

UNIVERSIDAD CENTROCCIDENTAL LISANDRO ALVARADO DECANATO
DE CIENCIAS Y TECNOLOGÍA



Modelos jerárquicos bayesianos espaciales en Epidemiología Agrícola

Ing. Nora Coromoto Monsalve Graterol, M.Sc.

Barquisimeto, Junio 2011.

Modelos jerárquicos bayesianos espaciales en Epidemiología Agrícola

Por Nora Coromoto Monsalve Graterol

Trabajo de Ascenso para optar a la Categoría de Agregado en el Escalafón del Personal
Docente y de Investigación

UNIVERSIDAD CENTROCCIDENTAL LISANDRO ALVARADO
DECANATO DE CIENCIAS Y TECNOLOGÍA
Barquisimeto, 2011.

Agradecimientos

A Dios que bajo sus distintas formas me ilumina y cuida de todos mis pasos.

A mi amado esposo por su apoyo incondicional, sus palabras de aliento, su espera, su paciencia y amor...

A mis padres y mi querida hermana, gracias por su amor infinito a lo largo de toda mi vida. Y al resto de familiares y amigos, gracias.

A la familia de Arnaldo (en especial a Doña Ana, Eyra y Wilmer), gracias por su apoyo y comprensión.

A la UCLA y a su personal, por su disposición y ayuda institucional.

Índice general

Índice de tablas	3
Índice de figuras	5
1. Introducción	1
1.1. Planteamiento del problema	4
1.2. Justificación e importancia	7
1.3. Objetivos	8
1.4. Motivación	8
1.5. Estructura del trabajo y metodología empleada	10
2. Modelos jerárquicos espaciales	12
2.1. Modelos Lineales Generalizados (GLM)	12
2.2. Definición de modelos jerárquicos	13
2.3. Modelos jerárquicos espaciales	14
2.4. Modelos jerárquicos bayesianos espaciales	21
3. Inferencia	27
3.1. Inferencia bajo el enfoque Bayesiano	27
3.2. Métodos Markov chain Monte Carlo (MCMC)	28
3.3. Inferencia en automodelos	34
3.4. Inferencia Bayesiana con MCMC	35
3.5. Criterios para selección de modelos	37
4. Análisis espacial de datos en Epidemiología Agrícola	43
4.1. Distribuciones previas para los parámetros del modelo	44
4.2. Estudio de Modelos Lineales Generalizados (GLM)	45
4.3. Estudio de modelos jerárquicos bayesianos espaciales	52
4.4. Estudio de modelos jerárquicos bayesianos espaciales dinámicos	60

ÍNDICE GENERAL

5. Conclusiones y Líneas Futuras	67
Apéndices	69
Bibliografía	111

Índice de tablas

1.1. Incidencia de CTV (%) en Municipios de CV en 1989, 1995 y 1998.	5
1.2. Incidencia de CTV en la Parcela de Burriana.	6
3.1. Criterios de Jeffreys sobre el BF para decidir sobre H_0	39
4.1. Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 10. . .	46
4.2. Resultados GLM de las modelizaciones propuestas con x_1 y x_2 a una distancia 10. . .	46
4.3. Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 20. . .	47
4.4. Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 30. . .	47
4.5. Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 40. . .	47
4.6. Probabilidades para los árboles (70, 75, 81) usando las estimaciones GLM a una distancia 10.	48
4.7. Estimaciones GLM usando OpenBUGS a distancia 10.	49
4.8. Probabilidades para los árboles (70, 75, 81) usando las estimaciones GLM obtenidas con OpenBUGS a distancia 10.	49
4.9. Autocorrelación de α y β para la modelización propuesta usando CODA.	51
4.10. Diagnóstico Heidelberger para Modelo GLM.2 usando CODA.	51
4.11. Resumen con la deviance para los MJBE (Año 95), considerando la primera estructura propuesta.	54
4.12. Resumen con la deviance para los MJBE (Año 96), considerando la primera estructura propuesta.	54
4.13. Resumen con la deviance para los MJBE (Año 97), considerando la primera estructura propuesta.	54
4.14. Resumen con la deviance para los MJBE (Año 98), considerando la primera estructura propuesta.	54
4.15. Resumen con la deviance para los MJBE (Año 95), considerando la segunda estructura propuesta.	55
4.16. Resumen con la deviance para los MJBE (Año 96), considerando la segunda estructura propuesta.	55

4.17. Resumen con la deviance para los MJBE (Año 97), considerando la segunda estructura propuesta.	55
4.18. Resumen con la deviance para los MJBE (Año 98), considerando la segunda estructura propuesta.	55
4.19. Estimaciones para el Modelo MJBE.1, correspondiente al año 97.	56
4.20. Probabilidades de algunos árboles para el Modelo MJBE.1, correspondiente al año 97.	57
4.21. Diagnóstico Geweke para α y β en el modelo MJBE.1, correspondiente al año 97.	59
4.22. Diagnóstico Heidelberger para modelo MJBE.1, correspondiente al año 97.	59
4.23. Resumen con la bondad de ajuste para los modelos MJBED, de acuerdo a la primera estructura.	62
4.24. Resumen con la bondad de ajuste para los modelos MJBED, de acuerdo a la segunda estructura.	62
4.25. Estimaciones para el modelo MJBED.1, correspondientes a $i=4$	63
4.26. Algunas probabilidades estimadas para el modelo MJBED.1, correspondientes al año 97.	64
4.27. Diagnóstico Geweke para α y β en el modelo MJBED.1, correspondiente al año 97.	66
4.28. Diagnóstico Heidelberger para modelo MJBED.1, correspondiente al año 97.	66
5.1. Sintaxis en OpenBUGS para el modelo MJBE.1 correspondiente al año 95, usando previas gammas.	74
5.2. Sintaxis en OpenBUGS para el modelo MJBE.1 correspondiente al año 95, usando previas uniformes.	75
5.3. Sintaxis en OpenBUGS para el modelo MJBED.1 usando previas gammas.	77
5.4. Sintaxis en OpenBUGS para el modelo MJBED.1 usando previas uniformes.	79

Índice de figuras

1.1. Representación de la evolución de CTV en la Parcela de Burriana.	6
1.2. Evolución temporal de la enfermedad de CTV en la Parcela de Burriana.	6
4.1. Convergencia de los parámetros α y β usando OpenBUGS.	50
4.2. Convergencia para la deviance del Modelo GLM.1 usando OpenBUGS.	50
4.3. Diagnóstico Geweke para α y β del Modelo GLM.2 usando CODA.	51
4.4. Convergencia de α y β en el modelo MJBE.1, correspondiente al año 97.	58
4.5. Convergencia de la deviance en el modelo MJBE.1, correspondiente al año 97. . .	59
4.6. Convergencia para α y β en el modelo MJBED.1, correspondiente al año 97. . . .	65
4.7. Convergencia de la deviance en el modelo MJBED.1, correspondiente al año 97. .	66
5.1. Convergencia para la deviance en el modelo MJBE para el año 95, usando previas gammas después de 10000 iteraciones.	71
5.2. Convergencia para la deviance en el modelo MJBE para el año 95, usando previas gammas después de 20100 iteraciones.	71
5.3. Convergencia para la deviance en el modelo MJBE para el año, 95 usando previas gammas después de 30000 iteraciones.	72
5.4. Convergencia para la deviance en el Modelo MJBE para el año 95, usando previas uniformes para 10000 iteraciones.	73
5.5. Convergencia para la deviance en el Modelo MJBE para el año 95, usando previas uniformes para 18000 iteraciones.	73
5.6. Convergencia para la deviance en el modelo MJBED.1 usando previas gammas (10000 iteraciones).	76
5.7. Convergencia para la deviance en el Modelo MJBED.1 usando previas gammas (después de 25000 iteraciones).	76
5.8. Convergencia para la deviance en el Modelo MJBED.1 usando previas uniformes (para 10000 iteraciones).	78
5.9. Convergencia para la deviance en el Modelo MJBED.1 usando previas uniformes (para 20000 iteraciones).	78

5.10. Convergencia para α en el modelo MJBE.1, correspondiente al año 95 (10000 iteraciones).	80
5.11. Convergencia para β en el modelo MJBE.1, correspondiente al año 95 (10000 iteraciones).	81
5.12. Densidades estimadas para α , β y deviance del modelo MJBE.1, correspondiente al año 95.	81
5.13. Convergencia, test Gelman y Rubin y densidad estimada para α del modelo MJBE.1, correspondiente al año 96.	82
5.14. Convergencia, test Gelman y Rubin y densidad estimada para β del modelo MJBE.1, correspondiente al año 96.	82
5.15. Convergencia, test Gelman y Rubin y densidad estimada para deviance del modelo MJBE.1, correspondiente al año 96.	83
5.16. Convergencia, test Gelman y Rubin y densidad estimada para θ [árbol=40] del modelo MJBE.1, correspondiente al año 96.	83
5.17. Convergencia, test Gelman y Rubin y densidad estimada para ϕ [árbol=40] del modelo MJBE.1, correspondiente al año 96.	84
5.18. Convergencia, test Gelman y Rubin y densidad estimada para p[árbol=40] en el modelo MJBE.1, correspondiente al año 96.	84
5.19. Convergencia, test Gelman y Rubin y densidad estimada para α en el modelo MJBE.1, correspondiente al año 98.	85
5.20. Convergencia, test Gelman y Rubin y densidad estimada para β en el Modelo MJBE.1, correspondiente al año 98.	85
5.21. Convergencia, test Gelman y Rubin y densidad estimada para deviance en el modelo MJBE.1, correspondiente al año 98.	86
5.22. Convergencia, test Gelman y Rubin y densidad estimada para θ [árbol=40] en el modelo MJBE.1, correspondiente al año 98.	86
5.23. Convergencia, test Gelman y Rubin y densidad estimada para ϕ [árbol=40] en el modelo MJBE.1, correspondiente al año 98.	87
5.24. Convergencia, test Gelman y Rubin y densidad estimada para p[árbol=40] en el modelo MJBE.1, correspondiente al año 98.	87
5.25. Convergencia, test Gelman y Rubin y densidad estimada para α en el modelo MJBE.3, correspondiente al año 97.	88
5.26. Convergencia, test Gelman y Rubin y densidad estimada para β en el modelo MJBE.3, correspondiente al año 97.	88
5.27. Convergencia, test Gelman y Rubin y densidad estimada para deviance en el modelo MJBE.3, correspondiente al año 97.	89

ÍNDICE DE FIGURAS

5.28. Convergencia, test Gelman y Rubin y densidad estimada para ϕ [árbol=40] en el modelo MJBE.3, correspondiente al año 97.	89
5.29. Convergencia, test Gelman y Rubin y densidad estimada para p [árbol=40] en el modelo MJBE.3, correspondiente al año 97.	90
5.30. Convergencia, test Gelman y Rubin y densidad estimada para α en el modelo MJBED.3.	90
5.31. Convergencia, test Gelman y Rubin y densidad estimada para β en el modelo MJBED.3.	91
5.32. Convergencia, test Gelman y Rubin y densidad estimada para deviance en el modelo MJBED.3.	91
5.33. Convergencia, test Gelman y Rubin y densidad estimada para ϕ [año=3, árbol=40] en el modelo MJBED.3	92
5.34. Convergencia, test Gelman y Rubin y densidad estimada para p [árbol=40, año=3] en el modelo MJBED.3	92

Resumen

El trabajo propone el uso de modelos jerárquicos bayesianos espaciales en un contexto dinámico, como herramienta práctica en estudios epidemiológicos relacionados con enfermedades contagiosas en cultivos. Debido a que España es el país donde el virus de la tristeza (CTV) ha causado mayor destrucción de árboles injertados sobre naranja ácida, debido al efecto combinado de la introducción no controlada de material vegetal infectado pero sin síntomas y a la extensión del vector *Aphis gossypii*, principal agente transmisor del virus, se analizan los árboles cultivados en la parcela Burriana ubicada en Castellón (CV) durante los años 1993-1998. Durante este período el número de árboles enfermos aumentó significativamente. Los modelos propuestos consideran la influencia tanto de covariables como de efectos aleatorios (espaciales y no espaciales); efectos que recogen el impacto de factores de riesgo ocultos. Los ajustes de los modelos se obtienen mediante técnicas con cadenas Markov Monte Carlo (MCMC) y se comparan con el criterio de información de deviance (DIC). La metodología propuesta sugiere la existencia de un patrón de contagio entre árboles ubicados a distancias menores a 10 metros, así como un patrón de movimiento en el principal vector transmisor *A. gossypii* a estas distancias.

Palabras Claves: cadenas Markov Monte Carlo (MCMC); criterio de información de deviance (DIC); vector *Aphis gossypii*; virus de la tristeza (CTV).

Capítulo 1

Introducción

La industria de cítricos en España tiene un importante valor económico y su producción anual está cerca de 6 millones de toneladas métricas, sobre *c.* 285000 hectáreas. España es el cuarto productor más importante del mundo y principal exportador de cítricos frescos. Esta industria ha sido socialmente importante porque proporciona puestos de trabajo e ingresos a pequeños cultivadores. La Comunidad Valenciana (CV), formada por las provincias Alicante, Castellón y Valencia, es una de las principales productoras de cítricos en el mundo. En esta región al este de España a lo largo de la costa mediterránea, más de 90 millones de árboles son cultivados sobre *c.* 185000 hectáreas.

El virus de la tristeza de los cítricos (CTV; Familia: *Closteroviridae*; Género: *Closterovirus*) (Karasev et al., 1995) es el causante de una de las enfermedades más dañinas y destructivas de los agrios (Bar-Joseph and Lee, 1989); las pérdidas ocasionadas por ella se estiman en más de 100 millones de árboles injertados sobre naranjo amargo (Román et al., 2004) (unos 38 millones de árboles en América, más de 55 millones en la cuenca mediterránea, especialmente en España, y unos 5 millones en otras zonas), a lo que hay que sumar la baja calidad del fruto y pérdida de producción que se da en varios millones de árboles injertados sobre patrones tolerantes a tristeza que han sido infectados con aislados agresivos de CTV causantes de acanaladuras en la madera, ya sea en el patrón y/o en la variedad (Bar-Joseph et al., 1989; Rocha-Peña et al., 1995; Cambra et al., 2000a).

CTV se transmite por varias especies de pulgones (Hemiptera, Aphididae) de manera semipersistente. De ellos, el vector más eficaz de CTV a nivel mundial es *Toxoptera citricida* (Meneghini, 1946; Bennet y Costa, 1949 ; Costa y Grant, 1951). Sin embargo, *Aphis gossypii* es el vector principal en España, en Israel, en algunas zonas de California (USA) y en todos aquellos lugares donde *T. citricida* está ausente (Dickson et al., 1951; Bar-Joseph y Loebenstein, 1972; Raccach et al., 1976; Hermoso de Mendoza et al., 1984; Yokomi et al., 1989; Gottwald et al., 1996; 1997; Cambra et al., 2000a; Marroquín et al., 2004). Además, se han descrito otras especies afídicas

como vectores menos eficaces de CTV: *A. spiraecola* (Norman y Grant, 1954; Hermoso de Mendoza et al., 1984), *T. aurantii* (Norman y Grant, 1956; Hermoso de Mendoza et al., 1984), *Myzus persicae* (Varma et al., 1960), *A. craccivora* y *Uroleucon jaceae* (Varma et al., 1965). Por otra parte, en ensayos paralelos con los mismos aislados del virus, *T. citricida* transmite CTV con una eficacia 6-25 veces superior a la de *A. gossypii* (Yokomi et al., 1994).

T. citricida es probablemente originario de China, y de aquí debió dispersarse a los demás países del este y el sur de Asia, Australia y África subsahariana. Parece así mismo bastante probable que *T. citricida* pasase de Sudáfrica a Brasil y Argentina con el material vegetal que introdujo en estos países la tristeza (Moreno, 1995). Tras las epidemias de los años 1930-40 en Brasil y Argentina (en que murieron 30 millones de árboles), *T. citricida* avanzó lentamente hacia el norte de América hasta alcanzar en 1976 Venezuela, donde en diez años causó la muerte de 6 millones de árboles (Rocha-Peña et al., 1995). En 1989, *T. citricida* se detectó en Costa Rica (Lastra et al., 1991) y posteriormente en Belice (Pollard, 1997), en Guatemala (Palmieri, comunicación personal), y en Yucatán (Michaud y Álvarez, 2000), habiendo ocupado además las islas del Caribe, mientras que en 1995 llegó a los Estados Unidos, concretamente a Florida (Moreno, 1995).

En cuanto a la cuenca del Mediterráneo, pese a haber cultivado los cítricos durante muchos siglos, no tuvo lugar la introducción de la tristeza hasta la década de 1920, y además se hizo sin la presencia de *T. citricida*, de manera que sufrió el problema del virus CTV de forma relativamente atenuada y esencialmente centrado en España e Israel. Pero en 1994 se detectó *T. citricida* en la isla de Madeira, a las puertas del Mediterráneo (Fernandes y Cruz de Boelpaepe, 1994). Se hicieron repetidas prospecciones en Portugal continental sin encontrar el pulgón (Cruz de Boelpaepe y Ferreira, 1998), hasta que en 2003 se detectó accidentalmente por primera vez en el norte de Portugal y en 2005 se supo que también se había encontrado en el norte de España (en Asturias desde 2002 y en Galicia en 2004) (Ilharco et al., 2005), lo que representaba la introducción de *T. citricida* en la cuenca mediterránea, con la gravísima amenaza que conlleva para las citriculturas del área, tanto en calidad de plaga como en su papel de principal vector del virus de la tristeza, y muy especialmente de los aislados severos del virus contra los que no sirve la actual estrategia de patrones tolerantes a razas no tan agresivas de CTV.

Las principales áreas españolas productoras de cítricos están situadas en el este y el sur de la Península Ibérica, es decir, bastante alejadas de las zonas del noroeste donde se había detectado *T. citricida*, pero el peligro de propagación del pulgón era evidente, y debido a ello se abordaron diversas acciones en España. Por un lado se desarrollaron dos proyectos : “Control biológico preventivo frente a la introducción de *T. citricida*” (INIA, 2005-2008) y “Prospección y estudios sobre *T. citricida* en la cornisa cantábrica” (IVIA, 2006-07). Además se realizaron desde 2005 varias prospecciones en el norte de España y diversas actuaciones de otro tipo (reuniones, visitas, etc.). Los objetivos de todas estas acciones eran, en primer lugar, estudiar la situación de *T.*

citricida en el norte de España (distribución geográfica, ciclo biológico y dinámica poblacional, composición y dinámica de la fauna de enemigos naturales, búsqueda de huéspedes alternativos a cítricos, y prospección de CTV) y, en segundo lugar, realizar estudios comparativos con otros pulgones de cítricos (dinámica poblacional, composición y dinámica de sus enemigos naturales) en Valencia, la principal zona citrícola española (situada en el este de la Península Ibérica), en la que de momento no ha aparecido *T. citricida* (Hermoso de Mendoza A. et al., 2008).

Desde 1935, las estimaciones basadas en estudios y en el número de árboles que han sido replantados, indican que cerca de 40 millones de árboles (principalmente de naranja dulce (*Citrus sinensis*) y mandarina (*Citrus reticulata*)) injertados sobre naranja ácida (*Citrus aurantium*) han disminuido sus rizomas y han sido matados por CTV. Las pérdidas causadas en Argentina (más de 10 millones de árboles), Brasil (más de 6 millones de árboles) y California (USA) (c. 3 millones de árboles) han sido reportados por Bar-Joseph et al., (1989). Esto significa que España es el país donde la enfermedad ha causado mayor destrucción de árboles injertados sobre naranja ácida. La diseminación inicial de la enfermedad fue debido al efecto combinado de la introducción no controlada de material vegetal infectado pero sin síntomas y a la extensión de los vectores áfidos ineficientes *Toxoptera aurantii* y *Aphis spiraecola*. Entre 1935 y 1989 aproximadamente cerca de 20 millones de árboles fueron removidos a causa del CTV. La población de áfidos cambió drásticamente y el vector eficiente y predominante *Aphis gossypii* fue el principal responsable de la extensión de CTV. Desde 1989 esto ha obligado al retiro de 20 millones de árboles injertados sobre naranja ácida.

El primer brote de la enfermedad al sur de Valencia fue seguido por otros investigadores en 1962 y 1968 (Martí, 1977) en diferentes áreas de la Comunidad Valenciana (CV). Desde 1968 el Ministerio de Agricultura Español había prohibido los injertos de naranja dulce, mandarina y pomelo sobre naranja ácida. La aparición del virus, hizo que la industria cítrica Española cambiará el empleo de rizomas tradicionales por rizomas tolerantes al CTV. Grandes esfuerzos científicos y económicos tanto públicos como privados se han hecho en España desde 1975 para desarrollar acciones y estrategias que permitan manejar el problema CTV. Un Programa de Mejora de Variedad en Cítricos se comenzó en 1975 con los objetivos de recuperar “plantas sin virus” por injerto de brotes in vitro (Navarro et al., 1975) y establecer un banco de germplasm de plantas sanas cítricas. En 1983 estos objetivos fueron expandidos con la importación de citrus budwood de otros países (Navarro et al., 1984). Otras estrategias de control fueron incluidas para el desarrollo de métodos sensitivos de diagnósticos confiables y especificación de reactivos para la detección de CTV en material vegetal y estudios epidemiológicos para seguir la incidencia y extensión de la enfermedad.

Entre 1986 y 1989 se realizó una importante revisión en la CV (Cambra et al., 1988, 1990a) y de cada mil árboles, se inspecciono uno al azar. Se prepararon mapas y listas con los porcentajes de árboles infectados en diferentes provincias de la CV (Cambra et al., 1990a). Esta revisión fue

hecha por otros autores en 1992-1995 y 1997-1998, en las cuales más de 30000 árboles fueron analizados. El muestreo de 0.1 % de todos los árboles dio una indicación muy exacta de la incidencia de CTV en áreas grandes. Fueron necesarias pruebas sobre el 10 % de los árboles para proporcionar una mejor incidencia de CTV en pequeñas áreas. Además, todos los árboles localizados en áreas ecológicas diferentes (entre 0.5-1.0 hectáreas) fueron estudiados cada año para seguir la extensión de la enfermedad y evaluar la exactitud de los estudios realizados en otras provincias de la misma región.

La tabla 1.1 muestra el porcentaje de infección en las áreas de cítricos de la CV. El promedio aproximado de incidencia CTV fue 9 % en 1989, 35 % en 1995 y 42 % en 1998. El incremento en la provincia de Alicante fue menos rápido que en las otras dos provincias, probablemente debido al mayor número de árboles de limón cultivados en esta provincia. Solamente 0.2 % de árboles de limón estaban infectados en 1989 (Cambra et al., 2000a). En la provincia de Valencia la extensión de la enfermedad por tristeza fue más rápida. En algunos municipios entre ellos Moncada, la incidencia total aumentó dramáticamente del 4.8 % en 1989 a casi el 100 % en 1998. La extensión de CTV en la provincia de Castellón entre 1989 y 1998, alarmó a los cultivadores de esta zona, donde muchas variedades de Clementina son cultivadas. El promedio de incidencia de árboles infectados en esta provincia fue 6 % en 1989 y 31 % en 1998 (tabla 1.1). En Villarreal (Castellón), se encontró material vegetal infectado y el porcentaje de árboles infectados era muy alto comparado con otros municipios de la provincia donde la extensión de la enfermedad era principalmente debido a vectores áfidos (Cambra et al., 2000a).

1.1. Planteamiento del problema

El virus de la tristeza se transmite desde árboles infectados (que pueden mostrar síntomas o no) a otros árboles a través del injerto y de pulgones. Los controles sistemáticos del material vegetal empleado en los injertos constituyen la principal herramienta para impedir un rápido contagio de la enfermedad. Pero los métodos habituales de lucha contra los pulgones se han mostrado ineficaces para prevenir nuevos casos de árboles infectados (Cambra et al., 2000a).

La velocidad de difusión de la enfermedad ha aumentado preocupantemente en los últimos años debido entre otras causas al incremento de la población del pulgón *Aphis gossypii*, eficaz vector de transmisión. Este trabajo intenta analizar la difusión del virus CTV en su evolución espacial, considerando los árboles plantados durante los años 1994 a 1998 en la Parcela de Burriana. Se sabe que el pulgón visita los cítricos en primavera y otoño, durante el período de brotación, permaneciendo en ellos unas pocas semanas mientras los brotes están frescos. Dado que este es el tiempo de transmisión del virus, cada árbol es analizado anualmente a principios de invierno, empleando métodos ELISA con anticuerpos monoclonales específicos (ELISA convencional e inmunopresión ELISA) (Garnsey y Cambra, 1991).

Provincia	1989	1995	1998
<i>Municipios Alicante (n = 10)</i>			
Incidencia mínima	0.1	1.3	3.0
Incidencia máxima	14.0	39.5	43.5
Promedio de incidencia	3.0	14.0	20.0
<i>Municipios Castellón (n = 10)</i>			
Incidencia mínima	1.2	11.0	17.0
Incidencia máxima	20.4	75.0	83.0
Promedio de incidencia	6.0	28.0	31.0
<i>Municipios Valencia (n = 16)</i>			
Incidencia mínima	1.1	13.5	17.5
Incidencia máxima	70.4	98.0	99.9
Promedio de incidencia	17.0	63.0	75.0
Total árboles analizados	Aprox.66000	Aprox.23000	Aprox.7000

Tabla 1.1: Incidencia de CTV (%) en Municipios de CV en 1989, 1995 y 1998.

Varias parcelas de cultivo de diferentes variedades citrícolas han sido examinadas durante los años transcurridos desde su plantación. Los datos han sido recogidos en la parcela de Burriana, ubicada en Castellón. No se realizó durante el tiempo de estudio ningún cambio, arranque o reinjerto en esta parcela (“parcela controlada”). El material vegetal inicial usado se comprobó estaba libre del virus.

La parcela está constituida por 300 árboles de naranjo, plantados en 1993 y que permanecieron libres de CTV hasta 1994. Los resultados de la prospección se muestran en la Figura 1.1.

Año	Árboles infectados	Proporción de Infect.	Nuevos Casos	Sanos
1994	1	0.0	1	300
1995	7	0.02	6	299
1996	30	0.10	23	294
1997	50	0.17	20	277
1998	115	0.38	65	280

Tabla 1.2: Incidencia de CTV en la Parcela de Burriana.

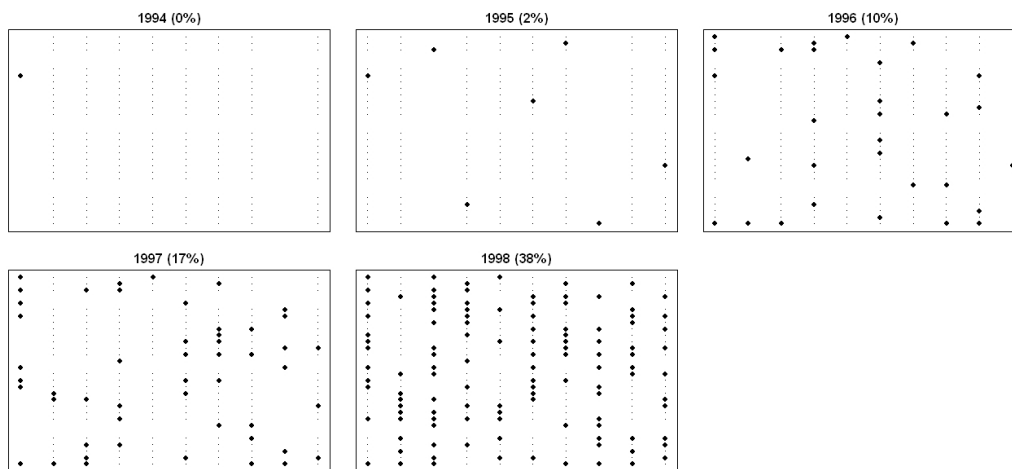


Figura 1.1: Representación de la evolución de CTV en la Parcela de Burriana.

La incidencia global del virus en la parcela, expresada como proporción de árboles infectados, permite visualizar el crecimiento de la enfermedad en los últimos años. Los datos se muestran en la tabla 1.2 y en la Figura 1.2.

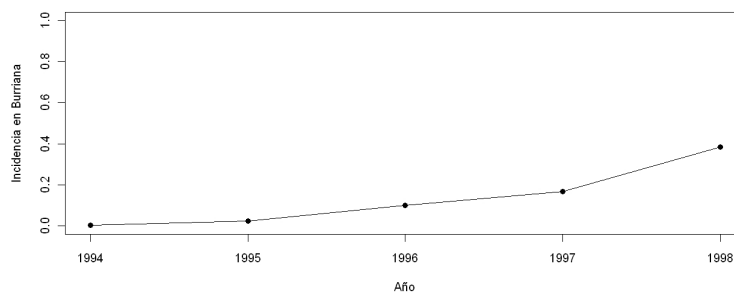


Figura 1.2: Evolución temporal de la enfermedad de CTV en la Parcela de Burriana.

1.2. Justificación e importancia

Los estudios epidemiológicos proporcionan información básica para desarrollar medidas de reducción en las tasas de infección de cualquier enfermedad y al mismo tiempo permiten proponer modelos estadísticos para el estudio de la incidencia, extensión y evolución espacial de una enfermedad contagiosa. Estos estudios han sido muy importantes en la reducción del impacto de la enfermedad causada por el virus CTV y ha ayudado a disminuir sus consecuencias sociales y económicas en la citricultura Española. La información epidemiológica obtenida de diferentes áreas de la CV ha facilitado el desarrollo de recomendaciones específicas para cultivadores (Cambra et al., 1990a) que implican el retiro de viejas plantaciones y el cambio a una moderna citricultura basada en rizomas injertados tolerantes a CTV. Todas estas políticas han resultado en un reemplazamiento progresivo y sistemático de los árboles afectados con el virus de tristeza que habían muerto lenta o rápidamente debido a la enfermedad. Más de 95 millones de árboles certificados libres de patógenos cítricos han sido plantados, principalmente aquellos injertados sobre Carrizo o Troyer citranges (*C. sinensis* x *Poncirus trifoliata*). Así, el 75 % de la citricultura Española y al menos el 90 % de la citricultura de la CV, es tolerante a la presencia aislada del CTV.

A pesar de los esfuerzos realizados para combatir la presencia de CTV y sus consecuencias, la enfermedad sigue presentándose, y el incremento de nuevos árboles infectados en estos últimos años (ver Figura 1.1) va en aumento. Es necesario continuar realizando estudios epidemiológicos con propuestas de nuevos modelos capaces de estudiar el comportamiento espacial de la enfermedad y que permitan distinguir la movilidad del agente transmisor (pulgón).

Los datos obtenidos de la parcela Burriana constituyen uno de los tantos datos que pueden encontrarse en estudios epidemiológicos de Agricultura. La información contenida en estos datos permitirá ilustrar el uso de la metodología propuesta sin llegar a resolver la problemática en la parcela.

A través de la metodología propuesta se intenta encontrar herramientas prácticas que ayuden a modelizar el comportamiento espacial del virus CTV y determinar el efecto de árboles infectados sobre árboles sanos. Los árboles plantados en la parcela Burriana inicialmente estaban libres del virus y durante el tiempo de estudio no se hizo ningún cambio (arranque o reinjerto). Por lo tanto, se espera que la aparición de nuevos árboles infectados este asociada a la presencia de pulgones infectados con el virus. Así, al analizar la distribución espacial de la enfermedad, se estará estudiando tanto el posible patrón de movimiento del vector *Aphis gossypii* como el incremento en su población.

1.3. Objetivos

Objetivo general:

Proponer y ajustar un modelo jerárquico bayesiano con estructura espacial, que permita estudiar el comportamiento epidemiológico de la enfermedad e identificar la dinámica del vector transmisor del virus CTV.

Objetivos específicos:

- Construir la covariable que recoge los factores ocultos con estructura espacial no incluidos explícitamente en los modelos formulados.
- Ajustar modelos GLM bajo el enfoque clásico y bayesiano por año, que incluyan la covariable construida.
- Definir la estructura de vecindad que se incluirá en el modelo condicional autoregresivo (CAR) que recoge la influencia espacial de árboles enfermos sobre árboles sanos.
- Ajustar modelos jerárquicos bayesianos espaciales usando técnicas de cadenas de Markov Monte Carlo (MCMC).
- Validar los modelos y seleccionar el mejor modelo (menor DIC).

1.4. Motivación

La modelización espacial de riesgos ha hecho uso repetidamente de distintas herramientas para conferir estructura de dependencia espacial a las observaciones objeto de modelización (Botella, 2010). Es común encontrar estudios de Disease Mapping aplicados en estudios de diversas áreas. Las áreas en las que más aplicaciones se encuentran son en medicina, epidemiología en salud pública, restauración de imágenes, epidemiología ecológica y veterinaria. En general, en esta metodología suele ser habitual considerar Modelos Lineales Generalizados (GLM) de Poisson para modelizar el número de casos observados en cada unidad de la región de estudio (Botella, 2010).

Son pocos los trabajos que utilizan modelos jerárquicos espaciales en epidemiología agrícola, sobre todo en casos donde la respuesta no es de conteo, desde una perspectiva bayesiana. En este trabajo nos concentraremos en este tipo de modelos, pero es posible encontrar un número importante de trabajos bajo el enfoque frecuentista, aunque en este campo no hayan sido extensivamente investigados. Trabajos de autores como por ejemplo, Cressie y Chan (1989), Besag et al. (1991), Clayton et al. (1993), Diggle et al. (1998), Leroux et al. (1999), Mugglin et al. (2000),

Knorr-Held y Raßer (2000), Kelsall y Wakefield (2002), Kelsall y Diggle (1995a, 1995b, 1998), Fernández y Green (2002), Christensen y Waagepetersen (2002) y Green y Richardson (2002). Richardson (2003) proporciona una revisión excelente de literatura en la que se proponen otros modelos residuales espaciales.

La metodología propuesta intenta encontrar modelos con los cuales sea posible modelizar la dinámica compleja de enfermedades de tipo contagioso en plantas. Como toda metodología presenta ventajas y desventajas que comentaremos de forma muy general.

Entre las ventajas a resaltar se puede mencionar que los modelos bayesianos espaciales aplicados en estudios epidemiológicos específicos, pueden ser de enorme ayuda en la vigilancia de enfermedades ya que permiten tener un conocimiento de la variabilidad del riesgo residual espacial y de aquella variabilidad no espacial en ausencia “de un punto caliente” ó área de alto riesgo. De manera similar, ayudarán a la regresión si se tiene conocimiento “previo” sobre la magnitud y las formas de la variabilidad no espacial y espacial de fondo. Esta información previa facilita el reconocimiento de dependencia en las variables, que en las técnicas de análisis convencionales sería inadecuado considerar.

La modelización de la dependencia espacial es muy difícil, ya que la posición espacial actúa como un sustituto de covariables no observadas. Por lo tanto, es necesario escoger un modelo apropiado para la dependencia espacial, incorporando indirectamente los efectos espaciales como covariables que tratamos de estudiar. Varios autores han propuesto diferentes formas de definir la estructura de vecindad, por ejemplo, Cressie y Chan (1989) asumen la estructura de vecindad como una función de la distancia entre centroides de área. Besag et. al (1991) sugiere un modelo que incluye efectos aleatorios espaciales y no espaciales y asigna a los efectos aleatorios espaciales una distribución condicional autoregresiva normal (CAR).

Los modelos bayesianos espaciales propuestos usarán los Modelos Lineales Generalizados (GLM) con vínculo logístico para modelizar la probabilidad de un árbol enfermar en función a covariables latentes (variables no observadas), variabilidad aleatoria no espacial y efectos aleatorios espaciales. A estas variabilidades se les asignarán distribuciones normales ya que con ellas es posible incorporar de forma sencilla dependencia entre observaciones y de esta forma resulta posible incorporar la estructura espacial de la localización o área de estudio a la covarianza de los datos (Botella, 2010).

Proponer modelos espaciales que incluyan sólo la estructura CAR sin considerar la variabilidad no espacial, podría conducir a inferencias erróneas si la mayoría de la variabilidad no es espacial. Leroux et al. (1999) demostró haciendo un estudio de simulación que si los datos eran realmente independientes, un modelo con sólo efectos CAR sin efectos no espaciales, produciría una seria sobrestimación en el parámetro de precisión del modelo CAR conduciendo a una pobre estimación de los coeficientes de regresión. Cuando se usa el modelo CAR, siempre deberán incluirse los efectos aleatorios no espaciales, ya que el modelo CAR no puede tomar una forma

restrictiva que permita incorporar la variabilidad no espacial.

En un modelo de efectos aleatorios la incorporación de la dependencia espacial proporciona una mejor especificación para explicar el modelo, teniendo como desventaja que no se pueda predecir el futuro o extrapolar a áreas sin datos observados. Este problema surge porque la dependencia espacial no puede ser generalizada, ya que cada unidad de muestreo observada tiene su propia estimación de parámetro.

Hoy en día gracias a metodologías como estas es posible abordar problemas complejos que combinan fuentes de información de diversa naturaleza y proveen herramientas prácticas para la toma de decisiones.

1.5. Estructura del trabajo y metodología empleada

En el capítulo 2, se describe el marco conceptual en el que se fundamenta el trabajo de investigación, se introducen los Modelos Lineales Generalizados (GLM), los modelos jerárquicos espaciales y se formulan los modelos más usados para datos espaciales en redes de localización (áreas pequeñas). Al final del capítulo se presentan los modelos jerárquicos bayesianos espaciales.

El capítulo 3, presenta los conceptos básicos de inferencia, introduce los métodos Markov chain Monte Carlo (MCMC) y cómo llevarlos a la práctica. Además, se presentan algunos métodos para la inferencia en campos aleatorios Markovianos basados en procedimientos MCMC desde un enfoque frecuentista. Se introduce la inferencia bayesiana y la forma habitual de hacer estimación a partir de la distribución posterior (principales algoritmos y software).

El capítulo 4, presenta los resultados obtenidos después de aplicar la metodología propuesta en los capítulos 2 y 3, a un conjunto de datos. Se presentan modelos ajustados desde la perspectiva frecuentista y bayesiana. Se comenta la bondad de ajuste de los modelos al incorporar y eliminar tanto los efectos de heterogeneidad como los efectos aleatorios espaciales. Por último, en el capítulo 5 se presentan las conclusiones y líneas futuras de investigación.

A continuación se describe de manera general, la metodología empleada en el presente trabajo:

- Se construyen covariables que recogen la estructura espacial oculta, considerando el efecto de los árboles enfermos en el pasado y el presente.
- Las covariables se incorporan en la estructura definida en el Modelo Lineal Generalizado (GLM) para determinar sus significancias, bajo el enfoque clásico y bayesiano.
- Se construye la estructura de vecindad que se incorpora en el modelo condicional autoregresivo CAR.

- Se ajustan modelos jerárquicos bayesianos espaciales por año, incorporando en el modelo, la influencia de alguna de las covariables construidas, junto con el efecto de heterogeneidad y el efecto aleatorio espacial.
- Se construyen modelos jerárquicos bayesianos espaciales dinámicos (conjuntos) y se comparan sus estimaciones con las halladas en los modelos ajustados por año bajo esta misma metodología.
- Los modelos propuestos se construyen al incorporar o eliminar la influencia de alguno de los efectos definidos en $\text{logit}(p_i) = \alpha + \sum_{j \sim i} \beta y_j + \theta_j + \phi_j$ (ver secciones 2.3.2 y 2.4).
- La estructura probabilística de los modelos jerárquicos bayesianos considerados, se completa asignando previas uniformes a los parámetros de precisión (Gelman, 2006).
- Se calculan las probabilidades finales de algunos árboles con el modelo de mejor ajuste.
- Se recomendará el mejor modelo (menor DIC).

Capítulo 2

Modelos jerárquicos espaciales

Este capítulo contiene la teoría que servirá de marco conceptual para el desarrollo del capítulo 4, se introducen los Modelos Lineales Generalizados (GLM), se presentan los modelos y técnicas espaciales desde una perspectiva jerárquica, en especial para el caso de datos en redes de localización. La última sección del capítulo está dedicada a los modelos jerárquicos bayesianos con estructura espacial. Se hace especial énfasis en la formulación de modelos jerárquicos bayesianos en redes de localización, ya que será el referente de aplicación y en el cual se fundamenta el capítulo 4.

2.1. Modelos Lineales Generalizados (GLM)

Los Modelos Lineales Generalizados constituyen una extensión de los modelos lineales y un caso ilustrativo de modelos jerárquicos. Los GLM comprenden aquellas distribuciones de familia exponencial uniparamétrica que recogen aditivamente los efectos fijos como una transformación monótona de la media. Esta familia permite modelizar una gran variedad de situaciones, con observaciones tanto continuas como discretas.

Un GLM se compone con una distribución concreta de la familia exponencial, los regresores que determinan el predictor lineal y la función vínculo, monótona y diferenciable, que define la relación entre la media y el predictor lineal. Estas componentes caracterizan el modelo y permiten una gran flexibilidad en la construcción del mismo (López, 2010). Además, la distribución empleada puede tener un parámetro secundario, denominado parámetro de dispersión, que afecta a la variabilidad. Los ejemplos de GLM más importantes y usados en la práctica son la regresión logística y la regresión de Poisson, además del propio modelo lineal normal. La estructura general de un GLM viene dada por

$$p(y_i|\lambda_i, \varphi) = \exp\left\{\frac{y_i\lambda_i - b(\lambda_i)}{a_i(\varphi)}\right\} + c(y_i, \varphi) \quad (2.1)$$

donde, y_i son variables aleatorias independientes con función de densidad dada por (2.1), λ_i es el parámetro natural o canónico, φ es un parámetro adicional de escala o dispersión, $a_i(\cdot)$, $b(\cdot)$ y $c(\cdot)$ son funciones específicas.

Si φ es conocido se tiene un modelo de la familia exponencial lineal y si es desconocido se tiene un modelo de dispersión exponencial.

En un GLM se quiere modelizar $\mu_i = E[y_i]$ en términos del predictor lineal $\beta' x_i$ formado por un conjunto de p variables $\beta' x_i = \beta_0 + \beta_1 x_i^1 + \dots + \beta_p x_i^p$.

Las componentes que determinan un GLM son:

1. Conjunto de n variables respuesta independientes, de una distribución de la familia exponencial.
2. Un vector de parámetros β y una matriz del modelo X , determinando el predictor lineal de cada variable $\beta' x_i$.
3. Una función vínculo monótona y diferenciable que define la relación entre μ_i y su predictor lineal $\beta' x_i$.

La elección del vínculo depende de la familia de distribuciones, del tipo de respuestas y de la aplicación.

La sobredispersión es un fenómeno que ocurre en aplicaciones con distribuciones que poseen varianza poco flexible, como el caso Binomial y el Poisson. Al añadir un parámetro de dispersión φ , se modifica la varianza $V[y] = a(\varphi)b''(\lambda)$. Esta varianza puede representar una heterogeneidad no observada o una correlación positiva entre respuestas individuales. Denominada también extravarianza (López, 2008).

Una importante extensión de este tipo de modelos es el Modelo Mixto Lineal Generalizado (GLMM), que incorpora en el predictor lineal un conjunto de variables latentes. Cuando estas variables provienen de un *proceso espacial* se obtiene un Modelo Espacial Lineal Generalizado (GLSM).

2.2. Definición de modelos jerárquicos

Son modelos probabilísticos para colecciones de variables formulados como combinaciones de diversas componentes denominadas niveles, capas o etapas. Esta estrategia es útil especialmente en la construcción de modelos complejos. Este tipo de modelización permite enlazar modelos provenientes de diversas ciencias (ambientales, médicas, sociales, biológicas, educativas y económicas, entre otras), combinando diferentes fuentes de información y empleando relaciones entre las variables estudiadas. Esta capacidad de adaptarse a situaciones complejas y gracias al

desarrollo de técnicas inferenciales asequibles mediante simulación (Möller, 2003) han permitido que se conviertan en una herramienta principal en la modelización estadística de problemas epidemiológicos.

La construcción de un modelo jerárquico se hace a través de las distribuciones condicionales, con las cuales, se construye un encadenamiento de dependencias lo que ayuda a flexibilizar y potenciar la conexión entre modelos complejos. Este tipo de metodología permite introducir no sólo factores de confusión indeterminados sino combinar fuentes de variabilidad y unir modelos parciales. La terminología para designar los elementos de un modelo jerárquico difiere según el enfoque empleado, frecuentista o Bayesiano.

2.3. Modelos jerárquicos espaciales

Se puede aplicar el término Geostatística basada en modelos, acuñado por Diggle, Tawn y Moyeed(1998) para denominar la aplicación de modelos estocásticos paramétricos explícitos y métodos formales de inferencia en problemas geostatísticos. La complejidad que se deriva de estas estructuras estocásticas dificulta la inferencia en este tipo de modelos. Estas dificultades pueden resolverse planteando el problema bajo el enfoque de modelos jerárquicos espaciales.

La incorporación de asociación espacial en alguna de las capas del modelo conduce a un proceso espacial jerárquico. Dicha asociación espacial puede venir modelizada mediante un proceso espacial continuo, un campo markoviano o un proceso puntual. Por ejemplo, la inclusión de un modelo autonormal genera un *modelo Gaussiano jerárquico espacial*, que está siendo ampliamente utilizado tanto con observaciones continuas como discretas. Siendo un caso concreto, el problema de suavización de mapas de riesgo de enfermedad, que puede ser empleado para estudiar las complicaciones que conlleva el análisis de estos modelos con estructura compleja. La inferencia que se hace a partir de los *modelos jerárquicos espaciales* dependerá de la perspectiva usada y del tipo de dato espacial involucrado.

Los datos espaciales se pueden clasificar en tres grupos fundamentales según el contexto de observación del que provienen: observaciones de un fenómeno continuo en el espacio (datos geoestadísticos), datos en una red fija de localizaciones y datos de sucesos que ocurren en el espacio proporcionando un conjunto aleatorio de puntos llamado patrón puntual (López, 2010). Estos tipos de datos diferenciados dan origen a formas distintas de modelización, y por tanto, de análisis estadístico (Cressie, 1993). En este trabajo sólo se presentará y aplicará la teoría relacionada con datos en redes de localización. Sin embargo, se presenta de manera muy general los elementos más sobresalientes a destacar en análisis que involucran datos geoestadísticos o de patrón puntual.

2.3.1. Datos Geoestadísticos

El análisis geostatístico más sencillo consiste en la estimación de los parámetros del modelo y la sustitución de dichas estimaciones en las ecuaciones del predictor como si fueran los verdaderos valores. Esto tiende a ser optimista en el sentido que conduce a una subestimación de la incertidumbre de predicción, ignorando la variabilidad entre las estimaciones de los parámetros y los verdaderos valores desconocidos. Esta aproximación se conoce como kriging y es ampliamente manejada (Banerjee et al., 2004).

La formulación básica de un proceso estocástico se concreta a la situación espacial tomando como conjunto de índices una determinada región continua D del espacio.

$$\{Y(s) : s \in D\}$$

Un proceso espacial estacionario es de la forma:

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + w(\mathbf{s}) + \epsilon(\mathbf{s}) \quad (2.2)$$

donde, la estructura de la media $\mu(\mathbf{s}) = x^T(\mathbf{s})\beta$. Los errores están particionados en dos partes, una espacial y la otra no espacial. Los $w(\mathbf{s})$ son asumidos como realizaciones de un proceso espacial estacionario no centrado en 0, el cual captura la asociación espacial residual. Mientras que los $\epsilon(\mathbf{s})$ son considerados términos de error puro. Por tanto, los $w(\mathbf{s})$ introducen el alféizar (σ^2) y el rango (ϕ) parcial y los $\epsilon(\mathbf{s})$ adicionan al modelo el efecto pepita (τ^2).

Tomando en cuenta la definición (2.2), se puede formular un proceso estocástico espacial como sigue:

$$Y \sim N(\mu, \sigma^2 H(\phi) + \tau I) \quad (2.3)$$

donde $\mu = X\beta$ recoge la variabilidad a gran escala y H es una matriz de correlaciones de la forma $H_{ij} = \rho(s_i - s_j \text{ phi})$ siendo ρ una función isotrópica de correlación en \mathbb{R}^2 indexada por los parámetros ρ .

2.3.2. Datos en redes de localización

Esta sección está dedicada al estudio de situaciones en que las observaciones provienen de un conjunto fijo de localizaciones. La predicción en otros puntos del espacio no tiene sentido cuando el fenómeno observado únicamente ocurre en las localizaciones de una red, o cuando es observado agregadamente.

Los modelos estadísticos para este tipo de datos tienen que expresar el hecho de que las observaciones próximas tienden a ser parecidas. Por tanto, deben incorporar la relación existente entre las observaciones de localizaciones vecinas. La especificación de estas relaciones a partir de las distribuciones condicionales origina los campos aleatorios markovianos.

Una red de localizaciones o retículo es una colección finita de localizaciones espaciales, que pueden estar distribuidas espacialmente de forma regular o irregular (López, 2010). Identificaremos una red de localizaciones con el conjunto de coordenadas de los puntos:

$$D \equiv \{(x_i, y_i) : i = 1, \dots, n\}$$

Para trabajar con modelos de esta naturaleza, es necesario definir un campo aleatorio Markoviano. Desde esta perspectiva se sabe que las distribuciones condicionales completas para Y_i dependen solamente de los vecinos $j \neq i$. Esta estructura condicional nos permite redefinir la relación de vecindad, a la vez que introducimos algunos conceptos necesarios en la especificación de los modelos. Una localización j es vecina de la localización i si la distribución condicional de Y_i , dados los valores en las demás localizaciones, depende funcionalmente de Y_j , para $j \neq i$.

Una clique es un conjunto de localizaciones en el que cada una de ellas es vecina de todas las demás. Un *campo aleatorio markoviano* es una medida de probabilidad cuyas distribuciones condicionales determinan una estructura de vecindad $\{V_i \subseteq D : i = 1, \dots, n\}$, es decir que para cada localización i se verifica que

$$P(y_i | y_{-i}) = P(y_i | y_{-i} v_i) \quad (2.4)$$

La estructura de probabilidad de un *campo aleatorio markoviano* viene caracterizada por su función potencial negativa, también denominada función negpotencial. Supongamos que $0 \in \zeta$. Definimos la función potencial negativa $Q(\cdot)$ como

$$Q(y) \equiv \log \left\{ \frac{P(y)}{P(0)} \right\}, \mathbf{y} \in \zeta \quad (2.5)$$

Conocer la función $Q(\cdot)$ equivale a conocer $P(\cdot)$, ya que

$$P(y) = \frac{\exp(Q(y))}{\sum_{z \in \zeta} \exp(Q(z))} \quad (2.6)$$

El teorema de Hammersley-Clifford (Clifford, 1990) establece que la función potencial negativa se descompone en sumas de términos correspondientes a las cliques definidas por la estructura de vecindad en la red de localizaciones. Sea k una clique y definimos:

$$y_k \equiv (y_i : i \in k)$$

$$U_k(y_k) \equiv G_k(y_k) \prod_{i \in k} y_i$$

Entonces,

$$Q(y) = \sum_{M_k} U_k(y_k) \quad (2.7)$$

Su importancia en la modelización espacial radica en el hecho de que la especificación condicional involucra una cantidad pequeña de funciones no nulas. El resultado en sentido opuesto

también es importante, y mantiene que la función potencial negativa conduce a una única función de probabilidad conjunta bien definida, siempre que se cumpla la condición de sumabilidad,

$$\sum_{y \in \zeta} \exp(Q(y)) < \infty$$

Con la teoría anterior, se pueden definir *automodelos* que pueden construirse como *modelos jerárquicos*. La construcción de los *campos aleatorios markovianos* permite establecer una gran variedad de *automodelos* para datos discretos y continuos.

La formulación de los *automodelos* propuesta por Besag (1974) permite modelizar *campos aleatorios markovianos*, tomando a las distribuciones condicionales en una familia exponencial lineal y limitando la interacción espacial a cliques de tamaño 2. En el caso discreto,

$$P(y_i|y_{-i}) = \exp[A_i(y_{-i})B_i(y_i) + C_i(y_i) + D_i(y_{-i})], \quad i = 1, \dots, n \quad (2.8)$$

donde $\{B_i(\cdot)\}$ y $\{C_i(\cdot)\}$ tienen formas específicas, siendo $\{A_i(\cdot)\}$ y $\{D_i(\cdot)\}$ funciones de los valores observados en las localizaciones vecinas de i . La forma en que se concreta la interacción entre las localizaciones vecinas es una implicación directa del *teorema Hammersley-Clifford*:

$$A_i(y_{-i}) = \alpha_i + \sum_{j=1}^n \beta_{ij} B_j(y_j), \quad i = 1, \dots, n, \quad (2.9)$$

donde $\beta_{ji} = \beta_{ij}$, $\beta_{ii} = 0$ y $\beta_{ik} = 0$ para to $k \notin V_i$. Para los automodelos, la función Q se puede simplificar en

$$Q(y) = \sum_{i=1}^n \{ \alpha_i y_i + C_i(y_i) + \sum_{1 \leq i < j \leq n} \beta_{ij} y_i y_j \} \quad (2.10)$$

Los automodelos ofrecen la posibilidad de incluir en el modelo la influencia de covariables relacionadas a las localizaciones como variables explicativas del proceso espacial. La formulación para el caso continuo es similar a (2.8), sólo que $P(y_i|y_{-i})$ es cambiada por la función de densidad $f(y_i|y_{-i})$.

1. **Caso discreto:** Se tienen los siguientes automodelos:

- **Autologístico:** En presencia de datos binarios la distribución condicional es necesariamente de forma logística. El modelo *autologístico* generaliza la regresión logística introduciendo la dependencia espacial entre las localizaciones. Esta definición puede verse como un *modelo jerárquico espacial*.

$$\begin{aligned} y_i|y_{-i} &\sim \text{Binomial}(1, p_i) \longmapsto \text{1er. nivel} \\ \text{logit}(p_i) &= \alpha + \sum_{j \sim i} \beta y_j \longmapsto \text{2do. nivel} \end{aligned}$$

Esto implica que la función potencial es

$$Q(y) = \alpha \sum_i y_i + \beta \sum_{j \sim i} y_i y_j$$

- **Autopoisson:** Cuando los datos espaciales surgen como conteos, la forma natural de modelizar el problema es empleando la distribución de Poisson. Esta definición también puede verse como un *modelo jerárquico espacial*.

$$\begin{aligned} y_i | y_{-i} &\sim \text{Poisson}(\lambda_i) \mapsto \text{1er. nivel} \\ \log(\lambda_i) &= \alpha + \sum_{j \sim i} \beta y_j \mapsto \text{2do. nivel} \end{aligned}$$

Conduce a una función potencial de la forma:

$$Q(y) = \alpha \sum_{i=1} y_i + \beta \sum_{j \sim i} y_i y_j - \sum_i \log(y_i!) \quad (2.11)$$

Una importante aplicación del *autopoisson* es la modelización de la incidencia regional de una determinada enfermedad. A menudo la distribución Poisson es una aproximación de la binomial, que puede ser empleada dando lugar al modelo *autobinomial*. Visto como un *modelo jerárquico* es de la forma:

$$\begin{aligned} y_i | y_{-i} &\sim \text{Binomial}(n_i \phi_i) \mapsto \text{1er. nivel} \\ \text{logit}(\phi_i) &= \alpha + \sum_{j \sim i} \beta y_j \mapsto \text{2do. nivel} \end{aligned}$$

cuya función potencial es:

$$Q(y) = \alpha \sum_i y_i + \beta \sum_{i, j \text{ vecinos}} y_i y_j + \sum_i \log \binom{n_i}{y_i}$$

Estos modelos han encontrado aplicación en el análisis de imágenes y la detección con datos tomados desde satélite (Gelfand, et al. 2005).

2. **Caso Continuo:** El modelo de *campo aleatorio markoviano* más empleado para datos continuos es el *autonormal* o *autogaussiano*,

$$y_i | y_{-i} \sim N\left(\mu_i + \sum_{j=1}^n c_{ij}(y_j - \mu_j), \alpha_i^2\right) \quad (2.12)$$

En este caso la constante de normalización puede ser evaluada, alcanzando un conocimiento exacto de la distribución de probabilidad conjunta,

$$Y \sim (\mu, (I - C)^{-1}M)$$

donde C es una matriz $n \times n$ con elementos c_{ij} tal que $c_{ij}\alpha_j^2 = c_{ji}\alpha_i^2$, y $c_{ii} = 0$, mientras que $M = \text{diag}(\alpha_1^2, \dots, \alpha_n^2)$. El caso Gaussiano presenta mayores posibilidades al facilitar una forma cerrada de la distribución de probabilidad conjunta. La construcción del

modelo se realiza mediante probabilidades conjuntas o probabilidades condicionales. Pero existen diferencias en ambas formulaciones, para ilustrar estas diferencias es conveniente la comparación con un proceso temporal. Si la autoregresión espacial, se hace a través de la expresión condicional de la probabilidad del proceso en cada localización (estructura de vecindad), entonces se construye el *modelo condicional autoregresivo espacial* (CAR). Mientras que si la autoregresión se incorpora mediante una matriz de dependencia espacial (análogo al término empleado en los modelos de series temporales, expresando la interrelación mutua entre localizaciones vecinas), entonces se tiene un *modelo simultáneo autoregresivo espacial* (SAR).

Cuando la distribución condicional no es normal, la constante de normalización es habitualmente intratable, al no tener una expresión analítica. Una desviación del caso Gaussiano puede conducir al planteamiento de otros modelos para datos continuos, como por ejemplo, el autogamma, el autoexponencial y el autobeta.

Definición de Modelo CAR:

Se hará especial énfasis en el *modelo* CAR, ya que después de 30 años de haber sido introducido por Besag (1974), ha tenido un resurgimiento importante en los últimos años por su empleo en el contexto del muestreo Gibbs y en el ajuste de cierta clase de modelos jerárquicos espaciales. Un *modelo* CAR viene dado por

$$Y_i|y_j, j \neq i \sim N\left(\sum_j b_{ij}y_j, \tau_i^2\right), \quad i = 1, \dots, n \quad (2.13)$$

De estas condicionales usando el Lema de Brook's se puede obtener la distribución conjunta, dada por

$$p(y_1, \dots, y_n) \propto \exp\left(-\frac{1}{2}y'D^{-1}(I - B)y\right) \quad (2.14)$$

La expresión (2.14) conduce a una distribución Normal Multivariante que bajo ciertas condiciones ($b_{ij} = \frac{w_{ij}}{w_i^+}$ y $\frac{\tau_i^2}{w_i^+}$), permite formular el *modelo condicional* CAR como:

$$Y_i|y_j, j \neq i \sim N\left(\sum_j \frac{w_{ij}y_j}{w_i^+}, \frac{\tau_i^2}{w_i^+}\right), \quad i = 1, \dots, n \quad (2.15)$$

donde $\sum_j w_{ij} = w_i^+$ y $w_{ij} = 1$ si j es vecino de i y 0 en caso contrario. La distribución conjunta a partir de la ecuación (2.15) puede escribirse como

$$p(y_1, \dots, y_n) \propto \exp\left(-\frac{1}{2\tau^2}y'(D_w - W)y\right) \quad (2.16)$$

donde D_w es una matriz diagonal con $(D_w)_{ii} = w_i^+$. Como \sum_y^{-1} es singular, la distribución (2.16) es impropia. La siguiente expresión permite obtener una forma más general de (2.14)

$$p(y_1, \dots, y_n) \propto \exp\left(-\frac{1}{2\tau^2} \sum_{i \neq j} w_{ij}(y_i - y_j)^2\right) \quad (2.17)$$

La expresión (2.17) sigue siendo impropia, pero imponiendo la condición $\sum_i Y_i = 0$ permitirá que todas las condicionales sean propias, aún cuando la conjunta sea impropia. Si el modelo CAR se usa para modelizar los efectos espaciales aleatorios (*en el contexto de modelo jerárquico*), esta condición puede ser fácilmente incorporada en el método Gibbs. Existen otras formas de remediar el carácter impropio de la distribución conjunta, ver Carlin y Banerjee (2003). A (2.17) se le conoce como *modelo autoregresivo intrínsecamente* (IAR) (Banerjee et al., 2004).

Si Y_i se piensa como un vector de unidades de área dependientes o en el contexto de efectos aleatorios, como un vector de efectos aleatorios asociados con una unidad de área, entonces Y_i se puede definir a través de un *modelo condicional autoregresivo multivariante MCAR*. Por otro lado, si los Y_i son vistos como medidas asociadas con una unidad de área i en el tiempo T , entonces se tendrán *modelos Espacio-Temporales*.

2.3.3. Procesos Puntuales

Un *proceso puntual espacial* es un proceso estocástico cuyas realizaciones consisten en conjuntos numerables de puntos de una región plana (Diggle, 2003). Los conceptos de *proceso estacionario* y de *proceso isotrópico* caracterizan, respectivamente, un proceso puntual invariante por traslación y por rotación. Para un proceso puntual es fundamental determinar si verifica las condiciones de estacionariedad e isotropía. Además, habitualmente se supondrá que el proceso es *ordenable*, es decir no pueden encontrarse dos sucesos en la misma localización. Los modelos para procesos puntuales más usados por su sencillez o su flexibilidad consituyen un conjunto básico de procesos para la modelización y análisis de un patrón puntual. A continuación se presentan algunos de ellos.

1. **Proceso de Poisson:** Es el fundamento sobre el que se basa la teoría de procesos puntuales espaciales. Representa una forma idealizada de aleatoriedad espacial completa, donde la *intensidad* es constante y cada uno de los sucesos aparece de forma independiente a los demás, siguiendo una distribución uniforme sobre la región considerada.
2. **Proceso de Poisson no homogéneo:** La hipótesis de estacionariedad puede ser poco realista en muchas aplicaciones. El proceso de Poisson no homogéneo es el proceso puntual no estacionario más sencillo. Su definición es como la del proceso de Poisson, excepto que la *intensidad* no es constante. De esa forma, cada suceso sigue una distribución con densidad proporcional a la función de intensidad.
3. **Proceso de agrupación de Poisson:** Estos procesos incorporan una forma explícita de agrupación espacial, proporcionando una herramienta importante para la modelización de patrones agregados. Pueden entenderse como derivados de un proceso de Poisson cuyos

sucesos generasen agrupaciones en su entorno según una distribución determinada (denominados también procesos puntuales de padres e hijos).

4. **Procesos de Cox:** Son procesos puntuales doblemente estocásticos, generalizan los procesos de Poisson no homogéneos, permitiendo que la *función de intensidad* sea también aleatoria. Se emplean frecuentemente para modelizar fenómenos en los que la distribución espacial de los sucesos puede ser el resultado de la variabilidad de factores ambientales. Considerar a la *intensidad* como variable aleatoria permitirá que este tipo de procesos puntuales pueda estudiarse en el contexto de *modelos jerárquicos*. Tanto los procesos de Cox, como los procesos de agrupación de Poisson, tienden a producir patrones agrupados.
5. **Procesos de inhibición simple:** Los patrones regulares surgen de forma natural al imponer una distancia mínima entre los sucesos, que representa el tamaño físico de los objetos, o en una situación de competencia o territorialidad. Éstos son los denominados *procesos de inhibición simple*, que pueden formularse de diversas formas atendiendo a su dinámica de construcción.
6. **Procesos puntuales de Markov:** Son aquéllos cuya intensidad condicional en un punto, dada la realización del proceso en el resto del plano, sólo depende de los sucesos que aparecen en el entorno próximo a ese punto. Así, un *procesos de inhibición simple* puede entenderse como un caso particular de proceso de Markov.

También existen patrones puntuales multivariantes y aparecen en estudios ecológicos y ambientales, debido a la aparición de fenómenos simultáneos relacionados entre sí, como la observación de individuos de distintas especies, la aparición de varios contaminantes, o el registro de mortalidad o morbilidad por diversas causas. Habitualmente, el objetivo del estudio de un patrón multivariante es la detección de una posible interacción entre los fenómenos, bien sea de asociación o de inhibición.

2.4. Modelos jerárquicos bayesianos espaciales

Desde una perspectiva bayesiana las capas de los *modelos jerárquicos* presentados en la sección (2.3) son vistas cada una, como un proceso estocástico compuesto de observaciones, factores ocultos y parámetros a estimar. A través del “*Teorema de Bayes*” (ver 3.1 y 3.2) es posible que la información de un dato se transfiera a factores asociados a otro dato, para esto se requiere incorporar incertidumbre (estructura probabilística) tanto en las observaciones como en los parámetros de interés. El proceso de aprendizaje a través de la distribución posterior es enorme y constante, es así, que los *modelos jerárquicos bayesianos* se convierten en una herramienta potencial para el análisis de problemas complejos. Bajo la perspectiva bayesiana e

incorporando asociación espacial en alguna de las capas de los modelos descritos en la sección (2.3) se tendrán modelos que llamaremos jerárquicos bayesianos espaciales.

Gracias al análisis bayesiano es posible transferir la información de los datos a factores asociados a otro conjunto de datos a través del aprendizaje sobre los parámetros. Esta estrategia permite construir modelos jerárquicos con capas complejas que contienen observaciones, factores ocultos y parámetros del modelo. Cuando los datos son recogidos de muchas unidades que son de algún modo similares, como sujetos, animales, ciudades, etcétera, el problema estadístico es combinar la información de varias unidades para entender mejor el fenómeno en estudio. Por lo general, hay variabilidad entre las unidades y un modo natural de acercarse al problema es construyendo un modelo en etapas “*modelo jerárquico*” y luego usarlo para hacer inferencias, si estas inferencias se hacen a través de la distribución posterior entonces se emplea un “*modelo jerárquico bayesiano*”.

El uso de modelos jerárquicos bayesianos se ha generalizado en los tres tipos de datos espaciales. Primero se extendió el modelo de Besag et al. (1991) para el problema de suavización de mapas de riesgo de enfermedad o cartografía de enfermedades en áreas pequeñas. En los últimos años ha irrumpido la generalización de la geostatística a partir del trabajo publicado por Diggle et al. (1998). Y de forma todavía más reciente los procesos de Cox log-gaussianos introducidos por Möller et al. (1998) para analizar datos puntuales.

En este trabajo se aplicará la teoría de modelos jerárquicos bayesianos dedicada sólo a datos espaciales ubicados en una red fija de localizaciones. La formulación de este tipo de modelos permite representar la correlación espacial entre las observaciones, atendiendo a la idea de que observaciones de unidades geográficas próximas se parecerán más que las observaciones de unidades geográficas más distantes (Botella, 2010). Así, la idea de este tipo de modelización es incorporar la relación existente entre las observaciones o los riesgos de localizaciones vecinas usando estructura probabilística en las capas del modelo.

El uso de modelos jerárquicos bayesianos espaciales, en adelante se abreviarán como MJBE, se ha generalizado a problemas donde la variable respuesta no es normal, a continuación se presentan algunos de los modelos más usados en este contexto.

2.4.1. Modelo Poisson-Gamma:

Un área común y de interés en Epidemiología y Bioestadística es el estudio de enfermedades y su comportamiento espacial. Generalmente, se tienen casos por área y se definen cantidades como:

$$Y_i = \text{número de casos de enfermos observados en el área } i$$

$$E_i = \text{número de casos de enfermos esperados en el área } i$$

con $i = 1, \dots, n$. Si E_i no es muy grande (es decir, la enfermedad es rara o la región i son suficientemente pequeñas), en el enfoque clásico se modeliza a Y_i como un modelo Poisson dado por la siguiente ecuación:

$$Y_i | \eta_i \stackrel{ind}{\sim} Poisson(E_i \eta_i) \quad (2.18)$$

donde η_i es la verdadera tasa de riesgo de enfermedad en la región i . Esta modelización del riesgo (η_i) es excelente para detectar sobre-dispersión (extra variabilidad) en las tasas observadas, pero no permite estimar y trazar un mapa con la superficie de riesgos relativos adyacentes, a partir de los cuales, puedan suavizarse los riesgos observados. En este caso, se puede pensar en un *Modelo de Efectos Aleatorios* para los η_i , ya que parece natural pensar que todos los riesgos verdaderos (η_i) provengan de una distribución común subyacente.

Un *modelo jerárquico* usado para la suavización de los riesgos relativos en su primer nivel es:

$$Y_i | \eta_i \stackrel{ind}{\sim} Poisson(E_i \eta_i) \quad (2.19)$$

En el segundo nivel, la especificación para los riesgos relativos, viene dada por

$$\eta_i \stackrel{iid}{\sim} G(a, b) \quad (2.20)$$

donde $G(a, b)$ es una distribución gamma con $\mu = \frac{a}{b}$ y varianza $\sigma^2 = \frac{a^2}{b}$. La distribución posterior resultante es también una gamma de la forma $G(y_i + a, E_i + b)$. Así una estimación puntual de η_i puede ser obtenida de la media posterior,

$$E(\eta_i | y) = E(\eta_i | y_i) = \frac{y_i + a}{E_i + b} = \frac{y_i + \frac{\mu^2}{\sigma^2}}{E_i + \frac{\mu}{\sigma^2}} \quad (2.21)$$

$$= \frac{E_i \left(\frac{y_i}{E_i}\right)}{E_i + \frac{\mu}{\sigma^2}} + \frac{\left(\frac{\mu}{\sigma^2}\right)\mu}{E_i + \frac{\mu}{\sigma^2}} \quad (2.22)$$

$$= w_i SMR_i + (1 - w_i)\mu \quad (2.23)$$

donde $SMR =$ Proporción de Mortalidad o Morbilidad Estandarizada en la región i , $w_i = \frac{E_i}{E_i + \frac{\mu}{\sigma^2}}$, $SMR_i = \frac{Y_i}{E_i}$ y $0 \leq w_i \leq 1$. Así, la estimación puntual bayesiana (2.23) es un promedio ponderado de los SMR basados en los datos para la región i y la media de la previa μ . Esta estimación es aproximadamente igual a SMR_i cuando w_i esta cerca de 1 (es decir, cuando E_i es grande, entonces los datos son fuertemente informativos, o cuando σ^2 es grande, entonces la previa es débilmente informativa). Por otra parte, (2.23) será aproximadamente igual a μ cuando w_i esta cerca de 0 (es decir, cuando E_i es pequeño, entonces los datos son escasos, o cuando σ^2 es pequeña, la previa será fuertemente informativa).

2.4.2. Modelo Poisson-Lognormal:

El *modelo Poisson-Gamma* descrito anteriormente es muy conveniente computacionalmente, pero falla al tener en cuenta la correlación espacial entre los riesgos (η_i). Para considerar la relación espacial, es necesario usar la versión multivariante de la distribución gamma, tales estructuras existen, pero son conceptual y computacionalmente complicadas. Un enfoque usual es usar una distribución normal multivariante en los $\psi_i \equiv \log(\eta_i)$, el logaritmo de los riesgos relativos.

Considerando esta transformación en los ψ_i se tiene un *modelo jerárquico* determinado por los niveles:

$$Y_i | \psi_i \stackrel{ind}{\sim} Po(E_i e^{\psi_i}) \quad (2.24)$$

donde $\psi_i = x'_i + \theta_i + \phi_i$ y las x_i son covariables espaciales explicativas. Estas variables son ecológicas y se espera que expliquen (“casi todo”) el comportamiento espacial en las Y_i .

Los θ_i capturan la heterogeneidad entre las regiones. A esta estructura probabilística se le asigna una previa normal,

$$\theta_i \stackrel{iid}{\sim} N(0, \frac{1}{\tau_h}) \quad (2.25)$$

donde τ_h es la precisión (recíproco de la varianza), que controla la magnitud de los θ_i . Los θ_i son efectos que capturan la variabilidad *Extra-Poisson* en los log-riesgos relativos sobre la región de estudio completa.

Finalmente, los ϕ_i son los parámetros que hacen de esta formulación, un *modelo realmente espacial*, ya que capturan las similitudes regionales (agrupamientos). Esta componente modeliza la variabilidad *Extra-Poisson* de los log-riesgos relativos que varían localmente, haciendo que regiones cercanas tengan tasas más similares.

Se requerirá una distribución normal multivariante para los ϕ . Del mismo modo, se necesita una noción de distancias entre las unidades de área. Con unidades aproximadamente del mismo tamaño definidas en una grid regular, la distancia intercentroidal puede ser apropiada. Pero, en unidades espaciales muy irregulares, tales distancias tienen poco sentido.

Debido a las dificultades comentadas, se tiene como resultado, emplear el *modelo CAR* para especificar la estructura probabilística de ϕ y volver a las nociones de vecino (proximidad). En este sentido,

$$\phi \sim CAR(\tau_c) \quad (2.26)$$

En esta formulación se reemplaza y_i por ϕ_i , τ^2 es reemplazado por $\frac{1}{\tau_c}$ y los pesos w_{ij} son $0 - 1$ (matriz de adyacencia).

Comparado con los *modelos para datos geostadísticos*, los *modelos CAR* son muy convenientes computacionalmente, ya que el método Gibbs sampling permite el ajuste tanto de θ como de ϕ

a partir de las condicionales completas. La distribución condicional completa para ϕ_i es

$$p(\phi_i | \phi_{j \neq i}) = N\left(\sum_{j \in \partial_i} \frac{\phi_j}{m_i}, \frac{\tau_c^2}{m_i}\right) \quad (2.27)$$

donde m_i = número de vecinos de la región i , ∂_i denota el conjunto de vecinos de i . La ecuación (2.27) es de la forma $p(y_i | y_j, j \neq i) = p(y_i | y_j, j \in \partial_i)$. Demostrando que no es necesario trabajar con la distribución conjunta de todos los ϕ . El enfoque con condicionales elimina la necesidad de invertir alguna matriz de covarianzas.

A pesar de que existen dificultades teóricas y computacionales para trabajar con los *modelos CAR* en la práctica, algunas discutidas en la sección (2.3), se pueden resolver usando la expresión (2.27) y seleccionando previas adecuadas para τ_c y τ_h , ya que estas no pueden elegirse arbitrariamente (Gelman, 2006).

2.4.3. Modelo Binomial con vínculo logit:

En muchos problemas se tienen observaciones cuyos valores son 0 ó 1 y se desea conocer por ejemplo, la probabilidad de un individuo enfermar. Esta modelización se requiere en estudios epidemiológicos de diversa índole, en especial, para modelizar enfermedades de contagio, cuya dinámica esta determinada por la proximidad de individuos cercanos. Generalmente se asigna un modelo Binomial (1,p) en la capa de las observaciones, incorporando de forma lineal el logit de la media; en la segunda capa se incluyen las componentes espaciales y temporales autoregresivos así como efectos aleatorios de heterogeneidad.

En este apartado se presenta la definición de la estructura probabilística base para la formulación de los modelos MJEB propuestos en el capítulo 4 (secciones 4.3, 4.4) y que en Epidemiología Agrícola han sido poco utilizados.

Bajo esta estructura jerárquica se establece que la probabilidad de un árbol enfermar, dependerá de la influencia tanto de la covariable x_1 (“efecto del pasado”), como de la influencia de efectos aleatorios no espaciales (θ_j) y espaciales (ϕ_j) (“efecto del presente”); efectos aleatorios que capturan la variabilidad no observada debido a la presencia de factores de riesgo ocultos. Por lo tanto, al combinar estas fuentes de información en un sólo modelo, se tiene la siguiente definición

$$y_j | p_j \sim Binomial(1, p_j) \quad (2.28)$$

$j = 1, \dots, \text{nro-árboles en el año } i$, considerando el efecto de los árboles enfermos del año $i - 1$, con $i = 2, \dots, 5$.

$$\text{logit}(p_j) = \alpha + \beta x_1 + \theta_j + \phi_j, \quad (2.29)$$

$$\theta \sim Normal(0, \sigma_\theta^2) \quad (2.30)$$

$$\phi_{i|\phi_{j \neq i}} \sim N\left(\sum_j \frac{w_{ij}\phi_j}{w_i^+}, \frac{\sigma_\phi^2}{w_i^+}\right) \quad (2.31)$$

donde $\sum_j w_{ij} = w_i^+$ y $w_{ij} = 1$ si j es vecino de i y 0 en caso contrario. A las variables α y β se les asignarán distribuciones normales con media 0 y varianza 0.001, cuando se quieran mantener constantes. Sin embargo, cuando se quieran incluir en el modelo como efectos aleatorios, se les asignarán distribuciones normales con la siguiente estructura probabilística:

$$\alpha \sim Normal(0, \sigma_a^2), \sigma_a \sim Unif(0, 1) \quad (2.32)$$

$$\beta \sim Normal(0, \sigma_b^2), \sigma_b \sim Unif(0, 1) \quad (2.33)$$

A las desviaciones estándar que definen a las precisiones de los parámetros, θ_j y ϕ_j , se asignan distribuciones como en 2.32 y 2.33.

Con esta modelización es posible considerar diferentes fuentes de variabilidad que con otras metodologías sería difícil considerar. En el contexto que vamos a trabajar no se tendrán datos de conteo (modelo Poisson), sino variables respuesta que asumen sólo dos valores (modelo Binomial). Es así, que el trabajo se centra sólo en el estudio de modelos jerárquicos bayesianos espaciales con la estructura definida en 2.28-2.33.

Capítulo 3

Inferencia

En este capítulo se introducen los conceptos de inferencia bayesiana, las técnicas Markov chain Monte Carlo (MCMC), así como la forma de llevarlas a la práctica. Se presentan los principales algoritmos usados en inferencia bayesiana, el software disponible y los criterios para la selección de modelos usados bajo el enfoque bayesiano.

3.1. Inferencia bajo el enfoque Bayesiano

Bajo esta perspectiva, la incertidumbre o falta de información sobre el parámetro θ puede ser incorporada a través de distribuciones previas, considerando este parámetro como una variable aleatoria. Sea $\pi(\theta|\lambda)$ la distribución previa, donde λ es un vector de hiperparámetros. Si λ es conocida, la inferencia sobre θ se hace a partir de la distribución posterior $p(\theta|y, \lambda)$, que se obtiene gracias a “Teorema de Bayes” que combina la previa y la verosimilitud.

$$p(\theta|y, \lambda) = \frac{p(y, \theta)|\lambda}{p(y|\lambda)} = \frac{p(y, \theta)|\lambda}{\int p(y, \theta|\lambda)d\theta} = \frac{f(y|\theta)\pi(\theta|\lambda)}{\int f(y|\theta)\pi(\theta|\lambda)d\theta} \quad (3.1)$$

En la práctica, λ no es conocido y por tanto, es necesario definir un segundo estado para los hiperparámetros (distribución para $p(\lambda)$), quedando (3.1) como:

$$p(\theta|y) = \frac{p(y, \theta)}{p(y)} = \frac{\int f(y|\theta)\pi(\theta|\lambda)h(\lambda)d\lambda}{\int \int f(y|\theta)\pi(\theta|\lambda)h(\lambda)d\theta d\lambda} \quad (3.2)$$

Alternativamente, se puede reemplazar λ por el estimador $\hat{\lambda}$ obtenido al maximizar la distribución marginal $p(y|\lambda) = \int f(y|\theta)\pi(\theta|\lambda)d\theta$, visto como una función de λ . La inferencia puede estar basada en el estimador de la distribución posterior $p(\theta|y, \hat{\lambda})$ al reemplazar $\hat{\lambda}$ en la ecuación (3.1). Este enfoque es conocido como *Análisis Empírico Bayes*; ver Berger (1985), Maritz y Lwin (1989), Carlin y Louis (2000) para más detalles de esta metodología.

- *Regla de Bayes*: Se debe comenzar con un modelo que provea una distribución conjunta para θ y y . La función de densidad conjunta es escrita como un producto de dos densidades

que son referidas frecuentemente como la distribución a priori $p(\theta|\lambda)$ y la verosimilitud $p(y|\theta)$ respectivamente:

$$p(\theta, y) = p(\theta)p(y|\theta) \quad (3.3)$$

Al condicionar en el valor conocido de los datos y , usando la *regla de Bayes* se obtiene la distribución posterior:

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)} \quad (3.4)$$

donde, $p(y) = \sum_{\theta} p(\theta)p(y|\theta)$, en el caso discreto o $p(y) = \int p(\theta)p(y|\theta)d\theta$ en el caso continuo. Una forma equivalente de (3.4) omite el factor de $p(y)$, el cual no depende de θ y con y fijo puede ser considerado como una constante. El lado derecho de (3.4) puede escribirse como:

$$p(\theta|y) \propto p(\theta)p(y|\theta) \quad (3.5)$$

En el enfoque bayesiano toda la inferencia se hace a partir de la distribución posterior $p(\theta|y)$.

3.2. Métodos Markov chain Monte Carlo (MCMC)

Las técnicas MCMC (generación de cadenas de Markov para usarlas en estudios Monte Carlo) están diseñadas con la finalidad de poder estudiar empíricamente las características de distribuciones complejas. Básicamente consisten en lo siguiente: Dada una distribución $p(\theta)$ completamente conocida, salvo quizá por su constante de proporcionalidad, se trata de generar una o varias realizaciones de una cadena de Markov cuya distribución estacionaria sea $p(\theta)$. Esas realizaciones se utilizarán posteriormente para obtener, por Monte Carlo, aproximaciones a todas las características de $p(\theta)$ que se desean conocer (Bermúdez, 2010).

Desde la perspectiva de las técnicas MCMC, el estudio de las cadenas de Markov es completamente distinto. Se parte suponiendo que la distribución que deseamos es invariante, $\pi(\theta)$, y se pretende construir un núcleo de transición que produzca una cadena de Markov con distribución estacionaria $\pi(\theta)$, para la que las medias ergódicas sean estimadores consistentes y podamos aplicar alguna versión del Teorema del Límite.

La primera utilización documentada de estas técnicas es Metropolis et al. (1953), donde se introduce un método que posteriormente es generalizado por Hastings (1970), y que ahora se denomina *algoritmo de Metropolis-Hastings*. Sin embargo, esos trabajos pioneros pasan prácticamente desapercibidos en la literatura estadística durante mucho tiempo. Más recientemente se introduce el *algoritmo de Gibbs* (Geman y Geman, 1984) y el *algoritmo de Data Augmentation* (Tanner y Wong, 1987), pero es tras la publicación de Gelfand y Smith (1990) cuando este tipo de métodos se convierten en una herramienta indispensable en la aplicación del paradigma bayesiano (Bermúdez, 2010).

Una referencia obligada en el estudio de estas técnicas es la monografía editada por Gilks, Richardson y Spiegelhalter (1996). También es de destacar el texto de Gamerman (1997), el artículo de Brooks (1998) y la monografía de Robert y Casella (1999). A continuación se detallan los algoritmos Gibbs sampling y Metropolis-Hastings.

3.2.1. Muestreador Gibbs

Entre las técnicas MCMC el algoritmo de Gibbs es uno de los métodos más fáciles de aplicar y, sin duda por ello, el más conocido y utilizado. El artículo de Casella y George (1992) constituye una introducción clara y concisa de este método, y en Gelfand et al. (1990) se presentan numerosas aplicaciones del mismo (Bermúdez, 2010).

El muestreador Gibbs aproxima integrales que no pueden ser calculadas en forma cerrada generando cadenas de Markov Monte Carlo (MCMC), donde la transición del origen de la distribución $\pi(\theta)$ esta formada por las distribuciones condicionales completas ($\pi(\theta_i) = \pi_i(\theta_i|\theta_{-i})$). Se asume que la distribución de interés es $\pi(\theta)$, donde el vector puede descomponerse en $k \geq 2$ subvectores, $\theta = (\theta_1, \dots, \theta_k)$. Cada uno de los componentes θ_i de θ puede ser un escalar, un vector o una matriz. Se considera que las distribuciones condicionales completas $\pi_1(\theta_i) = \pi_1(\theta_i|\theta_{-i})$ están disponibles, pudiendo generar valores de las mismas sin excesivo coste computacional, siendo θ_i el vector $(\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$.

Estas condiciones se presentan habitualmente en el estudio de *modelos jerárquicos* con iniciales conjugadas y posiblemente, datos faltantes o incompletos. En estos casos, los datos aumentados pueden considerarse incluidos en el vector θ , conjuntamente con los parámetros e hiperparámetros del modelo.

El objetivo del algoritmo de Gibbs es obtener una muestra suficientemente grande de la distribución posterior. A partir de ella se podrá hacer inferencias sobre los momentos, las marginales, la distribución predictiva, o cualquier otra característica de la distribución posterior que sea de interés (Bermúdez, 2010). El problema que se debe resolver es cómo tomar una muestra de la distribución π , cuando los planes para la generación de las muestras son costosos, complicados o simplemente no se conoce el origen de la distribución π , pero es posible generar muestras de las distribuciones $\pi_i(\theta_i)$. El algoritmo Gibbs puede ser descrito de la manera siguiente:

1. Se inicializa el contador de la iteración de la cadena en $j = 1$ y se asignan valores iniciales para el vector $\theta^{(0)} \leftarrow (\theta_1^{(0)}, \dots, \theta_k^{(0)})$
2. Repetir hasta convergencia {
 Se obtiene un nuevo valor $\theta^{(j)} = (\theta_i^{(j)}, \dots, \theta_k^{(j)})'$ a partir de $\theta^{(j-1)}$ por la sucesiva generación de los valores:

$$\theta_1^{(j)} \sim \pi_1(\theta_1|\theta_2^{(j-1)}, \dots, \theta_k^{(j-1)})$$

$$\theta_2^{(j)} \sim \pi_k(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_k^{(j-1)})$$

⋮

$$\theta_k^{(j)} \sim \pi_k(\theta_k | \theta_1^{(j)}, \dots, \theta_{k-1}^{(j)})$$

3. Se actualiza el contador de j a $j + 1$ y se regresa al paso 2.}

Cuando la cadena converge, los valores resultantes de $\theta^{(j)}$ son una muestra de la distribución π . Se asume la convergencia de la cadena si la cadena se aproxima a una condición de equilibrio cuando el número de iteraciones se incrementa (Gamerman, 1997).

Las condiciones de convergencia para el muestreador Gibbs fueron establecidas por Robert y Smith (1994). Los resultados son presentados en términos de espacios paramétricos, pero pueden extenderse y combinar los parámetros continuos y discretos.

Un valor de la distribución de interés π es obtenido solamente cuando el número de iteraciones de la cadena se aproxima a infinito. En la práctica esto no es posible, y un valor obtenido en una iteración suficientemente grande es tomado como una muestra de la distribución π . La dificultad es la determinación de cuán grande debería ser el número de iteraciones. No hay respuesta simple a esta pregunta y los mayores esfuerzos se han orientado al estudio de las características de la convergencia de cadenas (Gelfand y Smith, 1990).

La muestra obtenida de θ (con $j \rightarrow \infty$) es una muestra correlacionada de la distribución posterior de la cual, se puede obtener cualquier cantidad de interés. Usando la estimación Monte Carlo es posible encontrar

$$\hat{E}(\theta_i | y) = \frac{1}{K - j_0} \sum_{j=j_0+1}^K \theta_i^{(j)}$$

La iteración de $j = 0$ a $j = j_0$ se conoce como período de inicialización (burn-in). En la práctica, se pueden simular paralelamente m cadenas, en este caso, el estimador posterior de la media sería

$$\hat{E}(\theta_i | y) = m \frac{1}{K - j_0} \sum_{i=1}^m \sum_{j=j_0+1}^K \theta_{i,j}^{(j)}$$

3.2.2. Muestreador Metropolis-Hastings

Este algoritmo consiste en generar valores de una cadena de Markov cuya distribución estacionaria, distribución marginal de la cadena de Markov, sea la distribución objetivo $\pi(\theta)$. El algoritmo de Metropolis-Hastings proporciona un método sencillo para construir innumerables cadenas de Markov con esa propiedad, lo que nos permitirá buscar entre ellas una que además posea otras propiedades complementarias: rapidez de convergencia a estacionaridad y no demasiada autocorrelación (Bermúdez, 2010).

Básicamente se trata de utilizar una cadena de Markov auxiliar, para la que disponemos de un generador eficiente de su núcleo de transición $Q(\theta, A)$ (que representa la distribución de probabilidades de pasar en una etapa del punto θ a la región A), y añadirle en cada etapa un mecanismo de aceptación-rechazo con probabilidad de aceptación dada por:

$$\alpha(\theta, \phi) = \min\left\{\frac{q(\phi, \theta)\pi(\phi)}{q(\theta, \phi)\pi(\theta)}, 1\right\}$$

de manera que si en la etapa i el valor obtenido es $\theta^{(i)}$, en la etapa siguiente se genera ϕ a partir de $q(\theta, \phi)$, que es la función de probabilidad (o densidad en su caso) asociada al núcleo $Q(\theta, A)$. Entonces $\theta^{(i+1)} = \phi$ con probabilidad $\alpha(\theta, \phi)$, o $\theta^{(i+1)} = \theta^{(i)}$ en otro caso.

La mecánica del algoritmo es la siguiente:

1. Se inicializa el contador de la iteración de la cadena en $i = 0$ y se asignan valores iniciales para el vector $\theta^{(0)} \leftarrow (\theta_1^{(0)}, \dots, \theta_k^{(0)})$
2. Repetir

$$\begin{aligned} & \{\phi \leftarrow \sim q(\theta^{(i)}, \phi) \\ & r \leftarrow \frac{q(\phi, \theta^{(i)})\pi(\phi)}{q(\theta^{(i)}, \phi)\pi(\theta^{(i)})} \\ & u \leftarrow \sim Un(0, 1) \\ & \text{si } u \leq r \text{ entonces } \phi \leftarrow \theta^{(i)} \\ & \theta^{(i+1)} \leftarrow \phi \\ & i \leftarrow i + 1\} \end{aligned}$$

A la familia de distribuciones $q(\phi, \theta)$ se le conoce como *distribución instrumental*.

Un caso particular es cuando el núcleo de transición de probabilidades instrumental es simétrico, $q(\theta, \phi) = q(\phi, \theta)$, entonces se obtiene el *algoritmo de Metropolis*.

Como la distribución objetivo $\pi(\theta)$ sólo aparece en el algoritmo a través de un cociente, no es necesario conocer su constante de integración. Por ello, el *algoritmo de Metropolis-Hastings* puede ser una herramienta muy útil en la aplicación de las técnicas bayesianas, donde la constante de integración de la distribución posterior no suele ser conocida.

3.2.3. Cómo llevar a la práctica las técnicas MCMC

Son muchas las preguntas que se plantean en el estudio sobre las características de la convergencia de las cadenas. Aunque todas ellas están bastante relacionadas, se comentarán por separado. Las ideas que se muestran a continuación fueron tomadas de Bermúdez (2010).

- *Cuántas cadenas generar:* En el artículo de Gelfand y Smith (1992) se propone utilizar solamente el último valor observado de la cadena, generando tantas cadenas como tamaño deseemos que tenga la muestra de la distribución objetivo. De esa forma se consigue una muestra aleatoria cuyo análisis estadístico es muy sencillo.

Alternativamente, si se utiliza sólo una cadena habrá que descartar las primeras etapas hasta conseguir equilibrio (distribución estacionaria), pero cuando está se haya alcanzado todas las nuevas etapas tienen como distribución marginal la distribución objetivo. La independencia en la muestra obtenida se pierde, por lo tanto, la muestra ya no es aleatoria con lo que se complica ligeramente el análisis estadístico, pero se gana enormemente en eficiencia. Además, los resultados sobre invariabilidad son asintóticos, luego cuanto más avancemos en la cadena debemos esperar resultados más fiables. Sin embargo, posiciones intermedias entre estos dos extremos pueden ser mucho más interesantes. Así, Gelman y Rubin (1992) propusieron la utilización de pocas cadenas, cada una de ellas empezadas en puntos muy distantes del espacio paramétrico, pero utilizando muchas etapas de cada cadena. Esta estrategia permite realizar un diagnóstico de la convergencia, comparando los resultados obtenidos con las distintas cadenas. En efecto, cualquier cadena puede quedarse atrapada en una moda de la distribución objetivo, dando la sensación de una buena convergencia que, sin embargo, no existe; así si fuese la única cadena generada, las conclusiones del estudio serían erróneas.

- *De qué punto partir:* En teoría, los resultados de un algoritmo MCMC son independientes del valor inicial del algoritmo, pero en la práctica, la elección del valor inicial va a influir en la rapidez con la que se alcance el equilibrio. Lo aconsejable es partir de valores iniciales que estén en zonas de alta probabilidad con respecto a la distribución objetivo: partir del estimador máximo verosímil, por ejemplo. Si se utilizan varias cadenas, es aconsejable que los valores iniciales estén bastante repartidos por el espacio paramétrico, para evitar que todas las cadenas se queden atrapadas en un mismo máximo local. Pueden utilizarse métodos de optimización para localizar la moda, o modas de la distribución objetivo. Gelman y Rubin (1992) propusieron localizar las regiones de alta densidad con respecto a la función objetivo, construir una mixtura de distribuciones t-Student cada una de ellas centrada en una de las regiones localizadas, y utilizar esa mixtura para simular los valores iniciales de las distintas cadenas.
- *Número de iteraciones hasta convergencia:* Posiblemente esta es la cuestión más importante, y más difícil de resolver, de todas las planteadas. Las primeras etapas del algoritmo todavía estarán influenciadas por el punto inicial, por lo que su uso introduciría un sesgo en los resultados. Esas primeras etapas, hasta que el algoritmo alcanza equilibrio, se conocen como inicialización (burn-in) y deben ser desechadas. Pero, ¿cómo saber que el algoritmo

ya ha convergido y está en equilibrio? ¿cuántas etapas debemos desechar?. Para responder a estas interrogantes, se mencionan brevemente algunos métodos de diagnósticos.

Gelfand y Smith (1990) sugirieron observar las gráficas de cuantiles y de autocorrelaciones. Las propias trazas de las series univariantes pueden ser de utilidad, una vez alcanzada la convergencia las trazas de los diversos parámetros deben estabilizarse alrededor de algún valor, sin mostrar ninguna tendencia definida.

Geweke (1992) propone comparar, en cada serie univariante, la media del primer tramo con la del último, utilizando estimadores espectrales de la varianza. Si se descubren diferencias significativas el primer tramo es descartado, en otro caso se considera que toda la serie ya está en equilibrio. Algo parecido proponen Heidelberger y Welch (1983), pero utilizan el test no paramétrico de Cramer von Mises.

Gelman y Rubin (1992) propusieron un test basado en el análisis de varianza, comparando varias cadenas.

Raftery y Lewis (1992) propusieron un método basado en los resultados teóricos sobre convergencia en las cadenas de Markov de dos estados. Para ello sugieren fijar algún cuantil, habitualmente de orden 0,025 ó 0,975 y sustituir la cadena por otra formada por ceros y unos, que ya será una cadena de Markov de dos estados.

- *Cuántas iteraciones después de la convergencia:* Una vez alcanzado el equilibrio, el número de etapas a utilizar es un problema de tamaño muestral, que puede resolverse de la forma habitual. Así, debemos establecer una cota sobre el error típico de los estimadores de los parámetros que se consideren más relevantes y, a partir de ella, obtendremos el tamaño muestral.

Si la muestra es aleatoria y se estima una característica univariante de la distribución objetivo mediante su media muestral, \bar{x} , el error típico del estimador es s/\sqrt{n} , siendo n el tamaño de la muestra y s su desviación típica muestral. Alternativamente, si la serie puede aproximarse por un proceso autorregresivo de primer orden, el error típico de la media ergódica es:

$$\frac{s}{\sqrt{n}} \sqrt{\frac{1+r}{1-r}}$$

siendo r un estimador de la autocorrelación de la serie. De esta forma también resulta sencillo trabajar con muestras no independientes. A $\sqrt{(1+r)/(1-r)}$ se le conoce como *factor de inflación* y puede ser inferior a uno si la autocorrelación es negativa, lo que es muy difícil que ocurra en un algoritmo MCMC.

Acotando el error típico por una cantidad dada $\epsilon > 0$, el tamaño muestral debe ser:

$$n = \text{Parte entera} \left(\frac{s^2}{\epsilon^2} \frac{1+r}{1-r} \right)$$

- *Cómo saber si la cadena se esta mezclando adecuadamente:* La cadena no sólo debe recorrer todo el soporte de la distribución objetivo, debe hacerlo con rapidez. En otro caso se dice que la cadena no se está *mezclando* bien, y se necesitaría un enorme número de etapas para poder extraer resultados fiables.

Para comprobar si la cadena se esta mezclando bien resultan muy útiles las gráficas de las trazas univariantes. Tendencias cíclicas en esas trazas indican que la cadena no se está mezclando bien. Otra herramienta diagnóstica interesante la constituyen los *factores de inflación*, pues si las autocorrelaciones son muy grandes la cadena tardará mucho en poder recorrer todo el soporte de la distribución objetivo.

La inclusión de parámetros de sintonización en el método MCMC utilizado permite cambiar de núcleo sin realizar cambios en el código programado. Así, durante las primeras etapas se debe probar con diversos valores de los parámetros de sintonización, hasta encontrar un valor para el que la cadena converja y se mezcle bien. De no conseguir resultados satisfactorios debe cambiarse el método MCMC, o buscar alguna reparametrización.

Aunque no existen resultados teóricos convincentes, suelen conseguirse mejores resultados cuando las correlaciones entre los parámetros son pequeñas.

- *Se utilizan todas las etapas o se adelgaza la salida:* Al utilizar una única cadena muy larga o varias cadenas no tan largas, la muestra obtenida no es independiente, lo que dificulta su análisis estadístico. Por ello se ha propuesto *adelgazar* la salida utilizando tan sólo una de cada k etapas, siendo k un natural no demasiado grande, de manera que las etapas usadas sean *aproximadamente independientes*. Pero al desechar etapas se está perdiendo información y no se gana gran cosa, pues la estimación utilizando medias ergódicas no es complicada.

Existe, sin embargo, otra razón que justifica en ocasiones el adelgazamiento de la salida: restricciones en la capacidad de memoria para el almacenamiento de la salida. Teniendo en cuenta que habitualmente se necesitan cientos de miles de etapas, la memoria necesaria para el almacenamiento de la salida puede constituir un gran problema que hay que considerar. Si tenemos una restricción acotando el tamaño máximo de la salida, como es mucho más informativa una muestra independiente que una muestra relacionada, será conveniente que la adelgacemos. También puede ser útil truncar los valores generados antes de almacenarlos, para utilizar menos cifras decimales.

3.3. Inferencia en automodelos

En general, la inferencia en campos aleatorios Markovianos (*automodelos*) puede hacerse a través de varios métodos, entre los que se pueden mencionar: método de codificación, método

de Pseudoverosimilitud y métodos MCMC (Scoring de Monte Carlo, aproximación estocástica, entre otros). En este trabajo se aplicará el método de Pseudoverosimilitud.

3.3.1. Pseudoverosimilitud:

Este método consiste en el producto de todas las distribuciones condicionales completas (Besag, 1974). Su algoritmo será de la forma

$$l^*(\theta; y) = \sum_{i=1}^s \log(P(y_i|y_{-i})) \quad (3.6)$$

Su maximización proporciona un estimador máximo pseudoverosímil. La pérdida de exactitud y eficiencia se ve compensada con la provisión de expresiones cerradas que evitan las dificultades propias de la verosimilitud. En el caso de un automodelo, la pseudoverosimilitud coincide con la verosimilitud de un GLM en el que las interdependencias con otras localizaciones se consideran como covariables.

3.4. Inferencia Bayesiana con MCMC

La inferencia en modelos jerárquicos puede realizarse mediante máxima verosimilitud, pero con frecuencia dicha verosimilitud no es totalmente conocida. La metodología Bayesiana, ha extendido el uso de los *modelos jerárquicos* gracias a que la distribución posterior puede ser muestreada por métodos MCMC. No obstante, diversas dificultades prácticas deben tenerse en cuenta para llegar a conclusiones adecuadas.

La complejidad de las estructuras estocásticas que se derivan a partir de la formulación de un *modelo jerárquico bayesiano* dificulta su inferencia. Esta dificultad se debe a la variedad de posibilidades para la especificación de la distribución previa y a la dificultad de resumir la distribución posterior resultante. Sin embargo, en la actualidad están disponibles algoritmos y algunas herramientas informáticas que permiten realizar tal inferencia a pesar de su complejidad.

Esta inferencia supone un desafío computacional ya que en problemas reales, las integrales requeridas para hacer las estimaciones son generalmente intratables al no tener una forma analítica cerrada. Este obstáculo numérico se resuelve usando métodos de integración MCMC, algunos de los más usados se presentaron en la sección 3.2 y empleando herramientas informáticas.

Una herramienta computacional que permite el desarrollo de inferencia Bayesiana usando Muestreo Gibbs es el BUGS (Spiegelhalter et al., 1994, 2003), este software tiene dos versiones, el WinBUGS y el OpenBUGS. El OpenBUGS representa la versión abierta del proyecto BUGS.

Mientras que el WinBUGS es una versión estable que se encuentra disponible, pero no en desarrollo. Las últimas versiones del OpenBUGS se han diseñado para ser al menos tan eficaces y fiables como las del WinBUGS.

El WinBUGS (Bayesian inference Using Gibbs Sampling for Windows, Spiegelhalter et al. 2002) es un sistema capaz de especificar una variedad de distribuciones previas para muchos modelos dados y de muestrear las condicionales completas resultantes. Este sistema consiste en un conjunto de funciones que permiten la especificación de modelos y las distribuciones de probabilidad para todas sus componentes aleatorias tanto para las observaciones como para los parámetros. La especificación de los modelos es sorprendentemente simple dada la complejidad de estos modelos.

Para cada combinación de datos y modelos, WinBUGS genera muestras de los parámetros de modelo para cada iteración $k \geq 1$ después de m iteraciones. Los valores de k y m , así como los parámetros muestreados para ser almacenados, son escogidos por el usuario. Además, el programa provee los estimados basados en muestras de la media posterior y el intervalo de credibilidad para los parámetros. Este sistema emplea para su entrada y salida la sintaxis del lenguaje S desarrollado en Bell Laboratories (AT & T) a finales de los 70 y principios de los 80 por Richard Becker, Jhon Chambers y Allan Wilks. Este lenguaje inicialmente fue diseñado para análisis exploratorio de datos y la mayor parte de la funcionalidad estadística fue agregada posteriormente.

El OpenBUGS (GNU, 2010) es un software para el análisis Bayesiano de modelos complejos utilizando los métodos MCMC. Es la variante de código abierto del WinBUGS. Una diferencia fundamental entre ambos es la forma en que el sistema experto selecciona el algoritmo de actualización a utilizar para cada clase de distribución condicional completa en cada nodo. Mientras WinBUGS define un algoritmo para cada clase posible, no hay límites en el número de algoritmos que OpenBUGS puede utilizar, lo que permite una mayor flexibilidad y extensibilidad. En OpenBUGS el usuario puede seleccionar el programa de actualización que se utilizará para cada nodo después de la compilación.

El resultado de toda técnica MCMC es una realización finita de una cadena de Markov multivariante, por lo que se resume en una matriz de datos. La aplicación informática **CODA** (Convergence Diagnostic and Output Analysis), está especialmente diseñada para el análisis de esa matriz y puede obtenerse desde la página web de BUGS, junto con su manual (Best et al., 1995). También existe una versión que funciona como un módulo de R.

CODA utiliza todo el resultado de una técnica MCMC, tanto si se ha obtenido sólo una cadena o varias cadenas en paralelo, para construir un objeto del tipo MCMC que es sobre el que trabaja: es el *input* de la aplicación.

Dentro de **CODA**, ese objeto MCMC puede manipularse con facilidad, incluyendo un adelgazamiento de la salida, puede resumirse mediante diversos estadísticos descriptivos y gráficas,

y pueden realizarse varios diagnósticos de convergencia.

Entre los estadísticos descriptivos proporciona las medias, desviaciones típicas y cuantiles de la distribución empírica. Los errores típicos de la media los calcula suponiendo independencia y mediante métodos de series temporales, para incorporar las correlaciones dentro de las cadenas. Entre las gráficas destacan las trazas de cada serie, y la estimación de la densidad de cada variable. También obtiene las gráficas de autocorrelación y de correlaciones cruzadas entre variables.

Los diagnósticos de convergencia de los cuales dispone, son los de Gelman y Rubin (1992) y Geweke (1992), ambos con sus gráficas asociadas, y los de Heidelberger y Welch (1983) y Raftery y Lewis (1992).

Cuando el WinBUGS no permite trabajar con ciertos tipos de modelos complejos, es necesario elaborar un código de programación específico y, habitualmente, ir modificándolo ligeramente (mediante la utilización de parámetros de sintonización o buscando reparametrizaciones adecuadas) hasta conseguir que funcione adecuadamente. La elaboración de estos códigos en R suele ser relativamente fácil y cómoda, pero su ejecución puede ser excesivamente lenta para el volumen de operaciones a realizar y pueden presentarse problemas de memoria.

R (Core Team, 2008) es un “entorno”, es decir, un sistema completamente diseñado y coherente, y no una agrupación incremental de herramientas muy específicas e inflexibles. Posee un lenguaje que fue implementado en base al lenguaje S por Ross Ihaka y Robert Gentleman (University of Auckland, Nueva Zelanda). A partir de 1995 comienza a ser distribuido gratuitamente bajo los términos de la licencia de GNU (Free Software Foundation) y desde entonces, el desarrollo de R ha sido un esfuerzo de colaboración internacional, con trabajo aportado por voluntarios. Desde 1997 la coordinación del desarrollo de R está a cargo de un “Core Team” compuesto por miembros de todas partes del mundo.

R puede ser extendido por medio de programas escritos por el usuario o mediante “bibliotecas” (packages) que pueden ser obtenidos via Internet en forma gratuita. En R pueden incluirse paquetes como CODA (diagnóstico de convergencia), BRugs, spBayes (análisis de modelos espaciales jerárquicos para datos geostadísticos), glmmBUGS, GLMMGibbs, R2WinBUGS (Gelman et al., 2005), entre muchos otros.

Tanto el WinBUGS como el OpenBUGS pueden ser ejecutados desde R usando los paquetes R2WinBUGS o BRugs, en especial para este trabajo se usará la librería R2WinBUGS.

3.5. Criterios para selección de modelos

La comparación de modelos es requerida en muchas áreas, incluyendo la selección de variables en modelos de regresión, la determinación del número de componentes en un modelo mixto o en la selección de familias paramétricas. Igual que ocurre en el enfoque frecuentista, la com-

paración de modelos desde la perspectiva bayesiana no nos dirá cuál es el modelo verdadero, pero nos acercará al mejor a la luz de los datos y de otras informaciones. En esta sección se han recogido algunos de los métodos más empleados desde el enfoque bayesiano para comparar modelos, ingrediente indispensable y necesario en la inferencia. En este trabajo se usará sólo el criterio DIC para comparar modelos formulados desde la perspectiva bayesiana y el criterio AIC como medida de bondad de ajuste en aquellos modelos formulados desde el enfoque frecuentista.

3.5.1. Test de Hipótesis

La aproximación bayesiana a las pruebas de hipótesis está basada en el cálculo de la probabilidad condicional de una hipótesis H_0 dada la información disponible, digamos I_0 , esto es, $p(H|I_0)$. Cuando la hipótesis nula es $H_0 : \theta \in \Theta_0$ y la alternativa $H_1 : \theta \in \Theta_1$, con $\Theta_0 \cap \Theta_1 = \emptyset$, son formuladas, hay creencias apriori sobre ambas, digamos $\xi(H_0|I_0) + \xi(H_1|I_0) = 1$. Por el teorema de la probabilidad total, la distribución apriori de θ es:

$$\xi(\theta|I_0) = \xi(\theta|H_0, I_0)\xi(H_0|I_0) + \xi(\theta|H_1, I_0)\xi(H_1|I_0)$$

donde $\xi(\theta|H_i, I_0)$, son las densidades apriori de θ , condicionadas en cada hipótesis.

La información muestral es usada para calcular los odds apriori:

$$\frac{\xi(H_0|I_0)}{\xi(H_1|I_0)}$$

los odds posteriores en favor de H_0 :

$$\frac{\xi(H_0|I_0)}{\xi(H_1|I_0)} = \frac{p(y|H_0)}{p(y|H_1)} \frac{\xi(H_0|I_0)}{\xi(H_1|I_0)}$$

de la cual se deriva la siguiente regla de decisión:

Si $\xi(H_0|I_0) < \xi(H_1|I_0)$ se rechaza H_0

Si $\xi(H_0|I_0) > \xi(H_1|I_0)$ se acepta H_0

Si $\xi(H_0|I_0) = \xi(H_1|I_0)$ indecisión acerca de H_0

3.5.2. Factor de Bayes

A la razón $\frac{p(y|H_0)}{p(y|H_1)}$ se le conoce como *factor de Bayes*, denotado por **BF** o $B_{01}(y)$. Si se quiere probar:

$$H_0 : \theta \in \Theta_0 \text{ vs. } H_1 : \theta \in \Theta_1$$

Sea $f(x|\theta)$ la verosimilitud de x dado θ . Tenemos las siguientes formas del factor de Bayes:

$$B_{01}(x) = \frac{f(x|\theta_0)}{f(x|\theta_1)} \text{ (Prueba simple vs. simple)}$$

$$B_{01}(x) = \frac{f(x|\theta_0)}{\int_{\Theta_1} f(x|\theta)\xi_1(\theta)d\theta} \text{ (Prueba simple vs. compuesta)}$$

$$B_{01}(x) = \frac{\int_{\Theta_0} f(x|\theta)\xi_0(\theta)d\theta}{\int_{\Theta_1} f(x|\theta)\xi_1(\theta)d\theta} \text{ (Prueba compuesta vs. compuesta)}$$

Jeffreys presenta los siguientes criterios sobre el Factor Bayes (BF) para decidir cuándo optar por H_0 :

Factor Bayes (BF)	Decisión
$1 < BF$	Hipótesis Nula se sostiene
$10^{-\frac{1}{2}} < BF < 1$	Evidencia contra H_0 , apenas para mencionar
$10^{-1} < BF < 10^{-\frac{1}{2}}$	Evidencia sustancial contra H_0
$10^{-\frac{3}{2}} < BF < 10^{-1}$	Evidencia fuerte contra H_0
$10^{-2} < BF < 10^{-\frac{3}{2}}$	Evidencia muy fuerte contra H_0
$BF < 10^{-2}$	Evidencia decisiva contra H_0

Tabla 3.1: Criterios de Jeffreys sobre el BF para decidir sobre H_0

Cuando las probabilidades apriori son iguales, el factor de Bayes determina la regla de decisión. La evaluación del factor de Bayes involucra el cálculo de

$$p(y|H_0) = \int p(y|H_0, \theta)\xi(\theta|H_0, I_0)d\theta$$

$$p(y|H_1) = \int p(y|H_1, \theta)\xi(\theta|H_1, I_0)d\theta$$

El Factor Bayes proporciona una indicación de cuánto cambian nuestras razones de probabilidad de una situación sin datos, a la luz de los datos, para favorecer un modelo. Puede verse como una medida de la evidencia proporcionada por los datos en favor de un modelo comparado con un competidor. El logaritmo del Factor Bayes ha sido llamado *el peso de la evidencia* proporcionada por los datos (De Santis y Spezzaferri, 1999).

El Factor Bayes puede verse como la versión bayesiana de la prueba clásica de la razón de verosimilitudes (De Santis y Spezzaferri, 1999). Si se asumen dos hipótesis simples, digamos θ_1 y θ_2 , el Factor Bayes se reduce a la razón de verosimilitud $f(y|\theta_1)/f(y|\theta_2)$.

Por muchos años el *Factor Bayes* fue considerado apropiado para comparar modelos, pero sólo es posible usarlo con previas propias y para modelos de baja dimensión. Por lo tanto, cuando los modelos son complejos (*modelos jerárquicos*) y en alguno de sus niveles existen previas impropias, esta metodología no puede utilizarse, ya que si $\pi_i(\theta_i)$ es impropia, entonces $p(y|M_i) = \int f(y|\theta_i, M_i)\pi_i(\theta_i)d\theta_i$ también lo será y BF no estará definido.

3.5.3. Los criterios BIC y AIC

Como el Factor Bayes es a menudo difícil o imposible de calcular, sobre todo en modelos con muchos parámetros o efectos arbitrarios o previas impropias, una aproximación al factor de Bayes es el BIC (Criterio de Información Bayesiana). El BIC es un método “de acceso rápido” muy popular, también conocido como el Criterio de Schwarz, permite conocer el cambio entre dos modelos que se comparan a partir de:

$$\Delta BIC = W - (p_2 - p_1) \log n,$$

donde p_i es el número de parámetros en el modelo M_i , $i = 1, 2$ y

$$W = -2 \log \left[\frac{\sup_{M_1} f(y|\theta)}{\sup_{M_2} f(y|\theta)} \right]$$

es el usual ratio de verosimilitudes. Schwarz (1978) demuestra que para *modelos no-jerárquicos* (con dos estados) y tamaños de muestra n grandes, el *BIC* se aproxima a $-2 \log BF$. Una alternativa al *BIC* es el *criterio de Información Akaike* (AIC), cuya expresión es

$$\Delta AIC = W - 2(p_2 - p_1)$$

Tanto el BIC como el AIC son criterios de bondad de ajuste basados en el ratio de la verosimilitud. En ambos, el segundo término representa una penalización corregida por la diferencia entre el número de parámetros de los modelos comparados (se piensa en M_2 como el modelo “saturado” y en M_1 como el modelo “reducido”).

3.5.4. Comparación Múltiple de Modelos

De Santis y Spezzaferri (1999) piensan en términos de modelos, digamos M_1, \dots, M_s , donde se asume que M_i está parametrizado por $\theta_i \in \Theta_i$, de dimensión d_i , y con función de densidad de probabilidad de los datos $f_i(y|\theta_i)$ y distribución apriori $\xi(\theta_i)$. Si se tienen las probabilidades apriori para los modelos p_1, \dots, p_s , por el *Teorema de Bayes* se tiene:

$$Pr(M_i|y) = \frac{p_i m_i(y)}{\sum_{j=1}^s p_j m_j(y)}$$

donde

$$m_i(y) = \int_{\Theta_i} f_i(y|\theta_i) \xi(\theta_i) d\theta_i, \quad i = 1, \dots, s$$

es la distribución marginal de los datos bajo el modelo M_i . La razón de las probabilidades posteriores nos permiten hacer una comparación entre modelos.

Para los modelos M_j y M_k se tiene:

$$\frac{Pr(M_j|y)}{Pr(M_k|y)} = \frac{p_j}{p_k} B_{jk}(y),$$

donde

$$B_{jk}(y) = \frac{m_j(y)}{m_k(y)}$$

es el Factor Bayes para el modelo M_j contra el modelo M_k a partir de los datos y .

3.5.5. El Criterio de Información de Deviance (DIC)

Spiegelhalter et al. (2002) propone una generalización del criterio AIC, ya que este último no es apropiado para *modelos jerárquicos* de 3 o más niveles. Esta generalización esta basada en la distribución posterior de la *deviance*,

$$D(\theta) = -2\log f(y|\theta) + 2\log h(y) \quad (3.7)$$

donde $f(y|\theta)$ es la función de verosimilitud y $h(y)$ es función estandarizada de los datos. Este autor sugiere resumir la bondad de ajuste del modelo por la esperanza posterior de la deviance

$$\bar{D} = E_{\theta|y}[D]$$

y la complejidad de un modelo por el número efectivo de parámetros p_D

En el caso de *modelos Gaussianos*, una definición razonable de p_D , viene dada por:

$$p_D = E_{\theta|y}[D] - D(E_{\theta|y}[\theta]) = \bar{D} - D(\bar{\theta}) \quad (3.8)$$

Por tanto, el *Criterio de Información de Deviance* (DIC) se define como:

$$DIC = \bar{D} + p_D = 2\bar{D} - D(\bar{\theta}) \quad (3.9)$$

El DIC puede ser calculado durante una corrida de MCMC monitoreando θ y $D(\theta)$, y al final de la corrida simplemente se toma la media muestral de los valores simulados de D, menos la estimación de la deviance usando las medias muestrales de los valores simulados de θ . Valores pequeños de DIC indican un modelo mejor ajustado. El DIC consta de dos términos, uno que representa la bondad de ajuste y la otra una penalidad por incrementar la complejidad del modelo.

Pese a la facilidad en la implementación del cálculo del DIC en cada corrida MCMC, presenta varios inconvenientes que obligan a tener cuidado en su aplicación. La comparación de modelos usando DIC no es invariante a la parametrización, por tanto, debe ser escogida de antemano. Cualquier parámetro de escala desconocido que sea reparametrizado puede conducir a cambios sutiles en el valor del DIC calculado.

El DIC dependerá de qué parte de la especificación del modelo, sea considerada verosimilitud y cuál no. Esta consideración es el centro del asunto, ya que es necesario determinar cuáles parámetros son de interés y cuáles serán contados en el cálculo de p_D . Una dificultad con esto,

es que, en casos donde la integral $p(y|\eta) = \int f(y|\theta)p(\theta|\eta)d\theta$ no tenga forma cerrada, la versión no-integrada es realmente la única opción factible. El WinBUGS y el OpenBUGS calculan el DIC enfocándose en los parámetros de nivel más bajo para evitar tal integración.

El empleo del DIC es aún un asunto en discusión. Su formulación original (3.9) es apropiada en problemas en los cuales se usan modelos lineales generalizados, pero falla en otros contextos.

Capítulo 4

Análisis espacial de datos en Epidemiología Agrícola

El objetivo general de este capítulo, es ilustrar el uso de la metodología propuesta en el capítulo 2, empleando para ello, los árboles infectados con el virus CTV en la parcela de Burriana desde el año 1994 hasta el año 1998. Además, se usará para el estudio de los modelos espaciales propuestos, el marco conceptual de inferencia presentado en el capítulo 3.

Se sabe que la enfermedad se produce por la alimentación en un árbol sano de un pulgón vector virulífero, que antes se ha alimentado en un árbol infectado y ha adquirido el virus. Por tratarse de individuos adultos alados, la distribución espacial de las nuevas infecciones producidas en un año determinado no es sencilla. La distribución debe tener en cuenta para cada árbol la proximidad de los árboles infectados a su alrededor.

El tiempo que tarda en extenderse el virus por todo el árbol es de varios meses y es poco probable que un pulgón que se alimente en un árbol recién infectado adquiera el virus. Dado que los pulgones permanecen en los cítricos mientras los brotes están frescos, durante unas semanas, la interrelación entre los nuevos casos ocurridos en un mismo año se considera como covariable, a pesar de saberse no relevante.

La pertinencia de un modelo que explique la probabilidad de un árbol infectarse en función de la cercanía de árboles enfermos puede ser evaluada por la descripción de dicha proximidad en los datos observados.

Para cada árbol se ha calculado el número de árboles vecinos infectados considerando los que están enfermos del año anterior (“efecto del pasado”) ubicados a una distancia concreta, denominada x_1 . Así mismo, se ha usado el método de pseudoverosimilitud al considerar el efecto de la covariable x_2 , calculada con la interrelación entre los nuevos casos ocurridos en un mismo año. En general, se han considerado diferentes distancias en los modelos estudiados, distancias que van desde 10 metros hasta 40 metros, para determinar el efecto de la proximidad entre

árboles enfermos y sanos.

En este capítulo, se formulan los Modelos Lineales Generalizados (GLM), vistos desde la perspectiva frecuentista y bayesiana. Se formulan y analizan los modelos jerárquicos bayesianos espaciales (MJBE) por año y en forma dinámica (MJBED) de acuerdo a la metodología propuesta en el capítulo 2 y siguiendo la estructura probabilística definida en las ecuaciones 2.28-2.33.

La inferencia en los modelos estudiados se hace a partir de la distribución posterior determinada a través de los métodos MCMC descritos en el capítulo 3 y gracias a la implementación de los algoritmos en software R y OpenBUGS.

4.1. Distribuciones previas para los parámetros del modelo

En este apartado se describen brevemente las razones de la elección de las previas usadas para los parámetros del modelo. En primer lugar, se debe comentar que a las variables aleatorias α y β se les asigna una distribución normal, ya que pueden asumir valores positivos y negativos. Esta elección en principio es razonable, no sólo por la naturaleza de los parámetros a estimar, sino, que al obtener las estimaciones de estos parámetros con la regresión logística frecuentista y bayesiana, se observa que sus valores están en el rango de los valores de una variable con distribución normal.

En segundo lugar, es importante destacar que para elegir cuáles serían las distribuciones para las desviaciones estándar de los parámetros que definen a la varianza se hizo un estudio comparativo de la bondad de ajuste de los modelos (deviance) usando previas gammas y uniformes. Se obtuvo que al asignar previas gammas en las desviaciones estándar, los valores de la deviance eran más elevados y la estabilidad en la convergencia se alcanzaba después de 30000 iteraciones. Mientras que al usar previas uniformes, la deviance alcanzaba más rápido la convergencia (sobre 10000 iteraciones). Para mayores detalles de estas consideraciones revisar apéndice 1.

Por otra parte, esta elección parece apropiada, ya que al asignar previas uniformes a las desviaciones estándar de los parámetros α y β , se observa que son menos sensibles a los valores iniciales escogidos para el comienzo del algoritmo Gibbs sampling, que al usar previas gammas. En la figura 1 del apéndice 1, se muestra dos gráficas que presentan en paralelo el recorrido de dos cadenas para la deviance de uno de los modelos ajustados, en el que se asigna previas gammas a las desviaciones estándar. En el primer gráfico de esta figura se aprecian saltos en la cadena de color rojo que hablan de una pobre convergencia. Mientras que en el segundo gráfico de la misma figura ninguna de las cadenas alcanza el equilibrio (test de diagnóstico propuesto por Gelman y Rubin, 1992).

De igual modo, al comparar las estimaciones de los modelos donde se asignan previas uniformes vs. a los que usan previas gammas, se tiene que las estimaciones tanto de α y β son muy similares, siendo los valores del DIC y p_D muy parecidos numéricamente. También se observa

un suavizamiento espacial más claro al usar previas uniformes que al usar previas gammas. Este comportamiento es deseable ya que al modelizar datos espaciales en redes de localización, se espera que existan factores de riesgo ocultos que abaquen más de una unidad de localización (árbol) y que hacen que localizaciones vecinas (árboles) tengan comportamientos similares.

Gelman (2006) demuestra que es mejor asignar distribuciones previas uniformes con valores muy pequeños en los parámetros de varianza. Los autores Hobert y Casella (1996) demuestran que en modelos mixtos la ignorancia sobre parámetros como β puede ser modelada usando una previa normal con varianza muy grande y una previa inversa gamma con valores muy pequeños para los parámetros de varianza y así garantizar distribuciones posteriores propias. Esto junto con lo propuesto por Gelman (2006) permitirá en adelante, formular modelos con previas propias en todos los parámetros de interés. A las desviaciones estándar que definen a la varianza se impondrán entonces previas uniformes (0,1).

4.2. Estudio de Modelos Lineales Generalizados (GLM)

En esta sección se presentan los resultados obtenidos al realizar la regresión logística desde la perspectiva frecuentista y bayesiana y se comentan los resultados obtenidos desde ambas metodologías.

En un GLM se quiere modelizar $\mu = E[y_i]$ en términos del predictor lineal βx_i formado por un conjunto de p covariables.

$$\beta' x_i = \beta_0 + \beta_1 x_i^1 + \dots + \beta_p x_i^p \quad (4.1)$$

con función monótona y diferenciable que define la relación entre μ_i y su predictor lineal $g(\mu_i) = \beta' x_i$ (sección 2.1). En este caso la función de vínculo usada es logit ($\log \frac{p}{1-p}$).

Se desea determinar si un árbol se infecta o no del virus CTV, por lo tanto, se modeliza la probabilidad que tiene un árbol sano de infectarse con el virus. La transformación “logit” de dicha probabilidad se explica mediante una expresión lineal de factores que pueden ser influyentes. Las covariables que se consideran como factores son, número de árboles próximos infectados considerando los árboles ya infectados del año anterior y el efecto de nuevos árboles infectados en cada año, a menos de una distancia.

A las covariables consideradas las denotaremos como x_1 y x_2 , respectivamente, con las cuales se podrá evaluar una posible agrupación de árboles recién infectados en torno a los infectados anteriormente (“efecto del pasado”) y por el otro, la interrelación entre nuevos casos (“efecto del presente”).

4.2.1. Análisis frecuentista

Al aplicar el modelo de regresión logístico sobre los árboles que están sanos para cada año ubicados a menos de 10 metros, empleando sólo como covariable a x_1 , se han encontrado los resultados que se presentan en la tabla 4.1.

Modelo	α	p-valor(α)	β	p-valor(β)	AIC
GLM.1	-3.87	<0.001	-13.69	0.99	62.58
GLM.2	-2.55	<0.001	-0.26	0.52	165.03
GLM.3	-2.59	<0.001	0.09	0.05	146.49
GLM.4	-0.87	<0.001	-0.19	0.43	288.51

Tabla 4.1: Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 10.

Al incorporar la covariable x_2 en el modelo autologístico (método de pseudoverosimilitud) se han obtenido los resultados que se muestran en la tabla 4.2.

Año	α	p-valor(α)	β	p-valor(β)	γ	p-valor(γ)	AIC
GLM.1	-3.71	<0.001	-1.58	0.99	<0.001	1.00	62.708
GLM.2	-2.41	<0.001	-0.26	0.68	0.24	0.74	167.02
GLM.3	-2.81	<0.001	0.09	0.75	0.31	0.43	146.55
GLM.4	-0.84	0.008	-0.19	0.18	0.20	0.22	290.51

Tabla 4.2: Resultados GLM de las modelizaciones propuestas con x_1 y x_2 a una distancia 10.

Los modelos definidos en la tabla 4.1 tienen la forma $GLM.i = \alpha + \beta x_1$. Y los modelos definidos en la tabla 4.2 se definen como, $GLM.i = \alpha + \beta x_1 + \gamma x_2$, con $i = 1$ (año 95), 2 (año 96), 3 (año 97), 4 (año 98).

Se observa como en las tablas 4.1 y 4.2, el parámetro α resulta ser significativo en todos los modelos y con valores negativos cercanos a cero a medida que avanzamos en el tiempo. β no resulta ser significativo en ninguno de los modelos. Cuando se incorpora la covariable x_2 en el automodelo, se tiene que γ tampoco resulta significativo en ninguno de los automodelos ajustados por año. Por otro lado, al comparar los AIC de los modelos presentados en las tablas 4.1 y 4.2 no se tienen diferencias importantes en la bondad de ajuste. De acuerdo a los resultados, las covariables x_1 y x_2 no aportan información significativa en los modelos.

Se realizan regresiones logísticas a diferentes distancias, sus resultados se muestran en las tablas 4.3, 4.4 y 4.5. Se observa que el p-valor para β en la tabla 4.4 para el primer año, coincide con el nivel de de significación del 5% establecido.

Modelo	α	p-valor(α)	β	p-valor(β)	AIC
GLM.1	-2.87	<0.001	-18.14	0.99	61.72
GLM.2	-2.48	<0.001	0.006	0.96	164.08
GLM.3	-2.93	<0.001	0.072	0.48	146.41
GLM.4	-1.07	0.03	0.0017	0.96	290.40

Tabla 4.3: Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 20.

Modelo	α	p-valor(α)	β	p-valor(β)	AIC
GLM.1	-2.94	<0.001	-1.55	0.05	62.74
GLM.2	-2.29	<0.001	-0.042	0.74	165.19
GLM.3	-3.03	<0.001	0.07	0.46	143.66
GLM.4	-1.11	0.04	0.004	0.90	289.90

Tabla 4.4: Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 30.

Modelo	α	p-valor(α)	β	p-valor(β)	AIC
GLM.1	-2.79	<0.001	-0.98	0.09	62.04
GLM.2	-1.55	0.03	-0.15	0.21	165.16
GLM.3	-1.69	0.04	-0.08	0.33	143.68
GLM.4	-0.99	0.08	-0.002	0.93	289.25

Tabla 4.5: Resultados GLM de las modelizaciones propuestas con x_1 a una distancia 40.

En las tablas 4.3, 4.4 y 4.5 se tiene que a diferentes distancias las regresiones logísticas presentan AIC similares. β no resulta significativo en ninguno de los casos y sus valores son negativos y cercanos a cero. Estos resultados no presentan diferencias numéricas importantes con las estimaciones encontradas con la regresión logística cuando se consideran distancias menores a 10 metros.

En la tabla 4.6 se muestran algunas probabilidades calculadas usando la transformación inversa del logit $p = \frac{\exp(\alpha + \beta x_1)}{1 + \exp(\alpha + \beta x_1)}$, en función a las estimaciones obtenidas con las regresiones logísticas (modelos GLM) para distancias menores a 10 metros. Se observa que la probabilidad de enfermar para los árboles 70, 75 y 81 son pequeñas en los primeros tres años, mientras que en el cuarto año es más elevada. En general, este comportamiento se observa en casi todos los árboles analizados, es decir, se tienen probabilidades pequeñas en los primeros años que aumentan notablemente en el último año.

Modelo	p[70]	p[75]	p[81]
GLM.1	2.365e-08	2.043e-02	2.043e-02
GLM.2	0.072	0.072	0.072
GLM.3	0.076	0.069	0.069
GLM.4	0.295	0.223	0.295

Tabla 4.6: Probabilidades para los árboles (70, 75, 81) usando las estimaciones GLM a una distancia 10.

Las estimaciones anteriores constituyen un análisis exploratorio de los datos y se compararán con las estimaciones que se obtengan usando regresión logística desde la perspectiva bayesiana. En adelante, los modelos propuestos en las secciones siguientes incluirán a la covariable x_1 construida considerando distancias menores a 10 metros.

4.2.2. Análisis Bayesiano

Desde esta perspectiva es necesario asignar distribuciones a los parámetros α y β para formular los GLM por año.

Para la definición de cada modelo desde este enfoque, sólo se consideran los parámetros α y β (coeficiente de x_1). A pesar de β no resultar significativo, se incluye en el modelo porque se quiere comparar la estimación frecuentista con la bayesiana. Se desea demostrar que a pesar de las diferencias en ambos paradigmas, los resultados que se obtienen siguen los mismos principios empleados en el enfoque clásico.

La estructura del Modelo Lineal Generalizado, presenta la siguiente definición probabilística

$$y_j | p_j \sim \text{Binomial}(1, p_j) \quad (4.2)$$

$$\text{logit}(p_j) = \alpha + \beta x_1 \quad (4.3)$$

$$\alpha \sim \text{Normal}(0, 0.001) \quad (4.4)$$

$$\beta \sim \text{Normal}(0, 0.001) \quad (4.5)$$

En la tabla 4.7 se presentan las estimaciones de las modelizaciones GLM propuestas. Tales estimaciones fueron obtenidas usando OpenBUGS desde la función bugs del paquete R2WinBUGS de R. Se observa como el DIC numéricamente es similar al encontrado al realizar las estimaciones GLM desde la perspectiva clásica. Así mismo, p_D indica que el número efectivo de parámetros

esta entre 1 y 2, coincidiendo este resultado con el hallado en la sección 4.2.1, donde sólo α resultó significativo estadísticamente.

Modelo	α	β	Deviance	DIC	p_D
GLM.1	-3.38	-0.35	60.87	62.03	1.15
GLM.2	-2.30	-0.097	163.0	164.6	1.50
GLM.3	-2.47	0.13	143.4	145.0	1.57
GLM.4	-1.003	-0.003	288.3	289.9	1.64

Tabla 4.7: Estimaciones GLM usando OpenBUGS a distancia 10.

La tabla 4.8 contiene las probabilidades de enfermar de algunos de los árboles a partir de las estimaciones presentadas en la tabla 4.7. Se observa que estas probabilidades aún cuando siguen siendo pequeñas para los tres primeros años, son un poco mayor a las encontradas con la estimación clásica. Sin embargo, en el último año se siguen teniendo probabilidades más altas que la de los años anteriores. Para el último año, el número de árboles enfermos con el virus CTV se ha incrementado y por tal razón, las probabilidades de enfermar son mayores.

Modelo	p[70]	p[75]	p[81]
GLM.1	0.034	0.034	0.034
GLM.2	0.092	0.084	0.092
GLM.3	0.080	0.080	0.080
GLM.4	0.268	0.268	0.268

Tabla 4.8: Probabilidades para los árboles (70, 75, 81) usando las estimaciones GLM obtenidas con OpenBUGS a distancia 10.

Las estimaciones de los parámetros α y β se hicieron para cada modelo ejecutando simultáneamente dos cadenas y realizando 10000 repeticiones hasta lograr la convergencia de las mismas, quemando las 5000 primeras iteraciones.

En las figuras 4.1 y 4.2 se muestra la convergencia para los parámetros α y β y el test de diagnóstico de Gelman y Rubin (1992) correspondientes al Modelo GLM.1 formulado bajo el enfoque bayesiano. Se observa que ambas variables alcanzan la convergencia (figura 4.1). La prueba propuesta por Gelman y Rubin (1992), esta basada en 2 o más cadenas paralelas, cada una comenzando en valores iniciales diferentes y esta prueba es superada cuando los parámetros se estabilizan alrededor del valor 1.0. Los mejores resultados de este test son obtenidos para parámetros cuya densidad marginal posterior es aproximadamente normal. La figura 4.2 muestra como la deviance supera esta prueba.

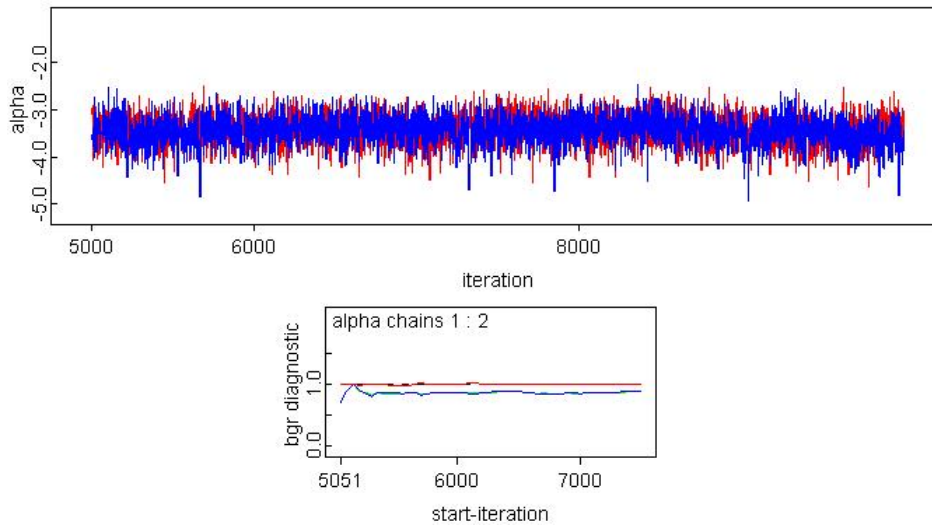


Figura 4.1: Convergencia de los parámetros α y β usando OpenBUGS.

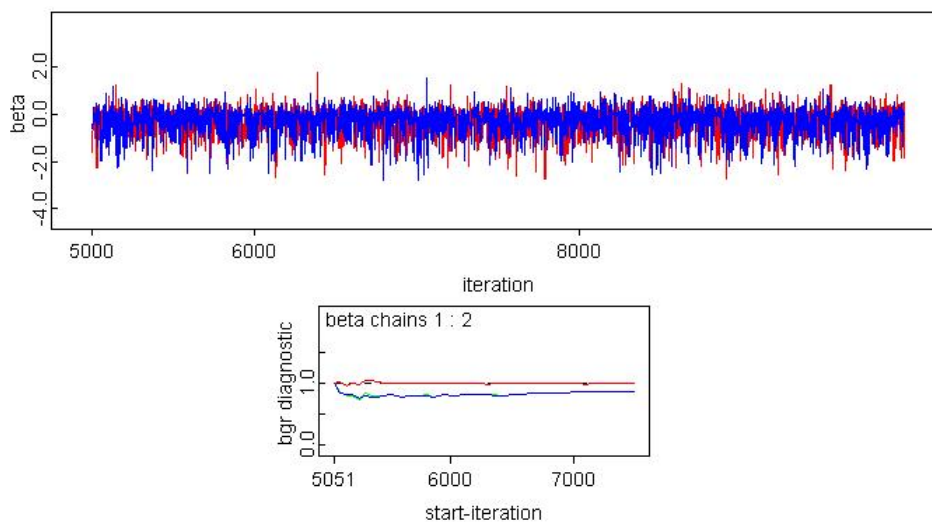


Figura 4.2: Convergencia para la deviance del Modelo GLM.1 usando OpenBUGS.

Para garantizar la convergencia de los parámetros estimados se realizan algunos test de diagnósticos propuestos en la teoría en el capítulo 3 (sección 3.2). Estos test se realizan partiendo del objeto creado al usar la función `bugs` desde R e invocando la librería CODA. CODA proporciona estadísticos descriptivos, cuantiles de la distribución empírica, errores de las medias calculados al suponer independencia y permite generar diferentes gráficas en las que se observa la convergencia de los parámetros. En la tabla 4.9 se muestra que existe poca correlación entre las variables α y β estimadas para los modelos propuestos. La figura 4.3 muestra que los parámetros superan el test de Geweke (1992), ya que cada variable está dentro de los límites de una dis-

Estimador	Lag 0	Lag 1	Lag 5	Lag 10	Lag 50
$\alpha_{95 94}$	1.00	-0.01	0.04	-0.02	0.007
$\beta_{95 94}$	1.00	-0.01	0.06	0.07	-0.05
$\alpha_{96 95}$	1.00	0.01	0.02	-0.01	0.009
$\beta_{96 95}$	1.00	-0.005	0.03	-0.05	0.01
$\alpha_{97 96}$	1.00	0.07	-0.02	0.04	-0.01
$\beta_{97 96}$	1.00	0.02	-0.06	0.05	-0.009
$\alpha_{98 97}$	1.00	0.02	0.03	0.005	0.02
$\beta_{98 97}$	1.00	0.02	0.01	-0.04	0.002

Tabla 4.9: Autocorrelación de α y β para la modelización propuesta usando CODA.

Estimador	Test de Estacionariedad	p-valor
α	pasado	0.198
β	pasado	0.56

Tabla 4.10: Diagnóstico Heidelberg para Modelo GLM.2 usando CODA.

tribución normal estandarizada. Así mismo, la tabla 4.10 muestra que α y β superan el test de Heidelberg y Welch (1983). Esta última prueba se basa en el estadístico de Cramer-von-Mises para probar la hipótesis nula de que los valores provienen de una distribución estacionaria. Se observa que ambos parámetros superan la prueba de estacionariedad.

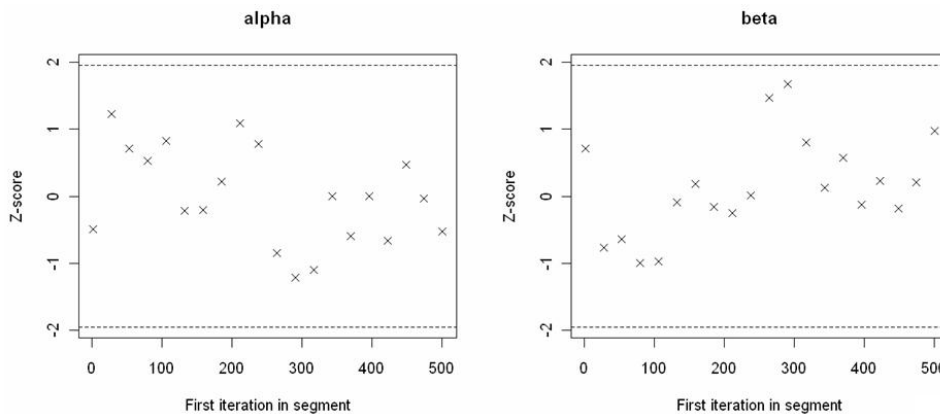


Figura 4.3: Diagnóstico Geweke para α y β del Modelo GLM.2 usando CODA.

Al comparar los resultados clásicos y los encontrados bajo el enfoque bayesiano no se observan diferencias numéricas importantes entre los criterios de bondad de ajuste y el número de parámetros efectivos en ambas metodologías. Los valores de α , en ambos casos, son muy

similares y β en la estimación bayesiana sigue presentando valores muy cercanos a cero, lo que coincide con la no significatividad de este parámetro en la estimación clásica. En ambos enfoques, se tiene que la bondad de ajuste (AIC y DIC, respectivamente) para el modelo del primer año es menor al obtenido para el modelo del último año (“peor ajuste”).

4.3. Estudio de modelos jerárquicos bayesianos espaciales

En este apartado se formula y analiza para cada año un modelo siguiendo la metodología propuesta en la sección 2.4. La estructura probabilística para estos modelos se presenta de dos formas. En una, se considera a las variables α y β iguales para cada año, es decir, se asigna a cada variable, una distribución normal con media 0 y varianza 0.001 (ecuaciones 4.8 y 4.9). En la otra, se cree que el efecto de α y β puede cambiar en cada árbol j , por esto en cada árbol j se asigna una previa como la definida en las ecuaciones 4.15 y 4.16.

Para la estructura probabilística de la primera forma se tiene lo siguiente

$$y_j|p_j \sim \text{Binomial}(1, p_j) \quad (4.6)$$

$$\text{logit}(p_j) = \alpha + \beta x_1 + \theta_j + \phi_j, \quad (4.7)$$

con $j=1, \dots, n$ -árboles, descontando el efecto de los árboles enfermos del año $i-1$; con $i=2, \dots, 5$.

$$\alpha \sim \text{Normal}(0, 0.001), \quad (4.8)$$

$$\beta \sim \text{Normal}(0, 0.001), \quad (4.9)$$

$$\theta_j \sim \text{Normal}(0, \sigma_\theta^2) \quad (4.10)$$

$$\phi_j|\phi_{i \neq j} \sim N\left(\sum_i \frac{w_{ji}y_i}{w_j^+}, \frac{\sigma_\phi^2}{w_j^+}\right) \quad (4.11)$$

donde $\sum_i w_{ji} = w_j^+$ y $w_{ji} = 1$ si i es vecino de j y 0 en caso contrario.

Para esta primera estructura, los hiperparámetros vienen dados por

$$\boxed{\sigma_\theta \sim \text{Unif}(0, 1), \quad \sigma_\phi \sim \text{Unif}(0, 1)} \quad (4.12)$$

Y para la segunda forma, se tiene la siguiente estructura probabilística

$$y_j|p_j \sim \text{Binomial}(1, p_j) \quad (4.13)$$

$$\text{logit}(p_j) = \alpha_j + \beta_j x_1 + \theta_j + \phi_j, \quad (4.14)$$

con $j= 1, \dots, n$ -árboles, descontando el efecto de los árboles enfermos del año $i - 1$; con $i = 2, \dots, 5$.

$$\alpha_j \sim \text{Normal}(0, \sigma_{\alpha_j}^2), \quad (4.15)$$

$$\beta_j \sim \text{Normal}(0, \sigma_{\beta_j}^2), \quad (4.16)$$

$$\theta_j \sim \text{Normal}(0, \sigma_{\theta_j}^2) \quad (4.17)$$

$$\phi_j|\phi_{i \neq j} \sim N\left(\sum_i \frac{w_{ji}y_i}{w_j^+}, \frac{\sigma_{\phi_j}^2}{w_j^+}\right) \quad (4.18)$$

donde $\sum_i w_{ji} = w_j^+$ y $w_{ji} = 1$ si i es vecino de j y 0 en caso contrario.

Los hiperparámetros para esta estructura se definen como

$$\sigma_a \sim \text{Unif}(0, 1) \quad (4.19)$$

$$\sigma_b \sim \text{Unif}(0, 1) \quad (4.20)$$

$$\sigma_\theta \sim \text{Unif}(0, 1) \quad (4.21)$$

$$\sigma_\phi \sim \text{Unif}(0, 1) \quad (4.22)$$

Las ecuaciones 4.11 y 4.18 corresponden a una distribución normal autoregresiva CAR. Los valores iniciales asignados a α y β para la ejecución del Gibbs, se fijaron considerando los resultados de las estimaciones obtenidas en los apartados anteriores.

Las tablas 4.11, 4.12, 4.13 y 4.14 contienen el resumen con la bondad de ajuste de los modelos formulados considerando la estructura definida en 4.6-4.12. Y las tablas 4.15, 4.16, 4.17 y 4.18 la bondad de ajuste de los modelos definidos de acuerdo a la estructura 4.13-4.22.

Modelo para Año 95	Descripción del Modelo	DIC	p_D
MJBE.1	$\alpha + \beta x_1 + \theta_j + \phi_j$	62.58	2.84
MJBE.2	$\alpha + \beta x_1 + \theta_j$	62.68	2.92
MJBE.3	$\alpha + \beta x_1 + \phi_j$	61.74	1.01

Tabla 4.11: Resumen con la deviance para los MJBE (Año 95), considerando la primera estructura propuesta.

Modelo para Año 96	Descripción del Modelo	DIC	p_D
MJBE.1	$\alpha + \beta x_1 + \theta_j + \phi_j$	164.3	8.42
MJBE.2	$\alpha + \beta x_1 + \theta_j$	165.0	7.29
MJBE.3	$\alpha + \beta x_1 + \phi_j$	162.3	0.80

Tabla 4.12: Resumen con la deviance para los MJBE (Año 96), considerando la primera estructura propuesta.

Modelo para Año 97	Descripción del Modelo	DIC	p_D
MJBE.1	$\alpha + \beta x_1 + \theta_j + \phi_j$	144.4	4.18
MJBE.2	$\alpha + \beta x_1 + \theta_j$	145.5	6.59
MJBE.3	$\alpha + \beta x_1 + \phi_j$	143.0	0.34

Tabla 4.13: Resumen con la deviance para los MJBE (Año 97), considerando la primera estructura propuesta.

Modelo para Año 98	Descripción del Modelo	DIC	p_D
MJBE.1	$\alpha + \beta x_1 + \theta_j + \phi_j$	289.8	17.0
MJBE.2	$\alpha + \beta x_1 + \theta_j$	290.4	12.06
MJBE.3	$\alpha + \beta x_1 + \phi_j$	288.7	1.73

Tabla 4.14: Resumen con la deviance para los MJBE (Año 98), considerando la primera estructura propuesta.

Modelo para Año 95	Descripción del Modelo	DIC	p_D
MJBE.4	$\alpha_j + \beta_j x_1 + \theta_j + \phi_j$	407.1	32.22
MJBE.5	$\alpha_j + \beta_j x_1 + \theta_j$	415.2	46.1
MJBE.6	$\alpha_j + \beta_j x_1 + \phi_j$	415.6	22.85

Tabla 4.15: Resumen con la deviance para los MJBE (Año 95), considerando la segunda estructura propuesta.

Modelo para Año 96	Descripción del Modelo	DIC	p_D
MJBE.4	$\alpha_j + \beta_j x_1 + \theta_j + \phi_j$	407.8	49.62
MJBE.5	$\alpha_j + \beta_j x_1 + \theta_j$	406.7	50.81
MJBE.6	$\alpha_j + \beta_j x_1 + \phi_j$	410.3	33.7

Tabla 4.16: Resumen con la deviance para los MJBE (Año 96), considerando la segunda estructura propuesta.

Modelo para Año 97	Descripción del Modelo	DIC	p_D
MJBE.4	$\alpha_j + \beta_j x_1 + \theta_j + \phi_j$	376.3	28.75
MJBE.5	$\alpha_j + \beta_j x_1 + \theta_j$	375.7	38.0
MJBE.6	$\alpha_j + \beta_j x_1 + \phi_j$	375.7	42.31

Tabla 4.17: Resumen con la deviance para los MJBE (Año 97), considerando la segunda estructura propuesta.

Modelo para Año 98	Descripción del Modelo	DIC	p_D
MJBE.4	$\alpha_j + \beta_j x_1 + \theta_j + \phi_j$	345.8	67.31
MJBE.5	$\alpha_j + \beta_j x_1 + \theta_j$	345.8	67.41
MJBE.6	$\alpha_j + \beta_j x_1 + \phi_j$	344.1	59.75

Tabla 4.18: Resumen con la deviance para los MJBE (Año 98), considerando la segunda estructura propuesta.

En este tipo de modelos, valores pequeños del DIC reflejarán un buen ajuste de los datos al modelo. Es así, que al analizar las tablas 4.11-4.18, se puede observar que de todos los modelos estudiados, los que presentan mejores ajustes corresponden a los modelos 1 y 3 para cada año

analizado. En estos modelos, esta presente la estructura espacial ϕ_j , lo que evidencia la presencia de un efecto espacial en el comportamiento del CTV y sugiere la existencia de un patrón epidemiológico de difusión de la enfermedad entre árboles vecinos.

El primer modelo es más complejo que el tercero (tablas, 4.11, 4.12, 4.13, 4.14), en este, se incluyen también los efectos aleatorios no espaciales θ_j . Los θ_j recogen la heterogeneidad existente en los árboles; valores elevados en la variabilidad de este efecto, indicarán que cada árbol aporta una variabilidad adicional al proceso.

A continuación se muestra información más detallada de las estimaciones encontradas para el modelo MJBE.1 correspondiente al año 97, así como algunas gráficas y test de diagnósticos obtenidas en CODA. No se muestra la convergencia para el resto de los modelos, porque el comportamiento resultante es similar al obtenido en el modelo MJBE.1 referenciado (detalles en apéndices 2 y 3).

Variable	Media	sd	Int.C 95 %
α	-2.53	0.26	[-3.08,-2.04]
β	0.12	0.21	[-0.28,0.60]
deviance	140.2	5.88	[124.7,149.1]
σ_θ	0.42	0.25	[0.06,0.95]
σ_ϕ	0.19	0.14	[0.07,0.64]
PV	0.35	0.18	[0.10,0.79]

Tabla 4.19: Estimaciones para el Modelo MJBE.1, correspondiente al año 97.

La tabla 4.19 muestra que tanto los efectos no espaciales (θ_j) como los espaciales (ϕ_j) son importantes y sugieren que la propagación del virus depende de la variabilidad propia de cada árbol y de la influencia de árboles vecinos enfermos. Aún cuando la desviación estándar para σ_θ y σ_ϕ sean elevadas al compararse con sus medias, el intervalo de credibilidad, muestra que en ambos casos, su presencia es relevante. Sin embargo, tanto σ_ϕ como la proporción de variación ($PV = \frac{\sigma_\phi}{\sigma_\theta + \sigma_\phi}$) al tener intervalos de menor longitud, muestran que la componente espacial ϕ tiene un efecto más claro que el efecto proporcionado por θ .

En cuanto a los valores de α se tiene un intervalo de credibilidad pequeño con valores negativos. Mientras que β asume valores que van desde negativos hasta positivos, coincidiendo con los resultados encontrados en las secciones 4.2. En el caso de β , el intervalo de credibilidad contiene al cero.

Probabilidad[árbol]	Media	sd	Int.C 95 %
p[40]	0.05	0.002	[0.04,0.09]
p[78]	0.1195	0.006	[0.038,0.298]
p[156]	0.13	0.07	[0.02,0.33]

Tabla 4.20: Probabilidades de algunos árboles para el Modelo MJBE.1, correspondiente al año 97.

La tabla 4.20 ilustra la aplicación de la transformación inversa del logit para el cálculo de probabilidades, empleando las estimaciones encontradas en el modelo MJBE.1. Se observa como el árbol en la posición 40 presenta una probabilidad más pequeña a la estimada para los árboles en las posiciones 78 y 156.

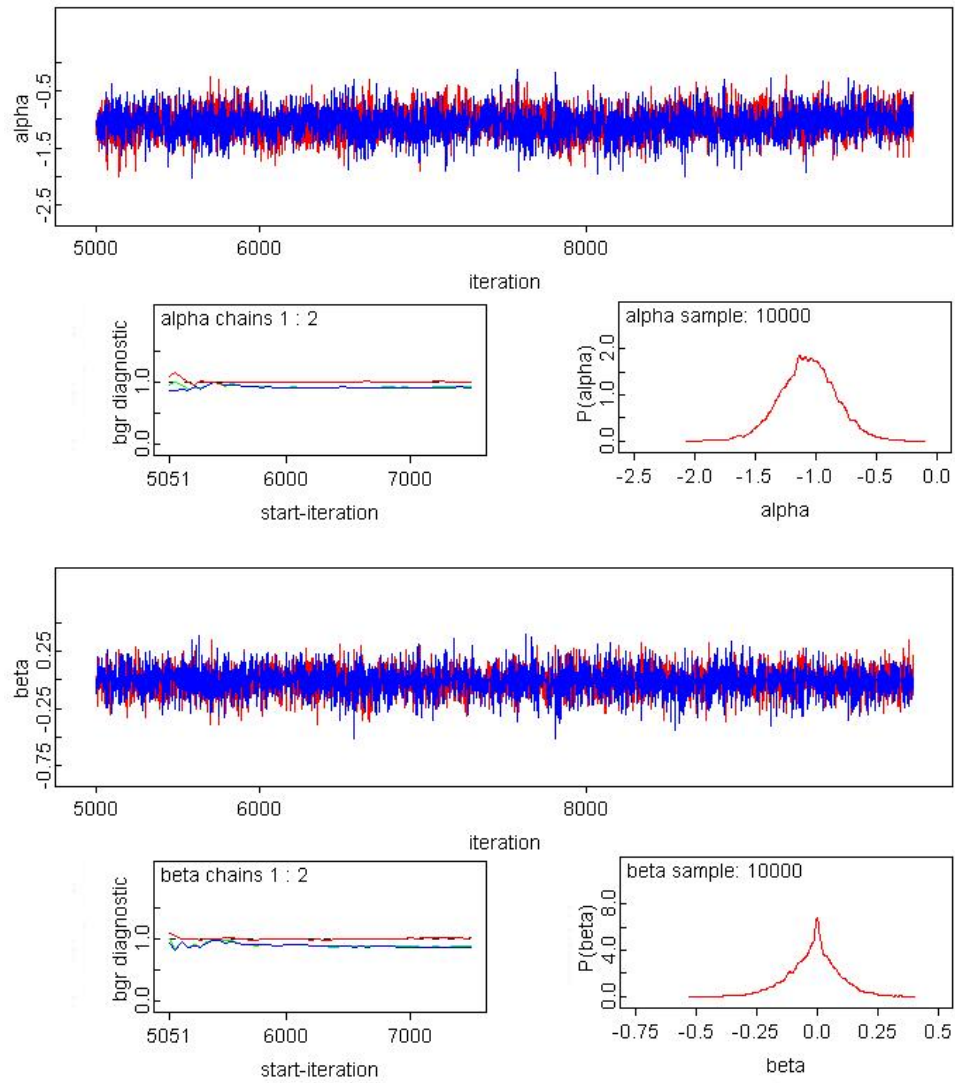


Figura 4.4: Convergencia de α y β en el modelo MJBE.1, correspondiente al año 97.

Estadístico Z para α	Estadístico Z para β
-0.77	1.31

Tabla 4.21: Diagnóstico Geweke para α y β en el modelo MJBE.1, correspondiente al año 97.

Variable	Test de Estacionariedad	p-valor
α	pasado	0.37
β	pasado	0.14

Tabla 4.22: Diagnóstico Heidelberger para modelo MJBE.1, correspondiente al año 97.

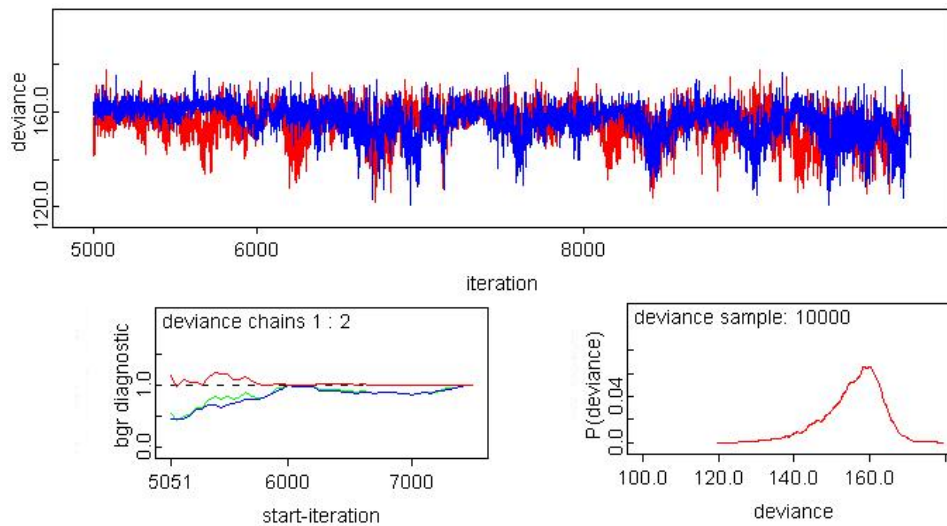


Figura 4.5: Convergencia de la deviance en el modelo MJBE.1, correspondiente al año 97.

La figura 4.4(b) muestra que tanto α como β alcanzan muy pronto la convergencia en las dos cadenas simuladas. Igualmente sucede con la convergencia de la deviance (figura 4.5). Se observa como la deviance supera el test de Gelman y Rubin (1992) ya que ambas cadenas se estabilizan cerca del valor 1.0.

Analizando los resultados de la tabla 4.21, se tiene que ambos parámetros superan el test de Geweke (1992), debido a que cada variable asume valores en el rango $[-3,3]$ correspondientes a una variable normal estandarizada. Así mismo, la tabla 4.22 muestra que α y β superan el test de Heidelberger y Welch (1983).

4.4. Estudio de modelos jerárquicos bayesianos espaciales dinámicos

Resulta tarea laboriosa hacer los análisis de los modelos por año, por lo tanto, se puede abordar el problema escribiendo modelos conjuntos que recojan tanto la información espacial como temporal de los datos. Con esta formulación se espera no sólo mejorar la estimación de los parámetros obtenidos en la sección previa al contar con más grados de libertad sino incrementar la bondad de ajuste en los modelos formulados.

Aún cuando la formulación conjunta considera el tiempo, no se aplica la teoría de series temporales propiamente (no se consideran periodos de retardo (“lag”) para modelizar cada observación), sólo se tiene en cuenta la información contenida en cada año y se incluye en el modelo. Esta generalización aprovecha la relación espacial que existe en cada instante de tiempo considerado e incorpora de alguna forma el concepto espacio-temporal en la definición de los modelos estudiados, es así, que en adelante nos referiremos a este tipo de modelos como modelos jerárquicos bayesianos espaciales dinámicos, abreviados como MJBED.

Es importante destacar, que tanto en los modelos bayesianos dinámicos como en los modelos bayesianos por analizados por año, el efecto del pasado se recoge en la covariable x_1 , mientras que el efecto del presente se considera al incluir la componente espacial ϕ .

La concatenación espacial y temporal requerirá de mayor esfuerzo computacional, pero gracias a las técnicas de simulación MCMC es posible estimar los parámetros de interés usando el programa OpenBUGS (detalles en apéndice 1).

Para estos modelos se plantean dos modificaciones en las estructuras probabilísticas consideradas y similares a la descrita en la sección anterior. Es decir, se estudiarán modelos dinámicos que van desde los más sencillos a los más complejos. La primera estructura será similar a la definida en las ecuaciones 4.6-4.12, sólo que ahora se tendrán dos sub-índices en las variables aleatorias de interés, con los cuales se indicarán conjuntamente el árbol y el tiempo, siendo j el sub-índice para árbol e i el indicador para el año. Con la segunda estructura probabilística, se intenta recoger el hecho de que los árboles puedan compartir una distribución común en cada año i a través de un efecto aleatorio.

La primera estructura considerada, será de la forma

$$y_{ji}|p_{ji} \sim \text{Binomial}(1, p_{ji}) \quad (4.23)$$

$$\text{logit}(p_{ji}) = \alpha + \beta x_{1ji} + \theta_{ji} + \phi_{ji} \quad (4.24)$$

$$\alpha \sim \text{Normal}(0, 0.001), \quad (4.25)$$

$$\beta \sim Normal(0, 0 \cdot 001), \quad (4.26)$$

$$\theta_{ji} \sim Normal(0, \sigma_{\theta}^2) \quad (4.27)$$

$$\phi_{ji} \sim CAR(\sigma_{\phi}^2) \quad (4.28)$$

Los hiperparámetros para esta estructura vienen dados por

$$\boxed{\sigma_{\theta} \sim Unif(0, 1), \quad \sigma_{\phi} \sim Unif(0, 1)} \quad (4.29)$$

donde x_{ji} representa el número de árboles infectados en torno al árbol_j observados en el año i . Esta covariable recoge el efecto del pasado, es decir, calcula cuántos árboles infectados existen en el año actual como consecuencia de los árboles ya infectados en años anteriores, considerando para ello distancias ≤ 10 metros. Con esta primera propuesta, se pretende modelizar la probabilidad de un árbol enfermar en función al efecto de la covariable x_{ji} y a la influencia de factores de riesgo ocultos (θ , ϕ). Al mantener α y β iguales en cada árbol durante los años estudiados, se intenta recoger el hecho de que los árboles puedan compartir parámetros comunes.

El número de árboles en cada año es diferente, pues dependerá del número de árboles sanos una vez eliminados los árboles enfermos del año inmediatamente anterior al año i analizado. Por tanto, con $j=1, \dots, n$ -árboles, descontando el efecto de los árboles enfermos del año $i-1$; con $i=2, \dots, 5$.

Y la segunda estructura, viene dada por

$$y_{ji}|p_{ji} \sim Binomial(1, p_{ji}) \quad (4.30)$$

$$logit(p_{ji}) = \alpha_i + \beta_i x_{1,ji} + \theta_{ji} + \phi_{ji} \quad (4.31)$$

$$\alpha_i \sim Normal(0, \sigma_{\alpha_i}^2), \quad (4.32)$$

$$\beta_i \sim Normal(0, \sigma_{\beta_i}^2), \quad (4.33)$$

$$\theta_{ji} \sim Normal(0, \sigma_{\theta_i}^2) \quad (4.34)$$

$$\phi_{ji} \sim CAR(\sigma_{\phi}^2) \quad (4.35)$$

Los hiperparámetros para esta segunda estructura se definen de la siguiente forma

$$\sigma_{a_i} \sim Unif(0, 1) \quad (4.36)$$

$$\sigma_{b_i} \sim Unif(0, 1) \quad (4.37)$$

$$\sigma_{\theta_i} \sim Unif(0, 1) \quad (4.38)$$

$$\sigma_{\phi} \sim Unif(0, 1) \quad (4.39)$$

Al proponer esta segunda reformulación, se desea incorporar mayor complejidad en los modelos anteriormente propuestos, al incrementar el número de parámetros a estimar. Bajo esta formulación, se considera que tanto α como β cambian en cada año i . Además, el efecto de heterogeneidad (θ) se supone igual para cada árbol $_j$ pero distinto entre los años i . Sin embargo para el efecto espacial (ϕ), se plantean dos variantes, una, donde la desviación estándar σ_{ϕ} se deja constante para cada año i , y otra, donde se asignan previas uniformes (0,1) distintas a cada año i (σ_{ϕ_i}).

Los modelos jerárquicos bayesianos espaciales dinámicos que surgen de las estructuras (4.23-4.29) y (4.30-4.39), se formulan para considerar diferentes escenarios, a partir de los cuales sea posible estudiar la dinámica de la enfermedad y el movimiento espacial del vector transmisor del virus CTV.

La tabla 4.23 contiene el resumen con los ajustes obtenidos para los modelos formulados considerando la estructura definida en 4.23-4.29. Y la tabla 4.24 presenta los ajustes para los modelos dados por las ecuaciones 4.30-4.39.

Modelo	Descripción del Modelo	DIC	p_D
MJBED.1	$\alpha + \beta x_{1_{ji}} + \theta_{ji} + \phi_{ji}$	622.7	20.71
MJBED.2	$\alpha + \beta x_{1_{ji}} + \theta_{ji}$	654.7	16.67
MJBED.3	$\alpha + \beta x_{1_{ji}} + \phi_{ji}$	625.5	11.18

Tabla 4.23: Resumen con la bondad de ajuste para los modelos MJBED, de acuerdo a la primera estructura.

Modelo	Descripción del Modelo	DIC	p_D
MJBED.4	$\alpha_i + \beta_i x_{1_{ji}} + \theta_{ji} + \phi_{ji}$	627.4	29.26
MJBED.5	$\alpha_i + \beta_i x_{1_{ji}} + \theta_{ji}$	668.3	28.12
MJBED.6	$\alpha_i + \beta_i x_{1_{ji}} + \phi_{ij}$	634.5	18.24
MJBED.7	$\alpha_i + \beta_i x_{1_{ji}} + \theta_{ji} + \phi_{ji}$	631.5	22.37
MJBED.8	$\alpha_i + \beta_i x_{1_{ji}} + \phi_{ji}$	642.4	15.5

Tabla 4.24: Resumen con la bondad de ajuste para los modelos MJBED, de acuerdo a la segunda estructura.

En las tablas 4.23 y 4.24 se recogen los valores de DIC y p_D de cada modelo dinámico estudiado, se observa como los MJBED.1 y MJBED.3 siguen siendo los que mejores ajuste presentan (menores DIC). Estos resultados coinciden con los encontrados en la sección 4.3. En ambos modelos se presenta la estructura espacial ϕ lo que evidencia que la probabilidad de un árbol enfermar dependerá de la existencia de vecinos enfermos. Así mismo, en el modelo MJBED.1 se tiene la influencia del efecto no espacial (θ), que incorpora la variabilidad no espacial en la probabilidad de un árbol enfermar.

Si se trabaja con el modelo MJBED.1 se tiene que la probabilidad de un árbol enfermar viene dada por la influencia de varios factores, uno, por la presencia de árboles enfermos de años anteriores, otro, por la variabilidad implícita en cada árbol y por último, de un efecto espacial que recoge la propagación del virus en el presente. Si por el contrario, se decide trabajar con el modelo MJBED.3, se tendrá que la probabilidad de un árbol enfermar viene determinada por la influencia de árboles enfermos de años anteriores y por el efecto espacial en el presente.

Los modelos MJBED.7 y MJBED.8 fueron obtenidos al asignar previas distintas en las desviaciones estándar que definen las varianzas de los parámetros de interés. Mientras que en los modelos MJBED.4, MJBED.5 y MJBED.6 las previas para α y β son distintas en cada año, pero la desviación estándar para el efecto ϕ se considera igual en todos los años.

Variable	Media	sd	Int.C 95 %
α	-2.61	0.16	[-2.96,-2.32]
β	0.34	0.10	[0.13,0.54]
deviance	602.0	19.55	[550.6,630.9]
σ_θ	0.307	0.24	[0.028,0.898]
σ_ϕ	0.53	0.11	[0.32,0.78]
PV	0.68	0.18	[0.35,0.95]

Tabla 4.25: Estimaciones para el modelo MJBED.1, correspondientes a $i=4$.

En la tabla 4.25 se observa como ambos efectos aleatorios son importantes (θ y ϕ). En especial, se observa como σ_ϕ y la proporción de variabilidad espacial respecto a la no espacial indican que existe un efecto espacial evidente y que la probabilidad de un árbol enfermar dependerá significativamente de la influencia de árboles enfermos cercanos.

Si se comparan los valores de α y β con los obtenidos en la sección 4.3 se tiene que β resulta de mayor magnitud, mientras que α mantiene valores similares. Con este resultado se evidencia que la estimación de β mejora al tener mayor información. β asume valores mayores a cero a diferencia de lo que sucedía en las secciones 4.2 y 4.3, por tanto queda claro la importancia del efecto del pasado en la descripción del comportamiento de la enfermedad.

Se debe mencionar que en el modelo MJBED.7 fueron asignadas previas uniformes (0,1)

diferentes en las desviaciones estándar que definen en la varianza en los primeros tres años, sin embargo, para el último año no fue posible asignar una previa uniforme, pues el algoritmo al generar las condicionales para ϕ_{ji} en la simulación generaba un error. Para eliminar tal error y ejecutar el algoritmo Gibbs se asignó una previa gamma(1,0.1) a la desviación estándar de ϕ en el último año. La asignación de esta previa menos informativa que la previa uniforme(0,1), evidencia una vez más la existencia de diferencias espaciales importantes en los datos.

En el trabajo de Spiegelhalter et al. (2002), los autores sugieren que modelos con valores del DIC que superan en menos de 3 unidades al ‘mejor’ (modelo con menor DIC) son ‘equivalentes’ en cuanto a su capacidad predictiva. Los modelos cuyo DIC supera entre 3 y 7 unidades al modelo con menor DIC se consideran ‘ligeramente inferiores’ y finalmente, los que superan en más de 7 unidades al menor DIC son considerados sustancialmente inferiores. En cuanto a los resultados presentados en la tabla 4.23 se puede concluir que los modelos MJBED.1 y MJBED.3 son equivalentes en cuanto a su capacidad de predicción, mientras que el modelo MJBED.4 es ligeramente inferior al mejor modelo encontrado (MJBED.1), siendo el resto de los modelos sustancialmente inferiores al mejor modelo.

Al sumar los menores DIC de los modelos MJBE con mejor ajuste por año y comparar este total con el DIC del mejor modelo dinámico respectivo, se tiene que numéricamente el criterio de información de deviance es menor en el modelo dinámico, por tanto, se comprueba numéricamente que al conjugar la información espacial y temporal se gana en bondad de ajuste.

Probabilidad[árbol]	Media	sd	Int.C 95 %
p[40,3]	0.05	0.03	[0.01,0.13]
p[78,3]	0.11	0.05	[0.03,0.24]
p[156,3]	0.07	0.04	[0.02,0.18]
p[257,3]	0.30	0.13	[0.097,0.62]
p[270,3]	0.36	0.17	[0.08,0.71]

Tabla 4.26: Algunas probabilidades estimadas para el modelo MJBED.1, correspondientes al año 97.

La tabla 4.26 muestra que los árboles observados en las posiciones 40, 78 y 156 en $i=3$ (año 97), presentan probabilidades similares a las obtenidas en la sección anterior. Además se incorporan probabilidades de árboles ubicados en las posiciones 257 y 270. Ambas probabilidades estimadas son altas, aún cuando sus desviaciones estándar también lo son.

A continuación se muestra información detallada de las estimaciones para el modelo MJBED.1 (tabla 4.25). Se ha elegido este modelo para ilustrar el comportamiento en la convergencia de los parámetros estimados, pero pudo escogerse cualquier otro modelo, ya que el comportamiento observado es similar (detalles en apéndice 4).

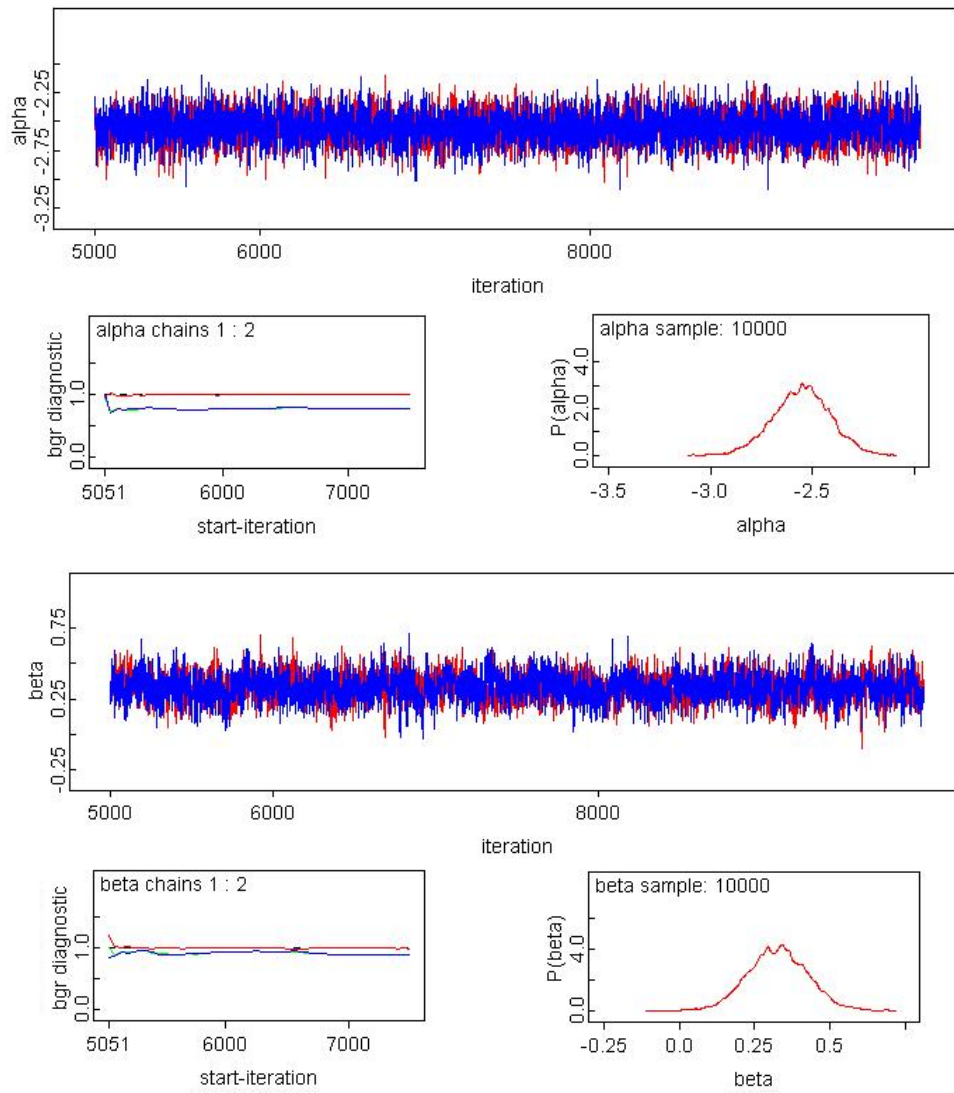


Figura 4.6: Convergencia para α y β en el modelo MJBED.1, correspondiente al año 97.

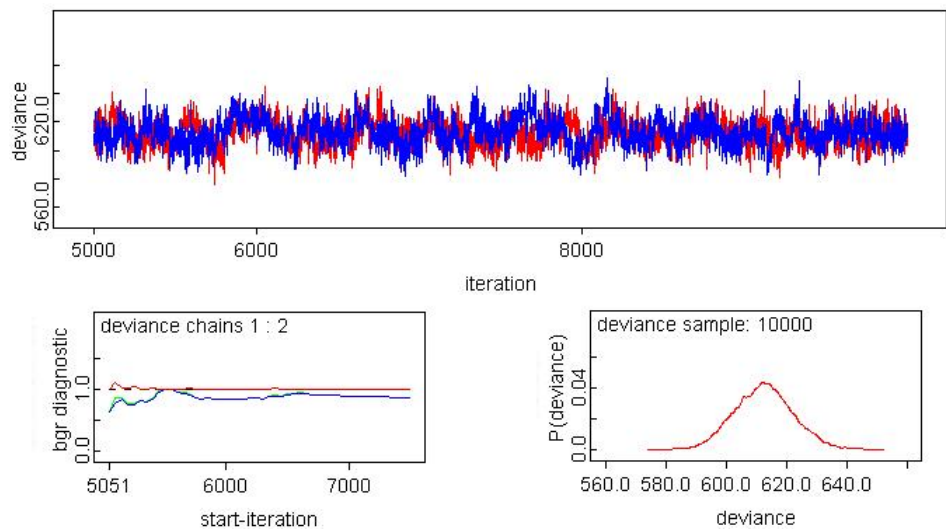


Figura 4.7: Convergencia de la deviance en el modelo MJBED.1, correspondiente al año 97.

La figura 4.6(b) muestra que tanto α como β alcanzan la convergencia en las dos cadenas simuladas. Se observa como la deviance supera el test de Gelman y Rubin (1992) ya que ambas cadenas se estabilizan alrededor del valor 1.0 (figura 4.7).

Analizando los resultados de la tabla 4.27, se tiene que ambos parámetros superan el test de Geweke (1992), ya que cada variable asume valores en el rango $[-3,3]$ correspondientes a una variable normal estandarizada. Así mismo, la tabla 4.28 muestra que α y β superan el test de Heidelberger y Welch (1983).

Estadístico Z para α	Estadístico Z para β
-2.44698	-0.08599

Tabla 4.27: Diagnóstico Geweke para α y β en el modelo MJBED.1, correspondiente al año 97.

Variable	Test de Estacionariedad	p-valor
α	pasado	0.0696
β	pasado	0.1213

Tabla 4.28: Diagnóstico Heidelberger para modelo MJBED.1, correspondiente al año 97.

Capítulo 5

Conclusiones y Líneas Futuras

5.1. Conclusiones

Al analizar los resultados y compararlos es posible determinar que el uso de modelos jerárquicos bayesianos espaciales constituyen una metodología de trabajo compleja y capaz de capturar mayor información que la recogida por los métodos clásicos aplicados en este trabajo. Se puede ver en la sección 4.3 que al ajustar los modelos bajo esta metodología los resultados son numéricamente similares a los encontrados en la sección 4.2, pero además, son capaces de incorporar fuentes de variabilidad no espacial y espacial que bajo el enfoque clásico sería muy difícil considerar.

Por otra parte, se tienen que las estimaciones alcanzadas con los modelos jerárquicos bayesianos espaciales dinámicos es más clara, es decir, los valores de β , σ_ϕ y la proporción de variación (PV) es mayor que las obtenidas en la sección 4.3. Se destaca la estimación de β en los modelos dinámicos ya que expresan que la covariable x_1 (“efecto del pasado”) tiene una importancia mayor a la observada en las estimaciones halladas tanto con la metodología clásica como con los modelos bayesianos estudiados por año.

De los modelos formulados en las secciones 4.3 y 4.4, los modelos 1 y 3 son los que menores DIC y p_D presentan. Estos modelos sugieren la existencia de una estructura espacial en el comportamiento de la enfermedad, indicando que durante los años estudiados ha estado activa la presencia del pulgón, principal vector transmisor del virus CTV. La poca diferencia en los valores del DIC en ambos modelos, sugiere que cualquiera de ellos es capaz de recoger la dinámica espacial de la infección y el comportamiento del pulgón.

Las técnicas MCMC junto con el programa OpenBUGS facilitan la implementación de estos modelos cuya estructura probabilística sería imposible de tratar en otros escenarios.

Después de analizar los resultados encontrados, se pueden mencionar las siguientes conclusiones generales:

- En las estimaciones usando Modelos Lineales Generalizados (GLM) desde el punto de vista

clásico y bayesiano, sólo se encuentra significativo el parámetro α . Es decir, la probabilidad global de un árbol enfermar depende únicamente de este parámetro.

- Los modelos jerárquicos bayesianos espaciales formulados por año, alcanzan valores para α y β similares a los obtenidos en los GLM. Sin embargo, encuentran que los efectos aleatorios (espaciales y no espaciales) inciden notablemente en la probabilidad de un árbol enfermar.
- Los modelos jerárquicos bayesianos espaciales dinámicos, mejoran las estimaciones de β y evidencian que la probabilidad de un árbol enfermar dependerá de la influencia tanto de árboles ya infectados a su alrededor como a la aparición de nuevos casos.
- La metodología empleada sugiere la presencia de patrones de contagio entre árboles ubicados a distancias menores a los 10 metros, así como un patrón de movimiento en el principal vector transmisor *A. gossypii* a estas distancias.
- La metodología propuesta permite considerar diversas fuentes de información y evidenciar gracias a la estructura espacial identificada, la presencia activa de la enfermedad durante el tiempo de estudio.
- Los modelos jerárquicos añaden cierta complejidad a los modelos tradicionales, permitiendo la construcción de estructuras flexibles mediante el encadenamiento condicional de modelos simples.
- Los modelos jerárquicos bayesianos espaciales pueden ser herramientas muy útiles en estudios epidemiológicos, con los cuales, sea posible estudiar la incidencia y la extensión de una enfermedad contagiosa en plantas.

5.2. Líneas futuras

Este trabajo es apenas un comienzo, los resultados encontrados pueden ampliarse incorporando nuevas fuentes de información y ajustando modelos más complejos. Después de aplicar la metodología propuesta se pueden señalar las líneas futuras de investigación siguientes:

- Incorporar la distancia como variable aleatoria, lo que supondría incrementar la complejidad de los modelos jerárquicos bayesianos espaciales.
- Combinar herramientas de modelización espacial con herramientas de modelización temporal, con lo cual, sea posible proponer técnicas de modelización espacio-temporal que permitan representar la dinámica compleja del virus CTV en cultivos agrícolas.

- Considerar otros tipos de datos espaciales y las metodologías propias de cada uno de ellos y su aplicación en problemas epidemiológicos relacionados con otras especies de plantas.
- Construir modelos que permitan predecir la evolución de la enfermedad en cítricos considerando la componente espacial y temporal.
- Proponer nuevas estructuras de vecindad e incorporarlas en la estructura del modelo jerárquico bayesiano, de acuerdo al tipo de dato espacial considerado.

Apéndices

Apéndice 1

En este apartado se presentan los gráficos de convergencia para la deviance de algunos modelos formulados usando por un lado previas gammas en las desviaciones estándar que definen a los parámetros de precisión y por el otro, usando previas uniformes. Con los gráficos que se muestran en este anexo se desea ilustrar el por qué se decidió trabajar con previas informativas (uniformes) y no con distribuciones gammas.

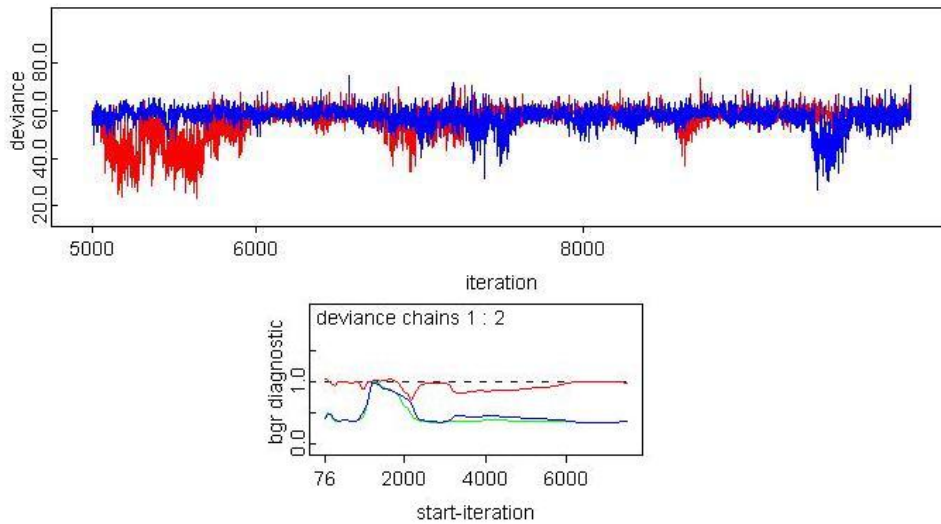


Figura 5.1: Convergencia para la deviance en el modelo MJBE para el año 95, usando previas gammas después de 10000 iteraciones.

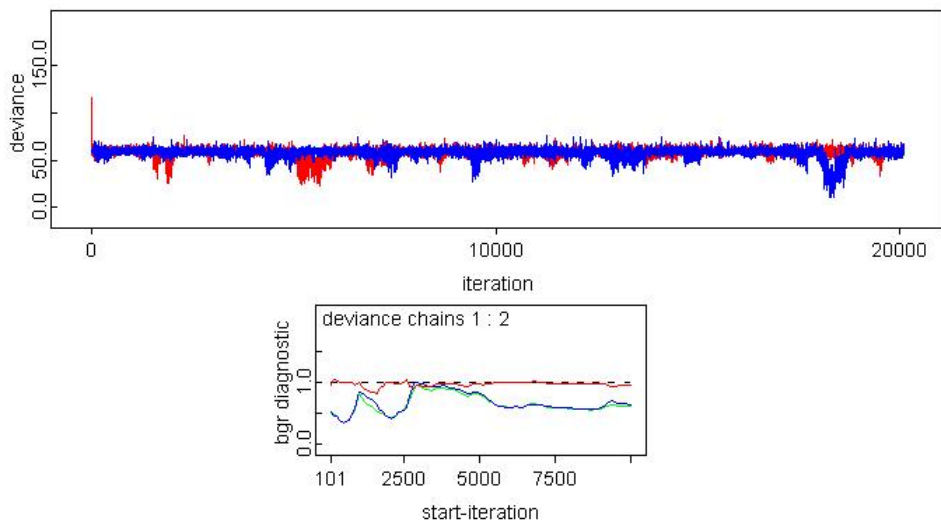


Figura 5.2: Convergencia para la deviance en el modelo MJBE para el año 95, usando previas gammas después de 20100 iteraciones.

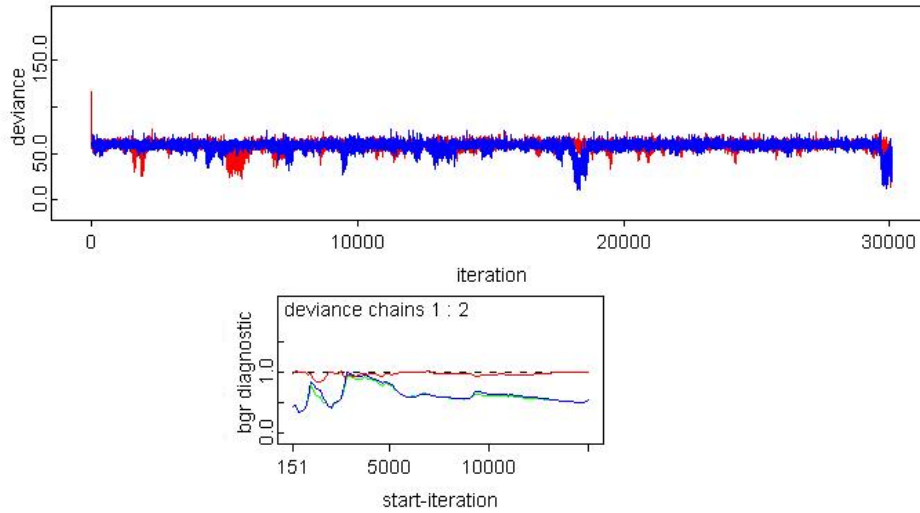


Figura 5.3: Convergencia para la deviance en el modelo MJBE para el año, 95 usando previas gammas después de 30000 iteraciones.

Las figuras 5.1, 5.2 y 5.3 muestran la evolución de la convergencia para la deviance del modelo MJBE.1 correspondiente al año 95, después de quemar las 5000 primeras iteraciones. Se observa en todos los casos que el diagnóstico propuesto por Gelman y Rubin (1992) no se supera y que la convergencia no se alcanza en la cadena de color azul. A la cadena de color azul se asignaron valores iniciales distintos de cero para α y β , mientras que a la cadena de color rojo se asignan ceros como valores iniciales a cada parámetro.

En la tabla 5.1 se muestra el código escrito en OpenBUGS para el modelo MJBE.1 correspondiente al año 95. Es importante destacar, que en el modelo definido en la tabla 5.1 se intentó asignar a τ_ϕ una distribución gamma más informativa usando la previa $\tau_\phi \sim \text{dgamma}(1,0.1)$, pero el nodo $\phi[1:\text{narb}] \sim \text{car.normal}(\text{adj}[\], \text{weights}[\], \text{num}[\], \tau_\phi)$ no pudo inicializarse, por tanto, se usó la previa $\tau_\phi \sim \text{dgamma}(0.5,0.005)$.

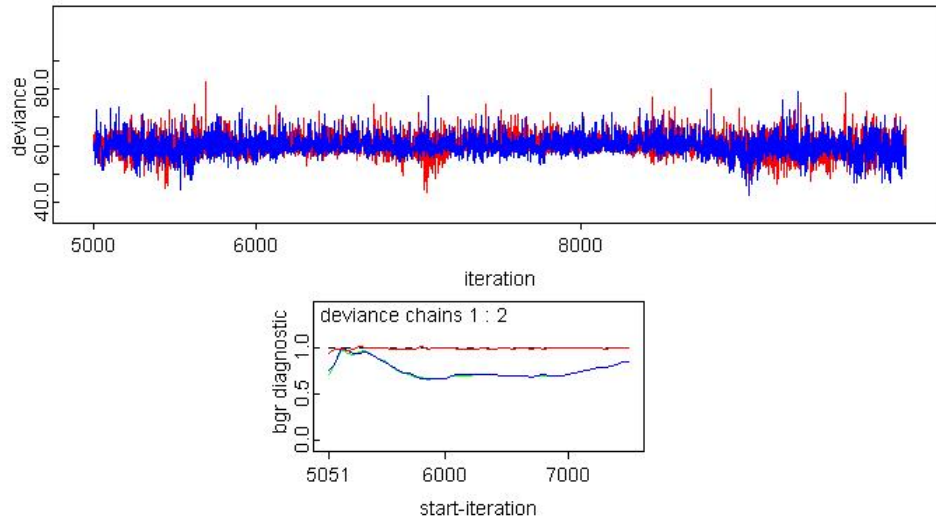


Figura 5.4: Convergencia para la deviance en el Modelo MJBE para el año 95, usando previas uniformes para 10000 iteraciones.

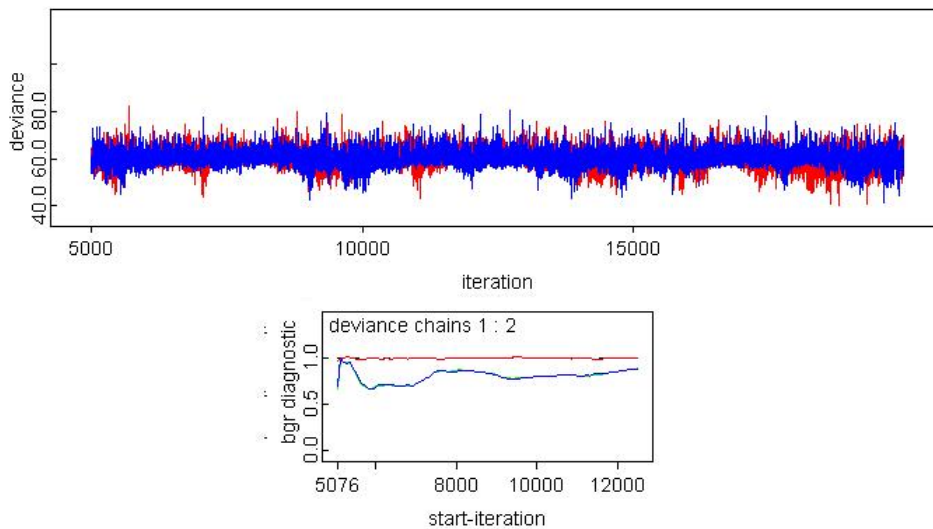


Figura 5.5: Convergencia para la deviance en el Modelo MJBE para el año 95, usando previas uniformes para 18000 iteraciones.

Las figuras 5.4 y 5.5 muestran una estabilidad en la convergencia más rápida (menos iteraciones) que las presentadas en las figuras 5.1, 5.2 y 5.3. Además el test de Gelman y Rubin (1992) en ambas figuras, muestra como la convergencia para la deviance se alcanza aproximadamente después de la iteración 8000.

En el modelo definido en la tabla 5.2 se induce el valor de la precisión requerida para cada parámetro. Un ejemplo de esto, $\tau_\phi = \frac{1}{\sigma_\phi^2}$ con $\sigma_\phi \sim \text{dunif}(0,1)$.

```

model{
for (j in 1 : narb) {#Nro. arboles considerados < 300
    O.t2[j]~ dbern(p[j]) # casos observados para t=1
                        #(año 95 dado enfermos 94)
    logit(p[j])<-alpha + beta*vauxt2[j] + theta[j] + phi[j]
    theta[j]~ dnorm(0.0,tau.y)
} # for de j

# CAR: distribucion previa para los efectos aleatorios espaciales
phi[1:narb] ~ car.normal(adj[], weights[], num[], tau.phi)

# radio de variacion espacial vs. variacion global
sd.t<-sd(theta[1:narb])
sd.phi<-sd(phi[1:narb])
prop.v<- sd.phi/(sd.t+sd.phi)
# prop.v = Representa la proporción de la variabilidad
# total que es atribuible a la componente espacial

# Asignaciones de previas
tau.phi ~ dgamma(0.5,0.005)
tau.y~dgamma(1,0.1)

alpha~dnorm(0.0,0.001)
beta ~ dnorm(0.0,0.001)
} # end model
# Valores iniciales
list(alpha = 0, beta = 0) # Cadena de color ROJO
list(alpha = 0.5, beta = -0.5)# Cadena de color AZUL

```

Tabla 5.1: Sintaxis en OpenBUGS para el modelo MJBE.1 correspondiente al año 95, usando previas gammas.

Para demostrar que los comportamientos observados en las figuras 5.1, 5.2, 5.3, 5.4 y 5.5 también se presentan en el resto de los modelos espaciales considerados por año y en los formulados en forma dinámica, se presenta a continuación la convergencia para la deviance del modelo MJBED.1 usando tanto previas gammas como uniformes en las desviaciones estándar que definen a las precisiones.

```
model
{
  for (j in 1 : narb) { #Nro. arboles considerados < 300
    O.t2[j]~dbern(p[j]) # casos observados para t=1
                        # (año 95 dado enfermos 94)
    logit(p[j])<-alpha + beta*vauxt2[j] + theta[j] + phi[j]
    theta[j]~dnorm(0.0,tau.y)
  } # for de j

# CAR: distribucion previa para los efectos aleatorios espaciales
phi[1:narb] ~ car.normal(adj[], weights[], num[], tau.phi)

# radio de variacion espacial vs. variacion global
sd.t<-sd(theta[1:narb])
sd.phi<-sd(phi[1:narb])
prop.v<- sd.phi/(sd.t+sd.phi)
# prop.v = Representa la proporción de la variabilidad
# total que es atribuible a la componente espacial

# Definicion de precisiones
tau.phi <- 1/ (sigma.phi*sigma.phi)
tau.y<- 1/ (sigma.y*sigma.y)

alpha~dnorm(0.0,0.001)
beta ~ dnorm(0.0,0.001)

# Asignaciones de previas para sigma
sigma.phi~dunif(0,1)
sigma.y~dunif(0,1)
} # end model

#Valores iniciales
list(alpha = 0, beta = 0) # Cadena de color ROJO
list(alpha = 0.5, beta = -0.5)# Cadena de color AZUL
```

Tabla 5.2: Sintaxis en OpenBUGS para el modelo MJBE.1 correspondiente al año 95, usando previas uniformes.

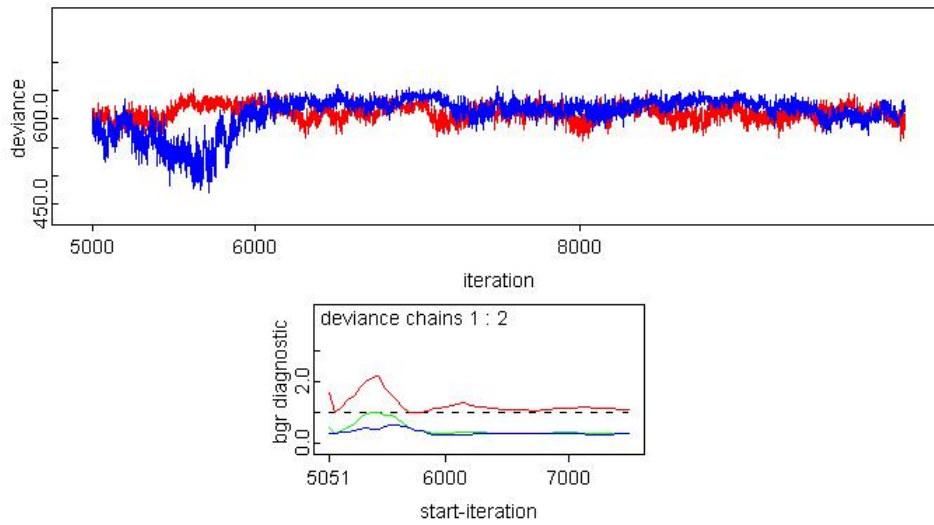


Figura 5.6: Convergencia para la deviance en el modelo MJBED.1 usando previas gammas (10000 iteraciones).

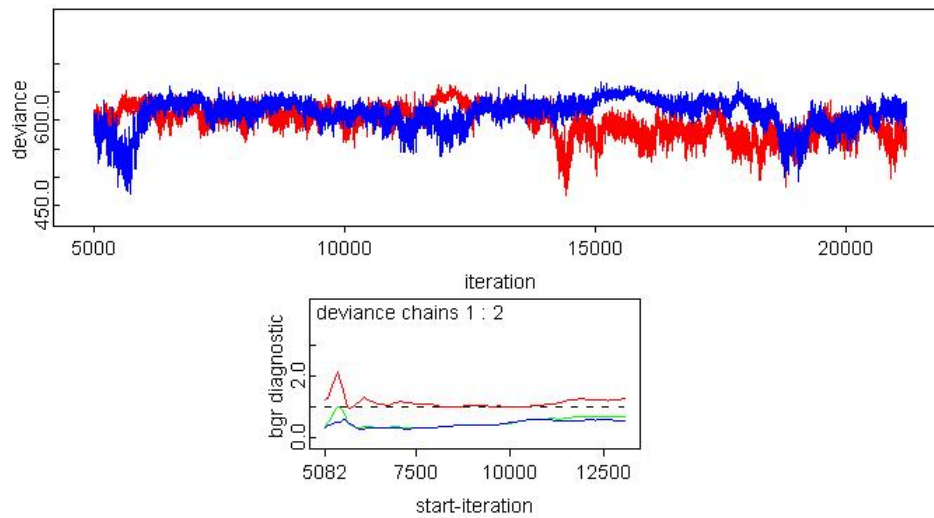


Figura 5.7: Convergencia para la deviance en el Modelo MJBED.1 usando previas gammas (después de 25000 iteraciones).

Las figuras 5.6 y 5.7 muestran como cuesta alcanzar la convergencia en la deviance del modelo MJBED.1 aún después de 25000 iteraciones.

El modelo definido en la tabla 5.3, se muestra que para definir a ϕ_1 y ϕ_3 se asignaron gammas diferentes a los parámetros de precisión en estas distribuciones, ya que al intentar iniciar la ejecución de los nodos $\phi_1[1,1:N[1]]$ y $\phi_3[1,1:N[3]]$ se generó un error en el algoritmo Gibbs. Por tanto, fue necesario establecer previas gammas diferentes en cada nodo $\phi[i,1:N[i]], i= 1, \dots, 4$.

```

model
{# Modelo jerárquico bayesiano espacial dinámico
 # se considera en forma conjunta tanto los datos
 # observados como el tiempo
for (t in 1:k[1]) {weights1[t]<-1}
for (t in 1:k[2]) {weights2[t]<-1}
for (t in 1:k[3]) { weights3[t]<-1}
for (t in 1:k[4]) {weights4[t]<-1}
# Distribuciones previas CAR (efectos aleatorios espaciales)
phi1[1,1:N[1]] ~ car.normal(adj1[], weights1[], num1[], tau.phi1)
phi2[1,1:N[2]] ~ car.normal(adj2[], weights2[], num2[], tau.phi)
phi3[1,1:N[3]] ~ car.normal(adj3[], weights3[], num3[], tau.phi3)
phi4[1,1:N[4]] ~ car.normal(adj4[], weights4[], num4[], tau.phi)

for (j in 1:N[1]) {phi[1,j]<-phi1[1,j]}
phi1[1,300]<-0
for (j in 1:N[2]) {phi[2,j]<-phi2[1,j]}
for (j in 1:7) {phi[2,j+N[2]]<-0}
for (j in 1:N[3]) {phi[3,j]<-phi3[1,j]}
for (j in 1:30) {phi[3,j+N[3]]<-0}
for (j in 1:N[4]) {phi[4,j]<-phi4[1,j]}
for (j in 1:50) {phi[4,j+N[4]]<-0}
for (i in 1:4){
  for (j in 1:N[i]) {
    theta[j,i]~dnorm(0.0,tau.y)
    O[j,i]~dbern(p[j,i])
    logit(p[j,i])<-alpha + (beta*vaux[j,i]) + theta[j,i] + phi[i,j]
  } # for de j
# radio de variacion espacial vs. variacion global
sd.t[i]<-sd(theta[1:N[i],i])
sd.phi[i]<-sd(phi[i,1:N[i]])
prop.v[i]<- sd.phi[i]/(sd.t[i]+sd.phi[i])
# prop.v = Representa la proporción de la variabilidad
# total que es atribuible a la componente espacial
} # for i

# Asignaciones de previas
tau.phi~dgamma(0.5,0.005)
tau.phi1~dgamma(1,0.01)
tau.phi3~dgamma(1,0.01)
tau.y~dgamma(0.5,0.005)

alpha ~ dnorm(0.0,0.001)
beta ~ dnorm(0.0,0.001)
} # end model

#Valores iniciales
list(alpha = 0, beta = 0) # Cadena de color ROJO
list(alpha = 0.5, beta = -0.5)# Cadena de color AZUL

```

Tabla 5.3: Sintaxis en OpenBUGS para el modelo MJBED.1 usando previas gammas.

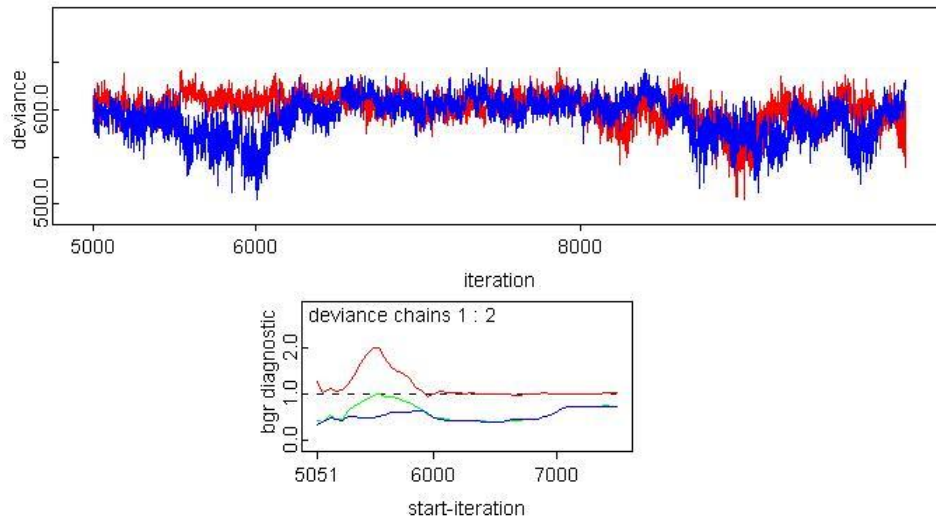


Figura 5.8: Convergencia para la deviance en el Modelo MJBED.1 usando previas uniformes (para 10000 iteraciones).

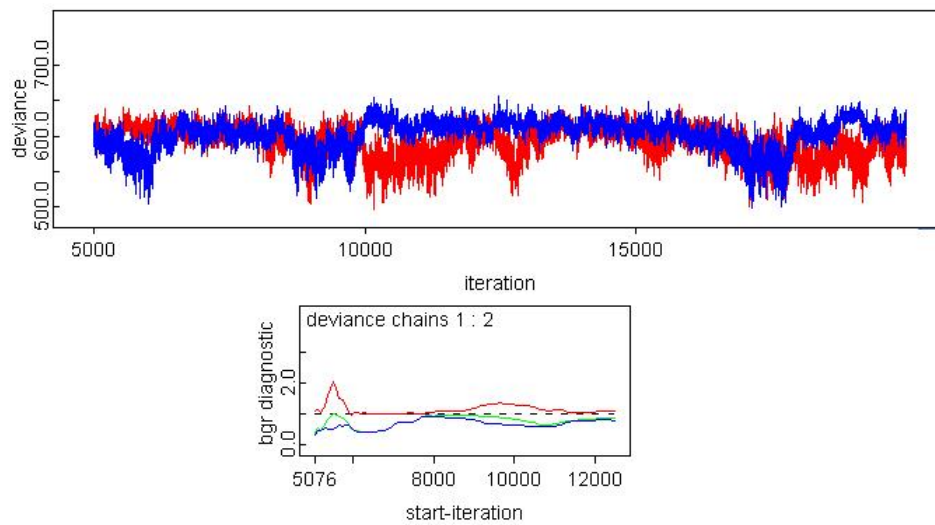


Figura 5.9: Convergencia para la deviance en el Modelo MJBED.1 usando previas uniformes (para 20000 iteraciones).

En la figura 5.9 se observa que las cadenas alcanzan la convergencia después de las 12000 iteraciones aproximadamente. Al comparar esta convergencia (ver figuras 5.8 y 5.9) con la convergencia mostrada en las figuras 5.6 y 5.7 se tiene que al usar previas uniformes, se logra una rápida y clara convergencia en la deviance del modelo dinámico formulado (figura ??).

```

model
{# Modelo jerárquico bayesiano espacial dinámico se considera
 # en forma conjunta tanto los datos observados como el tiempo
for (t in 1:k[1]) {weights1[t]<-1}
for (t in 1:k[2]) {weights2[t]<-1}
for (t in 1:k[3]) { weights3[t]<-1}
for (t in 1:k[4]) {weights4[t]<-1}

# Distribuciones previas CAR (efectos aleatorios espaciales)
phi1[1,1:N[1]] ~ car.normal(adj1[], weights1[], num1[], tau.phi)
phi2[1,1:N[2]] ~ car.normal(adj2[], weights2[], num2[], tau.phi)
phi3[1,1:N[3]] ~ car.normal(adj3[], weights3[], num3[], tau.phi)
phi4[1,1:N[4]] ~ car.normal(adj4[], weights4[], num4[], tau.phi4)

for (j in 1:N[1]) {phi[1,j]<-phi1[1,j]}
phi1[1,300]<-0
for (j in 1:N[2]) {phi[2,j]<-phi2[1,j]}
for (j in 1:7) {phi[2,j+N[2]]<-0}
for (j in 1:N[3]) {phi[3,j]<-phi3[1,j]}
for (j in 1:30) {phi[3,j+N[3]]<-0}
for (j in 1:N[4]) {phi[4,j]<-phi4[1,j]}
for (j in 1:50) {phi[4,j+N[4]]<-0}
for (i in 1:4){
  for (j in 1:N[i]) {
    theta[j,i]~dnorm(0.0,tau.y)
    O[j,i]~dbern(p[j,i])
    logit(p[j,i])<-alpha + (beta*vaux[j,i]) + theta[j,i] + phi[i,j]
  } # for de j
# radio de variacion espacial vs. variacion global
sd.t[i]<-sd(theta[1:N[i],i])
sd.phi[i]<-sd(phi[i,1:N[i]])
prop.v[i]<- sd.phi[i]/(sd.t[i]+sd.phi[i])
} # for i
# Asignaciones de previas
tau.phi<-1/(sigma.phi*sigma.phi)
tau.y <-1/(sigma.t*sigma.t)

alpha ~ dnorm(0.0,0.001)
beta ~ dnorm(0.0,0.001)
tau.phi4 ~ dgamma(1,0.01)

#Definicion de sigmas
sigma.phi~dunif(0,1)
sigma.t~dunif(0,1)
} # end model
#Valores iniciales
list(alpha = 0, beta = 0) # Cadena de color ROJO
list(alpha = 0.5, beta = -0.5)# Cadena de color AZUL

```

Tabla 5.4: Sintaxis en OpenBUGS para el modelo MJBED.1 usando previas uniformes.

Apéndice 2

En este apéndice se presentan las gráficas de convergencia, distribuciones de densidad estimadas y el test de diagnóstico de Gelman y Rubin (1992) para los parámetros de los modelos que no se presentaron en el capítulo 4 que resultaron con los menores DIC. Se verifica que en todos los modelos jerárquicos bayesianos espaciales (MJBE.1) formulados por año, se alcanza la convergencia en los parámetros estimados. En adelante las previas para las desviaciones estándar usadas en todos los modelos MJBE y MJBED serán distribuciones uniformes (0,1).

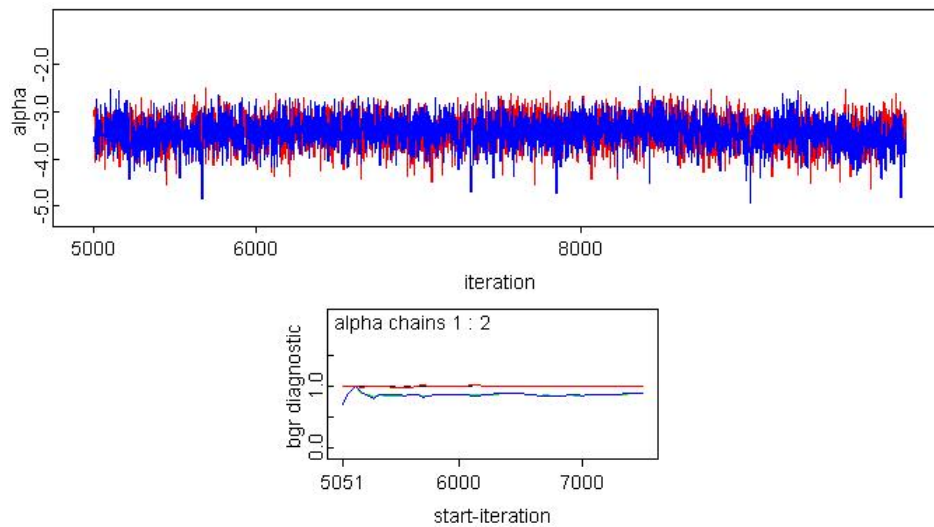


Figura 5.10: Convergencia para α en el modelo MJBE.1, correspondiente al año 95 (10000 iteraciones).

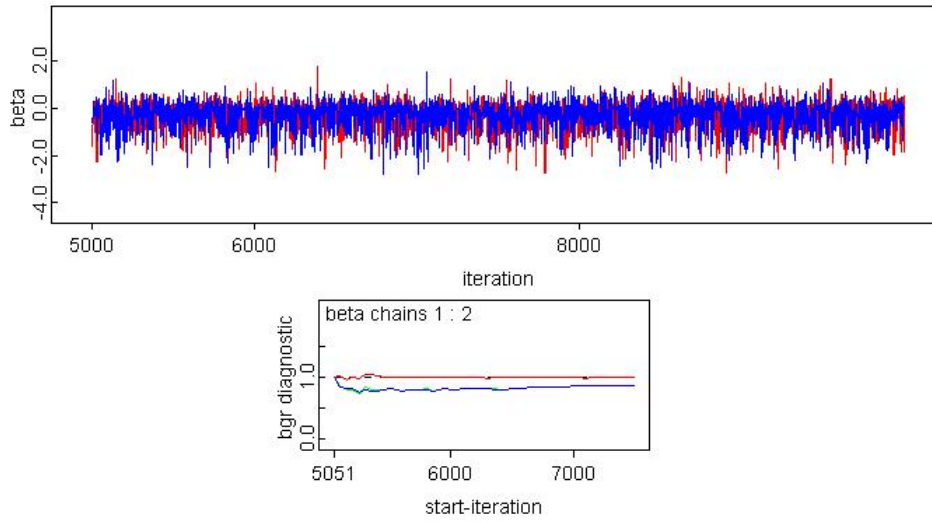


Figura 5.11: Convergencia para β en el modelo MJBE.1, correspondiente al año 95 (10000 iteraciones).

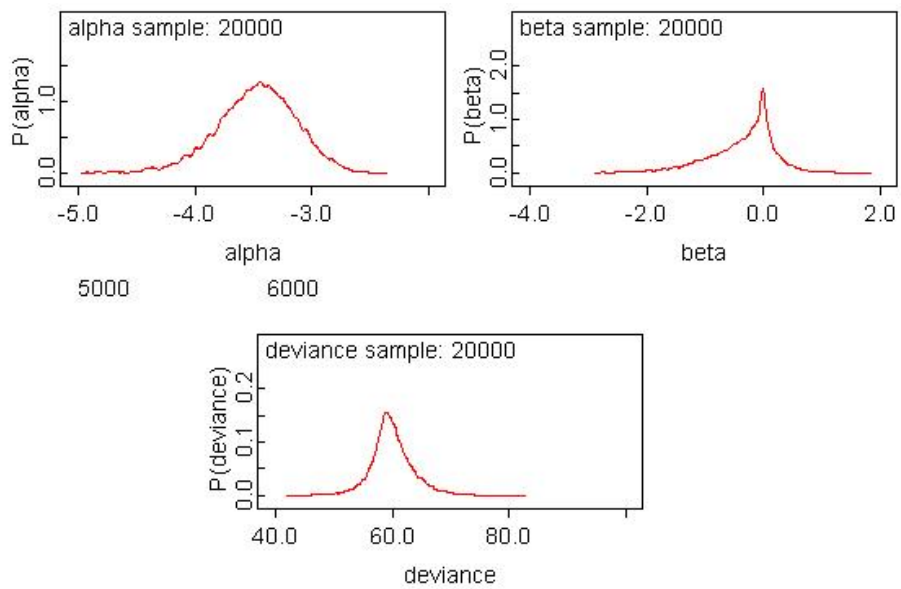


Figura 5.12: Densidades estimadas para α , β y deviance del modelo MJBE.1, correspondiente al año 95.

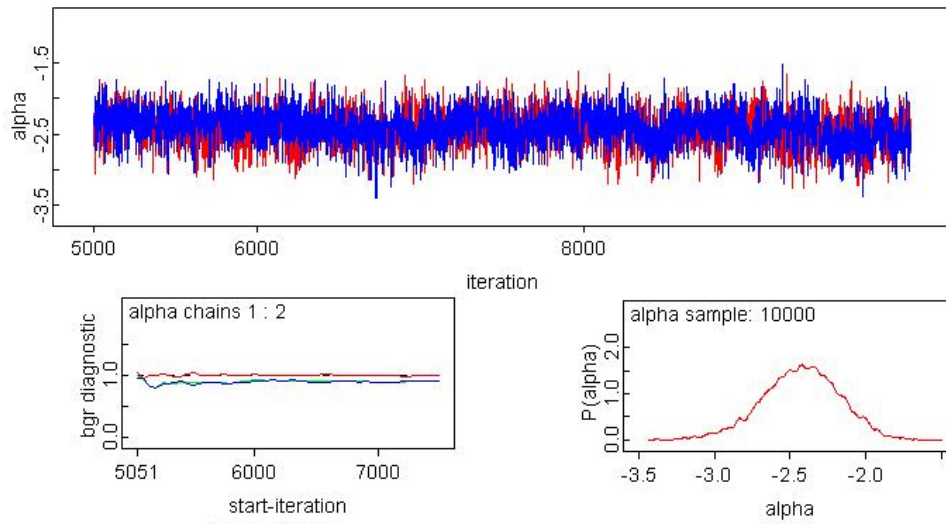


Figura 5.13: Convergencia, test Gelman y Rubin y densidad estimada para α del modelo MJBE.1, correspondiente al año 96.

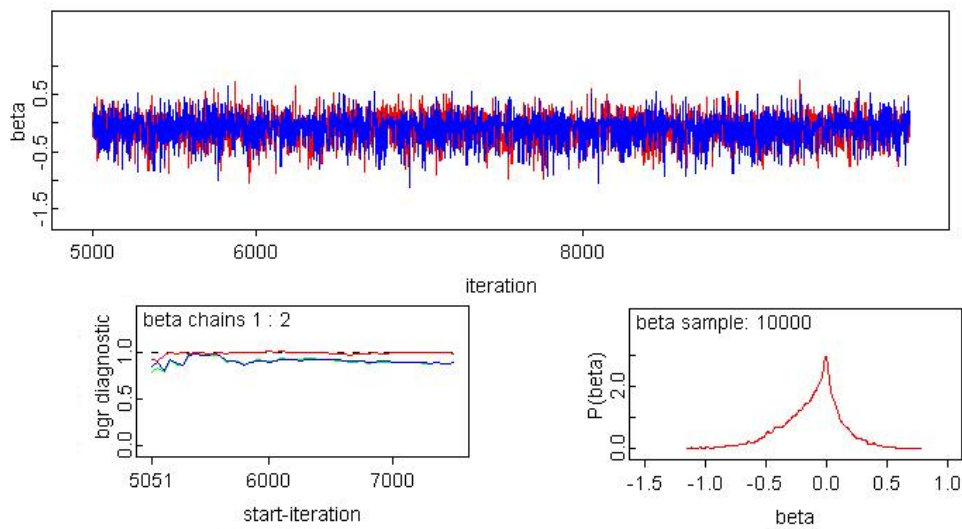


Figura 5.14: Convergencia, test Gelman y Rubin y densidad estimada para β del modelo MJBE.1, correspondiente al año 96.

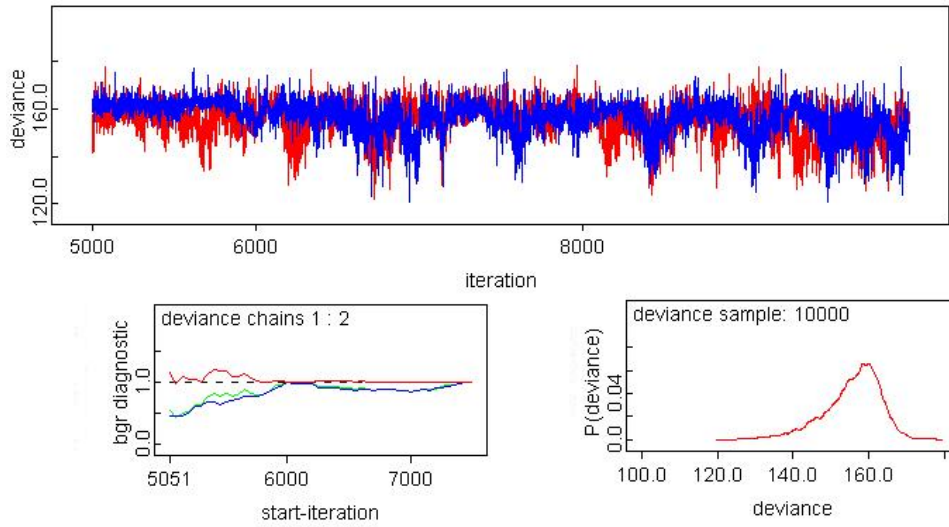


Figura 5.15: Convergencia, test Gelman y Rubin y densidad estimada para deviance del modelo MJBE.1, correspondiente al año 96.

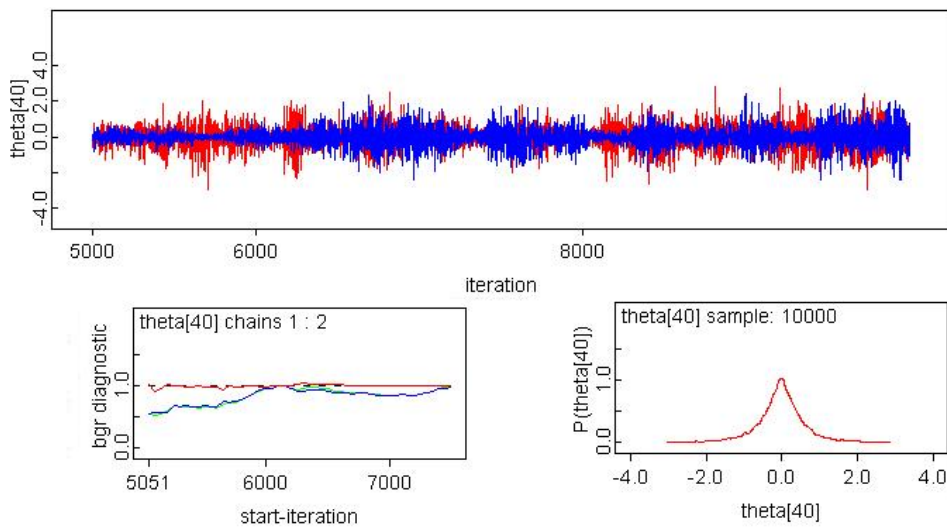


Figura 5.16: Convergencia, test Gelman y Rubin y densidad estimada para $\theta[\text{árbol}=40]$ del modelo MJBE.1, correspondiente al año 96.

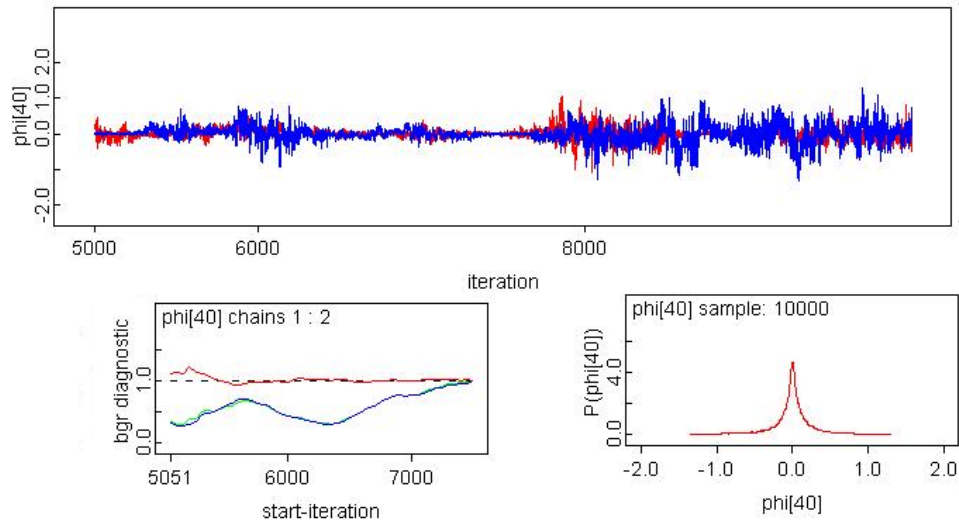


Figura 5.17: Convergencia, test Gelman y Rubin y densidad estimada para $\phi[\text{árbol}=40]$ del modelo MJBE.1, correspondiente al año 96.

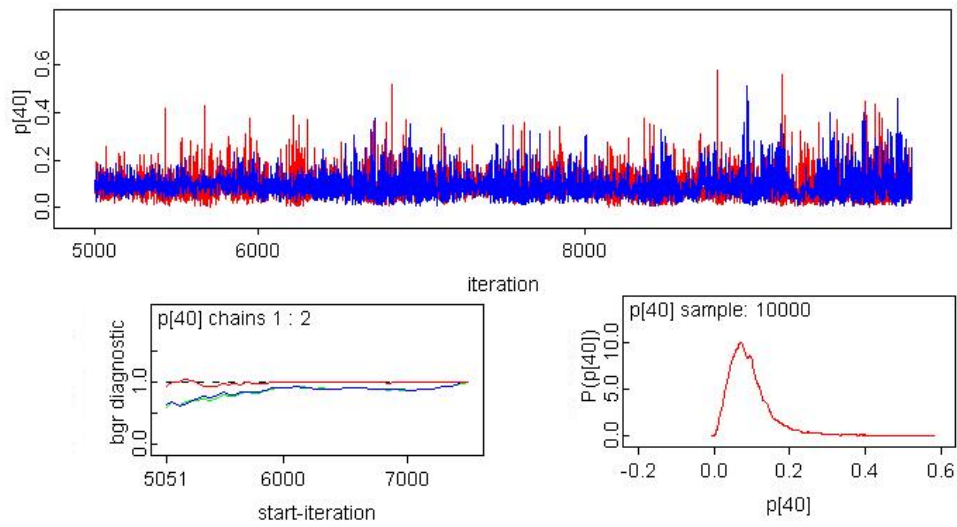


Figura 5.18: Convergencia, test Gelman y Rubin y densidad estimada para $p[\text{árbol}=40]$ en el modelo MJBE.1, correspondiente al año 96.

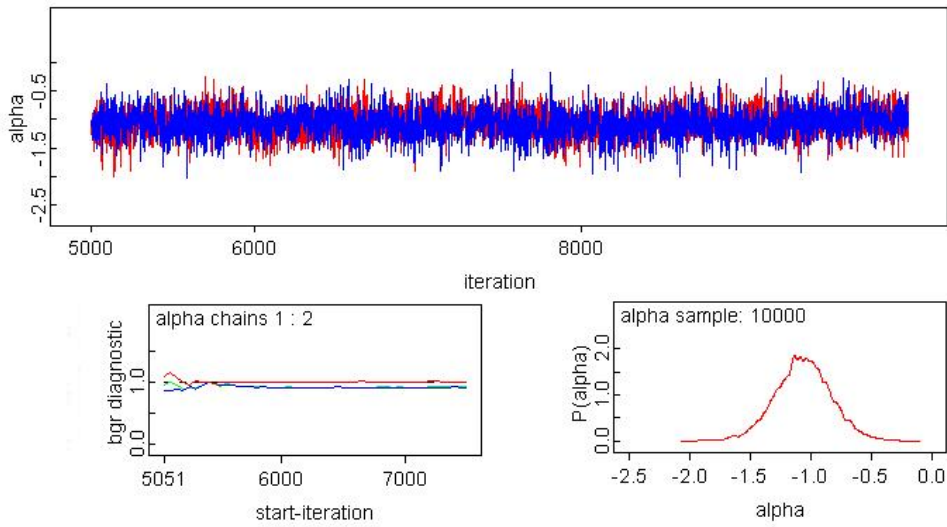


Figura 5.19: Convergencia, test Gelman y Rubin y densidad estimada para α en el modelo MJBE.1, correspondiente al año 98.

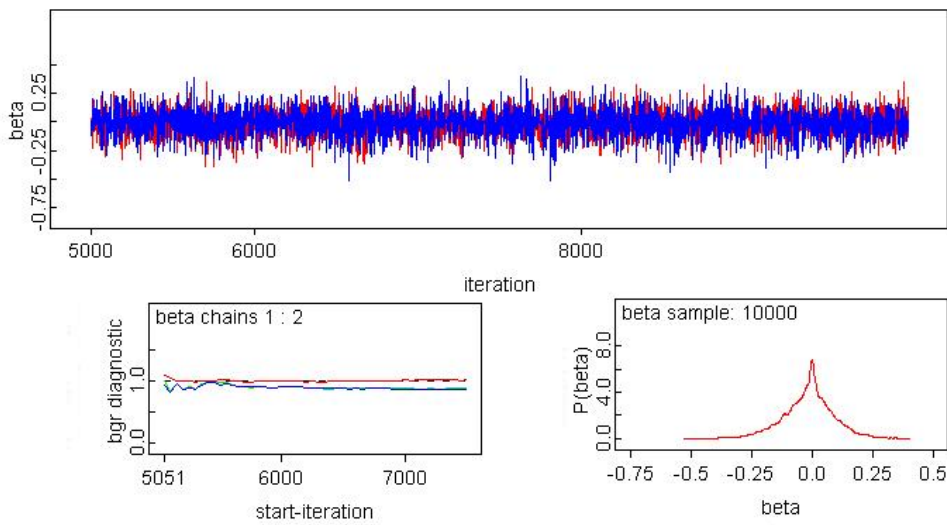


Figura 5.20: Convergencia, test Gelman y Rubin y densidad estimada para β en el Modelo MJBE.1, correspondiente al año 98.

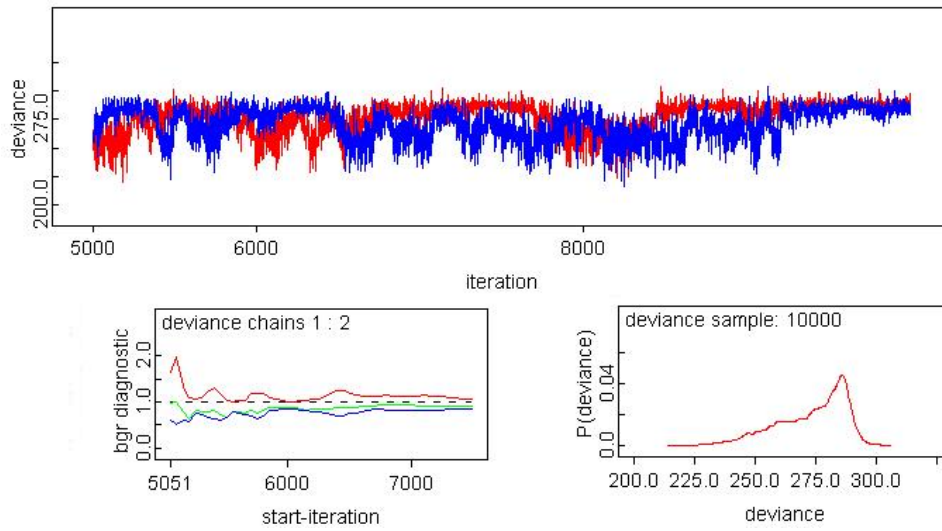


Figura 5.21: Convergencia, test Gelman y Rubin y densidad estimada para deviance en el modelo MJBE.1, correspondiente al año 98.

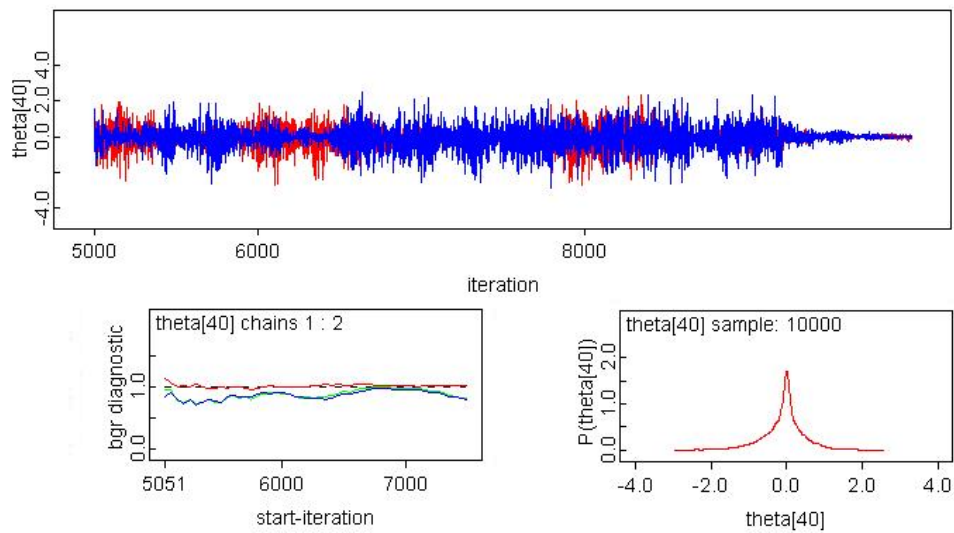


Figura 5.22: Convergencia, test Gelman y Rubin y densidad estimada para $\theta[\text{árbol}=40]$ en el modelo MJBE.1, correspondiente al año 98.

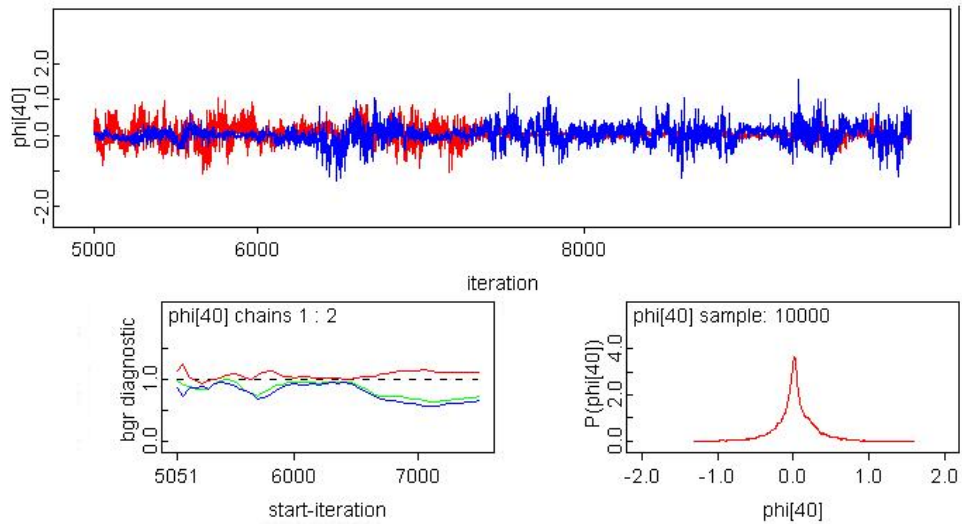


Figura 5.23: Convergencia, test Gelman y Rubin y densidad estimada para $\phi[\text{árbol}=40]$ en el modelo MJBE.1, correspondiente al año 98.

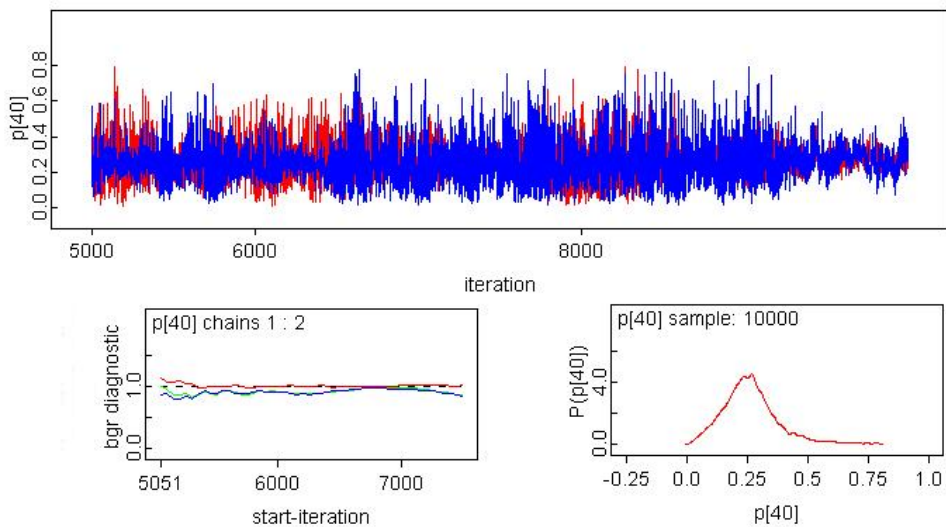


Figura 5.24: Convergencia, test Gelman y Rubin y densidad estimada para $p[\text{árbol}=40]$ en el modelo MJBE.1, correspondiente al año 98.

Apéndice 3

Se presentan las gráficas de convergencia, distribuciones de densidad estimadas y el test de diagnóstico de Gelman y Rubin (1992) para los parámetros del modelo jerárquico bayesiano espacial MJBE.3 correspondiente al año 97 que obtuvo el segundo mejor ajuste y no se presentó en el capítulo 4. DIC para este modelo, $\text{MJBE.3} = 143.0$ y $p_D = 0.34$.

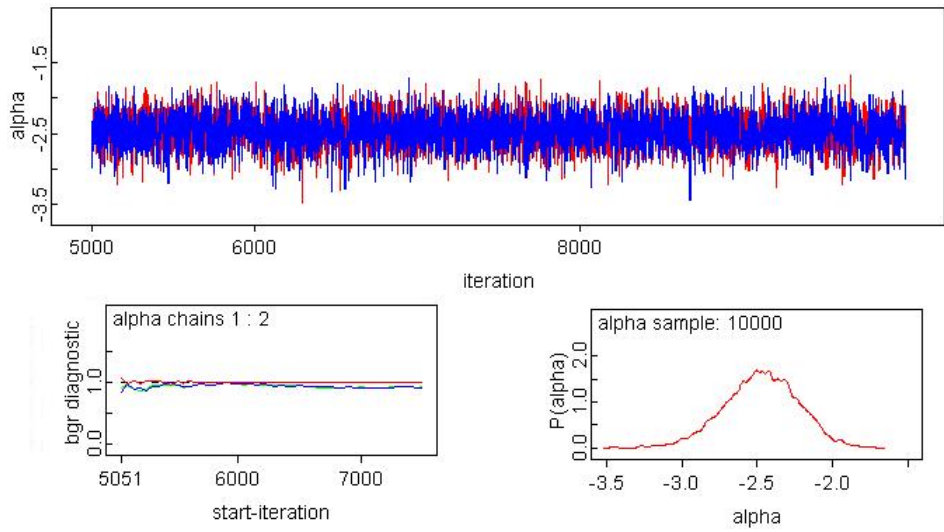


Figura 5.25: Convergencia, test Gelman y Rubin y densidad estimada para α en el modelo MJBE.3, correspondiente al año 97.

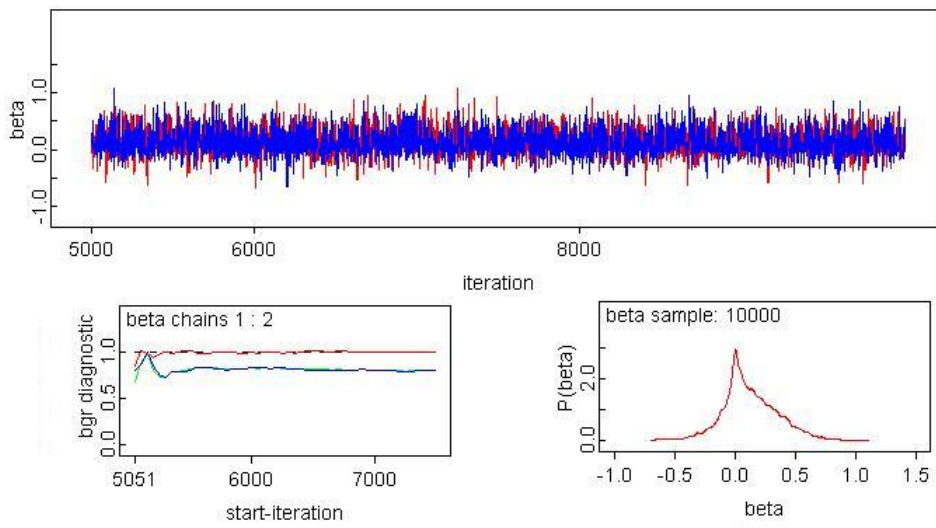


Figura 5.26: Convergencia, test Gelman y Rubin y densidad estimada para β en el modelo MJBE.3, correspondiente al año 97.

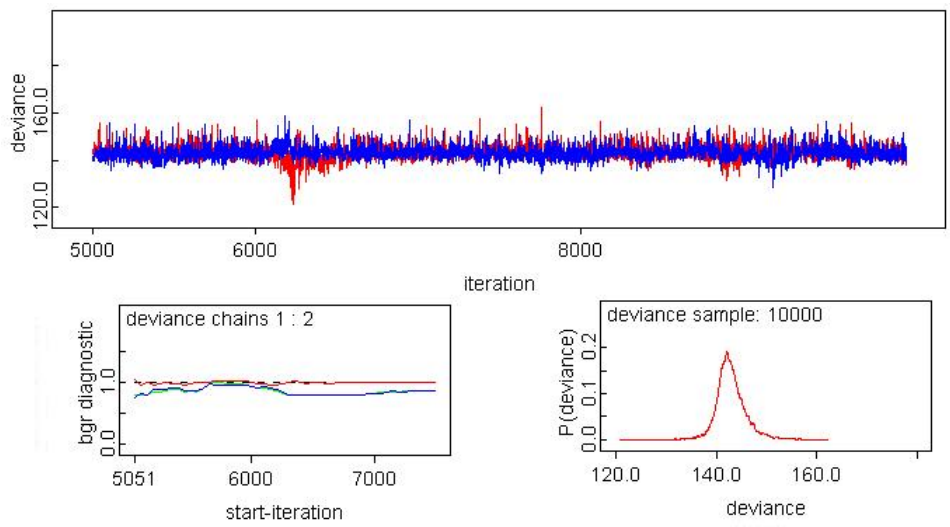


Figura 5.27: Convergencia, test Gelman y Rubin y densidad estimada para deviance en el modelo MJBE.3, correspondiente al año 97.

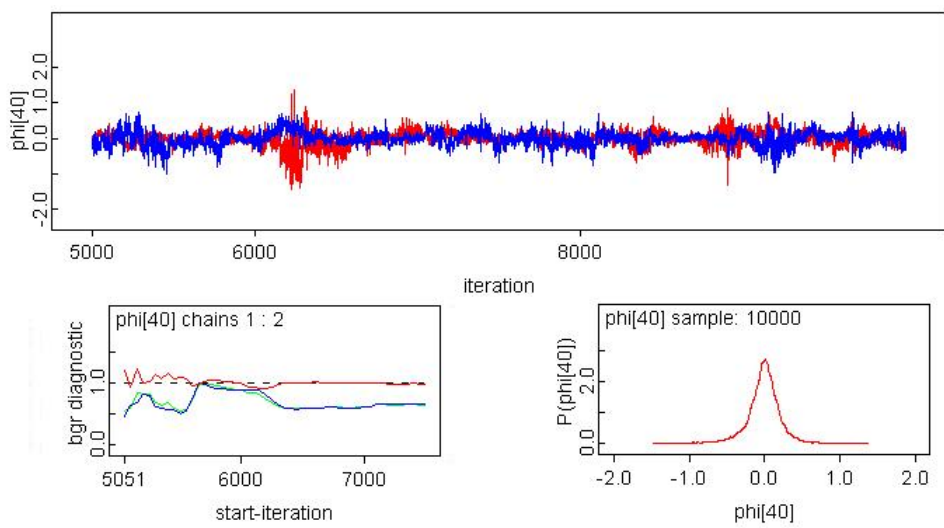


Figura 5.28: Convergencia, test Gelman y Rubin y densidad estimada para $\phi[\text{árbol}=40]$ en el modelo MJBE.3, correspondiente al año 97.

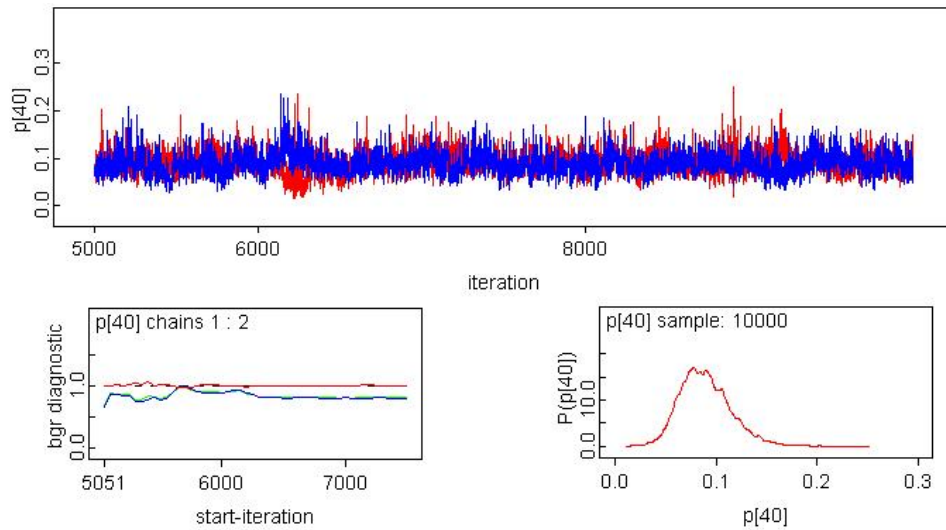


Figura 5.29: Convergencia, test Gelman y Rubin y densidad estimada para $p[\text{árbol}=40]$ en el modelo MJBE.3, correspondiente al año 97.

Apéndice 4

En este apartado se presentan las gráficas de convergencia, distribuciones de densidad estimadas y el test de diagnóstico de Gelman y Rubin (1992) para los parámetros del modelo jerárquico bayesiano espacial dinámico que obtuvo el segundo mejor ajuste (MJBED.3) cuyo DIC fue igual a 625.5 y $p_D = 11.18$.

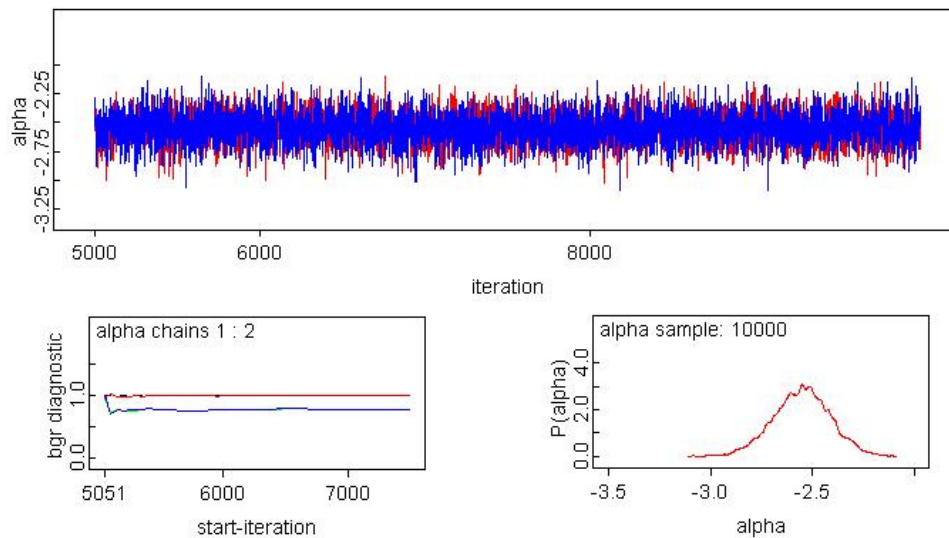


Figura 5.30: Convergencia, test Gelman y Rubin y densidad estimada para α en el modelo MJBED.3.

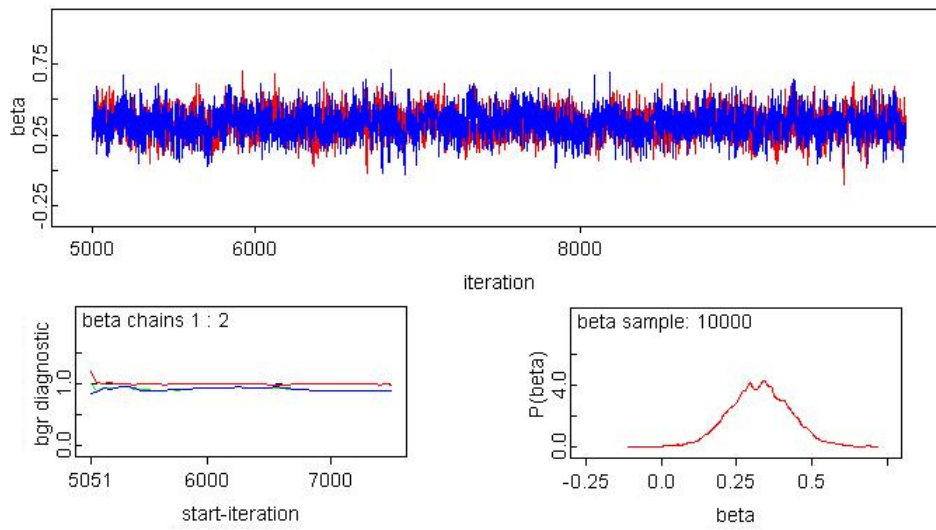


Figura 5.31: Convergencia, test Gelman y Rubin y densidad estimada para β en el modelo MJBED.3.

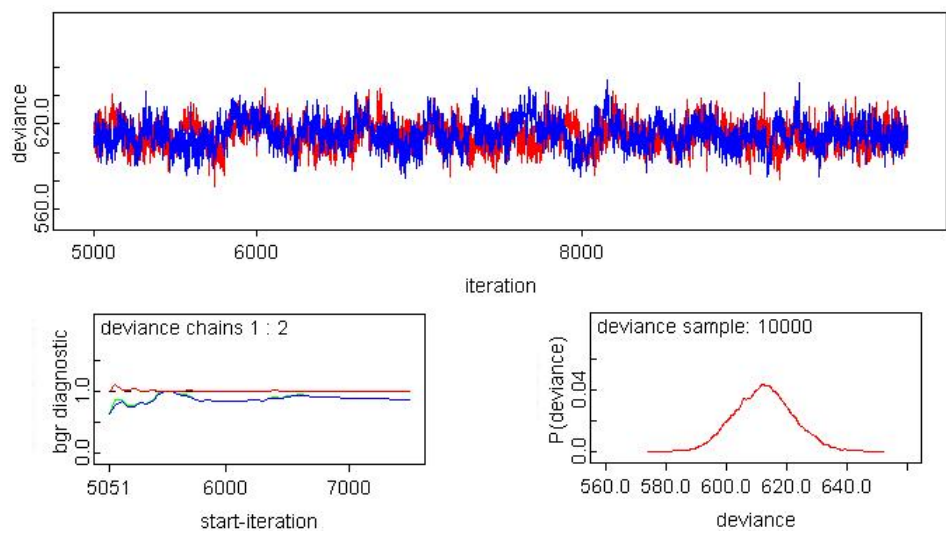


Figura 5.32: Convergencia, test Gelman y Rubin y densidad estimada para deviance en el modelo MJBED.3.

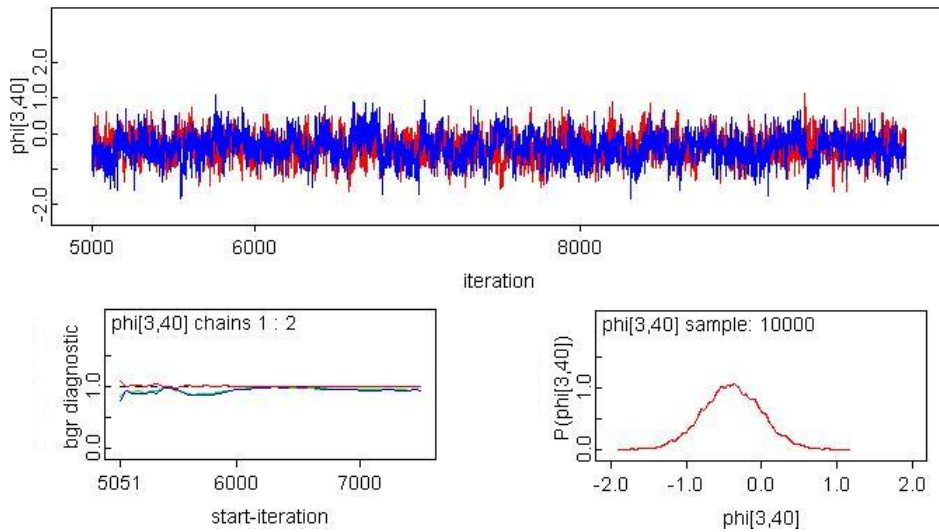


Figura 5.33: Convergencia, test Gelman y Rubin y densidad estimada para ϕ [año=3, árbol=40] en el modelo MJBED.3

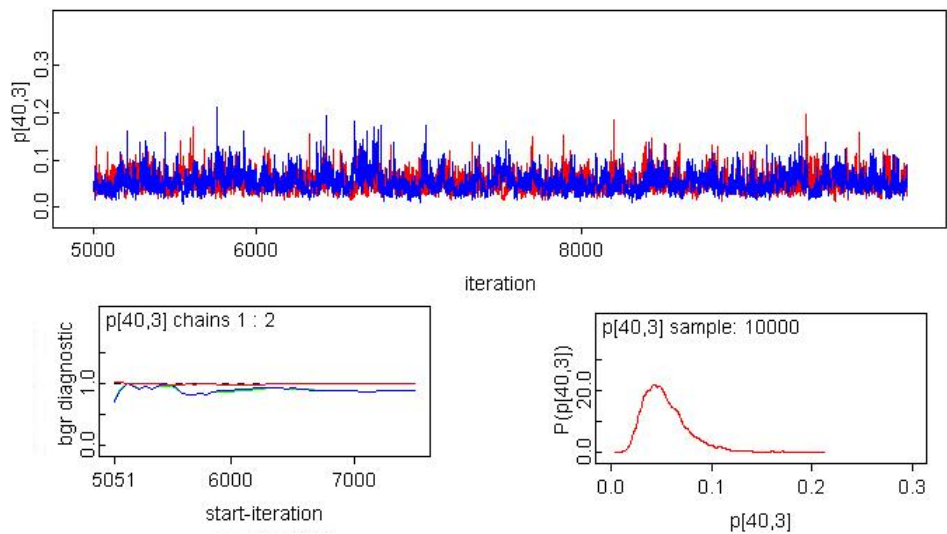


Figura 5.34: Convergencia, test Gelman y Rubin y densidad estimada para p [árbol=40, año=3] en el modelo MJBED.3

Apéndice 5

Se presentan las funciones escritas en el software R, con las cuales se obtuvieron las gráficas presentadas en el capítulo 1. Además se muestra el código escrito en R para las corridas de algunos de los modelos formulados usando la función bugs y openBUGS desde R. Así como la rutina con la cual se definieron las estructuras de vecindad incorporadas en la distribución CAR, así como el código escrito para otros analices realizados.

- Código fuente para obtener las salidas presentadas en el capítulo 1, las regresiones logísticas realizadas a diferentes distancias y las pruebas de hipótesis consideradas.

```
# Análisis descriptivo y gráfico de la evolución
# del CTV la parcela estudiada
# Creacion del conjunto de datos para la PARCELA: BURRIANA
bu<-matrix(scan("bu.txt"),300,5,byrow=T)

#Archivo con las localizaciones muestreadas en Burriana
buxy<-matrix(scan("burrianaN.txt"),300,2,byrow=T)

bu<-cbind(bu,buxy) # Datos de BURRIANA JUNTO con las localizaciones
anyobu<-c(1994,1995,1996,1997,1998)

#Calcula la proporcion de infectados en la parcela BURRIANA
apply(bu[,1:5],2,sum)
pinfbu<-apply(bu[,1:5],2,mean)

# Graficos con la configuracion del CTV en BURRIANA
burri<-function(i)
{
  ll<-(bu[,i]==1)
  plot(bu[,7],-bu[,6],type="n",xlab="",ylab="",xaxt="n",yaxt="n")
  points(bu[,7][!ll],-bu[,6][!ll],pch=".")
  points(bu[,7][ll],-bu[,6][ll],pch=16)
  title(paste(as.character(anyobu[i]),"
(",as.character(round(pinfbu[i]*100)),"\%)" ,sep=""))
}
par(mfrow=c(2,3),mar=c(1,1,2,1))
for (i in 1:length(anyobu)) {burri(i)}

# Grafica de incidencia en BURRIANA
plot(anyobu,pinfbu,ylim=c(0,1),xlab="Año",ylab="Incidencia",
pch=16,type="o")

# Funciones para el cálculo de las distancias entre los árboles
# de la parcela BURRIANA
dist2full <- function(dis)
{
  n <- attr(dis, "Size")
  full <- matrix(0, n, n)
  full[lower.tri(full)] <- dis
}
```



```

    full + t(full)
  }

  distabu<-dist2full(dist(bu[,6:7]))

#Calculo del Nro. de vecinos infectados a menos de una
#determinada distancia x en un tiempo determinado año para BURRIANA
#En la ultima columna se almacena el numero de vecinos a esa distancia.

nvecinbu<-function(x){
  cbind((distabu<x)\%*\%bu[,1:5]-bu[,1:5],apply((distabu<x),2,sum)-1)
}

### GLM's para BURRIANA
# GLM con el efecto del Nro. de vecinos infectados a
# una determinada distancia, considerando los árboles
# enfermos del año anterior
glm.bu1<-function(x){
  glm.bu<-matrix(0,11,4)
  nvec<-nvecinbu(x)
  for(i in 1:4)
  {
    vaux<-nvec[,i][bu[,i]==0]
    m<-glm(bu[,i+1][bu[,i]==0]~vaux,family=binomial)
    glm.bu[,i]<-c(coef(summary(m)),deviance(m),
    summary(m)$null.deviance,m$df.residual)
  }
  glm.bu
}

#calculo de los p-valores, significancia de la covariable = vaux
c.anyobu<-seq(6,55,2)
p.valor.glm.bu1<-sapply(c.anyobu,function(x){glm.bu1(x)[8,]})
colnames(p.valor.glm.bu1)<-c(rep("p.valor",length(c.anyobu)))
rownames(p.valor.glm.bu1) <- c("95|94","96|95","97|96","98|97")
# Tomar solo coeficientes estimados (intercepto y theta)
alpha<-sapply(15,function(x){glm.bu1(x)[1,]})
thetas<-sapply(15,function(x){glm.bu1(x)[2,]})
vparam<-c(apply(alpha,2,mean),apply(thetas,2,mean))

# GLM con el efecto de la proporcion de
# árboles enfermos a una determinada distancia,

```

```

# considerando los enfermos del año anterior
glm.bu2<-function(x){
  glm.bu<-matrix(0,11,4)
  nvec<-nvecinbu(x)
  for(i in 1:4)
    {vaux<-(nvec[,i][bu[,i]==0]/nvec[,6][bu[,i]==0])*100
  m<-glm(bu[,i+1][bu[,i]==0]~vaux,family=binomial)
  glm.bu[,i]<-c(coef(summary(m)),deviance(m),
  summary(m)$null.deviance,m$df.residual)
  }
  glm.bu
}
c.anyobu<-seq(6,35,2)
#calculo de los p-valores, significancia de la covariable = vaux
glmbu2<-glm.bu2(c.anyobu)
p.valor.glm.bu2<-sapply(c.anyobu,function(x){glm.bu2(x)[8,]})
colnames(p.valor.glm.bu2)<-c(rep("p.valor",length(c.anyobu)))
rownames(p.valor.glm.bu2) <- c("95|94", "96|95", "97|96", "98|97")

#Listado con Nro. de vecinos TOTALES por AÑO considerando distancias
# entre x e y
nvec.6<-nvecinbu(6)
nvec.10<-nvecinbu(10)
nvec.12<-nvecinbu(12)
nvec.20<-nvecinbu(20)
nvec.30<-nvecinbu(30)
nvec.40<-nvecinbu(40)
nvec.6a12<-nvec.12-nvec.6
nvec.10a20<-nvec.20-nvec.10
nvec.20a30<-nvec.30-nvec.20
nvec.30a40<-nvec.40-nvec.30

# GLM con el efecto de la proporción de
# árboles enfermos a varias distancias, considerando
# los enfermos del año anterior
glm.bu2<-function(){
  glm.bu<-matrix(0,5,4)
  for(i in 1:4)
    {vaux.6<-(nvec.6[,i][bu[,i]==0]/nvec.6[,6][bu[,i]==0])*100
    vaux.6a12<-((nvec.12[,i][bu[,i]==0]-nvec.6[,i][bu[,i]==0])/
    (nvec.12[,6][bu[,i]==0]-nvec.6[,6][bu[,i]==0]))*100
    }
  glm.bu
}

```

```

    vaux.10 a20<-((nvec.20[,i][bu[,i]==0]-nvec.10[,i][bu[,i]==0])/
(nvec.20[,6][bu[,i]==0]-nvec.10[,6][bu[,i]==0]))*100

    vaux.20 a30<-((nvec.30[,i][bu[,i]==0]-nvec.20[,i][bu[,i]==0])/
(nvec.30[,6][bu[,i]==0]-nvec.20[,6][bu[,i]==0]))*100

    vaux.30 a40<-((nvec.40[,i][bu[,i]==0]-nvec.30[,i][bu[,i]==0])/
(nvec.40[,6][bu[,i]==0]-nvec.30[,6][bu[,i]==0]))*100

    m<-glm(bu[,i+1][bu[,i]==0]~(vaux.6+vaux.6 a12+vaux.10 a20+
    vaux.20 a30+ vaux.30 a40),family=binomial)
    glm.bu[,i]<-c(coef(summary(m))[20],coef(summary(m))[21],
    coef(summary(m))[22],coef(summary(m))[23],coef(summary(m))[24])
    }
    glm.bu
}

#calculo de los p-valores, significancia de la covariable = vaux
p.valor.glm.bu3<-glm.bu2()
colnames(p.valor.glm.bu3)<-c(rep("p.valor",4))
rownames(p.valor.glm.bu3) <- c("6","6-12","10-20","20-30","30-40")

# Prueba de Hipotesis para medir diferencias significativas entre
# el nro. de casos de enfermos de un año determinado y el anterior
bu.test<-NULL
testbu<-function(x){
  for (i in 2:5){
    nvec<-nvecinbu(x)
    vaux<-nvec[,i][bu[,i]==0]
    bu.test<-c(bu.test,t.test(x = vaux[bu[,i-1]==0],
    y= vaux[bu[,i]==0], groups.p = T, mu = 0,
    alternative ="two.sided", t.paired = "Two-sample t",
    var.equal= T, conf.level = 0.95, print.object.p = T)$p.value)
  }
  bu.test
}

c.anyobu<-seq(6,50,4)
ttest.bu<-sapply(c.anyobu,function(x){testbu(x)})

```

- Código fuente para el desarrollo del Método Scoring de Monte Carlo.

```
## (1) Metodo de Scoring por Monte Carlo
```

```

# Funcion que determina el Nro. de vecinos infectados
# a menos de una distancia determinada(d) en el tiempo t
# En la ultima columna se almacena el Nro. total de vecinos
# a la distancia determinada(x)
nvecinbu.1<-function(x,t,distabu.1,b){
  cbind((distabu.1<x)\%*\%b-b)}

nvecinbu.2<-function(x,distabu.2,bu.2){
  cbind((distabu.2<x)\%*\%bu.2[,1]-bu.2[,1])}

nvecinbu.3<-function(x,t){
  cbind((distabu<x)\%*\%bu[,t]-bu[,t])}

# Funcion para estimar los parametros del modelo logístico
# considerando el total de vecinos infectados con
# CTV por año sin CONDICIONAR a enfermos del
# año anterior.
# parametros a estimar son: param[1]=alpha y param[2]=theta

# Función para calcular pi=probabilidad de enfermar
binombu<-function(i,param,vin){
  npot<-param[1]+(param[2]*vin)
  pi<-exp(npot)/(1+exp(npot))
  sim<<-rbinom(1,1,pi)
}

estadbu<-function(sim,t){
  estad<-rep(0,2)
  bu.2<-cbind(sim,buxy)
  distabu.1<-dist2full(dist(bu.2[,2:3]))
  nv<-nvecinbu.1(d,t,distabu.1,bu.2[,1])
  nv<-cbind(nv,apply(nv,1,sum))
  vinaux<-nv[,1] # Nro. de vecinos infectados por año
                # t a una cierta distancia
  ninf<-nv[,2]
  estad[1]<-sum(sim)
  estad[2]<-sum(sim*(vinaux/ninf))
  estad
}

# Inicializacion de valores
vparam<-c(-3,-1)

```

```

d<-25 # Distancias A PROBAR
t<-1
nv<-nvecinbu.1(d,t,distabu,bu[,t])
nv<-cbind(nv,apply(nv,1,sum))
vaux<-nv[,1] #Nro. de vecinos infectados por año t a una cierta distancia
ninf<-nv[,2]

for (j in 1:100){sim<-sapply(1:300,binombu,vparam,vaux/ninf)}#Calentamiento

N<-10 # Nro. de simulaciones del Muestreo Gibbs
params<-NULL
ini<-list(c(-3,-1),c(-2.4,0.01),c(-2,0.1),c(-2,0.5),c(-3,0.5))
for (t in 1:5){
  vparam<-ini[[t]]
  vparams<-vparam
  obs<-bu[,t]
  estobs<-estadbu(obs,t)
  nv<-nvecinbu.1(d,t,distabu,bu[,t])
  nv<-cbind(nv,apply(nv,1,sum))
  vaux<-nv[,1]
  ninf<-nv[,2]
  for (r in 1:N){ #Ciclo general
    estt<-matrix(0,100,2)
    for (j in 1:100){
      sim<-sapply(1:300,binombu,vparam,vaux/ninf)
      estt[j,]<-estadbu(sim,t)
    } # fin del for del MUESTREO
    # Actualizacion de los parametros
    mediat<-apply(estt,2,mean)
    vart<-var(estt)
    inc<-solve(vart,estobs-mediat)
    vparam<-vparam+inc
    vparams<-rbind(vparams,vparam)
  }### fin for de r
  params[[t]]<-vparams
} #fin del for t

# Estudio de la convergencia de los parametros estimados
# por METODO(1) Scoring Monte Carlo considerando como covariable,
# el nro. de vecinos infectados a una cierta distancia sin
# considerar a los arboles enfermos con CTV del año anterior
library(coda)

```

```

sol<-params[[1]] #parametros solo para t=1
colnames(sol)<-c("alpha","theta") # parametros estimados por
                                # Muestreo Gibbs

rownames(sol)<-NULL
sal<-mcmc(sol)
summary(sal)

# Para estudiar la cadena del Metodo (1)
plot(sal)
crosscorr(sal)
autocorr(sal)
autocorr.plot(sal)
geweke.plot(sal)
geweke.diag(sal)
heidel.diag(sal)
raftery.diag(sal)
# Fin del Analisis Output

#####
# (2) Scoring de Monte Carlo considerando
# como covariable, el Nro. de vecinos
# infectados a una cierta distancia(d)
# CONDICIONADOS a los arboles infectados
# del año anterior

estbu.2<-function(sim,t){
  estad<-rep(0,2)
  buxy.2<-bu[bu[,t-1]==0,6:7]
  bu.2<-cbind(sim,buxy.2)
  distabu.2<-dist2full(dist(bu.2[,2:3]))
  nv<-nvecinbu.2(d,t,distabu.2,bu.2) #Nro. de vecinos infectados por
                                #año t a una cierta
                                #distancia considerando SOLO las
                                #localizaciones de los arboles
                                #enfermos del año anterior

  nv<-cbind(nv,apply(nv,1,sum))
  vinaux<-nv[,1]
  ninf<-nv[,2]
  estad[1]<-sum(sim)
  estad[2]<-sum(sim*(vinaux/ninf))
  estad
}

```

```

# Valores iniciales
vparam<-c(-2.4,0.1)
d<-15 # Distancias A PROBAR
t<-2
nv<-nvecinbu.3(d,t)
nv<-cbind(nv,apply(nv,1,sum))
vaux<-nv[,1][bu[,t-1]==0] #Nro. de vecinos infectados por año
                                #t a una cierta distancia
                                #considerando los arboles enfermos
                                #del año anterior
ninf<-nv[,2][bu[,t-1]==0]

n<-length(vaux)

for(j in 1:100){sim<-sapply(1:n,binombu,vparam,vaux/ninf)}#Calentamiento

N<-10# Nro. de simulaciones del Muestreo Gibbs
params<-NULL
ini<-list(c(-2.4,0.1),c(-2,0.1),c(-2.1,0.1),c(-1.1,0))
for(t in 2:5){
  vparam<-ini[[t-1]]
  vparams<-vparam
  obs<-bu[,t][bu[,t-1]==0]
  estobs<-estbu.2(obs,t)
  nv<-nvecinbu.3(d,t)
  nv<-cbind(nv,apply(nv,1,sum))
  vaux<-nv[,1][bu[,t-1]==0]
  ninf<-nv[,2][bu[,t-1]==0]
  n<-length(vaux)
  for(r in 1:N){#Ciclo para controlar CONVERGENCIA
    estt<-matrix(0,100,2)
    for(j in 1:100){
      sim<-sapply(1:n,binombu,vparam,vaux/ninf)
      estt[j,]<-estbu.2(sim,t)
    } # fin del for del MUESTREO
  # Actualizacion de los parametros
  mediat<-apply(estt,2,mean)
  vart<-var(estt)
  inc<-solve(vart,estobs-mediat)
  vparam<-vparam+inc
  vparams<-rbind(vparams,vparam)

```

```

    } # fin del for r
  params[[t]]<-vparams
} #fin del for t

```

```

library(coda)
sol<-params[[2]]
colnames(sol)<-c("alpha","theta") # parametros estimados
                                     # por Muestreo Gibbs

rownames(sol)<-NULL
sal<-mcmc(sol)
summary(sal)

```

- Funciones con las cuales se obtuvieron algunos de los resultados presentados en el CAPÍTULO 4.

```

# Creacion del conjunto de datos para la PARCELA: BURRIANA
bu<-matrix(scan("bu.txt"),300,5,byrow=T)

```

```

#Archivo con las localizaciones muestreadas en Burriana
buxy<-matrix(scan("burrianaN.txt"),300,2,byrow=T)

```

```

bu<-cbind(bu,buxy) # Datos de BURRIANA JUNTO con las localizaciones

```

```

# Funciones para el cálculo de las distancias entre los árboles
# de la parcela BURRIANA

```

```

dist2full <- function(dis)
{
  n <- attr(dis, "Size")
  full <- matrix(0, n, n)
  full[lower.tri(full)] <- dis
  full + t(full)
}

```

```

dista<-dist2full(dist(bu[,6:7])) # calculo de todas las distancias

```

```

#####
#####
## Construcción de las covariables con los arboles
## sanos y enfermos del año t ubicados a cierta
## distancia x, CONDICIONADO solo a los sanos del
## año anterior (t-1) [es decir, se considera el efecto
## de los arboles enfermos del año anterior (t-1)]
nvecinbu.2<-function(x,distabu.2,bu.2){
  cbind((distabu.2<x)\%*\%bu.2[,1]-bu.2[,1])}

```



```

### distancias a PROBAR  x=10, x=15, x=20, x=25, x=30, x=35
for (t in 2:5){
  buxy.t<-bu[bu[,t-1]==0,6:7]
  bu.t<-cbind(bu[,t][bu[,t-1]==0],buxy.t)
  dista.t<-dist2full(dist(bu.t[,2:3]))
  nvinf<-nvecinbu.2(x,dista.t,bu.t)
  vaux.t<-nvinf[,1]
  if (t==2) {dput(file="vauxt2.txt", vaux.t)
             O.t2<-bu[,t][bu[,t-1]==0]
             vauxt2<-vaux.t}
  else
    if (t==3) {dput(file="vauxt3.txt", vaux.t)
               O.t3<-bu[,t][bu[,t-1]==0]
               vauxt3<-vaux.t}
  else
    if (t==4) {dput(file="vauxt4.txt", vaux.t)
               O.t4<-bu[,t][bu[,t-1]==0]
               vauxt4<-vaux.t}
  else
    if (t==5) {dput(file="vauxt5.txt", vaux.t)
               O.t5<-bu[,t][bu[,t-1]==0]
               vauxt5<-vaux.t}
}

### Construcción de las covariables con los nuevos casos de
#arboles enfermos para cada año t, quitando de la covariable
## el efecto de los árboles enfermos del año anterior(t-1)
## COVARIABLE con sólo NUEVOS casos en el año t
dput(file="vaux.actualt2", varsbuv[,2][bu[,1]==0])
dput(file="vaux.actualt3", varsbuv[,3][bu[,2]==0])
dput(file="vaux.actualt4", varsbuv[,4][bu[,3]==0])
dput(file="vaux.actualt5", varsbuv[,5][bu[,4]==0])

## Escritura de la Covariable con Arboles enfermos observados
#considerando el efecto del año anterior (compuesta SOLO por 0 y 1)
dput(file="obst2.txt", bu[,2][bu[,1]==0])
dput(file="obst3.txt", bu[,3][bu[,2]==0])
dput(file="obst4.txt", bu[,4][bu[,3]==0])
dput(file="obst5.txt", bu[,5][bu[,4]==0])

#####

```

```

write.table(as.matrix(varsbuy),"varsbuy.txt",row.names=T,
col.name=T,sep="\t")
write.table(as.matrix(varsbu),"varsbu.txt",row.names=T,
col.name=T,sep="\t")
# Tabla con los arboles enfermos del año actual t, quitando los
# arboles enfermos del año anterior [es decir, considerando el
# efecto de los arboles enfermos del año anterior t-1] y tabla
#con el efecto de los árboles enfermos de CTV del año anterior.

#####
## Construcción de las covariables con SOLO los arboles
## enfermos del año actual (t) ubicados a cierta
# distancia x, CONDICIONADO solo a los sanos del
## año anterior (t-1) [es decir, se considera el efecto de los
## arboles enfermos del año anterior (t-1)]

for (t in 2:5){
  buxy.t<-bu[bu[,t-1]==0,6:7]
  bu.t<-cbind(varsbuy[,t][bu[,t-1]==0]==0,buxy.t)
  dista.t<-dist2full(dist(bu.t[,2:3]))
  nvinf<-nvecinbu.2(x,dista.t,bu.t)
  vaux.t<-nvinf[,1]
  if (t==2) dput(file="vauxtact2.txt", vaux.t)
  if (t==3) dput(file="vauxtact3.txt", vaux.t)
  if (t==4) dput(file="vauxtact4.txt", vaux.t)
  if (t==5) dput(file="vauxtact5.txt", vaux.t)
}

#####
### Lanzamos los modelos por AÑO #####
### Modelo Uno: Para t=2 dado t=1 [es decir, año 95 dado 94]
narb<-length(O.t2)

#Leemos todos los datos para Modelo: M1t2 dado t1
datos<-list("narb","O.t2","vauxt2")
data<-c(datos)

# valores iniciales para Bugs
dput(file='init2.1.txt', list(alpha=0,beta=0))
dput(file='init2.2.txt', list(alpha=0.5,beta=-0.5))
init2.1<-source("init2.1.txt")
init2.2<-source("init2.2.txt")

```

```

# Valores iniciales definitivos
init1<-init2.1\$value
init2<-init2.2\$value
inits<-list (init1 , init2)

dput( file="datat2.txt" , list (narb ,O.t2 , vauxt2))

#parametros de interés
param<-c(" alpha" ," beta")

#Lanzamos el modelo: M1t2 dado t1
m1t2<-openbugs(data , inits , param,"M1t2 dado t1.txt" ,
n.chains=2,n.iter=10000)

## Estudio de la Cadenas con CODA
sal<-mcmc(m1t2\$sims.matrix[ ,1:2])
## Pruebas de CONVERGENCIA A LAS CADENAS
plot(sal)
crosscorr(sal)
autocorr(sal)
autocorr.plot(sal)
geweke.plot(sal)
geweke.diag(sal)
heidel.diag(sal)
raftery.diag(sal)
### Modelo Dos: Para t=3 dado t=2 [es decir , año 96 dado 95]
narb<-length(O.t3)

#Leemos todos los datos para Modelo: M1t3 dado t2
datos<-list (" narb" ,"O.t3" ," vauxt3")
data<-c(datos)

# valores iniciales para Bugs
dput( file='init3.1.txt ' , list (alpha=0,beta=0))
dput( file='init3.2.txt ' , list (alpha=0.5,beta=-0.5))
init3.1<-source(" init3.1.txt")
init3.2<-source(" init3.2.txt")

# Valores iniciales definitivos
init1<-init3.1\$value
init2<-init3.2\$value

```

```

inits<-list (init1 , init2)

dput( file="datat3.txt" , list (narb ,O.t3 , vauxt3))

#parametros de interés
param<-c("alpha" ,"beta")

#Lanzamos el modelo: M1t3 dado t2
m1t3<-openbugs(data , inits , param,"M1t3 dado t2.txt" ,
n.chains=2,n.iter=10000)

## Estudio de la Cadenas con CODA
sal<-mcmc(m1t3\%sims.matrix[ ,1:2])
summary(sal)
## Pruebas de CONVERGENCIA A LAS CADENAS
plot(sal)
crosscorr(sal)
autocorr(sal)
autocorr.plot(sal)
geweke.plot(sal)
geweke.diag(sal)
heidel.diag(sal)
raftery.diag(sal)

### Modelo Tres: Para t=4 dado t=3 [es decir , año 97 dado 96]
narb<-length(O.t4)

#Leemos todos los datos para Modelo: M1t4 dado t3
datos<-list("narb" ,"O.t4" ,"vauxt4")
data<-c(datos)

# valores iniciales para Bugs
dput( file='init4.1.txt' , list (alpha=0,beta=0))
dput( file='init4.2.txt' , list (alpha=0.5,beta=-0.5))
init4.1<-source("init4.1.txt")
init4.2<-source("init4.2.txt")

# Valores iniciales definitivos
init1<-init4.1\%value
init2<-init4.2\%value
inits<-list (init1 , init2)

```

```

dput( file="datat4.txt", list( narb, O.t4, vauxt4))

#parametros de interés
param<-c("alpha", "beta")

#Lanzamos el modelo: M1t4 dado t3
m1t4<-openbugs(data, inits, param, "M1t4 dado t3.txt",
n.chains=2, n.iter=10000)

## Estudio de la Cadenas con CODA
sal<-mcmc(m1t4$sims.matrix[, 1:2])
summary(sal)
## Pruebas de CONVERGENCIA A LAS CADENAS
plot(sal)
crosscorr(sal)
autocorr(sal)
autocorr.plot(sal)
geweke.plot(sal)
geweke.diag(sal)
heidel.diag(sal)
raftery.diag(sal)

### Modelo Cuatro: Para t=5 dado t=4 [es decir, año 98 dado 97]
narb<-length(O.t5)

#Leemos todos los datos para Modelo: M1t4 dado t3
datos<-list("narb", "O.t5", "vauxt5")
data<-c(datos)

# valores iniciales para Bugs
dput( file='init5.1.txt', list(alpha=0, beta=0))
dput( file='init5.2.txt', list(alpha=0.5, beta=-0.5))
init5.1<-source("init5.1.txt")
init5.2<-source("init5.2.txt")

# Valores iniciales definitivos
init1<-init5.1$value
init2<-init5.2$value
inits<-list(init1, init2)

dput( file="datat5.txt", list( narb, O.t5, vauxt5))

```

```

#parametros de interés
param<-c("alpha","beta")

#Lanzamos el modelo: M1t5 dado t4
mlt5<-openbugs(data,inits,param,"M1t5 dado t4.txt",
n.chains=2,n.iter=10000)

## Estudio de la Cadenas con CODA
sal<-mcmc(mlt5\${sims.matrix[,1:2]})
summary(sal)
## Pruebas de CONVERGENCIA A LAS CADENAS
plot(sal)
crosscorr(sal)
autocorr(sal)
autocorr.plot(sal)
geweke.plot(sal)
geweke.diag(sal)
heidel.diag(sal)
raftery.diag(sal)

#####
## Lanzamos los modelos que consideran el efecto de los
## arboles enfermos del año anterior y la componente
## aleatoria que recoge la heterogeneidad entre los árboles
## en toda la Parcela
Modelo 1: M1t2 dado t1 , más componente: M1t2 dado t1 hetArbol
narb<-length(O.t2)

#Leemos todos los datos para Modelo: M1t2 dado t1
datos<-list("narb","O.t2","vauxt2")
data<-c(datos)

# valores iniciales para Bugs
dput(file='init2.1.txt',list(alpha=0,beta=0,
arb=rnorm(narb,0,1)))
dput(file='init2.2.txt',list(alpha=0.5,beta=-0.5,
arb=rnorm(narb,0,1)))
init2.1<-source("init2.1.txt")
init2.2<-source("init2.2.txt")

# Valores iniciales definitivos
init1<-init2.1\${value}

```

```

init2<-init2.2\$value
inits<-list (init1 , init2)

# Construcción de los datos
dput( file="datat2.txt" , list (narb ,O.t2 , vauxt2))

#parametros de interés
param<-c(" alpha" , " beta" , " arb")

#Lanzamos el modelo: M1t2 dado t1 + componente aleatoria
m1t2<-openbugs(data , inits , param , "M1t2 dado t1 hetArbol.txt" ,
n.chains=2,n.iter=12000)

## Estudio de la Cadenas con CODA
sal<-mcmc(m1t2\$sims.matrix[ ,1:3])
## Pruebas de CONVERGENCIA A LAS CADENAS
plot(sal)
crosscorr(sal)
autocorr(sal)
autocorr.plot(sal)
geweke.plot(sal)
geweke.diag(sal)
heidel.diag(sal)
raftery.diag(sal)

## Creacion de las estructuras de adyacencias y de número
## de arboles vecinos para cada arbol , considerando el efecto
## de los enfermos del año anterior
#####
##### Construcción de la estructura de vecindad #####
##### considerada en la Distribución CAR #####
##### Para t=2 considerando enfermos del año t=1 ###

distmat<-as.matrix(dist(cbind(bu.t[ ,2] , bu.t[ ,3])))
dist.ind<-(distmat<=x)*1
diag(dist.ind)<-0
num<-as.vector(apply(dist.ind , 1 , sum))
n<-length(distmat[ ,1])
adj<-NULL
C<-NULL
for (i in 1:n){
  neigh<-as.vector((1:n)*dist.ind[i ,])

```

```

    neigh<-neigh[neigh>0]
    neigh.C<-as.vector(dist.ind[i,neigh])
    adj<-as.vector(c(adj,neigh))
    C<-as.vector(c(C,neigh.C))
  }
  list(adj=adj,num=num,weights=C)
  k2<-sum(num)
  narb<-length(O.t2)
  dput(file="data.txt",list(adj=adj,num=num,weights=C))

#####
## Lanzamos los modelos que consideran el efecto de los
## arboles enfermos del año anterior y la componente
## aleatoria que recoge la heterogeneidad entre los árboles
## en toda la Parcela Y el efecto aleatorio espacial
Modelo 1: M1t2 hetArbolEfecSpa: M1t2 dado t1 hetArbolEfecSpa
adj<-adj
num<-num
weights<-C
#Leemos todos los datos para Modelo: M1t2 dado t1 hetArbolEfecSpa
datos<-list("narb","O.t2","vauxt2","adj","num","weights")
data<-c(datos)

# valores iniciales para Bugs
phi1<-rep(0,narb)
phi2<-rep(1,narb)
dput(file='init2.1.txt',list(alpha=0,beta=0,arb=rnorm(narb,0,1),
phi=phi1))
dput(file='init2.2.txt',list(alpha=0.5,beta=-0.5,
arb=rnorm(narb,0,1),phi=phi2))
init2.1<-source("init2.1.txt")
init2.2<-source("init2.2.txt")

# Valores iniciales definitivos
init1<-init2.1$value
init2<-init2.2$value
inits<-list(init1,init2)

#parametros de interés
param<-c("alpha","beta","arb","phi","prop.v")

#Lanzamos el modelo: M1t2 dado t1 + componente aleatoria

```

```
m1t2<-openbugs(data ,inits ,param,"M1t2 dado t1 hetArbolEfecSpa1.txt",
n.chains=2,n.iter=5000)
```

```
## Estudio de la Cadenas con CODA
sal<-mcmc(m1t2\%sims.matrix[,1:2])
sal<-mcmc(m1t2\%sims.matrix[,3]) # cadena para efectoHeterog-Arbol
sal<-mcmc(m1t2\%sims.matrix[,405])# cadena para Efecto Spatial ramdon
## Pruebas de CONVERGENCIA A LAS CADENAS
plot(sal)
crosscorr(sal)
autocorr(sal)
autocorr.plot(sal)
geweke.plot(sal)
geweke.diag(sal)
heidel.diag(sal)
raftery.diag(sal)
```

Bibliografía

- Banerjee, S., Carlin, B.P. y Gelfand, A.E. (2004). Hierarchical Modelling and Analysis for Spatial Data. Monographs on Statistics and Applied Probability 101.
- Bar-Joseph, M. y Loebenstein, G. (1972). Effects of strain, source plant, and temperature on transmissibility of citrus tristeza virus by the melon aphid. *Phytopath*, 63, 716–720.
- Bar-Joseph, M., Marcus, R. y Lee, R.F. (1989). The continuous challenge of citrus tristeza virus control. *Phytopath*, 27, 291–316.
- Bennet, C.W. y Costa, A.S. (1949). Tristeza disease of citrus. *J. Agric. Res*, 78, 207–237.
- Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis. 2ª edición.
- Bermúdez, J. (2010). «Guía de asignatura: Métodos de Simulación». Universitat de València, Facultat de Matemàtiques, Dept. d'Estadística i Investigació Operativa.
- Besag, J. (1974). Spatial Interaction and the Statistical Analysis of Lattice Systems. *Journal of the Royal Statistical Society: Series B*, 36, 192–225.
- Besag, J., York, J.C. y Mollié, A. (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, 43, 1–59.
- Best, N.G., Cowles, M.K. y Vines, S.K. (1995). *CODA Manual version 0.30*. MRC Biostatistics Unit, Cambridge, U.K.
- Botella-Rocamora, P. (2010). *Suavización Espacio-Temporal en cartografía de enfermedades*. Tesis, Universitat de València, Facultat de Matemàtiques., Departament d'Estadística i Investigació Operativa.
- Broks, S.P. (1998). Markov chain Monte Carlo methods and its application. *Journal The American Statistician*, 47, 69–100.

BIBLIOGRAFÍA

- Cambra, M., Gorris, M.T., Marroquín, C., Román, M.P., Olmos, A., Martínez, M.C., Hermoso De Mendoza, A., López, A. y Navarro, L. (2000a). Incidence and epidemiology of Citrus tristeza virus in the Valencian Community of Spain. *Virus Res*, 71, 75–85.
- Cambra, M., Serra, J., Bonet, J.C. y Villalba, D. (1990a). «Situación de la tristeza de los cítricos en la Comunidad Valenciana». *Serie Fullets Divulgació pp. 3-90*, Generalidad Valenciana, Conselleria d'Agricultura i pesca, España.
- Cambra, M., Serra, J., Villalba, D. y Moreno, P. (1988). «Present situation of the citrus tristeza virus in the Valencian Community». *Informe técnico pp. 1-7*, Proceedings of the 10th Conference International Organization of Citrus Virologists, (Eds. by Garnsey, S.M., Timmer L.W., Navarro, L.), IOCV, Department of Plant Pathology, University of California, Riverside, USA.
- Carlin, B.P. y Banerjee, S. (2003). Hierarchical multivariate CAR models for spatio-temporally correlated survival data (with discussion). *In Bayesian Statistics*, 7, 45–63. Oxford: Oxford University Press.
- Carlin, B.P. y Louis, T.A. (2000). *Bayes and Empirical Bayes Methods for Data Analysis*. FL: Chapman and Hall/CRC Press, 2ª edición.
- Casella, G. y George, E.I. (1992). Explaining the Gibbs Sampler. *The American Statistician*, 46, 167–174.
- Christensen, O.F. y Waagepetersen, R. (2002). Bayesian prediction of spatial count data using generalised linear mixed models. *Biometrics*, 58, 280–286.
- Clayton, D., Bernardinelli, L. y Montomoli, C. (1993). Spatial correlation in ecological analysis. *International Journal of Epidemiology*, 22, 1193–1202.
- Clifford, P. (1990). *Markov random fields in statistics*. pp. 20-32. In Disorder in Physical Systems.
- Costa, A.S. y Grant, T.J. (1951). Studies on Transmission of the tristeza virus by the vector, *Aphis citricidus*. *Phytopathology*, 41, 105–113.
- Cressie, N. (1993). *Statistics for Spatial Data*. Rev. ed. New York.
- Cressie, N. y Chan, N.H. (1989). Spatial modelling of regional variables. *Journal of the American Statistical Association*, 84, 393–401.
- Cruz de Boelpaepe, M.O. y Ferreira, M.O. (1998). Survey of the brown citrus aphid, *Toxoptera citricida*, and other aphid vector of citrus tristeza virus in Continental Portugal. In: Nieto, J.M., Dixon, A. (Ed.). *Aphids in natural and managed ecosystems*, 1, 525–534. Universidad de León.

BIBLIOGRAFÍA

- De Santis, F. y Spezzaferri, F. (1999). Methods for robust and default Bayesian model comparison: the fractional Bayes factor approach. *Int. Statist. Rev.*, 67, 267–286.
- Diggle, P.J. (2003). *Statistical Analysis of Spatial Point Patterns*. p.104. London, 2ª edición.
- Diggle, P.J., Tawn, J.A. y Moyeed, R.A. (1998). Model-based geostatistics (with discussion). *J. Roy. Statist. Soc., Ser. C (Applied Statistics)*, 47, 299–350.
- Fernandes, J. y Cruz de Boelpalpe, M.O. (1994). «Programa de prospecção de organismos nocivos em citrinos. Sub-programa pragas I, Toxoptera citricidus». p.11, Centro nacional de Protecção da Produção Agrícola, Lisboa.
- Fernández, C. y Green, P.J. (2002). Modelling spatially correlated data via mixtures: a Bayesian approach. *Journal of the Royal Statistical Society: Series B*, 64, 805–826.
- Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall.
- Garnsey, S.M. y Cambra, M. (1991). «Enzyme-linked immunosorbent assay (ELISA) for citrus pathogens. In: Roistacher, C.N. (Ed.), Graft-transmissible Diseases of Citrus». *Handbook for detection and diagnosis*, FAO, Rome. pp. 193-216.
- Gelfand, A.E., Schmidt, A.M., Wu, S., Silander, J., Latimer, A. y Rebelo, A.G. (2005). Modelling species diversity through species level hierarchical modeling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(1), 1–20.
- Gelfand, A.E. y Smith, A.F.M. (1990). Sampling-based approach to calculating marginal densities. *Journal American Statistics Association*, 85, 398–409.
- Gelman, A. (2003). *Bugs.R: functions for calling Bugs from R*. www.stat.columbia.edu/~gelman/bugsR/.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Journal of the American Statistical Association*, 3, 515–533.
- Gelman, A. y Rubin, D. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.
- Gelman, A., Sturtz, S. y Ligges, U. (2005). R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software*, 12(3), 1–16.
- Geman, S. y Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *J. IEEE Trans. Pattn. Anal. Mach. Intell.*, 6, 721–741.

BIBLIOGRAFÍA

- Geweke, J. (1992). Evaluation the accuracy of sampling-based approaches to the calculation of posterior moments. *En Bayesian Statistics*, 4, 169–193. (J.M. Bernardo, A.F.M. Smith, A.P. Dawid and J.O. Berger editores). Oxford: Oxford University Press.
- Gilks, W.R., Richardson, S. y Spiegelhalter, D.J.(eds.). (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall, London.
- Gottwald, T.R., Cambra, M., Moreno, P., Camarasa, E. y Piquer, J. (1996). Spatial and temporal analyses of citrus tristeza virus in Eastern Spain. *Phytopathology*, 86, 45–55.
- Gottwald, T.R., Garnsey, S.M., Cambra, M., Moreno, P., Irej, M. y Borbon, J. (1997). Comparative effects of aphid vector species on increase and spread of citrus tristeza virus. *Fruits*, 52, 397–404.
- Green, P.J. y Richardson, S. (2002). Hidden Markov models and disease mapping. *Journal of the American Statistical Association*, 97, 1055–1070.
- Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57, 97–109.
- Heidelberger, P. y Welch, P.D. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, 31(6), 1109–1144.
- Hermoso de Mendoza, A., Ballester-Olmos, J.F. y Pina, J.A. (1984). «Transmission of citrus tristeza virus by aphids (Homoptera, Aphididae) in Spain». *Informe técnico pp. 23-27*, Proceedings of the 9th Conference International Organization of Citrus Virologists, (Eds. by Garnsey, S.M., Timmer L.W., Doods, J.A.), IOCV, Departament of Plant Pathology, University of California, Riverside, USA.
- Hermoso de Mendoza, A., Álvarez A., J.M., Michelena, González, P. y Cambra, M. (2008). Dispersión, biología y enemigos naturales de *Toxoptera citricida* (Kirkaldy) (Hemiptera, Aphididae) en España. *Boletín de Sanidad Vegetal Plagas*, 34, 77–87.
- Hobert, J.P. y Casella, G. (1996). The effect of improper priors on Gibbs Sampling in hierarchical linear mixed models. *Journal of the American Statistical Association*, 91, 1461–1473.
- Ilharco, F.A., Sousa-Silva, C.R. y Álvarez Álvarez, A. (2005). First report on *Toxoptera citricidus* (Kirkaldy) in Spain and continental Portugal (Homoptera-Aphidoidea). *Journal Agronomia Lusitana*, 51(1), 19–21.
- Karasev, A.V., Boyko, V.P., Gowda, S., Nikolaera, O.V., Hilf, M.E., Koonin, E.V., Niblett, C.L., Cline, K., Gumpf, D.J., Lee, R.F., Garnsey, S.M. y Dawson, W.O. (1995). Complete sequence of the citrus tristeza virus RNA genome. *Virology*, 208, 511–520.

BIBLIOGRAFÍA

- Kensall, J. E. y Diggle, P.J. (1995a). Kernel estimation of relative risk. *Bernoulli*, 1(1-2), 3–16.
- Kensall, J. E. y Diggle, P.J. (1995b). Non-parametric estimation of spatial variation in relative risk. *Statistics in Medicine*, 14, 559–573.
- Kensall, J. E. y Diggle, P.J. (1998). Spatial variation in risk: A non-parametric binary regression approach. *Applied Statistics*, 47, 559–573.
- Kensall, J. E. y Wakefield, J.C. (2002). Modeling spatial variation in disease risk: a geostatistical approach. *Journal of the American Statistical Association*, 97, 692–701.
- Knorr-Held, L. y Raßer, G. (2000). Bayesian detection of clusters and discontinuities in disease maps. *Biometrics*, 56, 13–21.
- Lastra, R., Menesses, R., Still, P.E. y Niblett, C.L. (1991). «The citrus tristeza situation in Central America». *Informe técnico pp. 146-149*, Proceedings 11th Conference International Organization of Citrus Virologists, (Eds. by R.H. Brlansky, R.F. Lee, L.W. Timmer), IOCV, Department of Plant Pathology, University of California, Riverside, USA.
- Leroux, B.G. (2000). Modeling spatial disease rates using maximum likelihood. *Statistics in Medicine*, 19, 2321–2332.
- License, GNU General Public (2010). *OpenBUGS, v. 3.1.1 rev 524*. OpenBUGS Foundation, <http://www.openbugs.info/w/>.
- López, A. (2008). «Guía de asignatura: Estadística Espacial en Epidemiología y Medio ambiente». Universitat de València, Facultat de Matemàtiques, Dept. d'Estadística i Investigació Operativa.
- López, A. (2010). «Guía de asignatura: Estadística Espacial y Medioambiental». Universitat de València, Facultat de Matemàtiques, Dept. d'Estadística i Investigació Operativa.
- Maritz, J.S. y Lwin, T. (1989). *Empirical Bayes Methods*. Chapman and Hall, London, 2^a edición.
- Marroquin, C., Olmos, A., Gorris, M.T., Bertolini, E., Martinez, M.C., Carbonell, E.A., Hermoso de Mendoza, A. y Cambra, M. (2004). Estimation of the number of aphids carrying Citrus tristeza virus that visit adult citrus trees. *Virus research*, 100, 101–108.
- Martí, F. (1977). La enfermedad de la tristeza de los agrios en España. In: I Congreso Mundial de Citricultura. *Inter. Soc. Citriculture*, 1, 531–536.
- McCulloch, C.E. (1997). Maximum likelihood algorithms for Generalized Linear Mixed Models. *Journal of the American Statistical Association*, 92, 162–170.

BIBLIOGRAFÍA

- Meneghini, M. (1946). Sobre a natureza e transmissibilidade da doença "Tristeza" dos Citrus. *O Biológico*, 12, 285–287.
- Metropolis, N., Rosenbluth, A.W., Sosenbluth, M.N., Teller, A.H. y Teller, E. (1953). Equations of state calculations by fast computing machines. *J. of Chem. Phys.*, 21, 1087–1092.
- Michaud, J.P. y Álvarez, R. (2000). Toxoptera citricida in México. *IOCV, newsletter*, 1, 2.
- Möller, J. (2003). *Spatial Statistics and Computational Methods*. Springer Verlag, New York.
- Möller, J., Syversveen, A.R. y Waagepetersen, R.P. (1998). Log gaussian cox processes. *Scandinavian Journal of Statistics*, 25, 451–482.
- Moreno, P. (1995). La tristeza y el pulgón pardo de los cítricos (*Toxoptera citricidus* Kirk.): Una amenaza inmediata para la citricultura en los países ribereños del Caribe. *Levante Agrícola*, 333(34), 316–325.
- Mugglin, A.S., Carlin, B.P. y Gelfand, A.E. (2000). Fully model-based approaches for spatially misaligned data. *Journal of the American Statistical Association*, 95, 877–887.
- Navarro, L., Juarez, J., Pina, J.A. y Ballester, J.F. (1984). «The Citrus quarantine station in Spain». *Informe técnico*, Proceedings of the 9th Conference International Organization of Citrus Virologists, (Eds. by Garnsey, S.M., Timmer L.W., Doods, J.A.).
- Navarro, L., Roistacher, C.N. y Murashige, T. (1975). Improvement of shoot-tip grafting in vitro for virus-free citrus. *J. Am. Soc. Hort. Sci.*, 100, 471–479.
- Norman, P.A. y Grant, T.J. (1954). Preliminary studies of aphid transmission of tristeza virus in Florida. *The Citrus Industry*, 35, 10–12.
- Norman, P.A. y Grant, T.J. (1956). Transmission of tristeza virus by aphids in Florida. *Proc. Fla. Hort. Soc.*, 69, 38–42.
- Pollard, G. (1997). Update on new pest introductions in the Caribbean. *Caraphin News*, 15, 11.
- R, Development Core Team (2008). *R: A Language and Environment for Statistical Computing*, v. 2.9.0. R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>. ISBN 3-900051-07-0.
- Raccah, B., Loebenstein, G., Bar-Joseph, M. y Oren, Y. (1976). «Transmission of tristeza by aphids prevalent on citrus, and operation on the tristeza suppression programme in Israel». *Informe técnico pp. 47-49*, Proceedings of the 7th Conference International Organization of

BIBLIOGRAFÍA

- Citrus Virologists, IOCV, Department of Plant Pathology, University of California, Riverside, USA.
- Raftery, A.E. y Lewis, S.M. (1992). How many iteraciones in the Gibbs sampler? *Bayesian Statistics*, 4, 763–773. Oxford: Oxford University Press. (J.M. Bernardo, A.F.M. Smith, A.P. Dawid and J.O. Berger editores).
- Richardson, S. (2003). *Spatial models in epidemiological applications*. capítulo Highly Structured Stochastic Systems, pp. 237–259. P.J. Green, N.L. Hjort, and S. Richardson (Eds.).
- Robert, C.P. y Casella, G. (1999). *Monte Carlo Statistical Methods*. Springer, New York.
- Robert, C.P. y Smith, A.F.M. (1994). *Simple condtions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms*. University of Minnesota.
- Rocha-Peña, M.A., Lee, R.F., Lastra, R., Niblett, C.L., Ochoa-Corona, F.M., Garnsey, S.M. y Yokomi, R.K. (1995). Citrus tristeza virus and its vector Toxoptera citricida. Threats to citrus production in the Caribbean and Central and North America. *Plant Disease*, 79, 437–445.
- Román, M.P., Cambra, M., Juárez, J., Moreno, P., Duran-Vila, N., Tanaka, F.A.O., Alves, E., Kitajina, E.W., Yamamoto, P.T., Basanezi, R.B., Teixeira, D.C., Jesús Junior, W. C., Ayres, A.J., Gimenes-Fernandes, N., Rabenstein, F., Giroto, L.F. y Bovo, J. M. (2004). Sudden death of Citrus in Brazil: A graft-transmissible bud union disease. *Plant Disease*, 88, 453–467.
- Schwarz, G. (1978). Estimating the dimension of a model. *Journal Annals of Statistics*, 6(2), 461–464.
- Spiegelhalter, D.J., Best, N., Carlin, B.P. y Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). , *Journal of the Royal Statistical Society: Series B*, 64, 583–639.
- Spiegelhalter, D.J., Thomas, A., Best, N., Gilks, W.R. y Lunn, D. (1994, 2003). *BUGS: Bayesian inference using Gibbs sampling, Version 0.50*. MRC Biostatistics Unit, Cambridge, England.
- Spiegelhalter, D.J., Thomas, A., Best, N. G. y Lunn, D. (2007). *OpenBUGS User Manual version 3.0.2*. MRC Biostatistics Unit, Cambridge, England.
- Tanner, M.A. y Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.*, 82, 528–550.
- Leroux, B.G. and Lei, X. and Breslow, N. (1999). *Estimation of disease rates in small areas: A new mixed model for spatial dependence*. *Statistical models in epidemiology, the environment and clinical trials*, , 179–192. Springer, Berlin Heidelberg. (Eds. Halloran M.E, Berry, D.).

BIBLIOGRAFÍA

- Varma, P.M., Rao, D.G. y Capoor, S.D. (1965). Transmission of tristeza virus by *Aphis craccivora* (Koch) and *Dactynotus jaceae* (L.). *Indian J. Entomol.*, 27, 67–71.
- Varma, P.M., Rao, D.G. y Vasudeva, R.S. (1960). Additional vectors of tristeza disease of Citrus in India. *Curr. Sci*, 29, 359.
- Yokomi, R.K., Garnsey, S.M., Civerolo, E. L. y Gumpe, D. (1989). Transmission of exotic citrus tristeza isolates by a Florida colony of *Aphis gossypii*. *Planta Disease*, 73, 552–556.
- Yokomi, R.K., Lastra, R., Stoetzel, M.B., Damsteegt, V.D., Lee, R.F., Garnsey, S.M., Gottwald, T.R., Rocha-Peña, M. A. y Niblett, C. N. (1994). Establishment of the brown citrus aphid *Toxoptera citricida* (Kirkaldy) (Homoptera:Aphididae) in Central American and the Caribbean basin and its transimission of citrus tristeza virus. *J. Econ. Entomol.*, 87, 1078–1085.

Otros aspectos a considerar

Este trabajo hasta ahora ha sido evaluado de dos formas:

1. Como trabajo de investigación, obteniendo una nota sobresaliente de 10/10 en el Programa de Doctorado: Estadística y Optimización, realizado en la Universidad Politécnica de Valencia (UPV), Valencia-España.
2. En defensa pública, ante tribunal designado por la UPV como requisito para optar al título de Estudios Avanzados en Investigación (DEA).

Así mismo, se anexa para cada forma de evaluación los soportes correspondientes. Se resalta que en cuanto a la segunda evaluación, además de agregar el acta, se anexan dos soportes correspondientes al certificado del título de Estudios Avanzados (DEA), detallados a continuación

- Diploma del DEA, Parte 1.
- Diploma del DEA, Parte 2, donde se detalla la carga académica cursada hasta llegar al trabajo de investigación.