este tipo de limitaciones dan lugar a los Modelos Lineales Generalizados (MLG) caracterizados porque sus componentes tienen una parte aleatoria estocástica y otra sistemática (o predictor lineal), junto con una función de enlace. En esta configuración el componente estocástico se relaciona con la variable respuesta, en tanto, el sistemático es una estructura que contiene a las variables explicatorias, tal como se observa en los modelos de regresión normal, donde usualmente esta composición es vista como una combinación lineal de dichas variables. Además se tiene a la función de enlace que es el nexo entre la variable respuesta, el predictor lineal y el resto de variables [39].

Por otro lado, está la opción para la reducción de restricciones sobre los residuos, manteniendo la normalidad y que adicionalmente contempla errores no independientes o heterogéneos, es decir, existe una ventana amplia para organizar la variancia de los residuos y obtener un Modelo Lineal Generalizado Mixto (MLGM) adoptando las propiedades de las dos propuestas.

Una de las mayores ventajas del enfoque bayesiano es la incorporación de la información adicional procedente de estudios anteriores, pruebas pilotos, opiniones de investigadores o de la intuición. Sin embargo, este aporte toma un sentido de controversia debido al proceso de transferencia de la información hacia una distribución matemática formal que se incluye en el modelo. Este paso se puede ver complicado en el sentido del coste que supone la subjetividad. En la práctica esto se soluciona con el uso de distribuciones a priori o no informativas que son aquellas que reflejan un nivel de conocimiento débil y que especialmente en muestras grandes aportan los mismos resultados que con los métodos frecuentistas. Se debe enfatizar que estas distribuciones son consideradas como un punto de referencia inicial sin limitar el potencial que pueden proporcionar otras fuentes de información [1].

En el marco de la toma de decisiones mientras que los métodos estadísticos convencionales se basan únicamente en la información dada por los datos y procesados por el decisor, el enfoque bayesiano permite incorporar al investigador antecedentes adicionales provenientes de otras fuentes para que el análisis gane robustez.

Por otro lado, el desarrollo de las cadenas de Markov ha permitido la estimación de modelos complejos que hasta entonces habían sido difíciles de resolver computacionalmente. En este mismo contexto estos procesos se han usado en estudios de tasas de mortalidad y prevalencia de enfermedades en pequeñas áreas ([19], [5]), en la incorporación de la heterogeneidad espacial en el estudio de mapas de enfermedades [29], la estimación de sistemas de información geográfica (GIS) [38], entre otras.

La implementación de los códigos de estos procesos ha sido posible con el desarrollo del programa WinBUGS14 que es de libre distribución que permiten aplicar desarrollos bayesianos a un coste de aprendizaje razonable, aunque la ganancia en sencillez lleva a una pérdida de control en el proceso de estimación lo que podría exponer al investigador a obtener resultados de sus estimaciones que no cumplen con los requisitos básicos como la ergodicidad de la cadena usada, la convergencia

de la misma e incluso la existencia de la distribución a priori [1].

A pesar de los aspectos positivos y las desventajas de incluir el enfoque bayesiano en el área de la salud, se pueden tener en cuenta las siguientes sugerencias a la hora de tomar decisiones:

- El análisis bayesiano debería comenzar con el uso de distribuciones de probabilidad a priori vagas o no informativas, ya que eliminan la subjetividad en el análisis y en muestras grandes los resultados son equivalentes a los proporcionados si es que se usara el enfoque frecuentista [2].
- La presentación de un análisis de sensibilidad que incluyan las diferentes distribuciones informativas a priori utilizadas junto con la variación que producen en los resultados [2].
- La contastación de la convergencia de las cadenas de Markov porque aunque el programa WinBUGS14 ha permitido la estimación sencilla de las mismas a través del muestreo de Gibbs, puede resultar conveniente repetir los procesos en el ambiente estadístico R con la función CODA que tiene inmersos los diferentes tests de convergencia, evitando resultados ineficientes y altamente sesgados [1].

## 2.1. Nociones Básicas

En esta sección describimos los conceptos básicos ([31], [6], [33]) en los que se apoya la teoría bayesiana y que conforman el marco teórico de nuestro estudio. Adicionalmente notaremos que siempre se tiene un conocimiento previo, por lo que se puede asumir que las probabilidades son condicionales.

**Probabilidad Condicional:** Dado un espacio de probabilidad  $(\Omega, \mathcal{F}, \mathbb{P})$  y los eventos  $A$  y  $B \in \mathcal{F}$  con  $Pr(B) > 0$ , la probabilidad condicional de  $A$  dado  $B$ , que se puede aplicar tanto a variables continuas como discretas es:

$$Pr(A|B) = \frac{Pr(A \cup B)}{Pr(B)} \quad (2.1)$$

**Probabilidad Total:** Dado el evento  $B$  y la partición de eventos  $A_1, \dots, A_n$  en el espacio muestral  $\Omega$ , se tiene:

$$P(B) = \sum_{i=1}^n Pr(B|A_i)Pr(A_i) \quad (2.2)$$

**Función de Verosimilitud:** Sean  $x_1, \dots, x_n$  una muestra aleatoria de la variable  $X$ , la función de verosimilitud se define como:

$$l(\theta|\mathbf{X}) = \prod_{i=1}^n Pr(x_i|\theta) \quad (2.3)$$

**Variable de Poisson:** Una variable aleatoria discreta se dice que sigue una distribución de Poisson de media  $\lambda > 0$  cuando la probabilidad de que tome un valor particular  $x = 0, 1, \dots$  se expresa mediante:

$$Pr(X = x) = p(x|\lambda) = \frac{\lambda^x}{x!} \exp(-\lambda), \quad x = 0, 1, \dots \quad (2.4)$$

donde  $E(X) = \lambda$ ,  $Var(X) = \lambda$ .

En los métodos bayesianos se debe tener en mente que la verosimilitud (2.3) no es lo más importante como sucede en el método clásico donde es el principal elemento en el que se basa la inferencia. En cambio, en este enfoque bayesiano la función de verosimilitud sirve como un instrumento para comunicar la distribución a priori con la posteriori.

La parte fundamental de esta teoría bayesiana toma como punto inicial el conocimiento a priori que posee el investigador, influenciado por estudios precedentes, reflexiones y experiencias resultantes los cuales son producto de un seguimiento previo del fenómeno en análisis. En adelante trataremos que los modelos transcriban las creencias hacia una aproximación de la realidad, de manera que al final tengamos una idea posterior de los parámetros que determinan nuestro estudio, para ello que introducimos el Teorema de Bayes [14].

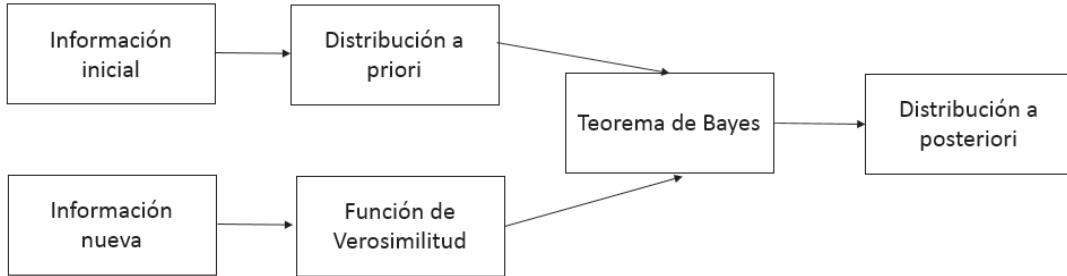
**Teorema de Bayes:** Sean  $A_1, A_2, \dots, A_n$  un conjunto de sucesos mutuamente excluyentes y de probabilidad no nula, tales que  $\bigcup_{i=1}^n A_i = \Omega$ .

Si  $B$  es un suceso en el espacio muestral  $\Omega$  de probabilidad no nula del que se conocen las probabilidades condicionales  $Pr(B|A_i)$ , entonces:

$$Pr(A_i|B) = \frac{Pr(B|A_i)Pr(A_i)}{\sum_{i=1}^n Pr(B|A_i)P(A_i)} \quad (2.5)$$

Mediante el teorema (2.5), la estadística bayesiana toma los datos generados de una

muestra (por medio de la función de verosimilitud), luego combina el juicio de criterios anteriores expresados a través de las distribuciones a priori y exhibe como resultado principal la distribución a posteriori. A través del esquema (2.1) observamos intuitivamente como suceden las conexiones de esta última descripción.



**Figura 2.1:** Esquema intuitivo del comportamiento de las distribuciones a priori y posteriori

Con esta breve descripción, consideraremos a  $\mathbf{X} = (x_1, x_2, \dots, x_n)^T$  un vector de  $n$  observaciones cuya distribución de probabilidad es  $Pr(x|\theta)$  que depende de  $k$  parámetros reunidos en el vector  $\theta = (\theta_1, \dots, \theta_k)$ . Entonces la distribución conjunta de  $\theta$  y  $\mathbf{X}$  es:

$$Pr(x, \theta) = Pr(x|\theta)Pr(\theta) = Pr(\theta|x)Pr(x) \quad (2.6)$$

de donde obtenemos la distribución de probabilidad condicional de  $\theta$  dado el vector de observaciones  $\mathbf{X}$ , con  $Pr(x) \neq 0$ :

$$Pr(\theta|x) = \frac{Pr(x|\theta)Pr(\theta)}{Pr(x)} \quad (2.7)$$

En la ecuación (2.7),  $Pr(x)$  es la distribución de probabilidad marginal de  $\mathbf{X}$ , que se puede expresar como [14]:

$$Pr(x) = \begin{cases} \int_{-\infty}^{\infty} Pr(x|\theta)Pr(\theta)d\theta, & \text{si } \theta \text{ es continuo} \\ \sum_{i=1}^n Pr(x|\theta_i)Pr(\theta_i), & \text{si } \theta \text{ es discreto} \end{cases}$$

donde la integral o la suma es tomada sobre el espacio paramétrico de  $\theta$ , de esta forma el Teorema de Bayes puede ser escrito como:

$$Pr(\theta|x) = cPr(x|\theta)Pr(\theta) \propto Pr(x|\theta)Pr(\theta) \quad (2.8)$$

En (2.8) los términos anteriores se denominan como:

- $Pr(\theta)$ : Es la distribución a priori de  $\theta$  que hace mención a lo que conocemos

previamente acerca del vector  $\theta$

- $Pr(\theta|x)$ : Representa la distribución a posteriori de  $\theta$  dado  $X$  que es lo que conocemos después de recolectar los datos.
- $c$ : Es una constante necesaria para que la distribución a posteriori se sume o se integre según corresponda.

Con las nociones establecidas se esperaría inferir sobre el parámetro  $\theta$ , pero en este punto se considera incluir el criterio del investigador, dotado por conocimientos anteriores, de tal forma que la incertidumbre se vea plasmada por medio de la distribución a priori que juntamente con la verosimilitud permiten hallar la distribución a posteriori.

Formalmente, se define el vector de observaciones conocidas  $X$  y a  $Pr(X|\theta)$  como una función únicamente de  $\theta$ , identificada como la función de verosimilitud que se denotará por  $l(\theta|x)$ , a través de la ecuación (2.9) :

$$Pr(\theta|x) \propto l(\theta|x)Pr(\theta) \quad (2.9)$$

A continuación planteamos un ejemplo para entender la interacción de estos conceptos.

*Ejemplo* Sea el parámetro  $\theta$  con una distribución a priori que corresponde a una distribución uniforme en el intervalo  $[0, 1]$  y la variable aleatoria  $Y$  con una distribución de probabilidad binomial con parámetros  $m$  y  $\theta$ , con  $m$  conocido. Entonces se tienen las siguientes funciones de distribución [12]:

$$Pr(\theta) = 1 \quad 0 \leq \theta \leq 1$$

$$Pr(y|\theta) = \binom{m}{y} \theta^y (1-\theta)^{m-y} \quad y = 0, 1, \dots, m$$

Para una muestra de tamaño aleatorio  $n$  la función de verosimilitud está dada por:

$$l(\theta|y) = \prod_{i=1}^n \binom{m}{y_i} \theta^{\sum_{i=1}^n y_i} (1-\theta)^{mn - \sum_{i=1}^n y_i} \quad y_i = 0, 1, \dots, m$$

Aplicando el Teorema de Bayes, la distribución a posteriori de  $\theta$  es:

$$Pr(\theta|y) = c \frac{nm!}{\prod_{i=1}^n y_i! \prod_{i=1}^n (m-y_i)!} \theta^{\sum_{i=1}^n y_i} (1-\theta)^{mn - \sum_{i=1}^n y_i}$$

que tiene la distribución Beta con parámetros  $(\sum_{i=1}^n y_i + 1)$  y  $(nm - \sum_{i=1}^n y_i + 1)$  y además la constante normalizadora se escribe:

$$c = \frac{\Gamma(nm + 2)}{\Gamma(\sum_{i=1}^n y_i + 1)\Gamma(nm - \sum_{i=1}^n y_i + 1)} \frac{\prod_{i=1}^n y_i! \prod_{i=1}^n (m - y_i)!}{nm!}$$

### 2.1.1. Distribución a Priori

La distribución a priori desempeña un rol importante en el análisis bayesiano pues brinda una visión inicial de los parámetros en estudio, pese a que con selecciones diferentes de prioris se generan cambios en la distribución a posteriori. Sin embargo, dado que todo indicio se puede traducir en un aporte a priori, tendremos pendiente que al evaluar algunas de ellas se tendrán tantos estados iniciales como prioris se hayan elegido, con lo que los resultados se verán estrechamente ligados a las convicciones e ideologías del investigador [48].

Para la selección de estas distribuciones conocemos tres puntos de vista diferentes: el primero que asume la falta de información de los parámetros de tal manera que se deba elegir alguna que no influya sobre los posibles parámetros en cuestión, el segundo criterio es escéptico, exponiendo una intuición subestimada de varias creencias y la tercera opción que parte de razones fundamentales para encontrar patrones, lo que enmascara la posible subjetividad del autor y la incertidumbre sobre su verdadero valor.

En el camino al uso de las distribuciones a priori el mayor obstáculo es la elección apropiada de las formas a priori, por lo que la alternativa más simple es dotar de información con poco contenido y en especial apoyar la selección de las mismas con la bibliografía.

## 2.2. Métodos Monte Carlo de Cadenas de Markov (MCMC)

En el análisis bayesiano el motor principal es la elección de las densidades a priori, el cálculo de las integrales que las definen y que son aquellas que aumentan la complejidad computacional en los procedimientos.

Previo a la introducción de los métodos MCMC, definimos a una *Cadena de Markov* como la secuencia de estados de un proceso estocástico los cuales dependen de las probabilidades de transición.

Una *Cadena de Markov* se define como  $\{X^0, X^1, X^2, \dots\}$  la secuencia de variables aleatorias tal que en todo momento  $t \geq 0$ , el siguiente estado  $X^{t+1}$  es muestreado de la distribución  $P(X^{t+1}|X^t)$ :

$$P(X^{t+1}|X^t) = P(X^{t+1}|X^t, X^{t-1}, \dots, X^1) \quad (2.10)$$

Es decir, la ecuación (2.10) depende sólo del estado actual de la cadena ( $X^{(t)}$ ) y no de la historia de la cadena  $\{X^0, X^1, X^2, \dots\}$ .

A este tipo de secuencia se le asigna el nombre de Cadena de Markov y  $P(\cdot|.)$  se conoce como el kernel de transición de la cadena y además se supone que es independiente de  $t$  [43].

Tendremos muy en cuenta que cada punto en la cadena va a tener correlación con el punto siguiente, con ello la idea intuitiva es ejecutar varias repeticiones que proporcionen una visión del comportamiento del sistema en conjunto, ayudados de la generación de números aleatorios [45].

Un Método Monte Carlo de Cadenas de Markov (MCMC), para la simulación de una distribución  $f$  es cualquier método que produce una Cadena de Markov ergódica ( $X^{(t)}$ ) cuya distribución estacionaria es  $f$ , y para la cual existe un algoritmo (eficiente) de simulación. [40]

La base principal de los algoritmos MCMC es que dado un valor inicial cualquiera  $X^0$ , la cadena ( $X^{(t)}$ ) se genera usando un kernel de transición con distribución estacionaria, digamos  $f$ , que garantiza la convergencia en distribución de ( $X^{(t)}$ ) a  $f$ . Dado que la cadena es ergódica, el valor inicial en principio no es importante.

Los métodos *Monte Carlo de Cadenas de Markov* (MCMC) son procedimientos de simulación usados en la generación de muestras de las distribuciones a posteriori que tienen como finalidad la estimación de cantidades que resultan de interés en los eventos siguientes. Esta técnica simula valores sucesivos de una densidad dada, donde cada uno de los valores generados dependen únicamente de los predecesores, resumiendo su sintaxis en las ideas siguientes.

Dada una distribución conocida  $Pr(\theta)$  se trata de generar una o múltiples realizaciones de una cadena de Markov con dicha distribución estacionaria. Las realizaciones ejecutadas servirán para obtener las aproximaciones a todas las características de  $Pr(\theta)$  que se deseen conocer.

Actualmente, en la teoría bayesiana este método es usado con éxito en paquetes computacionales, en especial, en WinBUGS14, en el que se implementarán los algoritmos de los modelos de las secciones venideras. En el procedimiento interno WinBUGS14 utiliza el Muestreador de Metropolis-Hastings y el de Gibbs.

### 2.2.1. Algoritmo Metropolis Hastings

El algoritmo de Metropolis-Hastings [14] es un método que crea una cadena de Markov convergente a una distribución límite igual a la distribución a posteriori de los parámetros de interés.

El algoritmo comienza con una distribución objetivo  $f(\theta) = Pr(\theta|x) \propto l(\theta|x)Pr(\theta)$  y una distribución propuesta  $q(.|x)$ , que satisface que :  $\frac{f(\theta_1)}{f(\theta_2)}$  es conocido para cualquier  $\theta_1$  y  $\theta_2$ .

El algoritmo de Metropolis Hastings asociado con la distribución objetivo y la densidad  $q$  obtiene una Cadena de Markov  $\theta^{(t)}$ .

#### Algoritmo de Metropolis Hastings .

1. Iniciar un valor  $\theta^{(0)}$
2. Considerar una densidad  $q(.|\theta^{(j-1)})$  para generar  $\phi_t$
3. Evaluar la probabilidad:

$$\alpha(\theta^{(j-1)}, \phi_t) = \min \left\{ 1, \frac{Pr(\phi_t)q(\phi_t|\theta^{(j-1)})}{Pr(\theta^{(j-1)})q(\theta^{(j-1)}|\phi_t)} \right\}$$

4. Aceptar  $\phi_t$  con la probabilidad  $\alpha$  del paso 3. Si  $\phi_t$  se acepta, entonces  $\theta^{(j)} = \phi_t$ , caso contrario:

$$\theta^{(j)} = \theta^{(j-1)}$$

5. Continuar con 2.

Notaremos que la probabilidad  $\alpha(\theta^{(j-1)}, \phi_t)$  es la probabilidad de aceptación del algoritmo y se aceptan los valores  $\phi_t$  siempre que existe un incremento en el cociente  $\frac{Pr(\phi_t)}{q(\phi_t|\theta^{(j-1)})}$ , obviamente, comparado con el término anterior  $\frac{Pr(\theta^{(j-1})}{q(\theta^{(j-1)}|\phi_t)}$ .

La probabilidad  $\alpha(\theta^{(j-1)}, \phi_t)$  está definida sólo cuando  $Pr(\theta^{(j-1)}) > 0$ , sin embargo, la cadena comienza con un valor inicial  $\theta^{(0)}$  tal que  $Pr(\theta^{(0)}) > 0$  y  $Pr(\theta^{(j-1)}) > 0$  para cualquier valor  $t \in \mathbb{N}$ .

Notaremos que los valores de  $\phi_t$ , tales que  $Pr(\phi_t) = 0$  conducen a  $\alpha(\theta^{(j-1)}, \phi_t) = 0$  y en este caso el algoritmo rechaza estos valores.

### 2.2.2. Muestreador de Gibbs

El algoritmo siguiente permite el análisis de los modelos complejos mediante la descomposición y el muestreo de las distribuciones condicionales completas, en las cuales la aproximación de integrales son difíciles de calcular en forma cerrada.

Este proceso genera cadenas de Markov Monte Carlo donde la transición del origen de la distribución  $Pr(\theta)$  está dada por las distribuciones completas.

Supongamos  $\theta = (\theta_1, \dots, \theta_i, \dots, \theta_k)^T$  un vector aleatorio con  $i \geq 1$ , además asumiremos las densidades condicionales  $Pr_1, \dots, Pr_k$ , donde:

$$\theta_i \sim Pr_i(\theta_i | \theta_1, \theta_2, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$$

Entonces el algoritmo de Gibbs asociado a la transición de  $\theta^{(t)}$  y  $\theta^{(t+1)}$  se describe en la siguiente secuencia:

#### Algoritmo de Gibbs

1. Iniciar la iteración  $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$
2. Actualizar la j-ésima iteración generando:
 
$$\begin{aligned} \theta_1^{(j)} &\sim Pr(\theta_1 | \theta_2^{(j-1)}, \dots, \theta_k^{(j-1)}) \\ \theta_2^{(j)} &\sim Pr(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)}) \\ &\vdots \\ \theta_k^{(j)} &\sim Pr(\theta_k | \theta_1^{(j)}, \dots, \theta_{k-1}^{(j)}) \end{aligned}$$
3. Ir al paso 2 e iterar hasta la convergencia.

Cuando se llega a la convergencia de la cadena, los valores obtenidos para  $\theta_j$  se convierten en una muestra de la distribución  $Pr$  y diremos que la cadena converge

cuando esta se aproxima a una condición de equilibrio, es decir, cuando el número de iteraciones aumenta.

Un valor de la distribución de interés solamente se obtiene cuando el número de iteraciones de la cadena se aproxima al infinito, en la práctica no se puede conseguir este resultado, pero si se puede obtener un valor suficientemente grande que se podrá tomar como muestra final de la distribución ( $Pr$ ).

En el programa estadístico WinBUGS14 los autores [39] sugieren que en un comienzo se realicen 2000 iteraciones para observar el comportamiento aproximado de la convergencia de la cadena, además de no descuidar el criterio informal basado en las técnicas gráficas.

## 2.3. Modelo Condicional Autorregresivo

Los *Modelos Condicionales Autorregresivos* (CAR) están fundamentados en la distribución condicional de los términos del error espacial, es decir, asume una dependencia espacial en la cual las estimaciones del riesgo en un área determinada dependen del riesgo de las áreas vecinas.

El modelo propuesto por Besag en 1991, define el vector  $\nu = (\nu_1, \nu_2, \dots, \nu_s)$  donde el modelo CAR propone como densidad conjunta de  $\nu_i | \nu_j$  con  $i \neq j$  a la expresión:

$$\exp \left[ -\frac{\gamma}{2} \left( a_i \nu_i - \sum_{i \neq j} W_{ij} \nu_j \right)^2 \right] \quad (2.11)$$

En la ecuación (2.11),  $W$  es la matriz que captura la influencia de  $\nu_j$  en la esperanza de  $\nu_i$ . Sin embargo,  $W$  se puede escribir como una matriz de adyacencia:

$$W_{ij} = \begin{cases} 1 & \text{si } i \text{ es adyacente a } j \\ 0 & \text{caso contrario} \end{cases}. \quad (2.12)$$

Para completar la definición en (2.12), entenderemos que un circuito  $i$  es *adyacente* a un circuito  $j$  si ambas localidades comparten en común algún tramo de frontera y además se asignará al término  $a_i$  como el tamaño muestral del circuito  $i$ .

Por otro lado, tenemos la reformulación propuesta nuevamente por Besag, donde  $\nu$

es proporcional a la expresión :

$$\exp \left[ -\frac{\gamma}{2} \nu^T B \nu \right] \quad (2.13)$$

En (2.13) la matriz B tiene como entradas  $B_{ii} = a_i$ ,  $B_{ij} = -a_i W_{ij}$  y bajo estas condiciones se propone el conjunto de vecinos  $\delta_i$  y el número de vecinos adyacentes al circuito i como  $m_i$ , de tal forma que la matriz W en (2.11) se reescribe así:

$$W_{ij} = \begin{cases} \frac{1}{m_i} & \text{si } j \in \delta_i \\ 0 & \text{caso contrario} \end{cases}.$$

Con los supuestos  $a_i = m_i$  y la simetría en B, se cumple que  $\bar{\nu}_i = \frac{1}{m_i} \sum_{j \neq i} \nu_j$ , de forma que se llega al siguiente resultado:

$$\nu_i | \nu_j, i \neq j \sim N \left( \bar{\nu}_i, \frac{1}{\tau_b m_i} \right) \quad (2.14)$$

En la ecuación (2.14),  $\tau_b$  representa el parámetro de escala y se concluye también que B es singular, y por ende la densidad conjunta es impropia.

Como un resultado de interés se considera que los modelos CAR son adecuados computacionalmente pues al parámetro  $\gamma$  en (2.13) es suficiente asignarle una distribución a priori Gamma, con lo que resulta que el muestreador de Gibbs en este caso es una herramienta útil que proporciona buenas estimaciones [48]

## 2.4. Criterios para la comparación de Modelos

Para el proceso de ajuste de un modelo se aborda principalmente en dos aspectos como la complejidad y la capacidad predictiva de cada uno, pues mientras más variables y relaciones estén involucradas más concretas pueden ser las predicciones, aunque sean menos generalizables [42].

De forma general, entre un conjunto de modelos a elegir se necesita medir la capacidad explicativa de un modelo y la penalización del grado de complejidad del mismo, que es la esencia de los criterios de información: de Akaike (AIC), Bayesiano (BIC) y de Devianza (DIC).

De estos tres criterios los dos primeros (AIC, BIC) utilizan la máxima verosimilitud

litud como criterio de bondad de ajuste y el número de parámetros como medida de la complejidad, mientras el *DIC* trabaja con la función completa de los parámetros que se construyen a través de las cadenas de Markov con el método de Monte Carlo.

El *DIC* es una generalización de los dos criterios anteriores que incorpora una medida mucho más completa del ajuste y la complejidad (número y grado de interrelación entre parámetros) [42].

#### 2.4.1. Criterio de Información de Akaike (AIC)

El *AIC* mide la bondad de ajuste a partir de la máxima verosimilitud ( $L$ ) del modelo y la complejidad a partir del número de parámetros  $k$ .

$$AIC = 2k - \ln(L) \quad (2.15)$$

Dependiendo de la forma del modelo y del grado de dispersión, cuanto más alto sea el valor de máxima verosimilitud mejor estará ajustado el modelo a los datos. En tanto  $L$  depende de la distancia de cada dato al modelo como del número de ellos [42].

#### 2.4.2. Criterio de Información Bayesiano (BIC)

De forma similar al anterior criterio, el *BIC* está dado por la ecuación siguiente, donde nuevamente,  $k$  es el número de parámetro y  $L$  la verosimilitud.

$$BIC = k \ln(n) - 2 \ln(L) \quad (2.16)$$

Como resultado se tiene que el *BIC* seleccionará un modelo sencillo que haga predicciones a menor detalle, pues en la ecuación (2.16) la medida de complejidad incorpora tanto  $k$  como  $\ln(n)$ .

Cuantitativamente, dado que el procedimiento de Schwarz difiere del de Akaike solamente en que la dimensión es multiplicada por  $\ln(n)$ , el *BIC* se inclina más que el *AIC* hacia modelos de dimensión más baja cuando el número de observaciones es  $n \geq 8$  [42].

Por otro lado, el *BIC* trata de seleccionar el modelo correcto con máxima probabilidad a posteriori y puede demostrarse que es un criterio consistente, de manera que la probabilidad de seleccionar el modelo correcto tiende a 1 cuando aumenta

el tamaño muestral. El criterio *AIC* no pretende seleccionar el modelo correcto, ya que admite que el modelo verdadero puede no estar entre los estimados y trata de obtener aquel que proporcione mejores estimaciones entre los existentes, sin embargo, en muestras pequeñas o medianas, el *AIC* tiende a seleccionar modelos con más parámetros de los necesarios [42].

### 2.4.3. Criterio de Información de Devianza (DIC)

El Criterio de Información de la Devianza (DIC) está formulado como una generalización del análisis de varianza para los modelos lineales generalizados. La devianza nos brinda una idea de la variabilidad de los datos y su utilidad queda vinculada con la selección de los modelos, donde la distribución a posteriori se ha obtenido a través de las simulaciones de las Cadenas de Markov con el método de Monte Carlo, por lo que se trata de una aproximación asintótica sustentada en muestras grandes cuando la distribución simulada a posteriori es aproximadamente normal:

$$DIC = D(\theta) = -2 \log(l(x|\theta)) + 2 \log(h(x)) \quad (2.17)$$

En la ecuación (2.17),  $l(x|\theta)$  representa la función de verosimilitud ,  $X$  el vector de datos,  $\theta$  es el vector de parámetros en el modelo y  $h(x)$  la función de estandarización.

También se propone considerar el ajuste de los modelos obtenidos mediante la esperanza posterior de la devianza:

$$\bar{D} = E_{\theta|x}[D] \quad (2.18)$$

Incluyendo el término  $p_D$  como una medida de complejidad, tomamos la devianza esperada menos la devianza evaluada esperada a posteriori , como se muestra en (2.19) :

$$p_D = E_{\theta|x}[D] - D(E_{\theta|x}[\theta]) = \bar{D} - D(\bar{\theta}) \quad (2.19)$$

Entonces el criterio de Información de la devianza queda expresado como:

$$DIC = \bar{D} + p_D = 2\bar{D} - D(\bar{\theta})$$

Y gracias a las simulaciones MCMC de  $\theta^i$ , el cálculo computacionalmente se estima por (2.20) :

$$\widehat{E}_{\theta|x}[D(\theta)] = \sum_{i=1}^N \frac{D(\theta^i)}{N} \quad (2.20)$$

Es decir, tenemos que el criterio DIC es dos veces la media de la devianza de la media de los parámetros, una vez que se ha logrado la convergencia de la cadena.

En síntesis el DIC explica un enlace fuerte entre la calidad del ajuste y la elección de un número de parámetros no muy grande, pues podemos caer en el caso de un modelo sobre ajustado. En los modelos que nos ocupan en este trabajo, el criterio DIC será proporcionado por el programa WinBUGS14, al igual que la cantidad relacionada con  $p_D$ .

Por supuesto el *DIC* no está a salvo de observaciones, pues el sistema por el que penaliza el número de parámetros no equivale a una forma de validación cruzada como ocurre en el *AIC*, lo que lleva a seleccionar modelos sobre-parametrizados. La forma de corregir esto es multiplicar por un factor el número efectivo de parámetros, tal que el indicador converja como lo hace el *AIC* y además los métodos se implementan en WinBUGS14 proporcionan fácilmente la medida *DIC* [42].

Para la comparación de los diferentes modelos (que estarán expuestos en la etapa de validación) se toma en cuenta la precisión respecto a sus estimaciones o la correspondencia con los datos reales empleando los criterios : (*DIC*) *Criterio de Información de Devianza* y ( $p_D$ ) *Pérdida Predictiva a Posteriori* que son obtenidos en el programa WinBUGS14.

Cuando hay dependencia espacial los circuitos cercanos tienen valores con mayor similitud que los distantes y la ausencia de correlación espacial se refleja en el hecho de que la distancia entre áreas no tiene relación con sus valores.

## 2.5. Índice I de Morán

El índice I de Morán ( $M$ ) es un indicador espacial que mide de manera global la presencia o ausencia de la autocorrelación espacial en un conjunto de observaciones. Este índice manifiesta la existencia de aglomeraciones espaciales alrededor de valores superiores o inferiores al valor de la media del conjunto de observaciones, señalando las áreas que son de mayor aporte a la autocorrelación global espacial. Para hacer uso del índice se construye la matriz de vecindad que nuevamente se define como  $W$  a través de la siguiente expresión:

$$W_{ij} = \begin{cases} 1 & \text{si } i \text{ es adyacente a } j \\ 0 & \text{caso contrario} \end{cases}. \quad (2.21)$$

Cabe señalar que en todo el período de estudio elegido no ha cambiado la división limítrofe entre los circuitos, así que la matriz definida en (2.21) permanece constante para los cálculos posteriores.

Definiremos en la ecuación (2.22) al coeficiente I de Morán para un mes en particular como:

$$M = \frac{C \sum_{i=1}^C \sum_{j=1}^C (r_i - \bar{r})(r_j - \bar{r})}{\sum_{i=1}^C \sum_{j=1}^C W_{ij} \sum_{i=1}^C (r_i - \bar{r})^2} \quad (2.22)$$

donde C representa el número total de circuitos indexados por  $i$  y por  $j$ , además  $r$  es la variable de residuos correspondiente a  $i$  y  $j$ ,  $\bar{r}$  es el valor promedio de  $r$  y  $W_{ij}$  es un elemento de la matriz (2.21).

Los valores estimados por la ecuación (2.22) toman valores entre -1 y 1 que indican autocorrelación negativa (perfecta dispersión) cuando hay cercanía a -1 y positiva (perfecta concentración) si se acercan a 1. En el caso de existir un patrón espacial completamente aleatorio los valores son cero, que es lo que esperaríamos en nuestro interés de estudio, dicho de otra forma, este indicador plantea la hipótesis nula como la ausencia de la asociación espacial entre los residuos frente a la hipótesis alternativa de la presencia del patrón de asociación espacial entre los mismos [8].

En vista de que el modelo que presentamos involucra el aspecto temporal por meses y el espacial por circuitos, generalizamos el cálculo del índice de Morán I en el programa R-project a través de los pasos siguientes [48]:

- Construir una matriz  $M$  de dimensiones  $k \times T$  que contiene los residuos originales ( $r_{it}$ ) del modelo seleccionado.
- Aplicar la ecuación (2.22) a cada una de las columnas de la matriz  $M$ .
- En el sentido de la interpretación, el segundo paso retorna el índice de Morán, los valores observado y esperado, conjuntamente con la desviación estándar y el p-valor.
- Este último resultado del tercer paso es esencial para no rechazar la hipótesis nula cuando el p-valor es menor al 5 %. En este caso la distribución espacial de los valores altos y bajos en el conjunto de datos se encuentra más agrupada espacialmente.

## 2.6. Estudio de los Residuos

Un paso posterior a la etapa de validación del modelo conlleva analizar el comportamiento de los residuos que provee el modelo. Según Carlin y Louis [48], se puede expresar los residuos bayesianos para cada circuito y cada año elegido de la

forma:

$$r_{st} = y_{st} - E_{st} \left( \frac{1}{M} \sum_{m=1}^M E(y_{st} | \lambda_{st}^m) \right) \quad (2.23)$$

donde  $E(y_{st} | \lambda_{st})$  es el valor esperado obtenido de la distribución predictiva posteriori y  $\lambda_{st}^m$  se constituye como el conjunto de los valores de los parámetros obtenidos de la distribución a posteriori.

En nuestro caso, la ecuación (2.23) nos permite expresar los residuos recordando que  $y_{st}$  sigue una distribución de Poisson con parámetro  $\lambda_{st}$  que es la esperanza de la variable, entonces tenemos la expresión siguiente para los residuos:

$$r_{st} = y_{st} - E_{st} \left( \frac{1}{M} \sum_{m=1}^M \lambda_{st}^m \right)$$

En el capítulo siguiente plantearemos la estructura para la construcción de los datos que serán usados en la búsqueda de los indicadores que ejerzan mayor influencia en la incidencia del dengue, analizando la temporalidad de cada uno de ellos y la espacialidad que estará organizada a nivel de circuitos, mientras que los residuos serán evaluados a través del estadístico I de Morán.

# **Capítulo 3**

## **Construcción de la Base de Datos**

En este capítulo se presenta el proceso de obtención, procesamiento y depuración de los datos utilizados en los modelos sometidos a la convergencia y ajuste de sus parámetros.

Para lograr los propósitos de esta sección se obtuvo el apoyo incondicional del Centro de Investigación y Referencia de Vectores (CIREV) y en la entrega de datos la intervención de la Plataforma Integrada de Epidemiología, Geomática, Bioinformática y Bioestadística (EPISIG). Ambas dependencias pertenecientes al Instituto Nacional de Investigación en Salud Pública (INSPI), quienes colaboraron en la emisión de criterios e información pertinente.

### **3.1. Clasificación de los casos de dengue en el período 2010-2015**

En materia de transmisión de enfermedades es primordial analizar la temporalidad de los eventos en períodos de fácil manejo. Comúnmente se tiene a disposición el calendario epidemiológico con 52 semanas que en algunas ocasiones no coinciden con las semanas del calendario ordinario.

Las semanas epidemiológicas inician en domingo y terminan en sábado, tomando en cuenta que en el conteo de la primera semana del año se ubica el primer sábado de enero que incluya en los días precedentes cuatro o más días del mes de enero, aún cuando esa primera semana se inicie en diciembre. Por ejemplo, en la siguiente ilustración definimos la primera semana epidemiológica para el año 2016.

Semana	Mes	D	L	M	M	J	V	S	Mes
1	Diciembre	29	30	31	1	2	3	4	Enero

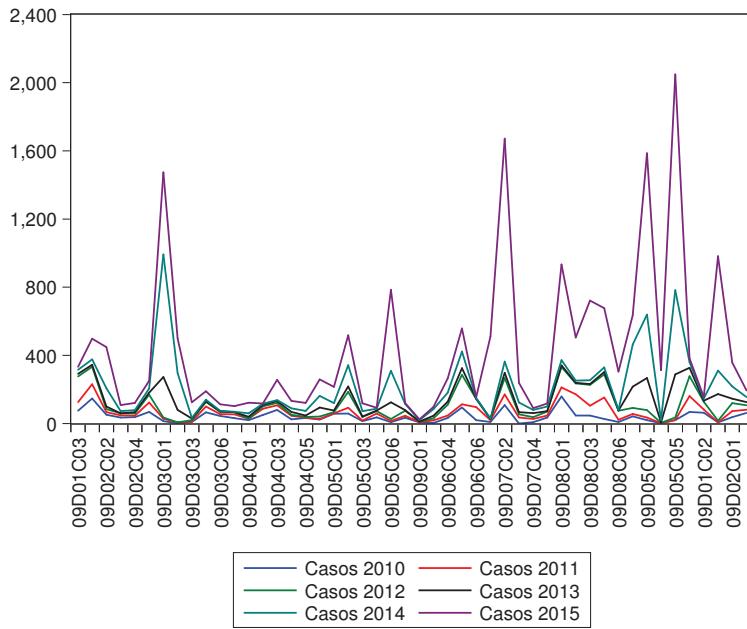
**Cuadro 3.1:** Ilustración de la primera semana epidemiológica del año 2016

Mencionamos estos conceptos puesto que los casos de dengue desde el 2013 al 2015 fueron colectados por semana epidemiológica, mientras que los valores correspondientes al período 2010-2012 fueron extraídos de forma anual.

En ambos lapsos de tiempo hubo pérdida de información, en especial en el segundo escenario en el que se encontró gran cantidad de ceros debido a que en los primeros meses del año 2010 no coinciden las fuentes de información para alimentar las bases de datos.

Se tomó la temporalidad mensual en todo el período de estudio, para lo cual fue efectivo ubicar las semanas epidemiológicas de cada mes que guardaron correspondencia al calendario epidemiológico de cada año para obtener el número de casos mensuales, en tanto, la información anual de los primeros años en cada circuito fue desagregada y estimada respecto del comportamiento de la misma localidad en años posteriores o tomando como referencia áreas vecinas o con características similares. Con esta reestructuración los nuevos datos faltantes fueron llenados con los valores promedio de las series respectivas, dado que era la única información oficial con la que se contaba.

Después de todas las estimaciones realizadas se desea visualizar el comportamiento de la enfermedad en cada unidad de estudio como en la figura (3.1). En los circuitos se observa un incremento en los casos de dengue a partir del año 2013, en especial un repunte hacia el 2015, dado que en este año coincide con una anomalía de calor. Este aumento surge desde el mes de abril a causa del impacto del fenómeno del Niño, que se considera como una posible explicación para esta situación [16].



**Figura 3.1:** Serie de casos de dengue por circuitos en la Ciudad de Guayaquil, en el período 2010-2015

Bastión Popular 2, Urdesa, Alborada, Centenario, Nueva Prosperina muestran altos valores para el número de casos y en las localidades: Kennedy y Chile, hay picos repetitivos considerables desde el 2013. Debido a la cercanía entre cada uno de los circuitos de la figura (3.1) es oportuna la inclusión de una componente espacial que explique la influencia de las localidades vecinas en un área específica, puesto que algunas de ellas comparten características similares en términos de calidad de vida o condiciones geográficas.

Es este marco se define la heterogeneidad espacial como uno de los efectos espaciales, aparte de la autocorrelación que guarda relación con la diferenciación espacial o regional de las unidades geográficas. A la vez se considera un concepto que define la ausencia de estabilidad en el espacio, tomando como precedente el comportamiento humano o de otras relaciones en estudio. Esto implica que, en los modelos espaciales, las formas funcionales y los parámetros variarán con la localización geográfica, no siendo homogéneos para toda la matriz de datos. [49]

### 3.2. Metodología para interpolar tamaños poblacionales.

Para organizar las poblaciones totales de cada año a nivel mensual y por circuitos usamos el método de interpolación con la finalidad de utilizar el resultado en los cálculos posteriores [44].

Para obtener una estimación en un tiempo  $t$  en el intervalo  $[0, T]$  aplicamos la técnica de la interpolación utilizando los datos de dos puntos observados (2013 y 2016). Entonces se divide el período en  $k$  subperíodos iguales, de forma que expresamos  $t$  como:  $\frac{T}{k}i$ , donde  $i$  nos indica el subperíodo desde  $1, \dots, k$ . Con esto se tiene que:

$$\begin{aligned} N(t) &= N(0) \exp^{\bar{r}t} \\ &= N(0) \exp^{\frac{(\bar{r}T)i}{k}} \end{aligned} \tag{3.1}$$

donde:  $N(t)$  representa la población en el tiempo  $t$ ,  $N(0)$  es la población inicial y  $\bar{r}$  es la tasa de variación media de la población en el periodo  $[0, T]$ , calculada con:

$$\bar{r} = \frac{1}{T} \ln \left( \frac{N(T)}{N(0)} \right) \tag{3.2}$$

### 3.3. Construcción de Variables

Las variables elegidas para el modelo bayesiano espacio temporal fueron tomadas del VII Censo de Población y VI de Vivienda realizado en el año 2010, las cuales fueron agrupadas para los indicadores de interés de acuerdo a las especificaciones explícitas en la sección *Anexos*.

De acuerdo con estudios anteriores [34], se revela que las poblaciones de *Aedes aegypti* están reguladas por el clima y las condiciones de pobreza, especialmente porque la reproducción de este vector está vinculada con la temperatura y la precipitación, que en este contexto tiene un efecto estadísticamente significativo sobre la variación mensual en la incidencia de esta enfermedad.

### 3.3.1. Indicadores sociales

La información recolectada para establecer los indicadores sociales toma en cuenta no sólo la bibliografía consultada sino también los criterios y sugerencias del área de investigadores del departamento CIREV, para tratar de condensar ciertos conceptos con la experiencia adquirida en el trabajo de campo en los distintos muestreos, en los que se recolecta información referente a vectores a nivel nacional y con especial atención en esta sectorización.

Para este primer análisis se tomaron en cuenta los siguientes indicadores, los mismos que están detallados en la sección *Anexos*.

- Tasa de analfabetismo (*anlf*).
- Razón persona-dormitorio (*hcnm*).
- Porcentaje de personas que habitan en viviendas en mal estado (*vstd*).
- Porcentaje de personas que habitan en viviendas sin acceso a la recolección de basura (*bsr*).
- Porcentaje de personas que habitan en viviendas sin acceso a alcantarillado (*alctr*).
- Porcentaje de personas que habitan en viviendas sin acceso a agua corriente (*agptbl*).
- Tasa de desempleo (*dsmpl*).

Este conjunto de indicadores sustentan el hecho de que la mayoría de problemas de servicios referentes a las viviendas son resultado de un proceso acelerado de urbanización y aglomeramiento. En el ambiente urbano las viviendas han debido modificarse por la densidad máxima de personas ocupando un cuarto de acuerdo al tamaño familiar y por el área de terreno que hacen que el hacinamiento sea considerado como un determinante en las condiciones de vida. En cuanto a los servicios de alcantarillado, agua potable, y servicios higiénicos es indispensable tener en mente que juntas forman parte del concepto de saneamiento ambiental de cualquier población. [22] [21]

A continuación, partiendo de un contraste de hipótesis buscaremos una pauta para decidir si es necesario reducir las variables a través de la técnica de *Componentes Principales* (ACP), ayudados de gráficos descriptivos que serán un apoyo visual

en este proceso.

Antes de aplicar un ACP verificamos si la correlación entre las variables de estudio es lo suficientemente fuerte, es decir, si esta relación es significativa para justificar la factorización de la matriz de coeficientes de correlación.

Partimos con el cálculo del test de esfericidad de Barlett que se expresa de la forma:

$$-\left[n - 1 - \frac{(2k + 5)}{6}\right] \ln |R| \sim \chi^2_{\frac{(k^2-k)}{2}} \quad (3.3)$$

Donde  $n$  es el tamaño de la muestra,  $k$  el número de variables y  $R$  la matriz de correlaciones.

	<b>agptbl</b>	<b>alctr</b>	<b>anlf</b>	<b>bsr</b>	<b>dsmpl</b>	<b>hcnm</b>	<b>vstd</b>
<b>agptbl</b>	1	0.8012	0.8217	0.8897	-0.2701	0.7845	0.6159
<b>alctr</b>	0.8012	1	0.7687	0.7181	-0.1617	0.7811	0.7345
<b>anlf</b>	0.8217	0.7687	1	0.8902	-0.2518	0.6627	0.5599
<b>bsr</b>	0.8897	0.7181	0.8902	1	-0.3433	0.6291	0.5021
<b>dsmpl</b>	-0.2701	-0.1617	-0.2518	-0.3433	1	0.1531	0.3936
<b>hcnm</b>	0.7845	0.7811	0.6627	0.6291	0.1531	1	0.8352
<b>vstd</b>	0.6159	0.7345	0.5599	0.5021	0.3936	0.8352	1

**Cuadro 3.2:** Matriz de correlaciones para las variables agptbl, alctr, anlf, bsr, dsmpl, vstd

Al aplicar la prueba de esfericidad de Barlett (3.3) el valor es de 356.664, con lo que aseguramos que existen variables correlacionadas debido a que el p-valor menor al 5 % provee esta significación perfecta. Algunos autores advierten que este test tiende a ser significativo cuando el tamaño muestral crece, por lo que debe ser aplicado cuando la razón  $n/k$  sea menor a 5, en nuestro caso el tamaño de  $n$  es grande por lo que es necesario calcular el índice de Kaiser-Meyer-Olkin (KMO), que para este fin el valor es igual a 0.779, superior a 0.75, por lo tanto también es adecuado basarse en el cálculo del determinante de la matriz de correlación, igual a 0.0002925648. Como esta cifra es muy cercana a cero, se verifica una fuerte redundancia en la información registrada en las variables, por lo tanto la reducción en componentes principales resultaría efectiva como muestran todas las pruebas realizadas.

Con el análisis de componentes principales pretendemos transformar el conjunto de datos inicial en otro conjunto de nuevas variables independientes que adquieren

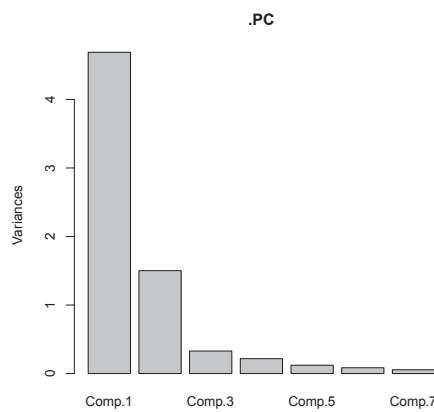
el nombre de componentes. Estas últimas serán combinaciones lineales de las anteriores y serán construidas de acuerdo al orden de importancia vinculado con la variabilidad total que recogen de la muestra, para lo cual usamos el paquete Rcmdr en el software R project y dado que las variables en estudio tienen diferentes escalas y medidas procederemos a la estandarización. Notaremos además que las variables originales al momento de escribir la ecuación fueron sustituidas por las variables estandarizadas y para la selección del número de componentes el criterio de Joliffe (1972,1973) propone reducir directamente las variables originales teniendo en cuenta a los componentes principales cuyos autovalores sean mayores que 0.7 en el caso de estandarización.

*Observación:* Las variables estarán estandarizadas cuando a sus valores se les resta la media y se dividen por su desviación estándar

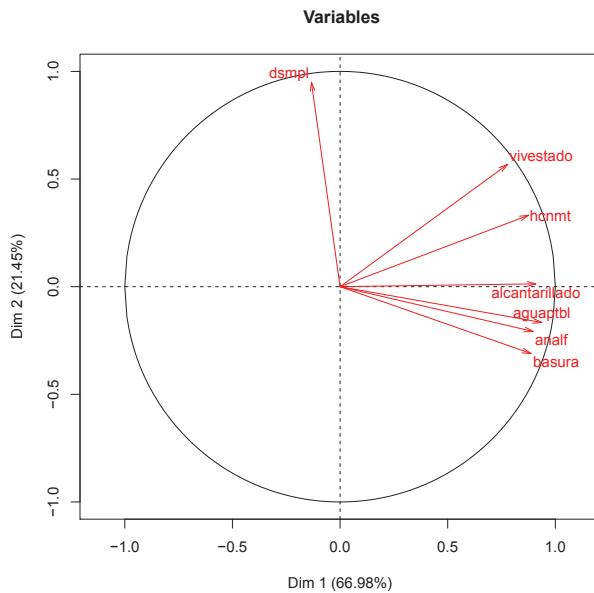
Del cuadro (3.3) y el gráfico (3.2) podemos ver que la elección adecuada encierra la toma de dos componentes principales, dado que hasta el segundo componente se agrupa el 88.44 % de la variación y también hasta ese punto los autovalores son mayores a uno.

	CP 1	CP2	CP3	CP4	CP5	CP6	CP7
Desviación estándar	2.1654	1.2255	0.5736	0.4662	0.3494	0.2915	0.2365
Proporción de la varianza	0.6698	0.2145	0.0470	0.0310	0.0174	0.0121	0.0079
Proporción acumulada	0.6698	0.8844	0.9313	0.9624	0.9798	0.9920	1

**Cuadro 3.3:** Importancia de los componentes principales, CP: Componente Principal



**Figura 3.2:** Proporción de variabilidad explicada para cada componente asociados los indicadores seleccionados



**Figura 3.3:** Círculo de correlaciones para las variables: agptbl, alcctr, anlf, bsr, dsmpl, vstd

Ahora con la información proporcionada en la tabla siguiente tenemos los coeficientes que acompañan a cada variable del grupo seleccionado, como se muestra.

Variable	CP1	CP2
agptbl	-0.4329	0.1356
alcctr	-0.4200	-0.0105
anlf	-0.4143	0.1697
bsr	-0.4098	0.2535
dsmpl	0.0613	-0.7746
hcmt	-0.4042	-0.2704
vstd	-0.3596	-0.4637

**Cuadro 3.4:** Coeficientes para CP1 y CP2, de acuerdo a cada variable.

En (3.4) los coeficientes de las ecuaciones para cada componente principal están explícitos y para la interpretación formal observamos los valores absolutos más altos en cada uno de ellos.

El primer componente acoge como variables principales al porcentaje de personas sin acceso a agua potable, a alcantarillado, a la recolección de basura o con analfabetismo, en tanto, el segundo componente se enfoca en el desempleo, el porcentaje

de personas que habitan en viviendas en mal estado y la razón persona-dormitorio (relativo al hacinamiento), que se ven expresados por las ecuaciones (3.4) y (3.5):

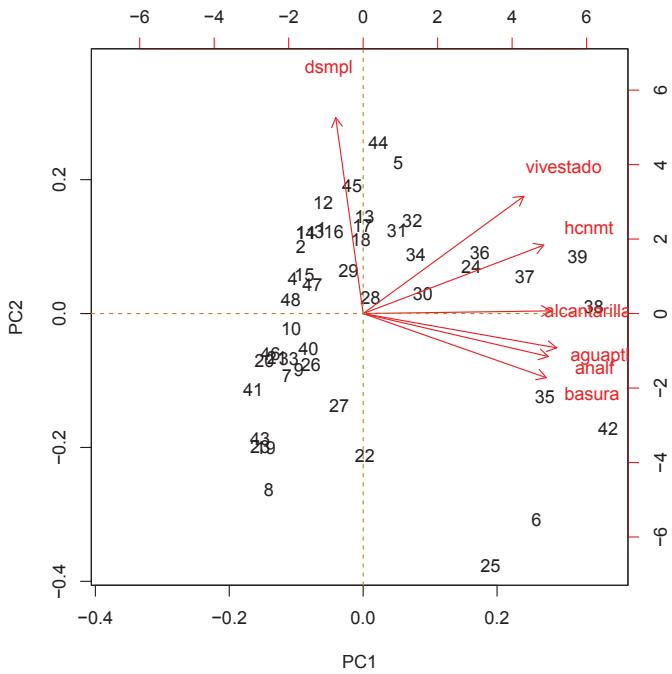
$$\begin{aligned} Z_1 = & -0,4329 \hat{agptbl} - 0,420 \hat{alctr} - 0,4143 \hat{anlf} - 0,4098 \hat{bsr} \\ & + 0,0613 \hat{dsmpl} - 0,4042 \hat{hcñmt} - 0,3596 \hat{vstd} \end{aligned} \quad (3.4)$$

Este primer componente acumula aproximadamente un 67 % de variación y los valores absolutos más elevados de sus coeficientes son los presentes en las variables descritas anteriormente, es decir, que dichas variables son las que contribuyen en gran medida a esa variación, y estas en nuestro caso están vinculadas con *las condiciones de salubridad*.

Para la segunda componente se agrupa un 21.45 % y los valores absolutos más altos corresponden a: desempleo, razón persona-dormitorio y porcentaje de personas que habitan viviendas en mal estado, que pueden asociarse con las *condiciones de pobreza*.

$$\begin{aligned} Z_2 = & 0,1356 \hat{agptbl} - 0,0105 \hat{alctr} + 0,1697 \hat{anlf} \\ & + 0,2535 \hat{bsr} - 0,7746 \hat{dsmpl} - 0,2704 \hat{hcñmt} - 0,4637 \hat{vstd} \end{aligned} \quad (3.5)$$

La utilidad del cálculo de las componentes principales permite determinar las puntuaciones (los nuevos valores que obtendrían las observaciones si pudiesen ser medidos directamente con cada componente) usando la matriz de datos. Los nuevos pares que se obtienen al reemplazar los valores en ambas componentes son proyecciones que se pueden graficar en un plano formado por las componentes  $Z_1$  y  $Z_2$ , donde se observa la mayor dispersión mediante el gráfico biplot [15].



En el gráfico anterior no se busca la proximidad entre las observaciones y las variables ya que las direcciones son más importantes, pero se concluye que la componente 2 positiva caracteriza las condiciones de vivienda relacionadas con el estado de la vivienda y la razón persona dormitorio, en tanto la componente 1 positiva guarda la conexión con las condiciones de salubridad, que en nuestro caso estas dos secciones muestran la insuficiencia de estos servicios.

En algunas aplicaciones el análisis de componentes principales constituye un objetivo en sí mismo y es objeto de interpretación, mientras en otras veces es simplemente una herramienta que permite reducir la dimensionalidad de un conjunto de datos que posteriormente puede ser objeto de un análisis estadístico posterior, como es la situación que nos interesa.

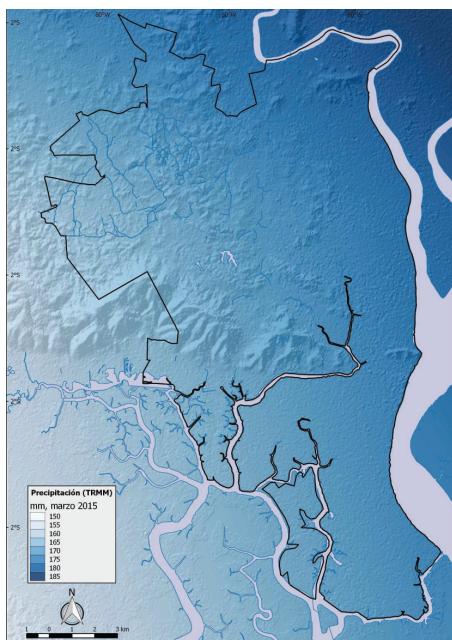
### 3.3.2. Indicadores climáticos

Para los datos relacionados con el clima tenemos las siguientes fuentes de extracción:

- **Precipitación mensual [(mm) milímetros]** : Las lecturas fueron a través de (*TRMM Tropical Rainfall Measuring Mission*), con conexión a la plataforma del EPISIG. La misión *TRMM* es una asociación conjunta entre la Administración

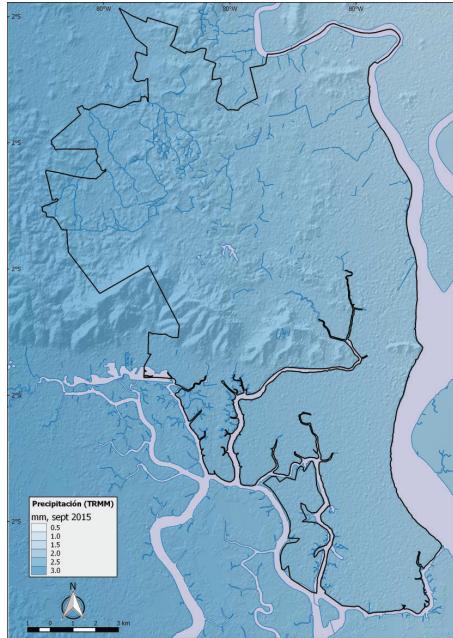
Nacional de la Aeronáutica y del Espacio (NASA) y la Agencia de Exploración Aeroespacial de Japón (JAXA) para estudiar las precipitaciones según el tiempo y el clima.

Esta plataforma de datos es un estándar espacial para la medición de la precipitación, y ha dado lugar a investigaciones que han mejorado la comprensión de la estructura de ciclones tropicales, las propiedades del sistema convectivo, las relaciones rayo-tormenta, el clima y los impactos en las precipitaciones.



**Figura 3.4:** Mapa de Precipitación, época lluviosa, Fuente:EPISIG-INSPI

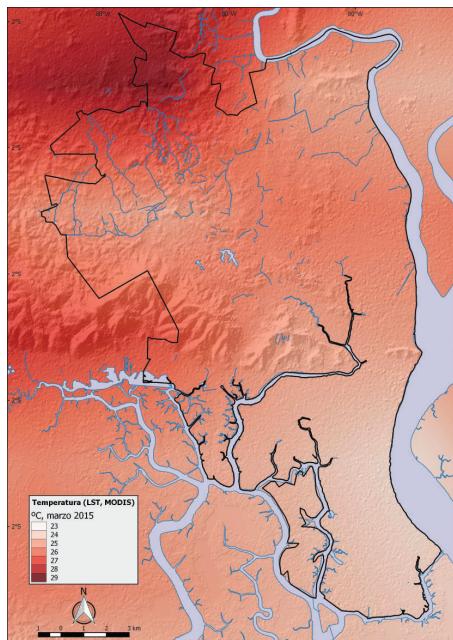
La escala de azul en los mapas de precipitación va de oscuro a claro correspondiendo a los lugares de mayor a menor precipitación. Se identifica que la zona noreste de la ciudad puede registrar valores más altos en la época lluviosa, mientras que en la época seca se marca con intensidad el relieve sureste.



**Figura 3.5:** Mapa de Precipitación, época seca, Fuente:EPISIG-INSPI

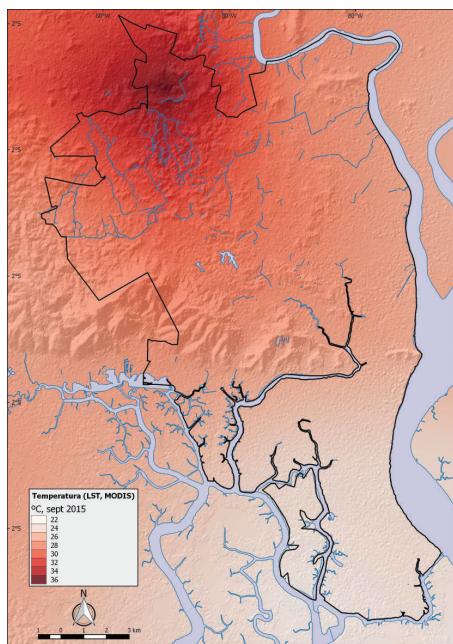
- **Temperatura mensual [( $^{\circ}$ C)]:** Las lecturas fueron tomadas de la fuente MODIS (*Moderate Resolution Imaging Spectroradiometer*). El instrumento MODIS opera a bordo de los satélites Terra y Aqua alcanzando una anchura de barrido de visualización de 2.330 kilómetros y abarcando la superficie de la Tierra cada uno o dos días.

Sus detectores miden 36 bandas espectrales entre 0.405 y 14.385 metros, abordando lecturas en tres resoluciones espaciales (250 m, 500 m y 1.000 m). Los productos de datos derivados de las observaciones de MODIS describen las características de la tierra, los océanos y la atmósfera que pueden ser utilizados para el estudio de los procesos y las tendencias en escalas locales a globales.



**Figura 3.6:** Mapa de Temperatura, época lluviosa, Fuente:EPISIG-INSPI

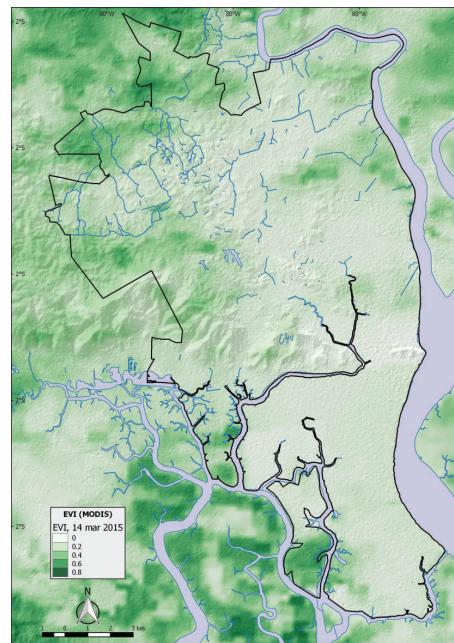
En los gráficos (3.6) y (3.7) la escala de color rojo va de tenue a intensa, donde la primera representa las zonas con mayor temperatura ubicadas en la región noroeste en ambas estaciones.



**Figura 3.7:** Mapa de Temperatura, época seca, Fuente:EPISIG-INSPI

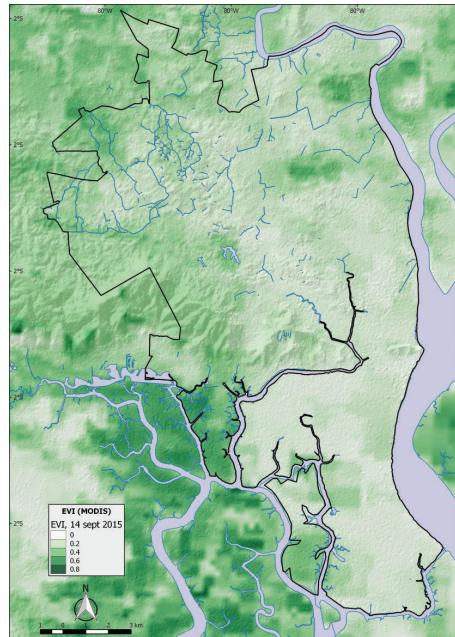
- **Índice de vegetación:** (*EVI*) Proporcionado por *MODIS*, exhibe la distribución temporal planteada para este estudio, con las observaciones de algunos vacíos

representados como valores negativos debido a píxeles incompletos. Aquí la periodicidad de la información es de 8 días, pero se transformarán las series para cada uno de los circuitos, a través del método de observaciones promedio ya que los faltantes fueron mínimos en cantidad.



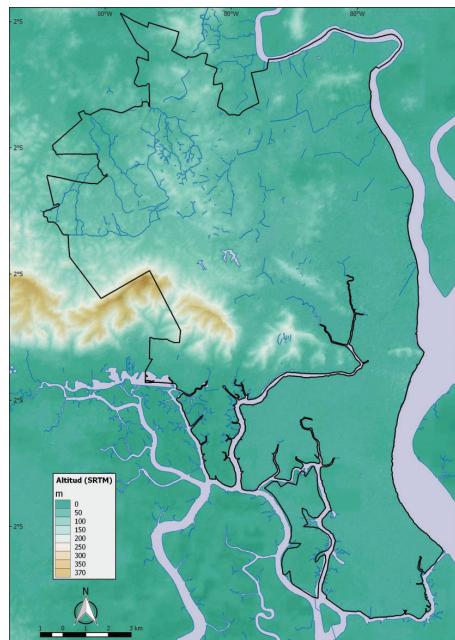
**Figura 3.8:** Mapa del Índice de vegetación, época lluviosa, Fuente:EPISIG-INSPI

De forma similar a las gráficas anteriores, en (3.8) y (3.9), la intensidad del color verde hace referencia a los lugares donde el índice de vegetación es saturado.



**Figura 3.9:** Mapa del Índice de vegetación, época seca, Fuente:EPISIG-INSPI

- **Altitud:** Es un valor fijo medido en metros (m) y considerado para nuestros planteamientos dado que el mosquito *Aedes aegypti* habita en regiones tropicales comprendidas entre los 45° de latitud norte y los 35° de latitud sur, en alturas que por lo general no superan los 1680 msnm, como la ciudad de Guayaquil que fue elegida para este estudio por estar en este rango.



**Figura 3.10:** Mapa de Altitud, Fuente: EPISIG-INSPI

En el desarrollo de este trabajo algunas series de tiempo relacionadas a la precipitación presentaron repeticiones que toman la forma de patrones. Para este hecho es conveniente retirar la estacionalidad de esta variable, es decir, haremos un ajuste que hace referencia al proceso de eliminación de los movimientos estacionales cíclicos de las series y la extracción de una posible tendencia subyacente en las mismas.

Una ventaja al desestacionalizar las series de datos es que al eliminar los factores implícitos ocasionados por situaciones ajenas a nuestro control se tiene un escenario más claro sobre el comportamiento del fenómeno en cuestión, además la comparación entre meses consecutivos y no consecutivos permite encontrar diferencias importantes porque en nuestro caso los valores de precipitación muestran cambios a lo largo del año y no solo en las comúnmente estaciones lluviosa y seca, dado que en nuestro país observamos en los datos algunas fluctuaciones interesantes cada 3 o 4 meses.

Las series desestacionalizadas permiten analizar la evolución de las series año tras año, puesto que algunas perturbaciones influyen en este proceso porque suceden simultáneamente. En nuestro caso el efecto calendario influye cuando las series mensuales obtenidas resultan de la agregación de datos diarios, sin embargo, no en todos los meses existen las mismas características por el número irregular de días. [25]

En el programa Eviews 9 tenemos la opción para eliminar la estacionalidad de las series con diferencias técnicas, sin embargo, elegimos la metodología Census X-Arima debido a su precisión y menor error medio de revisión para las series de precipitación, además que con este método no perdemos las primeras 12 observaciones del conjunto de datos. [28]

### 3.4. Uso de WinBUGS14

Una de las mayores dificultades en estadística bayesiana es la gran complejidad computacional presente en el cálculo de las cantidades a posteriori de interés. En los últimos años, mediante la simulación de Monte Carlo vía Cadenas de Markov se ha logrado salvar este problema.

El programa WinBUGS14 implementa las técnicas descritas en la sección anterior y permite el amplio uso para los usuarios de estadística bayesiana aplicada, de tal forma que diversas situaciones complejas intratables numéricamente son resueltas con rapidez y facilidad.

WinBUGS14 consiste en un sistema de funciones que permiten la especificación de modelos y de las distribuciones de probabilidad para las componentes aleatorias tanto de las observaciones como de los parámetros, de manera que el manejo resulta simple en comparación a la dificultad de lo planteado [46].

En la elección de las distribuciones a priori, podemos encontrar parámetros a los cuales se debe asignar distribuciones llamadas hiper-prioris, para lo cual debemos tener cuidado a la hora de elegir puntos iniciales, es decir, si contamos con datos numerosos entonces los datos dominarán el factor a priori y no será trascendental la elección de las distribuciones, en tanto una mayor distorsión de lo propuesto viene si los datos en estudio son escasos.

En el modelo que proponemos, que es a su vez una extensión del modelo de Besag y Clayton, para la parte que refleja la correlación espacial entre vecinos y el efecto condicional autoregresivo (CAR) se sugiere tomar las distribuciones hiper-prioris a través de la distribución gamma.

### 3.4.1. Análisis de Convergencia

Uno de los temas de interés relacionados con los algoritmos Metropolis Hastings y Gibbs es determinar la convergencia de los mismos, al igual que la asignación del número apropiado de iteraciones para alcanzarla. En la teoría se dice que el valor inicial de las cadenas es irrelevante puesto que ambos algoritmos convergen con cualquier valor inicial [40].

A pesar de esto, se encuentran inconvenientes para elegir un valor inicial correcto que garantice la rapidez de la convergencia, como es el caso del algoritmo de Metropolis-Hastings donde la velocidad de convergencia está fuertemente relacionada con la elección de la distribución instrumental.

Una solución para analizar la convergencia es a través de la exploración gráfica de cada uno de los parámetros que actúan en el modelo, con lo que se dice que la cadena alcanza el equilibrio cuando se estabiliza alrededor de un valor[40].

Dado que los algoritmos utilizados para el MCMC tienen como fundamento el uso de las cadenas de Markov, se deben monitorear los procesos para las pruebas de

convergencia que nos permitan tener garantía de valores convergentes y estables. Cuando existe la convergencia de la cadena se puede determinar la distribución objetivo posterior después de la llamada fase de calentamiento en el que monitoreamos el comportamiento de los parámetros.

Para llevar a cabo esta tarea es importante el análisis de ciertos indicios para la evaluación de la convergencia [48], por ejemplo:

- Una inspección visual de los gráficos de las cadenas generadas para cada parámetro son una forma intuitiva para este monitoreo, pues las diferentes trayectorias que el programa imprime muestran la existencia o no de la mezcla de las cadenas a medida que aumenta el número de iteraciones.
- Con los gráficos de autocorrelación podemos verificar si las muestras que se van generando con las iteraciones son independientes de los valores iniciales y a partir del histograma detectar la forma que toman las cadenas, es decir, estos gráficos deben casi desaparecer y con el análisis de la forma de la densidad a posteriori se espera suavidad de las curvas.
- En modelos con varios parámetros, no es común la evaluación de todos ellos, para lo cual se toma una selección aleatoria de parámetros y se comprueba la convergencia de este subconjunto.
- Mediante los gráficos de seguimiento de los valores de la muestra versus las iteraciones a través de la opción *trace plot* se busca evidencia de la estabilidad de la simulación y con el gráfico proporcionado por el comando *history plot* observamos la existencia de algún comportamiento aleatorio, caso contrario se acepta la convergencia de la cadena.
- En el caso de ejecutar varias cadenas a la vez, los gráficos anteriores toman diferentes colores para diferenciar cada muestra y aseguramos la convergencia de las cadenas si todos los gráficos están superpuestos.
- Un último criterio es el de Gelman y Rubin que se basa en el funcionamiento de las cadenas de valores paralelos a partir de valores iniciales dispersos. Se esperaría que el valor R que es una línea de color rojo en cada gráfico de cada parámetro sea mayor que 1 si los datos iniciales son adecuadamente dispersos, sin embargo, Brooks Y Gelman demuestran que un intervalo aceptable es con una ventana del 5 % alrededor de 1 para admitir convergencia [30].

**Observación:** Es importante que después de obtener cadenas convergentes se generen más iteraciones para producir una mejor estimación posterior y evaluar el error

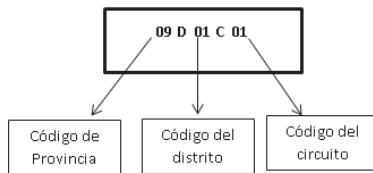
estándar de Monte Carlo para cada uno de los parámetros. WinBUGS14 nos indica que este valor es una expresión de la diferencia entre la media de los valores muestrados y la verdadera media posterior, donde para llegar a la eficiencia el error de Monte Carlo este valor debe ser pequeño en relación a la desviación estándar posterior [30].

### 3.5. Construcción de Mapas por Circuitos

El código geográfico que utiliza el Centro Nacional de Referencia e Investigación en Vectores (CIREV) es el que se encuentra establecido por la Policía Nacional del Ecuador, donde:

- Los Distritos están conformados por uno o varios Cantones.
- Los Circuitos están conformados por uno o varios Subcircuitos.

Los circuitos al ser elementos de los distritos, mantienen el código del distrito al que pertenecen, el número de circuito asignado, el código provincial (2 dígitos), la letra **D** del distrito, el número correspondiente al distrito dentro de la provincia (2 dígitos), la letra **C** del circuito y el número correspondiente al circuito dentro del distrito (2 dígitos), como en la figura siguiente:



**Figura 3.11:** Nomenclatura para definir el código de un circuito

En el caso de la ciudad de Guayaquil, las localidades siguen esta codificación:

Código	Número
Provincia del Guayas	09
Distrito Sur	01
Circuito Guasmo	01

**Cuadro 3.5:** Tabla de códigos para los circuitos de la Ciudad de Guayaquil

Con esta identificación para el diseño de mapas usamos el *Quantum Gis* (*Qgis*) que es un Software de Sistemas de Información Geográfica de libre acceso, que permite visualizar, manejar, editar y elaborar mapas a través de la estructura Shape (*.SHP Archivos compatibles con Sistemas de Información Geográfica*) para ubicar cíntomas y circuitos.

Con la determinación de los factores involucrados en la transmisión del dengue se procederá a armar la estructura para aplicar un modelo bayesiano espacio temporal teniendo en cuenta la variabilidad mensual de los datos y la distribución por circuitos. Además se planteará el cálculo de las distribuciones a posteriori para cada uno de los parámetros inmersos en el modelo a ajustarse.

## Capítulo 4

# Modelo Bayesiano Espacio Temporal

Para esta sección cubriremos la descripción de todos los modelos planteados que serán ingresados en el programa WinBUGS14 y monitoreados respecto de la convergencia a través de las cadenas markovianas generadas para cada parámetro que contemplen las estructuras bayesianas, además de analizar de forma gráfica y por el criterio de Gelman-Rubin la convergencia de las mismas.

La ubicación cartográfica de enfermedades conocida como *disease mapping* ha tomado fuerza como herramienta gráfica destinada a la localización de las diferentes distribuciones de los riesgos relativos que se encuentran influenciados por condiciones referentes al aspecto humano, territorial y ambiental.

La inclusión de los mapas de riesgo se vuelven necesarios al tratarlos como un sistema de información que permita el análisis periódico de sucesos, en nuestro caso los sucesos epidemiológicos están vinculados a la alerta temprana del brote del dengue en Guayaquil. En un ámbito informativo con esta herramienta se puede entender de manera gráfica la dinámica de la enfermedad, la misma que da indicios para expandir esta técnica hacia nuevas zonas geográficas o hacia otras patologías similares [23].

En este punto surge la inquietud para establecer una medida apropiada en cada nivel de riesgo que concuerde con la ocurrencia de eventos en áreas geográficas pequeñas, dado que al tomar la representación visual de los casos de dengue en bruto no tenemos información veraz al momento de tomar decisiones debido a probables distorsiones.

Una primera noción sugiere la representación geográfica por medio de las tasas

brutas como un identificador principal del riesgo haciendo una división simple entre el total de casos observados y el total de la población expuesta, por cada área considerada para el estudio.

Con este razonamiento se presentan algunas dificultades que enmascaran diferencias al variar la localidad, especialmente por la densidad territorial y el contexto asociado a las condiciones de vida. Estas limitantes se ven reflejadas ya que no se toman en cuenta las diferencias de los tamaños poblacionales que son variantes de un lugar a otro, por ejemplo, cuando las poblaciones son numéricamente pequeñas nos enfrentamos a resultados inestables pasando de cero hacia cifras altas cuando los casos se incrementan (aunque este crecimiento en cifras sea casi imperceptible). Otro punto de vista alumbría la posibilidad de que los casos esperados superen en número a los casos confirmados, dando como resultado una extra variabilidad.

Una opción para superar estos inconvenientes es la introducción en la modelación bayesiana donde los autores Clayton y Kaldor [48] demostraron que ya sea con técnicas de estandarización suavizadas o con la estimación de los riesgos relativos por el método bayesiano se obtiene diferentes valores numéricos pero coincidentes en términos absolutos si hablamos de rangos, y adecuados en términos relativos respecto del orden de riesgo.

Con este preámbulo definimos la tasa de incidencia estandarizada (*TIE*), como la razón entre el número de casos observados de la enfermedad y el número de casos esperados en cada lugar de estudio  $s = (1, \dots, S)$ . Este estimador se expresa con la ecuación:

$$TIE_s = \frac{y_s}{e_s} \quad (4.1)$$

En (4.1), un valor superior a la unidad nos alertaría de una mayor incidencia, observando que para los lugares donde la población es mayor se prevé un número creciente de casos y de forma directamente proporcional para las poblaciones de menor tamaño. Este resultado sugiere que el número de casos depende del tamaño de la población, pero otro aspecto importante es la cantidad de casos esperados que puede ser cercana a cero y que hace que la tasa de incidencia tienda a crecer y se vuelva una cantidad grande, sin importar la cifra de habitantes. Incluso es posible que existan áreas donde la incidencia sea cero, estos casos extremos nos presentan una perspectiva en la que el estimador *TIE* no distingue la variación entre los casos esperados y la influencia de los tamaños poblacionales. Por lo tanto, como la varianza del estimador es proporcional a  $(1/e_s)$  tenemos que la *TIE* es un estimador saturado del riesgo relativo que no refleja una situación acorde a la realidad [30].

Una manera de mejorar esta tasa *TIE* es a través del modelo bayesiano espacio temporal que reúne a los siguientes parámetros y variables.

Sea  $s = (1, \dots, S)$  el conjunto de circuitos,  $t = (1, \dots, T)$  los meses del período 2010-2015, entonces se establecen los conceptos:

- $Y_{st}$  es el número de casos modelados en el circuito  $s$  y mes  $t$ .
- $y_{st}$  representa los casos de dengue reportados en el circuito  $s$  en el mes  $t$ .
- $\lambda_{st}$  es el parámetro de la distribución de Poisson dado para cada circuito  $s$  en el mes  $t$ .
- $\psi_{st}$  es el riesgo relativo de contraer dengue en el circuito  $s$  y el mes  $t$
- $E_{st}$  es el número esperado de casos en el circuito  $s$  y tiempo  $t$  si  $P_{st}^*$  es la probabilidad de contraer la enfermedad cada circuito  $s$  en el tiempo  $t$ , expresada con:

$$P_{st}^* = \frac{\sum_{s=1}^S y_{st}}{\sum_{s=1}^S p_{st}} \quad (4.2)$$

- $p_{st}$  es la población en el circuito  $s$  en el mes  $t$ .

En la parte lineal identificaremos  $P$  variables climáticas (temperatura, precipitación, índice de vegetación,) y  $Q$  variables socioeconómicas que están vinculados con el analfabetismo, desempleo, razón persona -dormitorio (hacinamiento), agua potable, calidad de la vivienda, alcantarillado y recolección de basura, todos estos factores condensados en las variables transformadas a través del método de los componentes principales:  $Z_1$  (*condiciones de salubridad*) y  $Z_2$  (*condiciones de pobreza*).

Con estos fundamentos la formulación del modelo bayesiano establece las siguientes definiciones:

$$Y_{st} \sim Poisson(\lambda_{st}) \quad (4.3)$$

donde el parámetro para la distribución de Poisson está dado por (4.4):

$$\lambda_{st} = E_{st}\psi_{st} \quad (4.4)$$

Y el cálculo de  $E_{st}$  toma la siguiente expresión:

$$E_{st} = p_{st}P_{st}^* \quad (4.5)$$

Notemos que el número de casos esperados ( $E_{st}$ ) se consideran conocidos dado que tenemos como observaciones conocidas a:  $y_{st}$  y  $p_{st}$ .

En la ecuación (4.7) gracias a la función de enlace expresada a través del logaritmo en (4.6) y la estructura para el predictor lineal, incluimos una constante  $\alpha_t$  para cada mes  $t$ ,  $\beta_t^j$  y  $\gamma_s^k$  como los coeficientes de regresión asociados a las variables climáticas y socioeconómicas, respecto de cada mes y cada circuito, como corresponda.

$$\hat{\mu}_{st} = \ln \left( \alpha_t + \sum_{j=1}^P \beta_t^j x_{st}^j + \sum_{k=1}^Q \gamma_s^k Z_s^k + \mu_{st} + \nu_{st} \right) \quad (4.6)$$

Tenemos también la presencia del término  $\mu$  como el coeficiente que capta la heterogeneidad espacial entre circuitos y  $\nu$  que recoge el comportamiento de los vecinos alrededor de cada circuito. Por lo tanto  $\lambda_{st}$  toma la forma (4.7).

$$\lambda_{st} = E_{st} * \exp \left( \alpha_t + \sum_{j=1}^P \beta_t^j x_{st}^j + \sum_{k=1}^Q \gamma_s^k Z_s^k + \mu_{st} + \nu_{st} \right) \quad (4.7)$$

## 4.1. Distribuciones a priori

Para completar el diseño del modelo (4.7) definiremos las distribuciones a priori basándonos en el apoyo bibliográfico [32] y sobretodo porque estas son tomadas para las simulaciones dado que tienen densidades planas y a su vez cubren una amplia gama de valores que son generados en la parte interna del programa WinBUGS14.

Para todo circuito  $s = (1, \dots, S)$ , todo mes  $(t = 1, \dots, T)$ , para toda variable climática  $(j = 1, \dots, P)$ , y toda variable socioeconómica  $(k = 1, \dots, Q)$  se toman las distribuciones siguientes:

- $\alpha_t \sim U(-\infty, \infty)$
- $\beta_t^j \sim N(0, 1 \times 10^{-6})$
- $\gamma_s^k \sim N(0, 1000)$
- $\mu_{st} \sim N(0, \frac{1}{\tau_u})$
- $\tau_\mu \sim Gamma(a_\mu, b_\mu)$ , donde  $a_\mu = 0,5$  y  $b_\mu = 0,0005$
- $\nu_{st} | \nu_{-st} \sim N(\bar{\nu}_{st}, \frac{1}{\tau_v m_{st}})$
- $\tau_v \sim Gamma(c_v, d_v)$ , donde  $c_v = 0,5$  y  $d_v = 0,0005$

## 4.2. Distribuciones a posteriori

Para la construcción de las distribuciones condicionales a posteriori partimos con la función de densidad para una variable de Poisson de parámetro  $\lambda$ :

$$P(\hat{y}|\lambda) = \frac{\lambda^{\hat{y}}}{\hat{y}!} \exp(-\lambda) \quad (4.8)$$

Luego, para todas las observaciones  $y_{st}$  y con la ecuación (4.8) construimos la función de verosimilitud que luego tomará la notación  $l(y|\lambda)$ , como se muestra en (4.9):

$$\begin{aligned} Pr(y|\lambda) &= \left( \prod_{t=1}^T \prod_{s=1}^S \frac{\lambda_{st}^{y_{st}} \exp(-\lambda_{st})}{y_{st}!} \right) \\ &= \prod_{t=1}^T \left( \prod_{s=1}^S \frac{\lambda_{st}^{y_{st}} \exp(-\lambda_{st})}{y_{st}!} \right) \\ &= \underbrace{\left( \prod_{t=1}^T \prod_{s=1}^S \frac{\lambda_{st}^{y_{st}}}{y_{st}!} \right)}_{A} \underbrace{\left( \prod_{t=1}^T \prod_{s=1}^S \exp(-\lambda_{st}) \right)}_{B} \end{aligned} \quad (4.9)$$

A continuación describimos la parte A y B de la ecuación (4.9), pues en los procesos siguientes serán de utilidad.

$$A = \prod_{t=1}^T \left( \prod_{s=1}^S \frac{E_{st}^{y_{st}}}{y_{st}!} \exp \left( \alpha_t + \sum_{j=1}^p \beta_t^j x_{st}^j + \sum_{k=1}^Q \gamma_s^k Z_s^k + \mu_{st} + \nu_{st} \right) \right) \quad (4.10)$$

$$B = \exp \left( - \sum_{t=1}^T \sum_{s=1}^S \lambda_{st} \right) = \exp \left( - \sum_{t=1}^T \sum_{s=1}^S E_{st} \exp \left( \alpha_t + \sum_{j=1}^p \beta_t^j x_{st}^j + \sum_{k=1}^Q \gamma_s^k Z_s^k + \mu_{st} + \nu_{st} \right) \right) \quad (4.11)$$

Con este paso previo en la sección siguiente detallaremos los cálculos de las distribuciones a posteriori de los parámetros del modelo (4.7), notando a  $c$  con la expresión siguiente:

$$c = \prod_{t=1}^T \left( \prod_{s=1}^S \frac{E_{st}^{y_{st}}}{y_{st}!} \right) \quad (4.12)$$

#### 4.2.1. Distribución condicional para $\alpha_t$

Para todo  $t \in 1, 2, \dots, T$ , y todo circuito  $s \in 1, 2, \dots, S$ :

$$\begin{aligned}
Pr(\alpha|y, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) &= l(y|\alpha, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) Pr(\alpha) \\
&= c * \exp\left(\sum_{t=1}^T \sum_{s=1}^S \alpha_t y_{st}\right) \exp\left(-\sum_{t=1}^T \sum_{s=1}^S E_{st} \exp(\alpha_t) Pr(\alpha)\right) \\
&= c_2 * \exp\left(\sum_{t=1}^T \sum_{s=1}^S \alpha_t y_{st} - E_{st} \exp(\alpha_t)\right) \\
&= \exp(\ln(c_2)) * \exp\left(\sum_{t=1}^T \sum_{s=1}^S \alpha_t y_{st} - E_{st} \exp(\alpha_t)\right) \\
&= \exp\left(\left(\sum_{t=1}^T \sum_{s=1}^S \alpha_t y_{st} - E_{st}\right) \exp(\alpha_t) + \ln(c_2)\right)
\end{aligned} \tag{4.13}$$

donde:  $c_2 = c * c_1 = c * Pr(\alpha)$  dado que  $Pr(\alpha) \sim U(-\infty, \infty)$

#### 4.2.2. Distribución condicional para $\beta_t^j$

Para todo  $t \in 1, 2, \dots, T$ , y todo circuito  $s \in 1, 2, \dots, S$ , y toda  $j \in 1, \dots, P$

$$\begin{aligned}
Pr(\beta_t^j | \alpha_t, y, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) &= l(y|\alpha_t, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) Pr(\beta_t^j) \\
&= c_3 * \exp\left(\sum_{t=1}^T \sum_{s=1}^S (y_{st} \sum_{j=1}^P \beta_t^j x_{st}^j) \exp(-\sum_{t=1}^T \sum_{s=1}^S E_{st} \exp(\sum_{j=1}^P \beta_t^j x_{st}^j))\right) \\
&= c_3 * \exp\left(\sum_{t=1}^T \sum_{s=1}^S \left(\sum_{j=1}^P y_{st} \beta_t^j x_{st}^j\right) - E_{st} \exp\left(\sum_{j=1}^P \beta_t^j x_{st}^j\right)\right)
\end{aligned} \tag{4.14}$$

donde:  $c_3 = Pr(\beta_t^j)$  con la probabilidad  $N(0, 10^{-6})$ .

#### 4.2.3. Distribución condicional para $\gamma_s^k$

Para todo  $t \in 1, 2, \dots, T$ , y todo circuito  $s \in 1, 2, \dots, S$ , y toda  $k \in 1, \dots, Q$

$$\begin{aligned} Pr(\gamma_s^k | \alpha_t, \beta_t^j, y, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) &= l(y | \alpha_t, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) Pr(\gamma_s^k) \\ &= c_4 * \exp \left( \sum_{t=1}^T \sum_{s=1}^S \left( y_{st} \sum_{k=1}^Q \gamma_s^k z_s^k \right) \exp \left( - \sum_{t=1}^T \sum_{s=1}^S E_{st} \right) \exp \left( \sum_{k=1}^Q \gamma_s^k z_s^k \right) \right) \\ &= c_4 * \exp \left( \sum_{t=1}^T \sum_{s=1}^S \left( \sum_{k=1}^Q \gamma_s^k z_s^k - E_{st} \exp \left( \sum_{k=1}^Q \gamma_s^k z_s^k \right) \right) \right) \end{aligned} \quad (4.15)$$

donde:  $c_4 = Pr(\gamma_s^k)$  con la probabilidad  $N(0, 1000)$

#### 4.2.4. Distribución condicional para $\nu_{st}$

Para todo  $t \in 1, 2, \dots, T$ , y todo circuito  $s \in 1, 2, \dots, S$ ,

$$\begin{aligned} Pr(\nu_{st} | \nu_{-st} \alpha_t, \beta_t^j, y, \mu_{st}, \tau_\mu, \gamma_s^k, \tau_\nu) &= l(y | \alpha, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) Pr(\nu_{st} | \tau_\nu) \\ &= c * \exp \left( \sum_{t=1}^T \sum_{s=1}^S (y_{st} \nu_{st}) \exp \left( - \sum_{t=1}^T \sum_{s=1}^S E_{st} \right) \exp(\nu_{st}) * \exp \left( - \frac{\tau_{nu} \nu_{st}}{2} \sum_{s \neq t}^S (\nu_{st} - \bar{\nu}_{st})^2 \right) \right) \end{aligned} \quad (4.16)$$

#### 4.2.5. Distribución condicional para $\mu_{st}$

Para todo  $t \in 1, 2, \dots, T$ , y todo circuito  $s \in 1, 2, \dots, S$ ,

$$\begin{aligned} Pr(\mu_{st} | \mu_{-st} \alpha_t, \beta_t^j, y, \nu_{st}, \tau_\mu, \gamma_s^k, \tau_\nu) &= l(y | \alpha, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) Pr(\nu_{st} | \tau_\nu) \\ &= c * \exp \left( \sum_{t=1}^T \sum_{s=1}^S (y_{st} \mu_{st}) \exp \left( - \sum_{t=1}^T \sum_{s=1}^S E_{st} \right) \exp(\mu_{st}) * \exp \left( - \frac{\tau_{mu} \mu_{st}}{2} \sum_{s=1}^S (\mu_{st})^2 \right) \right) \end{aligned} \quad (4.17)$$

#### 4.2.6. Distribución condicional para $\tau_\mu$

Para todo  $t \in 1, 2, \dots, T$ , y todo circuito  $s \in 1, 2, \dots, S$ ,

$$\begin{aligned} Pr(\tau_\mu | \alpha_t, \beta_t^j, y, \nu_{st}, \tau_\mu, \gamma_s^k, \tau_\nu) &= l(y | \alpha, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, \nu_{st}, \tau_\nu) Pr(\mu_{st} | \tau_\nu) Pr(\tau_\mu) \\ &= \prod_{t=1}^T \tau_\mu^{\frac{s}{2} + a_u - 1} \exp(-\tau_\mu(b_\mu + \frac{1}{2} \sum_{s=1}^C \mu_{st})) \end{aligned} \quad (4.18)$$

#### 4.2.7. Distribución condicional para $\tau_v$

Para todo  $t \in 1, 2, \dots, T$ , y todo circuito  $s \in 1, 2, \dots, S$ ,

$$\begin{aligned} Pr(\tau_v | \alpha_t, \beta_t^j, y, v_{st}, \tau_\mu, \gamma_s^k, \tau_v) &= l(y | \alpha, \beta_t^j, \gamma_s^k, \mu_{st}, \tau_\mu, v_{st}, \tau_v) Pr(v_{st} | \tau_v) Pr(\tau_v) \\ &= \prod_{t=1}^T \tau_v^{\frac{S}{2}} \exp\left(\frac{\tau_v}{2}(v_{st} - \bar{v}_{st})^2\right) \tau_v^{c_\mu-1} \exp(-d_\nu \tau_\mu) \end{aligned} \quad (4.19)$$

# Capítulo 5

## Resultados

En este capítulo presentaremos los resultados obtenidos del ajuste de todas las estructuras consideradas a partir de los criterios de información de Devianza  $DIC$  y pérdida predictiva a posteriori  $p_D$ , implementados en los códigos monitoreados con el software WinBUGS14. Además se presentan las tablas de riesgo relativo para cada temporada y circuitos en los años 2010 al 2015, así como los mapas de riesgo correspondientes.

### 5.1. Descripción de los modelos

Para obtener los valores iniciales que pide por defecto el programa WinBUGS14 generamos las siguientes estructuras, cada una con 10.000 iteraciones.

- $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1)$
- $\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2)$
- $\lambda_{st} = E_{st} \exp(\beta_t^3 x_{st}^3)$
- $\lambda_{st} = E_{st} \exp(\gamma_s^1 z_s^1)$
- $\lambda_{st} = E_{st} \exp(\gamma_s^2 z_s^2)$

Donde las variables anteriores se describen a continuación:

$\beta_t^1$ : Representa el coeficiente que acompaña a la variable temperatura  $x_{st}^1$  definida para cada circuito  $s$  y cada mes  $t$ .

$\beta_t^2$ : Representa el coeficiente que acompaña a la variable precipitación  $x_{st}^2$  definida para cada circuito  $s$  y cada mes  $t$ .

$\beta_t^3$ : Representa el coeficiente que acompaña a la variable índice de vegetación  $x_{st}^3$  definida para cada circuito  $s$  y cada mes  $t$ .

$\gamma_s^1$ : Representa el coeficiente que acompaña a la variable condiciones de salubridad  $Z_s^1$  definida para cada circuito  $s$ .

$\gamma_s^2$ : Representa el coeficiente que acompaña a la variable condiciones de pobreza  $Z_s^2$  definida para cada circuito  $s$ .

Para cada una de las formas anteriores se revisó en cada paso la convergencia usando los criterios mencionados en la sección anterior, con lo que el programa provee los valores iniciales para las variables  $\beta_t^1$ ,  $\beta_t^2$  y  $\beta_t^3$  para cada mes  $t$ , al igual que para los coeficientes asociadas a los indicadores del INEC como son:  $\gamma_s^1$  y  $\gamma_s^2$ , para cada circuito  $s$ .

Para los modelos presentados a continuación tomamos la distribución Normal (0, 1000) apoyados en [20], ya que cubre una amplia gama de valores que se generan internamente en cada iteración que simula WinBUGS14, y para las variables climáticas las sugerencias dispuestas en la bibliografía [48].

## 5.2. Modelos espacio-temporales monitoreados en WinBUGS14

Los modelos que se describen en esta sección fueron monitoreados en el software WinBUGS14 incluyendo para cada estructura la componente de heterogeneidad  $\mu$  y la parte que captura la influencia de los vecinos de cada circuito a través de  $\nu$

1.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \gamma_s^1 z_s^1 + \gamma_s^2 z_s^2)$
2.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3)$
3.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2)$
4.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^3 x_{st}^3)$
5.  $\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3)$
6.  $\lambda_{st} = E_{st} \exp(\gamma_s^1 z_s^1 + \gamma_s^2 z_s^2)$
7.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \gamma_s^1 z_s^1)$
8.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \gamma_s^2 z_s^2)$

9.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \gamma_s^1 z_s^1)$
10.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \gamma_s^2 z_s^2)$
11.  $\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \gamma_s^1 z_s^1)$
12.  $\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \gamma_s^2 z_s^2)$
13.  $\lambda_{st} = E_{st} \exp(\beta_t^3 x_{st}^3 + \gamma_s^1 z_s^1)$
14.  $\lambda_{st} = E_{st} \exp(\beta_t^3 x_{st}^3 + \gamma_s^2 z_s^2)$
15.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \gamma_s^1 z_s^1)$
16.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \gamma_s^2 z_s^2)$
17.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^3 x_{st}^3 + \gamma_s^1 z_s^1)$
18.  $\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^3 x_{st}^3 + \gamma_s^2 z_s^2)$
19.  $\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^1 + \beta_t^3 x_{st}^3 + \gamma_s^1 z_s^1)$
20.  $\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^1 + \beta_t^3 x_{st}^3 + \gamma_s^2 z_s^2)$

Todos las ecuaciones anteriores fueron ingresadas con dos series de valores iniciales generados por las estimaciones de los primeros modelos, luego de que se iteraron 5.000 simulaciones para indagar la convergencia de las cadenas se procedió a implementar la fase final con 10.000 y hasta 50.000 iteraciones, en algunos casos.

En la mayoría de los modelos la componente  $v$  impidió la convergencia de las cadenas del resto de parámetros inmersos en el modelo correspondiente, lo que nos da una idea de que como los circuitos no varían de ubicación con el pasar de los años ni se han cambiado las áreas limítrofes no tiene mayor influencia como la que refleja la componente de heterogeneidad espacial explicada por  $\mu$ . Con estas restricciones presentamos en (5.1) la comparación de los modelos basados en el  $DIC$  y  $p_D$

### 5.3. Comparación de los Modelos

En esta tabla (5.1) presentamos los modelos monitoreados apoyados de los criterios de devianza  $DIC$  y  $p_D$ , donde para elegir el mejor modelo ajustado tomaremos aquel que tenga el menor valor  $DIC$ , observando que la cercanía entre ellos nos induce a tener en cuenta un segundo enfoque que es el de  $p_D$  junto con las ideas planteadas en los objetivos de este proyecto.

Modelo	Ajuste			
	Dbar	Dhat	pD	DIC
$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^3 x_{st}^3 + \mu_{st})$	9123.56	7190.63	1932.93	11056.5
$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \mu_{st} + \nu_{st})$	9061.39	7121.87	1939.52	1000.9
$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^3 x_{st}^3 + \gamma_s^2 Z_s^2 + \mu_{st})$	9139.87	7195.06	1944.81	11084.7
$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \gamma_s^2 Z_s^2 + \mu_{st})$	9116.9	7162.08	1954.82	11071.7
$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \gamma_s^1 Z_s^1 + \mu_{st})$	9119.8	7163.66	1956.14	11075.9
$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \gamma_s^1 Z_s^1 + \mu_{st})$	9101.48	7133.98	1967.49	11069
$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \gamma_s^2 Z_s^2 + \mu_{st})$	9130.18	7142.73	1987.44	11117.6
$\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \mu_{st})$	9170.42	7159.48	2010.94	11181.4
$\lambda_{st} = E_{st} \exp(\beta_t^3 x_{st}^3 + \gamma_s^2 Z_s^2 + \mu_{st})$	9205.39	7192.72	2012.67	11218.1
$\lambda_{st} = E_{st} \exp(\beta_t^3 x_{st}^3 + \gamma_s^1 Z_s^1 + \mu_{st})$	9231.48	7203.45	2028.03	11259.5
$\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \gamma_s^2 Z_s^2 + \mu_{st})$	9207.19	7170.8	2036.39	11243.6
$\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \gamma_s^1 Z_s^1 + \mu_{st})$	9211.4	7173.57	2037.83	11249.2
$\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \gamma_s^2 Z_s^2 + \mu_{st})$	9293.1	7220.98	2072.13	11365.2
$\lambda_{st} = E_{st} \exp(\beta_t^2 x_{st}^2 + \gamma_s^1 Z_s^1 + \mu_{st})$	9321.22	7239.17	2082.05	11403.3
$\lambda_{st} = E_{st} \exp(\gamma_s^1 Z_s^1 + \gamma_s^1 Z_s^1 + \mu_{st})$	9555.18	7322.03	2233.15	11788.3

Cuadro 5.1: Ajuste bayesiano para los modelos monitoreados en WinBUGS14

## 5.4. Modelo elegido

En la tabla (5.1) existen dos modelos que cubren las necesidades planteadas, sin embargo en la ecuación (5.1) vemos la existencia de una relación entre el riesgo relativo y las condiciones de temperatura, precipitación e índice de vegetación incluyendo la influencia espacial y la de los vecinos cercanos. Haremos también la observación que la constante no fue incluida por ser un motivo de sobre saturación donde no se alcanza la convergencia.

$$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \beta_t^3 x_{st}^3 + \mu_{st} + \nu_{st}) \quad (5.1)$$

Por otro lado (5.2) fusiona el efecto del escenario climático y los indicadores correspondientes a las condiciones de salubridad, incluyendo el efecto de la autocorrelación espacial.

$$\lambda_{st} = E_{st} \exp(\beta_t^1 x_{st}^1 + \beta_t^2 x_{st}^2 + \gamma_s^1 z_s^1 + \mu_{st}) \quad (5.2)$$

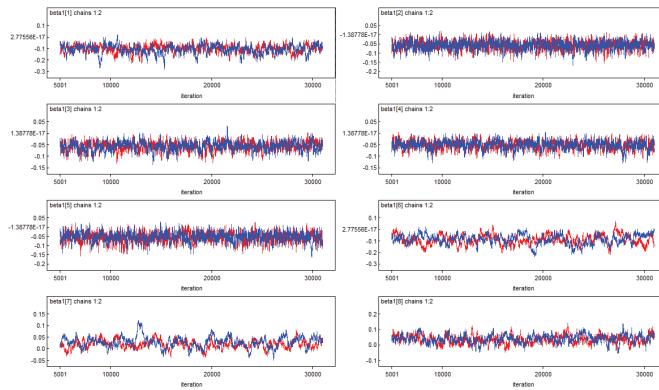
## 5.5. Monitoreo de la Convergencia para el Modelo Ele-gido

En esta sección presentamos los gráficos relativos a la inspección de la convergencia para los parámetros  $\beta_t^1, \beta_t^2, \gamma_s^1, \mu, \tau_\mu$ , para cada mes  $t$  y cada circuito  $s$ . Tendremos en mente los criterios descritos en el sección *Convergencia de WinBUGS14* y para la componente de heterogeneidad mencionaremos la convergencia mediante la prueba de Gelman y Rubin, sin incluir los gráficos mensuales y por circuitos debido a la redundancia de información para concluir la convergencia de las cadenas generadas para dichos parámetros.

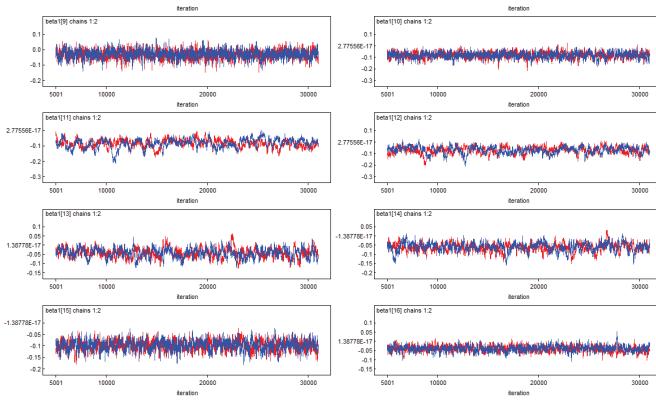
Para las tablas que contienen a los estadísticos a posteriori se agruparon los meses de acuerdo a las estaciones lluviosa y seca y se hizo un promedio en cada parámetro para obtener la media, la desviación estándar, error de Monte Carlo, la mediana y los intervalos del confianza, al igual que en las tablas donde se expresa el riesgo relativo.

### 5.5.1. Convergencia para el parámetro $\beta_t^1$

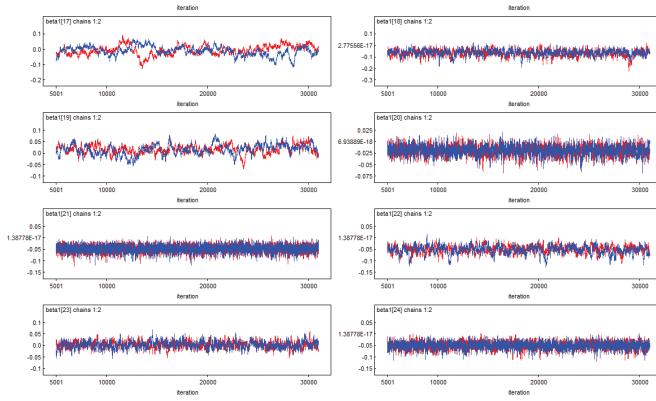
A través del gráfico *history plot* vemos el intervalo de iteraciones que va desde la número 5.000 a la 30.000 después de haber tenido un período burn-in de 10.000 de ellas hasta que se stabilicen las cadenas, organizadas por los períodos que se muestran en las leyendas de cada figura.



**Figura 5.1:** History plot para la variable  $\beta_t^1$ , desde enero 2010 - agosto 2010

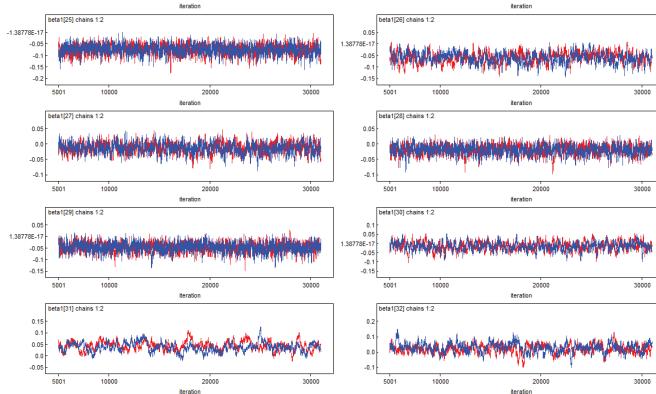


**Figura 5.2:** History plot para la variable  $\beta_t^1$ , desde septiembre 2010 - abril 2011

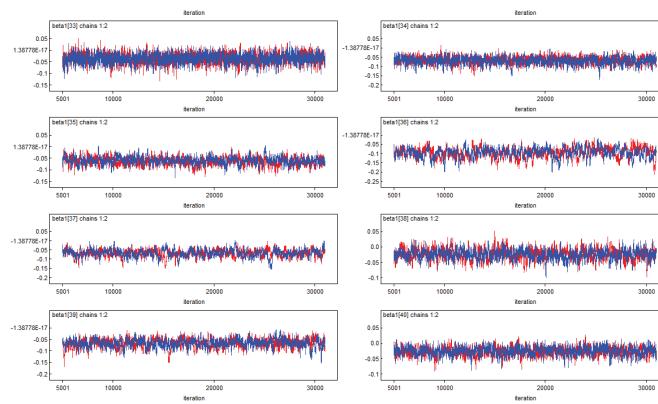


**Figura 5.3:** History plot para la variable  $\beta_t^1$ , desde mayo 2011 - diciembre 2011

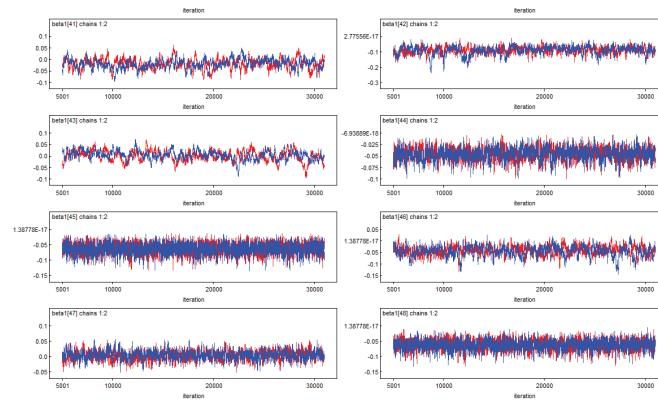
En cada una de las gráficas podemos observar que las cadenas se superponen unas sobre otras, lo que nos da un indicio de convergencia.



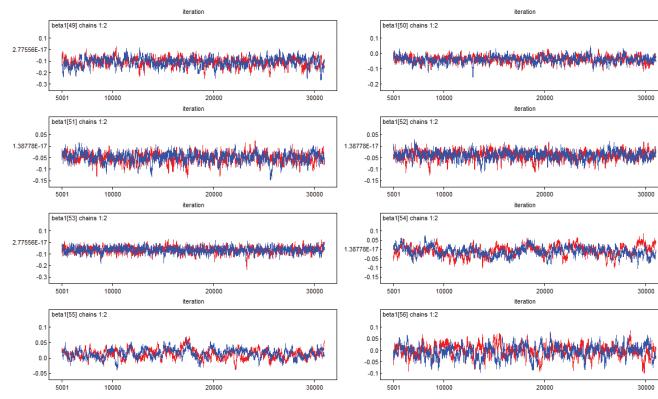
**Figura 5.4:** History plot para la variable  $\beta_t^1$ , desde enero 2012 - agosto 2012



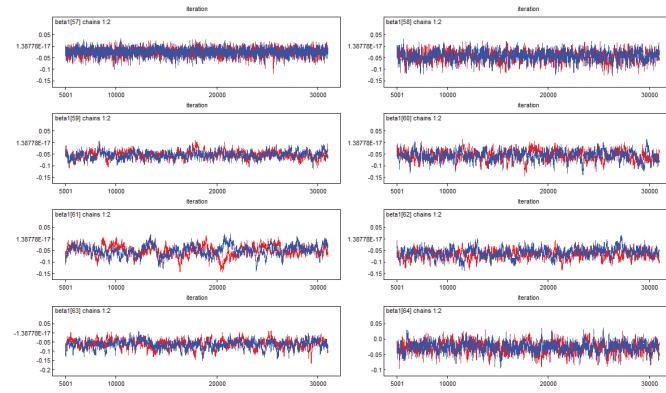
**Figura 5.5:** History plot para la variable  $\beta_t^1$ , desde septiembre 2012 - abril 2013



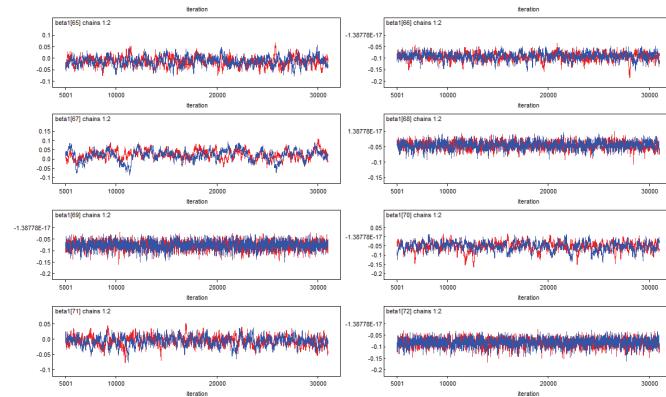
**Figura 5.6:** History plot para la variable  $\beta_t^1$ , desde mayo 2013 - diciembre 2013



**Figura 5.7:** History plot para la variable  $\beta_t^1$ , desde enero 2014 - agosto 2014

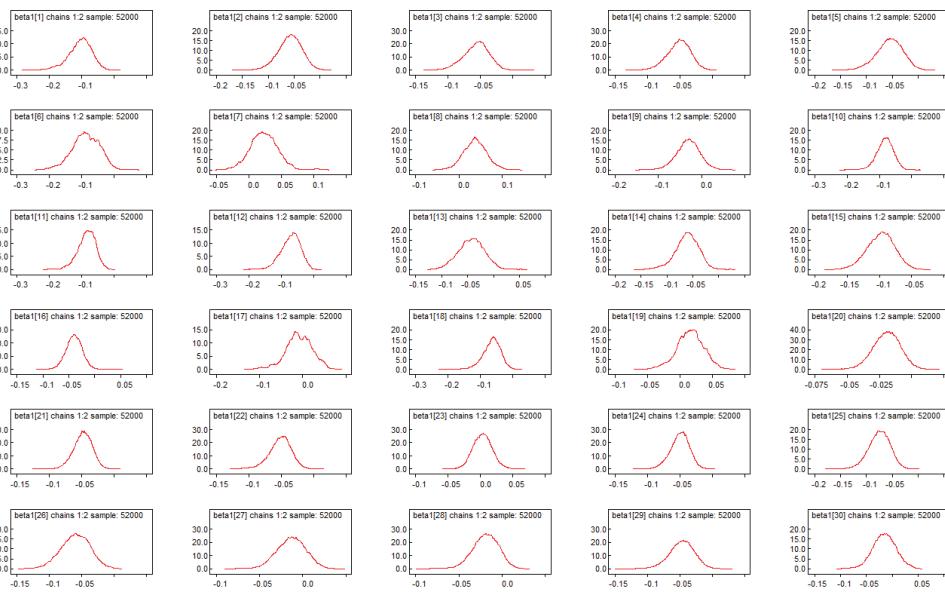


**Figura 5.8:** History plot para la variable  $\beta_t^1$ , desde septiembre 2014 - abril 2015

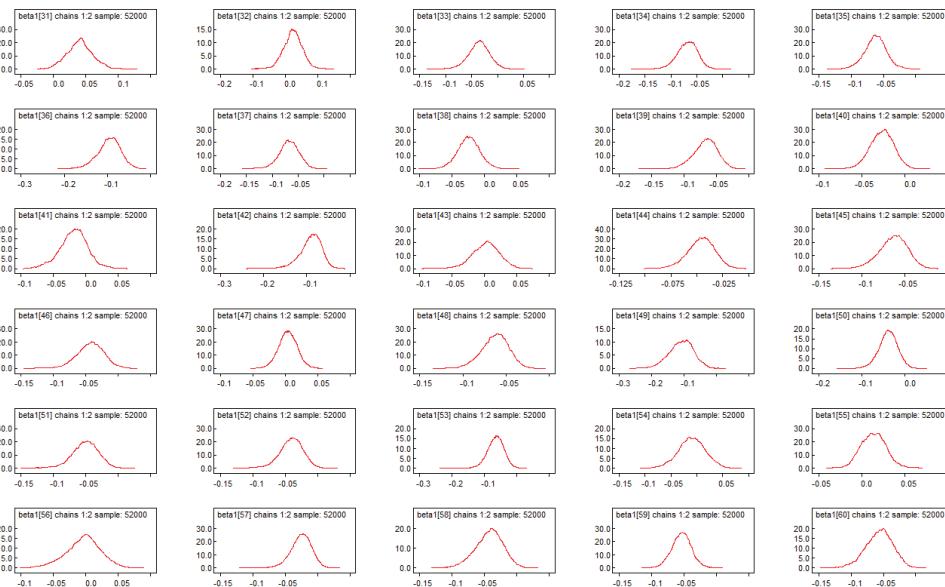


**Figura 5.9:** History plot para la variable  $\beta_t^1$ , desde mayo 2015 - diciembre 2015

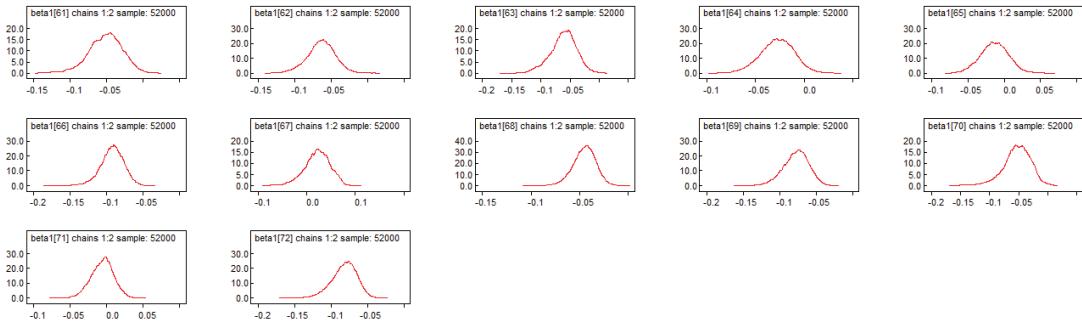
En los gráficos de las densidades tenemos que la mayor parte de las curvas guardan suavidad, es decir, no tenemos mayor irregularidad y se puede decir también que algunas de ellas son simétricas.



**Figura 5.10:** Gráfico de la densidad para la variable  $\beta_t^1$ , desde enero 2010- junio 2012



**Figura 5.11:** Gráfico de la densidad para la variable  $\beta_t^1$ , desde julio 2012-diciembre 2014



**Figura 5.12:** Gráfico de la densidad para la variable  $\beta_t^1$ , desde enero 2015- diciembre 2015

En los gráficos de correlación para esta variable se dibujaron histogramas casi imperceptibles, es decir, que las frecuencias decrecieron conforme aumentaban las iteraciones, al igual que el test de Gelman y Rubin en la que solo un subconjunto pequeño de alrededor 8 casos de todos los betas estuvieron fuera del rango de convergencia que es entre 0.95 y 1.05.

En las tablas siguientes se presentan los estadísticos a posteriori organizados para los años 2010 al 2015 y por las estaciones lluviosa y seca.

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^1$	-0.066104	0.024536	0.00138706	-0.064694	[-0.118068,-0.021752]
Seca	$\beta^1$	-0.0394	0.02778571	0.00119926	-0.051866	[-0.102268,-0.0129296]

**Cuadro 5.2:** Estadísticos a posteriori para  $\beta^1$ , año 2010, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^1$	-0.058402	0.020724	0.00114562	-0.05734	[-0.10376,-0.0210886]
Seca	$\beta^1$	-0.02642171	0.01955286	0.00147012	-0.0486928	[-0.099144,-0.0045726]

**Cuadro 5.3:** Estadísticos a posteriori para  $\beta^1$ , año 2011, estaciones lluviosa y seca

En los cuadros (5.2), (5.3), (5.4), (5.5), (5.6) y (5.7) verificamos que el error de Monte Carlo (MC) es menor que la desviación estándar a posteriori, lo que asegura la convergencia y el ajuste final para este parámetro.

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^1$	-0.051774	0.020416	0.00105442	-0.050898	[-0.094524,-0.0141936]
Seca	$\beta^1$	-0.02396429	0.01978143	0.00089915	-0.0344825	[-0.072895,-0.000762]

**Cuadro 5.4:** Estadísticos a posteriori para  $\beta^1$ , año 2012, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^1$	-0.04512	0.01872	0.00104272	-0.044102	[-0.084462,-0.010352]
Seca	$\beta^1$	-0.04377829	0.01825714	0.00107038	-0.0484575	[-0.0903125,-0.016042]

**Cuadro 5.5:** Estadísticos a posteriori para  $\beta^1$ , año 2013, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^1$	-0.057688	0.02497	0.00149176	-0.056622	[-0.111246,-0.0120636]
Seca	$\beta^1$	-0.02564871	0.02020143	0.0011687	-0.046825	[-0.0909575,-0.0061945]

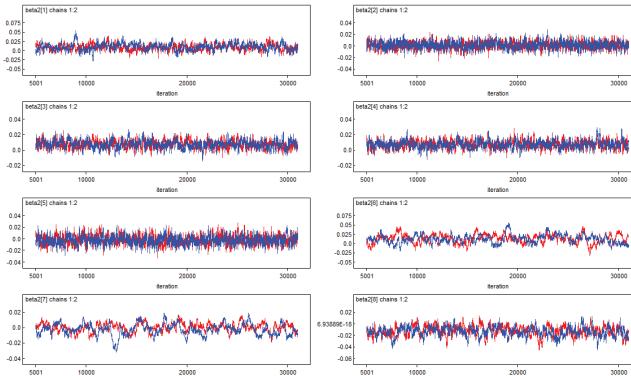
**Cuadro 5.6:** Estadísticos a posteriori para  $\beta^1$ , año 2014, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^1$	-0.057038	0.019344	0.0010967	-0.05613	[-0.09833,-0.0213132]
Seca	$\beta^1$	-0.03716057	0.01836714	0.00123728	-0.04124	[-0.0864075,-0.0040265]

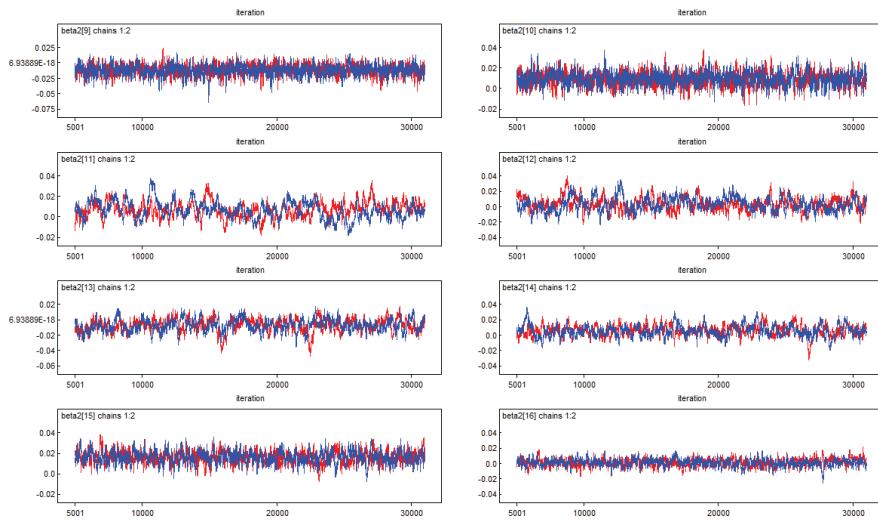
**Cuadro 5.7:** Estadísticos a posteriori para  $\beta^1$ , año 2015, estaciones lluviosa y seca

### 5.5.2. Convergencia para el parámetro $\beta_t^2$

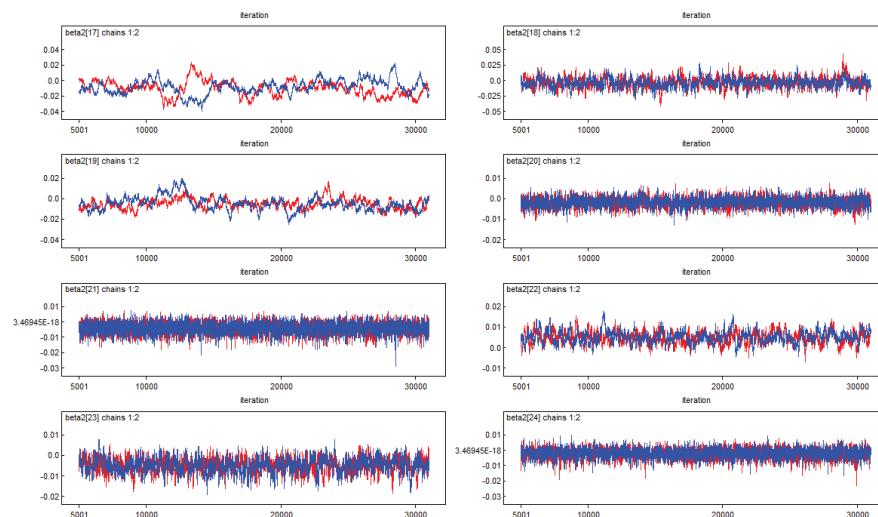
En esta sección seguiremos de cerca la convergencia para el coeficiente  $\beta_t^2$  que es el que acompaña a la variable precipitación, organizado en los intervalos de tiempo que se muestran a continuación.



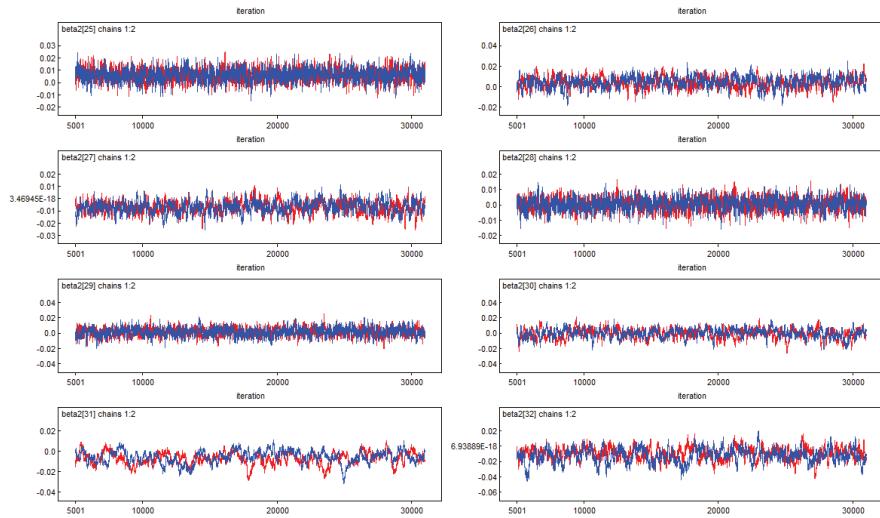
**Figura 5.13:** History plot para la variable  $\beta_t^2$ , desde enero 2010 - agosto 2010



**Figura 5.14:** History plot para la variable  $\beta_t^2$ , desde septiembre 2010 - abril 2011

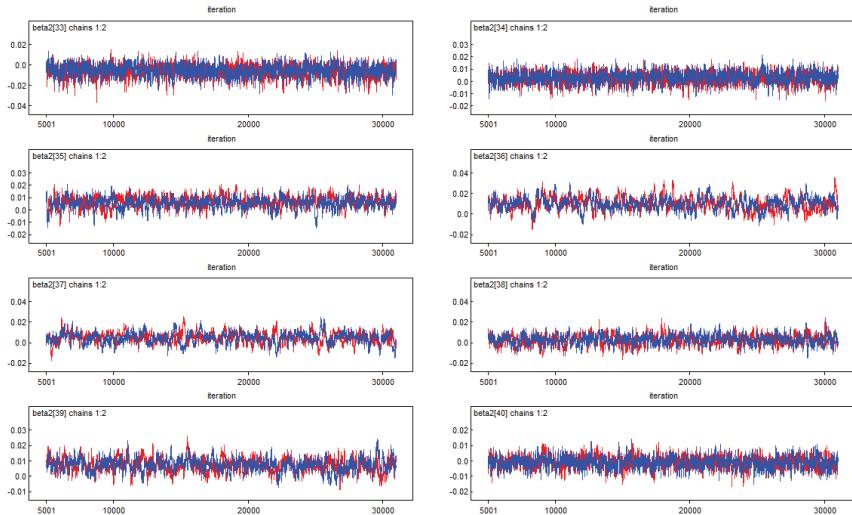


**Figura 5.15:** History plot para la variable  $\beta_t^2$ , desde mayo 2011 - diciembre 2011

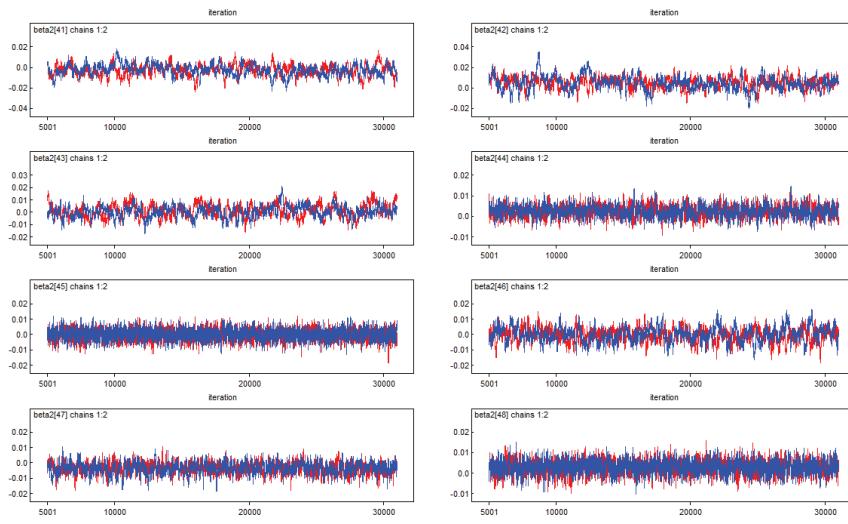


**Figura 5.16:** History plot para la variable  $\beta_t^2$ , desde enero 2012 - agosto 2012

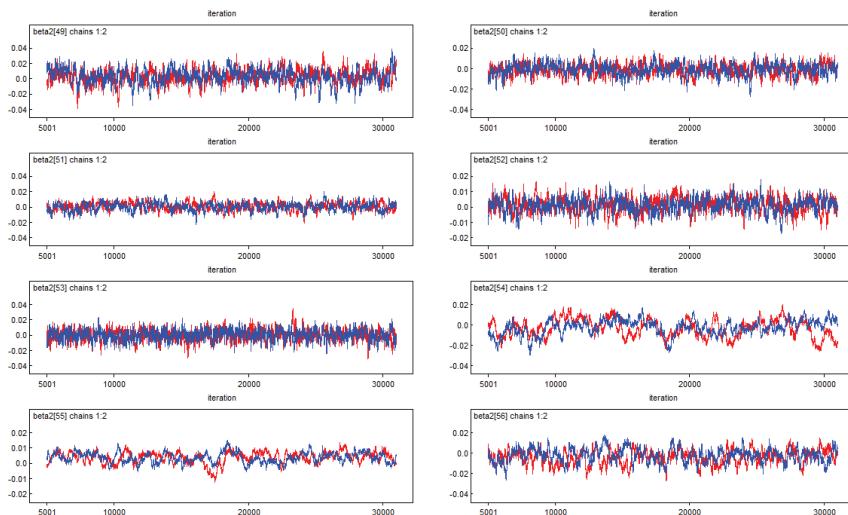
Al igual que el coeficiente anterior, miramos que las cadenas se mezclan en la mayoría de recuadros que hacen referencia al coeficiente  $\beta_t^2$  para cada mes, a excepción de algunas figuras en las que hay problemas en las iteraciones. Esto se comprueba con los histogramas cuando no decrecen y el estadístico de Gelman y Rubin sale del rango establecido, sin embargo, se acepta la convergencia para  $\beta_t^2$  pues hacemos uso del concepto de monitorear un subconjunto del parámetro.



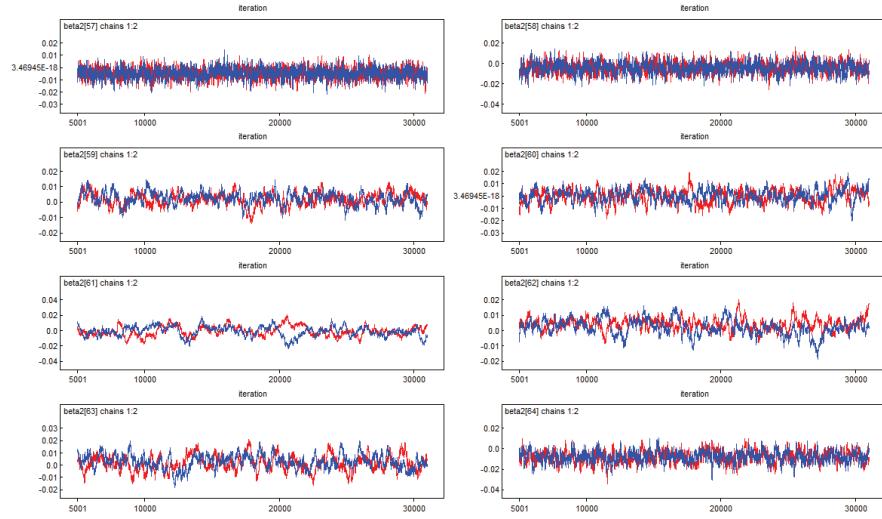
**Figura 5.17:** History plot para la variable  $\beta_t^2$ , desde septiembre 2012 - abril 2013



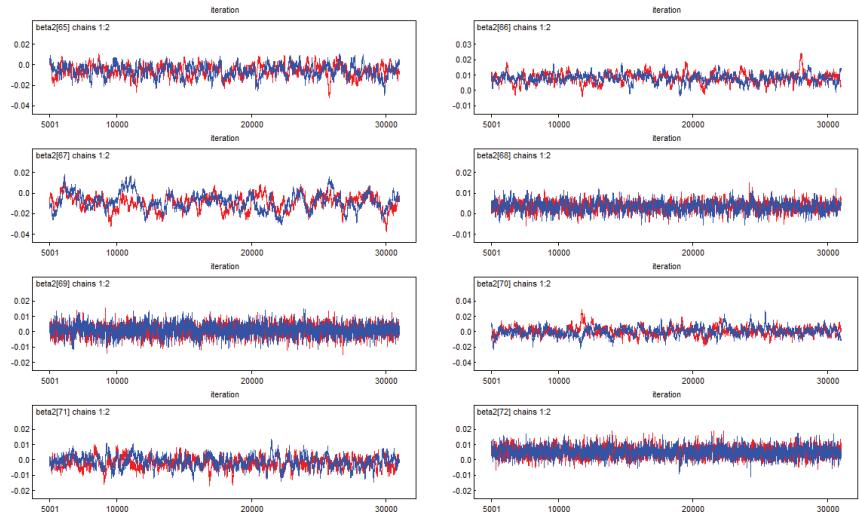
**Figura 5.18:** History plot para la variable  $\beta_t^2$ , desde mayo 2013 - diciembre 2013



**Figura 5.19:** History plot para la variable  $\beta_t^2$ , desde enero 2014 - agosto 2014

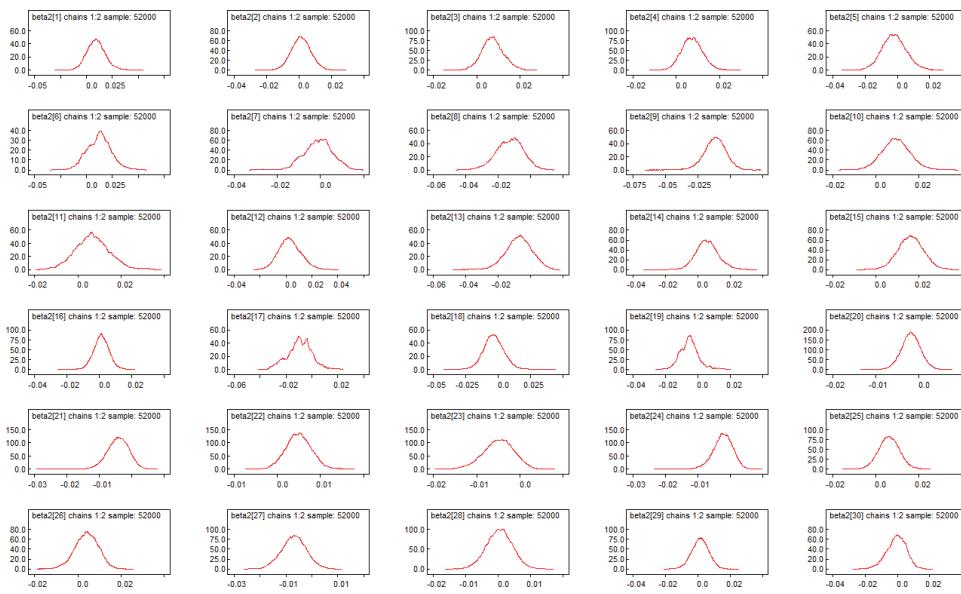


**Figura 5.20:** History plot para la variable  $\beta_t^2$ , desde septiembre 2014 - abril 2015

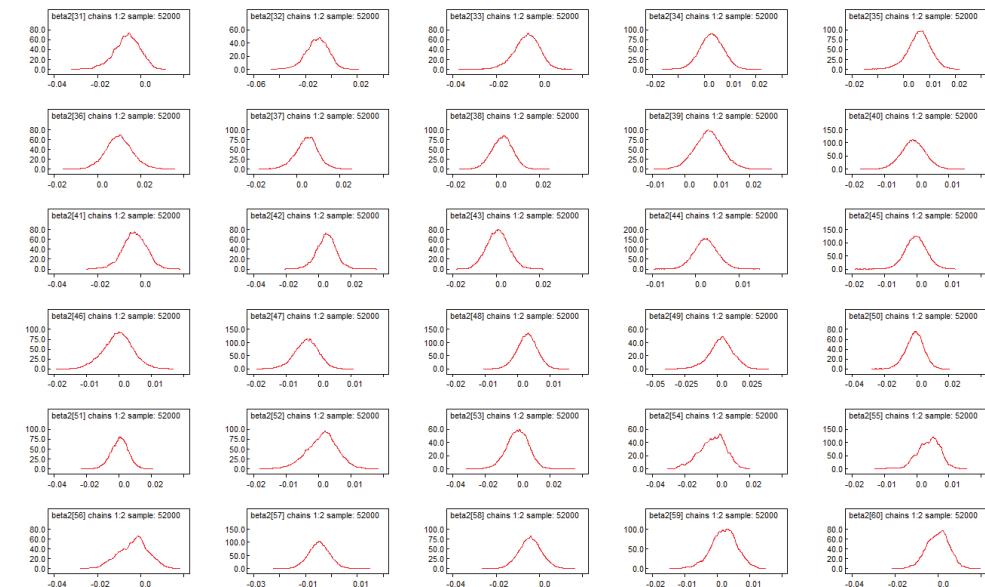


**Figura 5.21:** History plot para la variable  $\beta_t^2$ , desde mayo 2015 - diciembre 2015

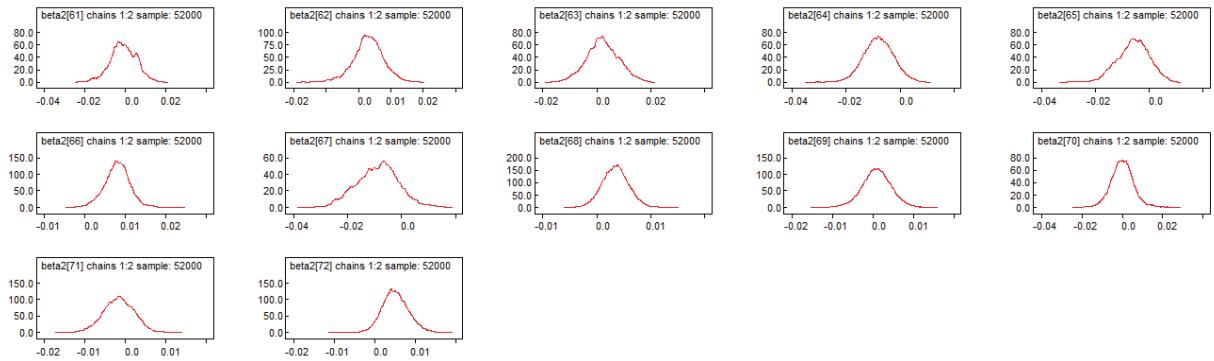
Para las gráficas (5.22), (5.23),y (5.24) observamos que algunas de las curvas mantienen la simetría y en las restantes existe la presencia de colas pronunciadas a la derecha o bien a la izquierda, como sucede en aquellos meses que presentaron distorsiones en los picos de estas curvas.



**Figura 5.22:** Gráfico de la densidad para la variable  $\beta_t^2$ , desde enero 2010- junio 2012



**Figura 5.23:** Gráfico de la densidad para la variable  $\beta_t^2$ , desde julio 2012-diciembre 2014



**Figura 5.24:** Gráfico de la densidad para la variable  $\beta_t^2$ , desde enero 2015- diciembre 2015

A continuación los estadísticos a posteriori se presentan de forma anual en las dos estaciones, además en cada tabla se cumple que el error MC es menor a la desviación estándar a posteriori, un claro signo de convergencia.

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^2$	0.00504	0.00676	0.00038	0.00499	[-0.0080766,0.018758]
Seca	$\beta^2$	0.00005	0.00815	0.00029	0.00337	[-0.00781225,0.0157275]

**Cuadro 5.8:** Estadísticos a posteriori para  $\beta^2$ , año 2010, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^2$	0.00274	0.00602	0.00035	0.00281	[-0.009384,0.015021]
Seca	$\beta^2$	-0.00312	0.00501	0.00043	0.00313	[-0.0106105,0.01634]

**Cuadro 5.9:** Estadísticos a posteriori para  $\beta^2$ , año 2011, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^2$	0.00257	0.00511	0.00027	0.00258	[-0.0074926,0.012807]
Seca	$\beta^2$	-0.00172	0.00559	0.00024	-0.00007	[-0.00979175,0.00920875]

**Cuadro 5.10:** Estadísticos a posteriori para  $\beta^2$ , año 2012, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^2$	0.00329	0.00414	0.00021	0.00337	[-0.00517702,0.0111136]
Seca	$\beta^2$	0.00023	0.00443	0.00024	0.00139	[-0.007489275,0.00986775]

**Cuadro 5.11:** Estadísticos a posteriori para  $\beta^2$ , año 2013, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	median	IC
Lluviosa	$\beta^2$	0.00025	0.00645	0.00040	0.00057	[-0.0133552,0.0119076]
Seca	$\beta^2$	-0.00139	0.00544	0.00033	0.00015	[-0.0129665,0.0101345]

Cuadro 5.12: Estadísticos a posteriori para  $\beta^2$ , año 2014, estaciones lluviosa y seca

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\beta^2$	0.00025	0.00503	0.00032	0.00024	[-0.0099448,0.0098704]
Seca	$\beta^2$	-0.00070	0.00466	0.00039	-0.00210	[-0.01383025,0.0089825]

Cuadro 5.13: Estadísticos a posteriori para  $\beta^2$ , año 2015, estaciones lluviosa y seca

### 5.5.3. Convergencia para el parámetro $\gamma_s^1$

Para el coeficiente  $\gamma_s^1$ , tenemos a simple vista la convergencia absoluta solo observando la mezcla perfecta de las cadenas en este intervalo de iteraciones. Además con la ejecución del histograma las diferentes clases desaparecían conforme las iteraciones aumentaban en número y en el test de Gelman y Rubin en todos los circuitos el rango estuvo entre 0.95 y 1.05, lo que demuestra una correcta convergencia, que también se ve reflejada en las curvas simétricas de (5.31) y (5.32) para cada uno de los 48 circuitos.

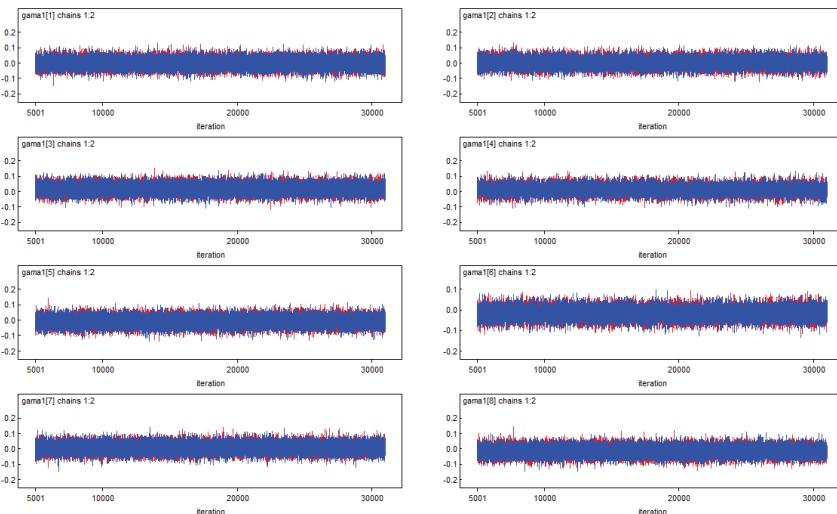
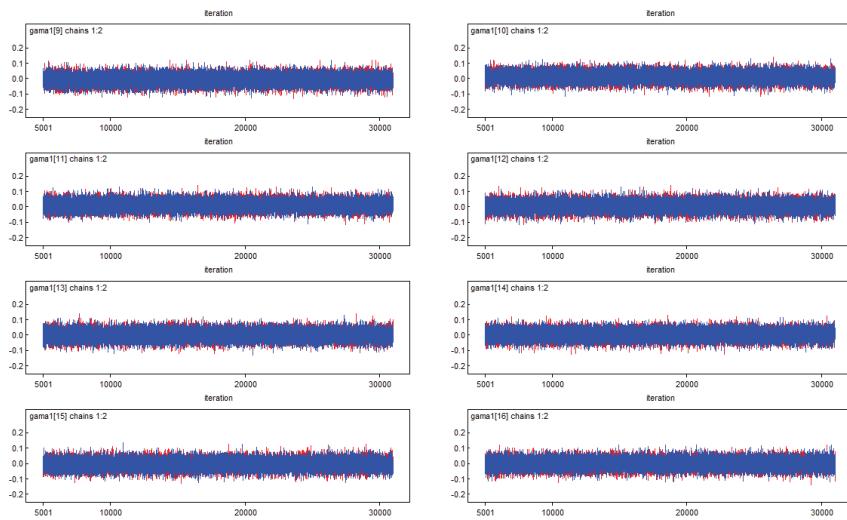
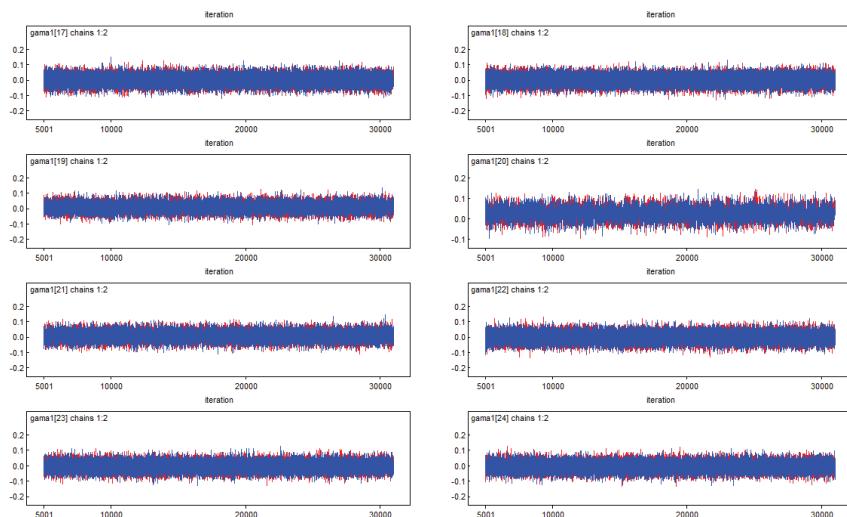


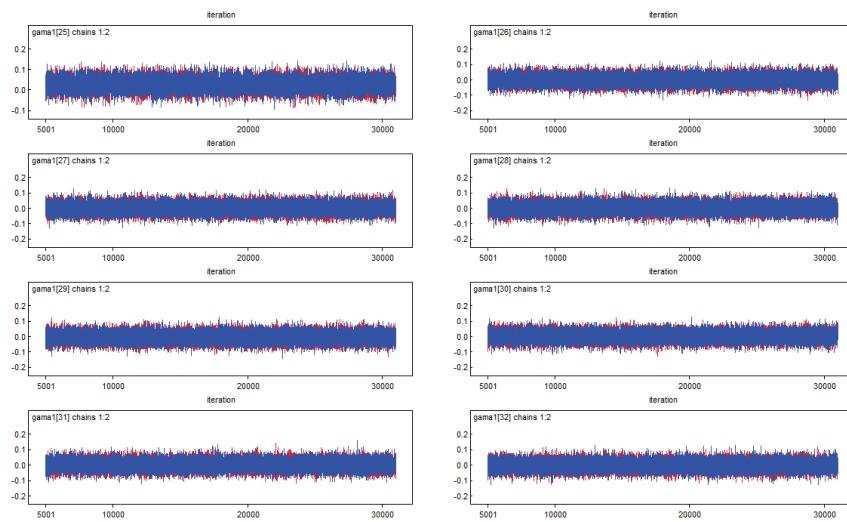
Figura 5.25: History plot para la variable  $\gamma_s^1$  para los circuitos: (1-8)



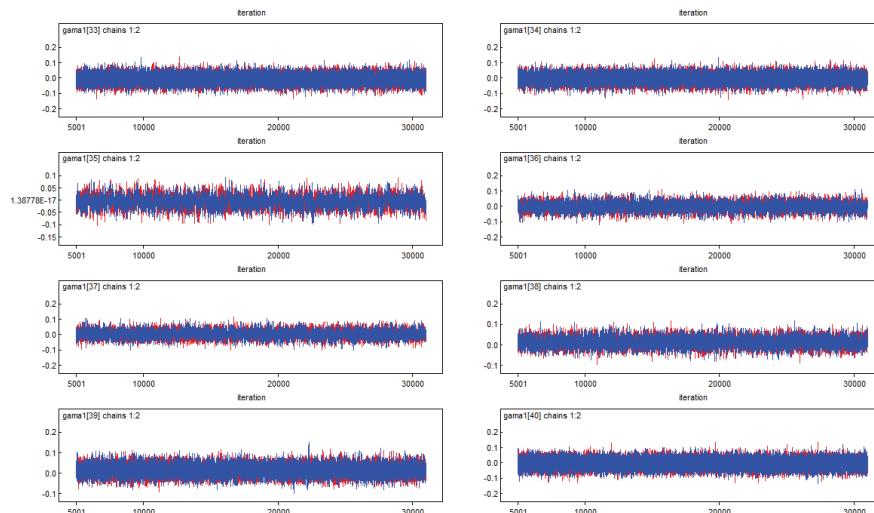
**Figura 5.26:** History plot para la variable  $\gamma_s^1$  para los circuitos: (9-16)



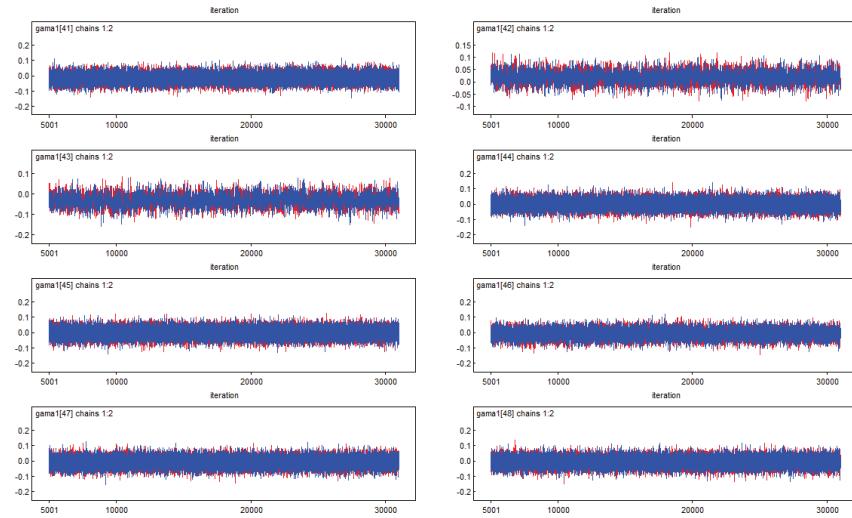
**Figura 5.27:** History plot para la variable  $\gamma_s^1$  para los circuitos: (17-24)



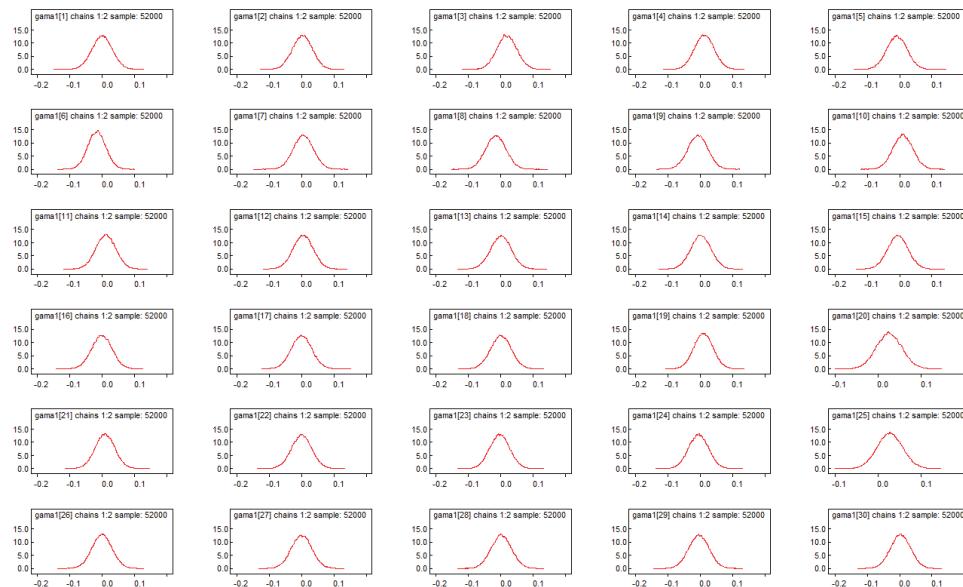
**Figura 5.28:** History plot para la variable  $\gamma_s^1$  para los circuitos: (25-32)



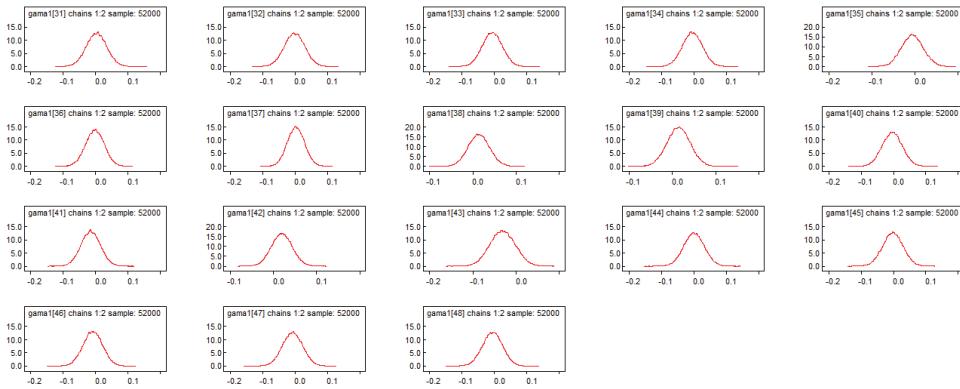
**Figura 5.29:** History plot para la variable  $\gamma_s^1$  para los circuitos:(33-39)



**Figura 5.30:** History plot para la variable  $\gamma_s^1$  para los circuitos: (40-48)



**Figura 5.31:** Gráfico de la densidad para la variable  $\gamma_s^1$ , para los circuitos: (1-24)



**Figura 5.32:** Gráfico de la densidad para la variable  $\gamma_s^1$ , para los circuitos: (25-48)

Circuitos	Media	Desv. Est.	MC error	Mediana	IC
09D01C01	-0.000120	0.031760	0.000292	-0.000066	[-0.06192,0.06228]
09D01C02	-0.002098	0.031800	0.000218	-0.002126	[-0.06386,0.06037]
09D01C03	-0.000259	0.031570	0.000211	-0.000237	[-0.06229,0.06121]
09D01C04	0.002822	0.030760	0.000231	0.002721	[-0.05714,0.06321]
09D01C05	-0.008670	0.030080	0.000360	-0.008623	[-0.06818,0.04951]
09D02C01	-0.007838	0.031090	0.000213	-0.007781	[-0.0679,0.05327]
09D02C02	0.018460	0.031230	0.000290	0.018270	[-0.04261,0.08003]
09D02C03	0.012120	0.030920	0.000449	0.012330	[-0.0492,0.07217]
09D02C04	-0.008248	0.031450	0.000231	-0.008605	[-0.06999,0.05372]
09D02C05	-0.016640	0.027800	0.000325	-0.016720	[-0.07092,0.03807]
09D03C01	0.005967	0.031180	0.000223	0.005984	[-0.05534,0.06729]
09D03C02	-0.015000	0.031230	0.000252	-0.015270	[-0.07666,0.04602]
09D03C03	-0.007451	0.030960	0.000212	-0.007909	[-0.06807,0.05297]
09D03C04	0.010060	0.030540	0.000293	0.010160	[-0.05014,0.07064]
09D03C05	-0.006331	0.031120	0.000258	-0.006345	[-0.06743,0.05459]
09D03C06	0.011160	0.031120	0.000283	0.011330	[-0.05037,0.0724]
09D03C07	0.004892	0.031270	0.000220	0.004797	[-0.05654,0.06645]
09D04C01	-0.000009	0.031770	0.000189	0.000211	[-0.06234,0.06201]
09D04C02	0.001945	0.031130	0.000222	0.001872	[-0.05933,0.0627]
09D04C03	-0.005250	0.031510	0.000241	-0.005502	[-0.06695,0.05789]
09D04C04	-0.001435	0.031580	0.000200	-0.001258	[-0.06324,0.06044]
09D04C05	-0.000018	0.031480	0.000198	0.000011	[-0.06185,0.06117]
09D04C06	0.000151	0.031650	0.000212	0.000070	[-0.06202,0.06249]
09D05C01	0.010150	0.029290	0.000375	0.010280	[-0.04715,0.06718]

**Cuadro 5.14:** Estadísticos a posteriori para  $\gamma_1$  para los circuitos: (1-24)

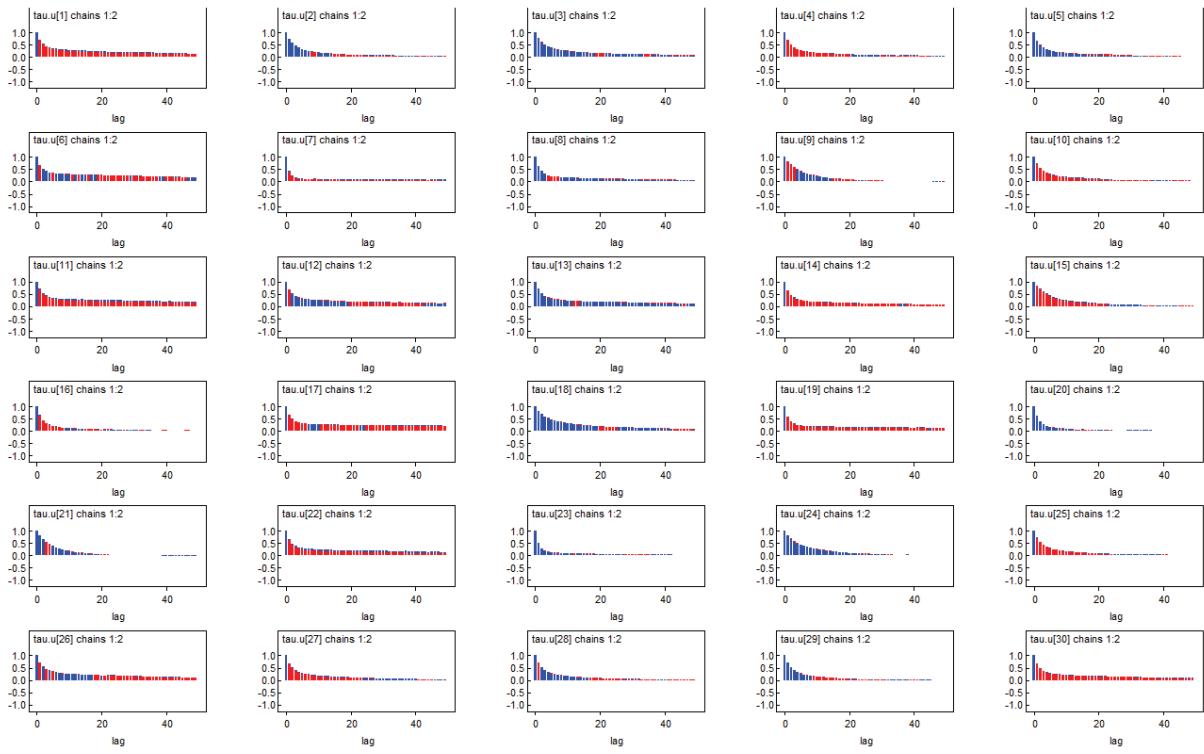
Circuitos	Media	Desv. Est.	MC error	Mediana	IC
09D05C02	0.026130	0.029350	0.000477	0.025860	[-0.03073,0.08484]
09D05C03	0.008791	0.030360	0.000336	0.008591	[-0.05129,0.0686]
09D05C04	-0.014650	0.029960	0.000291	-0.014740	[-0.07378,0.04432]
09D05C05	-0.032240	0.029000	0.000844	-0.031970	[-0.09017,0.02518]
09D05C06	-0.003414	0.030920	0.000237	-0.003473	[-0.06367,0.05735]
09D06C01	-0.000215	0.031580	0.000196	-0.000084	[-0.06166,0.06178]
09D06C02	-0.004020	0.031340	0.000246	-0.003922	[-0.06584,0.05723]
09D06C03	-0.000281	0.031180	0.000265	-0.000300	[-0.06171,0.06004]
09D06C04	0.000266	0.031430	0.000249	0.000542	[-0.06159,0.06179]
09D06C05	0.000714	0.031520	0.000191	0.000781	[-0.06123,0.06262]
09D06C06	-0.004007	0.031640	0.000207	-0.003839	[-0.06578,0.05766]
09D07C01	0.003050	0.031190	0.000245	0.003222	[-0.05833,0.06374]
09D07C02	0.003036	0.031410	0.000188	0.002928	[-0.05843,0.0644]
09D07C03	-0.001611	0.031210	0.000199	-0.001636	[-0.06309,0.05946]
09D07C04	-0.006802	0.030900	0.000245	-0.006872	[-0.06836,0.05375]
09D07C05	-0.004787	0.030880	0.000211	-0.004742	[-0.06515,0.05658]
09D07C06	-0.004515	0.030960	0.000330	-0.004701	[-0.06481,0.05688]
09D08C01	-0.005864	0.025030	0.000664	-0.005330	[-0.05553,0.04215]
09D08C02	-0.002038	0.028820	0.000586	-0.002079	[-0.05856,0.05503]
09D08C03	0.003026	0.026170	0.000546	0.003113	[-0.04861,0.05437]
09D08C04	0.020520	0.023610	0.000584	0.020250	[-0.02555,0.06665]
09D08C05	0.015130	0.024540	0.000452	0.014940	[-0.0325,0.06365]
09D08C06	0.014690	0.026540	0.000349	0.014610	[-0.03687,0.0662]
09D09C01	0.027300	0.029880	0.000315	0.027280	[-0.03069,0.08596]

**Cuadro 5.15:** Estadísticos a posteriori para  $\gamma_1$  para los circuitos: 25-48

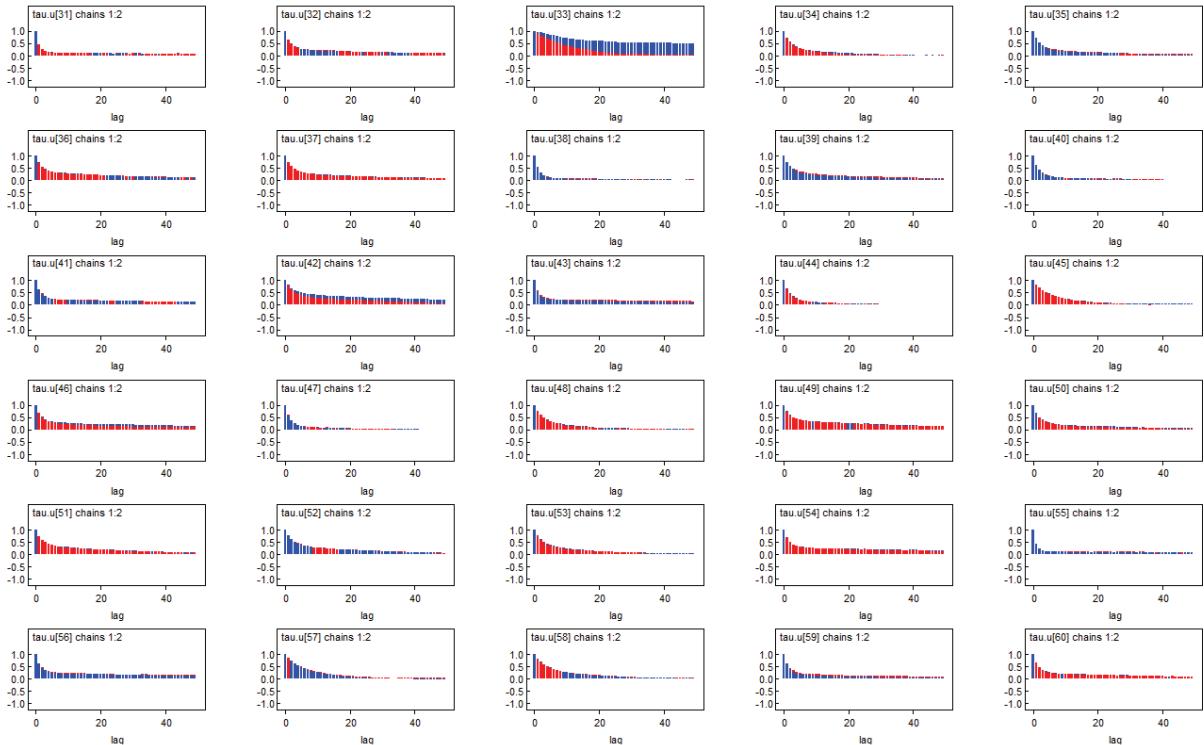
En estas dos tablas nuevamente comprobamos que el error MC es menor a la desviación estándar a posteriori, que es lo que se quería para concluir convergencia.

#### 5.5.4. Convergencia para $\tau_\mu$

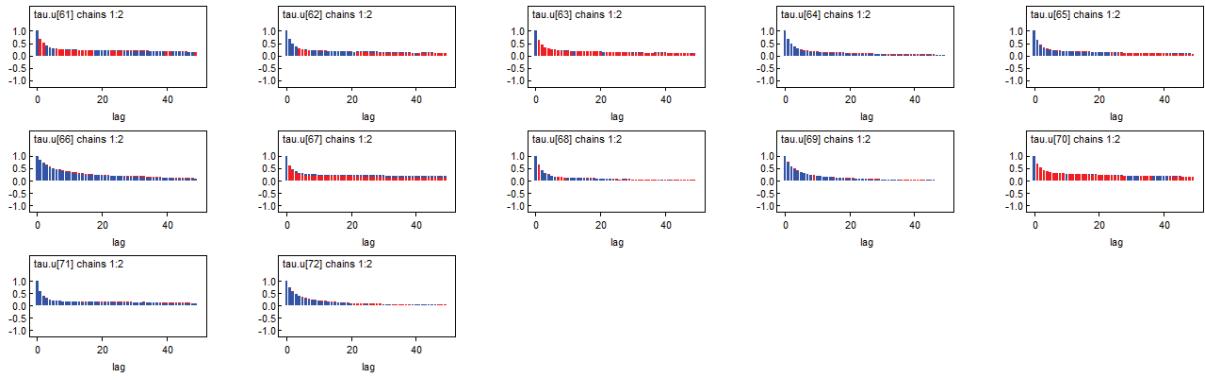
En el caso de la convergencia para las hiper prioris mostramos únicamente los gráficos de autocorrelación donde el histograma tiende a volverse más pequeño y casi imperceptible en la mayoría de casos. En cuanto al Gelman-Rubin se mantuvo el intervalo de 5 % alrededor de 1, lo que asegura la convergencia.



**Figura 5.33:** Gráfico de autocorrelación para la variable  $\tau_\mu$  para los meses: 1 al 30



**Figura 5.34:** Gráfico de autocorrelación para la variable  $\tau_\mu$  para los meses: 31 al 60



**Figura 5.35:** Gráfico de autocorrelación para la variable  $\tau_\mu$  para los meses: 61 al 72

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\tau_\mu$	0.68312	0.28000	0.00705	0.63352	[0.28728, 1.36462]
Seca	$\tau_\mu$	0.41193	0.15909	0.00809	0.74735	[0.33955, 1.611725]

**Cuadro 5.16:** Estadísticos a posteriori para  $\tau_\mu$ , año 2010

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\tau_\mu$	0.71638	0.31428	0.00732	0.65856	[0.29012, 1.46702]
Seca	$\tau_\mu$	0.64103	0.25775	0.00636	0.60973	[0.28465, 1.2935]

**Cuadro 5.17:** Estadísticos a posteriori para  $\tau_\mu$ , año 2011

En las tablas (5.16), (5.17), (5.18), (5.19), (5.20) y (5.21) se verifica que el error MC no supera a la desviación estándar a posteriori, para los años 2010 al 2015 distribuido en las épocas lluviosa y seca.

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\tau_\mu$	0.66406	0.25970	0.00627	0.62316	[0.2816, 1.28934]
Seca	$\tau_\mu$	0.56951	0.27336	0.00651	0.70543	[0.32395, 1.458975]

**Cuadro 5.18:** Estadísticos a posteriori para  $\tau_\mu$ , año 2012

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\tau_\mu$	0.69648	0.27054	0.00594	0.65256	[0.30424, 1.3505]
Seca	$\tau_\mu$	0.60166	0.25121	0.00513	0.59223	[0.2922, 1.161475]

**Cuadro 5.19:** Estadísticos a posteriori para  $\tau_\mu$ , año 2013

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\tau_\mu$	0.48684	0.19706	0.00563	0.45474	[0.200192,0.96098]
Seca	$\tau_\mu$	0.60963	0.27684	0.00672	0.54100	[0.2341,1.171325]

**Cuadro 5.20:** Estadísticos a posteriori para  $\tau_\mu$ , año 2014

Época	Parámetro	Media	Desv. Est	MC error	Mediana	IC
Lluviosa	$\tau_\mu$	0.46348	0.17536	0.00444	0.43640	[0.20084,0.87954]
Seca	$\tau_\mu$	0.62480	0.25798	0.00403	0.40385	[0.1936,0.788225]

**Cuadro 5.21:** Estadísticos a posteriori para  $\tau_\mu$ , año 2015

Finalmente para el seguimiento de la componente  $\mu$  se realizaron las pruebas del estadístico Gelman-Rubin y los resultados estuvieron entre el intervalo [0.95,1.05] que aseguran la convergencia, pero no se incluyeron los gráficos debido a que la convergencia era obvia.

## 5.6. Tablas y Mapas de Riesgo para el período 2010-2015

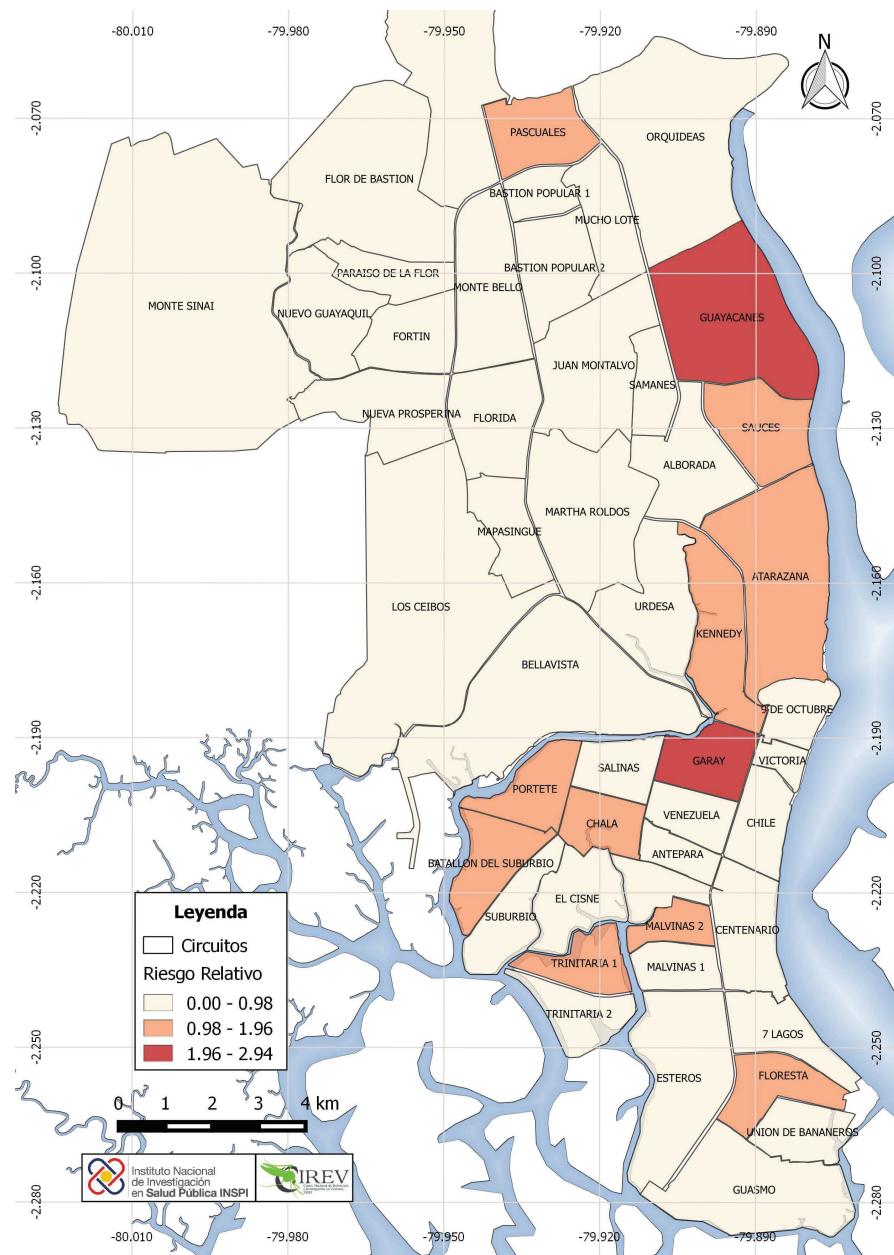
Para esta parte presentamos para cada año y por estaciones: lluviosa y seca las tablas de Riesgo Relativo (RR) calculadas con la ecuación del modelo que incorpora las variables climáticas y las condiciones de salubridad, para esto se hacen explícitos los códigos de los circuitos ordenados numéricamente. El riesgo relativo correspondiente a cada uno y el intervalo de confianza conjuntamente con los mapas de riesgo clasificados en tres categorías: Alto, Medio, y Bajo.

Para la inclusión de estas tablas de riesgo por circuito, para cada uno de ellos se calculó el riesgo relativo por meses. Posteriormente se agruparon los meses de acuerdo a la estación lluviosa o seca y se obtuvo el valor promedio de esos valores tanto para el riesgo relativo  $RR$  como para los intervalos de confianza procedentes del software WinBUGS14.

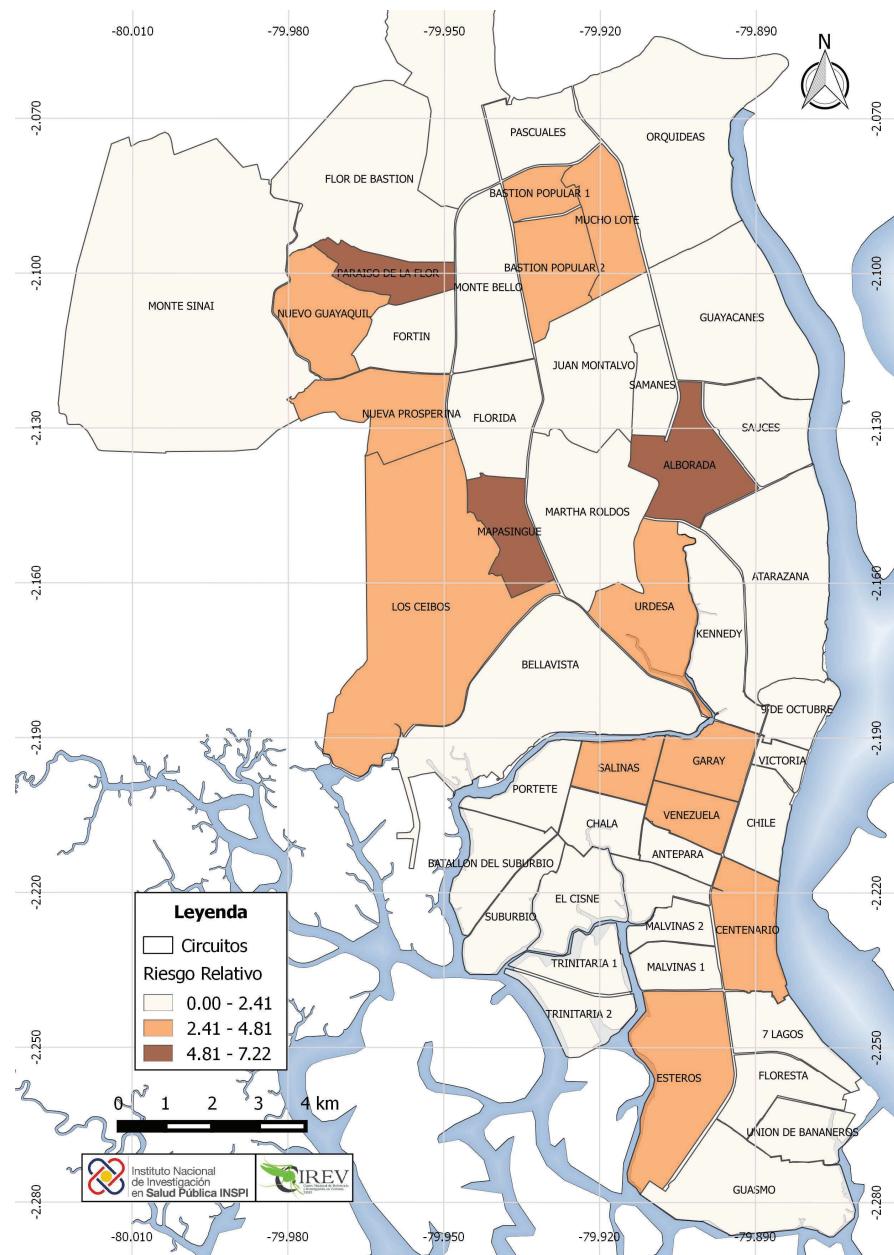
### 5.6.1. Riesgo Relativo para el año 2010

Circuitos	2010						Circuitos	2010						
	Época Lluviosa			Época Seca				Época Lluviosa			Época Seca			
	RR	IC		RR	IC			RR	IC		RR	IC		
		0.25	0.95		0.25	0.95			0.25	0.95		0.25	0.95	
09D01C01	0.41	0.28	0.54	2.06	1.47	2.28	09D05C02	1.43	0.76	2.00	1.01	0.65	2.04	
09D01C02	0.35	0.22	0.61	1.52	0.84	2.39	09D05C03	2.16	1.08	3.90	0.76	0.14	2.10	
09D01C03	1.84	0.14	3.82	1.34	0.02	4.91	09D05C04	0.23	0.01	0.80	5.55	1.32	7.74	
09D01C04	0.84	0.14	2.69	1.40	0.49	4.29	09D05C05	0.25	0.12	0.49	2.99	1.45	3.75	
09D01C05	0.45	0.12	0.84	4.32	2.65	6.94	09D05C06	1.62	0.46	5.48	1.97	0.44	8.28	
09D02C01	0.95	0.45	3.43	3.74	1.00	8.64	09D06C01	0.69	0.33	1.84	0.52	0.07	2.30	
09D02C02	0.91	0.28	1.56	0.65	0.14	1.89	09D06C02	0.11	0.01	0.84	0.22	0.05	1.64	
09D02C03	1.24	0.66	1.70	1.46	0.51	2.17	09D06C03	0.58	0.09	1.34	0.26	0.04	1.38	
09D02C04	1.11	0.54	2.82	1.15	0.31	2.40	09D06C04	0.23	0.10	0.81	0.63	0.09	1.42	
09D02C05	0.73	0.14	1.81	0.96	0.24	3.69	09D06C05	0.38	0.08	0.67	0.64	0.11	1.54	
09D03C01	0.72	0.02	1.88	0.51	0.05	4.67	09D06C06	0.40	0.02	0.59	5.26	3.10	8.04	
09D03C02	0.41	0.01	1.60	0.53	0.02	3.39	09D07C01	0.27	0.02	0.93	0.34	0.02	1.29	
09D03C03	0.64	0.02	1.52	0.69	0.04	4.36	09D07C02	0.55	0.02	1.68	2.50	0.82	9.16	
09D03C04	2.94	0.73	3.79	2.52	0.79	7.51	09D07C03	0.80	0.04	1.97	2.80	0.34	4.19	
09D03C05	0.30	0.02	1.00	4.33	2.59	9.38	09D07C04	0.19	0.05	1.55	2.64	1.01	6.99	
09D03C06	0.72	0.35	2.25	0.86	0.06	1.86	09D07C05	0.37	0.02	2.40	2.22	0.63	10.06	
09D03C07	0.45	0.30	1.34	1.89	0.82	2.66	09D07C06	1.03	0.26	1.87	0.38	0.03	1.11	
09D04C01	0.46	0.10	1.28	0.59	0.23	3.07	09D08C01	0.38	0.22	0.82	3.97	2.37	5.26	
09D04C02	1.57	0.87	3.49	1.00	0.24	3.05	09D08C02	0.30	0.19	0.95	0.48	0.12	0.97	
09D04C03	0.67	0.06	2.36	4.32	1.65	8.95	09D08C03	0.58	0.22	1.01	7.22	3.45	9.32	
09D04C04	1.91	0.08	3.46	0.39	0.03	2.25	09D08C04	0.28	0.04	1.03	0.55	0.08	1.51	
09D04C05	1.22	0.24	4.13	2.15	0.08	4.89	09D08C05	0.47	0.05	0.95	0.32	0.12	1.10	
09D04C06	0.88	0.10	2.03	0.91	0.25	2.77	09D08C06	0.96	0.17	2.24	3.73	1.39	8.46	
09D05C01	1.71	0.62	2.32	1.26	0.14	1.73	09D09C01	0.29	0.02	1.08	4.25	1.30	8.45	

**Cuadro 5.22:** Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2010, RR=riesgo relativo, IC= Intervalo de confianza



**Figura 5.36:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2010, época lluviosa



**Figura 5.37:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2010, época seca

En el año 2010 los sectores Garay y Guayacanes pertenecen al nivel alto de riesgo superando un valor de 2, en tanto en la época seca tenemos 3 localidades : Mapasingue, Paraíso de la Flor y Alborada, que en conjunto superan un valor de 5.2 para la escala de riesgo.

Pese a que la distribución de los casos es casi nula en los primeros meses, influye cierto desfase que se encuentra acumulado hasta en un mes posterior, además que el número de casos esperados tiene una dispersión parecida en todos los meses y con

especial atención en marzo y abril donde la temperatura toma valores alrededor de 24 grados Celsius y la precipitación se concentra en rangos de 62 a 83 mm.

En cuanto a la influencia de los indicadores que están condensados en  $Z^1$ , tiene un peso sustancial las similitudes en los aspectos de desempleo, hacinamiento, analfabetismo, exceptuando el caso del porcentaje de personas sin acceso a un buen sistema de alcantarillado que reflejan de 3 a 9 enfermos esperados. Por último en el estrato de riesgo bajo están ubicados los 36 circuitos restantes en los que se puede observar que presentan un rango entre 0.30 y 0.91 debido a que el registro de los casos es casi nulo.

Para ese mismo año en la época seca se han registrado en el primer nivel a: Mapasique, Paraíso de la Flor y Alborada, en la que la mayor dispersión de casos está concentrada en los meses de mayo, junio y julio, tiempo en el que inicia la época seca y se extiende hasta noviembre, esperando de 3 a 15 casos en esos meses.

Observaremos también que en estos sectores empieza un descenso de temperatura desde mayo a agosto, pero aún con esta temperatura los hospederos de los mosquitos pueden eclosionar. Desde septiembre a noviembre encontramos otro descenso en la temperatura de aproximadamente 2 grados. De igual forma se relaciona con la precipitación que toma un rango de 87 a 221 mm en los meses de mayo a julio, porque posterior a ello existe un descenso de lluvia hasta el mes de diciembre. Notemos que los 12 circuitos del nivel medio también comparten estas características dado que son los sectores aledaños a estos puntos de riesgo, en cambio para las 33 localidades de riesgo bajo el conteo de casos es casi nulo.

En cuanto a los indicadores sociales para esos sectores en el riesgo alto las características del desempleo tienen un rango entre 7 y 9 % , y especialmente en el caso de Paraíso de la Flor donde existe hasta un 90 % de personas que no tienen acceso al agua potable y 99 % de personas que no tienen acceso al alcantarillado.

Para el riesgo medio los 12 circuitos que conforman este estrato forman clústers en las zonas norte, centro y sur de la ciudad de Guayaquil, especialmente donde los indicadores sociales más importantes que afectan al reporte de casos tiene que ver con el porcentaje de personas sin acceso al agua corriente, el alcantarillado y la recolección de la basura, junto con la razón persona dormitorio que está casi al límite de 3 e indica una posición de hacinamiento.

## 5.6.2. Riesgo Relativo para el año 2011

Circuitos	2011						Circuitos	2011						
	Epoca Lluviosa			Epoca Seca				Epoca Lluviosa			Epoca Seca			
	RR	IC		RR	IC			RR	IC		RR	IC		
		0.25	0.95		0.25	0.95			0.25	0.95		0.25	0.95	
09D01C01	0.73	0.56	0.91	3.46	3.18	3.95	09D05C02	1.96	0.93	2.69	0.94	0.29	1.77	
09D01C02	0.15	0.06	0.39	0.67	0.21	1.76	09D05C03	0.20	0.02	0.57	0.46	0.09	1.30	
09D01C03	0.33	0.08	2.87	0.99	0.15	4.64	09D05C04	0.14	0.02	0.78	1.39	0.59	3.70	
09D01C04	0.50	0.03	1.52	0.46	0.02	1.76	09D05C05	0.04	0.01	0.17	1.67	1.05	2.54	
09D01C05	0.86	0.46	1.45	2.26	1.23	3.11	09D05C06	0.25	0.04	2.24	1.22	0.02	2.57	
09D02C01	0.61	0.01	1.12	0.18	0.03	1.97	09D06C01	0.77	0.08	1.54	1.14	0.17	2.04	
09D02C02	0.80	0.32	1.60	1.21	0.43	2.25	09D06C02	0.49	0.15	1.67	4.58	2.21	6.81	
09D02C03	0.92	0.35	1.21	0.44	0.20	0.95	09D06C03	0.46	0.20	1.41	3.59	1.94	5.74	
09D02C04	1.05	0.44	2.14	0.78	0.26	2.63	09D06C04	0.47	0.10	0.85	2.06	1.24	3.22	
09D02C05	0.61	0.11	1.76	0.80	0.13	1.93	09D06C05	0.51	0.19	0.97	2.65	1.96	4.10	
09D03C01	0.91	0.07	2.79	1.75	0.47	5.78	09D06C06	0.37	0.03	0.56	0.52	0.11	1.50	
09D03C02	0.27	0.03	2.16	1.01	0.02	2.32	09D07C01	0.94	0.38	1.98	3.26	1.95	5.65	
09D03C03	0.56	0.02	1.78	1.56	0.03	2.89	09D07C02	0.41	0.03	1.73	2.13	0.07	3.46	
09D03C04	1.28	0.49	2.74	0.38	0.03	1.59	09D07C03	0.17	0.04	1.25	3.88	1.26	7.07	
09D03C05	0.50	0.08	1.59	1.33	0.21	2.61	09D07C04	0.23	0.02	0.97	0.86	0.23	2.98	
09D03C06	0.52	0.15	1.44	0.29	0.00	1.78	09D07C05	0.26	0.03	2.24	1.82	0.05	3.93	
09D03C07	0.50	0.28	1.29	0.49	0.25	1.30	09D07C06	0.90	0.34	1.49	1.86	1.07	3.69	
09D04C01	0.97	0.20	1.63	0.38	0.00	1.97	09D08C01	0.23	0.05	0.42	1.80	1.12	2.50	
09D04C02	1.58	0.74	3.54	0.32	0.02	1.21	09D08C02	0.62	0.31	0.92	2.63	2.13	3.70	
09D04C03	0.30	0.02	1.38	1.08	0.04	3.06	09D08C03	0.26	0.08	0.65	1.24	0.68	2.10	
09D04C04	0.83	0.09	2.83	0.61	0.05	2.03	09D08C04	0.35	0.05	1.14	4.61	2.52	6.51	
09D04C05	0.69	0.06	2.59	0.34	0.04	2.43	09D08C05	0.41	0.17	0.99	3.02	1.87	4.05	
09D04C06	2.02	0.67	4.12	0.32	0.02	3.54	09D08C06	0.22	0.02	1.18	1.07	0.13	2.47	
09D05C01	0.87	0.24	1.29	0.48	0.04	1.39	09D09C01	0.24	0.01	0.91	0.64	0.38	3.57	

**Cuadro 5.23:** Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2011, RR=Riesgo Relativo, IC=Intervalo de confianza

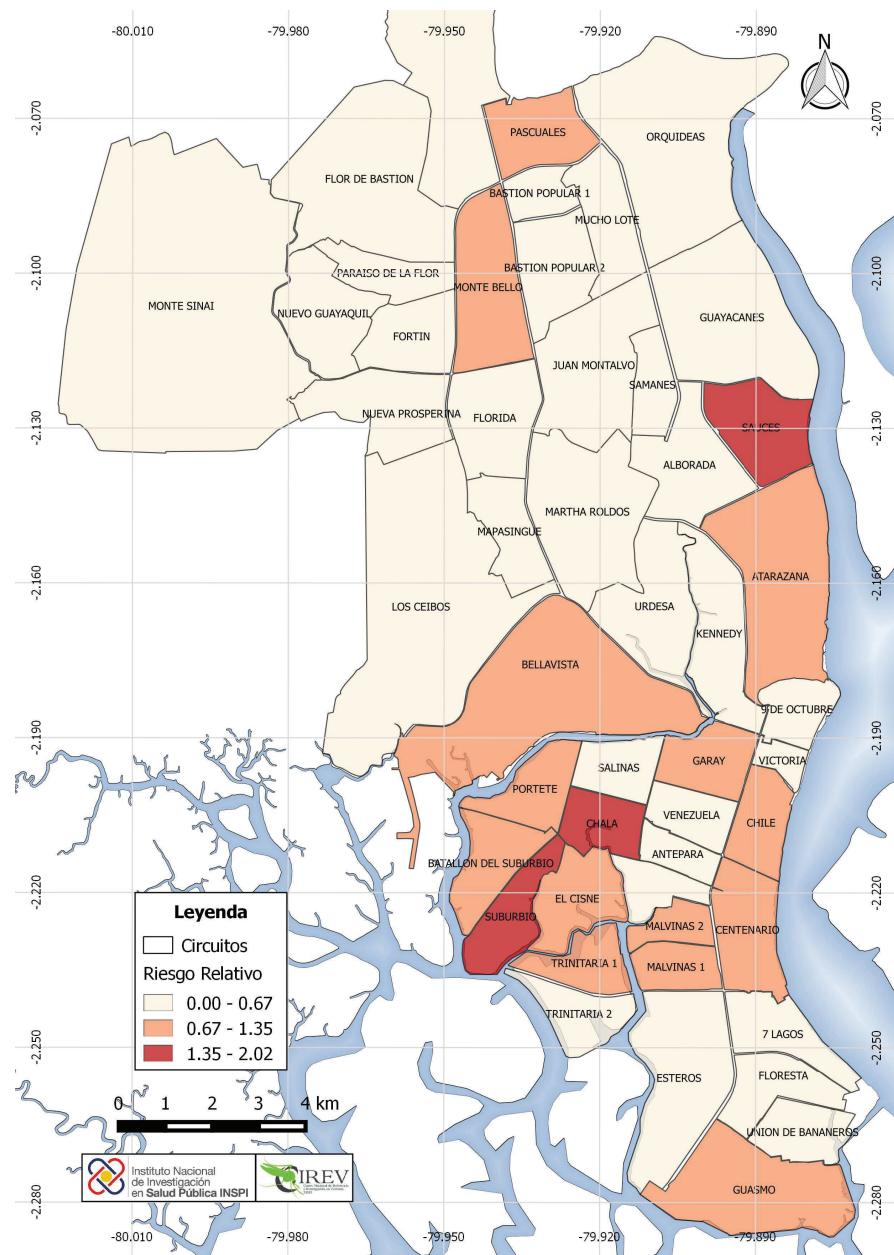
En la época lluviosa para el año 2011 el intervalo de riesgo alto oscila entre 1.58 y 2.02, el medio de 0.69 a 1.28 y el bajo entre 0.04 a 0.62, mientras en la época seca tenemos los intervalos de 3.26 a 4.61 para el alto, en el medio 1.56 a 3.02 y finalmente en el bajo llega a valores de 0.18

Respecto del esparcimiento de los casos para el primer nivel se reportan en mayor cantidad en los meses de marzo y abril (desde 1 hasta 12). Una observación

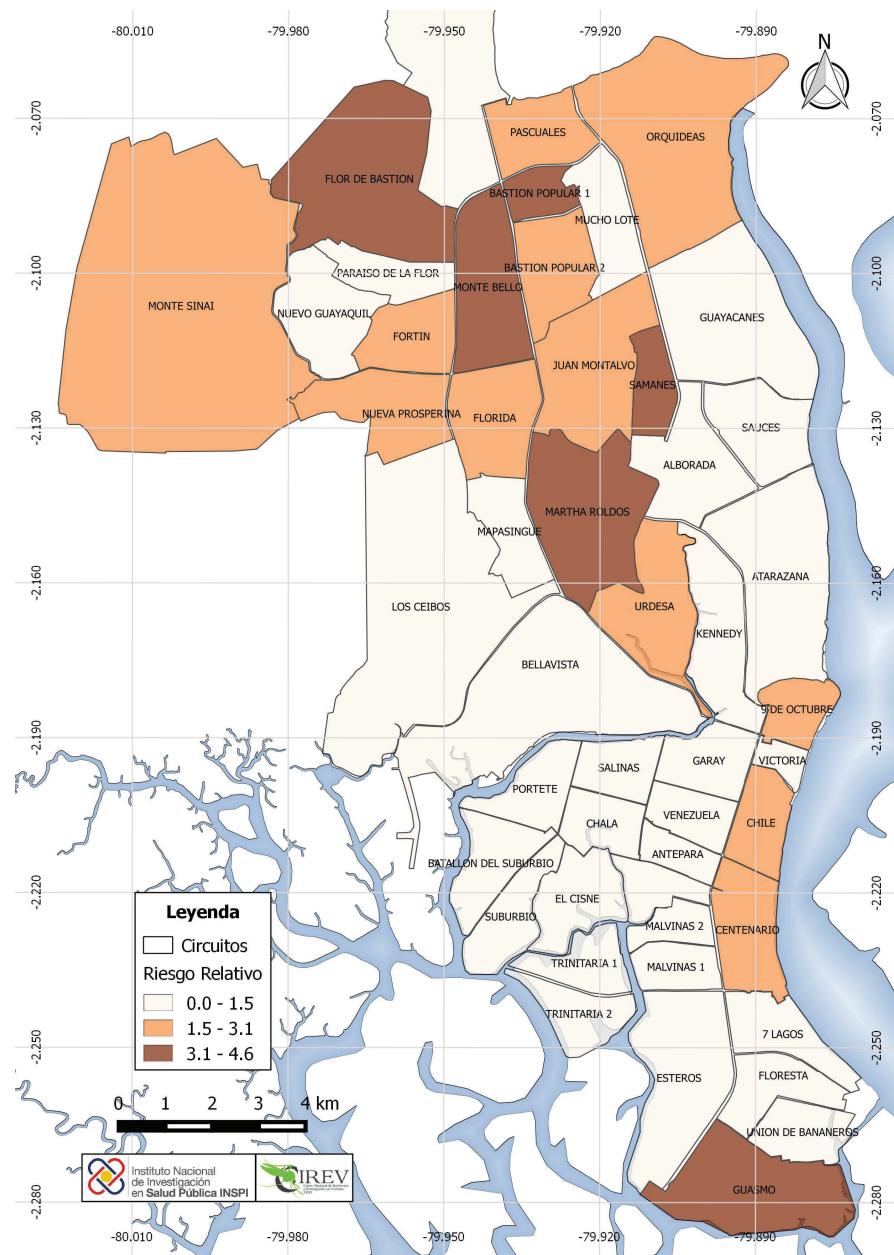
importante es que a pesar de que existen sectores con mayor porcentaje de necesidades, vemos que la dispersión de los casos y sobretodo la nulidad en los primeros meses del año son situaciones que influyen en el momento del cálculo del riesgo.

Para el nivel de riesgo medio en comparación con el 2010, en el segundo caso encontramos reportes desde febrero hasta abril que se contrastan con los meses donde se tiene igualdad con número de esperados y con los meses donde hay precipitaciones de hasta 70 mm.

Un escenario parecido a los dos años anteriores se manifiesta en algunos sectores como Nueva Prosperina, Fortín, Monte Sinaí, en donde hay más casos que en los circuitos del nivel de riesgo alto, sin embargo, en la clasificación a través del modelo elegido se toma en consideración la población por localidad y el descenso de los reportes de dengue en el resto del año.



**Figura 5.38:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2011, época lluviosa



**Figura 5.39:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2011, época seca

Enfocándonos en la época seca desde mayo hacia agosto tenemos la mayor concentración de reportes de dengue, asociados con el número esperado y a su vez registrados en el período en que la temperatura y la precipitación son más intensas que en los otros meses del período, pero no tan fuertes como en la época lluviosa.

Para el riesgo alto tenemos 6 circuitos en los cuales el porcentaje de personas sin acceso al alcantarillado varía desde el 30 % hasta aproximadamente un 97 %, el

59.26 % vive sin acceso a la recolección de basura y la razón persona-dormitorio supera el 3 extendiéndose hasta 5, lo que se considera un claro factor de hacinamiento.

Para el nivel que sigue la mayoría de enfermos de dengue están en mayo y junio donde tenemos un ligero aumento de temperatura y descenso de la precipitación. En tanto en el nivel bajo contamos con 30 circuitos en los que el porcentaje de personas con necesidades de desempleo, restricción al agua potable, el alcantarillado y la recolección de basura, excepto en el caso de Nuevo Guayaquil y Paraíso de la Flor que tienen estos índices elevados hasta un 99 %. Para los primeros los rangos demuestran que las condiciones de salubridad son de mediana condición.

### 5.6.3. Riesgo Relativo para el año 2012

Circuitos	2012						Circuitos	2012						
	Epoca Lluviosa			Epoca Seca				Epoca Lluviosa			Epoca Seca			
	RR	IC		RR	IC			RR	IC		RR	IC		
		0.25	0.95		0.25	0.95			0.25	0.95		0.25	0.95	
09D01C01	0.65	0.46	0.83	0.99	0.71	1.59	09D05C02	0.86	0.44	1.37	0.29	0.05	1.23	
09D01C02	0.35	0.25	0.76	5.53	2.47	6.32	09D05C03	0.88	0.55	1.93	1.46	0.32	2.61	
09D01C03	1.07	0.25	4.36	0.51	0.04	3.66	09D05C04	0.56	0.15	1.47	0.71	0.07	2.35	
09D01C04	0.68	0.20	1.51	1.23	0.43	2.70	09D05C05	0.24	0.09	0.49	0.79	0.54	1.44	
09D01C05	1.06	0.20	1.71	3.42	0.93	6.95	09D05C06	0.39	0.02	1.69	1.08	0.07	5.03	
09D02C01	0.66	0.02	1.35	6.20	2.47	11.68	09D06C01	1.42	0.36	4.47	1.75	0.20	5.85	
09D02C02	0.82	0.38	1.81	1.07	0.62	2.64	09D06C02	0.42	0.06	1.86	4.44	1.55	10.65	
09D02C03	2.19	1.18	3.06	1.48	0.54	2.31	09D06C03	0.48	0.18	1.28	1.47	0.63	3.56	
09D02C04	0.77	0.26	1.89	0.75	0.37	2.14	09D06C04	0.56	0.25	1.21	0.41	0.15	1.00	
09D02C05	0.38	0.13	2.23	2.33	0.09	3.81	09D06C05	0.33	0.03	0.93	9.57	4.56	13.03	
09D03C01	0.41	0.08	2.85	0.83	0.23	4.00	09D06C06	0.55	0.10	1.01	0.31	0.00	1.38	
09D03C02	0.59	0.05	8.35	0.59	0.01	7.51	09D07C01	0.26	0.02	1.42	4.93	2.24	15.40	
09D03C03	4.07	0.88	5.56	1.70	0.24	7.71	09D07C02	0.27	0.02	1.71	0.97	0.18	4.09	
09D03C04	1.51	0.73	2.63	1.03	0.11	2.07	09D07C03	0.83	0.02	1.67	2.51	0.67	9.30	
09D03C05	0.56	0.24	3.28	4.50	0.83	9.53	09D07C04	0.62	0.12	1.90	0.80	0.08	2.26	
09D03C06	2.06	0.92	3.71	1.02	0.29	2.46	09D07C05	2.35	0.13	5.52	0.94	0.02	6.00	
09D03C07	1.10	0.78	2.32	0.59	0.19	1.38	09D07C06	0.50	0.17	0.78	4.72	3.01	6.47	
09D04C01	2.02	0.74	3.42	0.96	0.44	2.68	09D08C01	0.72	0.47	1.39	0.41	0.07	0.95	
09D04C02	0.38	0.03	2.10	2.25	0.43	8.11	09D08C02	0.76	0.42	1.27	5.76	2.83	7.05	
09D04C03	0.62	0.12	3.06	0.63	0.03	3.04	09D08C03	0.27	0.11	0.80	0.56	0.22	1.42	
09D04C04	0.51	0.04	2.82	1.12	0.37	6.92	09D08C04	0.49	0.18	0.93	2.95	2.16	5.51	
09D04C05	1.09	0.07	2.48	1.58	0.25	3.70	09D08C05	0.26	0.13	1.41	6.09	4.38	11.93	
09D04C06	1.14	0.50	2.20	0.74	0.17	1.67	09D08C06	0.57	0.03	1.59	0.76	0.13	2.55	
09D05C01	1.13	0.35	1.66	1.29	0.64	2.77	09D09C01	0.27	0.15	1.81	0.26	0.03	1.86	

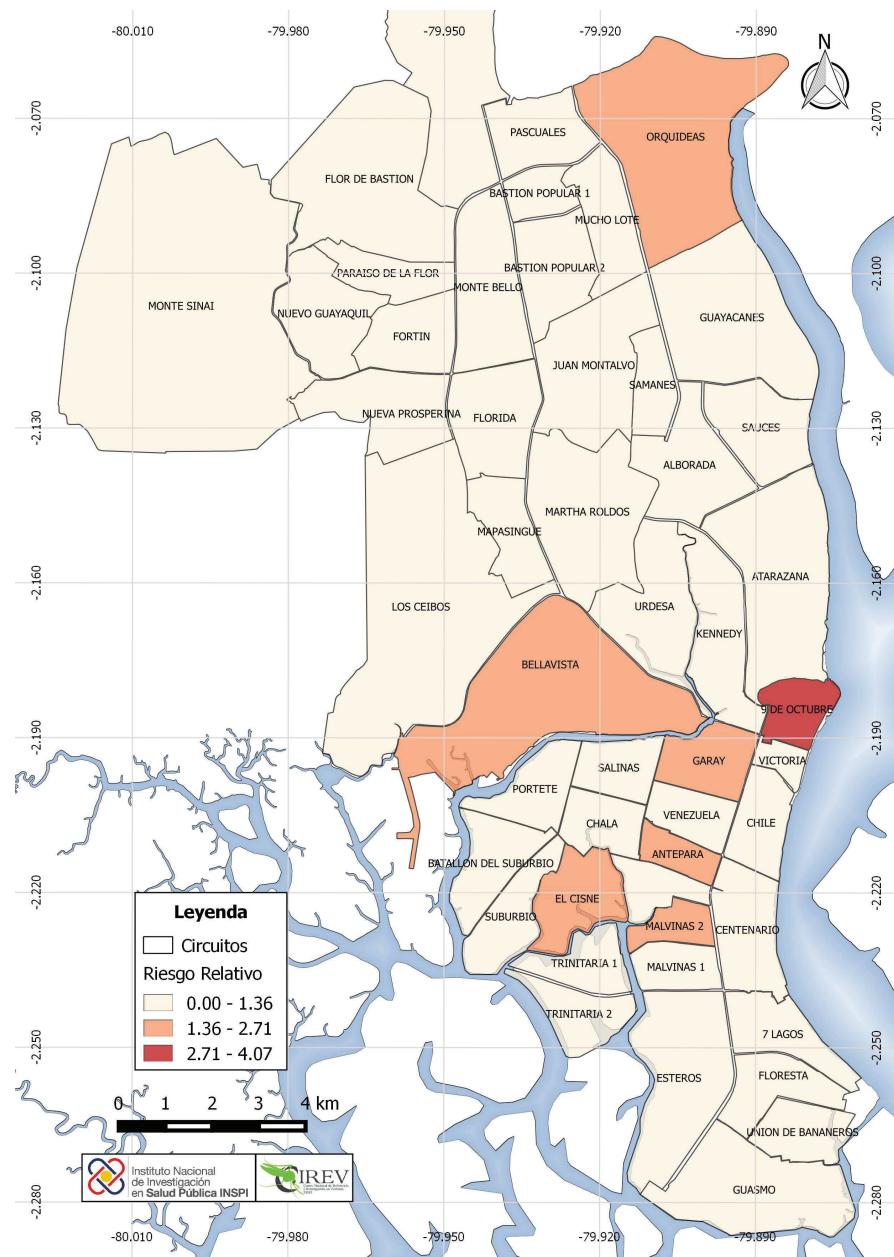
**Cuadro 5.24:** Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2012

Para este año hay un solo circuito de alto riesgo tanto en la época lluviosa como en la seca, que alcanza valores de 4.07 y 9.57 respectivamente. En el riesgo medio existen 6 circuitos en un rango de 1.42 a 2.35 en la estación lluviosa y en la época seca tenemos 9 sectores en un rango de 3.42 a 6.20 y para el riesgo bajo de 0.24 a 1.14 y de 0.26 a 2.95, correspondientemente.

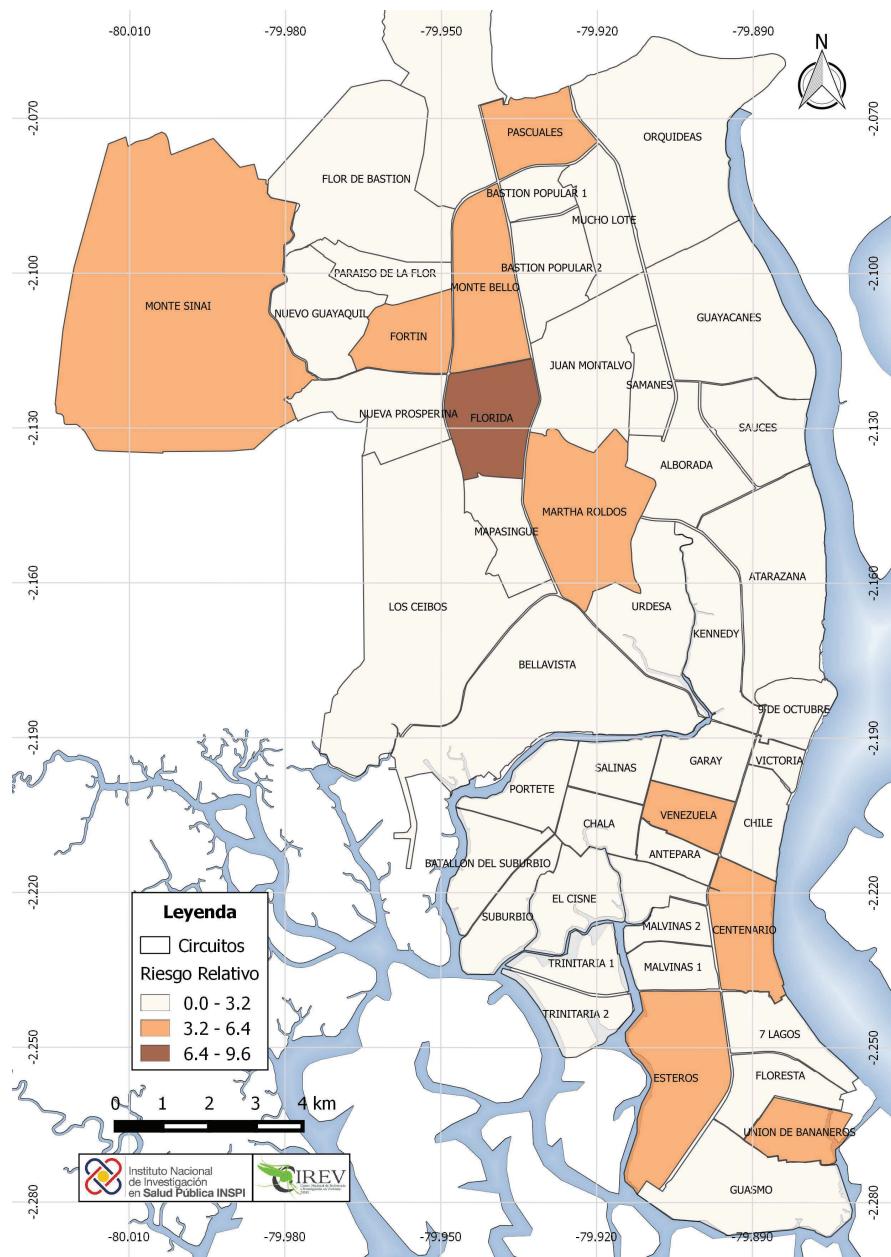
En el riesgo alto el circuito 09 de Octubre tiene casos registrados en enero y abril, que corresponde a los meses donde la temperatura toma valores de 21 y 25 grados

aproximadamente, además debemos tomar en cuenta que este año en este período la precipitación en toda la época se mantiene superior a 120 mm lo que no sucede en los años anteriores. En los meses de mayo a septiembre, la época seca empieza con 57 casos y para septiembre disminuye drásticamente (casi a la nulidad).

Notaremos que la influencia de los circuitos de riesgo medio se hace notoria porque en este grupo los circuitos de riesgo alto comparten límites de frontera. Para el nivel de riesgo bajo en la mayoría de localidades hay más enfermos de dengue en mayo y junio y posterior a ello tenemos reportes nulos de la enfermedad que guardan correspondencia con el inicio de la época seca.



**Figura 5.40:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2012, época lluviosa



**Figura 5.41:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2012, época seca

Como se ha visto con anterioridad, en la época lluviosa en el nivel medio y bajo los casos se concentran más desde febrero a abril, con mayor énfasis en los dos últimos meses dado que en este tiempo tenemos un crecimiento que supera los 23 grados Celsius y en combinación con el intervalo de precipitación que va en promedio desde 159 mm a 85 mm .

Observamos que en los 41 circuitos del riesgo bajo existe un comportamiento similar en el aspecto climático, dado que en el reporte de casos para enero y febrero en un

50 % de los circuitos existe únicamente al menos un caso, lo que nos indica nuevamente que en los primeros meses del año como en marzo y abril hay más eventos de contagio. Además en la época seca del 2012 hay ausencia de casos desde septiembre a noviembre donde las temperaturas desciden en promedio de 27.6 a 26.4 y la precipitación empieza un incremento de en promedio 61.84 mm a 99.89 mm .

#### 5.6.4. Riesgo Relativo para el año 2013

Circuitos	2013						Circuitos	2013						
	Época Lluviosa			Época Seca				Época Lluviosa			Época Seca			
	RR	IC		RR	IC			RR	IC		RR	IC		
		0.25	0.95		0.25	0.95			0.25	0.95		0.25	0.95	
09D01C01	0.48	0.38	0.74	0.40	0.24	0.93	09D05C02	0.15	0.04	0.55	0.56	0.06	1.53	
09D01C02	1.59	0.90	1.91	1.76	1.20	2.21	09D05C03	0.45	0.19	1.44	0.50	0.00	1.80	
09D01C03	1.31	0.09	2.62	0.48	0.02	2.24	09D05C04	0.29	0.05	1.10	4.49	2.27	6.58	
09D01C04	2.01	1.11	3.31	1.41	0.29	2.10	09D05C05	0.66	0.32	0.89	4.50	2.52	5.42	
09D01C05	0.21	0.01	0.79	0.38	0.05	2.27	09D05C06	0.61	0.02	1.43	0.54	0.05	2.81	
09D02C01	0.84	0.13	2.41	0.19	0.02	1.45	09D06C01	0.89	0.12	2.85	0.68	0.03	2.30	
09D02C02	0.59	0.19	1.19	1.11	0.46	1.88	09D06C02	0.34	0.02	1.19	0.63	0.16	3.08	
09D02C03	0.49	0.26	1.00	0.76	0.28	1.33	09D06C03	0.27	0.06	0.84	1.83	0.57	2.45	
09D02C04	0.87	0.17	1.48	0.54	0.15	1.37	09D06C04	0.48	0.24	1.03	4.20	2.81	5.50	
09D02C05	0.36	0.14	2.26	0.81	0.36	4.43	09D06C05	0.30	0.09	1.32	1.22	0.34	2.76	
09D03C01	1.42	0.39	3.60	0.52	0.02	2.04	09D06C06	0.85	0.53	1.58	4.09	1.90	5.01	
09D03C02	1.98	0.12	11.94	1.70	0.05	8.42	09D07C01	0.70	0.08	2.14	1.38	0.30	4.01	
09D03C03	1.21	0.67	4.97	0.50	0.02	2.00	09D07C02	0.75	0.10	2.36	2.25	0.58	6.41	
09D03C04	1.31	0.61	2.59	0.28	0.00	1.59	09D07C03	1.44	0.04	1.95	1.54	0.12	4.23	
09D03C05	0.36	0.02	1.47	1.16	0.06	3.10	09D07C04	0.57	0.44	2.38	1.59	0.66	4.20	
09D03C06	1.48	0.69	3.02	0.36	0.15	1.40	09D07C05	0.36	0.03	3.22	1.14	0.02	4.02	
09D03C07	1.13	0.28	1.41	0.37	0.07	1.28	09D07C06	0.30	0.14	0.69	1.72	0.68	2.50	
09D04C01	1.44	0.99	3.01	0.52	0.09	1.11	09D08C01	0.70	0.37	0.96	2.90	2.12	3.51	
09D04C02	0.76	0.11	3.43	2.87	0.11	4.46	09D08C02	0.28	0.10	0.59	1.31	0.70	2.36	
09D04C03	1.26	0.30	4.45	0.42	0.02	1.91	09D08C03	0.42	0.16	0.72	2.50	1.62	3.20	
09D04C04	0.45	0.02	2.30	0.62	0.02	3.00	09D08C04	0.06	0.01	0.38	1.65	0.83	3.09	
09D04C05	1.53	0.37	3.84	0.66	0.06	3.26	09D08C05	0.27	0.06	1.16	0.61	0.17	1.98	
09D04C06	0.77	0.20	1.47	1.02	0.26	2.01	09D08C06	0.84	0.28	2.12	2.48	1.03	5.09	
09D05C01	1.83	0.84	2.86	0.68	0.31	1.42	09D09C01	0.40	0.23	1.82	2.64	1.71	6.47	

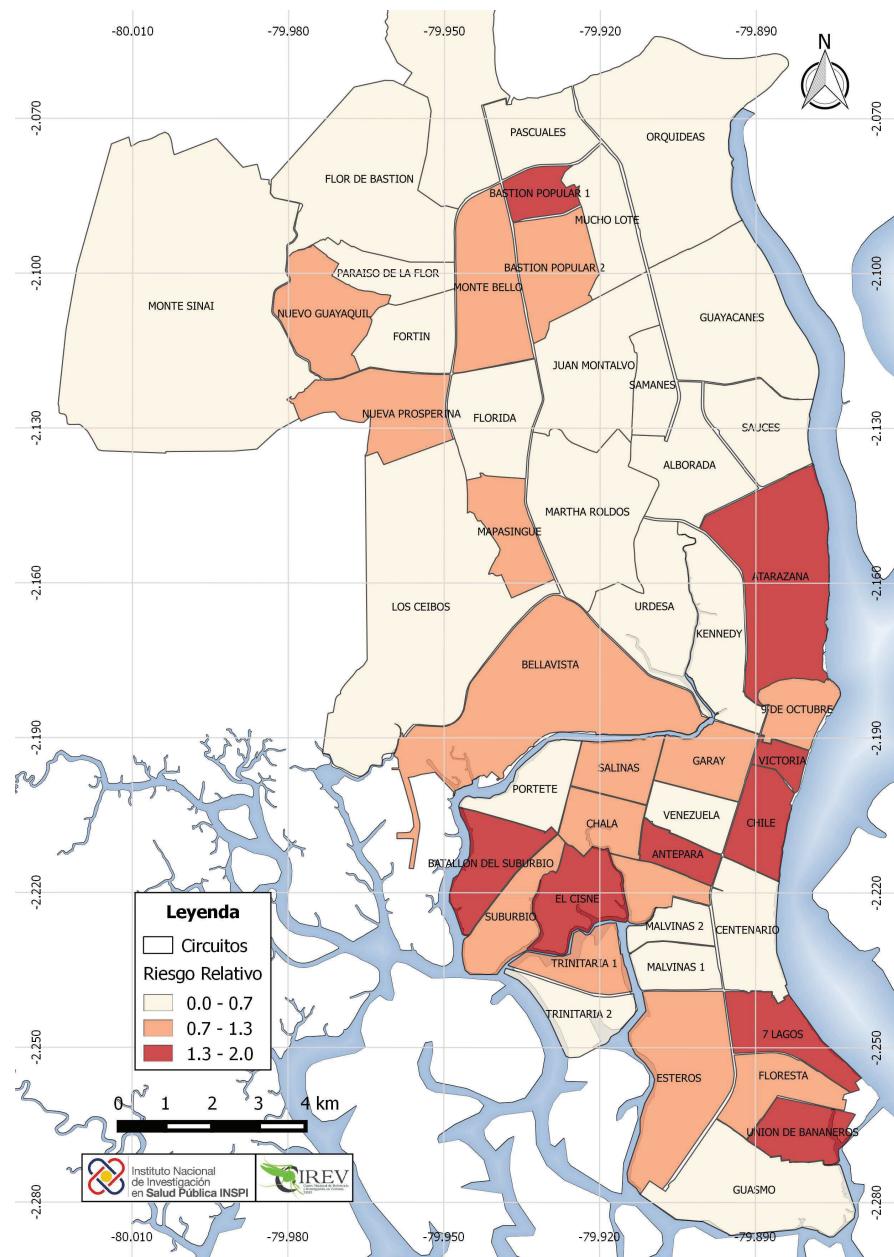
**Cuadro 5.25:** Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2013

Para el año 2013 se puede observar un incremento en el número de localidades que se añaden en el riesgo alto, ocupando un intervalo entre 1.31 y 2.01, en tanto

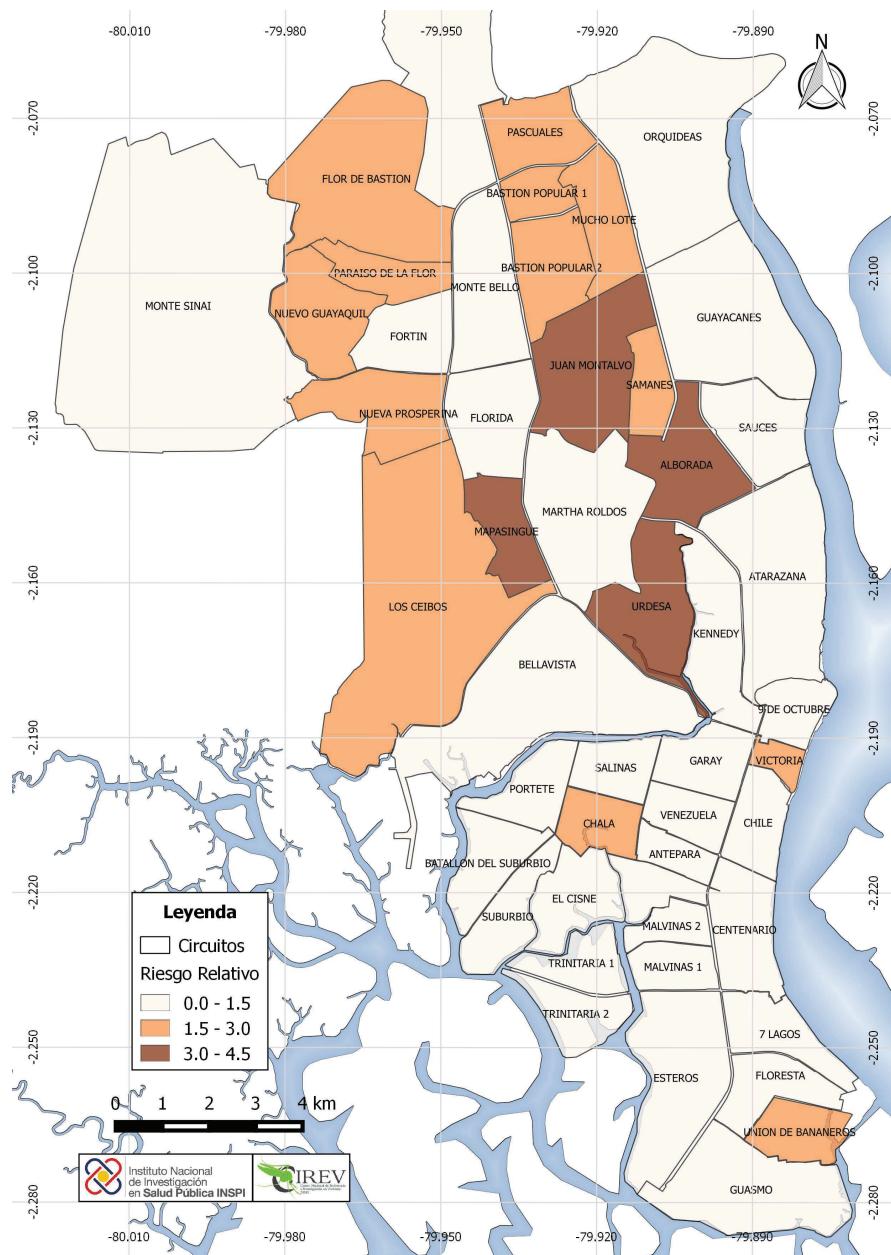
para el riesgo medio se cuenta con 11 localidades ubicadas entre 0.75 y 1.26 y en el nivel bajo entre 0.06 y 0.70.

Para la estación seca sucede algo similar, en el riesgo alto contamos con 4 localidades que toman valores para el riesgo relativo entre 4.20 y 4.50, en el nivel medio el rango es de 1.54 a 2.90, y se agrupan 13 localidades y para el resto de 31 circuitos el riesgo está entre 0.19 y 1.41.

Al igual que lo años anteriores los casos de dengue están dispersos en gran medida hacia los meses primeros del año, al igual que en el riesgo medio. En el nivel más bajo casi no existen reportes para este período, excepto en las localidades: Kennedy, Alborada, Urdesa y Centenario, lo que nos da la idea que en estos lugares las condiciones de salubridad son mejores.



**Figura 5.42:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2013, época lluviosa



**Figura 5.43:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2013, época seca

En el caso del riesgo medio, encontramos gran cantidad de ceros, excepto para los meses de mayo y junio en los que encontramos un número alto que después cambia radicalmente a uno o ceros nuevamente.

En el estrato bajo del riesgo, los reportes de enfermos por dengue han disminuido hacia cero especialmente desde agosto a septiembre, excepto en Chile y Centenario. El descenso empieza en mayo, donde se reportan casos que van reduciéndose

desde 63 a 2

Respecto al ámbito climatológico para este año en enero empieza el año marcando valores de precipitación alrededor de los 70 mm con un incremento en promedio hasta de 141 mm en el mes de marzo. Esta fluctuación llega a estabilizarse hacia el mes de abril (con 65 mm).

Diferente de la estación seca que en el período mayo-junio, en promedio las precipitaciones son de 27 mm hasta 39 mm siguiendo un pequeño incremento en la transición de julio a agosto, y finalmente, para septiembre a noviembre la acumulación de lluvia es de 30 mm.

#### **5.6.5. Riesgo Relativo para el año 2014**

En el año 2014 en la estación lluviosa en el riesgo alto clasifican 4 circuitos con un intervalo de 2.19 a 3.24, 10 localidades en el riesgo medio que están en un rango de 1.10 a 2.01 y en el bajo hay 34 circuitos con una escala de 0.13 a 1.06.

Para la estación seca, tenemos 5 circuitos con un riesgo que oscila entre 5.96 y 8.82 en el nivel alto, para el medio hay 9 circuitos entre 2.98 a 5.62, y en el bajo hay 34 circuitos con valores mucho menores.

Circuitos	2014						Circuitos	2014					
	Epoca Lluviosa			Epoca Seca				Epoca Lluviosa			Epoca Seca		
	RR	IC		RR	IC		RR	IC		RR	IC		RR
		0.25	0.95		0.25	0.95		0.25	0.95		0.25	0.95	
09D01C01	0.51	0.45	0.74	3.39	2.83	3.86	09D05C02	2.01	1.13	2.69	0.88	0.44	2.12
09D01C02	0.56	0.25	1.09	2.15	1.21	4.20	09D05C03	3.24	0.86	4.53	0.94	0.13	2.67
09D01C03	1.48	0.62	4.38	1.11	0.08	4.33	09D05C04	0.18	0.01	0.74	3.20	1.89	7.80
09D01C04	1.01	0.36	2.00	0.73	0.31	1.92	09D05C05	0.29	0.24	0.46	1.93	1.18	2.41
09D01C05	0.17	0.01	0.94	4.56	2.72	8.60	09D05C06	0.45	0.09	4.40	0.32	0.02	5.30
09D02C01	1.28	0.30	2.43	5.11	1.59	8.78	09D06C01	0.50	0.05	2.78	0.92	0.19	3.65
09D02C02	1.53	0.98	2.42	0.92	0.53	2.21	09D06C02	0.59	0.11	1.97	1.28	0.09	3.82
09D02C03	0.33	0.14	0.71	0.97	0.44	1.72	09D06C03	0.55	0.29	1.69	0.47	0.13	1.16
09D02C04	1.10	0.24	2.43	1.03	0.26	3.03	09D06C04	0.46	0.07	0.81	5.62	4.07	8.92
09D02C05	1.11	0.08	3.22	8.82	0.50	6.46	09D06C05	0.25	0.03	1.07	0.28	0.05	1.37
09D03C01	0.80	0.06	3.87	0.52	0.09	5.22	09D06C06	0.36	0.02	0.94	3.54	2.05	9.03
09D03C02	0.85	0.00	7.04	1.32	0.00	13.01	09D07C01	0.39	0.01	1.28	0.99	0.02	2.41
09D03C03	0.21	0.01	1.31	0.47	0.05	3.06	09D07C02	1.16	0.03	2.01	0.62	0.10	3.79
09D03C04	1.30	0.57	2.68	1.59	0.35	3.35	09D07C03	0.13	0.01	1.57	0.20	0.06	3.61
09D03C05	0.43	0.01	1.54	2.17	0.48	7.10	09D07C04	0.43	0.22	1.47	6.90	3.48	10.04
09D03C06	2.19	0.85	2.83	0.70	0.29	2.13	09D07C05	1.28	0.09	5.84	2.26	0.15	6.77
09D03C07	0.38	0.05	0.69	1.92	1.06	3.30	09D07C06	0.58	0.32	1.12	0.33	0.12	0.94
09D04C01	0.60	0.10	1.50	1.73	0.21	3.49	09D08C01	0.24	0.12	0.73	6.24	3.73	8.22
09D04C02	0.70	0.02	2.35	3.29	0.19	4.88	09D08C02	0.31	0.19	0.84	0.49	0.28	1.23
09D04C03	1.06	0.10	2.60	3.43	1.85	7.59	09D08C03	0.28	0.09	0.85	5.98	4.37	9.98
09D04C04	2.81	0.49	5.99	0.84	0.02	4.00	09D08C04	0.29	0.15	0.81	0.44	0.10	1.23
09D04C05	2.41	0.86	4.58	0.81	0.03	2.50	09D08C05	0.28	0.03	1.07	0.57	0.14	1.91
09D04C06	1.46	0.55	2.50	1.17	0.41	2.18	09D08C06	0.62	0.09	1.39	2.98	1.59	8.23
09D05C01	0.43	0.12	0.75	0.51	0.14	1.00	09D09C01	0.19	0.01	0.77	5.96	3.29	10.64

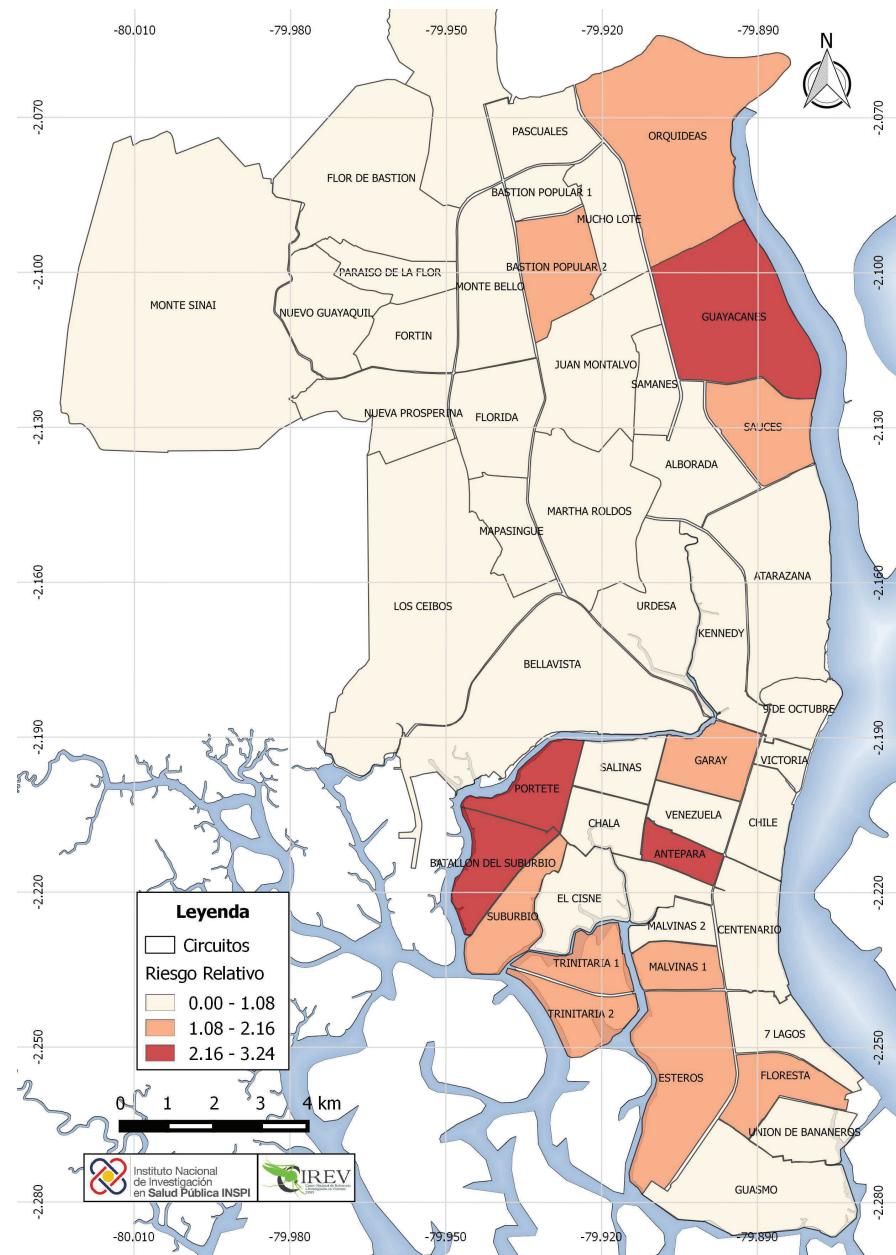
**Cuadro 5.26:** Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2014

En el 2014 la mayor cantidad de reportes es en el mes de marzo y abril , donde en comparación con el número esperado de enfermos hay 16 casos por mes para las localidades del riesgo alto en la estación lluviosa, en concordancia con los meses en los cuales la temperatura promedio es de 26 grados Celsius, pese a que la precipitación va de 52.21 mm a 73 mm. Si ponemos puntual atención en la cantidad de esperados para el nivel medio vemos que se reducen totalmente en comparación con los reales debido a la interacción de las variables climáticas e incluso de los indicadores socio-económicos.

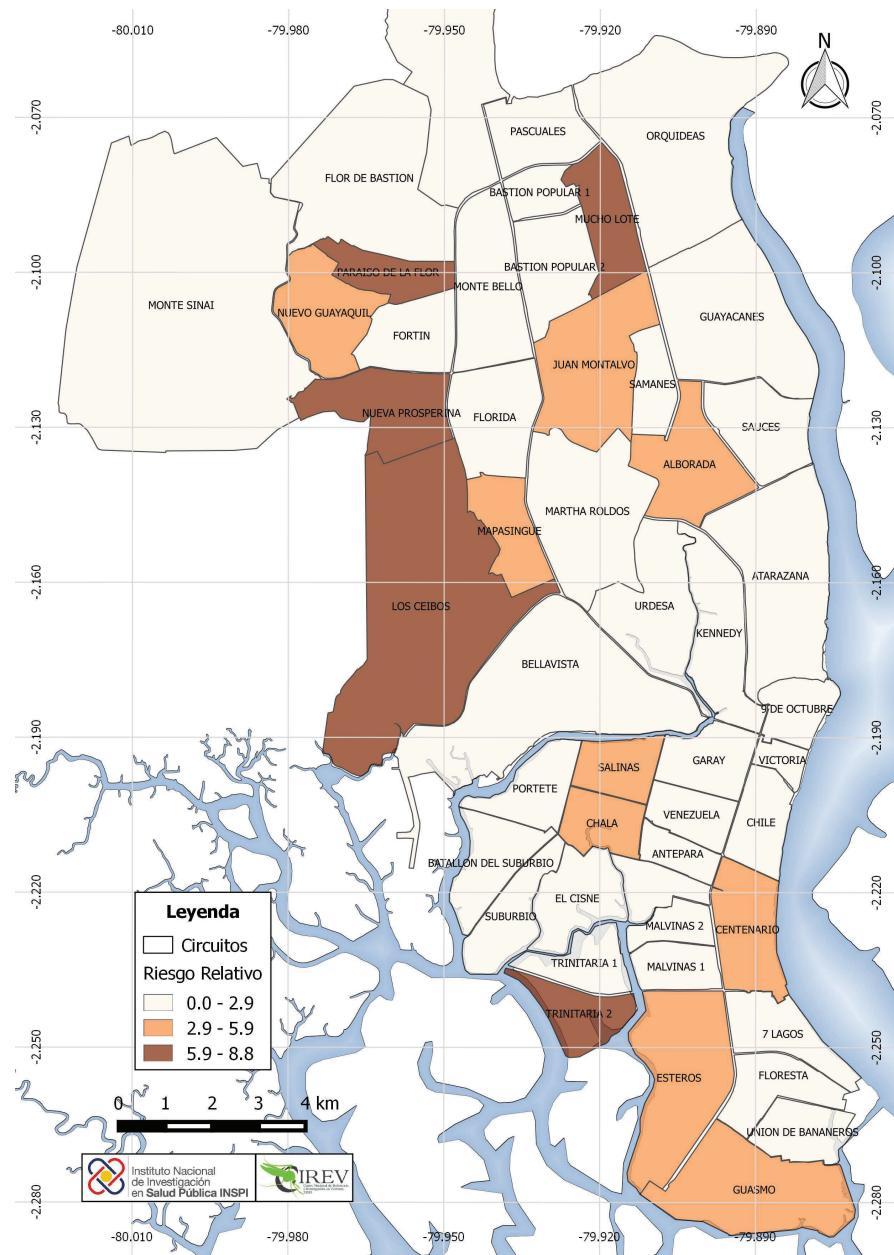
Por último en el riesgo bajo a pesar que inicialmente los datos reales reportan cifras altas en localidades como Urdesa, Alborada y Centenario, podemos ver que la influencia de los indicadores sociales de estos lugares pueden mitigar el aparecimiento

de casos propios de las localidades. También ejerce importancia en el modelo el número de casos esperados que guardan relación con la población que hay en cada circuito. En cambio, los reportes que en realidad fueron casi nulos en el riesgo bajo nos darían una perspectiva de que se esperaba más en comparación de los registrados, lo que quiere decir que la migración por atención de salud podría volverse un factor importante en el reporte de los casos a tiempo.

Para la estación seca, en los meses de mayo a julio los casos esperados influyen en el cálculo del riesgo porque en promedio considerando la población de cada localidad se tiene en mayo en promedio 11, en Junio 13, en Julio 9 y un decrecimiento hacia cero para los meses restantes, justo cuando en promedio la temperatura llega hasta los 23.8 Celsius pero la precipitación disminuye notablemente alrededor de 80 mm. En los estratos de riesgo medio y riesgo bajo tenemos los circuitos que tanto en los datos originales como en los esperados se reducen totalmente, en el primer caso el uno va a cero y el en segundo se esperan reportes de dengue respecto de la población que habita en esas localidades.



**Figura 5.44:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2014, época lluviosa



**Figura 5.45:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2014, época seca

### 5.6.6. Riesgo Relativo para el año 2015

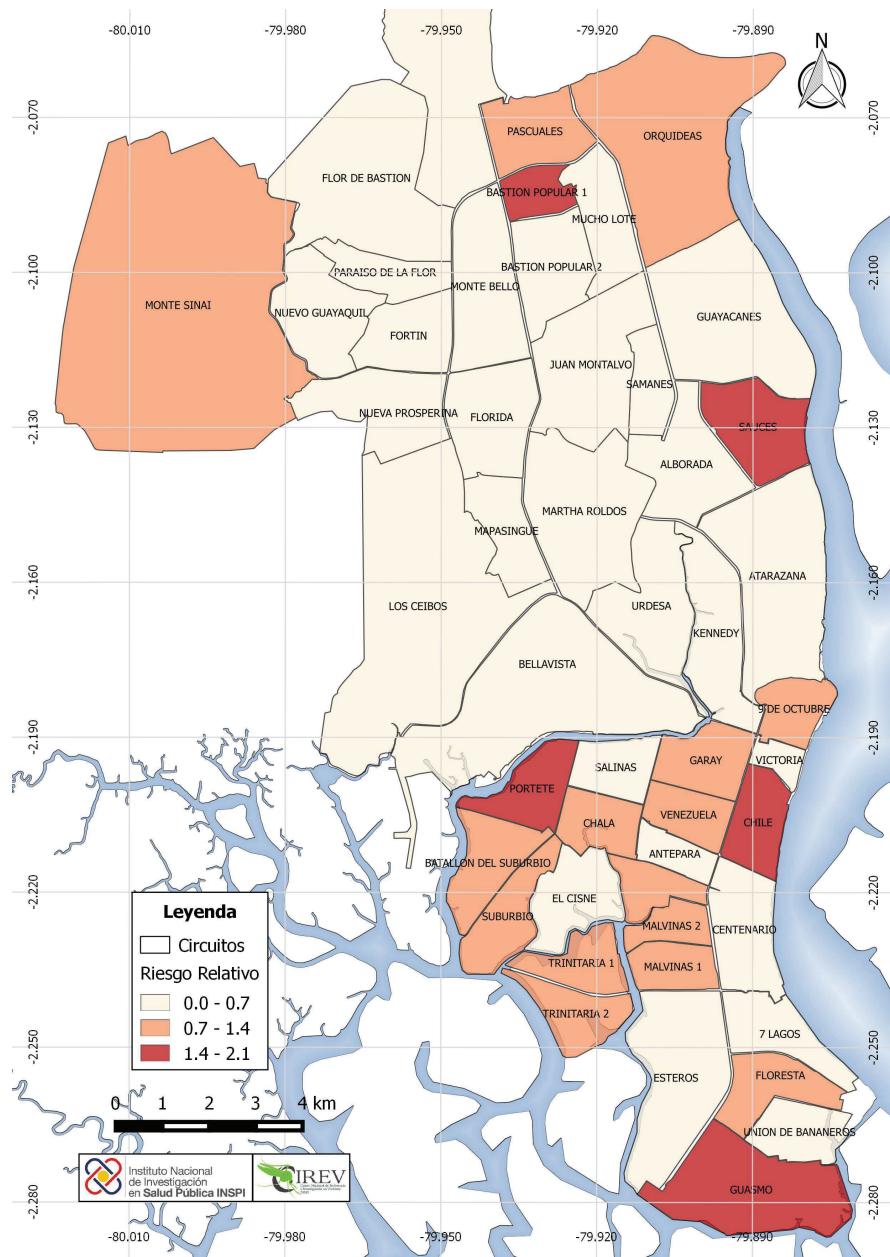
Circuitos	2015						Circuitos	2015						
	Época Lluviosa			Época Seca				Época Lluviosa			Época Seca			
	RR	IC		RR	IC			RR	IC		RR	IC		
		0.25	0.95		0.25	0.95			0.25	0.95		0.25	0.95	
09D01C01	1.85	1.45	2.03	1.44	1.21	1.77	09D05C02	2.13	1.04	2.76	0.67	0.23	1.47	
09D01C02	0.05	0.01	0.34	0.81	0.40	1.69	09D05C03	0.47	0.06	1.49	0.41	0.06	1.53	
09D01C03	1.34	0.83	4.47	0.32	0.02	2.09	09D05C04	0.20	0.01	0.62	2.09	0.70	3.50	
09D01C04	0.34	0.10	0.84	1.01	0.32	1.66	09D05C05	0.05	0.02	0.29	1.55	1.02	1.79	
09D01C05	0.56	0.25	1.85	0.86	0.43	2.71	09D05C06	0.14	0.01	2.17	0.57	0.02	3.31	
09D02C01	0.30	0.01	1.06	0.31	0.04	1.73	09D06C01	0.24	0.01	1.17	1.42	0.12	2.96	
09D02C02	0.72	0.40	1.44	0.63	0.29	1.27	09D06C02	0.36	0.12	2.12	2.98	1.07	5.39	
09D02C03	1.31	0.52	1.80	0.63	0.15	1.54	09D06C03	0.48	0.24	1.45	2.75	1.72	4.33	
09D02C04	0.77	0.29	2.28	1.02	0.33	3.11	09D06C04	0.24	0.03	0.56	1.96	0.73	2.53	
09D02C05	1.25	0.22	3.64	0.52	0.05	1.73	09D06C05	0.23	0.07	0.96	2.48	1.33	4.26	
09D03C01	1.82	0.35	5.91	2.00	0.20	5.90	09D06C06	0.35	0.06	1.16	0.67	0.12	1.90	
09D03C02	0.45	0.00	7.06	0.52	0.03	4.74	09D07C01	0.23	0.02	1.40	1.41	0.76	5.60	
09D03C03	1.21	0.29	3.02	0.86	0.12	3.12	09D07C02	0.53	0.01	1.49	2.19	0.34	4.69	
09D03C04	0.90	0.50	2.66	0.46	0.10	1.22	09D07C03	1.45	0.19	2.42	2.40	0.76	6.50	
09D03C05	1.38	0.08	3.24	1.86	0.25	4.38	09D07C04	0.10	0.01	0.55	1.08	0.20	1.92	
09D03C06	0.48	0.17	1.20	0.25	0.03	1.49	09D07C05	1.06	0.01	2.20	0.31	0.01	3.30	
09D03C07	0.91	0.27	1.25	1.13	0.54	1.88	09D07C06	0.95	0.35	1.27	2.18	1.73	3.25	
09D04C01	0.60	0.06	1.63	0.77	0.10	2.09	09D08C01	0.10	0.03	0.62	1.62	0.86	2.17	
09D04C02	1.23	0.33	4.98	0.25	0.01	2.06	09D08C02	0.43	0.21	0.64	2.59	1.82	3.24	
09D04C03	0.07	0.01	1.14	1.14	0.14	3.34	09D08C03	0.19	0.03	0.60	1.10	0.63	2.42	
09D04C04	1.76	0.38	5.11	0.25	0.01	2.29	09D08C04	0.29	0.10	0.79	3.84	2.33	5.16	
09D04C05	1.06	0.04	1.82	2.05	0.32	3.86	09D08C05	0.77	0.10	1.48	2.87	1.10	3.94	
09D04C06	1.18	0.56	2.60	0.76	0.12	1.79	09D08C06	0.28	0.00	0.74	0.98	0.51	3.47	
09D05C01	0.17	0.03	0.50	0.56	0.29	1.14	09D09C01	0.20	0.01	0.68	1.88	0.41	2.70	

**Cuadro 5.27:** Tabla de Riesgos Relativos, a nivel de Circuitos para la Ciudad de Guayaquil, época lluviosa y seca, año 2015

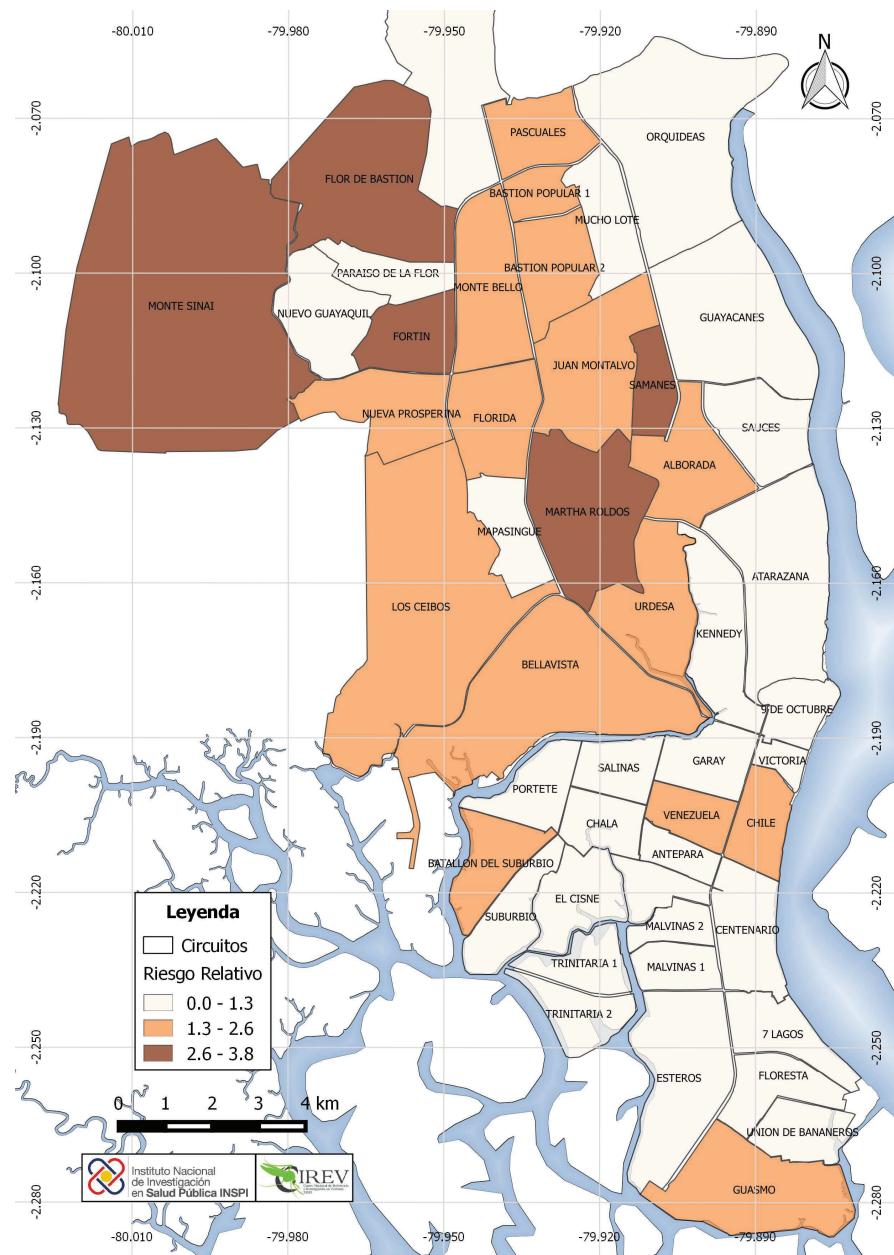
En la época lluviosa tenemos 5 circuitos, entre 1.4 y 2.13, 15 circuitos de riesgo medio entre 0.72 y 1.38, y el resto de localidades están en un rango de 0.05 a 0.60, mientras que en la época seca en el nivel alto tenemos 4 circuitos con un rango de 2.75 a 3.84, 16 circuitos en nivel medio en un intervalo de 1.41 a 2.59, y finalmente el nivel bajo entre los valores 0.25 a 1.14

En este año se tiene un mayor número de casos de dengue en los meses marzo y abril que coindicen con el comportamiento climatológico y que concuerda con el

número de casos que se espera, especialmente en este período donde la temperatura varía entre los 24 °C, y que también son los meses en los que la precipitación está entre 75 mm y 92 mm en promedio.



**Figura 5.46:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2015, época lluviosa



**Figura 5.47:** Mapa de Riesgos Relativos en niveles: Alto, Medio, Bajo, para el año 2015, época seca

De mayo a agosto tenemos la mayor acumulación de casos de dengue, pero en el riesgo alto influye nuevamente el número de casos esperados en los circuitos dado que en mayo se espera en promedio 102 enfermos, en junio alrededor de 89 casos donde los eventos citados guardan conexión con temperaturas de 26 °C y precipitación de 97 y 177 mm en promedio.

## 5.7. Análisis de Residuos

Para el comportamiento de los resultados se procedió a calcular cada uno de ellos para el período 2010-2015 en el ambiente estadístico R, ocupando la librería *ape* y usando como matriz a W que esta descrita en la parte de los *Anexos*.

Con la verificación de la convergencia si el modelo se ajusta al riesgo relativo los residuos deben presentar independencia, la cual se corrobora con el estadístico I de Moran [8]. Para el propósito se calcularon los valores de este estadístico para cada mes de cada año y dado que en cada uno el p-valor obtenido fue menor al 5 % se puede rechazar la hipótesis nula que es la presencia de autocorrelación espacial, como se verifica en las tablas presentes en esta sección.

Mes	Valor observado	Valor esperado	Desviación Estándar	p-valor
Enero	0.0992	-0.0213	0.0694	8.24E-03
Febrero	0.2245	-0.0213	0.0815	2.56E-03
Marzo	0.2410	-0.0213	0.0818	1.36E-03
Abril	0.2692	-0.0213	0.0853	6.59E-04
Mayo	0.2976	-0.0213	0.0791	5.58E-05
Junio	0.3400	-0.0213	0.0730	7.53E-07
Julio	0.1196	-0.0213	0.0697	4.31E-02
Agosto	0.1957	-0.0213	0.0721	2.63E-03
Septiembre	0.3208	-0.0213	0.0831	3.87E-05
Octubre	0.0571	-0.0213	0.0866	3.66E-03
Noviembre	0.2285	-0.0213	0.0520	1.55E-06
Diciembre	0.1128	-0.0213	0.0839	1.10E-03

Cuadro 5.28: Residuos del modelo elegido, año 2010

Mes	Valor observado	Valor esperado	Desviación Estándar	p-valor
Enero	0.3884	-0.0213	0.0765	8.52E-08
Febrero	0.2558	-0.0213	0.0844	1.03E-03
Marzo	0.1821	-0.0213	0.0825	1.37E-02
Abril	0.1358	-0.0213	0.0857	6.67E-03
Mayo	0.2380	-0.0213	0.0499	2.02E-07
Junio	0.1226	-0.0213	0.0787	6.74E-03
Julio	0.2097	-0.0213	0.0766	2.58E-03
Agosto	0.1628	-0.0213	0.0857	3.17E-02
Septiembre	0.0549	-0.0213	0.0873	3.83E-02
Octubre	0.2152	-0.0213	0.0380	5.07E-10
Noviembre	0.2095	-0.0213	0.0769	2.69E-03
Diciembre	0.1602	-0.0213	0.0720	1.18E-02

Cuadro 5.29: Residuos del modelo elegido, año 2011

Mes	Valor observado	Valor esperado	Desviación Estándar	p-valor
Enero	0.2328	-0.0213	0.0818	1.91E-03
Febrero	0.3288	-0.0213	0.0735	1.91E-06
Marzo	0.2549	-0.0213	0.0871	1.51E-03
Abril	0.1035	-0.0213	0.0839	1.37E-02
Mayo	0.2339	-0.0213	0.0815	1.73E-03
Junio	0.2184	-0.0213	0.0750	1.40E-03
Julio	0.3071	-0.0213	0.0841	9.53E-05
Agosto	0.3655	-0.0213	0.0847	4.95E-06
Septiembre	0.1629	-0.0213	0.0847	2.97E-02
Octubre	0.1917	-0.0213	0.0838	1.10E-02
Noviembre	0.2167	-0.0213	0.0860	5.66E-03
Diciembre	0.4648	-0.0213	0.0834	5.49E-09

Cuadro 5.30: Residuos del modelo elegido, año 2012

Mes	Valor observado	Valor esperado	Desviación Estándar	p-valor
Enero	0.2084	-0.0213	0.0824	5.31E-03
Febrero	0.1528	-0.0213	0.0844	3.92E-02
Marzo	0.2094	-0.0213	0.0828	5.37E-03
Abril	0.3113	-0.0213	0.0854	9.89E-05
Mayo	0.2035	-0.0213	0.0738	2.33E-03
Junio	0.3659	-0.0213	0.0826	2.79E-06
Julio	0.1210	-0.0213	0.0824	8.43E-03
Agosto	0.3082	-0.0213	0.0807	4.48E-05
Septiembre	0.2210	-0.0213	0.0850	4.37E-03
Octubre	0.2321	-0.0213	0.0656	1.12E-04
Noviembre	0.3236	-0.0213	0.0757	5.20E-06
Diciembre	0.2912	-0.0213	0.0816	1.29E-04

Cuadro 5.31: Residuos del modelo elegido, año 2013

Mes	Valor observado	Valor esperado	Desviación Estándar	p-valor
Enero	0.2823	-0.0213	0.0786	1.12E-04
Febrero	0.3128	-0.0213	0.0750	8.42E-06
Marzo	0.1307	-0.0213	0.0815	6.22E-03
Abril	0.1231	-0.0213	0.0849	8.90E-03
Mayo	0.1317	-0.0213	0.0764	4.53E-02
Junio	0.2964	-0.0213	0.0821	1.08E-04
Julio	0.2133	-0.0213	0.0773	2.42E-03
Agosto	0.0776	-0.0213	0.0832	2.35E-02
Septiembre	0.1284	-0.0213	0.0867	8.41E-03
Octubre	0.2675	-0.0213	0.0828	4.86E-04
Noviembre	0.1507	-0.0213	0.0824	3.69E-02
Diciembre	0.2348	-0.0213	0.0514	6.34E-07

Cuadro 5.32: Residuos del modelo elegido, año 2014

Mes	Valor observado	Valor esperado	Desviación Estándar	p-valor
Enero	0.2510	-0.0213	0.0343	2.00E-15
Febrero	0.2198	-0.0213	0.0744	1.20E-03
Marzo	0.1941	-0.0213	0.0844	1.07E-02
Abril	0.2093	-0.0213	0.0851	6.75E-03
Mayo	0.1731	-0.0213	0.0728	7.56E-03
Junio	0.2390	-0.0213	0.0394	4.01E-11
Julio	0.1608	-0.0213	0.0825	2.74E-02
Agosto	0.2789	-0.0213	0.0854	4.39E-04
Septiembre	0.3423	-0.0213	0.0839	1.48E-05
Octubre	0.2334	-0.0213	0.0783	1.14E-03
Noviembre	0.1989	-0.0213	0.0791	5.38E-03
Diciembre	0.2682	-0.0213	0.0804	3.20E-04

**Cuadro 5.33:** Residuos del modelo elegido, año 2015

A lo largo de este proceso uno de los objetivos en mente fue la identificación de los factores incidentes en la propagación del dengue y las zonas de mayor atención. Hemos evaluado los casos en el período 2010 al 2015 a nivel de circuitos a través de una estructura con variables climáticas y socioeconómicas para llegar a una posible estimación para el año 2016, sin embargo, la distribución de los datos tan dispersa, la técnica para el relleno de faltantes para la información epidemiológica, la poca variabilidad en la serie de temperatura y los indicadores del INEC que son fijos desde el año 2010, que fue el último censo, impidieron la estimación de una estructura a través de los métodos ARIMA para las series de datos que se encuentran inmersas en el modelo estudiado. Otro factor involucrado es la forma de almacenamiento de los datos climatológicos que en cada pixel de la imagen ráster registra información casi considerada como un vector y al refinar la extracción por circuitos se tiende a reducir y perder datos, debido a que se toma la media de cada pixel que se encuentra en la unidad de estudio.

Estos motivos incentivan a la fusión de la modelización bayesiana para detectar una estructura para simular el riesgo relativo en cada área pero la implementación se sugiere ejecutarla en la plataforma TerrSET que se considera un sistema para el análisis y seguimiento de los datos geoespaciales, ya que se puede añadir las funcionalidades de variables temporales sin tener que usar solo la estimación promedio.

# **Capítulo 6**

## **Limitaciones, Conclusiones, Recomendaciones, y Trabajo Futuro**

### **6.1. Limitaciones**

- Para las estimaciones de los datos faltantes en las series climáticas relacionadas con el Índice de Vegetación tuvimos como herramienta fundamental la información almacenada en los píxeles que proporcionan los satélites MODIS cada 8 días, sin embargo, en el relleno de los vacíos se utilizaron valores promedio ya que no se podían comparar con estaciones metereológicas para completar incluir otros métodos de estimación.
- No todos los casos reales constan para el uso de los modelos, puesto que en la base de datos únicamente se registran aquellos que pasaron previamente por la detección clínica a través de pruebas de laboratorio.
- Las semanas epidemiológicas no coinciden con las semanas normales del calendario y el comportamiento de la aparición de reportes de dengue no concuerda con las estaciones definidas, dado a la irregularidad de los registros de precipitaciones en todo el comportamiento anual por el cambio de un mes a otro.

### **6.2. Conclusiones**

- La ciudad de Guayaquil fue elegida por ser una zona ubicada bajo el rango de altitud (1680 metros) donde encontramos presencia del vector *Aedes aegypti* y de casos de contagio de dengue.

- La implementación bayesiana fue realizada para cada uno de los 48 circuitos tomando en cuenta la temporalidad mensual de las variables temperatura, precipitación e índice de vegetación, retirando la estacionalidad de la segunda variable por la presencia de ciclos desde enero a abril, de mayo a julio, y de agosto a diciembre.
- En el monitoreo de los modelos propuestos se tomaron como distribuciones a priori aquellas sugeridas en la bibliografía, las mismas que son el resultado de investigaciones anteriores, al igual que la sugerencia de las variables climáticas que influyen el desarrollo del vector transmisor. [41]
- Las condiciones geográficas de las diferentes localidades altamente urbanas y los lugares donde aún se perciben dificultades en las condiciones socio económicas y de vivienda son un factor clave para encaminar las actividades de control y disminuir los índices de infestación para el mosquito transmisor del dengue.
- El hallazgo de los casos de dengue guarda relación con las limitaciones en el acceso a agua corriente y la necesidad de las personas para almacenarla, puesto que en el momento en que ocurre este hecho se crean grandes ventanas de probabilidad para el aparecimiento de criaderos del mosquito.
- El factor vinculado a la recolección de desechos sólidos también aporta relevancia, especialmente si no se tienen los cuidados apropiados con aquellos de mayor dimensión como: llantas, recipientes plásticos, tarros, floreros, otros, que acumulan agua y son hospederos del *Aedes aegypti*.
- La ausencia del alcantarillado se vuelve otro indicador prioritario puesto que con un incorrecto almacenamiento de agua para el consumo y limpieza se recae en crear el ambiente idóneo para la reproducción del vector.
- El incremento de los casos de dengue se combina con la temporalidad de los picos de precipitación que ocurren en abril, en el cambio de junio a julio y en la fase de diciembre a febrero.
- La única forma de prevenir es la educación para difundir entre las personas que el almacenamiento de agua debe ser en recipientes herméticos sin oportunidad de que el mosquito pueda poner huevos dentro de ella, debido a que las fumigaciones se conocen como medida de control más no de erradicación.
- Las simulaciones realizadas en WinBUGS14 para determinar la convergencia tomaron como primera instancia la generación de 5000 iteraciones para conseguir cadenas estables y proceder al monitoreo de los parámetros en estudio.

- Los criterios gráficos basados en el *history plot*, las figuras relativas a la autocorrelación, la densidad relativa, y el estadístico de Gelman y Rubin con un intervalo alrededor de [0.95,1.05] fueron la base fundamental para probar la convergencia de las cadena de Markov. En especial cuando los parámetros a posteriori presentados reportaron que el error MC es menor a la desviación a posteriori, lo que verifica el ajuste adecuado del modelo seleccionado.
- En el análisis de los residuos se introdujo el coeficiente I de Morán para contrastar la hipótesis de la autocorrelación espacial y detectar la ausencia de la misma.
- Los mapas de riesgo presentados de forma anual y por estación lluviosa y seca según [47] fueron elaborados en el software QGis, de forma que internamente con un número fijo de clases de riesgo se definieron los umbrales de separación para todas las localidades.
- El concepto de la búsqueda del umbral sigue siendo subjetivo y relacionado con los intereses del investigador, en nuestro caso definimos tres clases para clasificar el riesgo en QGis y pintar las áreas de acuerdo al nivel.
- Este estudio nos permitió encontrar los factores que contribuyen en el contagio y la dispersión de la enfermedad del dengue adaptando el efecto de heterogeneidad entre las localidades y la influencia de la vecindad.

### 6.3. Recomendaciones

- Para mejora el modelo se debería tomar en cuenta en la dinámica de la enfermedad el serotipo circulante, dado que el tipo uno y tres son los menos agresivos. Por otro lado el serotipo 2 es el que menor respuesta inmune produce y también depende de la exposición de la gente, es decir, si la población susceptible ya presentaba anticuerpos para algún serotipo en particular hace que el hallazgo de casos disminuya.
- Para depurar y acercar más el modelo a la realidad podemos realizar primero un análisis para evaluar la selección de prioris, pues para este trabajo partimos de la bibliografía seleccionada [34], [36], ya que no fue objetivo central el enfoque en el estudio de la elección de estas distribuciones.
- Gracias a la determinación de los llamados lugares focales vinculados a los factores del riesgo relativo modelado (desempleo, analfabetismo, condiciones de salubridad, precipitación temperatura) se puede implementar simulaciones en

esas localidades con un modelo SIR (Susceptibles, Infectados, y Recuperados), de modo que esta técnica sustituya a la fase de completación de datos faltantes y la parte de predicciones.

- Con la técnica mencionada en el apartado anterior se pueden realizar predicciones y validar el modelo con la información de la base de datos, puesto que en este trabajo solo se pudo explorar el cambio del riesgo relativo del período 2010 al 2015 identificando los factores de esencial atención.
- Es de importante conocimiento que se puede fusionar la dinámica del vector y la propagación del dengue tomando en cuenta la influencia del transporte expresada a través de la movilidad de una localidad a otra, sustentada especialmente cuando existe alto movimiento poblacional como en días feriados o vacaciones.

## 6.4. Trabajo Futuro

- Se desea ampliar el uso de modelos bayesianos para aplicar el modelamiento del riesgo relativo en diferentes zonas del país y extender la generación de simulaciones a enfermedades como Chikungunya y Zika para las cuales el mosquito *Aedes aegypti* es el mismo transmisor.
- Una buena opción puede ser considerada como la búsqueda no subjetiva para establecer los umbrales de riesgo, dado que los presentados en este proyecto fueron determinados automáticamente por el programa QGIs definiendo un número fijo de clases.
- Se puede aprovechar la clasificación de los circuitos por riesgos para aplicar específicamente un modelo que considere la interacción vector-humano como el que se puede aplicar teniendo en cuenta la población humana el modelo epidemiológico SIR con los compartimentos susceptibles, infectados y recuperados. En tanto, la población del vector se modela con una estructura básica SI (susceptibles, infectados), tratando de incluir la capacidad vectorial, la carga viral, la tasa de infección humano-mosquito y mosquito-humano junto con los factores que hemos indagado en el modelo elegido en este proyecto.

# Capítulo 7

## Anexos

### 7.1. Definición de las Variables

Cada una de las variables descritas en este apartado fueron tomadas de las Fichas Metodológicas del Instituto Nacional de Estadística y Censos (INEC).

- **Hacinamiento:** Se considera que un hogar está hacinado si cada uno de los dormitorios con los que cuenta sirve, en promedio, a un número de miembro mayor a tres. Se define como dormitorio a los cuartos o espacios dedicados sólo para dormir; no se incluyen otros espacios disponibles para habitar (como salones, comedor, cuartos de uso múltiple, etc.) que pueden dedicarse ocasional o parcialmente para dormir, más como las cocinas, baños, pasillos, garajes y espacios destinados a fines profesionales. (La vivienda inadecuada o de mala calidad y el hacinamiento figuran entre las manifestaciones más visibles de pobreza. De hecho, la superpoblación es un reflejo indirecto de las condiciones sociales, económicas y sanitarias del país.)
- **Desempleo:** Este grupo está compuesto por personas de 15 años y más, que en el período de referencia, no tuvieron empleo y presentan simultáneamente ciertas características i) No tuvieron empleo la semana pasada y están disponibles para trabajar y ii) buscaron trabajo (o no) o realizaron gestiones concretas para conseguir empleo o para establecer algún negocio en las cuatro semanas anteriores.
- **Tasa de analfabetismo:** Comprende el número de personas que no saben leer y/o escribir de 15 años o más, expresado como porcentaje de la población total de la edad de referencia.  
Variable de agrupación: P19, opción 1.

- **Porcentaje de hogares hacinados:** Número de personas que viven en condiciones de hacinamiento, expresado como porcentaje del total de personas.  
*Variable de agrupación: V14.*
- **Porcentaje de viviendas en mal estado:** Número de personas que habitan en viviendas cuyas condiciones habitacionales se consideran irrecuperables a partir de la combinación materiales predominantes y el estado de aquellos materiales, expresado como porcentaje del total de viviendas.  
*Variable de agrupación: V02, V04, V06, se escogen los campos que no tienen el código 1.*
- **Porcentaje de viviendas con recolección de basura:** Número de viviendas que utilizan el servicio de carro recolector para eliminar la basura, expresado como porcentaje del total de viviendas.  
*Variable de agrupación: V13, se escoge los campos que no tienen el código 1.*
- **Porcentaje de viviendas con alcantarillado:** Número de viviendas que están conectadas a una red pública de alcantarillado o pozo séptico, expresado como porcentaje del total de viviendas.  
*Variable de agrupación: V09, se escogen los campos que no tienen el código 1.*
- **Porcentaje de viviendas con agua corriente:** Número de viviendas con servicio de agua por red pública en su interior, expresado como porcentaje del total de viviendas.  
*Variable de agrupación: V07, V08, se escogen los campos que no tienen el código 1.*
- **Tasa de desempleo:** Número de personas desempleadas, expresada como porcentaje del total de la población.  
*Variable de agrupación: P27.*

## Variables del Censo 2010 del INEC

Elaboración . E. Quentin  
Apoyo técnico . grupo EpiSIG

octubre de 2015

### Contenido

1. Introducción .....	2
2. Conformación de los datos disponibles .....	2
2.1. Censo Población.....	2
2.2. Censo Hogar .....	2
2.3. Censo Vivienda.....	3

## 1. Introducción

### 2. Conformación de los datos disponibles

#### 2.1. Censo Población

14483499 registros

101 - 105 : ubicación (prv, cnt, prq, zna, sct)

P01 (sexo 1=M, 2=F), P03 (P03M mes de nacimiento, P03A año de nacimiento)

P03, P04M, P04A Años cumplidos Mes que nació Año que nació	P19 Sabe leer y escribir	P27 ¿Qué hizo la semana pasada?	P28 Si no ha trabajado
Número	1-Si 2-No	1-Trabajo al menos 1 hora 2-No trabajó pero si tiene trabajo 3-Al menos 1 hora fabricó algún producto 4-Al menos 1 hora ayudó en algún negocio o trabajo de un familiar 5-Al menos 1 hora realizó labores agrícolas o cuidó animales 6-Es cesante, Buscó trabajo habiendo trabajado antes y está disponible para trabajar 7-No trabajó	1-Busca trabajo por primera vez y está disponible para trabajar 2-Es rentista 3-Es jubilado o pensionista 4-Es estudiante 5-Quehaceres domésticos 6-Discapacitado 7-Otro

#### 2.2. Censo Hogar

3815527 registros

Total de personas : 14468401 (faltan 15098; 7165216 hombres, 7302608 mujeres)

4979 hogares sin datos (0.1%)

Una vivienda puede albergar varios hogares (curiosamente el censo reporta menos hogares que vivienda)

H03, H04 Servicio higiénico Lugar para bañarse	H06 Proceso aplicado al agua para beberla	H05 Combustible o energía para cocinar
1-Uso exclusivo 2-Compartido 3-No hay	1-Tal cual 2-Hervir 3-Cloro 4-Filtro 5-Comprar agua	1-Gas (tanque, cilindro) 2-Gas centralizado 3-Electricidad 4-Leña, carbón 5-Residuos

		6-Otro (gasolina...)
		7-No cocina

### 2.3. Censo Vivienda

4654309 registros  
 Nota : 19.5% de las viviendas (905390) no tienen esta información (vacío). 19.3% de las viviendas (900411) tienen 0 personas (no hay vacío).

V01 a V06 material

V01 Techo o cubierta	V03 Paredes exteriores	V05 Piso	V02,V04,V06 Estado
1-Hormigón (losa, cemento)	1-Hormigón	1-Duela, parquet, tablón o piso flotante	1-Bueno
2-Asbesto (Eternit, Eurolit)	2-Ladrillo o bloque	2-Tabla sin tratar	2-Regular
3-Zinc	3-Adobe o tapia	3-Cerámica, baldosa, vinil o mármol	3-Malo
4-Teja	4-Madera	4-Ladrillo o cemento	
5-Palma, paja, hoja	5-Caña revestida o bahareque	5-Caña	
6-Otro	6-Caña no revestida	6-Tierra	
	7-Otro	7-Otro	

V07 a V09, V13 agua basura

V07 Proveniencia del agua	V08 Acceso agua	V09 Evacuación	V13 Basura
1-Red pública	1-Tubería dentro	1-Conexión alcantarillado	1-Carro recolector
2-Pozo	2-Tubería cerca	2-Conexión séptico	2-Arrojada terreno
3-Río	3-Tubería fuera	3-Conexión pozo ciego	3-Quemada
4-Carro repartidor	4-No por tubería	4-Descarga al río	4-Enterrada
5-Otro (lluvia)		5-Letrina	5-Arrojada río
		6-No tiene	6-Otro

V10 a V12B energía

V10 Servicio de luz	V11 Medidor de energía	V12A, V12B Focos ahorradores Focos convencionales
1-Red pública	1-Uso exclusivo	Número
2-Panel solar	2-Compartido	
3-Generador (planta)	3-No tiene medidor	
4-Otro		
5-No tiene		

V14 Cuartos	TOTDOR	PERCUA Personas por cuarto	PERDOR Personas por cuarto
Cuantos cuartos tiene la vivienda. Sin contar la cocina, el baño y cuartos de negocio.	Total de dormitorios de la vivienda	1- Menos de 2 " 2- De 2 a 3 3 De 4 a 5 4 Más de 5	1-Viviendas sin dormitorios 2-Menos de 2 " 3-De 2 a 3 4-De 4 a 5 5-Más de 5



```

u10[ i ]~dnorm(0 ,tau .u[10])
u11[ i ]~dnorm(0 ,tau .u[11])
u12[ i ]~dnorm(0 ,tau .u[12])
u13[ i ]~dnorm(0 ,tau .u[13])
u14[ i ]~dnorm(0 ,tau .u[14])
u15[ i ]~dnorm(0 ,tau .u[15])
u16[ i ]~dnorm(0 ,tau .u[16])
u17[ i ]~dnorm(0 ,tau .u[17])
u18[ i ]~dnorm(0 ,tau .u[18])
u19[ i ]~dnorm(0 ,tau .u[19])
u20[ i ]~dnorm(0 ,tau .u[20])
u21[ i ]~dnorm(0 ,tau .u[21])
u22[ i ]~dnorm(0 ,tau .u[22])
u23[ i ]~dnorm(0 ,tau .u[23])
u24[ i ]~dnorm(0 ,tau .u[24])
u25[ i ]~dnorm(0 ,tau .u[25])
u26[ i ]~dnorm(0 ,tau .u[26])
u27[ i ]~dnorm(0 ,tau .u[27])
u28[ i ]~dnorm(0 ,tau .u[28])
u29[ i ]~dnorm(0 ,tau .u[29])
u30[ i ]~dnorm(0 ,tau .u[30])
u31[ i ]~dnorm(0 ,tau .u[31])
u32[ i ]~dnorm(0 ,tau .u[32])
u33[ i ]~dnorm(0 ,tau .u[33])
u34[ i ]~dnorm(0 ,tau .u[34])
u35[ i ]~dnorm(0 ,tau .u[35])
u36[ i ]~dnorm(0 ,tau .u[36])
u37[ i ]~dnorm(0 ,tau .u[37])
u38[ i ]~dnorm(0 ,tau .u[38])
u39[ i ]~dnorm(0 ,tau .u[39])
u40[ i ]~dnorm(0 ,tau .u[40])
u41[ i ]~dnorm(0 ,tau .u[41])
u42[ i ]~dnorm(0 ,tau .u[42])
u43[ i ]~dnorm(0 ,tau .u[43])
u44[ i ]~dnorm(0 ,tau .u[44])
u45[ i ]~dnorm(0 ,tau .u[45])
u46[ i ]~dnorm(0 ,tau .u[46])
u47[ i ]~dnorm(0 ,tau .u[47])
u48[ i ]~dnorm(0 ,tau .u[48])
u49[ i ]~dnorm(0 ,tau .u[49])
u50[ i ]~dnorm(0 ,tau .u[50])
u51[ i ]~dnorm(0 ,tau .u[51])
u52[ i ]~dnorm(0 ,tau .u[52])
u53[ i ]~dnorm(0 ,tau .u[53])
u54[ i ]~dnorm(0 ,tau .u[54])
u55[ i ]~dnorm(0 ,tau .u[55])
u56[ i ]~dnorm(0 ,tau .u[56])
u57[ i ]~dnorm(0 ,tau .u[57])
u58[ i ]~dnorm(0 ,tau .u[58])
u59[ i ]~dnorm(0 ,tau .u[59])
u60[ i ]~dnorm(0 ,tau .u[60])
u61[ i ]~dnorm(0 ,tau .u[61])
u62[ i ]~dnorm(0 ,tau .u[62])
u63[ i ]~dnorm(0 ,tau .u[63])
u64[ i ]~dnorm(0 ,tau .u[64])
u65[ i ]~dnorm(0 ,tau .u[65])
u66[ i ]~dnorm(0 ,tau .u[66])
u67[ i ]~dnorm(0 ,tau .u[67])
u68[ i ]~dnorm(0 ,tau .u[68])
u69[ i ]~dnorm(0 ,tau .u[69])
u70[ i ]~dnorm(0 ,tau .u[70])
u71[ i ]~dnorm(0 ,tau .u[71])
u72[ i ]~dnorm(0 ,tau .u[72])
}
for(i in 1:N)
{
u[ i ,1]<-u1[ i ]
u[ i ,2]<-u2[ i ]
u[ i ,3]<-u3[ i ]
u[ i ,4]<-u4[ i ]
u[ i ,5]<-u5[ i ]
u[ i ,6]<-u6[ i ]
u[ i ,7]<-u7[ i ]
u[ i ,8]<-u8[ i ]
u[ i ,9]<-u9[ i ]
u[ i ,10]<-u10[ i ]
u[ i ,11]<-u11[ i ]
u[ i ,12]<-u12[ i ]
u[ i ,13]<-u13[ i ]
u[ i ,14]<-u14[ i ]
u[ i ,15]<-u15[ i ]
}

```

```

u[ i,16]<-u16[i]
u[ i,17]<-u17[i]
u[ i,18]<-u18[i]
u[ i,19]<-u19[i]
u[ i,20]<-u20[i]
u[ i,21]<-u21[i]
u[ i,22]<-u22[i]
u[ i,23]<-u23[i]
u[ i,24]<-u24[i]
u[ i,25]<-u25[i]
u[ i,26]<-u26[i]
u[ i,27]<-u27[i]
u[ i,28]<-u28[i]
u[ i,29]<-u29[i]
u[ i,30]<-u30[i]
u[ i,31]<-u31[i]
u[ i,32]<-u32[i]
u[ i,33]<-u33[i]
u[ i,34]<-u34[i]
u[ i,35]<-u35[i]
u[ i,36]<-u36[i]
u[ i,37]<-u37[i]
u[ i,38]<-u38[i]
u[ i,39]<-u39[i]
u[ i,40]<-u40[i]
u[ i,41]<-u41[i]
u[ i,42]<-u42[i]
u[ i,43]<-u43[i]
u[ i,44]<-u44[i]
u[ i,45]<-u45[i]
u[ i,46]<-u46[i]
u[ i,47]<-u47[i]
u[ i,48]<-u48[i]
u[ i,49]<-u49[i]
u[ i,50]<-u50[i]
u[ i,51]<-u51[i]
u[ i,52]<-u52[i]
u[ i,53]<-u53[i]
u[ i,54]<-u54[i]
u[ i,55]<-u55[i]
u[ i,56]<-u56[i]
u[ i,57]<-u57[i]
u[ i,58]<-u58[i]
u[ i,59]<-u59[i]
u[ i,60]<-u60[i]
u[ i,61]<-u61[i]
u[ i,62]<-u62[i]
u[ i,63]<-u63[i]
u[ i,64]<-u64[i]
u[ i,65]<-u65[i]
u[ i,66]<-u66[i]
u[ i,67]<-u67[i]
u[ i,68]<-u68[i]
u[ i,69]<-u69[i]
u[ i,70]<-u70[i]
u[ i,71]<-u71[i]
u[ i,72]<-u72[i]
}
}

```

### 7.3.1. Datos

```

list(N=48, T=72,
y=structure(
.Data=c(3,0,0,0,0,0,0,0,1,0,0,1,0,0,3,0,0,0,2,2,0,0,0,0,0,4,1,0,0,1,0,0,0,1,1,0,2,0,1,0,0,0,0,0,0,0,0,8,0,1,0,
0,0,0,0,3,0,0,0,0,0,0,9,0,0,0,0,0,0,0,0,0,3,2,0,0,8,0,0,0,2,0,0,1,0,0,0,0,0,0,0,0,0,0,0,4,0,0,0,1,12,2,4,0,0,0,4,
7,4,24,4,3,8,6,4,4,1,1,1,0,1,4,9,3,0,13,0,1,1,7,8,0,0,0,8,3,0,3,12,0,0,12,0,2,0,6,2,4,4,3,1,0,0,2,9,3,10,2,6,
12,8,6,6,3,2,3,2,1,1,8,15,7,1,19,0,3,2,23,4,3,6,1,10,7,1,7,25,0,3,4,0,21,37,20,11,4,21,3,1,4,17,15,3,1,5,22,8,9,
5,7,7,6,11,0,10,1,1,7,42,5,1,7,0,0,11,78,16,14,2,1,7,5,1,5,6,16,2,4,16,12,105,17,13,8,25,1,0,1,46,12,7,5,4,13,1,
0,1,14,14,0,13,2,13,1,0,0,12,0,2,30,0,2,13,34,8,20,1,2,8,0,0,0,12,45,0,8,44,12,5,5,4,5,8,1,0,0,2,5,3,4,1,13,1,0,
1,8,9,2,4,0,4,1,0,2,9,2,0,14,0,0,4,8,5,5,3,0,4,2,0,2,8,2,0,5,2,6,0,1,3,2,6,0,0,0,0,0,4,1,2,0,7,0,3,1,5,5,1,3,0,3,
0,0,1,4,2,2,5,0,0,3,3,3,1,3,2,3,2,0,2,4,0,1,2,0,4,0,0,0,2,0,0,1,0,0,0,2,0,0,0,4,0,0,4,0,4,0,2,0,2,0,1,0,0,3,5,0,
2,0,1,1,2,3,3,0,0,0,0,1,0,0,2,0,1,0,0,0,0,1,0,0,0,0,0,0,2,1,0,0,1,4,4,0,0,0,0,0,1,0,2,0,0,0,4,0,0,0,0,0,1,6,0,0,
0,0,0,0,0,0,0,1,0,1,1,0,2,0,0,0,0,0,0,3,0,4,0,0,2,2,0,2,0,1,0,0,0,0,0,0,0,1,1,0,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,1,0,0,0,0,0,0,0,0,0,1,1,1,0,0,1,0,0,0,0,0,0,0,0,2,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,3,2,0,0,1,0,0,8,2,8,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,3,0,5,

```























# Bibliografía

- [1] Jorge E Araña. Métodos bayesianos en economía de la salud. *Cuadernos económicos de ICE*, (67):299–320, 2004.
- [2] Jorge E Arana y Carmelo J León. Willingness to pay for health risk reduction in the context of altruism. *Health Economics*, 11(7):623–635, 2002.
- [3] Sudipto Banerjee, Bradley P Carlin, y Alan E Gelfand. *Hierarchical modeling and analysis for spatial data*. Crc Press, 2014.
- [4] Julian Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, págs. 192–236, 1974.
- [5] Nicola G Best, Katja Ickstadt, y Robert L Wolpert. Spatial poisson regression for health and exposure data measured at disparate resolutions. *Journal of the American statistical association*, 95(452):1076–1088, 2000.
- [6] William M Bolstad. *Introduction to Bayesian statistics*. John Wiley & Sons, 2013.
- [7] Oliver J Brady, Peter W Gething, Samir Bhatt, Jane P Messina, John S Brownstein, Anne G Hoen, Catherine L Moyes, Andrew W Farlow, Thomas W Scott, y Simon I Hay. Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl Trop Dis*, 6(8):e1760, 2012.
- [8] J. Bruin. newtest: command to compute new test @ONLINE, 2011.
- [9] Francisco Caballero. *Selección de modelos mediante criterios de información en análisis factorial. Aspectos teóricos y computacionales*. PhD thesis, Universidad de Granada, 2011.
- [10] Katty C Castillo, Birthe Körbl, Anna Stewart, Javier F Gonzalez, y Facundo Ponce. Application of spatial analysis to the examination of dengue fever in guayaquil, ecuador. *Procedia Environmental Sciences*, 7:188–193, 2011.
- [11] David Clayton y John Kaldor. Empirical bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, págs. 671–681, 1987.

- [12] Mario Fernando Coba Cisneros. *Modelización multinivel con estimación Bayesiana mediante Cadenas de Márkov*. PhD thesis, QUITO: 2014, 2014.
- [13] Jaramillo A (EL COMERCIO). Los casos de dengue se registran en zonas cada vez más altas@ONLINE, 2015.
- [14] Peter Congdon. *Bayesian statistical modelling*, volume 704. John Wiley & Sons, 2007.
- [15] Universidad de Granada. Prácticas con spss @ONLINE, 2016.
- [16] Instituto de la Armada. Estación lluviosa con condiciones normales en la costa del ecuador según el erfen. <http://http://www.inocar.mil.ec/web/index.php/comunicamos/341-estacion-lluviosa-con-condiciones-normales-en-la-costa-del-ecuador-segun-el-erfen>, 2015.
- [17] Organizacion Mundial de la Salud. Dengue y dengue grave @ONLINE, 2015.
- [18] Organización Panamericana de la Salud. Estrategia de gestión integrada para la prevención y control del dengue en ecuador@ONLINE, 2007.
- [19] Petros Dellaportas, Adrian FM Smith, y Photis Stavropoulos. Bayesian analysis of mortality data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 164(2):275–291, 2001.
- [20] Francesca Dominici. Lab 8: Introduction to winbugs. <http://http://www.biostat.jhsph.edu/~fdominic/teaching/-bio656/labs/Lab8.IntroWinBUGS.pdf>, 2014.
- [21] Werther Fernando Fernández y JOSÉ IANNACONE. Variaciones de tres índices larvarios de aedes aegypti (l.) (diptera: Culicidae) y su relación con los casos de dengue en yurimaguas, perú, 2000-2002. *Parasitología latinoamericana*, 60(1-2):3–16, 2005.
- [22] FLACSO. El desarrollo urbano de guayaquil y los problemas de vivienda y servicios. <http://www.flacsoandes.edu.ec/biblio/catalog/resGet.php?resId=16807>.
- [23] Montserrat García Gómez. Los mapas de riesgos: concepto y metodología para su elaboración. *Revista de sanidad e higiene pública*, 68(4):443–453, 1994.
- [24] Roxy González, Saba Infante, y Aracelis Hernández. Modelos jerárquicos espacio temporales para mapear riesgos relativos de dengue, en el municipio girardot, estado aragua, venezuela. *Boletín de Malariología y Salud Ambiental*, 52(1):33–45, 2012.

- [25] Víctor M Guerrero. *Desestacionalización de series de tiempo económicas*. Banco de México, Subdirección de Investigación Económica, 1983.
- [26] Kourí Gustavo. El dengue, un problema creciente de salud en las américa. *Revista Panamericana de Salud Pública*, 19(3):143–145, 2006.
- [27] INEC. Inec presenta resultados definitivos del censo de población y vivienda @ONLINE, 2016.
- [28] A Kikut y A Ocampo. Ajuste estacional de series económicas con tramo/seats y census x12-arima (nota técnica). *Costa Rica: Banco Central de Costa Rica*, 2005.
- [29] Leonhard Knorr-Held y Nikolaus Becker. Bayesian modelling of spatial heterogeneity in disease maps with application to german cancer mortality data. 1999.
- [30] Andrew B Lawson, William J Browne, y Carmen L Vidal Rodeiro. *Disease mapping with WinBUGS and MLwiN*, volume 11. John Wiley & Sons, 2003.
- [31] Peter M Lee. *Bayesian statistics: an introduction*. John Wiley & Sons, 2012.
- [32] Rachel Lowe, Trevor C Bailey, David B Stephenson, Richard J Graham, Caio AS Coelho, Marilia Sa Carvalho, y Christovam Barcellos. Spatio-temporal modelling of climate-sensitive disease risk: Towards an early warning system for dengue in brazil. *Computers & Geosciences*, 37(3):371–381, 2011.
- [33] Scott M Lynch. *Introduction to applied Bayesian statistics and estimation for social scientists*. Springer Science & Business Media, 2007.
- [34] N Monsalve, Y Rubio, y M Pérez. Modelaje bayesiano espacio-temporal de factores asociados con la incidencia del dengue en el área metropolitana de maracay, venezuela. *Bol. Mal. Salud Amb*, 50:219–232, 2010.
- [35] Nora Monsalve Graterol. Modelaje bayesiano espacio-temporal de factores asociados con la incidencia del dengue en el área de maracay. In *Actas del XXX Congreso Nacional de Estadística e Investigación Operativa y de las IV Jornadas de Estadística Pública*, pág. 190, 2007.
- [36] Nora Monsalve Graterol. *Modelos jerárquicos bayesianos espaciales en epidemiología agrícola*. PhD thesis, 2013.
- [37] MSP. Gaceta epidemiológica ecuador sive-alerta, 2010-2015.
- [38] Andrew S Mugglin, Bradley P Carlin, Liz Zhu, y Erin Conlon. Bayesian areal interpolation, estimation, and smoothing: an inferential approach for geographic information systems. *Environment and Planning A*, 31(8):1337–1352, 1999.

- [39] Ioannis Ntzoufras. *Bayesian modeling using WinBUGS*, volume 698. John Wiley & Sons, 2011.
- [40] Diego Andrés Pérez Ruiz. Gibb direccional optimo: Distribución normal truncada. 2011.
- [41] Bedoya Pilozo, Sergio A Pincay Montes, et al. Estudio comparativo del ciclo de desarrollo y sobrevivencia del aedes aegypti durante la época seca y época lluviosa en la ciudad de guayaquil. 2015.
- [42] Rubalcaba R. Modelización y estadística en biología @ONLINE, 2016.
- [43] Agustín Alonso Rodríguez. El resurgir de thomas bayes. *Anuario jurídico y económico escurialense*, (41):327–360, 2008.
- [44] Mauricio Sadinle et al. Metodología para interpolar tamaños poblacionales. *Documentos de CERAC*, 7, 2008.
- [45] Melvin Morera Salas. Modelos bayesianos espacio-temporales aplicados a las hospitalizaciones por problemas de salud susceptibles de cuidados ambulatorios en costa rica. *Economía y salud: boletín informativo*, (72):12–13, 2011.
- [46] David Spiegelhalter, Andrew Thomas, Nicky Best, y Dave Lunn. Winbugs user manual, 2003.
- [47] Merritt R Stevenson. *Variaciones estacionales en el Golfo de Guayaquil, un estuario tropical*. Instituto Nacional de Pesca, 1981.
- [48] Desireé Egleé Villalta. Modelaje bayesiano espacio-temporal de la incidencia de malaria en el estado sucre-venezuela. 2009.
- [49] Coro Chasco Yrigoyen. Modelos De Heterogeneidad Espacial. *Econometrics* 0411004, EconWPA, November 2004.