

**UNIVERSIDAD NACIONAL DE INGENIERÍA  
FACULTAD DE INGENIERÍA ECONÓMICA Y CCSS  
Escuela Profesional de Ingeniería Estadística**



**DETERMINACIÓN DE UN MODELAMIENTO ESPACIO-TEMPORAL PARA EL  
MAPEO DE LA INCIDENCIA DE MALARIA EN LA SELVA AMAZONICA DEL  
DEPARTAMENTO DE LORETO**

**TESIS**

**PARA OPTAR EL TITULO PROFESIONAL DE  
INGENIERO ESTADÍSTICO**

**Bach. Victor Omar Varas Villafuerte**

**Lima – Perú  
2011**

**Digitalizado por:**

**Consortio Digital del  
Conocimiento MebLatam,  
Hemisferio y Dalse**

**DETERMINACIÓN DE UN MODELAMIENTO ESPACIO-TEMPORAL PARA EL  
MAPEO DE LA INCIDENCIA DE MALARIA EN LA SELVA AMAZONICA DEL  
DEPARTAMENTO DE LORETO**

**Tesis**

**Para obtener el título profesional de Ingeniero Estadístico  
en la Escuela Profesional de Ingeniería Estadística**

**Por**

**Bach. Victor Omar Varas Villafuerte**

**-2011 -**

## Resumen

La presente investigación se avoca a proponer un modelo Espacio Temporal para la predicción de la Tasa de Incidencia de la Malaria en el Departamento de Loreto en función al nivel de precipitaciones, a la zona geográfica del área a analizar y su evolución en el tiempo de manera que podamos realizar un mapeo de la tasa de incidencia de la enfermedad que sea interpretable con respecto a las variaciones en la intensidad de precipitaciones pluviales y épocas del año dado que nuestra base bibliográfica nos señala una clara relación causal de éstas variables.

Para esto hemos recurrido al Modelamiento Bayesiano mediante modelos Log-Poisson para las distribuciones a priori, con variantes en su estructura asociada a las variables señaladas, los cuales serán comparados con el Criterio de Mínima pérdida de Información (DIC) a fin de establecer el más óptimo, recurriendo al método MCMC para la obtención de los intervalos de Credibilidad para los parámetros que han de utilizarse en el modelo.

**Palabras Clave:** Riesgo Relativo, Métodos MCMC, Modelos Log-Poisson, Intervalos de Credibilidad.

A mis padres por apoyarme  
en las decisiones decisivas de mi vida.  
A mi hermana como ejemplo para ella.  
A mis compañeros de Estudios  
a los cuales los tengo siempre presente  
y en especial a mi profesor y amigo  
Lic. Luis Huamanchumo de la Cuba  
que me motivo a terminar este trabajo  
de manera exitosa.

## INDICE

<b>INDICE</b>	<b>6</b>
<b>INDICE DE CUADROS</b>	<b>8</b>
<b>INDICE DE GRÁFICOS</b>	<b>9</b>
<b>INDICE DE FIGURAS</b>	<b>10</b>
<b>INTRODUCCIÓN</b>	<b>11</b>
<b>CAPITULO I: Marco Teórico</b>	<b>15</b>
<b>1.1 Conceptos Generales</b>	<b>15</b>
<b>1.2 Términos Básicos</b>	<b>39</b>
<b>CAPITULO II: Planteamiento del Problema</b>	<b>42</b>
<b>2.1 Identificación y Determinación del Problema</b>	<b>42</b>
<b>2.2 Formulación del Problema</b>	<b>43</b>
<b>2.3 Importancia y Alcances de la Investigación</b>	<b>47</b>
<b>2.4 Limitaciones de la Investigación</b>	<b>48</b>
<b>CAPITULO III: Metodología</b>	<b>50</b>
<b>3.1 Propuesta de Objetivos</b>	<b>50</b>
<b>3.2 Sistema de Hipótesis</b>	<b>51</b>

<b>3.3 Operacionalización de las Variables</b>	<b>51</b>
<b>3.4 Tipo y Método de la Investigación</b>	<b>52</b>
<b>3.4 Diseño de la Investigación</b>	<b>53</b>
<b>3.6 Población y Muestra</b>	<b>58</b>
<b>CAPITULO IV: Resultados y Análisis</b>	<b>60</b>
<b>4.1 Análisis de Datos</b>	<b>60</b>
<b>4.2 Comparación de Modelos Ajustados</b>	<b>62</b>
<b>4.3 Estimación de la Incidencia de Malaria</b>	<b>63</b>
<b>4.4 Estimación de Riesgos Relativos</b>	<b>67</b>
<b>4.5 Estimación de los Coeficientes de la Intensidad de Precipitaciones</b>	<b>72</b>
<b>4.6 Estimación del Efecto de las Áreas circundantes</b>	<b>75</b>
<b>4.7 Análisis de Residuos</b>	<b>77</b>
<b>CONCLUSIONES</b>	<b>81</b>
<b>RECOMENDACIONES</b>	<b>83</b>
<b>BIBLIOGRAFÍA</b>	<b>84</b>
<b>ANEXO A</b>	<b>89</b>
<b>ANEXO B</b>	<b>98</b>

## INDICE DE CUADROS

<b>CUADRO 4.2</b>	<b>63</b>
<b>CUADRO 4.5</b>	<b>72</b>

## INDICE DE GRÁFICOS

<b>GRÁFICO 4.1.1</b>	<b>61</b>
<b>GRÁFICO 4.1.2</b>	<b>62</b>
<b>GRÁFICO 4.2</b>	<b>66</b>
<b>GRÁFICO 4.4</b>	<b>71</b>
<b>GRÁFICO 4.5.1</b>	<b>73</b>
<b>GRÁFICO 4.5.2</b>	<b>74</b>
<b>GRÁFICO 4.6</b>	<b>76</b>
<b>GRÁFICO 4.7.1</b>	<b>77</b>
<b>GRÁFICO 4.7.2</b>	<b>78</b>
<b>GRÁFICO 4.7.3</b>	<b>79</b>
<b>GRÁFICO 4.7.4</b>	<b>80</b>

## INDICE DE FIGURAS

<b>FIGURA 1.2.3</b>	<b>40</b>
<b>FIGURA 3.6</b>	<b>59</b>
<b>FIGURA 4.1</b>	<b>61</b>

## INTRODUCCIÓN

El presente trabajo de investigación pretende dar luces sobre el modelamiento en un espacio temporal de la incidencia de la malaria en el Perú.

La malaria, conocida también como paludismo o terciana, es una enfermedad febril producida por un parásito del género *Plasmodium* que se transmite por mosquitos del género *Anopheles*.

Por lo regular la malaria se presenta en la región selva y en la costa norte del país. Tiene un comportamiento endémico y epidémico en determinadas épocas del año, incluso de manera cíclica. Estas regiones se caracterizan por sus altas temperaturas y humedad relativa que presentan pocas variaciones diarias debido a las frecuentes **precipitaciones pluviales**, lo que favorece la **sobrepoblación de mosquitos** y la transmisión de la enfermedad.

La Malaria cobra cada vez mayor importancia en la salud pública peruana por su elevada tasa de incidencia, sobre todo en las regiones de la selva

amazónica y de la selva central. Ello se debe a sus variaciones climatológicas y ecológicas y a los movimientos de las poblaciones humanas.

La Malaria se distribuye aproximadamente en el 75% del área geográfica del País. Se han llegado a notificar casos de esta enfermedad en las regiones de la costa, en los valles interandinos de la vertiente occidental y oriental y en la selva amazónica.

Durante el 2005 se reportaron aproximadamente 87 mil 564 casos de malaria a nivel nacional de los cuales el 82.9% (72 mil 611) correspondieron a la malaria vivax, con un índice de plasmodium vivax acumulado (IVA) de 2.6 por cada 1000 habitantes. El 17.1% de casos (14 mil 953) correspondieron a la malaria falciparum, con un índice de Plasmodium falciparum acumulado (IFA) de 0.54 por cada 1000 habitantes.

Comparado con los últimos cuatro años, se observó un ligero aumento en el número de casos (80 mil por año). En el 2005 se registró un 7% más de casos de malaria que en el 2004. En los últimos años la malaria se ha presentado con mayor incidencia en la selva. Durante el 2005 la mayoría de casos se registraron en la selva amazónica (76.5% del total) y en la selva central (17.1%). 28 Direcciones de Salud notificaron casos de malaria vivax y 10 de ellas reportaron malaria falciparum. Esta enfermedad se presentó en 105 provincias y 375 distritos, con una población en riesgo de 9 millones de habitantes. El área comprometida

con la transmisión de la malaria vivax involucro a 197 distritos en condición de bajo riesgo, 83 en mediano riesgo, 68 en alto riesgo y 27 en muy alto riesgo.

Tres Direcciones de Salud que notificaron caso de malaria vivax estuvieron clasificadas en el nivel de alto riesgo, 8 en mediano riesgo y 13 en bajo riesgo. En relación a la malaria falciparum, solamente Loreto estuvo en el nivel alto de riesgo de transmisión.

La región amazónica fue la región más afectada en los últimos cuatro años. La enfermedad se presentó con más frecuencia en Loreto, que reporto el 61.8% de la malaria del país. El 60% de los casos fueron notificados desde los distritos de Iquitos, Punchana, San Juan Bautista y Belén.

La región de la costa norte, desde Lambayeque hasta tumbes, reporto en el último año solamente el 1.2% del total de casos a nivel nacional, cifra similar a la de los últimos 4 años.

En menor Magnitud -con un comportamiento **epidémico cíclico** y un predominio al 100% de malaria vivax- se presento esta enfermedad en las direcciones de Salud de la selva central (Ayacucho, Cusco y Junín), donde las tareas de prevención y control resultaron difíciles debido a las características geográficas de la zona. Andahuaylas, Pasco (Oxapampa) y Cajamarca (Contumazá) reportaron brotes de malaria vivax que tuvieron importancia local en el transcurso del 2005. Los casos de malaria falciparum que se presentaron en

las direcciones de salud de Madre de Dios y Jaén fueron reportados desde la misma ciudad de Jaén.

Nuestro tipo de Investigación es experimental de manera que podamos determinar cómo algunas variaciones de efectos aleatorios van evolucionando en el tiempo mediante un modelo dinámico de incidencia de precipitaciones. Tenemos como objetivo obtener un modelo que, relacionando las precipitaciones y el área geográfica, nos permita establecer Riesgos Relativos de Incidencia de Malaria en el departamento de Loreto en un determinado momento, dado que toda nuestra revisión bibliográfica nos muestra evidencia de una importante relación causal entre estas variables.

La importancia de este estudio es que nos permite establecer un mapeo adecuado de la enfermedad en el departamento de Loreto, de manera que se fundamente en las relaciones geográficas, climatológicas y temporales que influyen en la incidencia de la enfermedad. Por otro lado se propone un modelamiento original y simple como precedente a otros modelos Bayesianos que mediante el método MCMC permitan obtener resultados para la toma de decisiones.

## **CAPITULO I MARCO TEÓRICO**

### **1.1 Conceptos Generales**

#### **1.1.1 Modelamiento e Inferencia Bayesiana**

El Desarrollo de la Inferencia Bayesiana se basa en la verosimilitud de los datos. La verosimilitud es la distribución conjunta de los datos evaluada en los valores muestreados. Se puede también considerar como una función descriptora de la dependencia de los parámetros en los valores muestreados. Entonces podría haber 2 interpretaciones para esta función. En efecto el principio de verosimilitud por el que las observaciones entran en juego a través de la función de verosimilitud y solo mediante esta, es una parte fundamental del Paradigma Bayesiano.

Según Lawson, A (2008) el paradigma bayesiano implica que la información contenida en la data está totalmente expresada por esta función de verosimilitud. Más aun, el principio de verosimilitud implica que un evento que

no haya ocurrido no tendría efecto en una inferencia, ya que si un evento no realizado afecta a la inferencia, entonces es que existe alguna información no contenida en la función de verosimilitud.

### 1.1.2 Modelos de Verosimilitud

La verosimilitud para la data  $\{y_i\}$ ,  $i = 1, 2, \dots, m$  está definida por

$$L(y|\theta) = \prod_{i=1}^m f(y_i|\theta) \quad (1.1)$$

Donde  $\theta$  es un vector  $p$  dimensional  $\theta = \{\theta_1, \theta_2, \theta_3, \dots, \theta_p\}$  y  $f(\cdot)$  es una función de masa(densidad) de probabilidad. La asunción hecha aquí es que los valores "muestrales" de  $y$  y dados los parámetros son independientes, y a partir de aquí es posible tomar el producto de las contribuciones individuales, desde que la data es asumida **condicionalmente independiente**. Nótese que en aplicaciones espaciales (Lawson, A : *ibid.*, pp 102-105) los datos pueden no ser incondicionalmente independientes y estarían en efecto correlacionados.

Esta Independencia condicional es una asunción fundamental e importante para muchas aplicaciones de mapeo de enfermedades. La log-verosimilitud es también muy usada en el desarrollo del modelo y es definida como:

$$l(y|\theta) = \text{Ln}(L(y|\theta)) = \sum_{i=1}^m \text{Ln}(f(y_i|\theta)) \quad (1.2)$$

### **1.1.3 Correlación Espacial**

En aplicaciones espaciales es usual encontrar que exista correlación entre unidades espaciales. Esta Correlación Geográfica es relativa a la idea básica de que localizaciones cercanas en el espacio usualmente tienen valores similares de variables resultantes mientras que localizaciones lejanas son usualmente diferentes. Esta correlación Espacial (autocorrelación como es llamada a veces) debe tenerse en cuenta en análisis espaciales. Pueden tener impacto en la estructura y forma del modelo de verosimilitud, tal y como lo señala Lawson, A (2008)

Upton y Fingleton (1985) la definen como la “propiedad de un conjunto de datos situados en un mapa geográfico que muestran un patrón de organización”. Sin embargo, esta definición no está exenta de una cierta subjetividad, ya que la expresión “patrón de organización” puede recibir diversas interpretaciones. Los propios autores tratan de acotar el concepto a través de las siguientes aproximaciones al concepto: “la variación espacial sistemática de los valores leídos a lo largo de un mapa”, o “patrones en los valores registrados en diferentes localizaciones”. Si valores relativamente altos (bajos) de la magnitud elegida en una determinada localización vienen acompañados de valores relativamente altos (bajos) de la misma magnitud en las localizaciones vecinas, puede hablarse de la existencia de correlación positiva. Si, por el contrario, valores relativamente altos (bajos) van

alternándose geográficamente con valores relativamente bajos (altos) de la misma magnitud en localizaciones adyacentes, la correlación será negativa.

#### **1.1.4 Independencia Condicional**

En algunas circunstancias es posible considerar la independencia condicional de la data dados los parámetros de los niveles altos de la jerarquía. Por ejemplo en ejemplos de datos de recuento  $y_i$  de la  $i$ -ésima área podría pensarse que son independientes de otros resultados, dada la información del modelo de parámetros.

En el caso simple de dependencia en el vector de parámetros  $\theta$ , tenemos que condicionando en los parámetros puede permitirse que  $y_i | \theta$  sea asumida como una contribución independiente.

Esto simplifica los casos en que la dependencia solo existe incondicionada (efectos no observados pueden inducir dependencia) Esto sucede a menudo en el mapeo de Enfermedades en donde los factores de confusión que tienen una expresión espacial podrían o no ser medidos en un estudio y su exclusión podría dejar una correlación residual en la data. Nótese que ésta aproximación de correlación no es completamente tomada en cuenta para efectos espaciales como lo pueden ser los efectos de correlación residual luego de la inclusión de confusiones.

Estos efectos podrían ser debido a confusiones no observadas o desconocidas. Alternativamente podrían deberse a la correlación intrínseca del proceso. Por lo tanto la asunción de independencia condicional debe **solo ser válida** si es que la correlación fuese representada en algún lugar dentro del modelo, tal y como lo menciona Lawson A. (2008)

La idea de incluir la correlación espacial en el nivel jerárquico superior a la verosimilitud es una asunción fundamental usualmente usada en modelos Bayesianos de salud en áreas pequeñas. Esto significa que la correlación aparece en la distribución a priori mucho más que en la verosimilitud en sí. A menudo los parámetros son dados así como la priori y se asume que la independencia condicional se aplica en la verosimilitud.

### **1.1.5 Distribuciones Conjuntas con Correlaciones**

Existen situaciones donde la correlación espacial puede ser incorporada incluyendo una distribución conjunta de la data. Por ejemplo si un proceso espacial continuo es observado en lugares de medición (tales como contaminantes del aire, concentración de los químicos del suelo, calidad del agua) entonces a menudo un Proceso Espacial Gaussiano (**PEG**) será asumido. Este proceso asume que cualquier realización del proceso es Normal Multivariado con covarianza espacialmente definida, dentro de esta especificación. Por lo tanto, si estos datos fueran unos datos de salida observados, luego la densidad conjunta debe incluir correlación espacial.

$$[\theta | y, \alpha, \beta] = \frac{\beta^{\alpha^*}}{\Gamma(\alpha^*)} \theta^{\alpha^*-1} \exp(-\theta \beta^*) \quad (1.12)$$

Donde:

$$\alpha^* = \sum y_i + \alpha, \beta^* = \sum e_i + \beta \quad (1.13)$$

.....

Un ejemplo si la data de verosimilitud es Poisson, y existe un parámetro de riesgo relativo común con una distribución priori gamma singular:

$$p(\theta | y) \propto L(y | \theta)g(\theta) \quad (1.6)$$

Donde  $g(\theta)$  es gamma con parámetros  $\alpha, \beta$  es decir  $G(\alpha, \beta)$  y

$$L(y | \theta) = \prod_{i=1}^m (e_i \theta)^{y_i} \exp(-e_i \theta) \quad (1.7)$$

Una notación compacta para este modelo es:

$$y_i | \theta \sim \text{Poisson}(e_i \theta) \quad (1.8)$$

$$\theta \sim G(\alpha, \beta) \quad (1.9)$$

Esto nos lleva a una distribución posterior de parámetros fijos  $\alpha, \beta$  de

$$[\theta | \{y_i\}, \alpha, \beta] = L(y | \theta, \alpha, \beta)p(\theta)/C \quad (1.10)$$

$$\text{donde } C = \int L(y | \theta, \alpha, \beta)p(\theta)d\theta \quad (1.11)$$

En este caso la constante C puede ser calculada directamente y esto nos lleva a otra distribución Gamma:

Alternativamente, cabe considerar datos discretos de resultado donde la correlación es explícitamente modelada. Los modelos autologísticos y Poisson fueron desarrollados para datos de lálices con correlación espacial incluida por la adicción en una vecindad espacial (Besag y Tantrum 2003) En esta aproximación la normalización de la verosimilitud es computacionalmente prohibitiva y un recurso que se hace a menudo es una aproximación de la verosimilitud. (Andrew B. Lawson 2008)

#### **1.1.6 Distribuciones Aprioris**

Todos los parámetros dentro de los modelos Bayesianos son estocásticos y con distribuciones de probabilidad apropiadas asignadas. Por lo tanto un valor único del parámetro es simplemente una realización posible de los valores posibles del parámetro, probabilidad que está definida por la distribución *a priori*.

La distribución *a priori* es una distribución asignada al parámetro antes de ver los datos. Note también que la decodificación de distribuciones aprioris es que proveen datos adicionales para un problema y así es que pueden usarse para mejorar estimación o identificación de parámetros. Para un único parámetro  $\theta$  la distribución priori puede ser denotada por  $g(\theta)$ , mientras que para un vector de parámetros  $\theta$ , la distribución priori conjunta es  $g(\theta)$ .

### 1.1.7 Prioris Impropias

Es posible que una distribución priori pueda ser impropia. La Impropiedad es definida como la condición de integración de la distribución priori de la variable aleatoria  $\theta$  sobre su rango ( $\Omega$ ) no sea finita:

$$\int_{\Omega} g(\theta) d\theta = \infty \quad (1.3)$$

Una distribución priori es impropia si su constante normalizadora es infinita. Mientras la impropiedad es una limitación para cualquier distribución priori, no es necesariamente el caso que una priori impropia nos lleve a la impropiedad de la distribución posterior. La distribución posterior puede ser a menudo propia incluso con la especificación de una priori impropia. (Andrew B. Lawson 2008)

### 1.1.8 Prioris No Informativas

A menudo las distribuciones prioris son asumidas con la condición de que no harán una fuerte preferencia sobre los valores de las variables. Estas son algunas veces es conocidas como distribuciones prioris vagas, o de referencia o plana o no informativa. Usualmente cuando tienen una forma relativamente plana acercándose a una preferencia uniforme para diferentes valores de las variables.

Esto tiende a significar que para cualquier análisis posterior las prioris posteriores tendrán un impacto pequeño comparado con la verosimilitud de la data. Las prioris de Jeffrey fueron desarrolladas en un esfuerzo por encontrar tales prioris de referencia.

La decisión de una priori no informativa puede ser hecha a menudo con algún entendimiento general del rango y del comportamiento de la variable.

Por supuesto que algunas veces es importante ser informativos con las distribuciones prioris. La Identificabilidad es un asunto relativo a la habilidad de distinguir entre parámetros dentro de un modelo paramétrico. En particular, si un rango restringido puede ser asumido para permitir que un número de variables sean identificadas, entonces puede ser importante especificar distribuciones que proveerán tal soporte.

Por último si la verosimilitud tiene poco o nada de información sobre la separación de los parámetros entonces la separación o identificación puede solo venir de una especificación priori. En general si una distribución priori propia es asumida para los parámetros, entonces ellos están identificados en la distribución posterior. Sin embargo cuan identificados se encuentren dependerá de la variabilidad asumida. ( Andrew B. Lawson 2008)

### 1.1.9 Distribución Posterior

La distribución priori y la verosimilitud proporcionan dos fuentes de Información acerca de algún problema. La verosimilitud informa acerca del parámetro vía la data, mientras que la distribución priori informa vía priori creencias y asunciones. Cuando hay una gran cantidad de datos, es decir, el tamaño de muestra es grande, la verosimilitud contribuirá más a la estimación del riesgo relativo. Cuando la data obtenida es pobre, la distribución priori dominará el análisis. (Andrew B. Lawson 2008)

El producto de la verosimilitud con la distribución priori es llamada la distribución posterior. Esta distribución describe el comportamiento del parámetro luego de que la data es observada y las asunciones prioris fueran hechas. La distribución posterior se define por:

$$p(\boldsymbol{\theta} | y) = L(y | \boldsymbol{\theta})g(\boldsymbol{\theta})/C \quad (1.4)$$

Donde  $g(\boldsymbol{\theta})$  es la distribución conjunta del vector de parámetros  $\boldsymbol{\theta}$ . Alternativamente esta distribución puede ser especificada como la proporcionalidad:

$$p(\boldsymbol{\theta} | y) \propto L(y | \boldsymbol{\theta})g(\boldsymbol{\theta}) \quad (1.5)$$

Un ejemplo si la data de verosimilitud es Poisson, y existe un parámetro de riesgo relativo común con una distribución priori gamma singular:

$$p(\theta | y) \propto L(y | \theta)g(\theta) \quad (1.6)$$

Donde  $g(\theta)$  es gamma con parámetros  $\alpha, \beta$  es decir  $G(\alpha, \beta)$  y

$$L(y | \theta) = \prod_{i=1}^m (e_i \theta)^{y_i} \exp(-e_i \theta) \quad (1.7)$$

Una notación compacta para este modelo es:

$$y_i | \theta \sim \text{Poisson}(e_i \theta) \quad (1.8)$$

$$\theta \sim G(\alpha, \beta) \quad (1.9)$$

Esto nos lleva a una distribución posterior de parámetros fijos  $\alpha, \beta$  de

$$[\theta | \{y_i\}, \alpha, \beta] = L(y | \theta, \alpha, \beta)p(\theta)/C \quad (1.10)$$

$$\text{donde } C = \int L(y | \theta, \alpha, \beta)p(\theta)d\theta \quad (1.11)$$

En este caso la constante C puede ser calculada directamente y esto nos lleva a otra distribución Gamma:

$$[\theta | y, \alpha, \beta] = \frac{\beta^{\alpha^*}}{\Gamma(\alpha^*)} \theta^{\alpha^*-1} \exp(-\theta\beta^*) \quad (1.12)$$

Donde:

$$\alpha^* = \sum y_i + \alpha, \beta^* = \sum e_i + \beta \quad (1.13)$$

### 1.1.10 Distribución Predictiva

La distribución posterior resume nuestra comprensión sobre los parámetros dada la data observada y juega un rol importante en el modelamiento bayesiano.

Sin embargo podemos incluso examinar otras distribuciones relacionadas que son a menudo utilizadas cuando es requerida la predicción de nuevo datos (futuros). Definase una nueva observación de  $y$  cómo  $y^*$ . Podemos determinar la distribución predictiva de  $y^*$  de dos formas. En general la distribución predictiva es definida como:

$$p(y^* | y) = \int L(y^* | \theta) p(\theta | y) d\theta \quad (1.14)$$

Aquí la predicción está basada en la marginalización sobre los parámetros en la verosimilitud de los nuevos datos  $L(y^* | \theta)$  usando la distribución posterior

$p(\theta | y)$  para definir la contribución de la data observada en la predicción. Esta es denominada la distribución predictiva posterior.

Una variante de ésta definición usa la distribución priori en lugar de distribución posteriori:

$$p(y^* | y) = \int L(y^* | \theta)p(\theta)d\theta \quad (1.15)$$

Esto hace hincapié en la predicción basado en la distribución priori (antes de observar la data). Nótese que esta distribución es solo la distribución marginal de  $y^*$  (Andrew B. Lawson 2008)

### **1.1.11 Intervalo de Credibilidad**

El Intervalo de Credibilidad del  $(1-\alpha)\%$  para un parámetro  $\theta$  se calcula tomando los valores centrales(alrededor de la moda) que den una probabilidad  $(1-\alpha)\%$  en la distribución Final. Gráficamente se interpreta como un corte horizontal a la distribución posterior de manera que el área comprendida entre la función de distribución y dicho corte sea de  $(1-\alpha)$ ; los puntos de corte serán los limites superior o inferior del intervalo o de los intervalos en caso de formas bimodales. (H. S. Migon and Gamerman 1999)

La interpretación teórica de estos intervalos es que existe una probabilidad de  $(1-\alpha)$  de que el parámetro  $\theta$  se encuentre en el intervalo muy diferente a la interpretación con el Intervalo de Confianza ampliamente utilizado con el Enfoque Clásico en el que se trabaja con errores del Tipo I y II.

#### **1.1.12 Modelamiento Bayesiano Jerárquico**

En el modelamiento Bayesiano el parámetro tiene una distribución. Esta distribución controla la forma del parámetro y es especificada por el investigador basado en, usualmente, en sus creencias prioris concerniente a su comportamiento. En el contexto de mapeo de enfermedades es usual asumir una distribución priori para  $\theta$  en un modelo de verosimilitud Poisson. (Andrew B. Lawson 2008)

El modelado Jerárquico consiste en que la distribución priori del modelo tiene parámetros gobernando su comportamiento los cuales pueden asumir valores pero usualmente el investigador no tiene establecidos valores con certeza para estos parámetros. El investigador requiere estimar estos parámetros a partir de la data por lo que se les asignan distribuciones denominadas hiperprioris, las cuales contendrán parámetros denominados como hiperparámetros.

### 1.1.13 Modelo Lineal Dinámico Normal

El Modelo Lineal Dinámico normal general es definido para un vector de observación  $Y_t$ , evaluando para el caso usual escalar.

Sea  $Y_t$  es un vector de observación ( $r \times 1$ ), para  $t=1,2,\dots$ .

Donde presenta la siguiente estructura

$$Y_t = F_t' \theta_t + v_t \quad v_t \sim N(0, V_t) \quad (1.16)$$

$$\theta_t = G_t \theta_{t-1} + w_t \quad w_t \sim N(0, W_t) \quad (1.17)$$

O también de la siguiente manera:

$$(Y_t / \theta_t) \sim N(F_t' \theta_t, V_t) \quad (1.18)$$

$$(\theta_t / \theta_{t-1}) \sim N(G_t \theta_{t-1}, W_t) \quad (1.19)$$

Para cada tiempo  $t$ , donde:

- $F_t'$  es una matriz conocida de ( $n \times r$ )
- $G_t$  es una matriz conocida de ( $n \times n$ )
- $V_t$  es una matriz varianza conocida de ( $r \times r$ )
- $W_t$  es una matriz varianza conocida de ( $n \times n$ )

Donde la secuencia de errores  $v_t$  y  $w_t$  son interna y mutuamente independientes, definiendo la distribución de muestro para  $y_t$  dado un parámetro  $\theta_t$ . Además los  $y_t$  son independientes de todas las observaciones y los valores de los parámetros. (Mike West and Jeff Harrison 1997)

- $F_t'$  es una matriz de diseño de valores conocidos de variables independientes
- $\theta_t$  es el estado
- $u_t = F_t' \theta_t$  es la respuesta media de un nivel
- $v_t$  es el error observacional
- $G_t$  es la evolución, matriz de estados
- $w_t$  error con evolución de varianza  $W_t$

#### **1.1.14 Simulación de Parámetros vía MCMC**

Los problemas de inferencia paramétrica bayesiana se centran en la obtención de una determinada cantidad a posteriori de interés, que será

función de los parámetros del problema. Si lo que se desea es la inferencia sobre:

$$\theta = (\theta_1, \theta_2, \dots, \theta_p) \in \Theta \quad (1.20)$$

o sobre cualquier transformación del parámetro,  $g(\theta)$ , desde el punto de vista bayesiano, la formulación ha de contemplar la información muestral, dada por los datos  $X = (x_1, x_2, x_3, \dots, x_n)$  con función de verosimilitud  $f(x|\theta)$ , y la elección de la distribución a priori sobre el parámetro  $\Pi(\theta)$ , que recoge la opinión subjetiva del experto.

La combinación de ambos ingredientes conduce a la obtención, vía Teorema de Bayes, de la distribución a posteriori,

$$\pi(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{\int_{\Theta} \pi(\theta)f(x|\theta)d\theta} \propto \pi(\theta)f(x|\theta), \quad (1.21)$$

bajo la que se pretende calcular la cantidad a posteriori de interés, que puede expresarse

$$E[g(\theta)|x] = \int_{\Theta} g(\theta)\pi(\theta|x)d\theta \quad (1.22)$$

De esta forma puede calcularse, por ejemplo, la media a posteriori, si

$$g(\theta) = \theta \quad (1.23)$$

o cualquier momento a posteriori, para

$$g(\theta) = \theta_1^r \cdot \theta_2^s \quad (1.24)$$

o la probabilidad a posteriori de un conjunto, para

$$g(\theta) = I_{\{\theta \in A\}} \quad (1.25)$$

El caso es que tanto la obtención de la densidad a posteriori, como el cálculo de la cantidad a posteriori, contienen integrales no factibles en la mayoría de los casos, que se solían resolver mediante diferentes métodos de integración numérica hasta que se empezaron a utilizar los métodos de simulación Monte Carlo de Cadenas de Markov, MCMC, que permiten simular una cadena que converja a la distribución a posteriori, de manera que cantidades muestrales obtenidas a partir de las realizaciones de dicha cadena converjan a la cantidad a posteriori de interés.

La creciente rapidez de cálculo de los equipos informáticos actuales posibilita la implementación efectiva de estos métodos, en concreto, los investigadores Spiegelhalter, Thomas y Best (1999) han desarrollado un

software para la inferencia bayesiana utilizando MCMC, WinBUGS, que hace que complejos métodos de simulación sean asequibles para los usuarios de la estadística bayesiana aplicada a las distintas disciplinas.

Este trabajo pretende también ser precedente del uso y utilidad del WinBUGS, como herramienta en diferentes problemas de la estadística bayesiana aplicada.

### **1.1.15 Métodos Montecarlo con Cadenas de Markov**

Desde hace aproximadamente 10 años, los métodos MCMC permiten la resolución de problemas que hasta entonces no eran analíticamente tratables y precisaban distintas aproximaciones numéricas para las integrales implicadas, que no tenían en cuenta la naturaleza aleatoria del problema, el carácter probabilístico de las funciones implicadas.

Estos métodos permiten muestrear la distribución a posteriori, aunque ésta sea desconocida, gracias a la construcción de una cadena de Markov cuya distribución estacionaria sea, precisamente,  $\Pi(\theta | x)$

Una cadena de Markov es una sucesión de variables aleatorias  $\{X_t\}$  tal que  $X_{t+1}$  sólo depende del estado actual  $X_t$ , mediante la probabilidad de transición de la cadena  $P(X_{t+1} | X_t)$ .

Bajo ciertas condiciones, la probabilidad de transición de la cadena no depende del instante 't', y converge a una distribución estacionaria  $\Phi$ , de manera que si:

$$X_t \xrightarrow{t \rightarrow \infty} X \quad (1.26)$$

Con distribución:

$$\Phi \Rightarrow \bar{g}_N = \frac{1}{N} \sum_{i=1}^N g(X_i) \xrightarrow{N \rightarrow \infty} E_{\Phi}[g(X)] \quad (1.27)$$

Se trata, por tanto, de simular una cadena de Markov  $\Theta$ ,  $\{\theta^{(t)}\}$  sobre cuya distribución estacionaria sea  $\Pi(\theta | x)$ , lo que permitirá que:

$$E[g(\theta) | x] \approx \frac{1}{N-m} \sum_{i=m}^N g(\theta^{(i)}) = \bar{g}_{N-m} \quad (1.28)$$

Donde 'm' indica el número de realizaciones de la cadena que se desechan para evitar correlación (muestra "burn in"), y 'N' debe ser suficientemente grande como para garantizar la convergencia (ver Gilks et al, 1996, o Robert y Casella, 1999).

Para diseñar la cadena  $\{\theta^{(t)}\}$ , se muestrearán (mediante simulación Monte Carlo) de distribuciones apropiadas (ya que no se puede hacer directamente de  $\Pi(\theta | x)$ ).

Cuando las distribuciones utilizadas en la simulación son las condicionadas a posteriori de cada parámetro con respecto al resto (porque tengan forma estándar, a diferencia de la distribución a posteriori) se habla de Muestreo de Gibbs. Mientras que el Algoritmo de Metrópolis-Hastings (que incluye el muestreo de Gibbs como caso particular) utiliza distribuciones más generales para la simulación de la cadena. Para un tratamiento detallado de estos tipos de muestreo, pueden consultarse Gilks et al (1996), Robert y Casella (1999), o Chen et al (2000).

El siguiente paso es la implementación informática de estos procedimientos, que puede llevarse a cabo mediante la programación específica de cada problema tratado, o haciendo uso del software.

#### **1.1.16 Criterio de Información de Devianza**

El criterio de Información de Devianza (DIC) es una generalización de modelado jerárquico de la AIC (Criterio de Información de Akaike) y BIC (criterio de información bayesiano, también conocido como el criterio de Schwarz). Es particularmente útil en problemas Bayesiano de selección de modelos donde la distribución posterior de los modelos se han obtenido por el método de simulación Montecarlo con Cadenas de Markov (MCMC).

Al igual que la AIC y el BIC es una aproximación asintótica cuando el tamaño de la muestra se hace grande. Sólo es válido cuando la distribución posterior es aproximadamente normal multivariada.

Definimos la pérdida de Información como

$$D(\theta) = -2 \log(p(y|\theta)) + C \quad (1.29)$$

Donde  $y$  son los datos,  $\theta$  son los parámetros desconocidos del modelo y  $p(y|\theta)$  es la función de verosimilitud.  $C$  es una constante que se anula en todos los cálculos que comparan los diferentes modelos, y que por tanto no necesita ser conocido. La esperanza  $\bar{D} = E^\theta(D(\theta))$  es una medida de qué tan bien el modelo se ajusta a los datos, mientras más grande sea, menor será el ajuste.

El número efectivo de parámetros del modelo se calcula como  $P_D = \bar{D} - D(\bar{\theta})$ , Donde  $\bar{\theta}$  es la esperanza de  $\theta$ . Mientras más grande sea éste, más fácil es para el modelo ajustarse a los datos. El criterio de información de Devianza se calcula como:

$$DIC = P_D - \bar{D} \quad (1.30)$$

La idea es que los modelos con menor DIC serán preferibles a los modelos con mayor DIC. Los modelos están penalizados tanto por el valor de  $\bar{D}$ , no solo favorece a un buen ajuste, sino también (en común con AIC y BIC) al número efectivo de los parámetros.

Desde que  $\bar{D}$  disminuirá a medida que el número de parámetros en un modelo de incrementos,  $p_D$  compensa este efecto, favoreciendo a los modelos con un menor número de parámetros.

La ventaja éste sobre otros criterios en el caso de la selección del modelo bayesiano es que el DIC es fácilmente calculado a partir de las muestras generadas por simulación Montecarlo con Cadenas de Markov.

AIC y el BIC requieren el cálculo de las probabilidades en su máximo valor de  $\theta$ , el cual no es fácilmente disponible a partir de la simulación MCMC. Pero para el cálculo de la DIC, simplemente calcular  $\bar{D}$  como el promedio de  $D(\theta)$  sobre las muestras de  $\theta$ , y  $D(\bar{\theta})$  como el valor de  $D$  evaluado en la media de las muestras de  $\theta$ . Luego, el DIC se desprende directamente de estas aproximaciones. ( Spiegelhalter D J, Best N G, Carlin B P y van der Linde A 2002)

### **1.1.17 Antecedentes Históricos de la Enfermedad**

La malaria ha infectado a los humanos durante más de 50.000 años, y puede que haya sido un patógeno humano durante la historia entera de nuestra especie. De cierto, especies cercanas a los parásitos humanos de la malaria se han encontrado en los chimpancés, pariente ancestral de los humanos. Se encuentran referencias de las peculiares fiebres periódicas de la malaria a lo

largo de la historia, comenzando desde 2700 a. C. en China. El término malaria proviene del italiano de la edad media: mala aria — "mal aire"; y se le llamó también paludismo, del latín "palus" (pantano).

Los estudios científicos sobre la malaria hicieron su primer avance de importancia en 1880, cuando el médico militar francés **Charles Louis Alphonse Laveran**, trabajando en Argelia, observó parásitos dentro de los glóbulos rojos de personas con malaria. Propuso por ello que la malaria la causaba un protozooario, la primera vez que se identificó a un protozooario como causante de una enfermedad. Por este y otros descubrimientos subsecuentes, se le concedió el Premio Nobel en Fisiología o Medicina en 1907. Al protozooario en cuestión se le llamó Plasmodium, por los científicos italianos Ettore Marchiafava y Angelo Celli. Un año después, Carlos Finlay, un médico cubano que trataba pacientes con fiebre amarilla en la Habana, sugirió que eran los mosquitos quienes transmitían la enfermedad de un humano a otro. Posteriormente, fue el británico **Sir Ronald Ross**, trabajando en la India, quien finalmente **demostró en 1898 que la malaria era transmitida por los mosquitos**. Lo probó al mostrar que ciertas especies del mosquito transmitían la malaria a pájaros y aislando los parásitos de las glándulas salivales de mosquitos que se alimentaban de aves infectadas. Por su aporte investigador, Ross recibió el premio Nobel de Medicina en 1902. Después de renunciar al Servicio Médico de la India, Ross trabajó en la recién fundada Liverpool School of Tropical Medicine y dirigió los esfuerzos por controlar la malaria en Egipto, Panamá, Grecia y Mauricio. Los hallazgos de Finlay y Ross fueron

confirmados luego por un comité médico dirigido por Walter Reed en 1900, y sus recomendaciones implementadas por William C. Gorgas en medidas de salud adoptadas durante la construcción del Canal de Panamá. Este trabajo salvó la vida de miles de trabajadores y ayudó a desarrollar los métodos usados en campañas de salud pública contra la malaria.

El primer tratamiento eficaz para la malaria fue la corteza del árbol Cinchona, que contiene el alcaloide quinina. Este árbol crece en las colinas de los Andes, en particular en Perú. Los habitantes del Perú usaban el producto natural para controlar la malaria, y los Jesuitas introdujeron esta práctica en Europa durante los años 1640, donde fue aceptada con rapidez. Sin embargo, no fue sino hasta 1820 cuando la quinina, el ingrediente activo, fue extraída de la corteza y nombrada por los químicos franceses Pierre Joseph Pelletier y Jean Bienaime Caventou.

A comienzos del siglo XX, antes de los antibióticos, los pacientes con sífilis eran intencionalmente infectados con malaria para crear una fiebre, siguiendo las investigaciones de Julius Wagner-Jauregg. Al controlar la fiebre con quinina, los efectos tanto de la sífilis como la malaria podían ser minimizados. Algunos de los pacientes murieron por la malaria, pero el riesgo era preferible por encima de la casi segura muerte por sífilis.

A pesar de que en el estadio sanguíneo y en el mosquito del ciclo de vida de la malaria se estableció en el siglo XIX y a comienzos del siglo XX, solo en 1980 se observó la forma latente hepática del parásito. Este descubrimiento

explicó finalmente por qué daba la impresión de que algunas personas se curaban de la enfermedad, para recaer años después de que el parásito hubiese desaparecido de su circulación sanguínea.

## **1.2 Términos Básicos**

### **1.2.1 Riesgo Relativo**

Se define como el cociente entre la tasa de incidencia de la enfermedad, en este caso la malaria, en una zona de estudio y la tasa de incidencia esperada sin presencia de comportamiento endémico. Obtener riesgos relativos mayores a uno indica un alto riesgo en la zona de estudio, por otro lado riesgos menores a uno indican una caída en la tasa de incidencia esperada lo cual hace que la zona sea menos riesgosa.

### **1.2.2 Tasa de Incidencia**

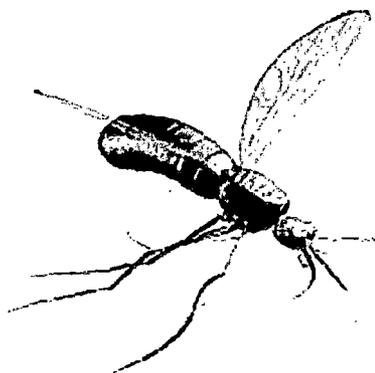
Se define como el cociente entre el número de enfermos en una zona determinada entre el número de habitantes de dicha zona, interpretable como la probabilidad bruta de adquirir la enfermedad.

### 1.2.3 Malaria

La malaria (mal aire), paludismo (laguna), fiebre intermitente, fiebre de los pantanos, fiebre palustre, es una enfermedad infecciosa, endémica en el Perú producida por protozoarios del género Plasmodium registrando para el Perú sólo las especies: P.Vivax con el 99% de las incidencias, P.Falciparum (ver figura N1.1), con el 0.2% de las incidencias P. Malariae con el 0.8% de las incidencias y transmitidas por el mosquito del género Anopheles.

Figura N°1.2.3

El mosquito transmisor de la Malaria P.Falciparum



El hombre es el único reservorio importante de la malaria humana aunque los monos de especies superiores pueden albergar el P. malariae. En la naturaleza los monos se infectan con muchas especies de malaria incluidos el P. Knowlesi, P. cynomolgi, P. brasilianum, P. inui, P. Schwetzi y P. simiun, todos los cuales pueden infectar al hombre, pero la transmisión habitual no es común.

Los síntomas característicos son fiebre, escalofríos, sudoración y dolor de cabeza. Además se puede presentar náuseas, vómitos, tos, heces con sangre, dolores musculares, ictericia, defectos de la coagulación sanguínea, shock, insuficiencia renal o hepática, trastornos del sistema nervioso central y coma.

Como es sabido, tan sólo las hembras de mosquitos son las que se alimentan de sangre para poder madurar los huevos, y por tanto los machos no pican y no pueden transmitir enfermedades ya que únicamente se alimentan de néctares y jugos vegetales. La fecundidad de la hembra, así como la fertilidad del macho, depende de la temperatura. Para el desarrollo de los huevos necesitan del agua, otros hidratos de carbono o bien proteínas. Las Hembras necesitan tomar sangre de vertebrados para permitir la maduración de los huevos.

## **CAPÍTULO II**

### **PLANTEAMIENTO DEL PROBLEMA**

#### **2.1 Identificación y Determinación del Problema**

Dado que en el País la enfermedad de la malaria tiene un relevante importancia por su naturaleza endémica, específicamente en la amazonia peruana constituye uno de los mayores focos infecciosos del país por sus naturales reservorios de humedad y por los niveles de temperaturas elevadas que se alcanzan en la zona, los cuales se constituyen en dos condiciones ideales como habitat del mosquito 'Anopheles' transmisor de la malaria.

Actualmente la incidencia de malaria en la región amazónica no tiene un patrón definido de comportamiento georeferenciado o que se encuentre fundamentado en los cambios ambientales como es la caída de lluvia. Existe una necesidad de un mapeo controlado en el tiempo adecuado de los riegos relativos de la enfermedad para que así podamos conocer que zonas se encuentran en niveles de riego altos.

## **2.2 Formulación del Problema**

### **2.2.1 Supuesto Principal**

El riesgo relativo de adquirir el virus, causante de la malaria falciparum o vivax, depende linealmente de los niveles de intensidad de lluvias (mm/h) y el efecto de las provincias circundantes, los cuales constituyen información a priori en su proceso de estimación

Este supuesto se basa en que el único medio de transmisión de la enfermedad es la picadura del mosquito del género anopheles el cual se multiplica en aguas estancadas propias de la región amazónica y cuyos niveles dependen de las precipitaciones pluviales y de la provincia en estudio así como las circundantes, esto se basa en las Investigaciones realizadas por Finlay y Ross.

### **2.2.2 Modelo Espacial**

En otros estudios de mapeo de enfermedades en un principio se asumió que la región de interés se ha dividido en  $n$  subregiones contiguas, y que  $Y_i$  representa el número de casos de la enfermedad para cada zona  $i = 1, \dots, n$ . En la literatura clásica, se supone que  $Y_i$  sigue una distribución de Poisson con media  $r_i * e_i$ , donde  $e_i$  es el conteo esperado de la enfermedad y  $r_i$  representa el riesgo relativo, es decir:

$$y_i = \text{Poisson}(e_i r_i) \quad (2.1)$$

Los conteos esperados es una cantidad conocida sobre la base de factores de riesgo comunes, y desempeñan un papel importante en la normalización de la información de las zonas de diferentes características. Por lo tanto, las cantidades desconocidas de interés son los riesgos relativos

Ya está bien documentado que los mapas de la enfermedad basada en la estimación de máxima verosimilitud (MLE) de  $r_i$ , que se denota aquí por  $r_i = \hat{y}_i/e_i$  **los ratios en crudo son engañosos** por ejemplo cuando la enfermedad es poco frecuente y/o las zonas son pequeños (Mollie, 1996).

Además, las estimaciones de Máxima Verosimilitud **no dan cabida** a la dependencia espacial posible. Otro aspecto a considerar es la homogeneidad dentro de cada área, donde los riesgos individuales se espera que sean iguales dentro de la zona.

Por lo tanto, es natural que se incluyan una variación de Poisson extra en el modelo para que los riesgos relativos estimados tengan en cuenta este efecto.

La literatura sobre la modelización de los riesgos relativos es considerablemente rica. Una referencia clásica es Besag et al. (1991), que supone que, para el área  $i$ , el registro de los riesgos relativos se modela a través

$$\log(r_i) = \alpha + \beta x_i + b_i + u_i \quad (2.2)$$

donde  $\alpha$  es un intercepto común para la región entera,  $x_i$  es el vector de covariables y  $u_i$  y  $b_i$  son términos de efectos aleatorios, tal que  $u_i \sim N(0, \sigma_u)$ ,  $i = 1, 2, \dots, n$

$$\frac{b_i}{b_j}, j \neq i \sim N\left(\frac{\sum_{j \in \delta_i} W_{ij} b_j}{\sum_{j \in \delta_i} W_{ij}}, \frac{\sigma_b}{\sum_{j \in \delta_i} W_{ij} b_j}\right), \quad i = 1, 2, \dots, n \quad (2.3)$$

donde  $\delta_i$  es el conjunto que contiene las zonas adyacentes a  $i$ , y  $W_{ij}$  es el peso que tiene la zona  $j$  vecinos en  $i$ . La más simple y utilizada estructura de vecinos se basa en las zonas adyacentes, donde las regiones vecinas comparten una frontera común. Este será el caso considerado en este trabajo.

Esto implica que  $W_{ij}=1$  si  $i$  y  $j$  comparten fronteras ( $i \sim j$ ) y 0 en caso contrario. La distribución anterior se conoce en la literatura como autoregresiva condicional (CAR) previa, y puede ser denotado aquí por  $CAR(\sigma_b)$ . Referencias adicionales utilizando distribuciones previas CAR para el mapeo de la enfermedad son Bernardinelli et al. (1995), Best et al. (1999), Kelsal y Wakefield (2001) y sus referencias. Mejor et al. (1999) investigar la elección de la estructura de vecindad y la especificación de las distribuciones hyperpriori.

Por otra parte, Keisal y Wakefield (2001) modeló los riesgos relativos a través de un campo de Gauss al azar donde se deriva la estructura de correlación de una superficie permanente de los riesgos subyacentes. Afirman que esto da lugar a estructuras de correlación más realistas a las zonas adyacentes

Sin embargo con estos antecedentes, nuestro problema es que no existe un modelo que además de captar la **correlación espacial** entre las áreas del mapeo geográfico, permita también estimar Riesgos Relativos como **componentes estacionales o tendenciales** dependientes de la época del año o de alguna otra variable correlacionada con la incidencia de la enfermedad como lo es la caída de lluvias que se relaciona con el hecho de que los mosquitos tienen más facilidad de reproducción cuando hay abundancia de agua en el caudal de los ríos, además del clima tropical propio de la región.

A todo esto tenemos las siguientes preguntas de Investigación:

**¿Será posible establecer un modelo que nos permita mapear adecuadamente los Riesgos Relativos de la enfermedad que capte las estructuras de correlación espacial y temporal además de su relación con las Precipitaciones Pluviales?**

**¿El riesgo relativo de adquirir el virus de la Malaria dependería de la intensidad de lluvias de la zona y sus áreas circundantes?**

**¿La incidencia de la enfermedad depende de los niveles de intensidad de lluvia de la zona y del efecto de las provincias circundantes?**

**¿Las precipitaciones pluviales tienen importancias equivalentes para todos los meses?**

**¿Los riesgos de adquirir la enfermedad son igual para todas las provincias del departamento de Loreto?**

### **2.3 Importancia y Alcances de la Investigación**

La importancia de la presente investigación está fundamentada por la relevancia de la enfermedad de la Malaria en la zona en estudio. Y por la necesidad de sumar esfuerzos contribuir la mejor forma de pronosticar la precipitaciones a relacionadas a la incidencia de la enfermedad de la Malaria en esta zona del departamento de Loreto, que califica como zona de Alto Riesgo para dicha enfermedad, según los criterios establecidos en el Programa de Control de Malaria

Esta enfermedad es de fácil y rápida transmisión, provocada por agentes patógenos. El ser vivo o agente patógeno que las produce recibe el nombre de

agente etiológico o causal. En algunas ocasiones para que se produzca la enfermedad es necesaria la intervención de otro organismo viviente llamado agente intermediario, transmisor o vector.

Los agentes patógenos de este tipo de enfermedades generalmente son virus o bacterias por lo que un modelo que establezca una relación cuantificable del efecto del agente intermediario puede ayudar a establecer medidas de control enfocadas a la prevención atacando dichos agentes.

La incidencia de enfermedades infecto-contagiosas, es una característica común en algunas zonas endémicas del territorio peruano. Por lo tanto, es de necesidad el desarrollo de modelos apropiados para el Monitoreo, Supervisión y Control de la incidencia de estas enfermedades.

En tal sentido la presente investigación brinda un aporte sustancial que suma esfuerzos, para la mejor atención de la zona de la selva amazónica de Loreto frente a la enfermedad de la Malaria.

#### **2.4 Limitaciones de la Investigación**

Debido al alto costo que significa realizar un estudio más detallado de todos los factores potenciales que contribuyen en cada zona a la proliferación de ésta enfermedad nuestro modelo es global y adecuado para zonas de extensa

envergadura por lo que no vamos a obtener un resultado que nos indique a detalle la problemática de cada zona.

La información de precipitaciones pluviales es disponible puntualmente por lo que asumiremos que en cada área hay niveles de lluvia muy similares.

## **CAPÍTULO III METODOLOGÍA**

### **3.1 Propuesta de Objetivos**

#### **3.1.1 Objetivo General**

- Obtener un modelo Espacio-Temporal que establezca una relación entre la enfermedad y las Precipitaciones Pluviales en el departamento de Loreto

#### **3.1.2 Objetivos Específicos**

- Determinar el nivel de Riesgo Relativo mensual por provincia.
- Estimar la incidencia de la enfermedad mensual para las provincias
- Estimar la importancia de las lluvias en cada mes
- Estimar el efecto mensual de las provincias circundantes en cada provincia.

## **3.2 Sistema de Hipótesis**

### **3.2.1 Hipótesis General**

- El riesgo relativo de adquirir el virus, causante de la malaria falciparum o vivax, depende linealmente de los niveles de intensidad de lluvias (mm/h) y el efecto de las provincias circundantes, los cuales constituyen información a priori en su proceso de estimación

### **3.2.2 Hipótesis Específicas**

- El riesgo relativo en una provincia específica depende de la intensidad de lluvia y sus áreas circundantes respectivas
- La incidencia de la enfermedad depende de los niveles de intensidad de lluvia de la zona y del efecto de las áreas circundantes.
- Existen diferencias entre los niveles de importancia de las lluvias para cada mes.
- Los efectos mensuales de las provincias circundantes son diferentes y tienen una estructura definida

## **3.3 Operacionalización de Variables**

Nuestras variables a analizar son las siguientes:

$Y_{it}$ : Número de casos reportados de Incidencias de malaria en la región  $i$  durante el mes  $t$

$X_{it}$ : Intensidad de lluvias en mm en la región  $i$  durante el mes  $t$

$e_{it}$ : valor estimado del número de casos reportados en el área  $i$  en el mes  $t$

$R_{it}$ : Riesgo Relativo en el área  $i$  en el mes  $t$

### **3.4 Tipo y métodos de investigación**

#### **3.4.1 Tipo de Investigación**

El tipo de investigación es experimental para poder determinar la naturaleza de algunas variaciones de efectos aleatorios que van evolucionando con el tiempo mediante un modelo Dinámico.

#### **3.4.2 Metodología de la Investigación**

Este tipo de investigación es aplicada dado que está interesada en la aplicación de los conocimientos en modelación dinámica bayesiana para determinar las precipitaciones en la zona de la selva amazónica de Loreto lo cual es la solución de un problema práctico importante.

El aporte metodológico de la presente investigación es el permite que algunas de las variaciones de efectos aleatorios vayan evolucionando con el tiempo de acuerdo a un modelo dinámico.

Así mismo, el cambio enfoque del problema causado por la combinación de conteos de malaria (por distrito) con la precipitación (por estación) está parcialmente resuelto por la interpolación de toda el área a través de un proceso de Gauss.

Se realizara el modelamiento con diferentes modelos para investigar cual es aquel que mejor pueda pronosticas la precipitaciones relacionadas con la incidencia de la enfermedad de la Malaria en la selva amazónica del departamento de Loreto.

### **3.5 Diseño de Investigación**

#### **3.5.1 Modelo Espacio Temporal**

Proponemos una generalización espacio-temporal de las prioris en el modelo (2) tal que, entre otras cosas, permite que  $\beta_t$  y  $b_{it}$  evolucione suavemente en el tiempo:

$$\log(r_{it}) = \beta_t x_{it} + b_{it} \quad (3.1)$$

$$\beta_t = \beta_{t-1} + w_t, \quad w_t \sim N(0, \tau_\beta^2) \quad (3.2)$$

para  $i = 1, \dots, n$  y  $t = 1, \dots, T$ . Como puede ser fácilmente reconocida, esta estructura se parece a un modelo lineal generalizado dinámico (West et al 1985).

Modelamos los términos de efectos aleatorios  $b_{it}$ , ya sea como independientes e idénticamente distribuidos normalmente con media cero y varianza  $\sigma_t^2$ ,  $b_{it} \sim N(0; \sigma_t^2)$  o como un CAR ( $\sigma_t^2$ ).

Además los  $\sigma_t^2$  pueden ser también independientes e idénticamente distribuidos como una gamma inversa con parámetros (fijos)  $a$  y  $c$ ,  $\sigma_t^2 \sim IG(a, c)$  para  $t = 1, \dots, T$ , o  $\log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \varepsilon_t$  con  $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$ .

Por lo tanto, tenemos las cuatro distribuciones prioris siguientes,

$$M1: b_{it} \sim N(0, \sigma_t^2) \text{ y } \sigma_t^2 \sim IG(a, c) \quad (3.4)$$

$$M2: b_{it} \sim CAR(\sigma_t^2) \text{ y } \sigma_t^2 \sim IG(a, c) \quad (3.5)$$

$$M3: b_{it} \sim N(0, \sigma_t^2) \text{ y } \log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \varepsilon_t \quad (3.6)$$

$$M4: b_{it} \sim CAR(\sigma_t^2) \text{ y } \log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \varepsilon_t \quad (3.7)$$

Está claro que los modelos M1 y M3 son muy similares. Se permite que los efectos aleatorios varíen independientemente a priori con varianzas que también varían independientemente en M1 pero evoluciona suavemente en el tiempo en M3. Igualmente, se permite que modelos M2 y M4 sean también similares, pero ahora, se asume que los efectos aleatorios siguen una priori espacialmente estructurada y la variabilidad de estos efectos permite que cambien con el tiempo.

A diferencia de los trabajos anteriores, nuestra estructura de modelado permite que los datos hablen, siempre que cualquier estructura en la evolución de la variabilidad de los efectos sea aparente.

### **3.5.2 Inferencia Posterior**

Es obvio, en nuestro contexto, que  $y_{it}$  y  $x_{it}$  son mal alineados, ya que  $x_{it}$  representaría la precipitación en el distrito  $i$ , mientras la única información disponible es la precipitación para varios puntos fijos (estaciones) en la región en estudio (el departamento de Loreto). Sin embargo, asumimos, en esta sección, que la covariable  $x_{it}$  es conocida para toda el área  $i$  y en el tiempo  $t$ .

Digamos que sea  $\theta$  el vector de parámetro que contiene todas las cantidades desconocidas del interés. Más exactamente:

$$\theta = (\beta, b, \sigma_t^2, \tau_\beta^2, \sigma_\varepsilon^2) \quad (3.8)$$

donde

$$\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_T), \quad (3.9)$$

$$b = (b_{(1)}, b_{(2)}, b_{(3)}, \dots, b_{(T)}), \quad (3.10)$$

$$b_t = (b_{1t}, b_{2t}, b_{3t}, \dots, b_{nt}) \quad \text{para } t = 1, \dots, T \text{ y} \quad (3.11)$$

$$\sigma^2 = (\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2, \dots, \sigma_T^2) \quad (3.12)$$

Para la ilustración los objetivos consideran M4 modelo como definido anteriormente. Por el teorema de las Bayes, la distribución posterior de  $\theta$  es proporcional a la priori por la función de probabilidad, que es:

$$\begin{aligned} p(\theta|y, x) \propto & \prod_{t=1}^T \prod_{i=1}^n \exp(\beta_t x_{it} + b_{it})^{y_{it}} \exp\{-e_{it} \exp(\beta_t x_{it} + b_{it})\} \\ & \times \prod_{t=1}^T [p(\beta_t | \beta_{t-1}, \tau_\beta^2) p(b_t | \sigma_t^2) p(\sigma_t^2 | \sigma_{t-1}^2, \sigma_\varepsilon^2)] \\ & \times p(\beta_0) p(\tau_\beta^2) p(\sigma_\varepsilon^2). \end{aligned} \quad (3.13)$$

donde  $Y = \{Y_{it}: i = 1, \dots, n; t = 1, \dots, T\}$  y  $x = \{X_{it}: i = 1, \dots, n; t = 1, \dots, T\}$ . Ahora estamos listos para adjudicar distribuciones prioris a los parámetros del modelo en el tiempo  $t=0$ .

Para adjudicamos una distribución priori normal con media cero y con el varianza grande (fija). Para los parámetros de varianza usamos distribuciones aprioris gamma inversas con parámetros fijados en algunos valores convenientes.

Claramente, la distribución posteriori es analíticamente sin solución. Sin embargo, varios algoritmos de Monte Carlo con Cadenas de Markov ha sido propuestos ya para simular  $P(\theta/y, x)$ . Más expresamente, dirigimos una simulación de Gibbs (Gelfand y Smith, 1990) un algoritmo basado en las distribuciones posteriores llenas de condicionales para cada uno de los parámetros.

Excepto los parámetros de escala:  $\tau_{\alpha}^2, \tau_{\beta}^2, \sigma^2$  y  $\sigma_{\epsilon}^2$  cuyas distribuciones sobrecondicionadas son distribuciones gamma inversas, todo que las demás tienen distribuciones sobrecondicionadas posteriores desconocidas. Sin embargo, ellas son logarítmicamente cóncavas y uno podría usar, por ejemplo, el método de rechazo adaptado (Gilks y Wild, 1992) para obtener muestras de ellos. Usamos el software la versión 1.42 de WinBugs (Spiegelhalter y en., 2002) para realizar toda la inferencia posterior para modelos M1, M2, M3 y M4.

### **3.5.3 La importancia de las precipitaciones pluviales y Efecto de las Áreas circundantes**

En la Ecuación 3.1 del presente Capítulo se definen los parámetros  $\beta_t$  y  $b_{it}$  como la importancia de las precipitaciones pluviales y el efecto de las áreas circundantes respectivamente, los cuales luego de ser estimados se utilizan para describir el comportamiento así como la interacción entre las precipitaciones y la estructura de correlación con las provincias circundantes.

El parámetro  $\beta_t$  que se define como la importancia de las precipitaciones en el mes  $t$  nos permitirá observar la evolución de la influencia de las lluvias sobre los riesgos relativos y por ende la incidencia de la Enfermedad.

El parámetro  $b_{it}$  que se define como el efecto de las áreas circundantes en el mes  $t$  para el área  $i$  nos permitirá observar la evolución de la influencia de las áreas sobre los riesgos relativos.

El modelado permite describir con detalle las relaciones importantes en la estructura del diseño, el cual se basa en la información de investigaciones sobre el comportamiento de la enfermedad, establecida por Finlay y Ross.

### **3.6 Población y Muestra**

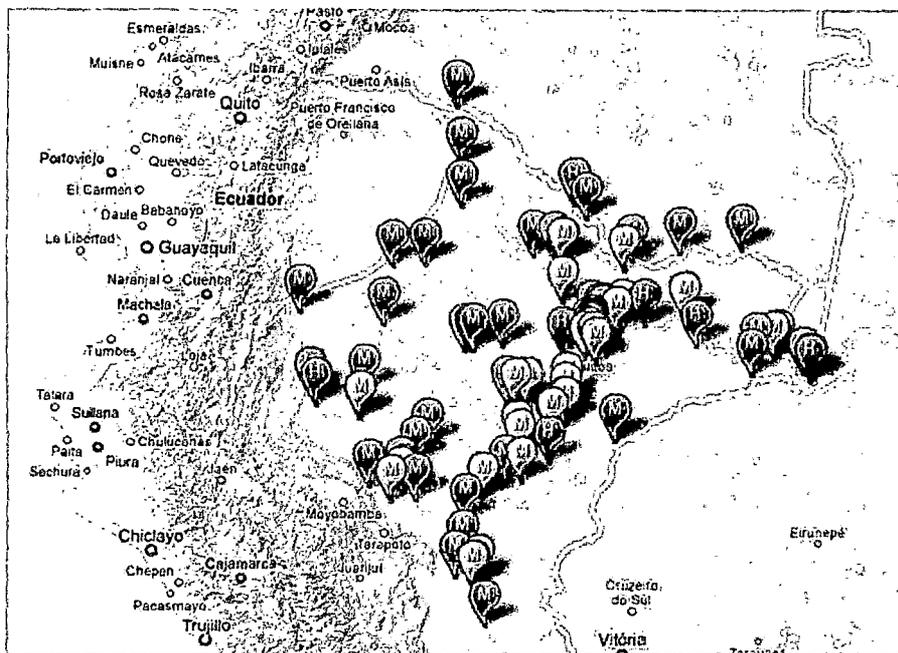
En el presente estudio nuestra unidad poblacional será la región o provincia  $i$  de una población de provincias que conforma el departamento de Loreto. De esta población de regiones gracias a la información otorgada por el departamento de Epidemiología del Ministerio de Salud así como del banco de datos de monitoreo de lluvias que posee el Senahmi hemos de tomar una muestra de datos mensual de la incidencia de malaria así como el volumen de caídas de lluvia en diversos puntos de cada provincia de Loreto durante el Periodo Enero-Diciembre 2009.

Esta muestra nos servirá para obtener la función de verosimilitud de nuestro modelo para que conjuntamente con la probabilidad priori nos permitan obtener predicciones de la enfermedad en algún tiempo determinado en la región i

**Figura N°3.6**

**Puntos de Monitoreo de las Precipitaciones Pluviales**

**Por el Senahmi**



## **CAPITULO IV**

### **RESULTADOS Y ANÁLISIS**

En este capítulo se presentarán los resultados obtenidos al ajustar mediante el enfoque bayesiano los modelos considerando los criterios DIC y D; así como también la selección del mejor modelo y la estructura desarrollada con el paquete estadístico de software libre OpenBugs del modelo seleccionado. Se presentarán las tablas de estudio por Provincia del comportamiento de los Riesgos Relativos, así como los Intervalos de Confiabilidad para las estimaciones de la incidencia de Malaria y de los parámetros del modelo.

#### **4.1 Análisis de Datos**

A partir de la información de Incidencias de la enfermedad, la caída de lluvias para cada área en cada mes hemos modelado mediante cuatro estructuras diferentes en cuanto al comportamiento del efecto de las áreas circundantes.

Podemos observar un alto número de incidencias en las Provincia de Maynas y Alto Amazonas las cuales tienen una baja densidad poblacional a comparación de las provincias de Ucayali, Requena y Mariscal Ramón Castilla lo cual sugiere la inclusión en el modelo de información poblacional de cada provincia a través de la variable  $e_{it}$  como el número de casos esperado ponderado con la cantidades poblacionales actuales.

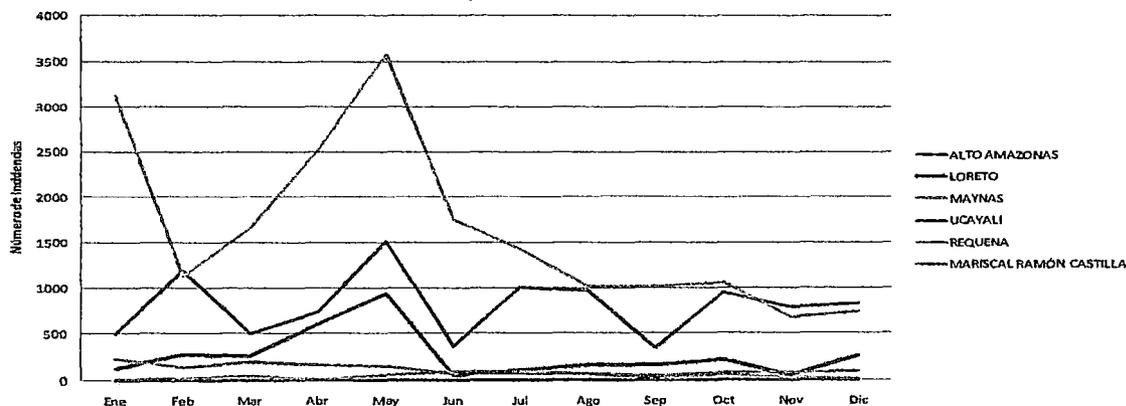
**Figura N°4.1**

**Puntos de Monitoreo de las Precipitaciones Pluviales Por el Senahmi**



**Gráfico N°4.1.1**

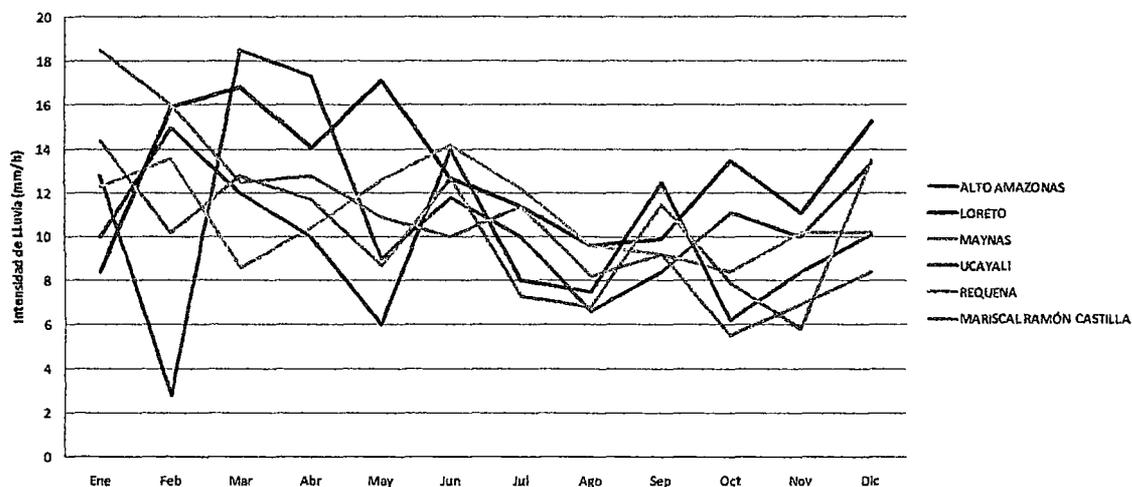
**Numero de Incidencias de Malaria por Provincia en el departamento de Loreto**



Fuente: Ministerio de Salud

**Gráfico N°4.1.2**

**Intensidad de Precipitaciones por Provincia  
en el departamento de Loreto  
(En milímetros por hora)**



Fuente: Datos Históricos Senahmi

En cuanto al nivel de lluvias para todas las provincias observamos un comportamiento correlativo especialmente en los meses de Junio, Julio, Agosto y Septiembre.

**4.2 Comparación de Modelos Ajustados**

En la sección 3.5.1 se especifico los cuatro modelos a tener en cuenta para el mapeo de la enfermedad:

$$M1: bit \sim N(0, \sigma_t^2) \text{ y } \sigma_t^2 \sim IG(a, c)$$

$$M2: bit \sim CAR(\sigma_t^2) \text{ y } \sigma_t^2 \sim IG(a, c)$$

$$M3: bit \sim N(0, \sigma_t^2) \text{ y } \log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \varepsilon_t$$

$$M4: bit \sim CAR(\sigma_t^2) \text{ y } \log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \varepsilon_t$$

Manteniendo todos estos una estructura similar. En el Cuadro N°1 se muestra la comparación de Ajuste de los cuatro modelos de los cuales se procede a elegir

Cuadro N°4.2

Comparación del Ajuste de los Modelos sencillos (sin CAR) y complejos (con CAR) con los criterios DIC y pD

Modelo	Dbar	Dhat	DIC	pD
Modelo 1	549.6	477.8	621.3	71.76
Modelo 2	1063	1046	1080	17.08
Modelo 3	547.7	475.2	620.1	72.43
Modelo 4	578.4	650.3	506.5	71.86

Elaboración: Propia

al modelo 4 como el modelo que más se ajusta a la información debido a que tiene el más bajo valor de DIC sin afectar el valor de pD (número aproximado de parámetros del modelo)

**4.3 Estimación de La Incidencia de Malaria**

Utilizamos el modelo 4 para estimar el número de incidencias de la enfermedad, para esto recurrimos al software WingBugs que nos permitirá simular los parámetros mediante el Método Montecarlo con Cadenas de Markov con un total de 100000 iteraciones podemos establecer los Intervalos de Credibilidad para el Número de Incidencias de la enfermedad.

En el cuadro B.1.1 del Anexo se observa que para el mes de Enero del año 2009 se espera que se produzcan 507 casos de Malaria en la Provincia de Alto Amazonas con una desviación de alrededor de 23 casos y podemos establecer un Intervalo de credibilidad del 95% con límite inferior y superior de 462 y 553 Infectados. Asimismo en la Provincia de Loreto se espera que se produzcan 134 casos de Malaria con una desviación de alrededor de 12 casos y podemos establecer un Intervalo de credibilidad del 95% con límite inferior y superior de 107 y 156 Infectados. En la provincia de Maynas, en donde se detectó la mayor cantidad de Infectados, se espera que se produzcan 3120 casos de Malaria con una desviación de alrededor de 59 casos y podemos establecer un Intervalo de credibilidad del 95% con límite inferior y superior de 3004 y 3237 Infectados, a su vez en la provincia de Ucayali en donde se detectó el menor número de Infectados, se estima que se produzcan 5 casos de Malaria y un intervalo de credibilidad al 95% con límite inferior y superior de 1 a 10 infectados.

En la provincia de Requena se espera que se produzcan 233 casos de Malaria con una desviación de alrededor de 20 casos y podemos establecer un Intervalo de credibilidad del 95% con límite inferior y superior de 194 y 276 Infectados.

En cuadros B.1.2 y B.1.3 del Anexo para los meses de Febrero y Abril observamos como nuestras estimaciones captan la caída en la Incidencia de Malaria en la provincia de Maynas en Febrero, así como la tendencia a incrementarse en el mes de Abril.

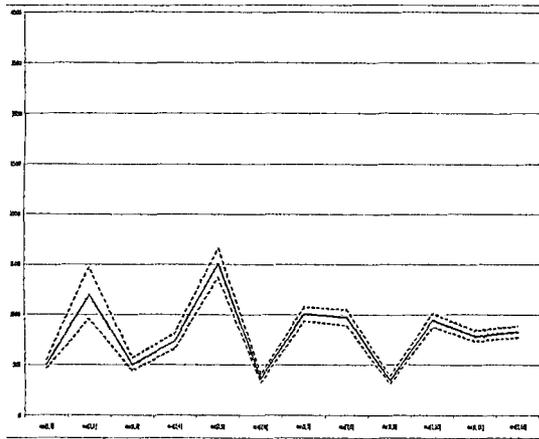
También podemos observar como para la provincia de Ucayali las estimaciones mantienen a un nivel de credibilidad del 95% casos de 1 a 10 infectados. Por otro lado en los meses de Abril y Mayo se observa que nuestras estimaciones captan nuevamente el incremento de Incidencias de la enfermedad en la provincia de Maynas, así como el incremento en la Provincia de Alto Amazonas.

Estos resultados mensuales nos permiten con ayuda del software WingBugs realizar mapeos de la Incidencia de la enfermedad para la estimación de la Tasa de Incidencia Mensual del Número de Incidencias en cada provincia del Departamento de Loreto y monitoreo como se muestran en los Gráficos B.2.1 B.2.2 y B.2.3.

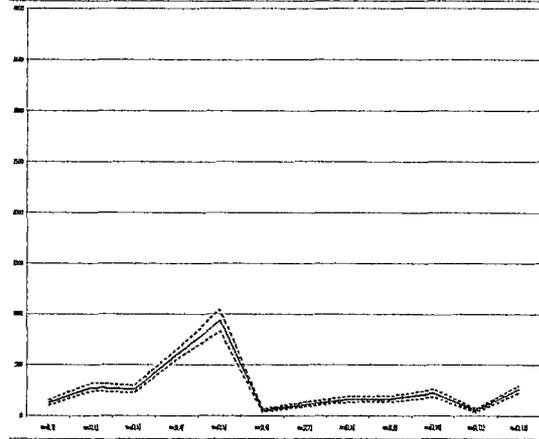
En el Grafico 4.3 que se muestra las series temporales del Numero de Incidencias por provincias y los respectivos intervalos de confiabilidad. Podemos observar en que hay mayores cantidades de Incidencias de Malaria en Zonas con baja densidad poblacional como son las provincias de Alto Amazonas, Loreto y Maynas como focos Geográficos de la Enfermedad en especial en la provincia de Maynas en donde las Incidencias llagan hasta los 3500 infectados mientras que en las demás provincias el número de incidencias oscilan entre los 0 y 300 Infectados Por lo que podemos decir que en las provincias de Alto Amazonas y Maynas se observan un gran número de Incidencias, en la provincia de Loreto y Requena se observan niveles relativamente moderados, mientras que en las provincias de Ucayali y Ramón Castilla las cantidades de Incidencias son Bajas.

**Gráfico N°4.3**

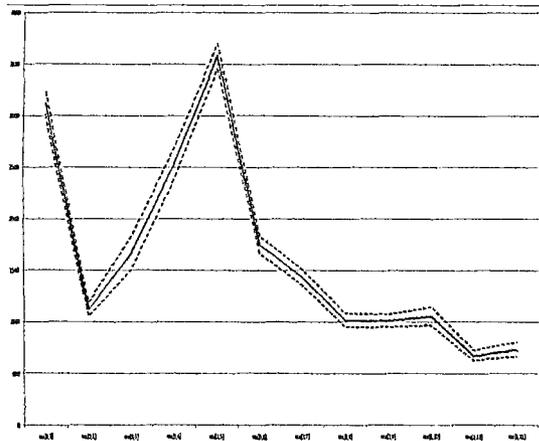
**Intervalos de Confiabilidad para el número de Incidencias de la Enfermedad para la provincia de Alto Amazonas (Enero-Diciembre 2009)**



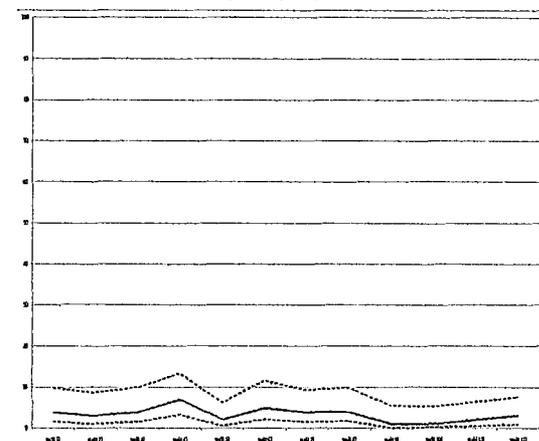
**ALTO AMAZONAS**



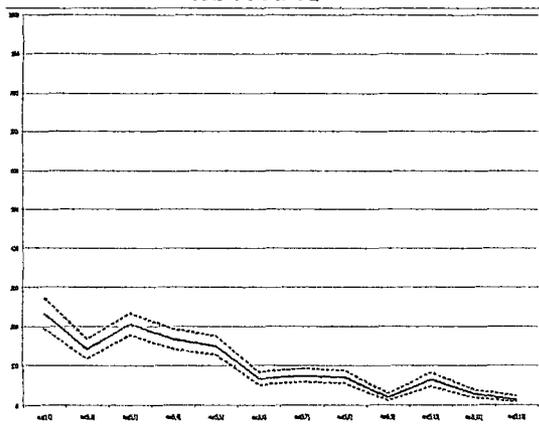
**LORETO**



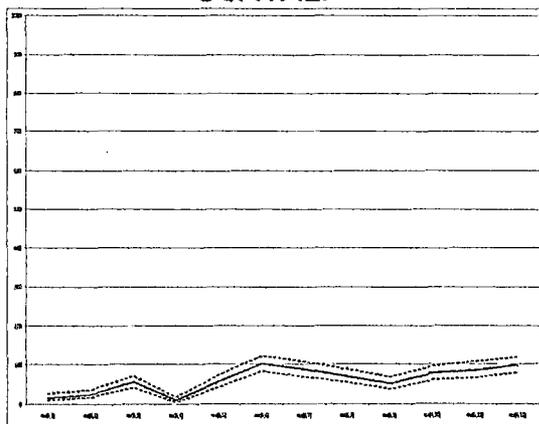
**MAYNAS**



**UCAYALI**



**REQUENA**



**RAMÓN CASTILLA**

#### **4.4 Estimación de Riesgos Relativos**

El riesgo Relativo es un Indicador que compara la tasa de Incidencia de la enfermedad con la tasa de incidencia esperada en la zona de estudio. En nuestra estructura modelar se relacionó esta tasa directamente con los efectos de las áreas circundantes así como también los efectos de las Precipitaciones pluviales.

Es así como mediante el software Wingbugs podemos establecer los niveles de Riesgo Relativo con un notable nivel de certeza.

Se observa en el cuadro B.3.1 del anexo que en el mes de Enero del 2009 en las provincias de Alto Amazonas, Loreto, Maynas y Requena los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en las provincias de Ucayali y Ramón Castilla estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifican como Zonas de Bajo Riesgo de Adquirir la Enfermedad.

Se observa en el cuadro B.3.2 del anexo que en el mes de Febrero del 2009 en las provincias de Alto Amazonas, Loreto, Maynas y Requena los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad. En cuanto a la provincia de Ramón Castilla debido a que el Riesgo oscila alrededor de la unidad se puede decir que es una zona de Riesgo Moderado.

Se observa en el cuadro B.3.3 del anexo que en el mes de Marzo del 2009 en las provincias de Alto Amazonas, Loreto, Maynas, Requena y Mariscal Ramón Castilla los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad.

Se observa en el cuadro B.3.4 del anexo que en el mes de Abril del 2009 en las provincias de Alto Amazonas, Loreto, Maynas y Requena los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad. En cuanto a la provincia de Ramón Castilla debido a que el Riesgo oscila alrededor de la unidad se puede decir que es una zona de Riesgo Moderado.

Se observa en el cuadro B.3.5 del anexo que en el mes de Mayo del 2009 en las provincias de Alto Amazonas, Loreto, Maynas, Requena y Mariscal Ramón Castilla los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad.

Se observa en el cuadro B.3.6 del anexo que en el mes de Junio del 2009 en las provincias de Alto Amazonas, Loreto, Maynas, Requena y Mariscal Ramón Castilla los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad.

Se observa en el cuadro B.3.7 del anexo que en el mes de Julio del 2009 en las provincias de Alto Amazonas, Loreto, Maynas, Requena y Mariscal Ramón Castilla los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad.

Se observa en el cuadro B.3.8 del anexo que en el mes de Agosto del 2009 en las provincias de Alto Amazonas, Loreto, Maynas, Requena y Mariscal Ramón Castilla los riesgos relativos obtenidos por simulación no son menores a 1 con un nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad.

Se observa en el cuadro B.3.9 del anexo que en el mes de Septiembre del 2009 en las provincias de Alto Amazonas, Loreto, Maynas, Requena y Mariscal Ramón Castilla los riesgos relativos obtenidos por simulación no son menores a 1 con un

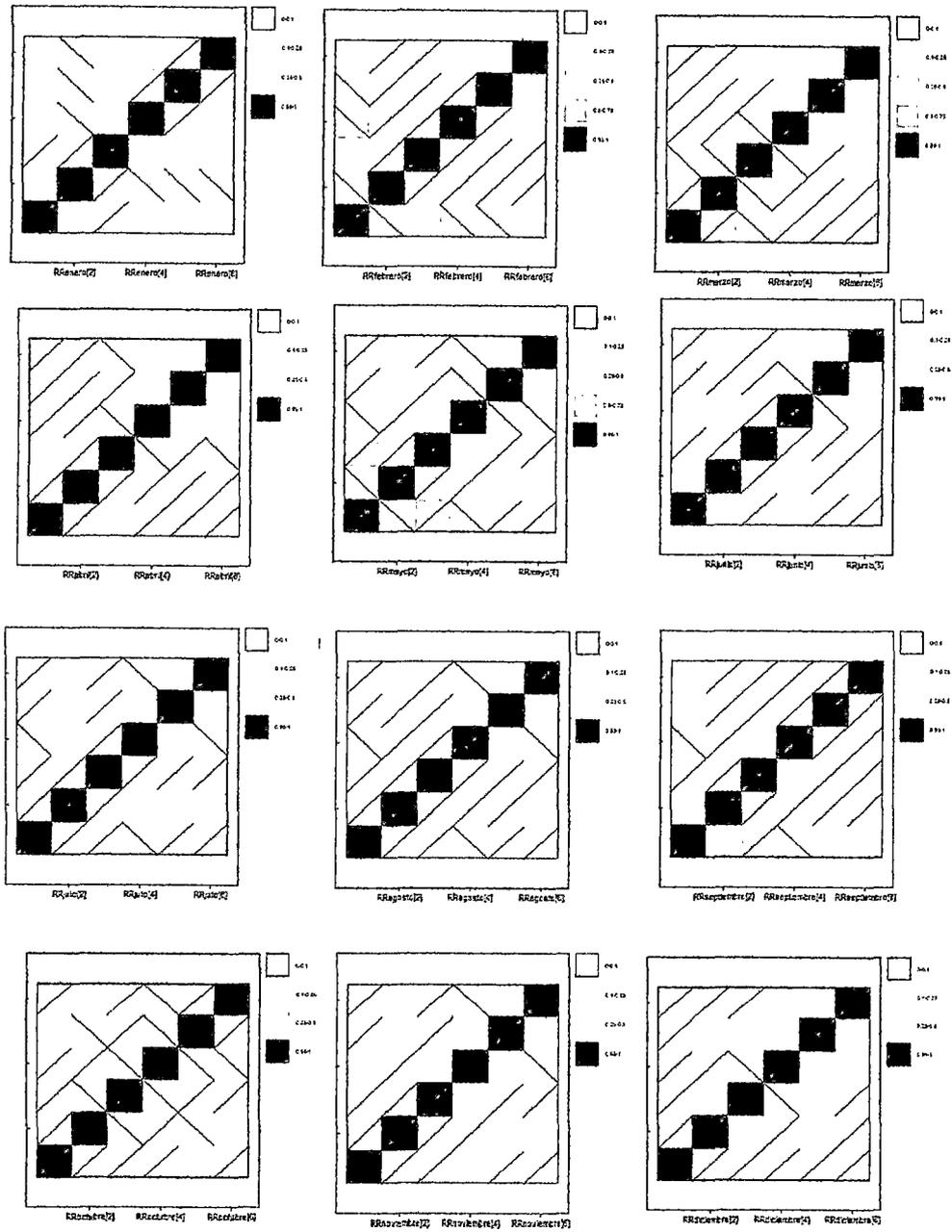
nivel de confiabilidad Mayor al 95% por lo que se clasificarían como zonas de Alto Riesgo de contraer la enfermedad, mientras que en la provincia de Ucayali estos riesgos Relativos no superan la unidad con un nivel de Confiabilidad del 95% por lo que se clasifica como Zona de Bajo Riesgo de Adquirir la Enfermedad.

Sin Embargo hay que tener en cuenta que los Riesgos Relativos para las provincias de Requena y Mariscal Ramón Castilla, oscilan en valores cercanos a la unidad por lo que habría que reconsiderarlas como zonas de Riesgo Moderado a Alto.

Se observa en los Gráficos B.4.1, B.4.2 y B.4.3 del anexo que los Riesgos Relativos para las provincias de Alto Amazonas, Loreto y Maynas son muy altos especialmente en los meses de Marzo, Abril, Mayo, Junio y Julio. Por otro lado observamos que la provincia de Ucayali no se ve afectada por estos Riesgos por lo que podemos decir con seguridad que es una zona segura teniendo en cuenta además el bajo número de Infectados que no supera los 20.

### Gráfico N°4.4

Matriz de Correlaciones de los Riesgos Relativos  
(Enero-Diciembre 2009)



#### 4.5 Estimación de coeficientes de la Intensidad de Precipitaciones

En el modelo se incluyen las magnitudes de Intensidad de Lluvias (en mm/h) en cada Provincia del departamento de Loreto y se procedió a estimar los coeficientes mensuales que con las magnitudes realizan una contribución al modelo. En el cuadro 25 se muestran los coeficientes estimados por el modelo así como sus respectivos intervalos de Confiabilidad al 95% procedentes de la simulación.

##### Cuadro N°4.5

Estimación de coeficientes de las Precipitaciones Pluviales por mes  
(Enero-Diciembre 2009)

Coefficiente	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	Interacciones
beta[Enero]	0.04748	0.007119	2.68E-05	0.03248	0.04788	0.06034	1	100000
beta[Febrero]	0.09386	0.007835	2.87E-05	0.07717	0.09439	0.1076	1	100000
beta[Marzo]	0.1299	0.006476	2.27E-05	0.1161	0.1304	0.1413	1	100000
beta[Abril]	0.1844	0.006449	2.47E-05	0.1711	0.1847	0.1962	1	100000
beta[Mayo]	0.2752	0.008894	3.84E-05	0.256	0.276	0.2903	1	100000
beta[Junio]	0.1693	0.005851	2.32E-05	0.157	0.1696	0.1799	1	100000
beta[Julio]	0.2411	0.007697	3.05E-05	0.2247	0.2416	0.2549	1	100000
beta[Agosto]	0.175	0.0101	3.81E-05	0.1534	0.1756	0.1931	1	100000
beta[Septiembre]	0.1399	0.01595	5.86E-05	0.104	0.1416	0.1665	1	100000
beta[Octubre]	0.1046	0.01398	6.90E-05	0.07293	0.1061	0.1276	1	100000
beta[Noviembre]	0.1419	0.012	4.85E-05	0.1156	0.1428	0.1626	1	100000
beta[Diciembre]	0.07833	0.008569	3.17E-05	0.05993	0.07886	0.09356	1	100000

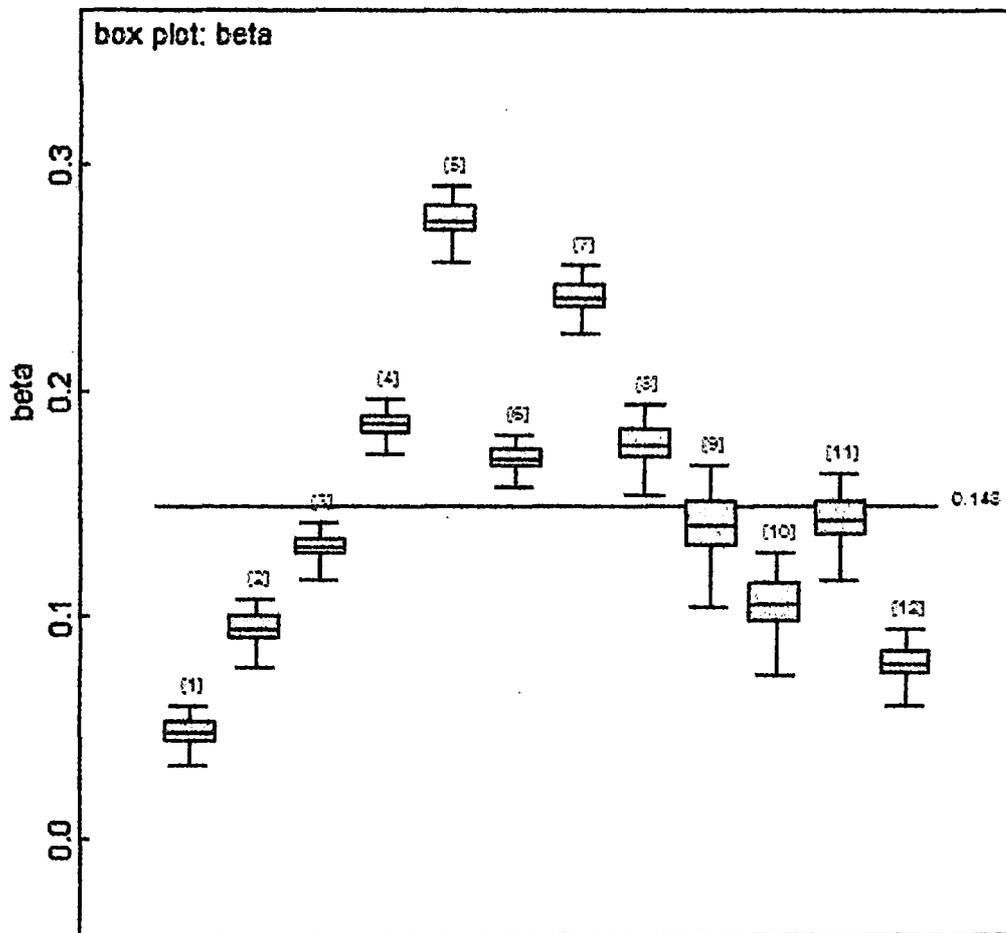
Elaboración: Propia

En el mes de Enero el coeficiente de la intensidad de lluvia registrada en dicho mes es de 0.04748 con un Intervalo de confiabilidad al 95% de límites Inferior y Superior de 0.03248 y 0.06034. Asimismo en el mes de Febrero el coeficiente de

la intensidad de lluvia registrada en dicho mes es de 0.09386 con un Intervalo de confiabilidad al 95% de límites Inferior y Superior de 0.07717 y 0.1076. Observamos en el Gráfico 12 que hay mayor importancia del efecto de las lluvias durante los meses de Abril, Mayo, Junio, Julio y Agosto

Gráfico N°4.5.1

Estimación del Efecto de las Precipitaciones Pluviales por Meses en el departamento de Loreto (Enero-Diciembre 2009)

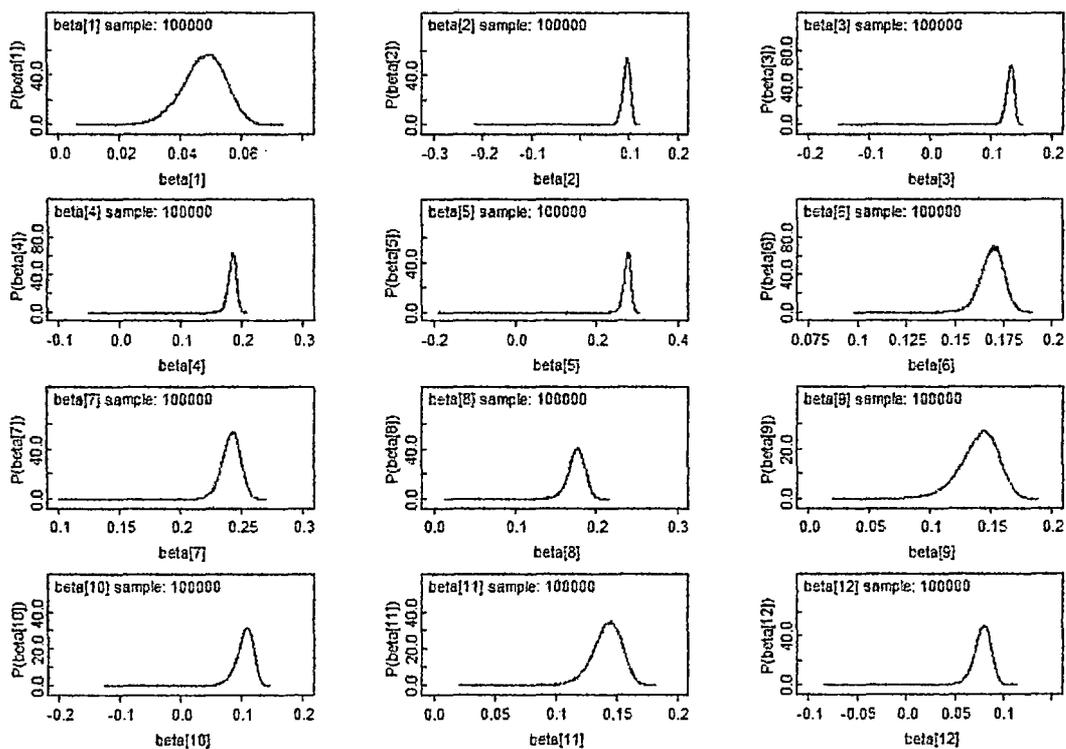


Podemos establecer una mayor importancia de las lluvias durante las épocas de Invierno, a pesar de que en la Región tiene un clima tropical y las intensidades de

lluvias son entre moderadas (mayor a 2 mm/h y menores a 15 mm/h) y fuertes (mayores a 15mm/h y menores a 30 mm/h) según Agencia Estatal de Meteorología de España.

**Gráfico N°4.5.2**

**Estimación de los coeficientes relativos a las Precipitaciones Pluviales por Meses en el departamento de Loreto (Enero-Diciembre 2009)**



Observamos como las distribuciones de los coeficientes relativos a las Precipitaciones pluviales presentan colas al lado izquierdo, es decir hay cierta probabilidad de que los coeficientes simulados caigan en puntos fuera de los intervalos de confiabilidad.

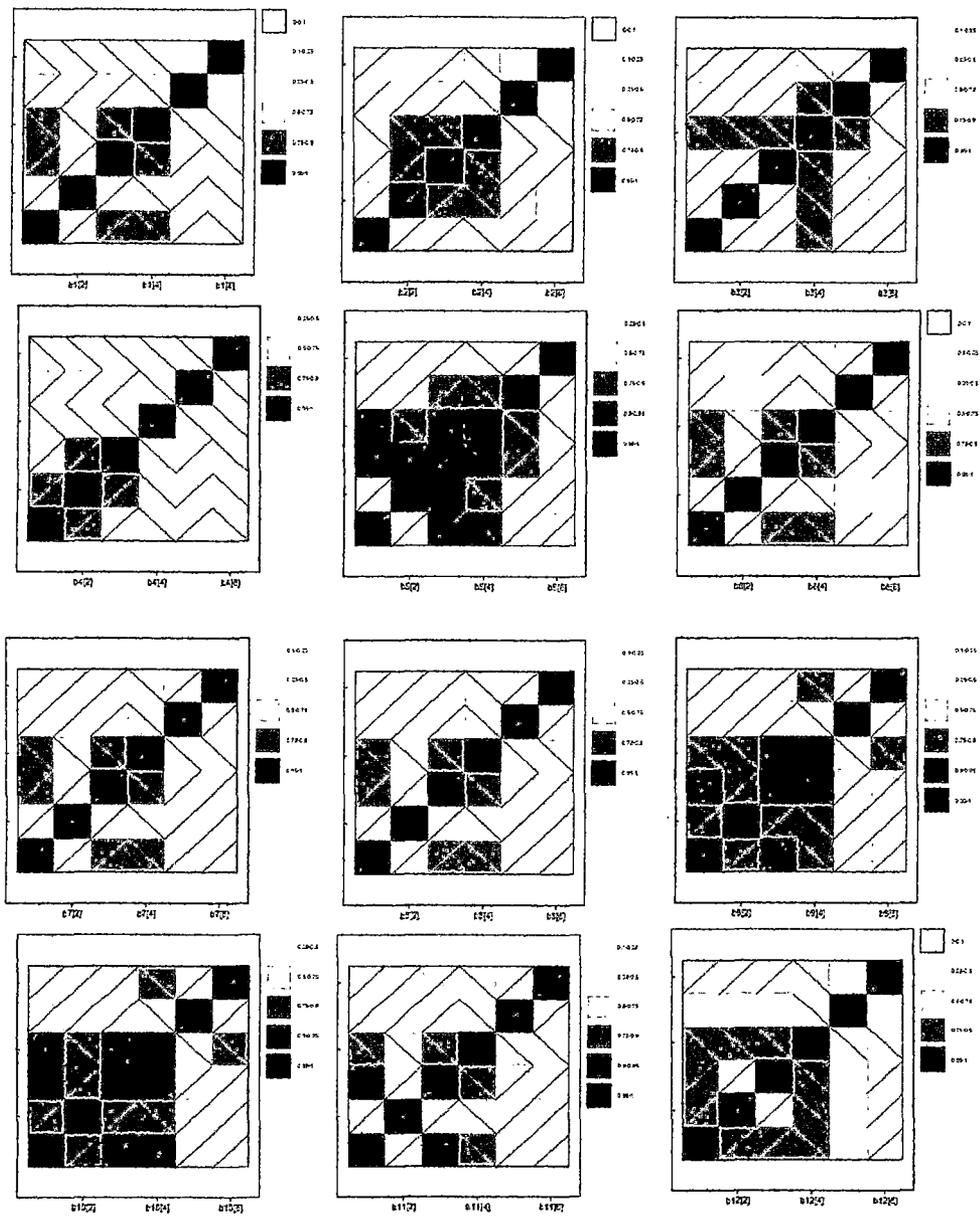
#### **4.6 Estimación de Efectos de las Áreas Circundantes**

En el modelo se incluyó el efecto de las áreas circundantes de cada área en cada mes de manera que siga un comportamiento correlativo a través de la distribución  $CAR(\sigma_b)$ . En el anexo B.5 se muestran los cuadros mensuales mediante los cuales se obtienen las estimaciones de los efectos así como los respectivos intervalos de Credibilidad.

Debido a que la suma de efectos de áreas circundantes debe ser

En el Gráfico 4.7 se observa como los niveles de correlación de todas las provincias se mantienen en niveles medianos y altos en varios meses a excepción de la provincia de Ramón Castilla ( $b_i(6)$ ) la cual presenta medianos a bajos niveles de correlación sin establecer una relación con ninguna de las otras provincias. El efecto de las áreas circundantes a provincia de Alto Amazonas mantiene una correlación positiva con el efecto para la provincia de Loreto, Maynas y Requena teniendo estas también correlaciones positivas entre si, por otro lado el efecto de las áreas circundantes de la provincia de Ucayali mantiene una correlación negativa con el efecto para las provincias de Alto Amazonas, Loreto, Maynas y Requena. Se establecen entonces estructuras de correlación para los efectos aleatorios determinadas lo cual

**Gráfico N° 4.6**  
**Matrices de Correlaciones del Efecto de las Áreas Circundantes**  
**por Meses en el departamento de Loreto**  
**(Enero-Diciembre 2009)**



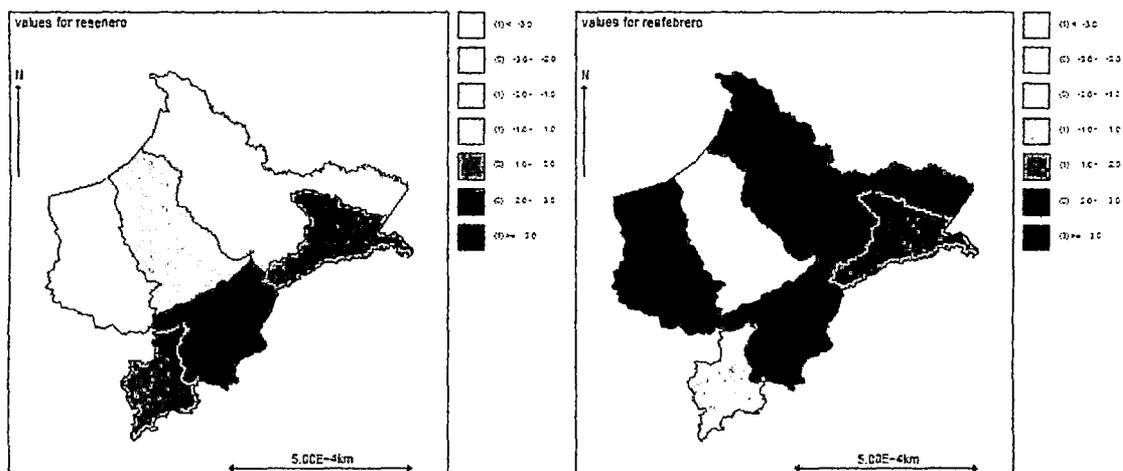
#### 4.7 Análisis de Residuos

A continuación se hace un análisis de los residuales con el fin de evaluar el ajuste del modelo. Los Residuales fueron calculados a partir de las estimaciones de Número de infectados menos los valores reales obtenidos en el Periodo. El Criterio para analizar los residuales es el de aleatoriedad.

Es así que podemos Observar en los Gráficos 4.7.1, 4.7.2, 4.7.3 y 4.7.4 que los residuales se ubican de manera aleatoria aparente en las provincias no existiendo dependencia temporal o espacial.

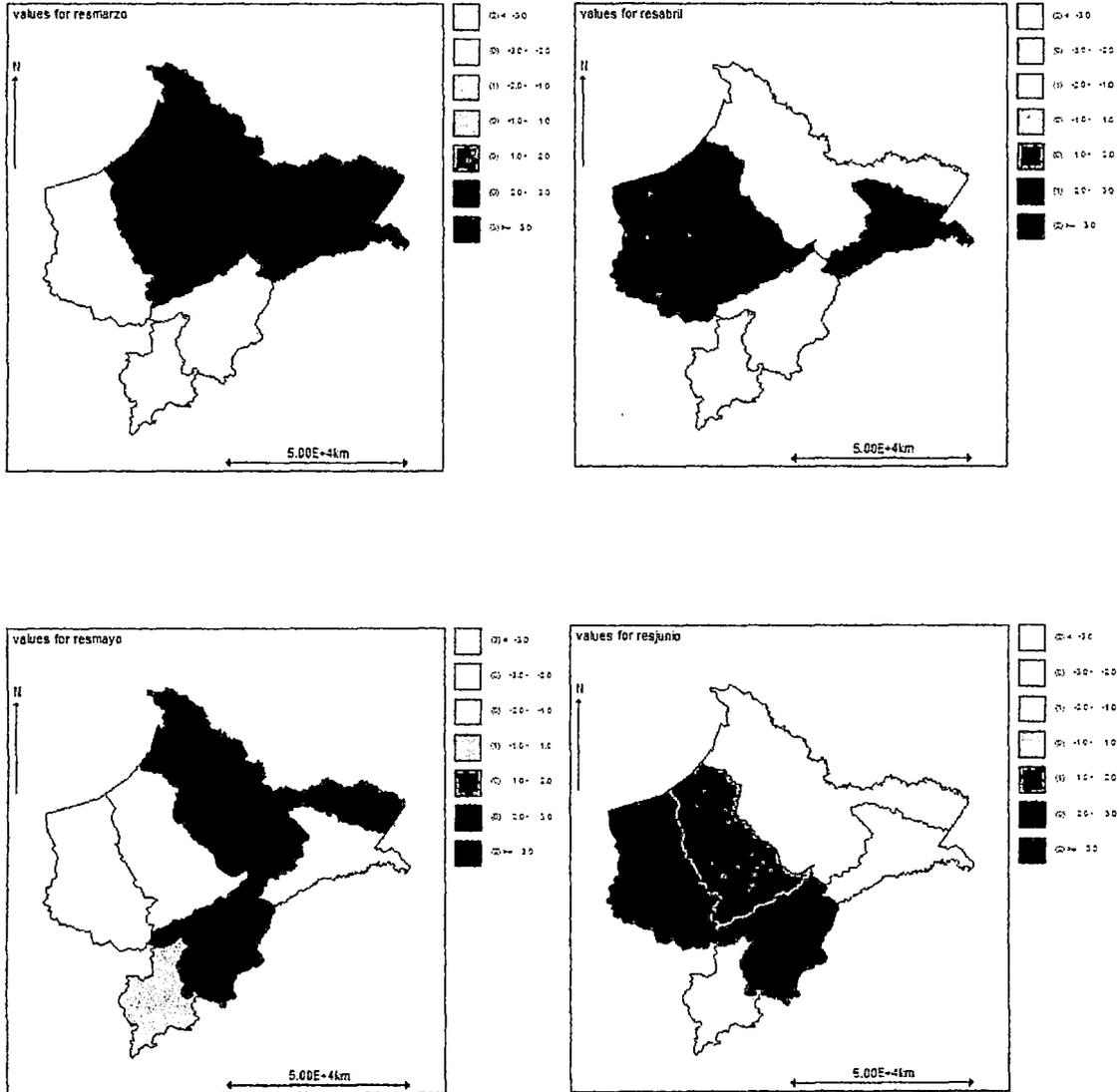
Gráfico N°4.7.1

Residuales por Meses en el departamento de Loreto  
(Enero-Febrero 2009)



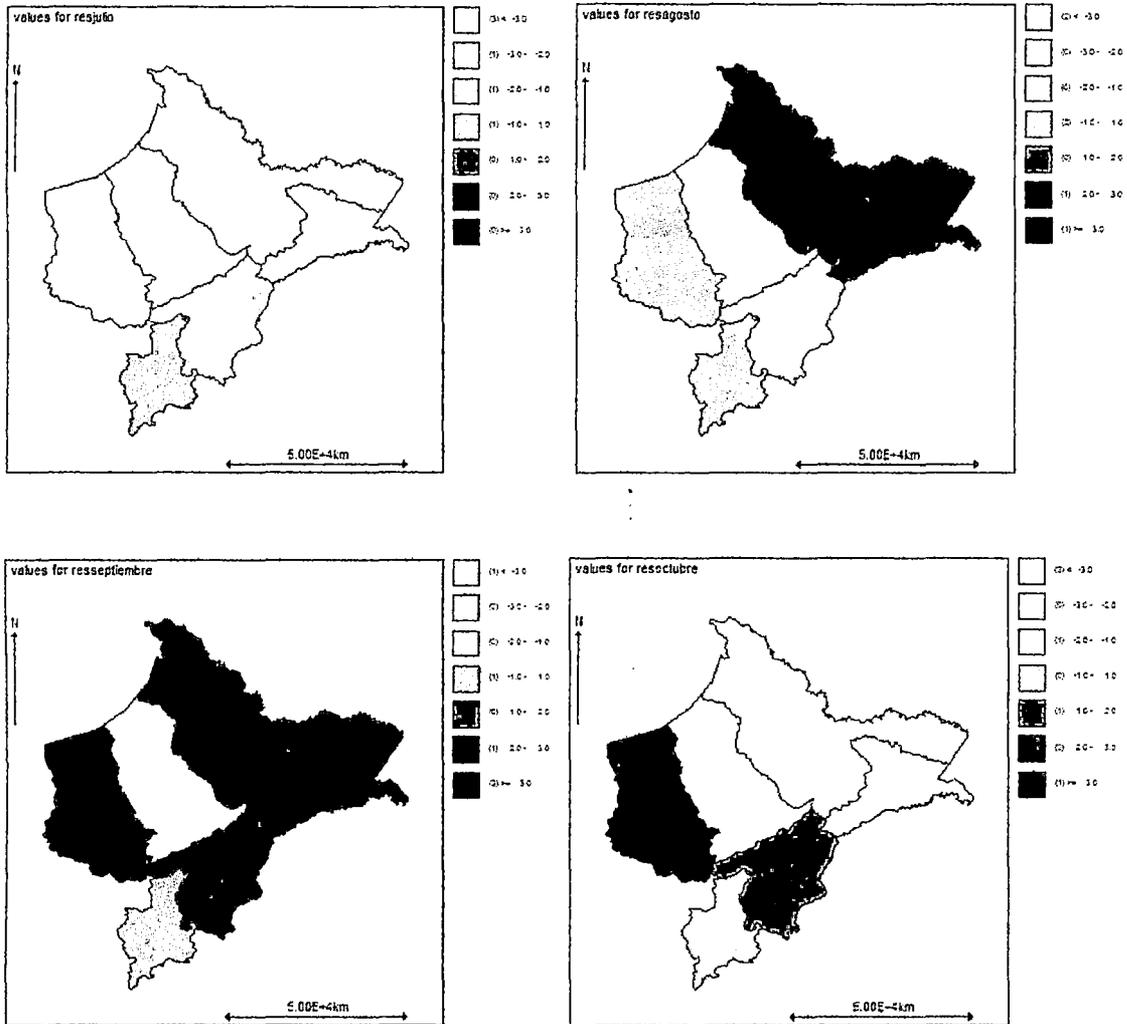
### Gráfico N°4.7.2

## Residuales por Meses en el departamento de Loreto (Marzo-Junio 2009)



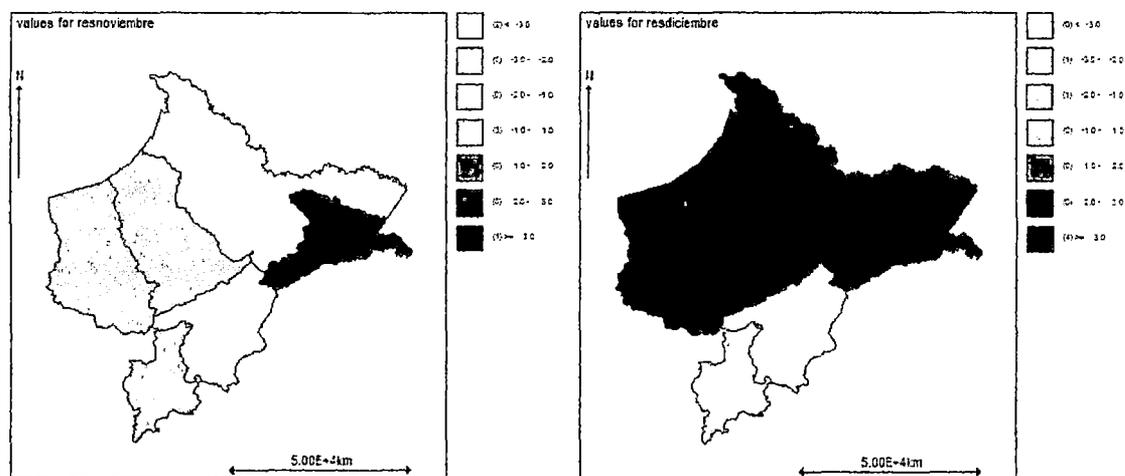
### Gráfico N°4.7.3

## Residuales por Meses en el departamento de Loreto (Julio-October 2009)



### Gráfico N°4.7.4

#### Residuales por Meses en el departamento de Loreto (Noviembre-Diciembre 2009)



La Distribución de los Residuales muestra que considerando los efectos de las áreas circundantes y efectos de las precipitaciones fluviales, modelando los riesgos relativos con valores esperados de incidencias se obtiene un modelo que se ajusta adecuadamente a la información introducida vía verosimilitud; por lo tanto el modelo es eficiente.

## CONCLUSIONES

- 1.- Se ha obtenido un modelo Espacio-Temporal que establece una relación entre la enfermedad y las precipitaciones pluviales en el departamento de Loreto, el cual fue escogido mediante el estudio del Criterio de Información de Devianza. El Modelo describe claramente las estructuras de correlación entre las áreas circundantes, detectando además en que temporadas hay mayor influencia de las lluvias sobre las Incidencias de Malaria, logra estimar adecuadamente los Riesgos Relativos de contagio de cada provincia y estima adecuadamente el Número total de incidencias de malaria por cada provincia, en síntesis un modelo completo.
- 2.- Se establecieron los niveles de Riesgos Relativos para cada zona en cada mes, estableciendo que la provincia de Ucayali es ajena a los efectos espaciales de las áreas adyacentes mostrando Riesgos Relativos menores a la unidad. Las provincias de Alto Amazonas, Loreto, Maynas y Requena presentan muy altos Riesgos Relativos especialmente durante los meses de Marzo, Abril, Mayo, Junio, Julio, Agosto y Septiembre en los que reportan Riesgos mayores a 7.5 mientras que la provincia de Mariscal Ramón Castilla solo los meses de Mayo, Junio, Julio y Noviembre presenta riesgos relativos mayores a 7.5, Es decir, todas las provincias a excepción de Ucayali presenta riesgos relativos de altos a muy altos lo cual evidencia el comportamiento endémico y epidémico de la enfermedad.
- 3.-Se logró establecer Intervalos de Confiabilidad para la Incidencia de la Enfermedad en cada área y en cada mes verificando que las estimaciones del

modelo se ajustan a la información de Incidencias en el departamento de Loreto. En las provincias de Alto Amazonas y Maynas se observan un gran número de Incidencias, en la provincia de Loreto y Requena se observan niveles relativamente moderados, mientras que en las provincias de Ucayali y Ramón Castilla las cantidades de Incidencias son Bajas.

4.- Se establecieron intervalos de Credibilidad para los efectos de las áreas circundantes logrando visualizar relaciones directas e inversas importantes de correlación entre las provincias a excepción de la provincia de Mariscal Ramón Castilla, el efecto de las áreas circundantes a provincia de Alto Amazonas mantiene una correlación positiva con el efecto para la provincia de Loreto, Maynas y Requena teniendo estas también correlaciones positivas entre sí. Por otro lado el efecto de las áreas circundantes de la provincia de Ucayali mantiene una correlación negativa con el efecto para las provincias de Alto Amazonas, Loreto, Maynas y Requena. Es decir, hay correlaciones espaciales importantes y determinadas.

5.-Finalmente, se realizó un análisis de los Residuales a partir de las estimaciones de la incidencia de la enfermedad con el criterio de Aleatoriedad encontrando que la distribución de los residuales presenta una aleatoriedad espacial y temporal. Además estos muestran que considerando los efectos de las áreas circundantes y efectos de las precipitaciones fluviales, modelando los riesgos relativos con valores esperados de incidencias, se obtiene un modelo que se ajusta adecuadamente a la información introducida vía verosimilitud; por lo tanto el modelo es eficiente.

## RECOMENDACIONES

A partir del análisis realizado del comportamiento de los diversos parámetros del modelo podemos decir que se pueden mejorar las estimaciones del modelo sin considerar la provincia de Ucayali dado que no establece relaciones espaciales con las demás provincias en nuestro análisis. Sin embargo, el modelo es adecuado dado que se estudia la problemática del departamento de Loreto.

Para un mejor análisis se recomendaría un diseño a nivel de distritos para así poder utilizar de manera más eficaz la intensidad de precipitaciones.

## BIBLIOGRAFÍA

Andrew B. Lawson (2008) Bayesian Disease Mapping Hierarchical Modeling in Spatial Epidemiology. 2, 19-27

H. S. Migon and Gamerman (1999) An Integrated Approach Statistical Inference 4, 99-103

Mike West and Jeff Harrison (1997) Bayesian Forecasting and Dynamic Models 2, 32-51

Assunção, R. M., Reis, I. A. and Oliveira, C. D. L. (2001) Diffusion and prediction of Leishmaniasis in a large metropolitan area in Brazil with a Bayesian space-time model. Statistics in Medicine, 20, 2319–2335.

Bernardinelli, L., Clayton, D. and Montomoli, C. (1995) Bayesian estimates of disease maps: How important are priors? Statistics in Medicine, 14, 2411–2431.

Besag, J., York, J. and Mollié, A. (1991) Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics, 43, 1–59.

- Best, N. G., Arnold, R. A., Thomas, A., Waller, L. A. and Conlon, E. M. (1999) Bayesian models for spatially correlated disease and exposure data. In Bayesian Statistics 6 (eds J. M. Bernardo, J. O. Dawid A. P. Berger and A. F. M. Smith), pp. 131–156. Oxford University Press.
- Cressie, N. A. C. (1993) Statistical for Spatial Data. New York: Wiley.
- Días, R. C. (2003) Uma contribuição ao estudo da malária no estado de Roraima e sua associação com a precipitação pluviométrica no período de 1985 a 1996. PhD thesis. Escola Nacional de Saúde Pública, Brasil (En Português).
- Gelfand, A.E. and Smith, A.F.M. (1990) Sampling-Based Approaches to Calculating Marginal Densities. Journal of the American Statistical Association, 85, 398–409.
- Gelfand, A.E., Zhu, L. and Carlin, B.P. (2001) On the change of support problem for spatio-temporal data. Biostatistics, 2, 31–45.
- Gilks, W.R. and Wild, P. (1992) Adaptive Rejection Sampling for Gibbs Sampling. Applied Statistics, 41, no. 2, 337–48.
- Kelsal, J. and Wakefield, J. (2001) Modeling spatial variation in disease risk: A geostatistical approach. Journal of the American Statistical Association, 97, 692–701.

Knorr-Held, L. and Besag, J. (1998) Modelling risk from a disease in time and space. *Statistics in Medicine*, 17, 2045–2060.

Knorr-Held, L. and Richardson, S. (2003) A hierarchical model for space-time surveillance data on meningococcal disease incidence. *Applied Statistics*, 52, 169–183.

Mollié, A. (1996) Bayesian mapping of disease. In *Markov Chain Monte Carlo in Practice* (eds W. R. Gilks, S. Richardson and D. J. Spiegelhalter), pp. 359–379. Chapman & Hall.

Nobre, A. A. (2003) A relação entre a malária e a chuva no estado do Pará: Uma análise espaco-temporal. Master's dissertation. Universidade Federal do Rio de Janeiro, Brasil (En Português).

O'Hagan, A. (1994) *Kendall's Advanced Theory of Statistics, Volume 2B, Bayesian Inference*. John Wiley & Sons, Inc.

Rodriguez-Iturbe, I. and Mejia, J. (1974) On the transformation of point rainfall to areal rainfall. *Water Resour. Research*, 10, 729–735.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*, 64, 583–639.

Spiegelhalter, D. J., Thomas, A. and Best, N. G. (2002) WinBUGS version 1.4 user manual. Research Report. Cambridge: Medical Research Council Biostatistics.

Stidd, C. K. (1973) Estimating the precipitation climate. *Water Resour. Research*, 9, 1235–1241.

Waller, L. A., Carlin, B. P. and Xia, A. E. H. Gelfand (1997) Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical Association*, 92, 607–617.

West, M. and Harrison, J. (1997) Bayesian Forecasting and Dynamic Models. New York: Springer - Second Edition.

West, M., Harrison, J. and Migon, H. (1985) Dynamic generalized linear models and Bayesian forecasting (with discussion). *Journal of the American Statistical Association*, 80, 73–97.

Zhu, L., Carlin, B.P. and Gelfand, A.E. (2003) Hierarchical regression with misaligned spatial data: relating ambient ozone and pediatric asthma ER visits in Atlanta. *Environmetrics*, 14, 537–557.

Upton, G.J. y B. Fingleton, Spatial data analysis by example, volume 1: Point pattern and quantitative data, (Toronto: Wiley), 1985.

Chen, M.-H., Shao, Q.-M. e Ibrahim, J. G. (2000): Monte Carlo methods in Bayesian Computation. Springer series in statistics. Springer-Verlag. New York.

Gilks, W. R., Richardson, S. y Spiegelhalter, D. J. (1996): Markov chain Monte Carlo in practice. Chapman and Hall. London.

Robert, C. P. y Casella, G. (1999): Monte Carlo statistical methods. Springer series in statistics. Springer-Verlag. New York.

Chen, M.-H., Spiegelhalter, D., Thomas, A. y Best, N. (1996-2000): "The BUGS Project". <http://www.mrc-bsu.cam.ac.uk/bug>. (MRC Biostatistics Unit, Cambridge, UK).

Spiegelhalter D J, Best N G, Carlin B P y van der Linde A (2002) Bayesian promedios Bayesianos de ajuste y complejidad modelar (con discusión). *J. Roy. Estadísticos. Soc. B.* 64, 583-640.

## ANEXO A

### **Programas en WINGBUGS**

## A.1 MODELO 1 (M1)

$$M1: bit \sim N(0, \sigma_t^2) \text{ y } \sigma_t^2 \sim IG(a, c)$$

```
model
{for (i in 1 : 6)
{
b1[i] ~ dnorm(0, tau.b[1])
b2[i] ~ dnorm(0, tau.b[2])
b3[i] ~ dnorm(0, tau.b[3])
b4[i] ~ dnorm(0, tau.b[4])
b5[i] ~ dnorm(0, tau.b[5])
b6[i] ~ dnorm(0, tau.b[6])
b7[i] ~ dnorm(0, tau.b[7])
b8[i] ~ dnorm(0, tau.b[8])
b9[i] ~ dnorm(0, tau.b[9])
b10[i] ~ dnorm(0, tau.b[10])
b11[i] ~ dnorm(0, tau.b[11])
b12[i] ~ dnorm(0, tau.b[12])
}
for (i in 1 :6)
{
for (t in 1:12)
{
Y[i,t] ~ dpois(mu[i,t])
}
log(mu[i, 1]) <- log(E[i, 1]) + beta[1]*chuva[i, 1] + b1[i]
log(mu[i, 2]) <- log(E[i, 2]) + beta[2]*chuva[i, 2] + b2[i]
log(mu[i, 3]) <- log(E[i, 3]) + beta[3]*chuva[i, 3] + b3[i]
log(mu[i, 4]) <- log(E[i, 4]) + beta[4]*chuva[i, 4] + b4[i]
log(mu[i, 5]) <- log(E[i, 5]) + beta[5]*chuva[i, 5] + b5[i]
log(mu[i, 6]) <- log(E[i, 6]) + beta[6]*chuva[i, 6] + b6[i]
log(mu[i, 7]) <- log(E[i, 7]) + beta[7]*chuva[i, 7] + b7[i]
log(mu[i, 8]) <- log(E[i, 8]) + beta[8]*chuva[i, 8] + b8[i]
log(mu[i, 9]) <- log(E[i, 9]) + beta[9]*chuva[i, 9] + b9[i]
log(mu[i, 10]) <- log(E[i, 10]) + beta[10]*chuva[i, 10] + b10[i]
log(mu[i, 11]) <- log(E[i, 11]) + beta[11]*chuva[i, 11] + b11[i]
log(mu[i, 12]) <- log(E[i, 12]) + beta[12]*chuva[i, 12] + b12[i]
RRenero[i] <- exp(beta[1]*chuva[i, 1] + b1[i])
RRfebrero[i] <- exp(beta[2]*chuva[i, 2] + b2[i])
RRmarzo[i] <- exp(beta[3]*chuva[i, 3] + b3[i])
RRabril[i] <- exp(beta[4]*chuva[i, 4] + b4[i])
RRmayo[i] <- exp(beta[5]*chuva[i, 5] + b5[i])
}
```

```

RRjunio[i] <- exp(beta[6]*chuva[i,6] + b6[i])
RRjulio[i] <- exp(beta[7]*chuva[i,7] + b7[i])
RRagosto[i] <- exp(beta[8]*chuva[i,8] + b8[i])
RRseptiembre[i] <- exp(beta[9]*chuva[i,9] + b9[i])
RRoctubre[i] <- exp(beta[10]*chuva[i,10] + b10[i])
RRnoviembre[i] <- exp(beta[11]*chuva[i,11] + b11[i])
RRdiciembre[i] <- exp(beta[12]*chuva[i,12] + b12[i])
muenero[i]<-mu[i,1]
mufebrero[i]<-mu[i,2]
mumarzo[i]<-mu[i,3]
muabril[i]<-mu[i,4]
mumayo[i]<-mu[i,5]
mujunio[i]<-mu[i,6]
mujulio[i]<-mu[i,7]
muagosto[i]<-mu[i,8]
museptiembre[i]<-mu[i,9]
muoctubre[i]<-mu[i,10]
munoviembre[i]<-mu[i,11]
mudiciembre[i]<-mu[i,12]
}
beta[1] ~ dnorm(0,tau.beta)
for (t in 2:12)
{
beta[t]- dnorm(beta[t-1],tau.beta)
}
for (t in 1:12)
{
tau.b[t] ~ dgamma (1, 1)
}
tau.beta ~ dgamma (1, 1)
}

list(Y=structure(.Data=c(507, 1192, 504, 738, 1507, 368, 1003, 968, 348, 943, 788, 828, 130, 281,
261, 603, 933, 55, 116, 162, 167, 222, 54, 266, 3121, 1124, 1658, 2529, 3581, 1748, 1433, 1020,
1016, 1056, 680, 736, 4, 3, 4, 7, 2, 5, 4, 4, 1, 1, 2, 3, 234, 142, 207, 168, 151, 68, 75, 71, 20, 65, 30,
14, 16, 25, 56, 10, 58, 102, 88, 71, 53, 78, 86, 99),.Dim=c(6, 12)),chuva=structure(.Data=c(12.8, 2.8,
18.5, 17.3, 9, 11.8, 10, 6.6, 8.4, 11.1, 10, 13.4, 8.4, 15.9, 16.8, 14.1, 17.1, 12.7, 11.4, 9.6, 9.9, 13.5,
11.1, 15.3, 12.3, 13.6, 8.6, 10.4, 12.6, 14.2, 12.2, 9.6, 9.2, 8.4, 10.2, 10.2, 10, 15, 12, 10, 6, 14, 8,
7.5, 12.5, 6.2, 8.4, 10.1, 18.5, 16, 12.5, 12.8, 10.9, 10, 11.4, 8.2, 9.2, 5.5, 6.9, 8.4, 14.4, 10.2, 12.8,
11.7, 8.7, 12.7, 7.3, 6.8, 11.5, 7.9, 5.8, 13.5),.Dim=c(6, 12)),E=structure(.Data=c(106.98, 64.08,
40.67, 23.58, 17.94, 19.82, 18.11, 47.66, 25.45, 69.19, 33.99, 59.27, 46.31, 27.74, 17.6, 10.21, 7.76,
8.58, 7.84, 20.64, 11.02, 29.97, 14.72, 25.68, 358.59, 214.83, 136.36, 79.07, 60.17, 66.48, 60.76,
159.95, 85.43, 232.24, 114.13, 199.03, 41.93, 25.11, 15.93, 9.23, 7.02, 7.75, 7.08, 18.64, 9.95,
27.04, 13.28, 23.15, 40.84, 24.44, 15.5, 8.98, 6.82, 7.53, 6.88, 18.1, 9.66, 26.23, 12.88, 22.44,
31.33, 18.77, 11.91, 6.91, 5.25, 5.81, 5.31, 13.98, 7.46, 20.3, 9.98, 17.4),.Dim=c(6, 12)))

```

## A.2 MODELO 2 (M2)

$$M2: bit \sim CAR(\sigma_t^2) \text{ y } \sigma_t^2 \sim IG(a, c)$$

```
model
{
```

```
  b1[1:6] ~ car.normal(adj[], weights[], num[], tau.b[1])
  b2[1:6] ~ car.normal(adj[], weights[], num[], tau.b[2])
  b3[1:6] ~ car.normal(adj[], weights[], num[], tau.b[3])
  b4[1:6] ~ car.normal(adj[], weights[], num[], tau.b[4])
  b5[1:6] ~ car.normal(adj[], weights[], num[], tau.b[5])
  b6[1:6] ~ car.normal(adj[], weights[], num[], tau.b[6])
  b7[1:6] ~ car.normal(adj[], weights[], num[], tau.b[7])
  b8[1:6] ~ car.normal(adj[], weights[], num[], tau.b[8])
  b9[1:6] ~ car.normal(adj[], weights[], num[], tau.b[9])
  b10[1:6] ~ car.normal(adj[], weights[], num[], tau.b[10])
  b11[1:6] ~ car.normal(adj[], weights[], num[], tau.b[11])
  b12[1:6] ~ car.normal(adj[], weights[], num[], tau.b[12])
```

```
  for (i in 1 :6)
```

```
  {
```

```
    for (t in 1:12)
```

```
    {
```

```
      Y[i,t] ~ dpois(mu[i,t])
```

```
    }
```

```
    log(mu[i,1]) <- log(E[i,1]) + beta[1]*chuva[i,1] + b1[i]
    log(mu[i,2]) <- log(E[i,2]) + beta[2]*chuva[i,2] + b2[i]
    log(mu[i,3]) <- log(E[i,3]) + beta[3]*chuva[i,3] + b3[i]
    log(mu[i,4]) <- log(E[i,4]) + beta[4]*chuva[i,4] + b4[i]
    log(mu[i,5]) <- log(E[i,5]) + beta[5]*chuva[i,5] + b5[i]
    log(mu[i,6]) <- log(E[i,6]) + beta[6]*chuva[i,6] + b6[i]
    log(mu[i,7]) <- log(E[i,7]) + beta[7]*chuva[i,7] + b7[i]
    log(mu[i,8]) <- log(E[i,8]) + beta[8]*chuva[i,8] + b8[i]
    log(mu[i,9]) <- log(E[i,9]) + beta[9]*chuva[i,9] + b9[i]
    log(mu[i,10]) <- log(E[i,10]) + beta[10]*chuva[i,10] + b10[i]
    log(mu[i,11]) <- log(E[i,11]) + beta[11]*chuva[i,11] + b11[i]
    log(mu[i,12]) <- log(E[i,12]) + beta[12]*chuva[i,12] + b12[i]
    RRenero[i] <- exp(beta[1]*chuva[i,1] + b1[i])
    RRfebrero[i] <- exp(beta[2]*chuva[i,2] + b2[i])
    RRmarzo[i] <- exp(beta[3]*chuva[i,3] + b3[i])
    RRabril[i] <- exp(beta[4]*chuva[i,4] + b4[i])
    RRmayo[i] <- exp(beta[5]*chuva[i,5] + b5[i])
    RRjunio[i] <- exp(beta[6]*chuva[i,6] + b6[i])
```



### A.3 MODELO 3 (M3)

$$M3: bit \sim N(0, \sigma_t^2) \text{ y } \log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \varepsilon_t$$

```

model
{for (i in 1 : 6)
  {
    b1[i] ~ dnorm(0, tau.b[1])
    b2[i] ~ dnorm(0, tau.b[2])
    b3[i] ~ dnorm(0, tau.b[3])
    b4[i] ~ dnorm(0, tau.b[4])
    b5[i] ~ dnorm(0, tau.b[5])
    b6[i] ~ dnorm(0, tau.b[6])
    b7[i] ~ dnorm(0, tau.b[7])
    b8[i] ~ dnorm(0, tau.b[8])
    b9[i] ~ dnorm(0, tau.b[9])
    b10[i] ~ dnorm(0, tau.b[10])
    b11[i] ~ dnorm(0, tau.b[11])
    b12[i] ~ dnorm(0, tau.b[12])
  }
for (i in 1 : 6)
  {
    for (t in 1:12)
      {
        Y[i,t] ~ dpois(mu[i,t])
      }
    log(mu[i,1]) <- log(E[i,1]) + beta[1]*chuva[i,1] + b1[i]
    log(mu[i,2]) <- log(E[i,2]) + beta[2]*chuva[i,2] + b2[i]
    log(mu[i,3]) <- log(E[i,3]) + beta[3]*chuva[i,3] + b3[i]
    log(mu[i,4]) <- log(E[i,4]) + beta[4]*chuva[i,4] + b4[i]
    log(mu[i,5]) <- log(E[i,5]) + beta[5]*chuva[i,5] + b5[i]
    log(mu[i,6]) <- log(E[i,6]) + beta[6]*chuva[i,6] + b6[i]
    log(mu[i,7]) <- log(E[i,7]) + beta[7]*chuva[i,7] + b7[i]
    log(mu[i,8]) <- log(E[i,8]) + beta[8]*chuva[i,8] + b8[i]
    log(mu[i,9]) <- log(E[i,9]) + beta[9]*chuva[i,9] + b9[i]
    log(mu[i,10]) <- log(E[i,10]) + beta[10]*chuva[i,10] + b10[i]
    log(mu[i,11]) <- log(E[i,11]) + beta[11]*chuva[i,11] + b11[i]
    log(mu[i,12]) <- log(E[i,12]) + beta[12]*chuva[i,12] + b12[i]
    RRenero[i] <- exp(beta[1]*chuva[i,1] + b1[i])
    RRfebrero[i] <- exp(beta[2]*chuva[i,2] + b2[i])
    RRmarzo[i] <- exp(beta[3]*chuva[i,3] + b3[i])
    RRabril[i] <- exp(beta[4]*chuva[i,4] + b4[i])
    RRmayo[i] <- exp(beta[5]*chuva[i,5] + b5[i])
    RRjunio[i] <- exp(beta[6]*chuva[i,6] + b6[i])
    RRjulio[i] <- exp(beta[7]*chuva[i,7] + b7[i])
  }
}

```

```

RRagosto[i] <- exp(beta[8]*chuva[i,8] + b8[i])
RRseptiembre[i] <- exp(beta[9]*chuva[i,9] + b9[i])
RRoctubre[i] <- exp(beta[10]*chuva[i,10] + b10[i])
RRnoviembre[i] <- exp(beta[11]*chuva[i,11] + b11[i])
RRdiciembre[i] <- exp(beta[12]*chuva[i,12] + b12[i])
muenero[i]<-mu[i,1]
mufebrero[i]<-mu[i,2]
mumarzo[i]<-mu[i,3]
muabril[i]<-mu[i,4]
mumayo[i]<-mu[i,5]
mujunio[i]<-mu[i,6]
mujulio[i]<-mu[i,7]
muagosto[i]<-mu[i,8]
museptiembre[i]<-mu[i,9]
muoctubre[i]<-mu[i,10]
munoviembre[i]<-mu[i,11]
mudiciembre[i]<-mu[i,12]
}
beta[1] ~ dnorm(0,tau.beta)
for(t in 1:12)
{
tau.b[t]<-exp(logtau.b[t])
}
for (t in 2:12)
{
beta[t]~ dnorm(beta[t-1],tau.beta)
logtau.b[t] ~ dnorm(logtau.b[t-1], 1)
}
logtau.b[1] ~ dnorm(0,1)
tau.beta ~ dgamma (1, 1)
}

```

```

list(Y=structure(.Data=c(507, 1192, 504, 738, 1507, 368, 1003, 968, 348, 943, 788, 828, 130, 281,
261, 603, 933, 55, 116, 162, 167, 222, 54, 266, 3121, 1124, 1658, 2529, 3581, 1748, 1433, 1020,
1016, 1056, 680, 736, 4, 3, 4, 7, 2, 5, 4, 4, 1, 1, 2, 3, 234, 142, 207, 168, 151, 68, 75, 71, 20, 65, 30,
14, 16, 25, 56, 10, 58, 102, 88, 71, 53, 78, 86, 99),.Dim=c(6, 12)),chuva=structure(.Data=c(12.8, 2.8,
18.5, 17.3, 9, 11.8, 10, 6.6, 8.4, 11.1, 10, 13.4, 8.4, 15.9, 16.8, 14.1, 17.1, 12.7, 11.4, 9.6, 9.9, 13.5,
11.1, 15.3, 12.3, 13.6, 8.6, 10.4, 12.6, 14.2, 12.2, 9.6, 9.2, 8.4, 10.2, 10.2, 10, 15, 12, 10, 6, 14, 8,
7.5, 12.5, 6.2, 8.4, 10.1, 18.5, 16, 12.5, 12.8, 10.9, 10, 11.4, 8.2, 9.2, 5.5, 6.9, 8.4, 14.4, 10.2, 12.8,
11.7, 8.7, 12.7, 7.3, 6.8, 11.5, 7.9, 5.8, 13.5),.Dim=c(6, 12)),E=structure(.Data=c(106.98, 64.08,
40.67, 23.58, 17.94, 19.82, 18.11, 47.66, 25.45, 69.19, 33.99, 59.27, 46.31, 27.74, 17.6, 10.21, 7.76,
8.58, 7.84, 20.64, 11.02, 29.97, 14.72, 25.68, 358.59, 214.83, 136.36, 79.07, 60.17, 66.48, 60.76,
159.95, 85.43, 232.24, 114.13, 199.03, 41.93, 25.11, 15.93, 9.23, 7.02, 7.75, 7.08, 18.64, 9.95,
27.04, 13.28, 23.15, 40.84, 24.44, 15.5, 8.98, 6.82, 7.53, 6.88, 18.1, 9.66, 26.23, 12.88, 22.44,
31.33, 18.77, 11.91, 6.91, 5.25, 5.81, 5.31, 13.98, 7.46, 20.3, 9.98, 17.4),.Dim=c(6, 12)))

```

#### A.4 MODELO 4 (M4)

$$M4: bit \sim CAR(\sigma_t^2) \text{ y } \log(\sigma_t^2) = \log(\sigma_{t-1}^2) + \varepsilon_t$$

```
model
{
```

```
b1[1:6] ~ car.normal(adj[], weights[], num[], tau.b[1])
b2[1:6] ~ car.normal(adj[], weights[], num[], tau.b[2])
b3[1:6] ~ car.normal(adj[], weights[], num[], tau.b[3])
b4[1:6] ~ car.normal(adj[], weights[], num[], tau.b[4])
b5[1:6] ~ car.normal(adj[], weights[], num[], tau.b[5])
b6[1:6] ~ car.normal(adj[], weights[], num[], tau.b[6])
b7[1:6] ~ car.normal(adj[], weights[], num[], tau.b[7])
b8[1:6] ~ car.normal(adj[], weights[], num[], tau.b[8])
b9[1:6] ~ car.normal(adj[], weights[], num[], tau.b[9])
b10[1:6] ~ car.normal(adj[], weights[], num[], tau.b[10])
b11[1:6] ~ car.normal(adj[], weights[], num[], tau.b[11])
b12[1:6] ~ car.normal(adj[], weights[], num[], tau.b[12])}
```

```
for (i in 1 :6)
```

```
{
```

```
for (t in 1:12)
```

```
{
```

```
Y[i,t] ~ dpois(mu[i,t])
```

```
}
```

```
log(mu[i, 1]) <- log(E[i, 1]) + beta[1]*chuva[i, 1] + b1[i]
```

```
log(mu[i, 2]) <- log(E[i, 2]) + beta[2]*chuva[i, 2] + b2[i]
```

```
log(mu[i, 3]) <- log(E[i, 3]) + beta[3]*chuva[i, 3] + b3[i]
```

```
log(mu[i, 4]) <- log(E[i, 4]) + beta[4]*chuva[i, 4] + b4[i]
```

```
log(mu[i, 5]) <- log(E[i, 5]) + beta[5]*chuva[i, 5] + b5[i]
```

```
log(mu[i, 6]) <- log(E[i, 6]) + beta[6]*chuva[i, 6] + b6[i]
```

```
log(mu[i, 7]) <- log(E[i, 7]) + beta[7]*chuva[i, 7] + b7[i]
```

```
log(mu[i, 8]) <- log(E[i, 8]) + beta[8]*chuva[i, 8] + b8[i]
```

```
log(mu[i, 9]) <- log(E[i, 9]) + beta[9]*chuva[i, 9] + b9[i]
```

```
log(mu[i, 10]) <- log(E[i, 10]) + beta[10]*chuva[i, 10] + b10[i]
```

```
log(mu[i, 11]) <- log(E[i, 11]) + beta[11]*chuva[i, 11] + b11[i]
```

```
log(mu[i, 12]) <- log(E[i, 12]) + beta[12]*chuva[i, 12] + b12[i]
```

```
RRenero[i] <- exp(beta[1]*chuva[i, 1] + b1[i])
```

```
RRfebrero[i] <- exp(beta[2]*chuva[i, 2] + b2[i])
```

```
RRmarzo[i] <- exp(beta[3]*chuva[i, 3] + b3[i])
```

```
RRabril[i] <- exp(beta[4]*chuva[i, 4] + b4[i])
```

```
RRmayo[i] <- exp(beta[5]*chuva[i, 5] + b5[i])
```

```
RRjunio[i] <- exp(beta[6]*chuva[i, 6] + b6[i])
```

```
RRjulio[i] <- exp(beta[7]*chuva[i, 7] + b7[i])
```



**ANEXO B**  
**Anexo Estadístico**

## B.1 Estimaciones del Número de Incidencias de Malaria por Provincia

### Cuadro N°B.1.1

Estimación del número de incidencias para el mes de Enero del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	506.9	22.48	463.8	506.6	552.1	100000
LORETO	130.5	12.18	107.8	130.1	155.6	100000
MAYNAS	3120	59.16	3004	3120	3237	100000
UCAYALI	4.917	2.111	1.668	4.623	9.806	100000
REQUENA	233	20.47	194.8	232.4	275.2	100000
MARISCAL RAMÓN CASTILLA	17.36	4.24	10.07	17	26.58	100000

Elaboración: Propia

### Cuadro N° B.1.2

Estimación del número de incidencias para el mes de Febrero del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	1196	129.8	960.5	1189	1471	100000
LORETO	281.1	19.23	241	280.7	340.9	100000
MAYNAS	1122	34.25	1055	1122	1191	100000
UCAYALI	3.823	2.002	0.9765	3.472	8.624	100000
REQUENA	141.6	13.11	117	141.2	169.3	100000
MARISCAL RAMÓN CASTILLA	25.37	4.884	16.73	25.04	35.8	100000

Elaboración: Propia

### Cuadro N° B.1.3

Estimación del número de incidencias para el mes de Marzo del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	504.5	32.58	430.5	503.8	617.1	100000
LORETO	261.2	18.41	226.3	260.7	298.5	100000
MAYNAS	1657	81.15	1482	1657	1824	100000
UCAYALI	4.876	2.174	1.586	4.553	9.994	100000
REQUENA	205.5	14.29	178.4	205.1	234.7	100000
MARISCAL RAMÓN CASTILLA	56.58	7.478	42.97	56.25	72.24	100000

Elaboración: Propia

### Cuadro N° B.1.4

Estimación del número de incidencias para el mes de Abril del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	738.8	39.56	573.2	738.1	961.6	100000
LORETO	602.3	25.3	553.8	601.8	653.8	100000
MAYNAS	2528	77.86	2282	2527	2703	100000
UCAYALI	7.52	2.612	3.286	7.23	13.44	100000
REQUENA	167.2	12.88	142.8	167	193.6	100000
MARISCAL RAMÓN CASTILLA	11.15	3.264	5.672	10.83	18.35	100000

Elaboración: Propia

### Cuadro N° B.1.5

Estimación del número de incidencias para el mes de Mayo del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	1506	74.43	1349	1505	1662	100000
LORETO	934.3	54.38	607.6	933.4	1368	100000
MAYNAS	3579	62.01	3197	3578	3870	100000
UCAYALI	2.74	1.482	0.6401	2.484	6.297	100000
REQUENA	150.8	12.32	127.6	150.5	175.8	100000
MARISCAL RAMÓN CASTILLA	58.19	7.649	44.26	57.84	74.17	100000

Elaboración: Propia

### Cuadro N° B.1.6

Estimación del número de incidencias para el mes de Junio del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	367.4	19.73	330.2	367	407.8	100000
LORETO	56.03	7.359	42.53	55.75	71.26	100000
MAYNAS	1748	41.61	1667	1747	1831	100000
UCAYALI	6.067	2.456	2.228	5.738	11.74	100000
REQUENA	67.02	8.001	52.24	66.67	83.69	100000
MARISCAL RAMÓN CASTILLA	102	10.01	83.28	101.6	122.7	100000

Elaboración: Propia

### Cuadro N° B.1.7

Estimación del número de incidencias para el mes de Julio del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	1002	34.72	934.3	1001	1072	100000
LORETO	116.5	10.72	96.3	116.2	138.5	100000
MAYNAS	1433	39.05	1357	1433	1515	100000
UCAYALI	4.674	2.02	1.585	4.397	9.369	100000
REQUENA	75.78	8.638	59.75	75.47	93.64	100000
MARISCAL RAMÓN CASTILLA	87.21	9.512	69.71	86.81	106.9	100000

Elaboración: Propia

### Cuadro N° B.1.8

Estimación del número de incidencias para el mes de Agosto del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	967	40.35	889.5	966.2	1048	100000
LORETO	162.2	13.14	137.4	161.8	188.8	100000
MAYNAS	1020	34.3	952.9	1019	1089	100000
UCAYALI	4.863	2.136	1.612	4.549	9.836	100000
REQUENA	71.14	8.325	55.81	70.8	88.42	100000
MARISCAL RAMÓN CASTILLA	70.77	8.312	55.47	70.42	88	100000

Elaboración: Propia

### Cuadro N° B.1.9

Estimación del número de incidencias para el mes de Septiembre del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	347.3	19.79	309.7	346.9	387.9	100000
LORETO	166.5	13.18	141.6	166.2	193.1	100000
MAYNAS	1015	31.9	953.1	1014	1078	100000
UCAYALI	1.849	1.424	0.1787	1.495	5.489	100000
REQUENA	20.84	4.444	13.04	20.53	30.35	100000
MARISCAL RAMÓN CASTILLA	53.07	8.136	10.47	52.65	108	100000

Elaboración: Propia

### Cuadro N° B.1.10

Estimación del número de incidencias para el mes de Octubre del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	942.1	33.15	877.7	941.8	1008	100000
LORETO	222.6	19.68	180.1	222	276	100000
MAYNAS	1056	46.47	963.9	1055	1149	100000
UCAYALI	1.989	1.325	0.2734	1.709	5.282	100000
REQUENA	64.59	8.88	48.66	64.07	83.55	100000
MARISCAL RAMÓN CASTILLA	78.02	9.006	61.44	77.69	96.65	100000

Elaboración: Propia

### Cuadro N° B.1.11

Estimación del número de incidencias para el mes de Noviembre del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	786.9	27.96	733	786.6	842.3	100000
LORETO	54.73	7.546	40.92	54.38	70.42	100000
MAYNAS	679.9	26.06	629.9	679.5	732.1	100000
UCAYALI	2.691	1.557	0.5449	2.398	6.487	100000
REQUENA	30.2	5.291	20.77	29.89	41.39	100000
MARISCAL RAMÓN CASTILLA	85.5	10.23	66.8	84.98	107.1	100000

Elaboración: Propia

### Cuadro N° B.1.12

Estimación del número de incidencias para el mes de Diciembre del 2009  
por Provincia

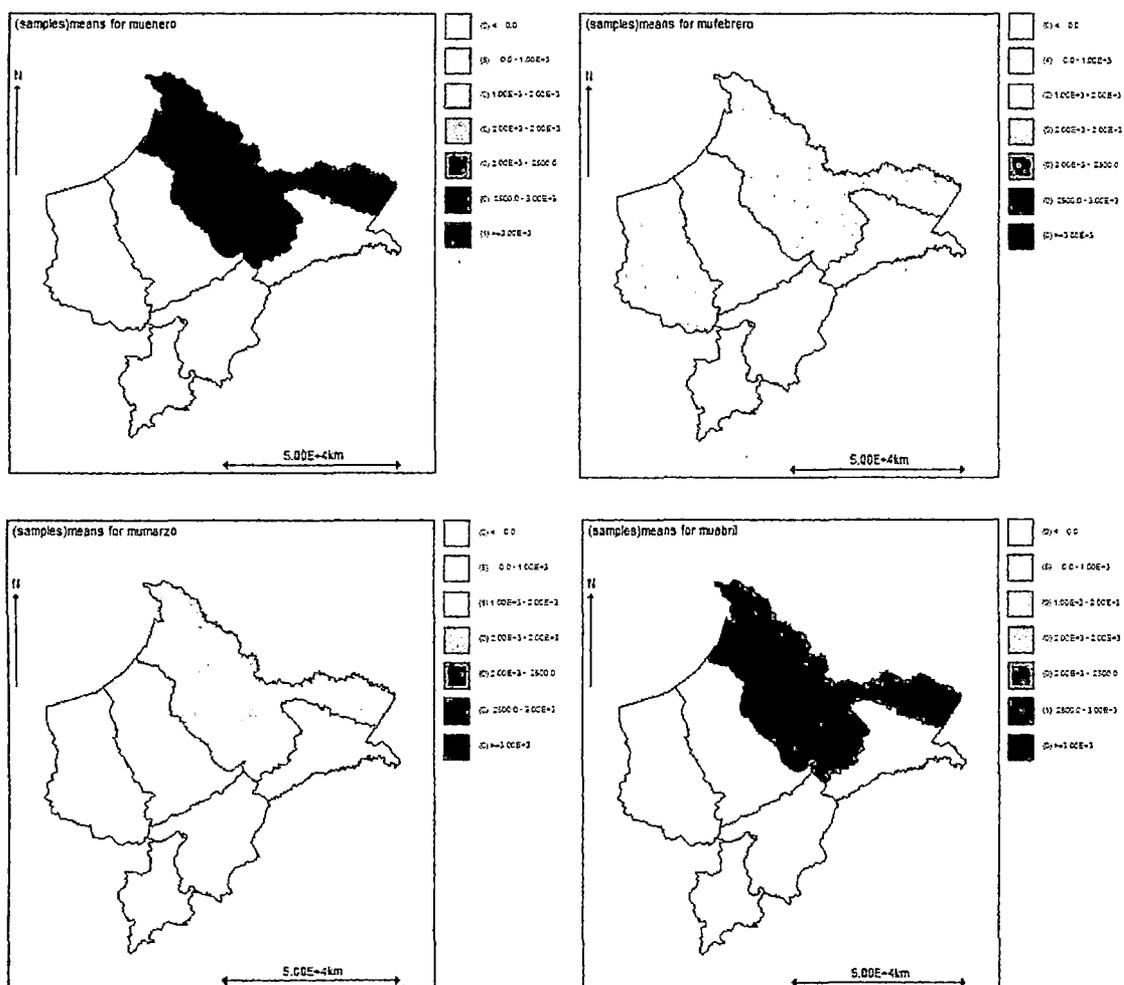
Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	827.6	29.05	771.4	827.2	885.3	100000
LORETO	265.6	18.07	228.5	265.3	306.4	100000
MAYNAS	736.4	34.8	669.3	735.9	807.1	100000
UCAYALI	3.408	1.742	0.8992	3.126	7.55	100000
REQUENA	15.48	3.743	9.002	15.2	23.56	100000
MARISCAL RAMÓN CASTILLA	98.58	9.973	79.95	98.23	119.1	100000

Elaboración: Propia

## B.2 Mapeo de Estimaciones de Incidencias de Malaria por Provincia

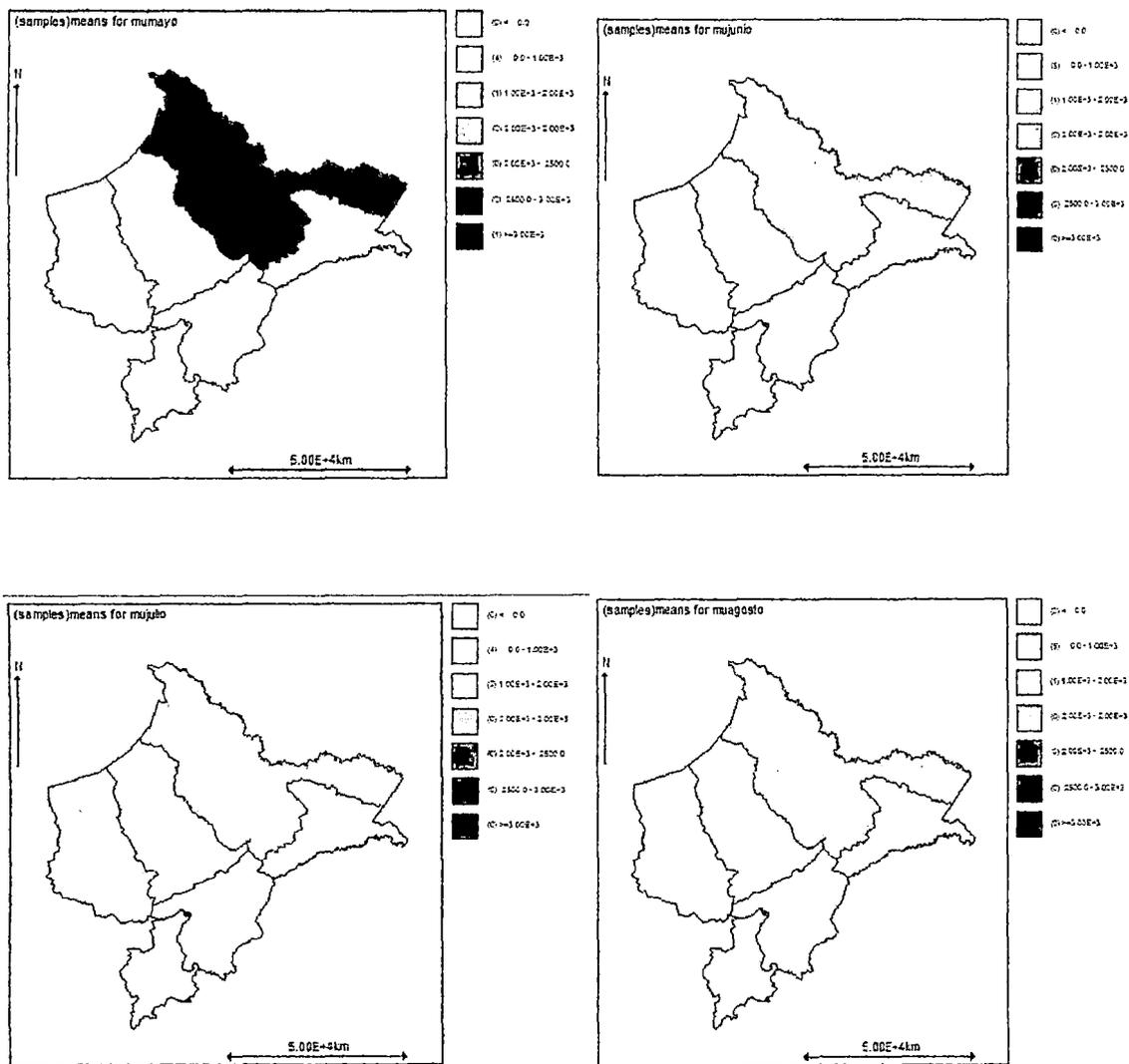
Gráfico N° B.2.1

Estimación del Número de Incidencias de Malaria por Provincia en el departamento de Loreto (Enero-Abril 2009)



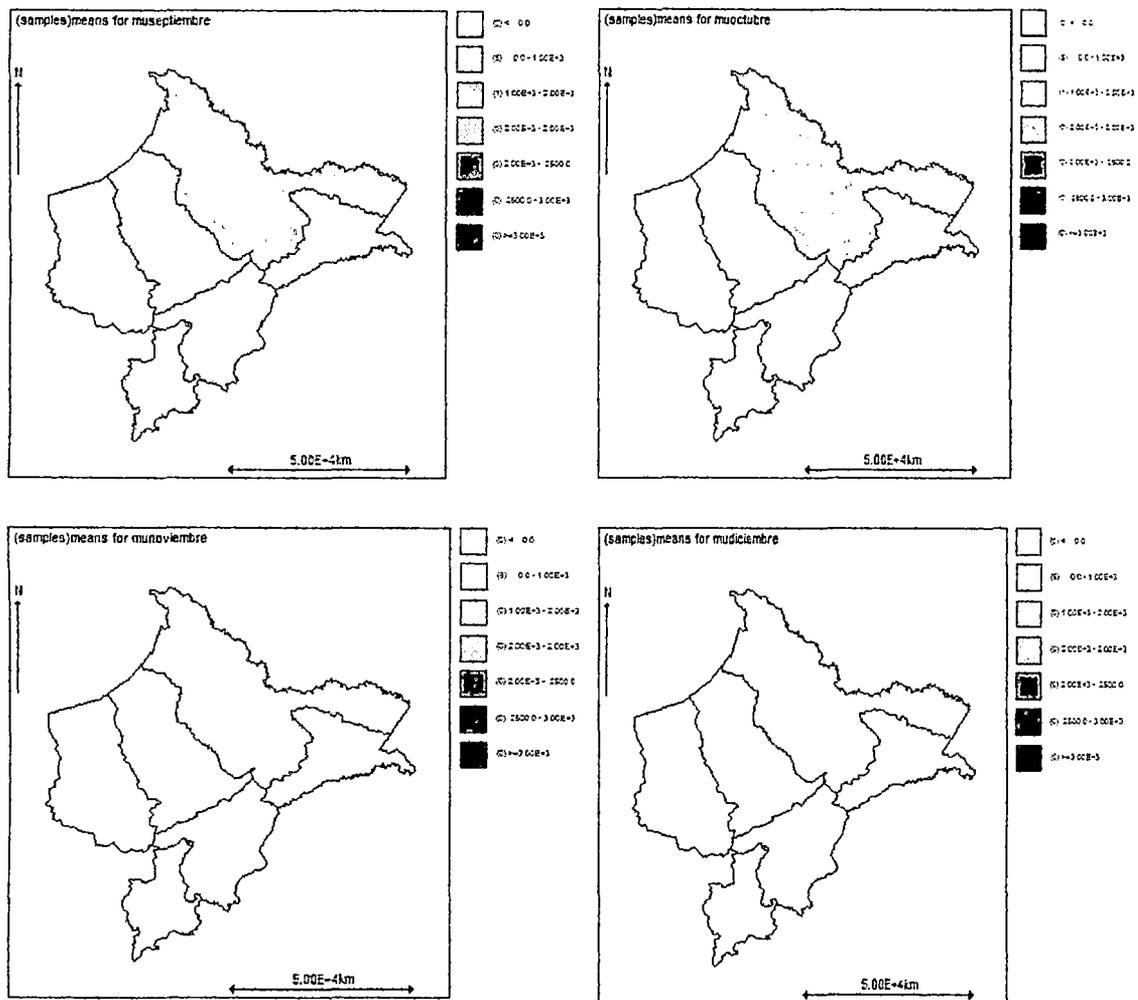
### Gráfico N°B.2.2

Estimación del Número de Incidencias de Malaria por Provincia en el departamento de Loreto (Mayo-Agosto 2009)



### Gráfico N°B.2.3

## Estimación del Número de Incidencias de Malaria por Provincia en el departamento de Loreto (Septiembre-Diciembre 2009)



### B.3 Estimaciones de los Riesgos Relativos por Provincia

Cuadro N°B.3.1

Estimación del Riesgo Relativo para el mes de Enero del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	4.739	0.2102	4.336	4.735	5.16	100000
LORETO	2.817	0.263	2.327	2.808	3.359	100000
MAYNAS	8.7	0.165	8.378	8.7	9.026	100000
UCAYALI	0.1173	0.05034	0.03979	0.1103	0.2341	100000
REQUENA	5.706	0.5011	4.77	5.688	6.734	100000
MARISCAL RAMÓN CASTILLA	0.554	0.1353	0.3211	0.5426	0.8479	100000

Elaboración: Propia

Cuadro N° B.3.2

Estimación del Riesgo Relativo para el mes de Febrero del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	18.66	2.025	15	18.55	22.95	100000
LORETO	10.13	0.6932	8.814	10.12	11.52	100000
MAYNAS	5.224	0.1594	4.916	5.223	5.538	100000
UCAYALI	0.1523	0.07971	0.03889	0.1383	0.3436	100000
REQUENA	5.793	0.5366	4.796	5.776	6.895	100000
MARISCAL RAMÓN CASTILLA	1.352	0.2602	0.8928	1.334	1.907	100000

Elaboración: Propia

### Cuadro N° B.3.3

Estimación del Riesgo Relativo para el mes de Marzo del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	12.4	0.8011	10.87	12.39	15.17	100000
LORETO	14.84	1.046	12.86	14.81	16.96	100000
MAYNAS	12.15	0.5951	11.01	12.15	13.34	100000
UCAYALI	0.3061	0.1365	0.09962	0.2858	0.6255	100000
REQUENA	13.26	0.922	11.51	13.23	15.14	100000
MARISCAL RAMÓN CASTILLA	4.75	0.6278	3.606	4.722	6.056	100000

Elaboración: Propia

### Cuadro N° B.3.4

Estimación del Riesgo Relativo para el mes de Abril del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	31.33	1.678	24.31	31.3	40.78	100000
LORETO	58.99	2.478	54.27	58.94	63.93	100000
MAYNAS	31.97	0.9847	30.07	31.96	33.9	100000
UCAYALI	0.8148	0.283	0.3563	0.7835	1.457	100000
REQUENA	18.62	1.434	15.92	18.6	21.55	100000
MARISCAL RAMÓN CASTILLA	1.613	0.4724	0.8204	1.567	2.655	100000

Elaboración: Propia

### Cuadro N° B.3.5

Estimación del Riesgo Relativo para el mes de Mayo del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	83.97	4.149	75.96	83.87	92.54	100000
LORETO	120.4	7.008	78.29	120.3	176.3	100000
MAYNAS	59.48	1.031	53.14	59.48	64.32	100000
UCAYALI	0.3903	0.2111	0.09116	0.3539	0.8967	100000
REQUENA	22.11	1.807	18.71	22.06	25.77	100000
MARISCAL RAMÓN CASTILLA	11.08	1.457	8.429	11.02	14.13	100000

Elaboración: Propia

### Cuadro N° B.3.6

Estimación del Riesgo Relativo para el mes de Junio del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	18.54	0.9956	16.65	18.52	20.54	100000
LORETO	6.53	0.8577	4.949	6.497	8.311	100000
MAYNAS	26.29	0.6259	25.07	26.28	27.53	100000
UCAYALI	0.7828	0.3169	0.2877	0.7405	1.514	100000
REQUENA	8.9	1.063	6.946	8.853	11.11	100000
MARISCAL RAMÓN CASTILLA	17.55	1.723	14.33	17.49	21.09	100000

Elaboración: Propia

### Cuadro N° B.3.7

Estimación del Riesgo Relativo para el mes de Julio del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	55.31	1.917	51.66	55.27	59.16	100000
LORETO	14.86	1.368	12.3	14.82	17.67	100000
MAYNAS	23.58	0.6426	22.33	23.58	24.85	100000
UCAYALI	0.6601	0.2853	0.2238	0.621	1.323	100000
REQUENA	11.01	1.255	8.686	10.97	13.62	100000
MARISCAL RAMÓN CASTILLA	16.42	1.791	13.13	16.35	20.14	100000

Elaboración: Propia

### Cuadro N° B.3.8

Estimación del Riesgo Relativo para el mes de Agosto del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	20.29	0.8467	18.68	20.27	21.98	100000
LORETO	7.859	0.6367	6.658	7.841	9.146	100000
MAYNAS	6.376	0.2144	5.962	6.373	6.803	100000
UCAYALI	0.2609	0.1146	0.08664	0.2441	0.5281	100000
REQUENA	3.93	0.46	3.083	3.911	4.885	100000
MARISCAL RAMÓN CASTILLA	5.062	0.5946	3.967	5.036	6.295	100000

Elaboración: Propia

### Cuadro N° B.3.9

Estimación del Riesgo Relativo para el mes de Septiembre del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	13.65	0.7774	12.17	13.63	15.22	100000
LORETO	15.11	1.196	12.85	15.07	17.53	100000
MAYNAS	11.88	0.3734	11.16	11.87	12.62	100000
UCAYALI	0.1858	0.1431	0.01791	0.1503	0.5523	100000
REQUENA	2.157	0.46	1.35	2.125	3.142	100000
MARISCAL RAMÓN CASTILLA	7.115	1.091	1.403	7.057	14.48	100000

Elaboración: Propia

### Cuadro N° B.3.10

Estimación del Riesgo Relativo para el mes de Octubre del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	13.62	0.4791	12.68	13.61	14.57	100000
LORETO	7.426	0.6566	6.197	7.406	8.771	100000
MAYNAS	4.547	0.2001	4.161	4.544	4.947	100000
UCAYALI	0.07356	0.049	0.01004	0.06322	0.1955	100000
REQUENA	2.462	0.3385	1.855	2.443	3.185	100000
MARISCAL RAMÓN CASTILLA	3.844	0.4437	3.028	3.827	4.758	100000

Elaboración: Propia

### Cuadro N° B.3.11

Estimación del Riesgo Relativo para el mes de Noviembre del 2009  
por Provincia

Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	23.15	0.8226	21.57	23.14	24.79	100000
LORETO	3.718	0.5126	2.78	3.695	4.787	100000
MAYNAS	5.957	0.2283	5.519	5.954	6.413	100000
UCAYALI	0.2027	0.1173	0.04107	0.1806	0.4887	100000
REQUENA	2.345	0.4108	1.614	2.321	3.212	100000
MARISCAL RAMÓN CASTILLA	8.567	1.025	6.705	8.515	10.73	100000

Elaboración: Propia

### Cuadro N° B.3.12

Estimación del Riesgo Relativo para el mes de Diciembre del 2009  
por Provincia

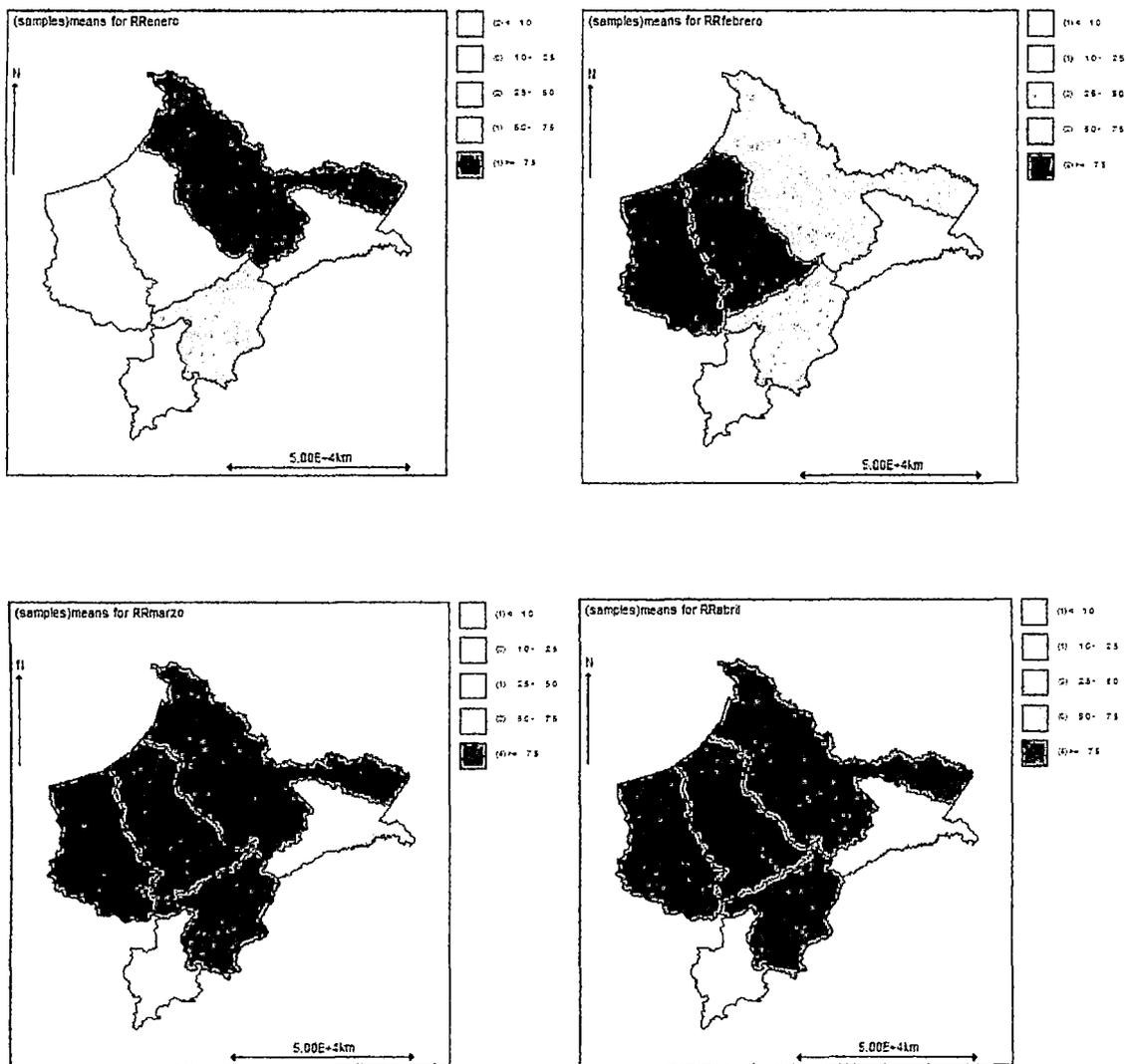
Provincia	media	desviación	val2.5	mediana	val97.5	interacciones
ALTO AMAZONAS	13.96	0.4901	13.02	13.96	14.93	100000
LORETO	10.34	0.7038	9.013	10.33	11.76	100000
MAYNAS	3.7	0.1749	3.366	3.697	4.051	100000
UCAYALI	0.1472	0.07523	0.03883	0.135	0.3264	100000
REQUENA	0.69	0.1668	0.4012	0.6774	1.051	100000
MARISCAL RAMÓN CASTILLA	5.665	0.5732	4.598	5.646	6.843	100000

Elaboración: Propia

## B.4 Mapeo de Estimaciones de los Riesgos Relativos por Provincia

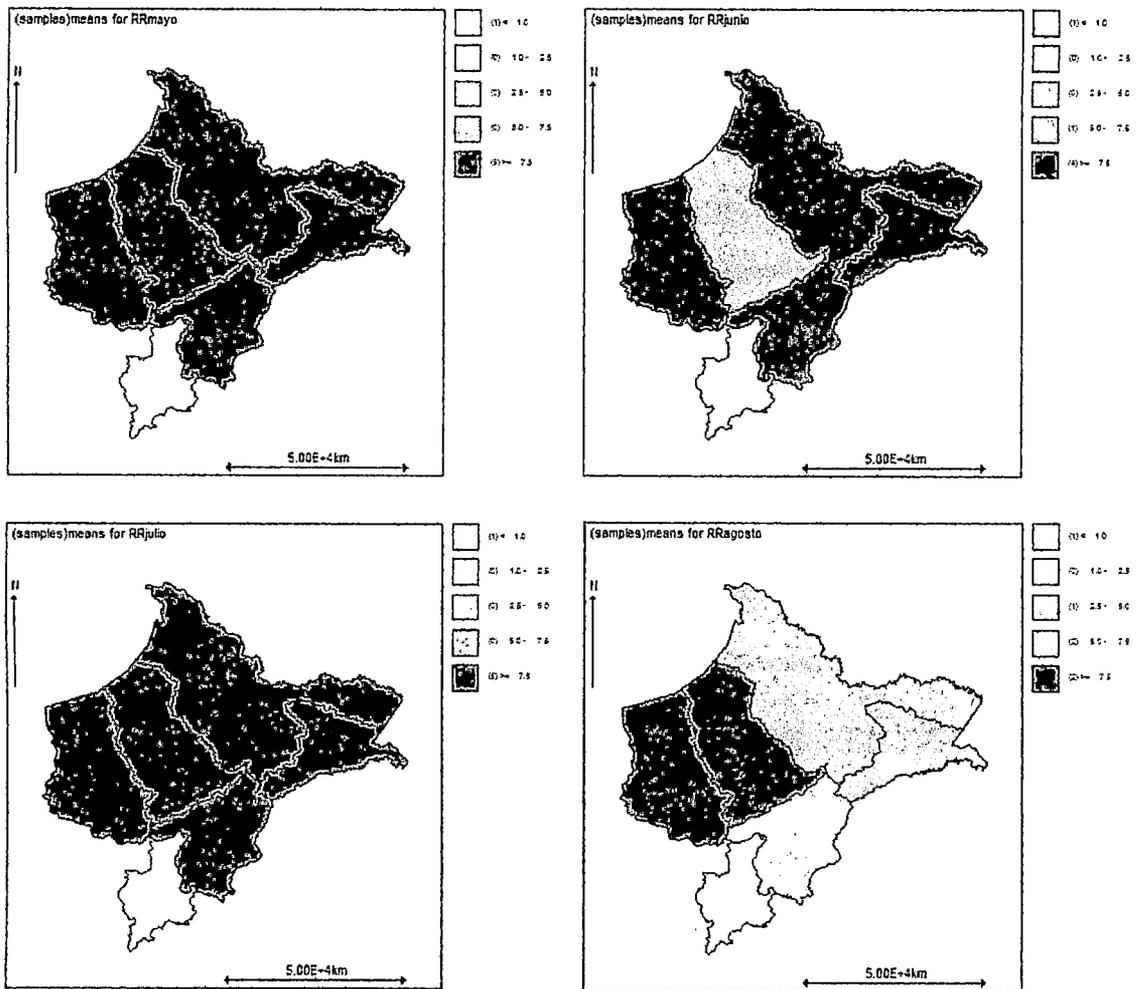
Gráfico N°B.4.1

Estimación del Riesgo Relativo por Provincia en el departamento de Loreto (Enero-Abril 2009)



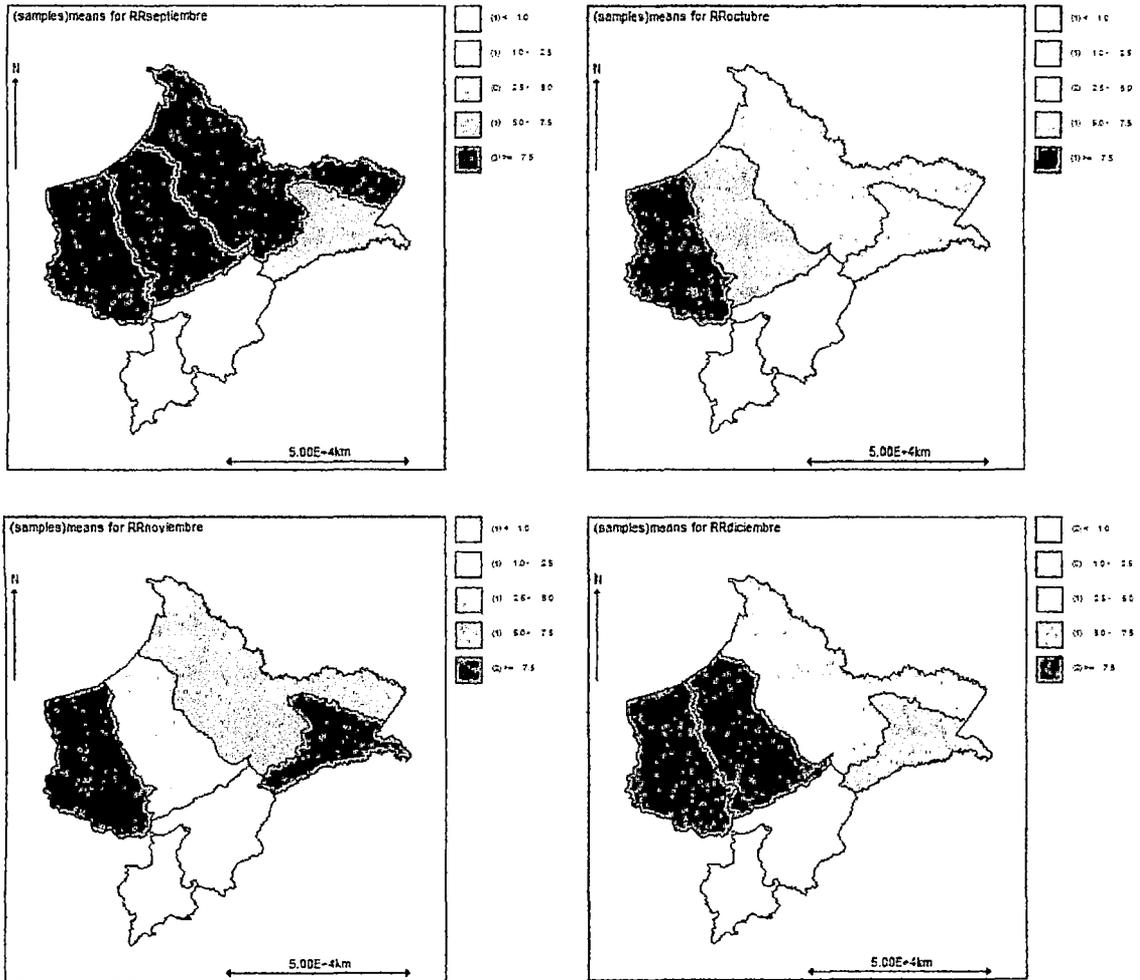
### Gráfico N°B.4.2

## Estimación del Riesgo Relativo por Provincia en el departamento de Loreto (Mayo-Agosto 2009)



### Gráfico N°B.4.3

## Estimación del Riesgo Relativo por Provincia en el departamento de Loreto (Septiembre-Diciembre 2009)



## B.5 Estimaciones del Efecto de las Áreas Circundantes por Provincia

### Cuadro N°B.5.1

Estimación del Efecto de áreas circundantes por Provincia durante el mes de Enero

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	0.947	0.098	0.0003611	0.7669	0.943	1.15	1	100000
LORETO	0.6326	0.1185	0.00032	0.4053	0.6302	0.8703	1	100000
MAYNAS	1.579	0.0922	0.0003305	1.412	1.574	1.773	1	100000
UCAYALI	-2.714	0.3925	0.001398	-3.567	-2.685	-2.028	1	100000
REQUENA	0.8593	0.1145	0.0005214	0.6437	0.856	1.095	1	100000
MARISCAL RAMÓN CASTILLA	-1.305	0.2159	0.0006872	-1.747	-1.298	-0.8971	1	100000

Elaboración: Propia

### Cuadro N°B.5.2

Estimación del Efecto de áreas circundantes por Provincia durante el mes de Febrero

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	2.658	0.128	0.0002095	2.423	2.655	2.905	1	100000
LORETO	0.8213	0.1161	0.0004834	0.6113	0.8169	1.06	1	100000
MAYNAS	0.3764	0.1028	0.0003825	0.1928	0.3706	0.5935	1	100000
UCAYALI	-3.432	0.4512	0.001513	-4.423	-3.395	-2.653	1	100000
REQUENA	0.2507	0.125	0.0004958	0.02042	0.2462	0.5054	1	100000
MARISCAL RAMÓN CASTILLA	-0.6745	0.1889	0.0005394	-1.051	-0.6722	-0.3097	1	100000

Elaboración: Propia

### Cuadro N°B.5.3

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Marzo

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	0.112	0.1006	0.0004366	-0.06927	0.1077	0.3182	1	100000
LORETO	0.5116	0.1019	0.0003853	0.3246	0.5076	0.7209	1	100000
MAYNAS	1.379	0.09294	0.0002216	1.212	1.375	1.57	1	100000
UCAYALI	-2.846	0.3979	0.001482	-3.708	-2.817	-2.152	1	100000
REQUENA	0.9576	0.1018	0.0003489	0.7669	0.9544	1.167	1	100000
MARISCAL RAMÓN CASTILLA	-0.1139	0.1371	0.0004436	-0.3821	-0.1143	0.1565	1	100000

Elaboración: Propia

### Cuadro N°B.5.4

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Abril

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	0.2522	0.09492	0.0004308	0.07694	0.2494	0.4438	1	100000
LORETO	1.476	0.08985	0.0003612	1.308	1.473	1.66	1	100000
MAYNAS	1.546	0.08528	0.0002603	1.389	1.543	1.721	1	100000
UCAYALI	-2.111	0.3137	0.001092	-2.777	-2.092	-1.545	1	100000
REQUENA	0.5605	0.1035	0.0003654	0.3618	0.5591	0.7674	1	100000
MARISCAL RAMÓN CASTILLA	-1.723	0.2604	0.0008657	-2.267	-1.712	-1.247	1	100000

Elaboración: Propia

### Cuadro N°B.5.5

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Mayo

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	1.952	0.1121	0.0003573	1.758	1.943	2.197	1	100000
LORETO	0.08251	0.1323	0.0006623	-0.1454	0.07361	0.3583	1	100000
MAYNAS	0.6175	0.1131	0.0004767	0.4232	0.6086	0.8644	1	100000
UCAYALI	-2.747	0.5357	0.002119	-3.936	-2.695	-1.844	1	100000
REQUENA	0.09249	0.1273	0.0004389	-0.1386	0.0855	0.362	1	100000
MARISCAL RAMÓN CASTILLA	0.00219	0.1533	0.0004628	-0.2885	-0.001188	0.3133	1	100000

Elaboración: Propia

### Cuadro N° B.5.6

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Junio

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	0.9208	0.08984	0.0003032	0.7528	0.9177	1.105	1	100000
LORETO	-0.2822	0.1337	0.0004391	-0.5459	-0.2816	-0.02104	1	100000
MAYNAS	0.8648	0.08244	0.0003335	0.7146	0.8609	1.038	1	100000
UCAYALI	-2.7	0.353	0.001314	-3.465	-2.676	-2.079	1	100000
REQUENA	0.486	0.1272	0.0004194	0.238	0.4858	0.7375	1	100000
MARISCAL RAMÓN CASTILLA	0.7101	0.1132	0.0003794	0.4915	0.7091	0.935	1	100000

Elaboración: Propia

### Cuadro N° B.5.7

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Julio

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	1.601	0.08942	0.0003258	1.44	1.596	1.791	1	100000
LORETO	-0.05443	0.1142	0.0004053	-0.2723	-0.05657	0.1761	1	100000
MAYNAS	0.2182	0.0905	0.0003797	0.05525	0.2132	0.4098	1	100000
UCAYALI	-2.441	0.3984	0.001451	-3.307	-2.411	-1.745	1	100000
REQUENA	-0.3563	0.1255	0.0004036	-0.5985	-0.3571	-0.1052	1	100000
MARISCAL RAMÓN CASTILLA	1.032	0.1244	0.0003635	0.7937	1.031	1.281	1	100000

Elaboración: Propia

### Cuadro N° B.5.8

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Agosto

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	1.854	0.08947	0.0002645	1.692	1.85	2.042	1	100000
LORETO	0.3784	0.1079	0.000411	0.1754	0.3756	0.5985	1	100000
MAYNAS	0.1719	0.09073	0.0003691	0.008047	0.1678	0.3622	1	100000
UCAYALI	-2.756	0.3936	0.001444	-3.609	-2.728	-2.071	1	100000
REQUENA	-0.07316	0.1272	0.0004217	-0.3188	-0.0747	0.1802	1	100000
MARISCAL RAMÓN CASTILLA	0.4248	0.1295	0.0004338	0.1742	0.4235	0.6822	1	100000

Elaboración: Propia

**Cuadro N° B.5.9**

**Estimación del Efecto de áreas circundantes por Provincia durante el mes de Septiembre**

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	1.436	0.1544	0.0005266	1.171	1.423	1.78	1	100000
LORETO	1.327	0.1613	0.0006141	1.046	1.315	1.679	1	100000
MAYNAS	1.187	0.1497	0.0005492	0.9338	1.173	1.524	1	100000
UCAYALI	-3.75	0.6892	0.002697	-5.338	-3.667	-2.643	1	100000
REQUENA	-0.5417	0.2213	0.0007192	-0.9727	-0.5434	-0.1002	1	100000
MARISCAL RAMÓN CASTILLA	0.3415	0.1887	0.0007721	-0.004529	0.3325	0.7402	1	100000

Elaboración: Propia

**Cuadro N° B.5.10**

**Estimación del Efecto de áreas circundantes por Provincia durante el mes de Octubre**

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	1.449	0.1483	0.0007603	1.202	1.435	1.782	1	100000
LORETO	0.5885	0.1699	0.0009405	0.2946	0.5743	0.9607	1	100000
MAYNAS	0.6344	0.1396	0.0005723	0.4037	0.6196	0.9494	1	100000
UCAYALI	-3.501	0.676	0.003166	-5.051	-3.42	-2.414	1	100000
REQUENA	0.3162	0.1702	0.0005316	0.00396	0.3082	0.6761	1	100000
MARISCAL RAMÓN CASTILLA	0.5131	0.1642	0.000654	0.2166	0.5034	0.8641	1	100000

Elaboración: Propia

### Cuadro N° B.5.11

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Noviembre

Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	1.723	0.1224	0.0004903	1.508	1.714	1.988	1	100000
LORETO	-0.271	0.1636	0.0006319	-0.5818	-0.2743	0.0606	1	100000
MAYNAS	0.337	0.1234	0.0005022	0.1202	0.3282	0.605	1	100000
UCAYALI	-2.965	0.5405	0.001978	-4.177	-2.911	-2.059	1	100000
REQUENA	-0.1419	0.1847	0.0005524	-0.5044	-0.1426	0.2235	1	100000
MARISCAL RAMÓN CASTILLA	1.318	0.1495	0.0004141	1.038	1.313	1.627	1	100000

Elaboración: Propia

### Cuadro N° B.5.12

#### Estimación del Efecto de áreas circundantes por Provincia durante el mes de Diciembre

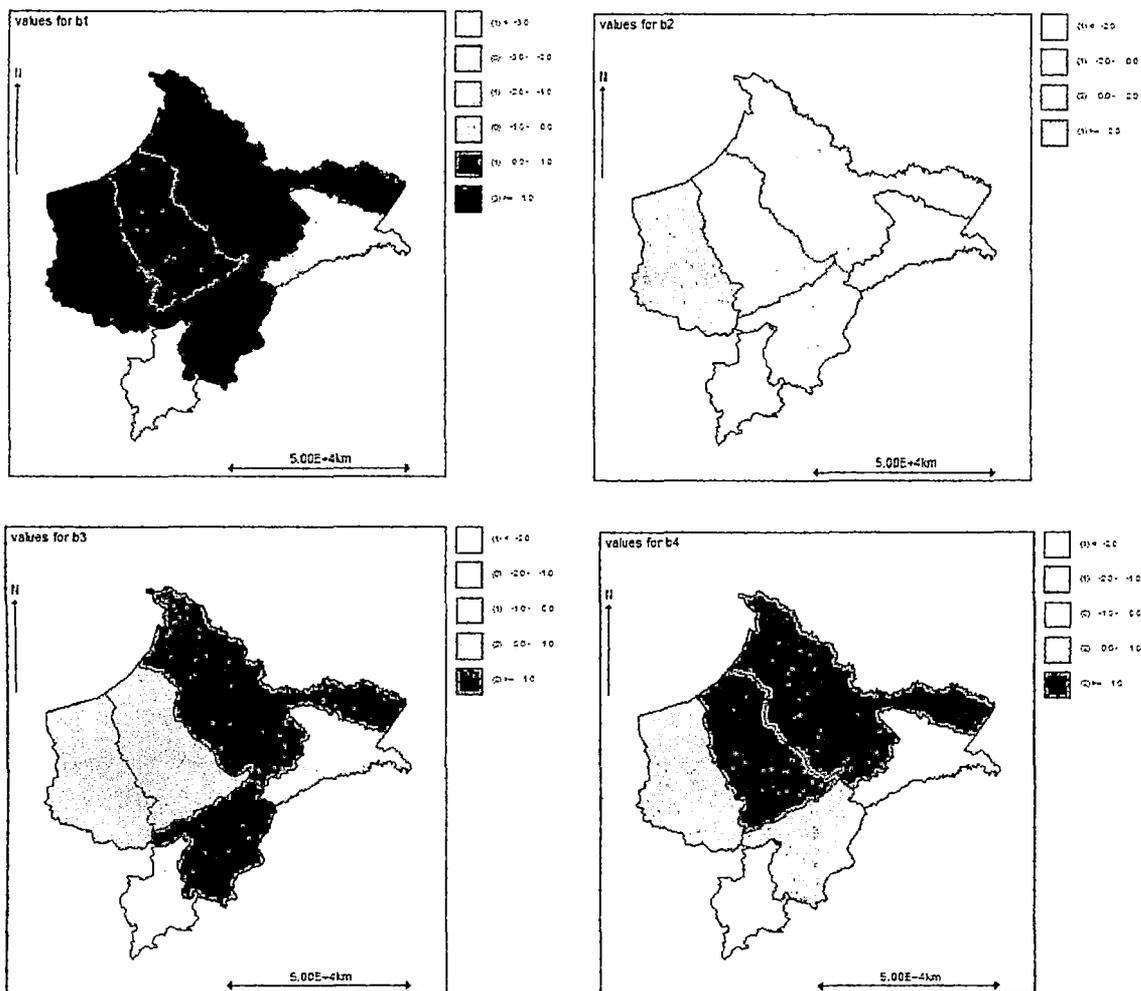
Provincia	media	desviación	Error MC	val2.5	mediana	val97.5	inicio	interacciones
ALTO AMAZONAS	1.586	0.113	0.000429	1.383	1.58	1.827	1	100000
LORETO	1.136	0.1242	0.0004984	0.9089	1.13	1.394	1	100000
MAYNAS	0.5082	0.1119	0.0003375	0.3068	0.5021	0.7474	1	100000
UCAYALI	-2.843	0.471	0.001546	-3.873	-2.801	-2.038	1	100000
REQUENA	-1.059	0.2313	0.0007873	-1.534	-1.051	-0.6239	1	100000
MARISCAL RAMÓN CASTILLA	0.6718	0.137	0.000495	0.4133	0.6685	0.9518	1	100000

Elaboración: Propia

## B.6 Mapeo del Efecto de las Áreas Circundantes por Provincia

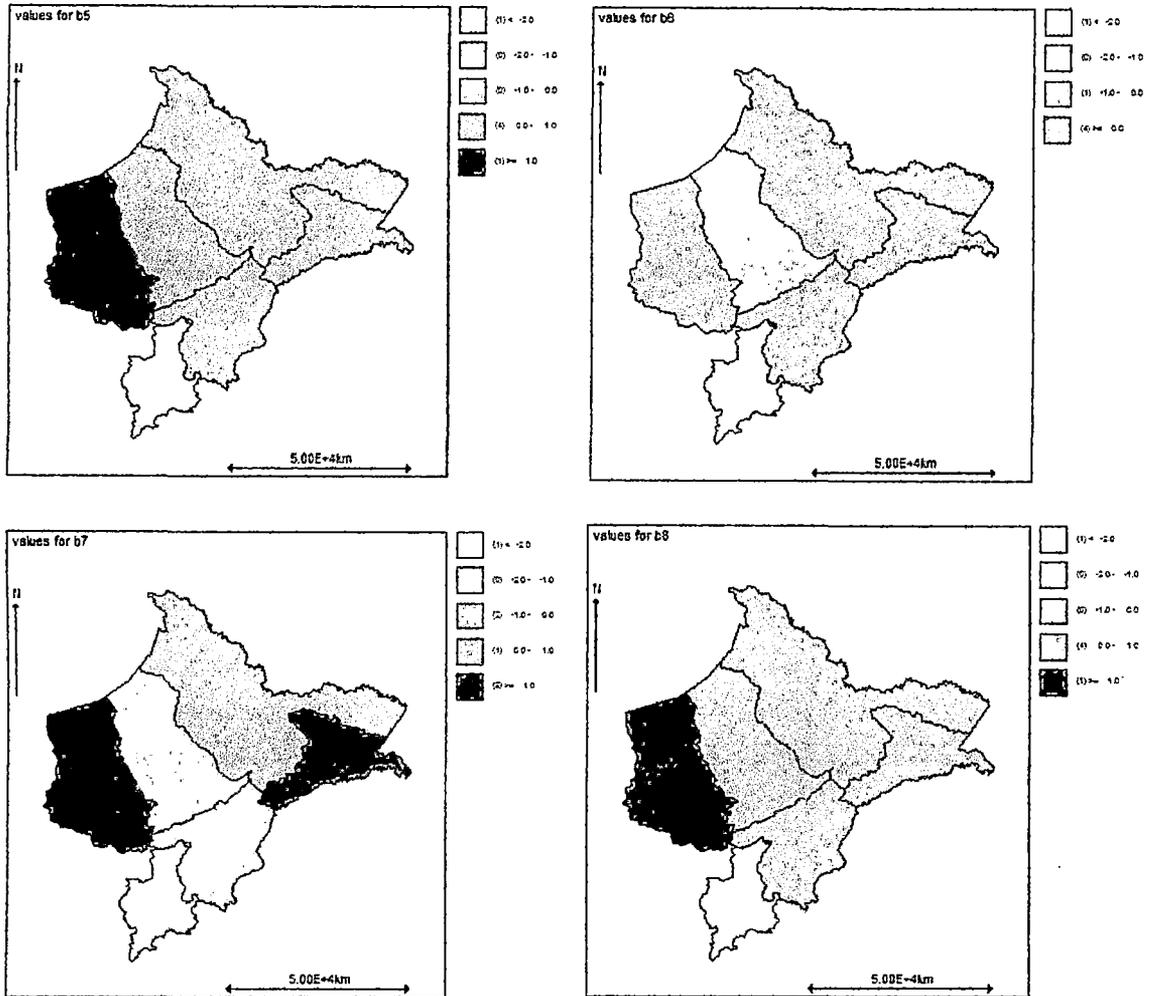
Gráfico N°B.6.1

Estimación del Efecto de las Áreas Circundantes  
por Meses en el departamento de Loreto  
(Enero-Abril 2009)



**Gráfico N°B.6.2**

**Estimación del Efecto de las Áreas Circundantes  
por Meses en el departamento de Loreto  
(Mayo-Agosto 2009)**



### Gráfico N°B.6.3

## Estimación del Efecto de las Áreas Circundantes por Meses en el departamento de Loreto (Septiembre-Diciembre 2009)

