



Universidad Autónoma de Guerrero

Facultad de Matemáticas

Maestría en Matemáticas Aplicadas

**Estimación bayesiana de los
parámetros en dos sistemas
dinámicos que describen la
transmisión de la enfermedad
del dengue.**

TESIS

PARA OBTENER EL GRADO DE:

Maestro en Matemáticas Aplicadas

PRESENTA:

ALIAN LI MARTÍN

DIRECTORES DE TESIS:

Dr. Ramón Reyes Carreto

M. en C. Cruz Vargas de León

Abril 2019.

Dedicatoria

Dedico esta tesis a mi padres, Maria Isabel Martín del Toro y Guillermo Li Cardoso y a mi hermano Alain Li Martín por haber creído en mi y que con sus palabras siempre me dieron las fuerzas para seguir adelante y no me rindiera en el camino. A mi esposa Yaneisy Enriquez Sánchez por todo su amor, esfuerzo y sacrificio que me ha brindado, y por apoyarme en los momentos más difíciles. A toda mi familia por todo su amor y apoyo incondicional que me han brindado durante este tiempo.

Agradecimientos

A mis padres y mi hermano, por sus valiosas enseñanzas y todos sus sacrificios que han hecho para darme la oportunidad de estudiar.

A mi esposa por siempre haberme acompañado durante este maravilloso viaje.

A toda mi familia que siempre me ha apoyado a pesar de la distancia.

A mis amigos y compañeros, por haber compartido conmigo todos sus conocimientos, alegrías y tristezas.

A todas las personas que nos han dado su apoyo en nuestra llegada a Mexico.

A mis asesores, por creer en mi y permitirme trabajar a su lado.

A todos los profesores que durante estos dos años nos brindaron sus conocimientos y enseñanzas para que fuésemos mejores profesionales.

A la Maestría de Matemáticas Aplicadas de la Facultad de Matemáticas de la Universidad Autónoma de Guerrero por darme la oportunidad de realizar aquí mis estudios de posgrados.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por haberme brindado el apoyo económico necesario para realizar mis estudios de posgrado.

Resumen

El dengue es una enfermedad viral transmitida por el mosquito hembra de la especie *Aedes aegypti*. Esta enfermedad se distribuye en países de África, Asia, Oceanía y América, donde existen condiciones geográficas, demográficas y epidemiológicas propicias para: el vector, el virus y el hospedero. Para el estudio de la dinámica del virus, los mecanismos implicados en la transmisión del virus entre los hospederos y los vectores, se han propuestos modelos matemáticos en el lenguaje de las ecuaciones diferenciales ordinarias. Para estimar los parámetros que contienen dichos modelos generalmente se ha utilizado el método de mínimos cuadrados. El objetivo de este trabajo consiste en utilizar el enfoque bayesiano para estimar los parámetros de dos modelos matemáticos que describen la transmisión de la enfermedad del dengue y comparar sus resultados utilizando datos de humanos infectados por dengue en la ciudad de Cali, Colombia. Se estudia el modelo clásico *Ross – Macdonald* para el caso del dengue y el modelo *Vector – Hospedero*, el cuál fue desarrollado para enfermedades con inmunidad permanente transmitida por vectores. Como estimaciones, para los parámetros de interés se tienen la mediana y los intervalos creíbles, así como la estimación por intervalo del número reproductivo básico. Se utiliza la distribución normal como función de verosimilitud y en la simulación Monte Carlos no se fijan valores iniciales para los parámetros, además se hace uso de distribuciones a priori no informativas. Se concluye que las estimaciones obtenidas de cada parámetro son muy semejantes para ambos modelos cuando se utilizan datos de casos de humanos infectados por dengue, y la metodología bayesiana es una alternativa eficiente para estimar los parámetros y el número reproductivo básico de los dos modelos. Finalmente, no existe evidencia en favor del modelo *Vector – Hospedero*, pero este brinda más información biológica sobre la dinámica de la transmisión de la enfermedad del dengue.

Abstract

Dengue is a viral disease transmitted by the female mosquito of the *Aedes aegypti*. This disease is distributed in countries of Africa, Asia, Oceania and America, where geographic, demographic and epidemiological conditions are favorable for: the vector, the virus and the host. For the study of the virus dynamics, the mechanisms involved in the transmission of virus between hosts and vectors, mathematical models have been proposed in the language of ordinary differential equations. To estimate the parameters of these models, the least squares method has been used in a number of situations. The goal of this work is to use the Bayesian approach to estimate the parameters of two mathematical models that describe the transmission of dengue disease and compare their results using data on the number of cases of human illness from dengue virus (DV) in the Cali City, Colombia. The classic *Ross – Macdonald* model is studied for the case of dengue and the *Vector – Host* model, which was developed for diseases with permanent immunity transmitted by vectors. As estimates, for the parameters of interest we have the median and the credible intervals, as well as the estimation by interval of the basic reproductive number. The normal distribution is used as a likelihood function and in the Monte Carlos simulation, no initial values are set for the parameters, and non-informative a priori distributions are used. It is concluded that the estimates obtained for each parameter are very similar for both models when using data on the number of cases of human illness from DV, and the Bayesian methodology is an efficient alternative to estimate the parameters and the basic reproductive number of the two models. Finally, although there is no evidence in favor of the *Vector – Host* model, but it provides more biological information on the dynamics of the transmission of dengue disease.

Índice general

Índice de Figuras	VIII
Índice de Cuadros	x
1. Introducción	1
1.1. Modelos Matemáticos	3
1.2. Antecedentes	4
1.3. Planteamiento del Problema	5
1.4. Objetivos	7
2. Marco Teórico	8
2.1. Modelo de <i>Ross – Macdonald</i>	8
2.1.1. Análisis cualitativo del modelo	9
2.2. Modelo <i>Vector – Hospedero</i>	12
2.2.1. Análisis cualitativo del modelo	14
2.3. Solución Numérica de Ecuaciones Diferenciales Ordinarias	17
2.4. Estimación de Parámetros	20
2.4.1. Enfoque Bayesiano	20
3. Metodología	29
3.1. Datos	29
3.2. Estimación Bayesiana	30
3.2.1. Modelo Estadístico	30
3.2.2. Función de Verosimilitud	30
3.2.3. Información a priori	31

3.2.4. Distribuciones a posteriori	32
4. Resultados y Discusión	34
4.1. Estimación bayesiana del modelo <i>Ross – Macdonald</i>	34
4.2. Diagnósticos de convergencia	39
4.3. Estimación de los parámetros.	40
4.4. Comparación de Modelos	44
4.5. Discusión	46
5. Conclusiones	48
Referencias	50

Índice de figuras

2.1. Diagrama compartamental del modelo <i>Ross – Macdonald</i>	9
2.2. Diagrama compartamental para el modelo <i>Vector – Hospedero</i>	13
3.1. Incidencia de dengue ocurrida en la ciudad de Cali, Colombia en el año 2010.	29
4.1. Ajuste del modelo bayesiano a la solución numérica del modelo <i>Ross – Macdonald</i>	38
4.2. Ajuste del modelo bayesiano a la solución numérica del modelo <i>Vector – Hospedero</i>	43
4.3. Ajuste de los modelos <i>Ross – Macdonald</i> y <i>Vector – Hospedero</i> , con las estimaciones obtenidas, a los casos semanales de humanos infectados de dengue.	45
5.1. Mezclado de las cadenas para cada parámetro del modelo <i>Ross – Macdonald</i> .	53
5.2. Distribución a posteriori marginal de cada cadena para los parámetros del modelo de <i>Ross – Macdonald</i>	54
5.3. Diagnóstico de Gelman -Rubin para los parámetros del modelo <i>Ross – Macdonald</i>	55
5.4. Autocorrelación para cada cadena de los parámetros del modelo <i>Ross – Macdonald</i> . Parte I.	56
5.5. Autocorrelación para cada cadena de los parámetros del modelo <i>Ross – Macdonald</i> . Parte II.	57
5.6. Autocorrelación para cada cadena de los parámetros del modelo <i>Ross – Macdonald</i> . Parte III.	58
5.7. Diagnóstico de convergencia de Geweke para cada cadena de los parámetros del modelo de <i>Ross – Macdonald</i> . Parte I.	59

5.8. Diagnóstico de convergencia de Geweke para cada cadena de los parámetros del modelo de <i>Ross – Macdonald</i> . Parte II.	60
5.9. Diagnóstico de convergencia de Geweke para cada cadena de los parámetros del modelo de <i>Ross – Macdonald</i> . Parte III.	61
5.10. Mezcla de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i> . . .	62
5.11. Distribución a posteriori de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i>	63
5.12. Diagnóstico de Gelman - Rubin de los parámetros del modelo <i>Vector – Hospedero</i>	64
5.13. Autocorrelación de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i> . Parte I.	65
5.14. Autocorrelación de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i> . Parte II.	66
5.15. Autocorrelación de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i> . Parte III.	67
5.16. Diagnóstico de Geweke de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i> . Parte I.	68
5.17. Diagnóstico de Geweke de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i>	69
5.18. Diagnóstico de Geweke de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i> . Parte II.	69
5.19. Diagnóstico de Geweke de las cadenas de los parámetros del modelo <i>Vector – Hospedero</i> . Parte III.	70

Índice de Cuadros

2.1. Escala según Jeffreys para el Factor Bayes.	27
3.1. Distribuciones a priori.	32
4.1. Resultados del diagnóstico de convergencia Gelman - Rubin para las estimaciones para el Sistema Dinámico(2.1)	35
4.2. Resultados del diagnóstico de convergencia de Geweke.	35
4.3. Estimación bayesiana de los parámetros de interés del modelo <i>Ross - Macdonald</i>	37
4.4. Resultados del diagnóstico de convergencia de Gelman - Rubin para las estimaciones para el Sistema dinámico (2.3).	39
4.5. Resultados del diagnóstico de convergencia de Geweke.	40
4.6. Estimación bayesiana de los parámetros de interés del modelo <i>Vector - Hospedero</i>	41
4.7. Comparación entre los resultados del Criterio de Devianza y Factor Bayes para los modelos <i>Ross - Macdonald</i> y <i>Vector - Hospedero</i>	46
4.8. Comparación de las estimaciones obtenidas con la de otras investigaciones.	47

Introducción

El dengue es una enfermedad viral transmitida por mosquitos hembras de la especie *Aedes aegypti*, estos se infectan al momento de alimentarse de la sangre de un humano infectado durante el período infeccioso de la enfermedad que dura entre 4 y 5 días. Un mosquito luego de que se infecta tiene un período mínimo de incubación extrínseca del virus de 8 a 14 días antes de que sea infeccioso. Cuando un mosquito en éste último estado pica a un humano susceptible, el virus es inyectado a su torrente sanguíneo y comienza un período de incubación intrínseco en el hospedero que toma de 5 a 7 días, con un promedio más cercano a los cinco días. Luego de este tiempo comienza el período infeccioso del hospedero comenzando de nuevo el proceso infeccioso.

Se conocen cuatro serotipos del virus del dengue: DEN-1, DEN-2, DEN-3 y DEN-4. Cuando una persona se recupera de la infección adquiere inmunidad permanente contra el serotipo que se infectó. La inmunidad cruzada a los otros serotipos es parcial y temporal, y las infecciones causadas por otros serotipos aumentan el riesgo de padecer el dengue grave¹. Se debe sospechar que una persona padece dengue cuando una fiebre elevada (40°C) se acompaña de dos de los síntomas siguientes: dolor de cabeza muy intenso, dolor detrás de los globos oculares, dolores musculares y articulares, náuseas, vómitos, agrandamiento de ganglios linfáticos o salpullido [1].

¹Es una etapa de la enfermedad del dengue potencialmente mortal pues cursa con extravasación de plasma, acumulación de líquidos, dificultad respiratoria, hemorragias graves o falla orgánica.

Las primeras epidemias² de dengue reportadas datan de 1779 - 1780 en Asia, África y América del Norte. Hoy en día, la enfermedad es endémica³ en más de 100 países de las regiones de África, América, el Mediterráneo Oriental, Asia Sudoriental y el Pacífico Occidental. De estas regiones, las más gravemente afectadas son: América, Asia Sudoriental y el Pacífico Occidental [2].

En el Continente Americano, el dengue data desde hace más de 200 años. Entre los años 1960 y 1970 se llevó a cabo una campaña de erradicación del mosquito *Aedes* con el objetivo de la interrupción de la transmisión del dengue. Sin embargo, no se mantuvieron las medidas de vigilancia y control del vector, y hubo reinfestaciones subsiguientes del mosquito, seguidas de brotes epidémicos en el Caribe, en América Central y América del Sur. Desde entonces el dengue se ha propagado con brotes cíclicos que ocurren cada 3 a 5 años [3].

En Colombia el dengue ha sido endémico y ocupa uno de los primeros lugares en la lista de países de América del Sur con mayor reporte de casos. Hasta el 11 de marzo del 2010, las Secretarías Departamentales de Salud habían informado al Instituto Nacional de Salud (INS) un total de 28280 casos de dengue, de los cuales 2624 (9 %) correspondieron a dengue grave o hemorrágico. En el municipio de Santiago de Cali, se presentó una tendencia de disminución de casos entre los años 2005 a 2009, siendo el 2008 un año sin mortalidad por dengue, según informes de la Secretaría Municipal de Salud. A partir del 2010, se han incrementado los casos afectando principalmente a la población infantil y adulto joven, generando mayor severidad en los mismos [4].

Hasta este año 2019, el método más eficaz para controlar o prevenir la transmisión del virus del dengue consiste en controlar la población de los vectores [1]:

- Evitar que los mosquitos encuentren lugares donde depositar sus huevos con el ordenamiento y la modificación del medio ambiente.
- Eliminar correctamente los desechos sólidos y los posibles hábitats artificiales.
- Cubrir, vaciar, y limpiar cada semana los recipientes donde se almacena agua para uso doméstico.

²Enfermedad que ataca a un gran número de personas o de animales en un mismo lugar y durante un mismo período de tiempo.

³[Enfermedad] que afecta habitualmente a una región o país o en cierta época determinada.

- Aplicar insecticidas adecuados a los recipientes en que se almacena agua a la intemperie.
- Utilizar protección personal en el hogar, como mosquiteros en las ventanas, usar ropa de manga larga, materiales tratados con insecticidas, espirales y vaporizadores.
- Durante los brotes epidémicos, las medidas de lucha antivectorial de emergencia pueden incluir la aplicación de insecticidas mediante el rociamiento.
- Vigilar activamente los vectores para determinar la eficacia de las medidas de control.

Una manera de estudiar esta enfermedad es con el desarrollo de modelos matemáticos y estadísticos, ya que describen la dinámica viral, los mecanismos implicados en la transmisión del patógeno entre los hospederos y los vectores, los factores que influyen en la incidencia del dengue, entre otros. Además la información derivada de ellos, como el número reproductivo básico, el tamaño final de la epidemia, la incidencia y la fuerza de infección, han demostrado ser una herramienta para la vigilancia y control de enfermedades infecciosas, especialmente en situaciones epidémicas.

1.1. Modelos Matemáticos

Desde hace muchos años la relación entre las matemáticas y ciencias como la biología y la medicina ha sido beneficiosa para ambas partes; para la matemática, porque permite resolver problemas que surgen en biología, dando posibilidad al desarrollo de nuevas ideas, aplicaciones y teorías; para los biólogos o médicos porque la modelación matemática es una herramienta de gran utilidad de experimentación o verificación de resultados.

- El primer modelo matemático para una enfermedad lo propuso Daniel Bernoulli aproximadamente en el año 1760, cuando trató de fundamentar la utilidad de la vacunación para la viruela.
- Para el caso de las epidemias, uno de los pioneros en utilizar modelos matemáticos para describir su comportamiento fue el médico escocés Sir Ronald Ross. En el año 1912, con el objetivo de modelar la transmisión del paludismo. Este fue el primer acercamiento entre los modelos matemáticos y los fenómenos epidemiológicos. A pesar de los importantes resultados obtenidos, el modelo propuesto por Ross no fue tomado

en cuenta sino hasta que George Macdonald retomó su trabajo; y en 1956 propuso una extensión al modelo de Ross, conocido en la literatura como el modelo de *Ross – Macdonald*.

- En los años 1927, 1932 y 1933, Kermack y McKendrick publicaron una serie de tres artículos sobre la teoría matemática de las epidemias. En 1927 propusieron el modelo actualmente conocido como SIR (Susceptible - Infectado - Recuperado). De acuerdo al trabajo de Ross para la malaria, se supuso que durante un intervalo de tiempo infinitesimal “ dt ”, la probabilidad de que ocurra una nueva infección es proporcional a la cantidad de susceptibles y a la cantidad de infectado; y la probabilidad de una nueva recuperación es proporcional a los infectados [5].
- En 1998, Esteva y Vargas propusieron un modelo *Vector – Hospedero* para la transmisión de la fiebre del dengue en una población humana constante y una población de vectores variables. En este modelo las personas pasan por los estados: susceptibles, infectados y recuperados (SIR), mientras que los vectores pasan por los estados susceptibles e infectados (SI) [6].

1.2. Antecedentes

Existen varias investigaciones que están relacionadas con la transmisión de la enfermedad del dengue. Aldila et al. [7], planteó un problema de control óptimo para la transmisión de la enfermedad del dengue en un modelo *Vector – Hospedero*, en el que toma en cuenta un tratamiento con repelente de mosquito para las personas adultas y niños. Los resultados numéricos obtenidos mediante el método del gradiente, para los controles óptimos y las dinámicas relacionadas se mostraron para el caso de las estrategias de prevención de epidemias y reducción de brotes. Buonomo et al. [8], realizaron un estudio sobre la eficacia de utilizar, por parte de los individuos, mosquiteros tratados con insecticidas combinados o no con la fumigación, para evaluar su efectividad en la reducción de la transmisión de la fiebre del dengue. Para dicho estudio consideraron un modelo matemático *Vector – Hospedero* para la transmisión del dengue. Se estableció un problema de control óptimo para minimizar el número de infecciones humanas y el costo de los esfuerzos realizados en la adopción y mantenimiento del mosquitero y la fumigación con insecticidas. Arias et al. [9], presentan en su trabajo el

ajuste de dos modelos matemáticos, para el cual tuvieron en cuenta características del área urbana de la ciudad de Cali, Colombia. Con base en el comportamiento natural del mosquito *Aedes aegypti* y con el uso del modelo presa - depredador, se estiman algunos de los parámetros teniendo en cuenta la literatura existente sobre el tema. Posteriormente, estimaron los parámetros del modelo *Ross - Macdonald* para el caso del dengue, a través de la solución de mínimos cuadrados para los casos reportados de dengue durante el año 2010. Pandey et al. [10], utilizaron la estimación bayesiana para ajustar los modelos *Vector - Hospedero* y SIR a los datos de la fiebre hemorrágica del dengue, y así estudiar el efecto de los supuestos del modelado de la dinámica del dengue en Tailandia. Utilizaron el Criterio de Información de Akaike para comparar ambos modelos y concluyeron que para estos datos el modelo SIR resultó ser el mejor. Khan et al. [11], realizaron un análisis retrospectivo de la epidemia de fiebre del dengue en Pakistán para evaluar la transmisibilidad de la enfermedad. Obtuvieron estimaciones del número reproductivo básico (R_0) a partir de datos de epidemias; para evaluar la solidez de la estimación utilizaron distintas metodologías aplicadas a diferentes modelos de epidemia. Estimaron los parámetros de un modelo *Vector - Hospedero* determinista, para la dinámica de transmisión del dengue de cepa única a los datos de epidemia, utilizando el método de mínimos cuadrados ordinarios (OLS) y mínimos cuadrados generalizado (GLS). También formularon un modelo estocástico de transmisión directa para la dinámica de transmisión del dengue y se obtuvieron estimaciones de los parámetros del modelo usando inferencia bayesiana a través de los métodos MCMC.

1.3. Planteamiento del Problema

Con frecuencia se busca relacionar los parámetros que caracterizan al modelo matemático con las observaciones dadas, y . Suponiendo que la dinámica que se busca explicar se captura adecuadamente con el modelo matemático, entonces los parámetros, θ , pueden relacionarse con los datos a partir de:

$$y = G(\theta) \tag{1.1}$$

En la práctica, y puede ser una función que depende del tiempo y/o espacio, o simplemente una colección de observaciones discretas. Es importante tener en cuenta que las observaciones

dadas tienen cierta cantidad de ruido, ϵ , debido a los instrumentos de medición y por el redondeo numérico. Por lo tanto:

$$\begin{aligned} y &= G(\theta_{true}) + \epsilon \\ &= y_{true} + \epsilon \end{aligned} \tag{1.2}$$

donde y_{true} satisface la Ecuación (1.1) para θ igual al verdadero, θ_{true} . Cuando θ y y son funciones, generalmente nos referimos a G como operador, el cuál es en algunos casos una ecuación diferencial ordinaria (EDO) o una ecuación en derivadas parciales (PDE). En otros casos, G es un sistema lineal o no lineal de ecuaciones algebraicas.

El **problema directo** (forward problem or forward map) es definido como la solución del sistema de ODEs (or PDEs, etc), es decir, encontrar y dado los parámetros θ ; sin embargo, el enfoque que se utiliza en esta investigación es el **problema inverso** (inverse problem). Para definir este problema hay que definir primero el problema directo y un método de solución, verificar que el modelo cumpla las condiciones de regularidad; y posteriormente, asumir distribuciones a priori para obtener las distribuciones a posteriori marginales de los parámetros. Es decir, la solución del problema inverso es la distribución posterior de las cantidades de interés.

En el presente trabajo se aborda la estimación de parámetros haciendo uso de la metodología bayesiana, en dos sistemas dinámicos no lineales que describen la transmisión de la enfermedad del dengue.

En epidemiología el número reproductivo básico, R_0 , es una medida útil, que depende de los parámetros del modelo, para entender la propagación de la enfermedad, pues a partir del valor umbral de la unidad se puede indicar si un brote seguirá propagándose o no. En la literatura es común utilizar métodos de optimización para estimar la estimación de los parámetros que están incluidos en el R_0 . Desde la perspectiva bayesiana se puede aprovechar el conocimiento a priori o empírico que se tiene o sabe de los parámetros para su estimación así como obtener el comportamiento del R_0 a través del uso de las cadenas MCMC de los parámetros involucrados.

1.4. Objetivos

Objetivo General

Utilizar el enfoque bayesiano para estimar los parámetros en dos sistemas dinámicos que describen la transmisión de la enfermedad del dengue.

Objetivos específicos

- Estimar el valor de los parámetros con sus respectivos intervalos creíbles y el número reproductivo básico del modelo de *Ross – Macdonald*.
- Estimar el valor de los parámetros con sus respectivos intervalos creíbles y el número reproductivo básico del modelo *Vector – Hospedero*.
- Seleccionar el mejor modelo que ajuste a un conjunto de datos dados.

Para dar cumplimiento a estos objetivos en el Capítulo 2 del presente trabajo se muestran: los marcos teóricos de los sistemas dinámicos utilizados, *Ross – Macdonald* y *Vector – Hospedero*, así como su análisis de estabilidad; el método de solución numérica utilizado para dar solución a dichos sistemas; la teoría de la inferencia bayesiana utilizada para encontrar la estimación de los parámetros de los modelos que describen los sistemas dinámicos y criterios de comparación de modelos para poder identificar cuál de los modelos utilizados se ajusta mejor a los datos utilizados en la presente investigación. En el Capítulo 3 se muestra la metodología utilizada para resolver el problema que se planteó y en el Capítulo 4 se muestran todos los resultados obtenidos al utilizar dicha metodología. En el Capítulo 5 se muestran las conclusiones que se obtuvieron al desarrollar este trabajo.

Marco Teórico

2.1. Modelo de *Ross – Macdonald*

Es un modelo compartamental basado en ecuaciones diferenciales ordinarias, que describe las interacciones entre las poblaciones de mosquitos y humanos. Cada población está dividida en dos compartimentos: susceptibles e infectados.

El modelo *Ross – Macdonald* se puede explicar de la siguiente forma: los humanos susceptibles se infectan con una fuerza $\frac{b\beta_h I_v}{N_h}$, pasando ser de humanos susceptibles a infectados, estos se recuperan de la enfermedad a una tasa γ . Estos humanos infectados al interactuar con los mosquitos susceptibles provocan que estos se infecten a una fuerza $\frac{b\beta_v I_h}{N_h}$, convirtiéndose así en mosquitos infectados. A su vez estos mosquitos infectados tienen una mortalidad natural μ_v y al interactuar con los humanos susceptibles provocarían que estos se conviertan en humanos infectados. Estas interacciones se observan en la Figura 2.1.

Este modelo se representa a través del siguiente sistema de ecuaciones diferenciales ordinarias:

$$\frac{dI_h}{dt} = \frac{b\beta_h I_v}{N_h} (N_h - I_h) - \gamma I_h \tag{2.1}$$

$$\frac{dI_v}{dt} = \frac{b\beta_v I_h}{N_h} (N_v - I_v) - \mu_v I_v$$

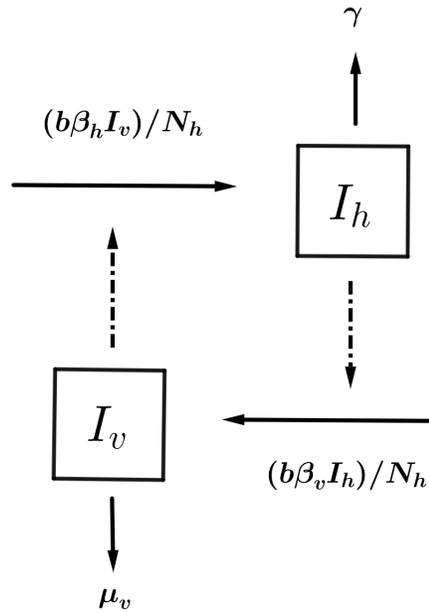


Figura 2.1: Diagrama compartamental del modelo *Ross - Macdonald*.

Las variables del modelo son I_h , el número de humanos infectados e I_v , el número de vectores infectados. Los parámetros del modelo son: tasa semanal de picadura del vector en los humanos, b , probabilidad de que un humano susceptible se contagie al ser picado por un vector infectado, β_h , probabilidad de que un vector susceptible se contagie al picar a un humano infectado, β_v , tasa semanal de recuperación de los humanos, γ , mortalidad natural de los vectores, μ_v , población total de los humanos, N_h y la población total de los vectores, N_v [4].

2.1.1. Análisis cualitativo del modelo

A continuación se determinan los puntos de equilibrio del Modelo (2.1) y posteriormente se estudiará la estabilidad local de cada uno.

- Puntos de Equilibrio

Para encontrar los puntos de equilibrio del Modelo (2.1) se iguala a cero el sistema de ecuaciones diferenciales. Al realizar este proceso se obtienen dos puntos, el punto que

representa la ausencia de la enfermedad, conocido como punto libre de enfermedad, $E_1 = [0; 0]$, y el punto endémico que representa el estado donde la epidemia progresa:

$$E_2 = \left[\frac{N_h (b^2 \beta_h \beta_v N_v - \mu_v N_h \gamma)}{b \beta_v (b \beta_h N_v + \gamma N_h)}; \frac{N_h (b^2 \beta_h \beta_v N_v - \mu_v N_h \gamma)}{b \beta_h (b \beta_v N_h + \mu_v N_h)} \right]$$

Para este modelo el número reproductivo básico está dado por:

$$R_0 = \frac{b^2 \beta_h \beta_v N_v}{N_h \mu_v \gamma}$$

Usualmente se escribe E_2 en términos del R_0 :

$$E_2 = [I_h^*; I_v^*]$$

donde

$$I_h^* = \frac{N_h \mu_v (R_0 - 1)}{b \beta_v + \mu_v R_0}$$

$$I_v^* = \frac{\gamma \mu_v N_h (R_0 - 1)}{\beta_h (\mu_v b + b^2 \beta_v)}$$

■ Estabilidad Local

Para analizar la estabilidad local de los puntos de equilibrio de (2.1), se utilizan los valores propios de la matriz jacobiana, que en este caso está dada por:

$$J = \begin{pmatrix} -\frac{b \beta_h I_v}{N_h} - \gamma & \frac{b \beta_h}{N_h} (N_h - I_h) \\ \frac{b \beta_v}{N_h} (N_v - I_v) & -\frac{b \beta_v I_h}{N_h} - \mu_v \end{pmatrix}$$

Evaluando E_1 en la matriz jacobiana se obtiene:

$$J(E_1) = \begin{pmatrix} -\gamma & b\beta_h \\ \frac{b\beta_v N_v}{N_h} & -\mu_v \end{pmatrix}$$

Luego el polinomio característico para E_1 es:

$$p_1(\lambda) = |J_{Libre} - \lambda I| = \begin{vmatrix} -\gamma - \lambda & b\beta_h \\ \frac{b\beta_v N_v}{N_h} & -\mu_v - \lambda \end{vmatrix}$$

$$p_1(\lambda) = \lambda^2 + \lambda(\gamma + \mu_v) + \gamma\mu_v - \frac{b^2\beta_h\beta_v N_v}{N_h}$$

Sea $A = \gamma + \mu_v$ y $B = \gamma\mu_v - \frac{b^2\beta_h\beta_v N_v}{N_h}$ si, $A, B > 0$, entonces

$$\gamma\mu_v > \frac{b^2\beta_h\beta_v N_v}{N_h}$$

luego

$$\frac{b^2\beta_h\beta_v N_v}{N_h\gamma\mu_v} < 1$$

de donde

$$R_0 < 1$$

Por el criterio de Routh - Hurwitz, si $R_0 < 1$ entonces E_1 es asintóticamente estable y si $R_0 > 1$, entonces E_1 es inestable [12].

Por otra parte, se sustituye el punto de equilibrio endémico E_2 en las expresiones

$$N_h - I_h^* = \frac{N_h\gamma I_h^*}{b\beta_h I_v^*}$$

$$N_v - I_v^* = \frac{N_h\mu_v I_v^*}{b\beta_v I_h^*}$$

deducidas al igualar (2.1) a cero. Ahora se sustituyen las dos expresiones anteriores en la matriz Jacobiana se tiene:

$$J(E_2) = \begin{pmatrix} \frac{-b\beta_h}{N_h} I_v^* - \gamma & \frac{\gamma I_h^*}{I_v^*} \\ \frac{\mu_v I_v^*}{I_h^*} & -\frac{b\beta_v}{N_h} - \mu_v \end{pmatrix}$$

El polinomio característico para E_2 es:

$$p_2(\lambda) = |J(E_2) - \lambda I| = \begin{vmatrix} \frac{-b\beta_h}{N_h} I_v^* - \gamma - \lambda & \frac{\gamma I_h^*}{I_v^*} \\ \frac{\mu_v I_v^*}{I_h^*} & -\frac{b\beta_v}{N_h} - \mu_v - \lambda \end{vmatrix}$$

Luego

$$p_2(\lambda) = \lambda^2 + \left(\frac{b\beta_v}{N_h} + \mu_v + \frac{b\beta_h}{N_h} I_v^* + \gamma \right) \lambda + \left(\frac{b\beta_h}{N_h} I_v^* + \gamma \right) \left(\frac{b\beta_v}{N_h} I_h^* + \mu_v \right) - \gamma \mu_v$$

Sea

$$A = \frac{b\beta_v}{N_h} + \mu_v + \frac{b\beta_h}{N_h} I_v^* + \gamma$$

y

$$B = \left(\frac{b\beta_h}{N_h} I_v^* + \gamma \right) \left(\frac{b\beta_v}{N_h} I_h^* + \mu_v \right) - \gamma \mu_v$$

Al sustituir I_v^* e I_h^* en las expresiones de A y B se obtiene que si $R_0 > 1$ entonces las expresiones $A > 0$ y $B > 0$, por lo tanto tienen dos valores propios con parte real negativa, por el criterio de Routh - Hurwitz el polinomio característico $p_2(\lambda)$ tiene dos valores propios con parte real negativa. En consecuencia E_2 es un punto local asintóticamente estable [12].

2.2. Modelo *Vector – Hospedero*

El modelo *Vector – Hospedero* para el caso del dengue consta de tres compartimentos para los hospederos (humanos), susceptibles (S_h), infectados (I_h) y recuperados (R_h), y dos compartimentos para los vectores (mosquitos), susceptibles (S_v) e infectados (I_v) [6].

La población total de humanos está dada por N_h , los cuales tiene una mortalidad natural μ_h perteneciendo entonces a la población de humanos susceptibles a la enfermedad del dengue. Estos humanos susceptibles también poseen una mortalidad natural μ_h y al ser picados por

los mosquitos infectados se infectan con una fuerza $\frac{\beta_h b I_v}{N_h + m}$. Una vez que estos humanos se infectan, también poseen una mortalidad natural μ_h y se recuperarán de la enfermedad a una tasa γ , pasan a formar parte ahora de la población de humanos recuperados y poseen además una mortalidad natural de μ_h . Por parte de la población de mosquitos; existe una tasa de reclutamiento, A , por semana, los cuales pasan a ser la población de mosquitos susceptibles. Estos tienen una mortalidad natural de μ_v y al picar a los humanos infectados entonces se infectan con una fuerza $\frac{\beta_v b I_h}{N_h + m}$. Una vez que estos mosquitos se infectan, tienen también una mortalidad μ_v y picarán nuevamente a los humanos susceptibles. Estas interacciones se observan en la Figura 2.2.

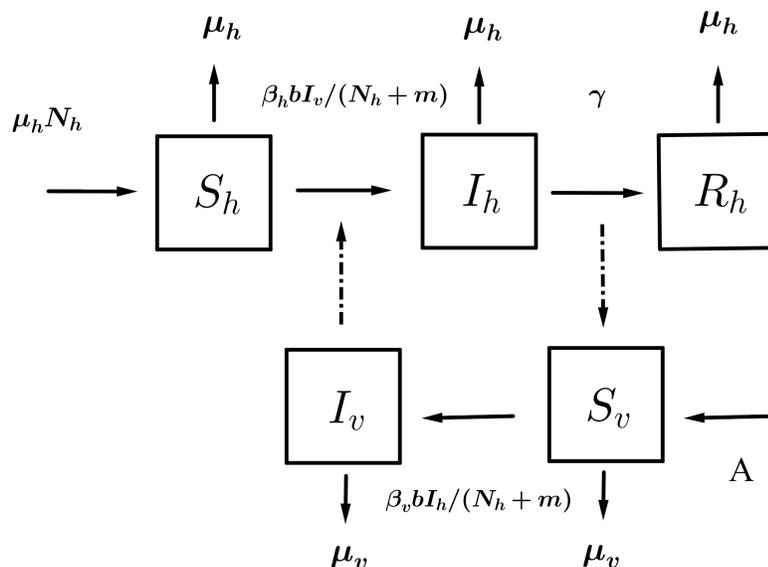


Figura 2.2: Diagrama compartamental para el modelo *Vector – Hospedero*.

Este modelo se representa a través del siguiente sistema de ecuaciones diferenciales ordinarias:

$$\begin{aligned}
\frac{dS_h}{dt} &= \mu_h N_h - \frac{\beta_h b}{N_h+m} S_h I_h - \mu_h S_h \\
\frac{dI_h}{dt} &= \frac{\beta_h b}{N_h+m} S_h I_h - (\mu_h + \gamma) I_h \\
\frac{dR_h}{dt} &= \gamma I_h - \mu_h R_h \\
\frac{dS_v}{dt} &= A - \frac{\beta_v b}{N_h+m} S_v I_h - \mu_v S_v \\
\frac{dI_v}{dt} &= \frac{\beta_v b}{N_h+m} S_v I_h - \mu_v I_v
\end{aligned} \tag{2.2}$$

Se asume que $S_h + I_h + R_h = N_h$ y $S_v + I_v = \frac{A}{\mu_v} = N_v$ son los tamaños de las poblaciones de humanos y vectores, respectivamente.

Sea $S_v = \frac{A}{\mu_v} - I_v$ y al sustituir en la última ecuación del Modelo (2.2), entonces la tercera y cuarta ecuación quedan desacopladas del Modelo (2.2), por lo que ya no son necesarias en el modelo *Vector - Hospedero*. Por lo que el sistema queda planteado de la siguiente forma:

$$\begin{aligned}
\frac{dS_h}{dt} &= \mu_h (N_h - S_h) - b\beta_h \frac{S_h I_v}{N_h+m} \\
\frac{dI_h}{dt} &= b\beta_h \frac{S_h I_v}{N_h+m} - (\gamma + \mu_h) I_h \\
\frac{dI_v}{dt} &= \frac{b\beta_v}{N_h+m} \left(\frac{A}{\mu_v} - I_v \right) I_h - \mu_v I_v
\end{aligned} \tag{2.3}$$

2.2.1. Análisis cualitativo del modelo

A continuación se determinan los puntos de equilibrio del Modelo (2.3) y posteriormente se realiza el análisis de su estabilidad local.

- Puntos de equilibrio del modelo.

Para obtener los puntos de equilibrio, igualamos todas las ecuaciones del Modelo (2.3) a cero.

$$\mu_h N_h - \frac{\beta_h b}{N_h + m} S_h I_v - \mu_h S_h = 0 \quad (2.4)$$

$$\frac{\beta_h b}{N_h + m} S_h I_v - \mu_h I_h - \gamma I_h = 0 \quad (2.5)$$

$$\frac{\beta_v b}{N_h + m} (N_v - I_v) I_h - \mu_v I_v = 0 \quad (2.6)$$

De la ecuación (2.4) y (2.6) se obtiene que:

$$S_h = \frac{N_h(BI_h + N_h)}{(B + R_0M)I_h + N_h} \quad (2.7)$$

$$I_v = \frac{BN_v I_h}{BI_h + N_h} \quad (2.8)$$

donde

$$B = \frac{b\beta_v N_h}{\mu_v(N_h + m)}, \quad M = \frac{\gamma + \mu_h}{\mu_h}, \quad R_0 = \frac{b^2\beta_h\beta_v N_h N_v}{(N_h + m)^2 \mu_v(\gamma + \mu_h)}$$

Al sustituir las Ecuaciones (2.7), (2.8) y $\frac{Bb\beta_h N_v}{N_h + m} = R_0(\gamma + \mu_h)$, en (2.5) se tiene que I_h se obtiene al resolver la siguiente ecuación cuadrática para I_h :

$$(B + R_0M)I_h^2 + (N_h - R_0N_h)I_h = 0 \quad (2.9)$$

de donde $I_h = 0$ e $I_h = \frac{N_h(R_0 - 1)}{B + R_0M}$, ahora al sustituir estos valores en las Ecuaciones (2.7) y (2.8) se obtienen los puntos de equilibrio del Modelo (2.3):

$$E_1 = (N_h, 0, 0) \text{ y } E_2 = (S_h^*, I_h^*, I_v^*)$$

donde

$$S_h^* = \frac{N_h(B + M)}{B + R_0M}, \quad I_h^* = \frac{N_h(R_0 - 1)}{B + R_0M}, \quad I_v^* = \frac{BN_v(R_0 - 1)}{R_0(B + M)}$$

E_1 es el punto libre de enfermedad y E_2 es el punto de equilibrio endémico.

- Estabilidad de los puntos de equilibrio.

La matriz jacobiana para este modelo tiene la siguiente forma:

$$J = \begin{pmatrix} -\left(\frac{\beta_h b}{N_h+m} I_v + \mu_h\right) & 0 & -\frac{\beta_h b}{N_h+m} S_h \\ \frac{\beta_h b}{N_h+m} I_v & -(\mu_h + \gamma) & \frac{\beta_h b}{N_h+m} S_h \\ 0 & \frac{\beta_v b}{N_h+m} (N_v - I_v) & -\left(\frac{\beta_v b}{N_h+m} I_h + \mu_v\right) \end{pmatrix}$$

Al evaluar esta matriz en el punto de equilibrio libre de enfermedad obtenemos:

$$J(E_1) = \begin{pmatrix} -\mu_h & 0 & -\frac{\beta_h b N_h}{N_h+m} \\ 0 & -(\mu_h + \gamma) & \frac{\beta_h b N_h}{N_h+m} \\ 0 & \frac{\beta_v b N_v}{N_h+m} & -\mu_v \end{pmatrix}$$

y los eigenvalores para este caso son:

$$-\mu_h; \frac{-(\mu_h + \gamma + \mu_v) \pm \sqrt{(\mu_h + \gamma + \mu_v)^2 - 4\mu_v(\gamma + \mu_h)(1 - R_0)}}{2}$$

Si $R_0 < 1$, E_1 es local asintóticamente estable dado que todos los eigenvalores son negativos [12]. Si $R_0 > 1$, E_1 es inestable dado que no todos los eigenvalores son negativos.

Ahora se evalúa E_2 en J obtenemos:

$$J(E_2) = \begin{pmatrix} -\mu_h \left(\frac{B+MR_0}{B+M}\right) & 0 & -\frac{\mu_h MR_0}{B} \frac{N_h}{N_v} \left(\frac{B+M}{B+R_0M}\right) \\ \frac{\mu_h M(R_0-1)}{B+M} & -\mu_h M & \frac{\mu_h MR_0}{B} \frac{N_h}{N_v} \left(\frac{B+M}{B+R_0M}\right) \\ 0 & \frac{\mu_v BN_v}{R_0 N_h} \left(\frac{B+R_0M}{B+M}\right) & -\mu_v R_0 \left(\frac{B+M}{B+R_0M}\right) \end{pmatrix}$$

El polinomio característico para este caso queda la siguiente forma:

$$p(\lambda) = \lambda^3 + A\lambda^2 + B\lambda + C \quad (2.10)$$

donde

$$A = -tr(J(E_2))$$

$$B = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

$$C = - | J(E_2) |$$

De donde se obtiene que:

$$A = \mu_h \left(\frac{B + MR_0}{B + M} \right) + \mu_h M + \frac{\mu_v R_0 (B + M)}{B + MR_0}$$

$$B = \mu_h^2 M \left(\frac{B + MR_0}{B + M} \right) + \mu_h \mu_v R_0 + \frac{\mu_h \mu_v MB (R_0 - 1)}{B + R_0 M}$$

$$C = \mu_h^2 \mu_v M (R_0 - 1)$$

Para $R_0 > 1$ los coeficientes de la Ecuación (2.10) son positivos. Además es fácil verificar también la tercera condición de Routh - Hurwitz:

$$AB > \mu_h^2 \mu_v MR_0 > C$$

Por lo que E_2 es local asintóticamente estable [12].

2.3. Solución Numérica de Ecuaciones Diferenciales Ordinarias

Un problema con valores iniciales es aquél en el que se busca determinar una solución a una ecuación diferencial, sujeta a condiciones sobre la función desconocida y sus derivadas

para un valor de la variable independiente. Este problema se puede representar de la siguiente forma:

$$\begin{aligned}\frac{dy}{dt} &= f(y, t), \quad a < t < b \\ y(\alpha_0) &= \alpha\end{aligned}\tag{2.11}$$

Una generalización de este problema es cuando y y f son vectores de dimensión n , como el siguiente sistema de ecuaciones de primer orden:

$$\begin{aligned}\frac{dy_1}{dt} &= f_1(y_1, y_2, \dots, y_n, t) \\ \frac{dy_2}{dt} &= f_2(y_1, y_2, \dots, y_n, t) \quad a < t < b \\ &\vdots \\ \frac{dy_n}{dt} &= f_n(y_1, \dots, y_n, t)\end{aligned}\tag{2.12}$$

$$y_1(\alpha_0) = \alpha_1, y_2(\alpha_0) = \alpha_2, \dots, y_n(\alpha_0) = \alpha_n$$

Pocas ecuaciones diferenciales tienen una solución analítica, frecuentemente es necesario realizar aproximaciones para estudiar el comportamiento del sistema bajo ciertas condiciones.

Para realizar estas aproximaciones, los métodos de Taylor tienen la propiedad de un error local de truncamiento de orden superior, pero poseen la desventaja de requerir el cálculo y la evaluación de las derivadas de $f_i(y_1, y_2, \dots, y_n, t)$. Esto resulta lento y complicado, en la mayoría de los problemas, razón por la cual, en la práctica casi no se utilizan. El método de Euler, lamentablemente requiere de un paso muy pequeño para una precisión razonable. Los métodos de Runge Kutta tienen el error local de truncamiento del mismo orden que los métodos de Taylor, pero prescinden del cálculo y evaluación de las derivadas de las funciones $f_i(y_1, y_2, \dots, y_n, t)$.

Métodos de Runge Kutta

El procedimiento de Runge-Kutta de orden 4 se puede programar fácilmente en los ordenadores, y además se emplea mucho en la práctica, debido a su exactitud relativamente elevada de la solución aproximada de la Ecuación [13].

Sea el problema con valor inicial dado en (2.11), entonces se elige un tamaño de paso h , y se calculan cuatro números k_1, k_2, k_3, k_4 de acuerdo con el siguiente procedimiento:

$$\frac{dx}{dt} = f(x(t), t)$$

$$k_1 = hf(x(t), t)$$

$$k_2 = hf(x(t) + \frac{1}{2}k_1, t + \frac{1}{2}h)$$

$$k_3 = hf(x(t) + \frac{1}{2}k_2, t + \frac{1}{2}h)$$

$$k_4 = hf(x(t) + k_3, t + h)$$

$$x(t + h) = x(t) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

Según el procedimiento ordinario de Runge - Kutta, a partir del valor de x en el instante t se determina el valor de x en el instante $t + h$.

Para la resolución de sistemas de ecuaciones diferenciales de primer orden, el procedimiento de Runge - Kutta es igualmente efectivo. Supongamos el caso del problema de valores iniciales (2.12), para $n = 2$. El procedimiento es:

$$\frac{dx}{dt} = f(x(t), y(t), t)$$

$$\frac{dy}{dt} = g(x(t), y(t), t)$$

$$k_1 = hf(x(t), y(t), t)$$

$$l_1 = hg(x(t), y(t), t)$$

$$k_2 = hf(x(t) + \frac{1}{2}k_1, y(t) + \frac{1}{2}l_1, t + \frac{1}{2}h)$$

$$l_2 = hg(x(t) + \frac{1}{2}k_1, y(t) + \frac{1}{2}l_1, t + \frac{1}{2}h)$$

$$k_3 = hf(x(t) + \frac{1}{2}k_2, y(t) + \frac{1}{2}l_2, t + \frac{1}{2}h)$$

$$l_3 = hg(x(t) + \frac{1}{2}k_2, y(t) + \frac{1}{2}l_2, t + \frac{1}{2}h)$$

$$k_4 = hf(x(t) + k_3, y(t) + l_3, t + h)$$

$$l_4 = hg(x(t) + k_3, y(t) + l_3, t + h)$$

$$x(t + h) = x(t) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$$

$$y(t + h) = y(t) + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4)$$

Además de los cuatro números k_1, k_2, k_3, k_4 para la primera ecuación diferencial, se precisa de otros cuatro números l_1, l_2, l_3, l_4 para la segunda ecuación diferencial. A partir del valor de x en el instante t , se determina el valor de x en el instante $t + h$, y a partir del valor de y en el instante t se determina el valor de y en el instante $t + h$.

2.4. Estimación de Parámetros

El problema de la estimación de parámetros, consiste en determinar el valor de los parámetros desconocidos $\theta_1, \dots, \theta_k$ de un modelo matemático. En el contexto, donde se dispone de un modelo matemático y datos reales, es muy importante la estimación de los parámetros pues ayuda a analizar el comportamiento del modelo.

2.4.1. Enfoque Bayesiano

En los últimos años la estadística bayesiana se ha vuelto muy popular debido a su capacidad de resolver problemas que con otros métodos sería complicado, además que permite incorporar información que es útil en la solución del problema. Bajo este enfoque, la in-

certidumbre sobre los parámetros poblacionales se explica por medio de distribuciones de probabilidad.

En la aproximación bayesiana se tiene:

- La información que se tiene sobre un parámetro se considera en una distribución de probabilidad llamada *distribución a priori*.
- Los parámetros son considerados variables aleatorias.
- La información a priori puede provenir de:
 - Estudios previos.
 - Información subjetiva de los expertos.

Los bayesianos se basan en la interpretación subjetiva de la probabilidad, es decir, en el grado de creencia que se tiene sobre la ocurrencia de un suceso y tiene como base el Teorema de Bayes.

Para hacer afirmaciones de θ dado y , se debe comenzar con un modelo que proporcione una distribución conjunta para θ y y . Su función de densidad conjunta se puede escribir como un producto de dos densidades que frecuentemente se denomina distribución a priori $p(\theta)$ y la verosimilitud $p(y|\theta)$, respectivamente:

$$p(\theta, y) = p(\theta)p(y|\theta)$$

Usando la propiedad básica de probabilidad condicional conocida como regla de Bayes, se obtiene la densidad a posteriori:

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)}$$

donde $p(y) = \sum_{\theta} p(\theta)p(y|\theta)$, y la suma es sobre todos los valores posibles de θ (si θ es continuo $p(y) = \int p(\theta)p(y|\theta)d\theta$). Dado que $p(y)$ no depende de θ ,

$$p(\theta|y) \propto p(\theta)p(y|\theta)$$

Una vez obtenida la distribución a posteriori, se conoce toda la información de θ y sus características se resumen a través de la estimación bayesiana y sus intervalos creíbles.

El hecho de que la estadística bayesiana considera el conocimiento a priori genera controversia ya que considera la aportación subjetiva del investigador y esto podría llegar a ser

considerada como una manipulación de los datos. Sin embargo, esto no tiene por que ser así ya que una buena justificación de los conocimientos a través de una búsqueda bibliográfica adecuada, pueden garantizar la validez de la información inicial [14].

Cadenas MCMC

Generalmente en el enfoque bayesiano se requiere realizar integraciones sobre distribuciones de probabilidad de alta dimensión. Una solución a este problema es considerar un procedimiento Monte Carlo iterativo o Monte Carlo por Cadenas de Markov. La metodología MCMC es una herramienta de gran alcance para la modelación estadística y es la única manera de hacer inferencia bayesiana en los modelos estadísticos de gran complejidad.

Los métodos MCMC son algoritmos iterativos que se utilizan cuando el muestreo directo de la distribución a posteriori de interés $p(\boldsymbol{\theta}|\text{Datos})$ no es factible. En una cadena de Markov dada $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \dots, \boldsymbol{\theta}^{(T)}$, donde

$$\boldsymbol{\theta}^{(t+1)} \sim p(\boldsymbol{\theta}|\boldsymbol{\theta}^t)$$

p es el kernel de transición de la cadena de Markov. Así $\boldsymbol{\theta}^{t+1}$ depende solo de $\boldsymbol{\theta}^{(t)}$, y no de $\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(1)}, \dots, \boldsymbol{\theta}^{(t-1)}$ [14].

Existen dos principales problemas que rodean la implementación e inferencias con los métodos MCMC: la convergencia y la dependencia entre las muestras de la distribución a posteriori.

Las condiciones necesarias para que una cadena de Markov tenga una única distribución estacionaria son:

- Ser no periódica
- Irreducible
- Positiva recurrente

Si el algoritmo se implementa correctamente, la convergencia de la cadena está garantizado, independientemente de cuales sean los valores iniciales.

Diagnósticos de Convergencia

Una vez que se ejecuta un algoritmo MCMC, es importante examinar si los valores simulados $\theta^{(t)}$ convergen a la distribución estacionaria $p(\theta|y)$. Existen varios diagnósticos útiles para analizar los resultados de una cadena, y dado que ninguno puede garantizar la convergencia de las cadenas, se recomienda que se utilicen varios de estos.

Inspección visual

Una inspección visual de los gráficos de las cadenas es una forma simple de monitorear la convergencia de las cadenas. La observación de las trayectorias de diferentes cadenas partiendo de valores iniciales distintos permite verificar la convergencia, lo cual ocurre si existe una mezcla de las cadenas a medida que aumenta el número de iteraciones. A partir de los gráficos de autocorrelación, se verifica si las muestras generadas pueden ser consideradas independientes de los valores iniciales.

Prueba de Gelman y Rubin

Gelman et al.(1992) [15], proponen un enfoque general para monitorear la convergencia de la salida de MCMC en el que se ejecutan $m > 1$ cadenas paralelas con valores de inicio que están sobredispersos en relación con la distribución posteriori. La convergencia se diagnostica cuando las cadenas han *olvidado* sus valores iniciales y las salidas de todas las cadenas son indistinguibles. Esta prueba se basa en una comparación de las varianzas dentro y entre las cadenas y es similar al análisis de varianza clásico. Hay dos formas de estimar la varianza de una distribución estacionaria: la media de la varianza empírica dentro de la cadena (W), y la varianza empírica de todas las cadenas combinadas, que puede expresarse como:

$$\hat{\sigma}^2 = \frac{(n-1)}{n}W + \frac{1}{n}B$$

donde n es el número de iteraciones y $\frac{B}{n}$ es la varianza empírica entre las cadenas.

Si las cadenas han convergido, entonces ambas estimaciones de la varianza son insesgadas. De otra manera el primer método subestima la varianza, ya que las cadenas individuales no han tenido tiempo de llegar a la distribución estacionaria y el segundo método sobrestima la varianza, ya que los valores iniciales considerados fueron diferentes.

El diagnóstico de convergencia está dado por:

$$R = \sqrt{\frac{(d+3)\hat{V}}{(d+1)W}}$$

donde

$$\hat{V} = \hat{\sigma}^2 + \frac{B}{mn}$$

$$d = 2 \frac{\hat{V}}{Var(\hat{V})}$$

Valores sustancialmente mayores que 1 indican falta de convergencia (en la práctica valores inferior a 1.1 o 1.2 son aceptables) [16].

Prueba de Geweke

Geweke (1992) [17], sugirió un criterio para evaluar la convergencia en base a la comparación de medias, en intervalos diferentes, después de una fase de *calentamiento* del algoritmo. Si la convergencia se alcanza, las medias en esos intervalos deben ser semejantes.

Considere $\pi = t(\theta)$ una función real de un parámetro de interés θ , para el cuál fué obtenida una muestra a partir del algoritmo MCMC. Sea m el número de iteraciones correspondientes a la fase de *calentamiento* y $m+n$ el total de iteraciones. Se construyen las medias $\bar{\pi}_a$ y $\bar{\pi}_b$ en base a los grupos de iteraciones $n_a < n$ y $n_b < n$. Luego

$$z_G = \frac{\bar{\pi}_a - \bar{\pi}_b}{\sqrt{Var(\bar{\pi}_a) + Var(\bar{\pi}_b)}} \rightarrow N(0, 1)$$

De esta forma, los valores grandes de z_G indican falta de convergencia. Sin embargo, valores pequeños de esta estadística no significan que existe convergencia. Luego, la decisión debe ser tomada conjuntamente con otros criterios y con la inspección visual. Para implementar este criterio basta con efectuar solamente una cadena con un número grande de iteraciones del algoritmo MCMC.

Estimadores Bayesianos

Un enfoque de decisión teórica para la estimación de un parámetro desconocido θ introduce la función de pérdida $L(\theta, T)$ que, en términos generales, da el costo de decidir que el

parámetro tiene el valor T , cuando en realidad es igual a θ . La estimación T puede ser elegida de forma que minimice la *pérdida posteriori esperada*,

$$E(L(a|y)) = \int L(\theta, a)p(\theta|y)d\theta$$

y se llama *Estimación de Bayes* (T^B) para la función de pérdida L .

A continuación se muestran las siguientes funciones de pérdida:

- Si $L(\theta, a) = (\theta - a)^2$, entonces la media posteriori, $E(\theta|y)$, si existe, es la única estimación de Bayes de θ .
- Si $L(\theta) = |\theta - a|$, entonces la mediana posterior de θ es una estimación de Bayes de θ .
- Si

$$L(\theta, a) = \begin{cases} 0, & \text{si } \theta = a \\ 1, & \text{si } \theta \neq a \end{cases}$$

entonces la estimación bayesiana es la moda de la distribución posteriori [18].

El estimador Bayes T^B es la solución de

$$T^B = \underset{T}{\text{mín}} E[L(T, \theta)]$$

Con la pérdida lineal absoluta, se tiene que:

$$\begin{aligned} E[L(T, \theta)] &= \int |T - \theta| f(\theta) d\theta \\ &= \int_{-\infty}^T (T - \theta) f(\theta) d\theta + \\ &\quad + \int_T^{\infty} (\theta - T) f(\theta) d\theta \end{aligned}$$

$$\begin{aligned} \frac{d}{dT} E[L(T|\theta)] &= (T - T) f(T) + \int_{-\infty}^T f(\theta) d\theta - \\ &\quad - (T - T) f(T) - \int_T^{\infty} f(\theta) d\theta \\ &= F(T) - (1 - F(T)) \\ &= 2F(T) - 1 \end{aligned}$$

Luego, dado que para encontrar el mínimo se debe igualar la derivada a cero, se obtiene que $F(T^B = \frac{1}{2})$ y el estimador Bayes es la mediana de la distribución de θ .

Intervalos Creíbles

El análogo bayesiano al intervalo de confianza es el intervalo creíble y, más generalmente, el conjunto creíble, que no tiene que ser contiguo. La mayor parte del tiempo, en la práctica se calcula exactamente de la misma manera que el intervalo de confianza para formas simétricas unimodales. Por ejemplo, el cálculo de un intervalo creíble del 95% bajo el supuesto de normalidad significa que se eliminan 1.96 errores estándar de la media en cualquier dirección, al igual que se crea un intervalo de confianza análogo. Sin embargo, para distribuciones asimétricas, este algoritmo produciría un intervalo creíble con colas desiguales y cobertura incorrecta.

Un intervalo creíble del $100(1 - \alpha)$, da la región del espacio de parámetros donde la probabilidad de cubrir θ es igual a $1 - \alpha$ (en realidad puede ser un poco más de $1 - \alpha$ para espacios de parámetros discretos, para garantizar al menos este nivel de cobertura).

En contraste, la aplicación de esta definición al intervalo de confianza, significa que la probabilidad de cobertura es cero o uno, ya que cubre el verdadero valor de θ o no lo hace.

Formalmente, un conjunto creíble de igual cola para la distribución posteriori se define como un subconjunto C del espacio de parámetros, θ , de manera que un intervalo creíble del $100(1 - \alpha)$ cumpla la condición:

$$1 - \alpha = \int_C \pi(\theta|\mathbf{X})d\theta \quad (2.13)$$

Es importante tener en cuenta que los intervalos creíbles no son únicos. Es decir, se puede definir fácilmente C en diferentes formas para cubrir partes de θ y aún así cumplir con la condición probabilista (2.13). Las diferencias importantes surgen en las distribuciones asimétricas y multimodales, y la convención es crear intervalos de colas iguales, independientemente de la forma de la distribución posteriori [18].

Factor Bayes

En la estadística bayesiana se puede usar los factores bayes como una alternativa a los contrastes de hipótesis clásicos.

Supongamos que se debe elegir entre dos posibles modelos M_1 y M_2 , una vez que se ha observado una muestra \mathbf{x} . El Factor Bayes, B , se define como:

$$B = \frac{f(\mathbf{x}|M_1)}{f(\mathbf{x}|M_2)}$$

donde $f(\mathbf{x}|M_i)$ es la verosimilitud marginal o verosimilitud integrada del modelo i [14].

Los modelos M_1 y M_2 se identifican mediante los vectores de parámetros $\boldsymbol{\theta}_1$ y $\boldsymbol{\theta}_2$. De este modo, B se puede expresar como:

$$B = \frac{f(\mathbf{x}|M_1)}{f(\mathbf{x}|M_2)} = \frac{\int_{\boldsymbol{\theta}_1} f(x|\boldsymbol{\theta}_1, M_1)\pi(\boldsymbol{\theta}_1|M_1)d\boldsymbol{\theta}_1}{\int_{\boldsymbol{\theta}_2} f(x|\boldsymbol{\theta}_2, M_2)\pi(\boldsymbol{\theta}_2|M_2)d\boldsymbol{\theta}_2}$$

Jeffreys estableció una escala de interpretación para B la cuál está dada en el Cuadro 2.1.

Cuadro 2.1: Escala según Jeffreys para el Factor Bayes.

B	Fuerza de la evidencia a favor de M_1
< 1	Negativa(apoya a M_2)
$1 - 3$	Muy escasa
$3 - 10$	Sustancial
$10 - 30$	Fuerte
$30 - 100$	Muy Fuerte
> 100	Decisiva

Estimación de la Verosimilitud Marginal

Un método para aproximar la verosimilitud marginal es utilizar los valores simulados, vía MCMC, de la distribución a posteriori mediante la siguiente ecuación:

$$\int \frac{1}{f(y|\boldsymbol{\theta}_m, m)} f(\boldsymbol{\theta}_m|y, m)d\boldsymbol{\theta}_m = \int \frac{1}{f(y|\boldsymbol{\theta}_m, m)} \frac{f(y|\boldsymbol{\theta}_m, m)f(\boldsymbol{\theta}_m|m)}{f(y|m)} d\boldsymbol{\theta}_m = \frac{1}{f(y|m)}$$

Newton y Raftery [19] propusieron una estimación de la verosimilitud marginal. Este estimador, se obtiene en base a la media armónica de las verosimilitudes calculadas en cada paso de un algoritmo MCMC y está definido como:

$$\hat{f}_1(y|m) = \left(\frac{1}{T} \sum \left\{ f(y|\boldsymbol{\theta}_m^{(t)}, m) \right\} \right)^{-1}$$

Gelfand y Dey [20] generalizaron esta idea con:

$$\hat{f}_2(y|m) = \left(\frac{1}{T} \sum \frac{g(\boldsymbol{\theta}_m^{(t)})}{f(y|\boldsymbol{\theta}_m^{(t)}, m) f(\boldsymbol{\theta}_m^{(t)}|m)} \right)^{-1}$$

En esta ecuación g desempeña una densidad de muestreo por importancia y las elecciones naturales para que coincidan con la parte posteriori, serían distribuciones *normales* o *t* multivariadas con media y covarianza calculadas a partir de $\boldsymbol{\theta}_m$.

Metodología

3.1. Datos

Los datos para este trabajo corresponden a conteos que representan el número de casos de dengue ocurridos durante las 52 semanas del año 2010 en el municipio Santiago de Cali, Colombia. La distribución de los datos se presentan en la Figura 3.1 [9].

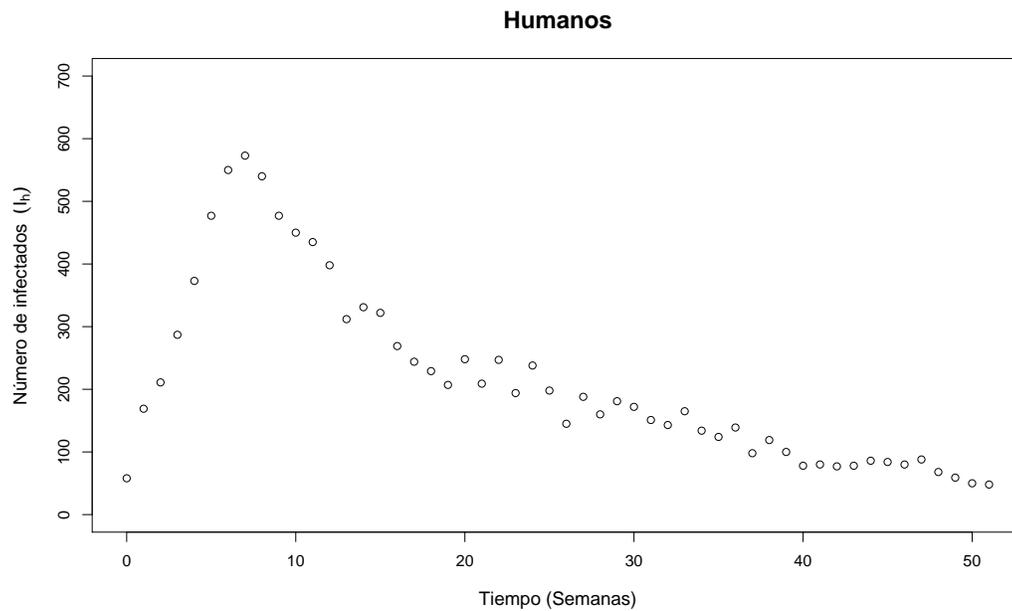


Figura 3.1: Incidencia de dengue ocurrida en la ciudad de Cali, Colombia en el año 2010.

3.2. Estimación Bayesiana

3.2.1. Modelo Estadístico

Para la estimación bayesiana de los parámetros en los Sistemas Dinámicos (2.1) y (2.3), el modelo estadístico que se considera es:

$$y(t_j) = V(t_j, \boldsymbol{\theta}) + \epsilon(t_j), \quad j = 1, \dots, 52$$

donde:

- $\boldsymbol{\theta}$ es el vector de los parámetros de los sistemas dinámicos, $\boldsymbol{\theta} = (\beta_h, \beta_v, \gamma, A, I_{V_{inic}})$.
- $y(t_j)$ son el número de casos de dengue que se reportaron en la semana t_j .
- $V(t_j, \boldsymbol{\theta})$ es la solución numérica de la ecuación correspondiente a la variable de estado “humanos infectados” de los Sistemas (2.1) y (2.3) en el tiempo t_j .
- $\epsilon(t_j)$ son los ruidos aleatorios que surgen de aproximar los datos $y(t_j)$ a través de $V(t_j, \boldsymbol{\theta})$ en el tiempo t_j . Se asume que estos ruidos tienen distribución normal, con media 0 y varianza σ^2 (desconocida).

Por consiguiente

$$y(t_j) \sim N(V(t_j, \boldsymbol{\theta}), \sigma^2)$$

3.2.2. Función de Verosimilitud

Para obtener la función de verosimilitud de los datos se requiere calcular su media, la cuál corresponde a la solución numérica de los sistemas (2.1) y (2.3), en este trabajo se obtiene mediante el método Runge Kutta de orden 4 y de las condiciones iniciales que se dan a continuación:

Condiciones Iniciales

- En el modelo de *Ross – Macdonald* el número casos de humanos infectados $I_h(t_0) = 58$ se obtuvo de la cantidad de casos que hubo en la primera semana de los datos registrados.

- En el modelo de *Vector – Hospedero* también $I_h(t_0) = 58$. La cantidad de humanos susceptibles $S_h(t_0) = 2244581$ se calcula a través de la relación $N_h = S_h + I_h + R_h$, donde la cantidad de humanos recuperados inicialmente era de 0 casos y se tenía una población total de 2244639 personas [21].

Valores iniciales de los parámetros

- La mortalidad natural de los mosquitos, en la región de donde provienen los datos y una temperatura en una temperatura promedio entre $[18^\circ C, 24^\circ C]$ [12], implica que es de dos semanas aproximadamente [8], por lo que $\mu_v = \frac{1}{2}$.
- La mortalidad natural de los humanos es $\mu_h = \frac{1}{76 \times 52}$, para lo cuál se toma en cuenta que el promedio de vida en Colombia en el año 2010 era de aproximadamente de 76 años.

3.2.3. Información a priori

Con base en la bibliografía revisada, se tiene la siguiente información sobre los parámetros de ambos sistemas dinámicos:

- La probabilidad de que un humano susceptible quede infectado al ser picado por un mosquito infectado (β_h) tiene como soporte $[0.1, 0.75]$ [22].
- La probabilidad de que un mosquito susceptible quede infectado al picar a un humano infectado (β_v) tiene como soporte $[0.5, 1]$ [22].
- La tasa de recuperación semanal de los humanos (γ) tiene como soporte $[1, 7/3]$ pues según la OMS la enfermedad está presente en los humanos, de 3 a 7 días aproximadamente [1].
- Para la tasa de reclutamiento de los mosquitos A se toma como soporte $[3.0 \times 10^5, 3.3 \times 10^5]$ pues según Alcalá et al. [23], obtienen que el índice de pupa *Aedes Aegypti* por cada persona en Colombia puede tomar valores entre 0.3 y 0.4. Tomando en cuenta estos valores para una población de 2244639 personas se obtienen los valores mencionados.

- Se tuvo en cuenta el soporte $[0, 5000]$ para el número inicial de vectores infectados $I_{V_{inic}}$, dado que Aldana et al. [24] reporta que la proporción inicial de mosquitos infectados es de 0.019, usando este resultado y la población de mosquitos (A/μ_v) se llega a que este rango del parámetro.

Distribuciones a priori

Con base a estos soportes para cada parámetro, en este trabajo se utilizan los siguientes supuestos:

- Independencia entre las distribuciones a priori de los parámetros.
- Distribuciones a priori no informativas.

En el Cuadro 3.1 se muestran las distribuciones a priori para los parámetros de los Sistemas Dinámicos (2.1) y (2.3).

Cuadro 3.1: Distribuciones a priori.

Parámetros	Distribuciones a priori
β_h	$U(0.1, 0.75)$
β_v	$U(0.5, 1)$
γ	$U(1, \frac{7}{3})$
A	$U(3.0 \times 10^5, 3.3 \times 10^5)$
$I_{V_{inic}}$	$U(0, 5 \times 10^3)$

- Para el caso de la varianza se toma en cuenta la precisión definida por $\tau = \frac{1}{\sigma^2}$ donde se asume que tiene una distribución $\Gamma(1.0 \times 10^{-3}, 1.0 \times 10^{-3})$ [25].

3.2.4. Distribuciones a posteriori

Para calcular las distribuciones a posteriori marginales se utilizaron las cadenas MCMC, se generaron 3 cadenas para cada parámetro con 300000 iteraciones cada una, de las cuales se descartaron las primeras 250000 y se realizó una selección con un adelgazamiento de tamaño 10.

La estimación puntual bayesiana de los parámetros se obtuvo con la mediana de la distribución a posteriori marginal de los parámetros y los intervalos creíbles se calcularon con los percentiles 2.5 % y 97.5 %. El número reproductivo básico de cada sistema se estimó a partir de la distribución a posteriori obtenida para cada uno de sus parámetros.

Los cálculos se hicieron mediante la función *jags* de la librería *R2jags* [26] del software estadístico R [27]. Se hizo uso también de las librerías *deSolve* [28] y *coda* [29].

Resultados y Discusión

En este capítulo se muestran los resultados de la estimación puntual e intervalos creíbles de los parámetros de los Sistemas Dinámicos (2.1) y (2.3). Así mismo se presenta la estimación bayesiana y los intervalos creíbles del número reproductivo básico (R_0).

Las trazas, las densidades a posteriori marginales y los diagnósticos de convergencias para cada uno de los parámetros se presentan en los Anexos 1 y 2.

4.1. Estimación bayesiana del modelo *Ross – Macdonald*

Diagnósticos de convergencia

Se realizaron los diagnósticos de Gelman - Rubin y Geweke, además de la inspección visual, para comprobar la convergencia de las cadenas de cada parámetro.

En el Cuadro 4.1 se muestran los resultados del diagnóstico de convergencia de Gelman - Rubin. Se puede observar que todos los valores obtenidos del factor de reducción de escala potencial se encuentran muy cercanos a uno, lo cual indica la convergencia de la cadena a la distribución estacionaria, para los parámetros de interés del modelo *Ross – Macdonald* [18,30].

Cuadro 4.1: Resultados del diagnóstico de convergencia Gelman - Rubin para las estimaciones para el Sistema Dinámico(2.1)

	Estimación puntual	I.C Superior
β_h	1.00	1.00
β_v	1.00	1.00
γ	1.00	1.01
A	1.00	1.00
I_{Vinic}	1.01	1.02
R_0	1.01	1.02

En el Cuadro 4.2 se muestran los resultados del diagnóstico de convergencia de Geweke. En estos se puede apreciar que para casi todos los parámetros $|Z| < 2$, por lo que las medias para las primeras y últimas iteraciones para cada parámetro son iguales, excepto para la segunda cadena del parámetro β_v . Este resultado también se puede ver reflejado en el Anexo 1 (Figuras 5.7, 5.8 y 5.9).

Cuadro 4.2: Resultados del diagnóstico de convergencia de Geweke.

Cadenas	Parámetros					
	β_h	β_v	γ	A	I_{Vinic}	R_0
1	0.91	-1.07	-0.20	0.12	-0.18	0.15
2	-1.78	2.89	-0.21	-1.60	0.31	-0.20
3	-0.41	0.47	0.98	0.27	1.06	-0.77

En el Anexo 1 se muestran las trazas, la distribución a posteriori marginal y autocorrelación de cada cadena de los parámetros que se estimaron del modelo de *Ross – Macdonald*. En las trazas se muestra que existe un buen mezclado entre las cadenas de cada parámetro. De la distribución a posteriori de las cadenas para cada parámetro se puede observar que están superpuestas una sobre otra. Las autocorrelaciones comienzan siendo altas pero rápidamente disminuyen y tienden a cero, aunque en algunos casos las autocorrelaciones tienden más rápido a cero.

Esta evidencia indica que las cadenas de cada parámetro si logran alcanzar la convergencia, por lo que partiendo de valores iniciales distintos estas alcanzan la misma distribución y por consiguiente las estimaciones obtenidas son confiables.

Estimación de los parámetros.

En el Cuadro 4.3 se presentan las estimaciones bayesianas e intervalos creíbles para los parámetros de interés del Sistema Dinámico(2.1).

En este cuadro se puede observar lo siguiente: el estimador bayesiano de la probabilidad de transmisión vector – humano (β_h) es de 0.739, esto indica que por cada 100 picaduras de los mosquitos infectados en los humanos susceptibles, aproximadamente 74 picaduras provocaron que los humanos quedaran infectados y existe una probabilidad de 0.95 de que este valor esté entre 72 y el 75 picaduras aproximadamente. Para la probabilidad de transmisión humano – vector (β_v) el estimador bayesiano es de 0.985, lo que señala entonces que por cada 100 picaduras de los mosquitos susceptibles en los humanos infectados, aproximadamente 98 picaduras provocaron que los mosquitos quedaran infectados y con una probabilidad de 0.95 de que este valor se encuentre entre 95 y 99 picaduras aproximadamente. La recuperación semanal de los humanos (γ), según la estimación bayesiana fue de 2.151, en otras palabras, la duración de la enfermedad en los humanos fue de 3.2 días aproximadamente y existe una probabilidad de 0.95 de que este valor se encuentre entre 3 y 4 días aproximadamente. La estimación bayesiana para el reclutamiento semanal de los mosquitos (A), indica que en la ciudad de Cali durante el año 2010 fue de 325198 mosquitos aproximadamente y existe una probabilidad de 0.95 de que este valor esté entre 315095 y 329794. También se obtuvo que inicialmente existía un total de 1366 mosquitos infectados $I_{V_{inic}}$, según la estimación bayesiana para este parámetro, y existe una probabilidad de 0.95 de que este valor se encuentre entre 1113 y 1872.

El número reproductivo básico, obtenido a través de la estimación bayesiana, fue de 0.859 y existe una probabilidad de 0.95 de que este valor se encuentre entre 0.831 y 0.880, indicando que no existirá brote epidémico por lo que el número de casos de humanos infectados de dengue desaparecerá con el tiempo.

Cuadro 4.3: Estimación bayesiana de los parámetros de interés del modelo *Ross – Macdonald*.

Parámetros	Unidad	Descripción	Estimación	Intervalos Creíbles	
				2.5 %	97.5 %
β_h	-	Probabilidad de transmisión vector – humano	0.739	0.716	0.750
β_v	-	Probabilidad de transmisión humano – vector	0.985	0.955	0.999
γ	sem^{-1}	Tasa de recuperación de los humanos	2.151	2.087	2.211
A	$mosq \times sem^{-1}$	Reclutamiento de los mosquitos	325198	315095	329794
I_{vini}	$mosq$	Condición inicial de los vectores infectados	1366	1113	1872
R_0	-	Número Reproductivo Básico	0.859	0.831	0.880

Ajuste del Modelo

Con las estimaciones obtenidas para cada parámetro y sus respectivos intervalos creíbles del 95 % se calcula la solución numérica del Sistema dinámico (2.1) y se obtiene la gráfica de la Figura 4.1.

En esta gráfica se observa que la solución numérica ajusta bien a 48 de los 52 datos de humanos infectados, aunque existen tres valores que se encuentran más alejados. Las bandas correspondientes a los percentiles 2.5 % y 97.5 % contienen casi el total de los datos.

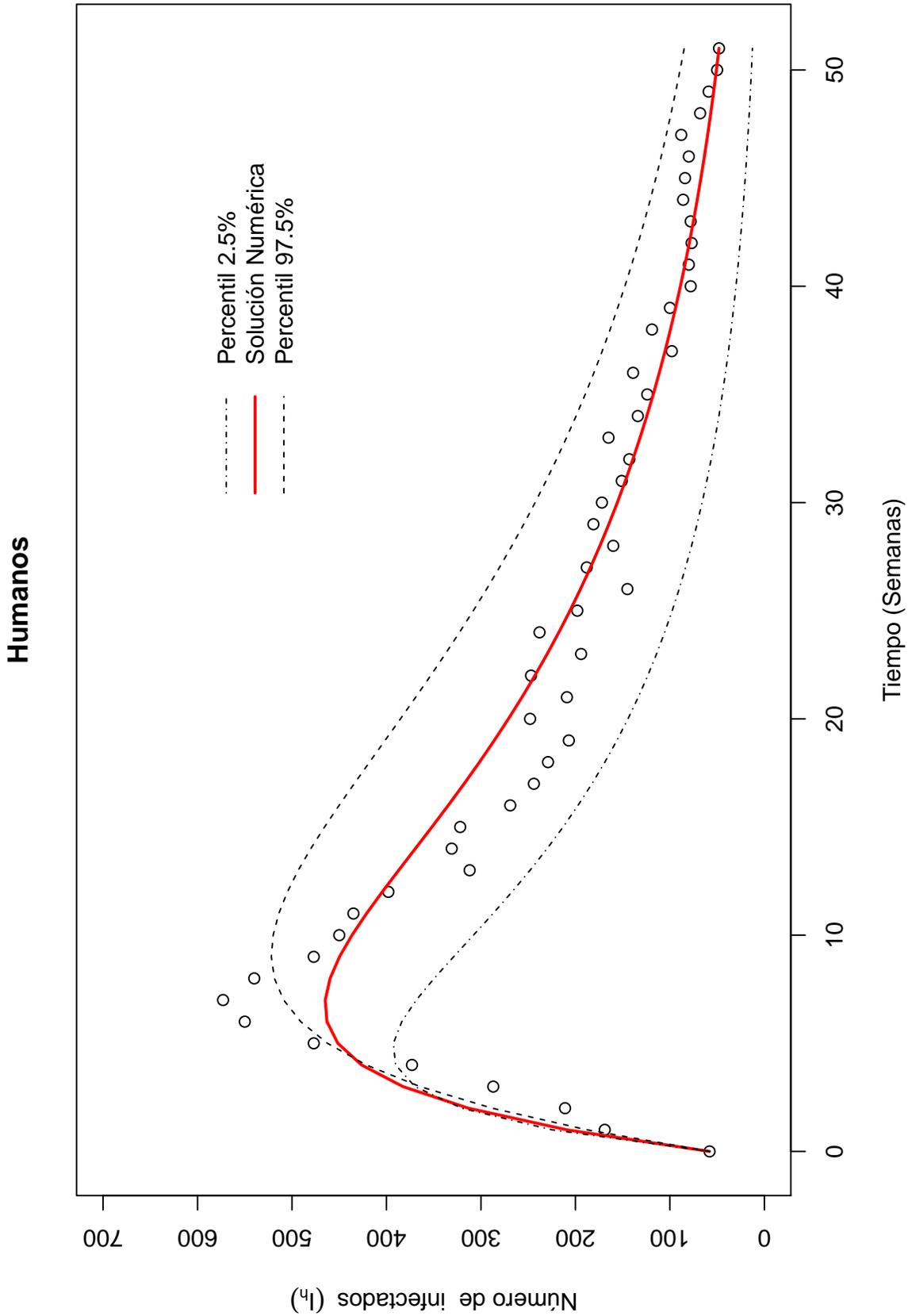


Figura 4.1: Ajuste del modelo bayesiano a la solución numérica del modelo Ross – Macdonald.

Modelo *Vector – Hospedero*

4.2. Diagnósticos de convergencia

Al igual que para el modelo anterior, se realizaron los diagnósticos de Gelman - Rubin y Geweke, además de la inspección visual para comprobar la convergencia de las cadenas de cada parámetro.

En el Cuadro 4.4 se muestran los resultados del diagnóstico de convergencia de Gelman y Rubin. Se puede observar que todos los valores obtenidos del factor de reducción de escala potencial se encuentran muy cercanos a uno, indicando esto la convergencia de las cadena a una distribución estacionaria, para los parámetros de interés en el modelo *Vector – Hospedero* [18, 30].

Cuadro 4.4: Resultados del diagnóstico de convergencia de Gelman - Rubin para las estimaciones para el Sistema dinámico (2.3).

	Estimación puntual	I.C Superior
β_h	1.00	1.00
β_v	1.00	1.00
γ	1.00	1.00
A	1.00	1.00
I_{Vinic}	1.06	1.07
R_0	1.01	1.01

En el Cuadro 4.5 se muestran los resultados del diagnóstico de convergencia de Geweke. En estos se muestra que para casi todos los parámetros $|Z| < 2$, por lo que las medias para las primeras y últimas iteraciones de cada parámetro son iguales, excepto para la primera cadena de los parámetros β_h y β_v . Este resultado también se encuentra en el Anexo 2 (Figuras 5.16, 5.17 y 5.18).

Cuadro 4.5: Resultados del diagnóstico de convergencia de Geweke.

Cadenas	Parámetros					
	β_h	β_v	γ	A	I_{Vinic}	R_0
1	3.13	-2.17	0.85	-0.20	0.65	-0.65
2	1.35	0.25	-0.36	-1.14	-0.05	0.66
3	0.17	0.31	-0.83	0.10	-0.89	1.01

En el Anexo 2 se muestran las trazas, la distribución a posteriori marginal y autocorrelación de cada cadena de los parámetros que se estimaron del modelo de *Vector – Hospedero*. Al igual que para el modelo de *Ross – Macdonald*, entre las trazas de los parámetros se muestra que existen un buen mezclado, en la distribución a posteriori de las cadenas para cada parámetro se observa que están superpuestas una sobre otra y las autocorrelaciones comienzan siendo alta y rápidamente disminuyen y tienden a cero, aunque en algunos casos las autocorrelaciones tienden más rápido a cero.

Esta evidencia indica que las cadenas de cada parámetro si logran alcanzar la convergencia, por lo que partiendo de valores iniciales distintos estas alcanzan la misma distribución y por consiguiente las estimaciones obtenidas son confiables.

4.3. Estimación de los parámetros.

En el Cuadro 4.6 se presentan las estimaciones bayesianas e intervalos creíbles para los parámetros de interés del Sistema dinámico (2.3).

Cuadro 4.6: Estimación bayesiana de los parámetros de interés del modelo *Vector – Hospedero*.

Parámetros	Unidad	Descripción	Estimación	Intervalos Creíbles	
				2.5 %	97.5 %
β_h	-	Probabilidad de transmisión vector – humano	0.740	0.720	0.750
β_v	-	Probabilidad de transmisión humano – vector	0.987	0.960	0.999
γ	sem^{-1}	Tasa de recuperación de los humanos	2.142	2.074	2.204
A	$mosq \times sem^{-1}$	Reclutamiento de los mosquitos	325785	316745	329818
$I_{V_{inic}}$	$mosq$	Condición inicial de los vectores infectados	1343	1073	1785
R_0	-	Número Reproductivo Básico	0.868	0.843	0.888

En el cuadro anterior se tiene lo siguiente: la estimación bayesiana de la probabilidad de transmisión vector – humano (β_h) es de 0.740, lo que indica que por cada 100 picaduras de los mosquitos infectados en los humanos susceptibles, 74 provocaron que los humanos se infectaran y existe una probabilidad de 0.95 de que este valor se encuentre entre 72 y 75 picaduras aproximadamente. Para la probabilidad de transmisión humano – vector (β_v) la estimación bayesiana es de 0.987, lo que señala entonces que por cada 100 picaduras de los mosquitos susceptibles en los humanos infectados, 98 provocaron que los mosquitos se infectaron y existe una probabilidad de 0.95 de que este valor se encuentre entre 96 y 99 picaduras aproximadamente. La recuperación semanal de los humanos (γ), según la estimación bayesiana, 2.142, indica que la duración de la enfermedad en los humanos fue de 3.2 días aproximadamente y existe una probabilidad de 0.95 de que este valor se encuentre entre 3 y 4 días aproximadamente. La estimación bayesiana para el reclutamiento semanal de los mosquitos (A), indica que en la ciudad de Cali, Colombia durante el año 2010 fue de 325785 mosquitos aproximadamente y existe una probabilidad de 0.95 de que este valor se encuentre entre 316745 y 329818.

También se obtuvo que inicialmente existía un total de 1343 mosquitos infectados $I_{V_{inic}}$, según la estimación bayesiana para este parámetro, y existe una probabilidad de 0.95 de que este valor se encuentre entre 1073 y 1785. El número reproductivo básico obtenido por la estimación bayesiana fue de 0.868, que por ser menor a la unidad, indica que no existirá brote epidémico, por lo contrario el número de casos de dengue desaparecerá con el tiempo. Además, existe una probabilidad de 0.95 de que este valor esté entre 0.843 y 0.888, y por no incluir a la unidad ratifica lo anteriormente planteado.

Ajustes de los modelos

Con las estimaciones obtenidas para cada parámetro y sus respectivos intervalos creíbles del 95 % se calcula la solución numérica del Sistema dinámico (2.3) a través del método de Runge Kutta de orden 4 y se puede observar en la Figura 4.2.

En esta gráfica se observa que la solución numérica ajusta bien a 46 de los 52 datos de humanos infectados, aunque existen tres valores que se encuentran más alejados. Las bandas correspondientes a los percentiles 2.5 % y 97.5 % contienen casi la totalidad de los datos.

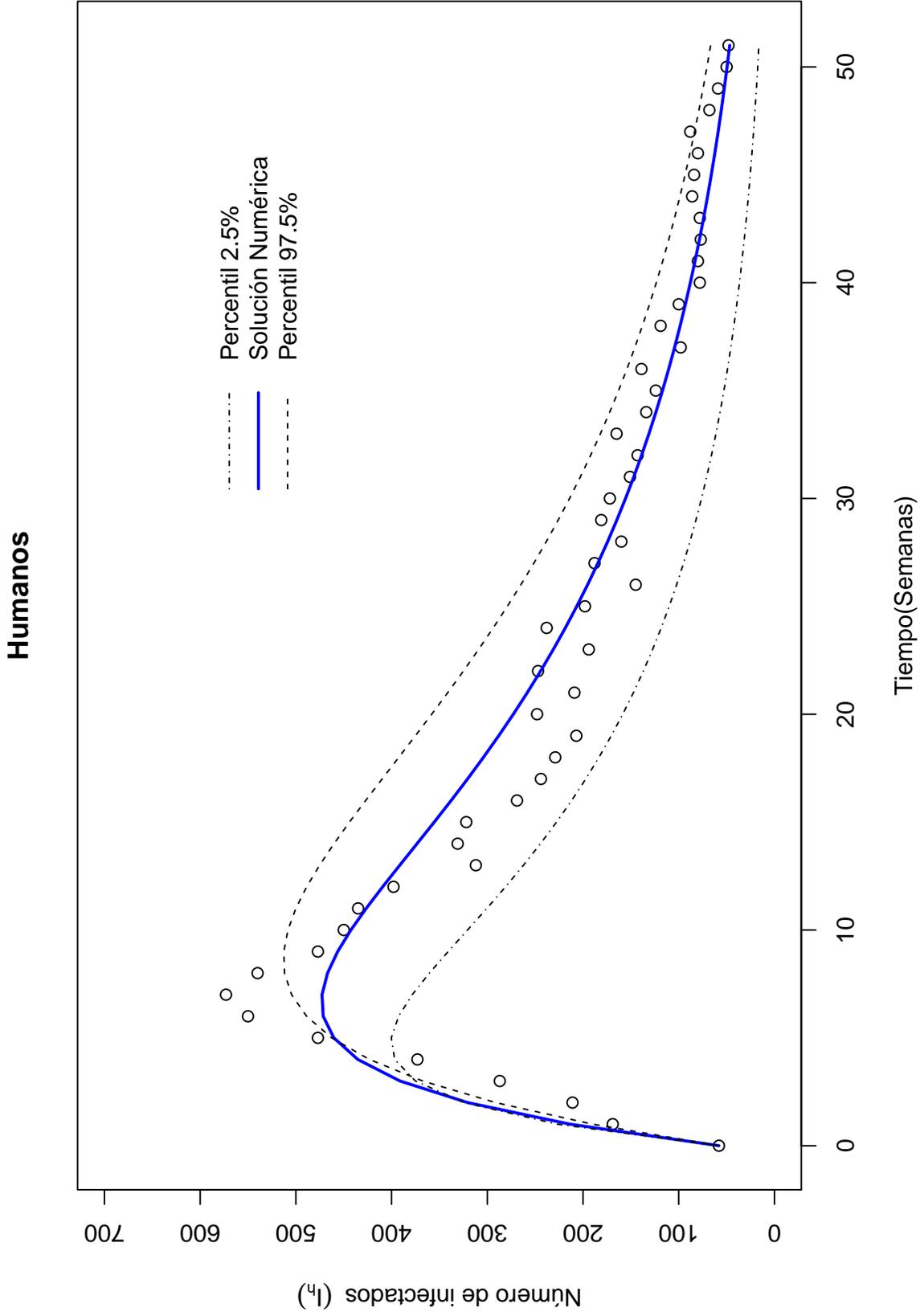


Figura 4.2: Ajuste del modelo bayesiano a la solución numérica del modelo *Vector - Hospedero*.

4.4. Comparación de Modelos

Con el objetivo de seleccionar el modelo que mejor ajusta a los datos de humanos infectados, primeramente se realiza una comparación entre los ajustes de ambos modelos con las estimaciones bayesianas obtenidas para los parámetros. Se puede observar que para todos los casos las estimaciones, a pesar de no ser iguales, si son muy similares. En la Figura 4.3 se muestra las gráficas de ajuste de las soluciones numéricas para los modelos *Ross – Macdonald* y *Vector – Hospedero* y los datos del número de casos semanales de dengue.

En la figura anterior se observa que con los resultados de las estimaciones bayesianas ambos modelos ajustan bien a los conteos de humanos infectados; se puede apreciar que ambos ajustes son muy parecidos gráficamente. Por tal razón resulta conveniente realizar la comparación de ambos modelos mediante los criterios de Devianza y Factor Bayes. Los resultados de la comparación se presentan en el Cuadro 4.7.

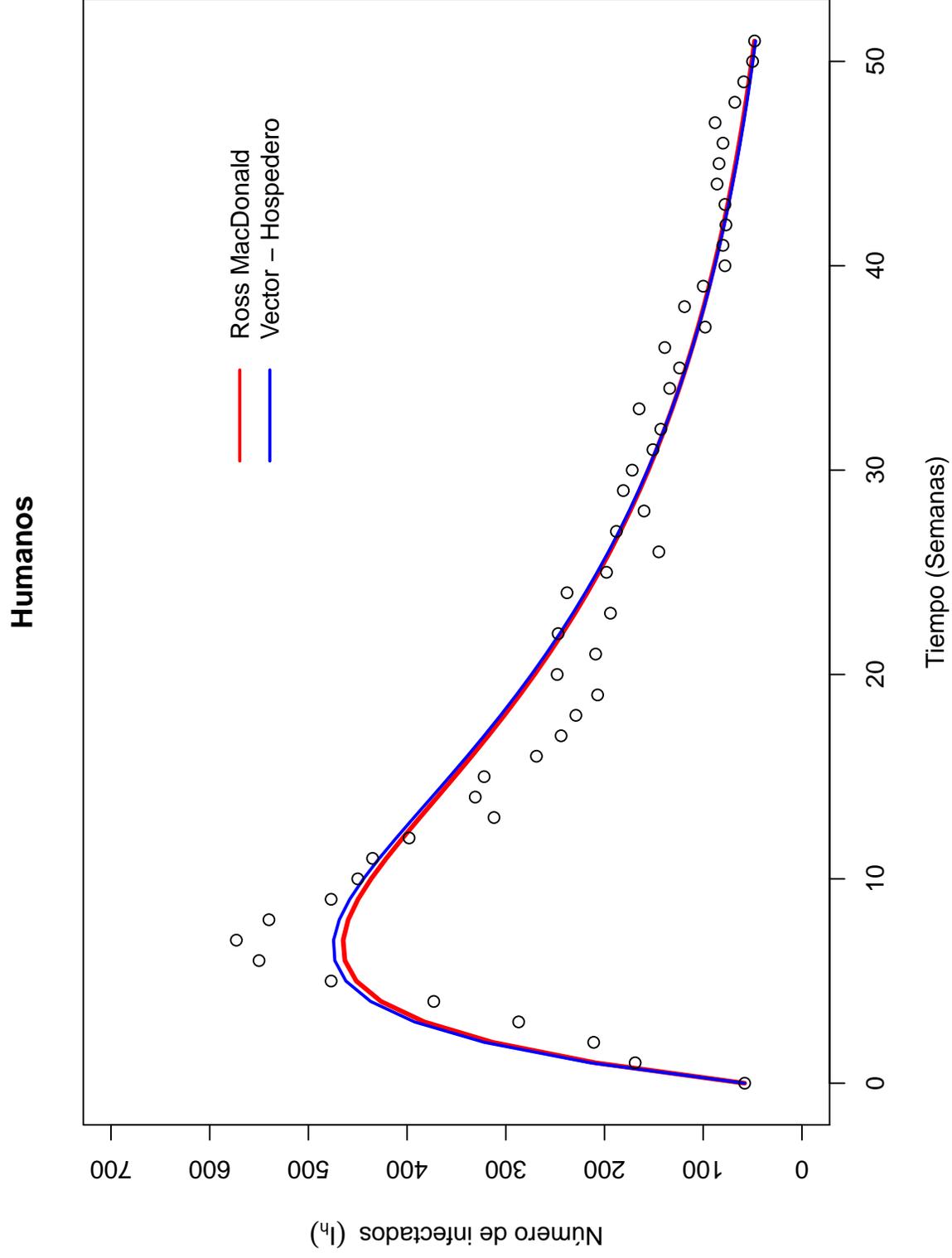


Figura 4.3: Ajuste de los modelos *Ross - Macdonald* y *Vector - Hospedero*, con las estimaciones obtenidas, a los casos semanales de humanos infectados de dengue.

Cuadro 4.7: Comparación entre los resultados del Criterio de Devianza y Factor Bayes para los modelos *Ross – Macdonald* y *Vector – Hospedero*.

Criterio de Comparación	Modelo <i>RM</i>	Modelo <i>VH</i>
DIC	545.3	546.2
Factor Bayes	1.12	

En este cuadro se observa que el valor del DIC para el modelo *Ross – Macdonald* es menor que para el *Vector – Hospedero*, lo cuál indica que el modelo *Ross – Macdonald* es mejor.

Con el criterio del Factor Bayes se observa que la prueba que muestra la fuerza de la evidencia a favor del modelo *Ross – Macdonald* tiene como resultado 1.12, lo que indica que existe escasa evidencia que el modelo *Ross – Macdonald* sea mejor al modelo *Vector – Hospedero*. Esta evidencia parece indicar que el modelo *Ross – Macdonald* es mejor para los datos de humanos infectados que se están trabajando, las condiciones iniciales y las distribuciones a priori que se están tomando en cuenta.

4.5. Discusión

Para los modelos *Ross – Macdonald* y *Vector – Hospedero* se obtuvieron las estimaciones de los parámetros $\beta_h, \beta_v, \gamma, A$ y de la condición inicial $I_{V_{inic}}$, para esto se tomó en cuenta distribuciones a priori no informativas e independientes, con soporte obtenido de literaturas relacionadas con el tema [8,22]. Para obtener las estimaciones bayesianas no fue necesario fijar valores iniciales a los parámetros de interés ni una de las condiciones iniciales de los modelos, y solo se utilizaron los datos correspondientes a la variable de estado, “humanos infectados”. A partir de las cadenas MCMC de los parámetros estimados, se obtiene la distribución del número reproductivo básico (R_0) para cada modelo. En ambos casos dicho número fue menor a la unidad, lo cuál indica que la epidemia del virus del dengue no se propagó.

En este trabajo las estimaciones bayesianas obtenidas de los parámetros son consistentes con los valores dados por Esteva et al. [6], Buonomo et al. [8]; sin embargo, difiere de otros trabajos que estiman los parámetros por el método de mínimos cuadrados Arias et al. [9].

Cuadro 4.8: Comparación de las estimaciones obtenidas con la de otras investigaciones.

	Ross - Macdonald	Vector - Hospedero	Esteva	Bounomo	Arias
β_h	0.73	0.74	0.75	0.95	0.24
β_v	0.98	0.98	1	0.99	0.08

Se compararon los modelos *Ross - Macdonald* y *Vector - Hospedero*, a través de los criterios de Devianza y Factor Bayes utilizando los datos de dengue durante el año 2010 en la ciudad de Cali, se obtuvo que aunque los modelos son muy semejantes existe una escasa evidencia que el modelo *Vector - Hospedero* no es mejor al de *Ross - Macdonald*. Sin embargo el modelo *Vector - Hospedero* proporciona mayor información sobre la dinámica de la transmisión del dengue.

Conclusiones

De acuerdo con los resultados de las estimaciones bayesianas de los parámetros para los modelos *Ross – Macdonald* y *Vector – Hospedero*, sin necesidad de fijar valores iniciales a los parámetros bajo estudio para obtener la función de verosimilitud, y considerando distribuciones a priori no informativas e independientes, se concluye que ambos modelos proporcionan resultados muy similares al utilizar los datos de humanos infectados de dengue durante el año 2010 en la ciudad de Cali, Colombia.

Los resultados obtenidos sobre los intervalos creíbles para la probabilidad de transmisión vector – humano (β_h), humano – vector (β_v) y la tasa recuperación de los humanos (γ) son consistentes con la información consultada en la literatura.

La metodología bayesiana es una alternativa eficiente para estimar el valor del número reproductivo básico, que significa el número promedio de nuevos casos infectados que surgen de introducir un humano infectado en una población de humanos totalmente susceptible, esto es posible al hacer uso de las distribuciones a posteriori de los parámetros estimados. Además, esta metodología ofrece la ventaja de obtener el comportamiento de esta medida, útil para entender la transmisión del dengue por medio de las distribuciones a posteriori marginales de los parámetros. En ambos modelos el resultado de la estimación fue menor a la unidad, lo cuál indica que no se propagó la epidemia del virus del dengue.

Al comparar los resultados de las estimaciones entre los modelos *Ross – Macdonald* y *Vector – Hospedero*, para los datos de humanos infectados, con los criterios estadísticos de Devianza y Factor Bayes, se concluye que existe escasa evidencia en favor del modelo *Ross –*

Macdonald.

El enfoque bayesiano para estimar los parámetros de los sistemas dinámicos *Ross – Macdonald* y *Vector – Hospedero*, que describen la transmisión de la enfermedad del dengue, así como para comparar los resultados de ambos sistemas, resulta ser una metodología eficiente para los datos de casos de humanos infectados. Otras ventajas destacadas de este enfoque son; la obtención de una distribución para cada parámetro de interés, que describe su variabilidad, y el cálculo de los intervalos creíbles.

Referencias

- [1] Dengue y dengue grave. OMS[Internet]. [23 de Diciembre de 2018]. <https://www.who.int/>
- [2] Aldana – Bermúdez E, Restrepo – Triviño M, Muñoz – Loaiza A. Ajuste de la fuerza de infección del dengue. *Revista de Salud Pública, SciELO Public Health*. 2017; (19):194-198.
- [3] San – Martin JL, Prado M. Percepción del riesgo y estrategias de comunicación social sobre el dengue en las Américas. *Revista Panamericana de Salud Pública, SciELO Public Health*. 2004; (15):135-139.
- [4] Sepúlveda LS, Vasilieva O, Romero JM, Castr JH. Ross – Macdonald: Un modelo para la dinámica del dengue en Cali, Colombia. *Revista de Salud Publica*. 2015; 17(5):749.
- [5] Moyano GE, Fernandez – Ferreyra D. Modelo de dinámica y control de epidemia de dengue con información a gran escala[tesis doctoral en internet]. [Córdoba] Universidad Nacional de Córdoba; 2016[citado 23 diciembre de 2018]. Recuperado a partir de: <http://www.famaf.unc.edu.ar/wp-content/uploads/2016/07/ DMat99.pdf>
- [6] Esteva L, Vargas C. Analysis of a dengue disease transmission model. *Mathematical biosciences*. 1998; 150(2):131–151.
- [7] Aldila D, Götz T, Soewono E. An optimal control problem arising from a dengue disease transmission model. *Mathematical biosciences*. 2013; 242(1): 9-16.
- [8] Buonomo B, Della – Marca R. Optimal bed net use for a dengue disease model with mosquito seasonal pattern. *Mathematical Methods in the Applied Sciences*. 2018; 41(2): 573-592.

- [9] Arias J, Martínez H, Sepulveda LS, Vasilieva O. Estimación de los parámetros de dos modelos para la dinámica del dengue y su vector en Cali, Colombia. *Ingeniería Ciencia*. 2018; 14(28):69–92.
- [10] Pandey A, Mubayi A, Medlock J. Comparing vector–host and SIR models for dengue transmission. *Mathematical biosciences*. 2013; 246(2): 252-259.
- [11] Khan A, Hassan M, Imran M. Estimating the basic reproduction number for single-strain dengue fever epidemics. *Infectious diseases of poverty*. 2014; 3(1):12.
- [12] Murray JD. *Mathematical Biology: I. An Introduction*. New York: Springer. 1993.507.
- [13] Blanes Zamora S, Ginestar Peiro D, Roselló Ferragud MD. *Introducción a los métodos numéricos para ecuaciones diferenciales*. Editorial Universitat Politècnica de València. 2014.
- [14] Gelman A, Carlin J, Stern H, Dunson D, Vehtari A, Rubin D. *Bayesian Data Analysis*. New York: Chapman and Hall/CRC. 2013.
- [15] Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical Science*. 1992; 7: 457-511.
- [16] Gelman A. Inference and monitoring convergence. *Markov chain Monte Carlo in practice*. 1996: 131-143.
- [17] Geweke J. Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian statistics*. 1992: 641-649.
- [18] Gill J. *Bayesian methods: A social and behavioral sciences approach*. Chapman and Hall/CRC. 2002.
- [19] Newton MA, Raftery AE. Approximate Bayesian Inference with the Weighted Likelihood Bootstrap. *Journal of the Royal Statistical Society. Series B(Methodological)*. 1994; 56(1): 3-48.
- [20] Gelfand AE, Dey DK. Bayesian model choice: asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B(Methodological)*. 1994; 501–514.

- [21] Salud en cifras. Secretaría de Salud Pública Municipal de Cali[Internet]. Recuperado de: [http://www.cali.gov.co/salud/publicaciones/33386/ publicaciones o boletines/](http://www.cali.gov.co/salud/publicaciones/33386/publicaciones%20o%20boletines/)
- [22] Andraud M, Hens N, Marais C, Beutels P. Dynamic epidemiological models for dengue transmission: a systematic review of structural approaches. *PloS one, Public Library of Science*. 2012; 7.
- [23] Alcalá L. Quintero J, González-Uribe C, Brochero H. Productividad de *Aedes aegypti*(L.)(Diptera: Culicidae) en viviendas y espacios públicos en una ciudad endémica para dengue en Colombia. *Biomédica*. 2015: 35(2).
- [24] Aldana-Bermúdez E, Restrepo-Triviño M, Muñoz-Loaiza A. Ajuste de la fuerza de infección del dengue. *Revista de Salud Pública*. 2017; 19: 194-198.
- [25] Plummer M. JAGS Version 4.0.0 user manual.2015. Recuperado a partir de: <https://sourceforge.net/projects/mcmc-jags/files/Manuals/4x>.
- [26] Yu-Sung Su, Yajima M. R2jags: Using R to Run “JAGS”. R package version 0.5-7.2015. Recuperado a partir de: [https://CRAN.R-project.org/ package = R2jags](https://CRAN.R-project.org/package=R2jags).
- [27] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013. URL <http://www.R-project.org/>.
- [28] Soetaert K, Petzoldt T, Woodrow-Setzer R. Solving Differential Equations in R: Package deSolve. *Journal of Statistical Software*. 2010;33(9):1-25.URL <http://www.jstatsoft.org/v33/i09/>.
- [29] Plummer M, Best N, Cowles K, Vines K. CODA: Convergence Diagnosis and Output Analysis for MCMC, *R News*. 2006; 6: 7-11.
- [30] Cowles MK, Carlin BP. Markov Chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*. 1996; 91(434): 883–904.

Anexos

Anexo 1

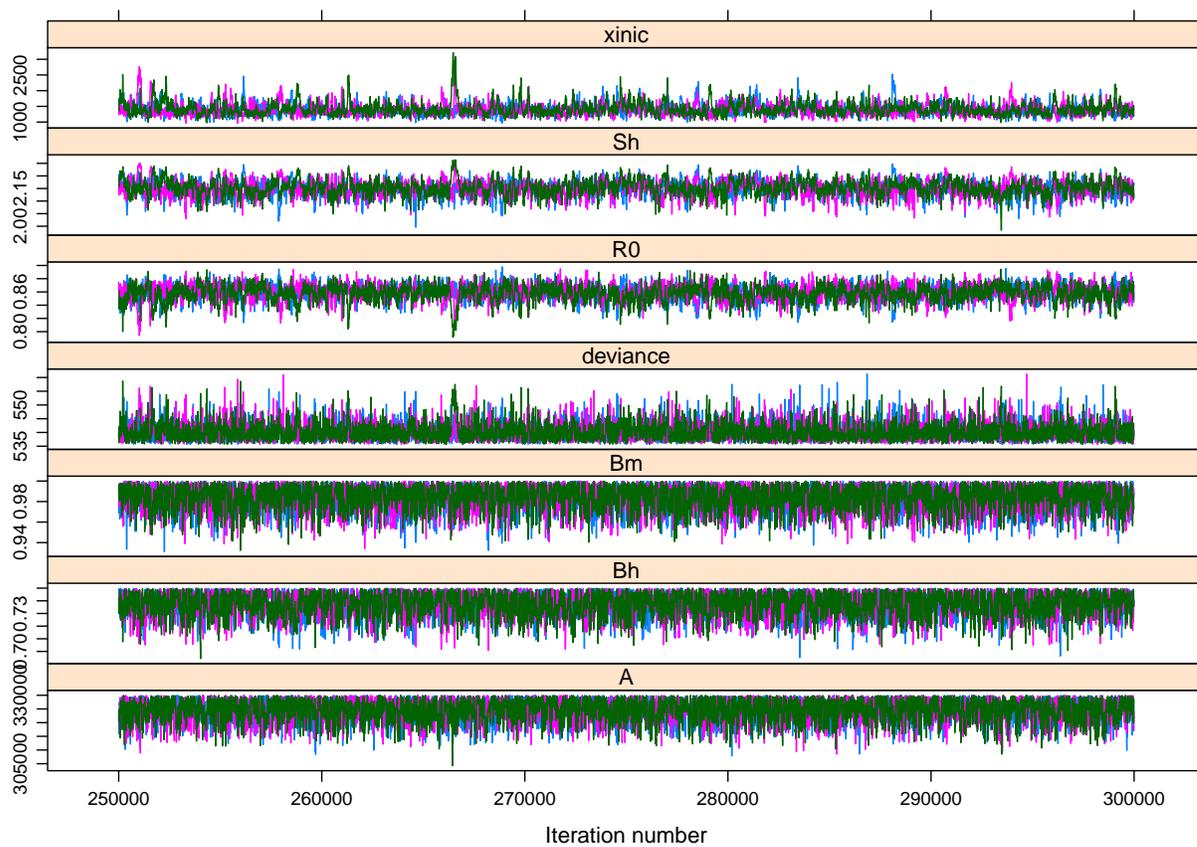


Figura 5.1: Mezclado de las cadenas para cada parámetro del modelo *Ross – Macdonald*.

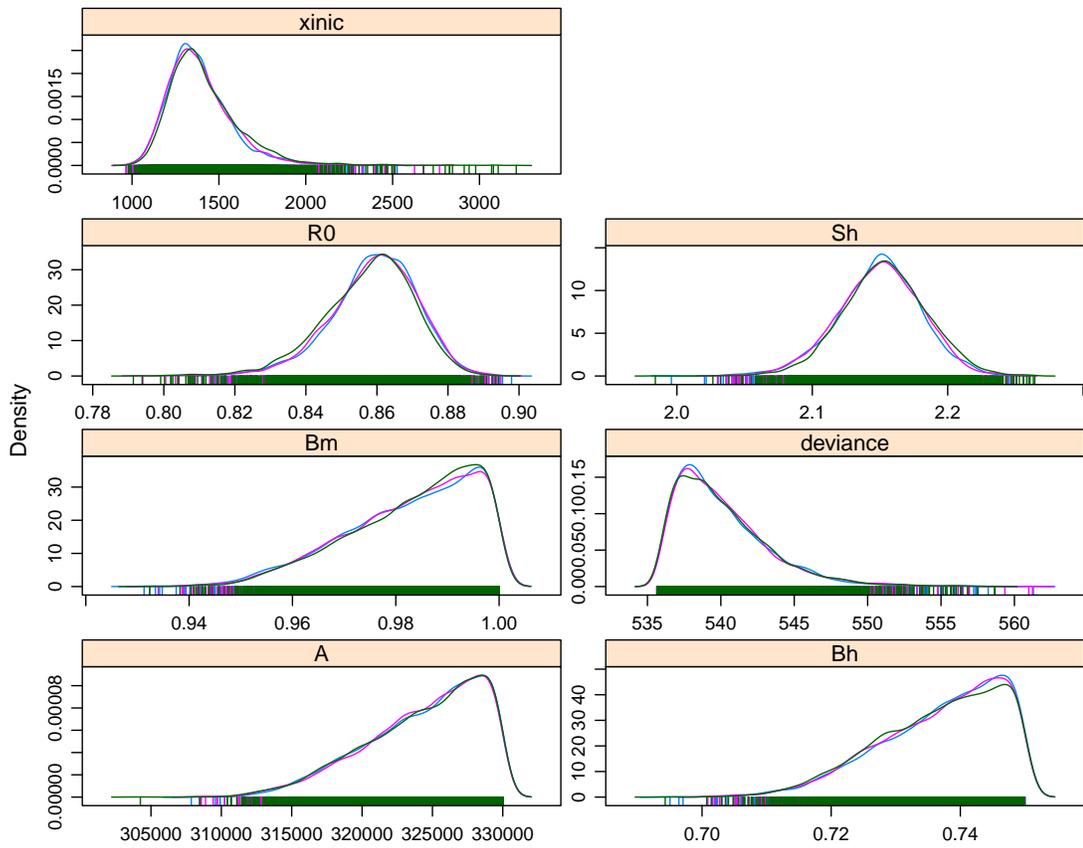


Figura 5.2: Distribución a posteriori marginal de cada cadena para los parámetros del modelo de *Ross – Macdonald*.

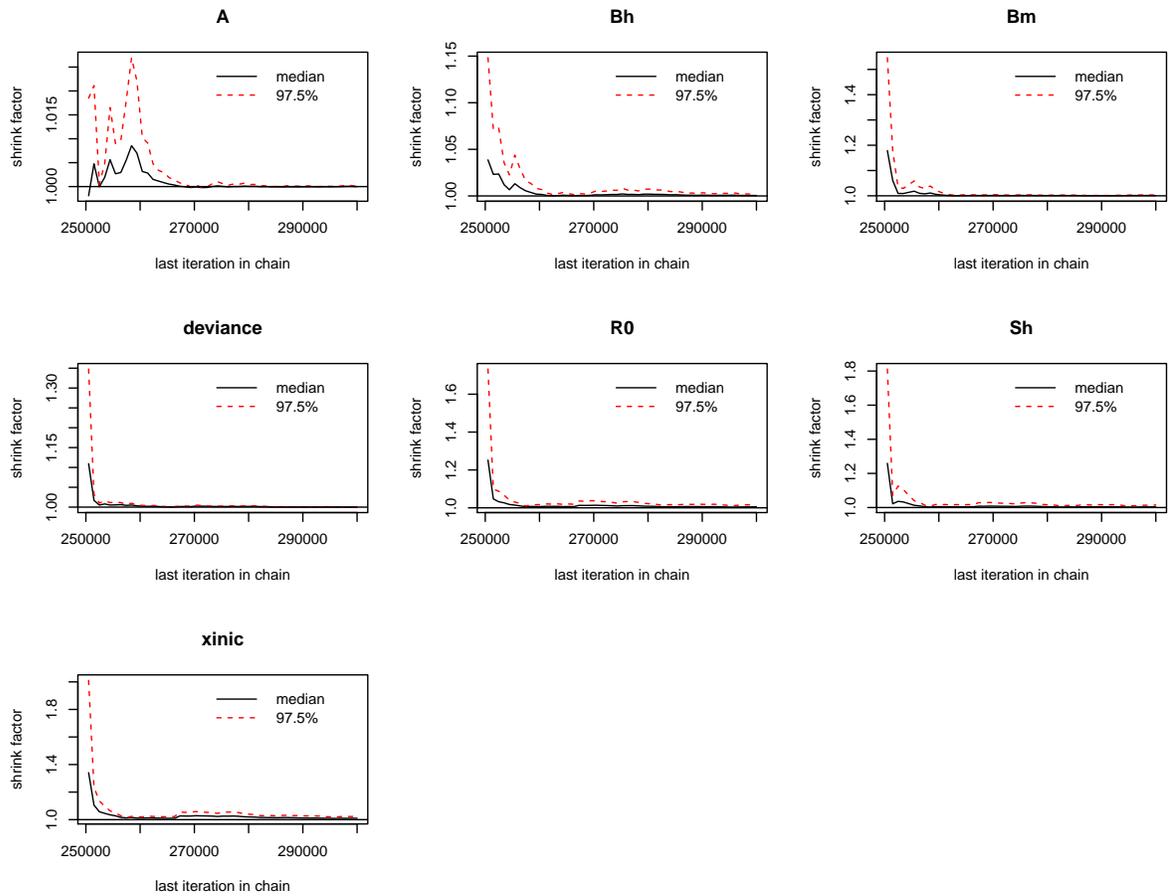


Figura 5.3: Diagnóstico de Gelman -Rubin para los parámetros del modelo *Ross – Macdonald*.

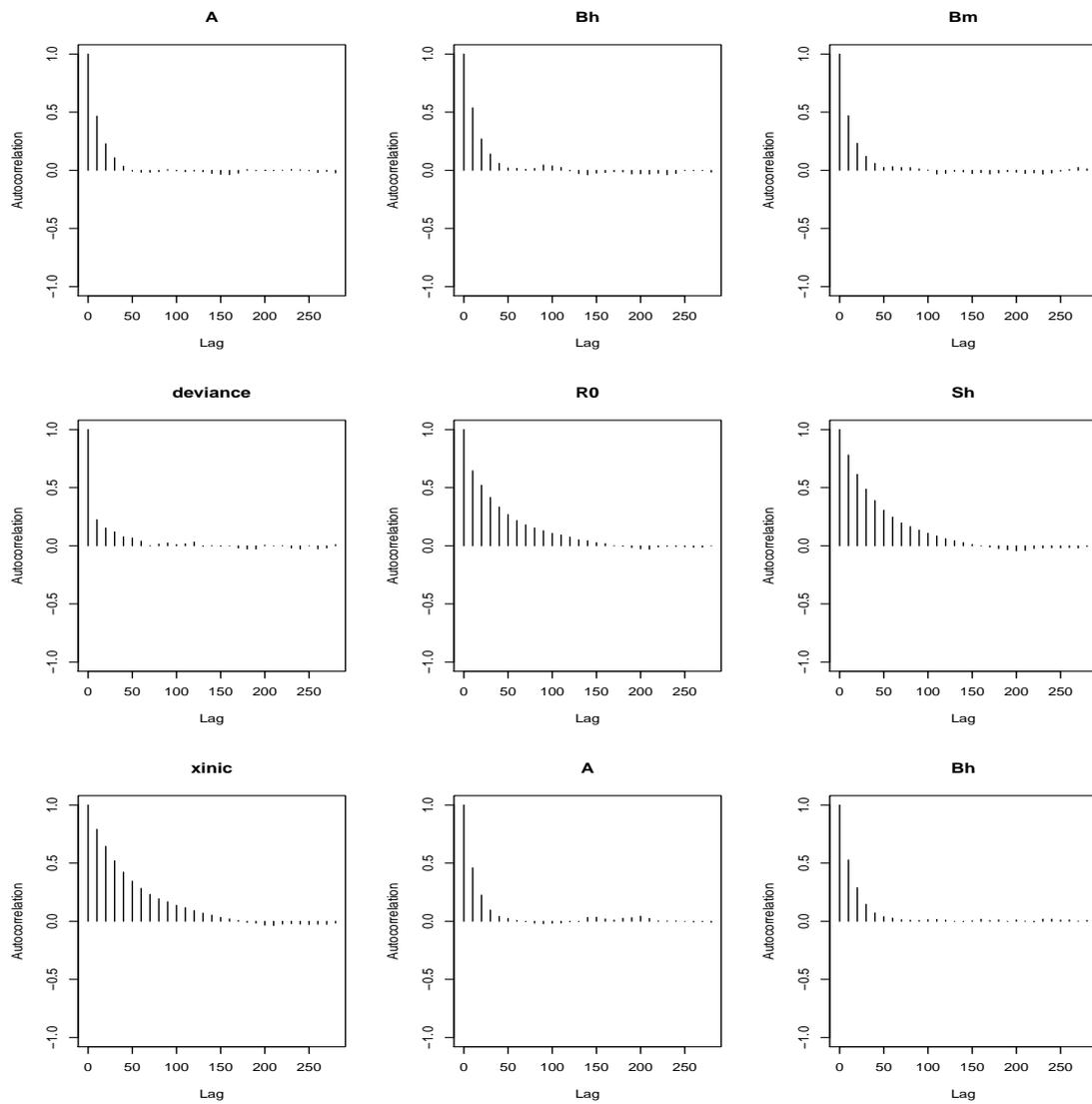


Figura 5.4: Autocorrelación para cada cadena de los parámetros del modelo *Ross – Macdonald*. Parte I.

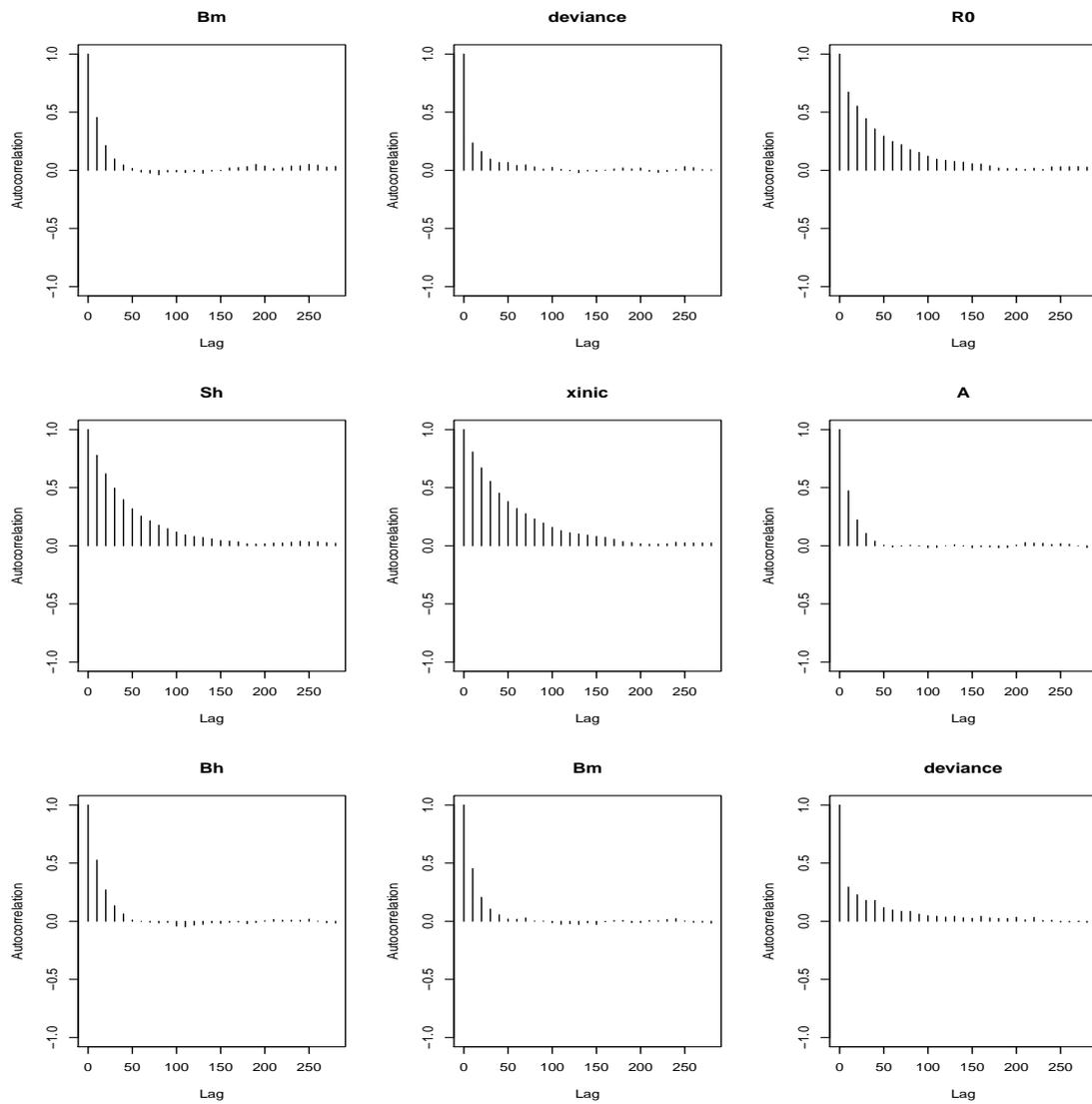


Figura 5.5: Autocorrelación para cada cadena de los parámetros del modelo *Ross – Macdonald*. Parte II.

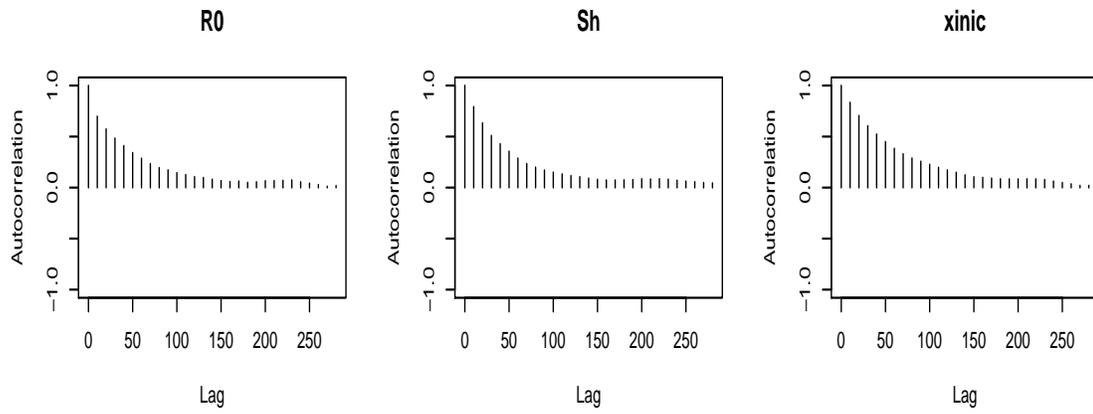


Figura 5.6: Autocorrelación para cada cadena de los parámetros del modelo *Ross – Macdonald*. Parte III.

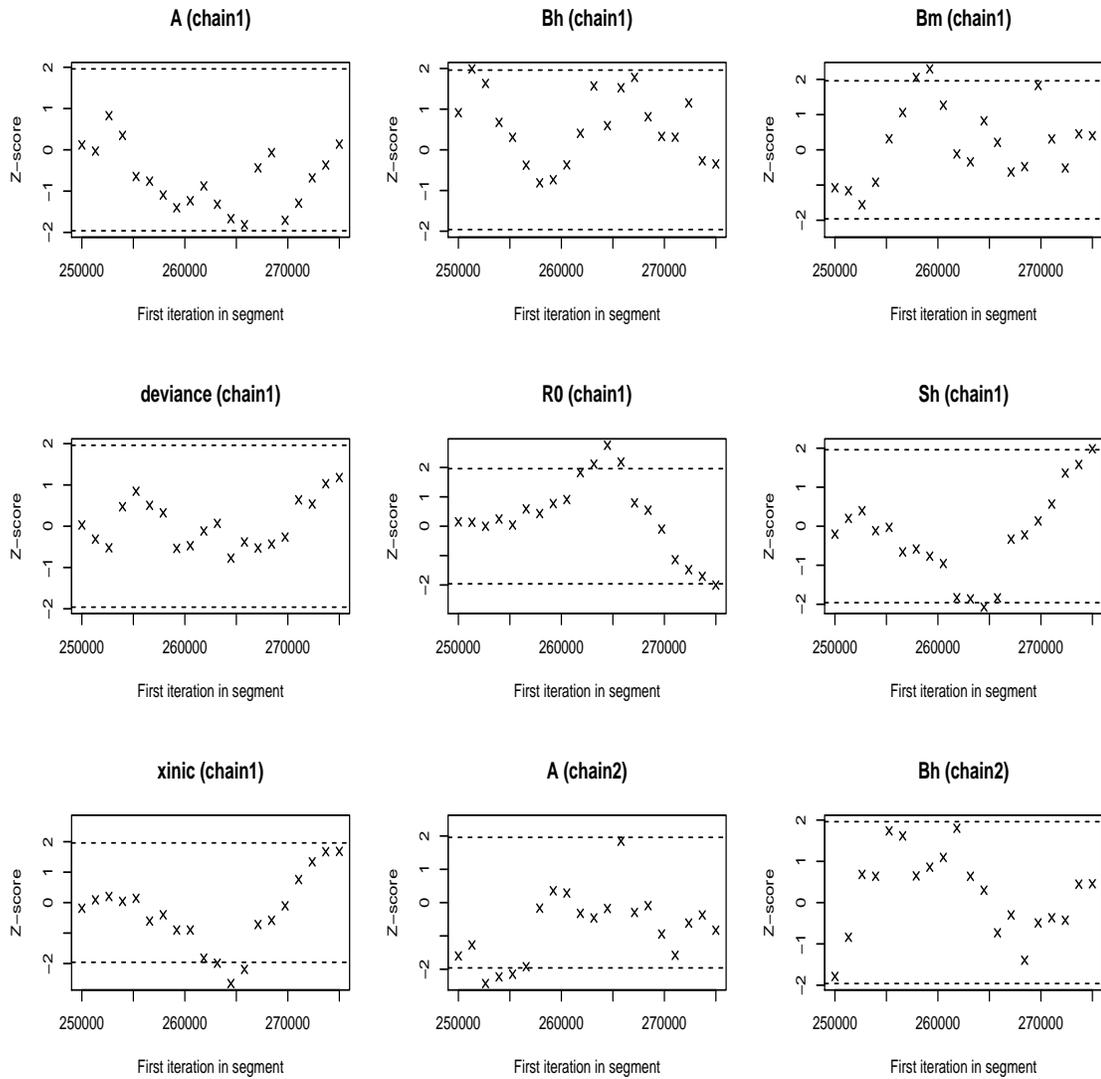


Figura 5.7: Diagnóstico de convergencia de Geweke para cada cadena de los parámetros del modelo de *Ross – Macdonald*. Parte I.

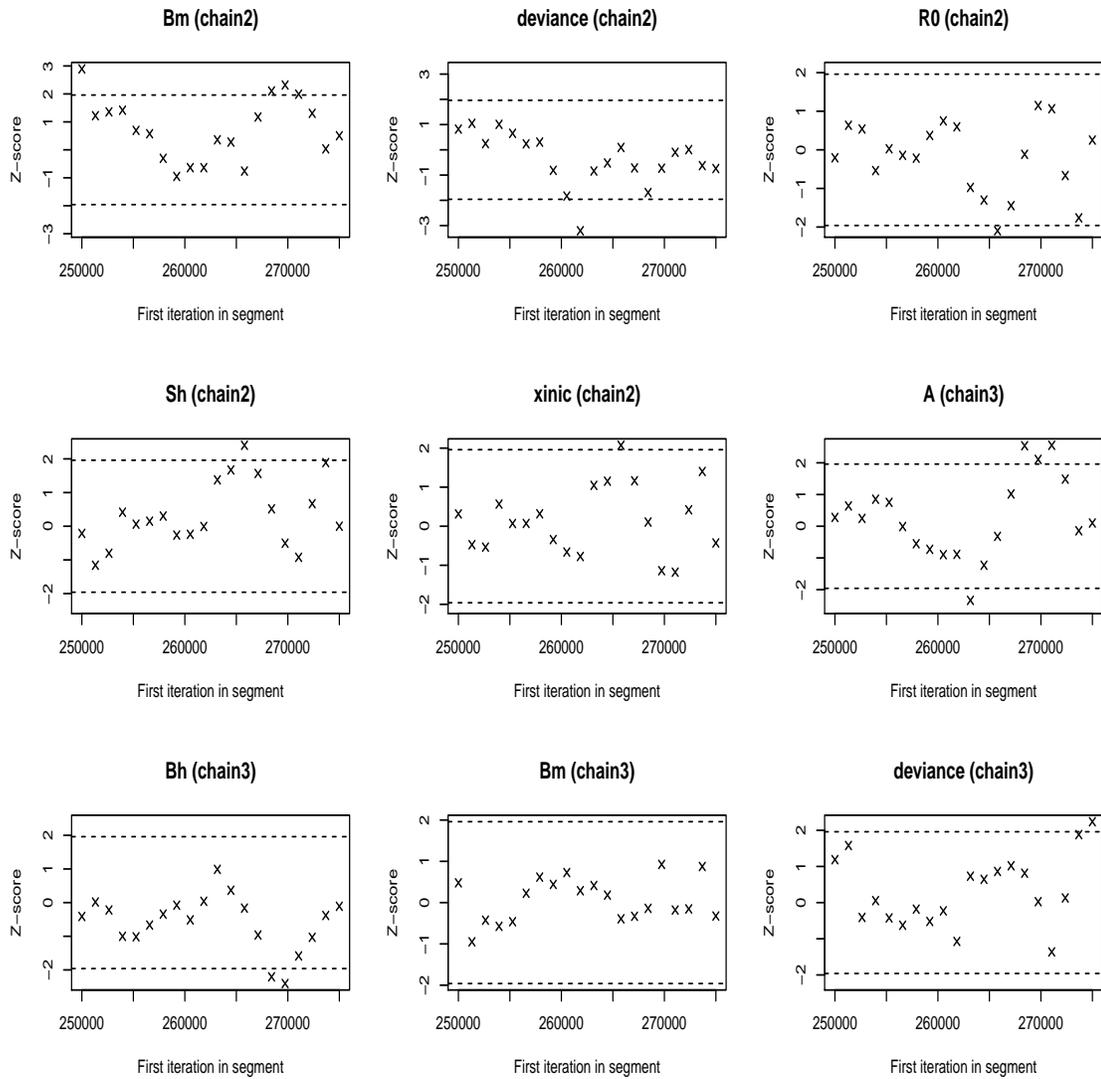


Figura 5.8: Diagnóstico de convergencia de Geweke para cada cadena de los parámetros del modelo de *Ross – Macdonald*. Parte II.

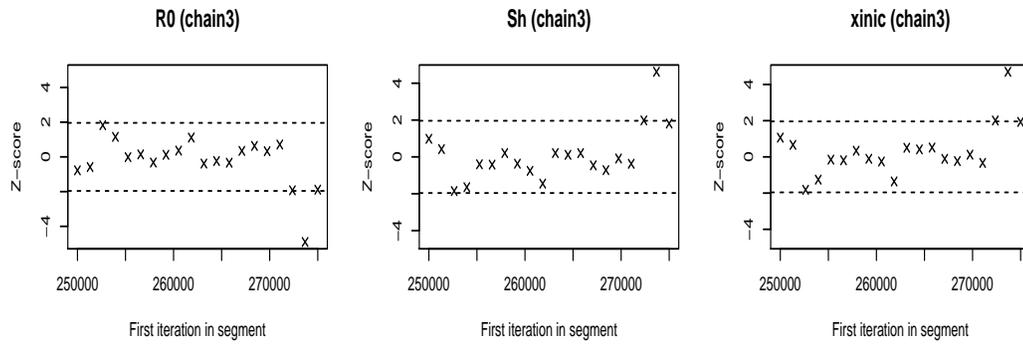


Figura 5.9: Diagnóstico de convergencia de Geweke para cada cadena de los parámetros del modelo de *Ross – Macdonald*. Parte III.

Anexo 2

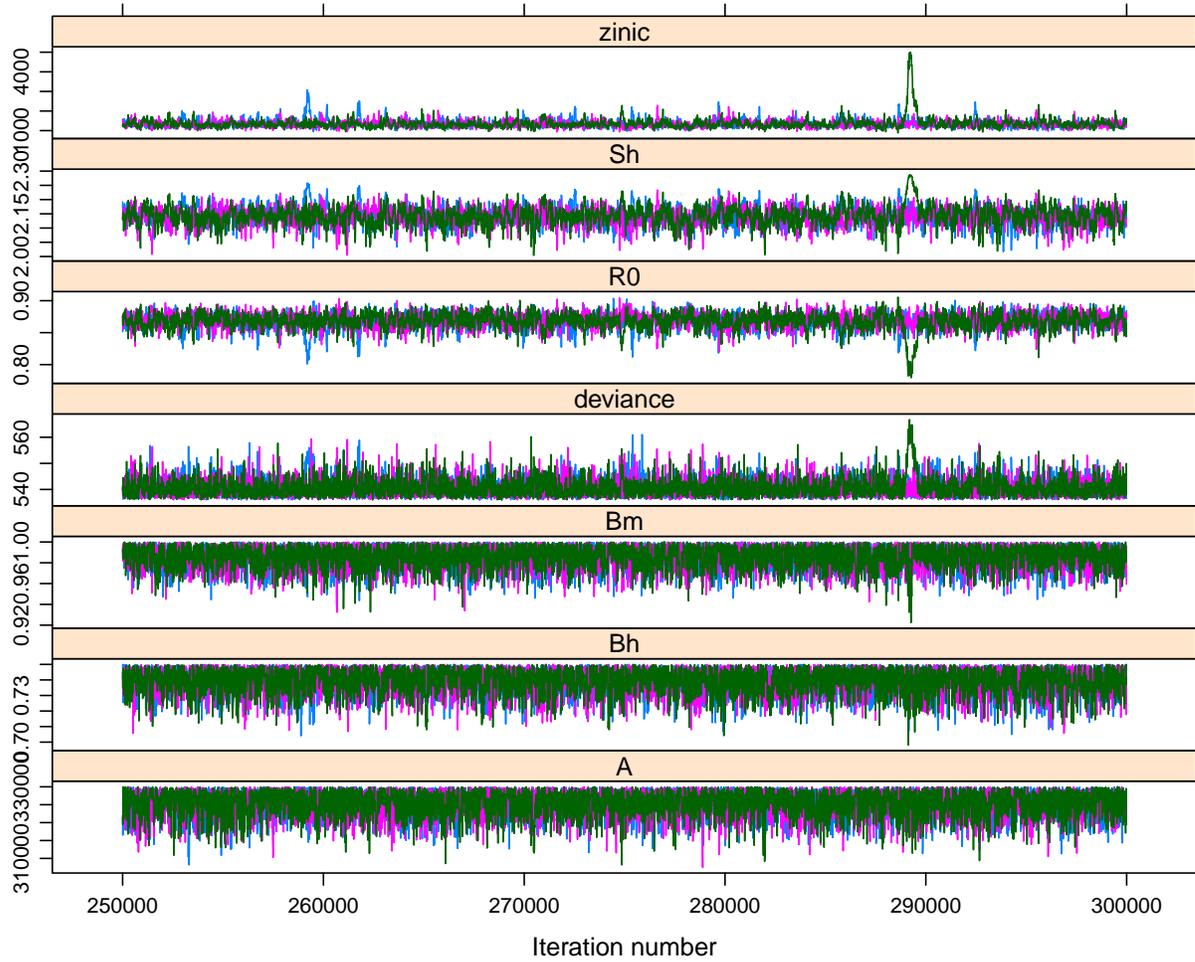


Figura 5.10: Mezcla de las cadenas de los parámetros del modelo *Vector – Hospedero*.

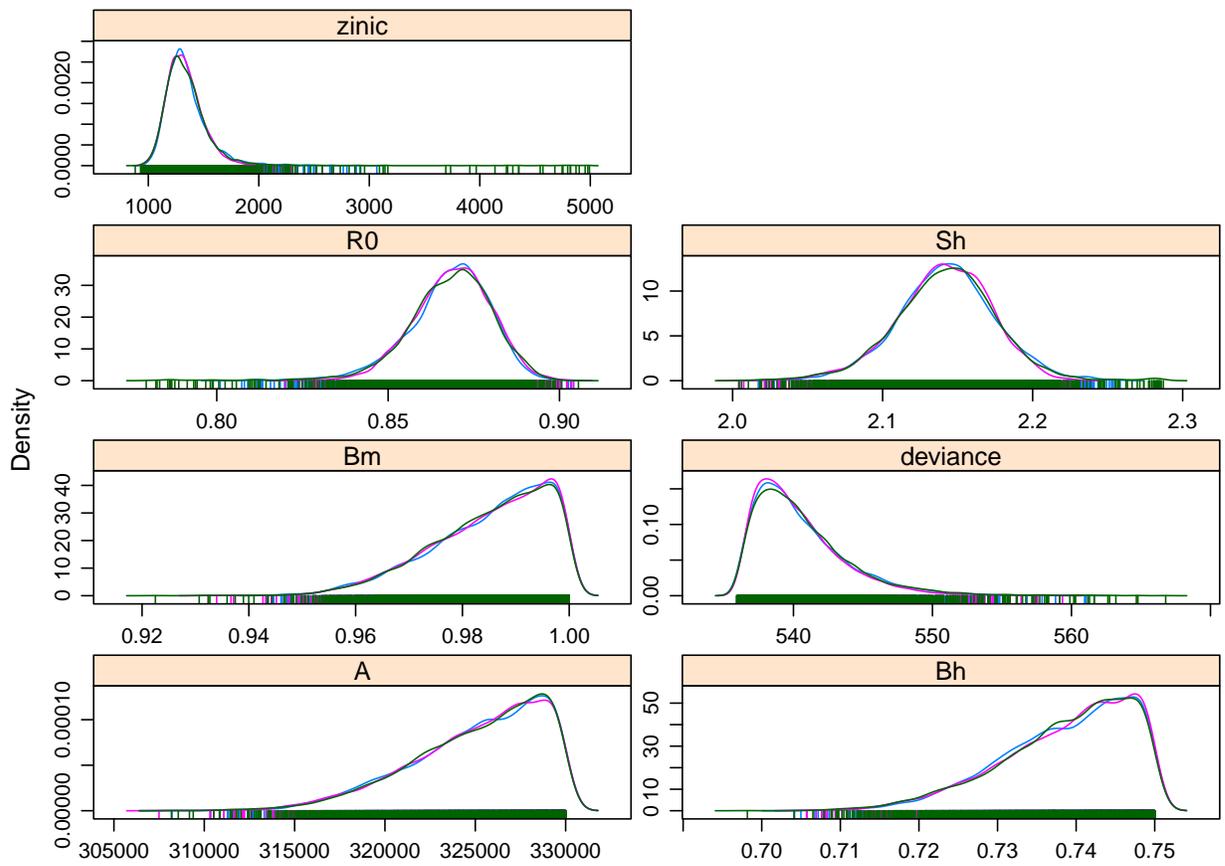


Figura 5.11: Distribución a posteriori de las cadenas de los parámetros del modelo *Vector – Hospedero*.

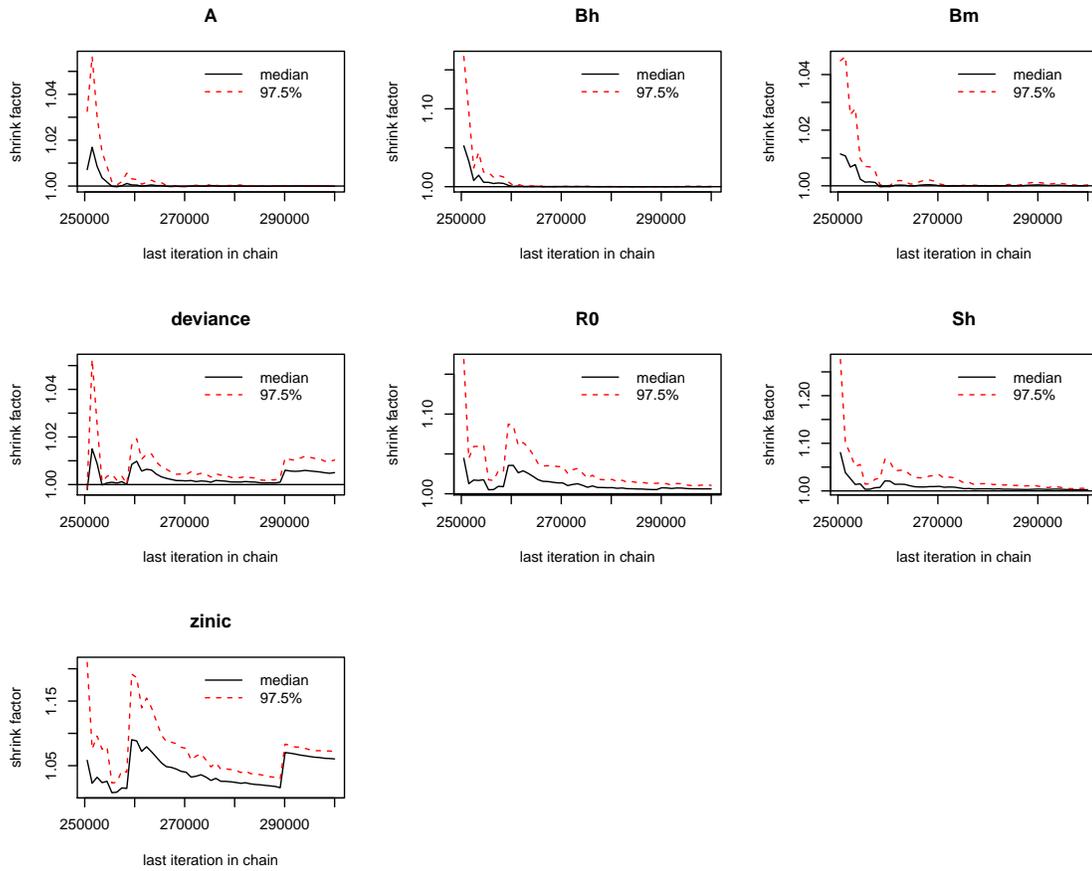


Figura 5.12: Diagnóstico de Gelman - Rubin de los parámetros del modelo *Vector - Hospedero*.

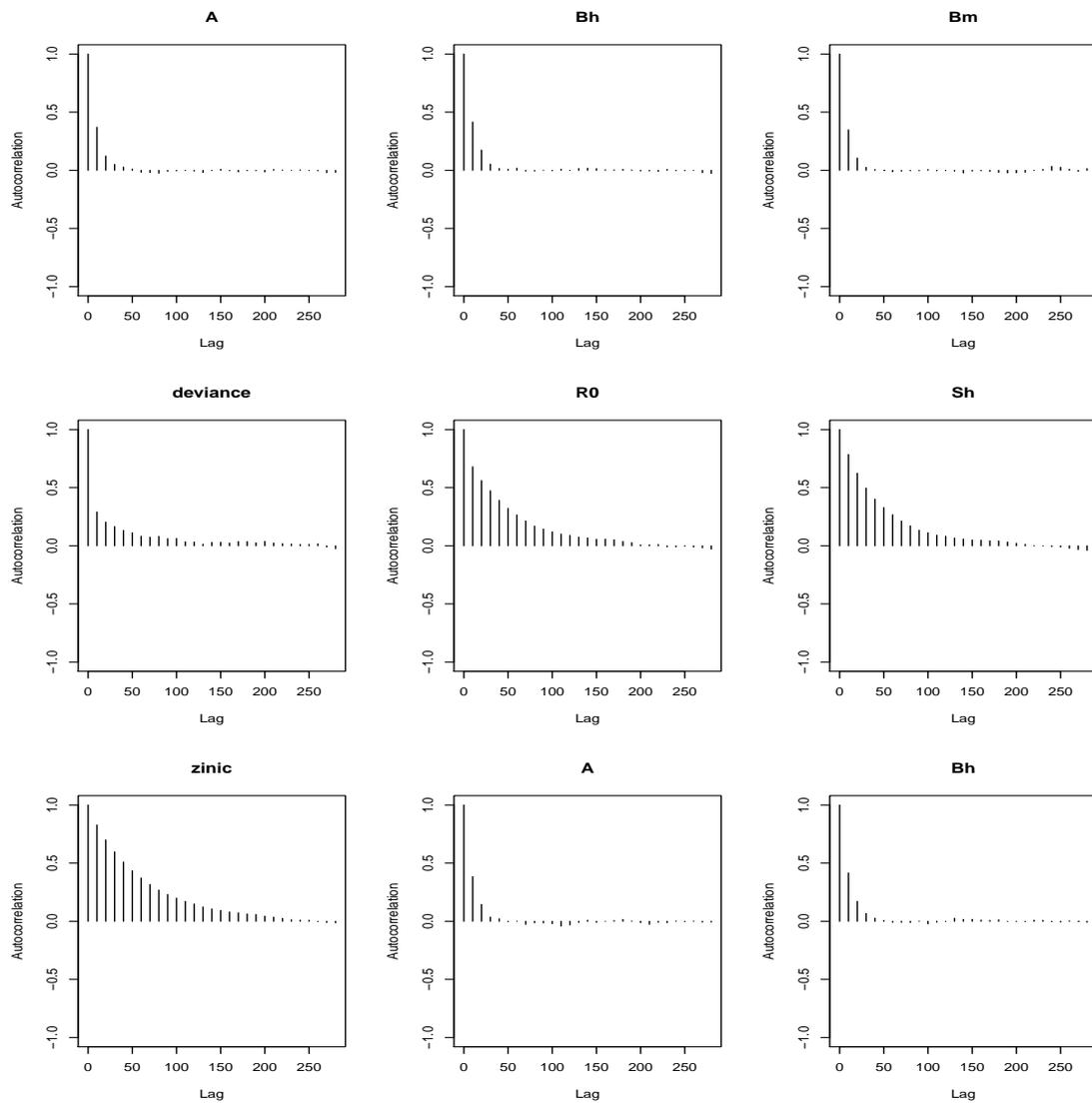


Figura 5.13: Autocorrelación de las cadenas de los parámetros del modelo *Vector – Hospedero*.

Parte I.

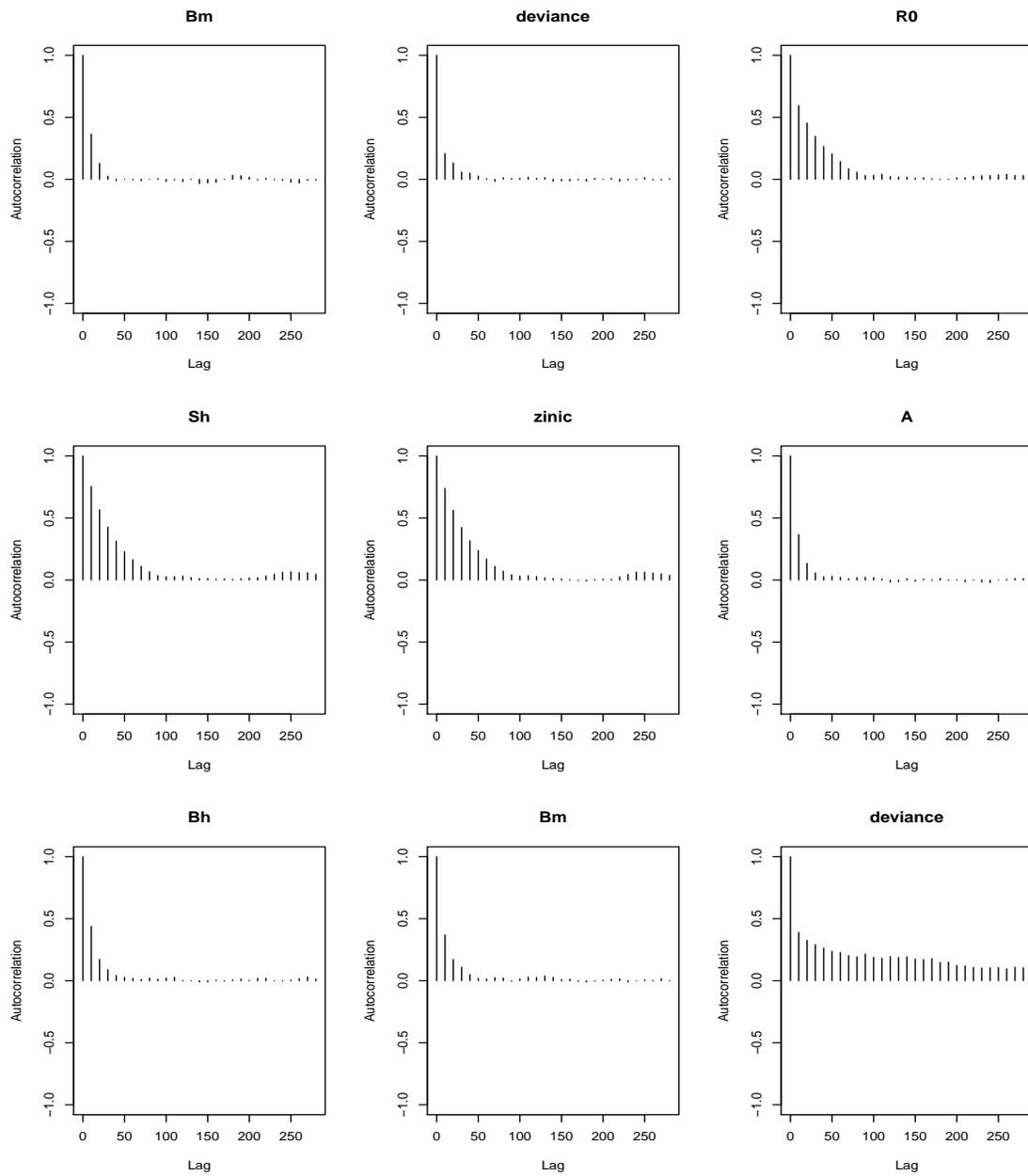


Figura 5.14: Autocorrelación de las cadenas de los parámetros del modelo *Vector – Hospedero*.
Parte II.

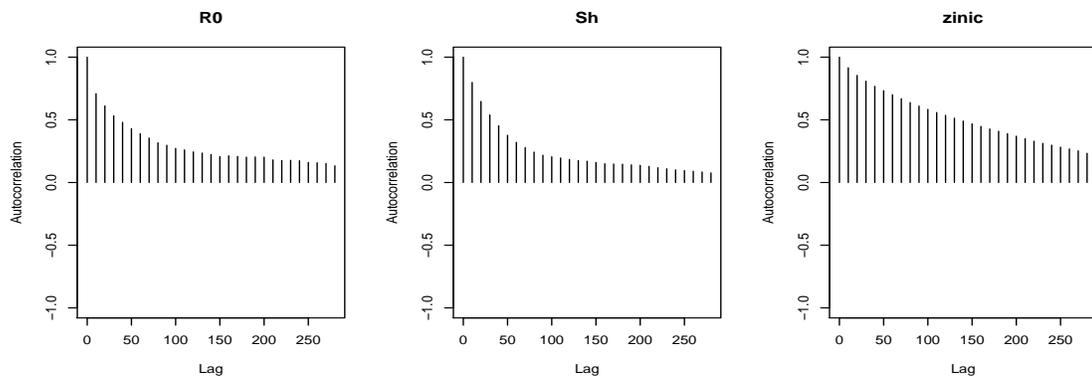


Figura 5.15: Autocorrelación de las cadenas de los parámetros del modelo *Vector – Hospedero*.
 Parte III.

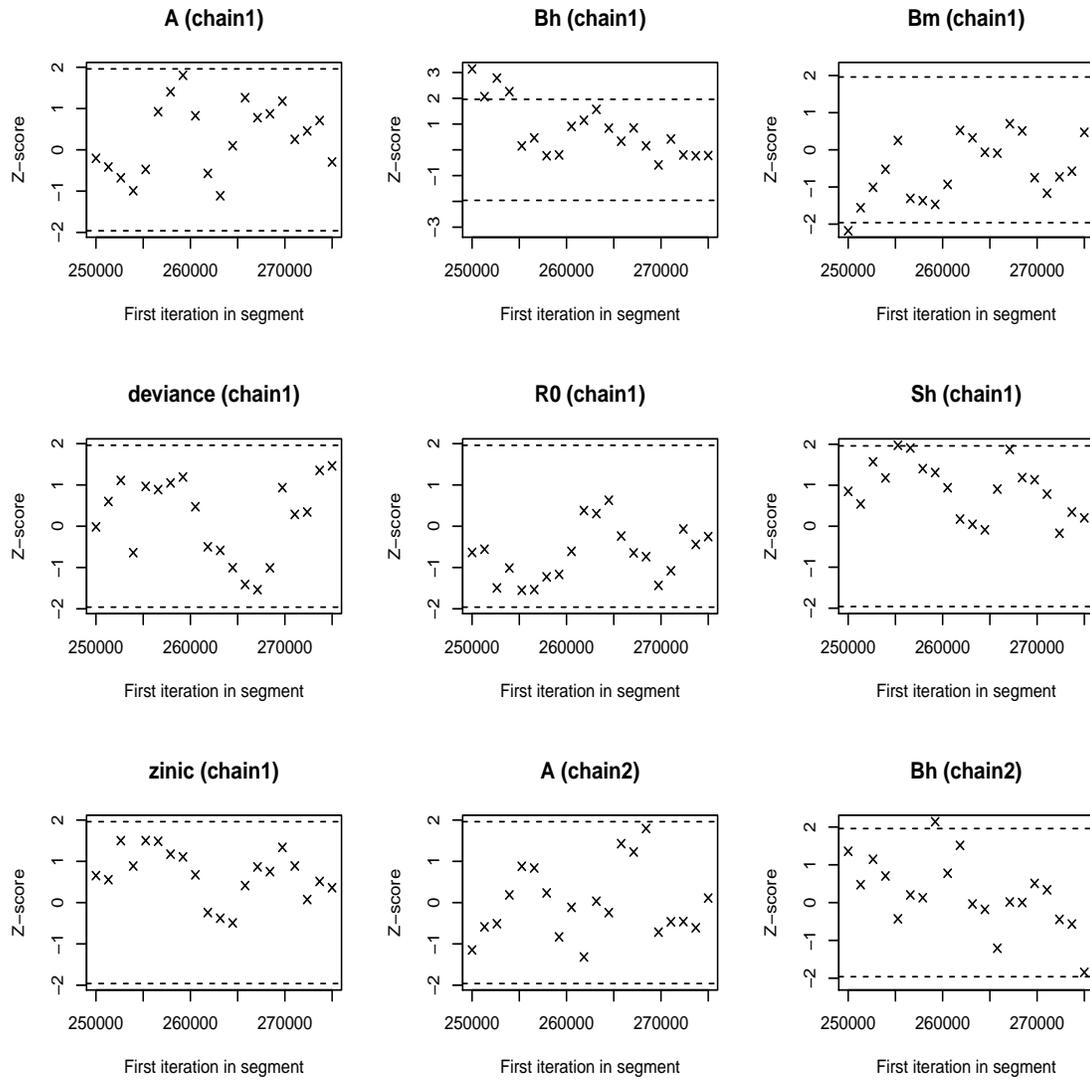


Figura 5.16: Diagnóstico de Geweke de las cadenas de los parámetros del modelo *Vector – Hospedero*. Parte I.

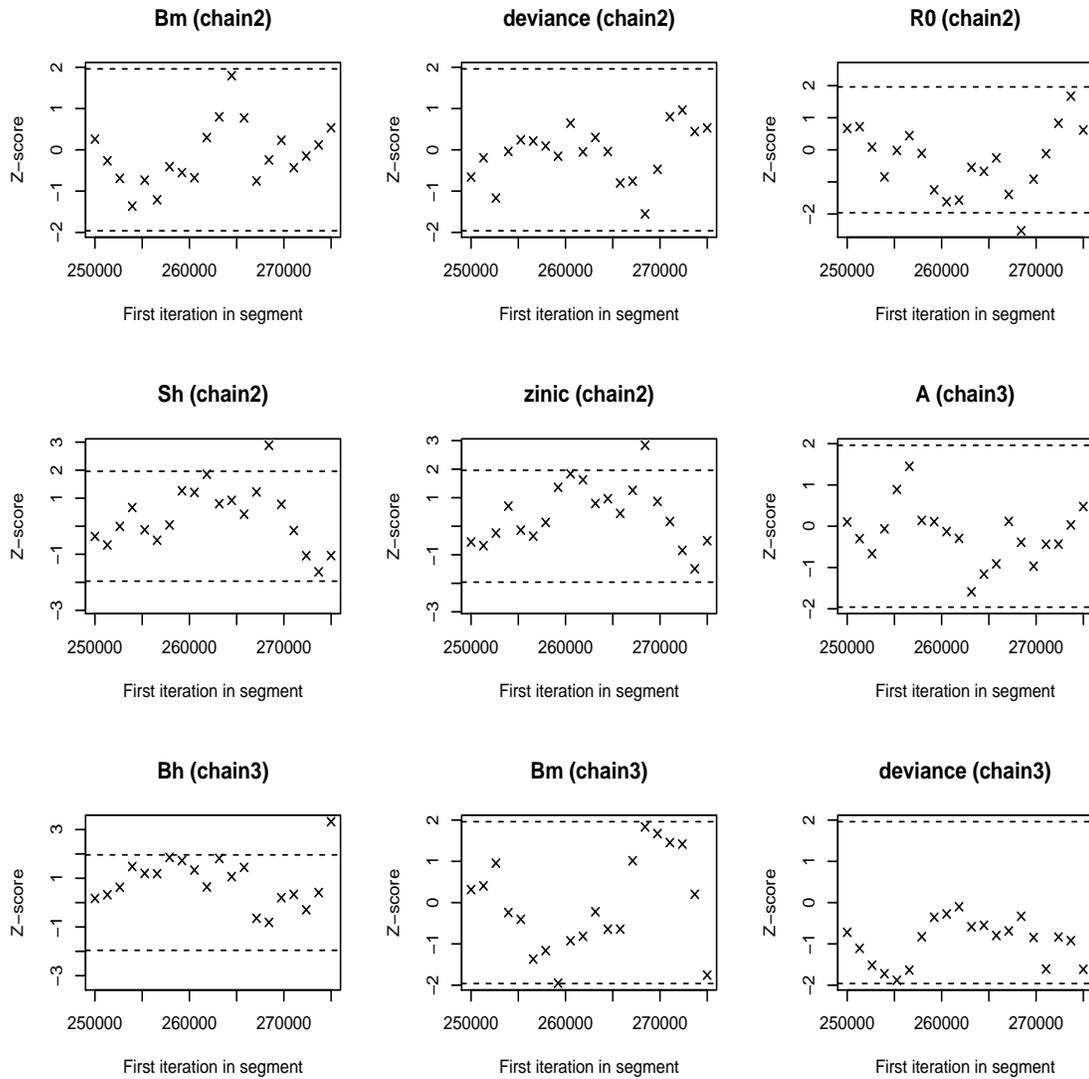


Figura 5.17: Diagnóstico de Geweke de las cadenas de los parámetros del modelo *Vector – Hospedero*.

Figura 5.18: Diagnóstico de Geweke de las cadenas de los parámetros del modelo *Vector – Hospedero*. Parte II.

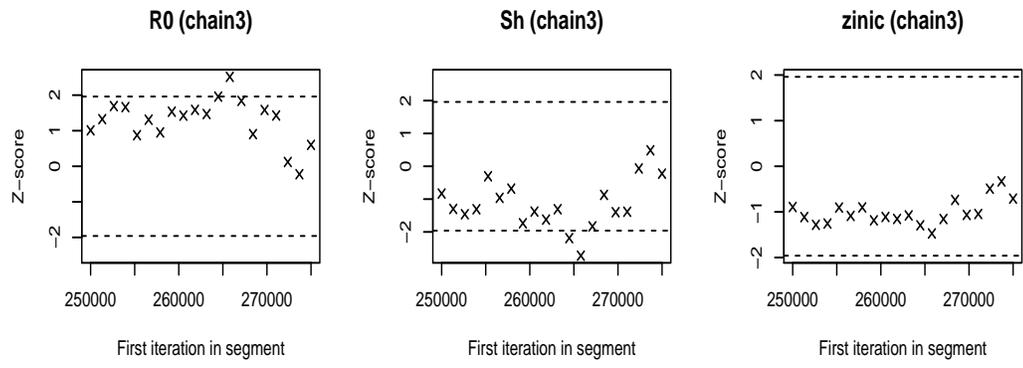


Figura 5.19: Diagnóstico de Geweke de las cadenas de los parámetros del modelo *Vector – Hospedero*. Parte III.