

Universidade Federal do Rio de Janeiro

**Modelos Bayesianos Semi-paramétricos:
uma aplicação a dados de saúde e meio ambiente**

Carolina Pereira Ornelas

Rio de Janeiro

2006

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Departamento de Métodos Estatísticos
Instituto de Matemática
Universidade Federal do Rio de Janeiro

Carolina Pereira Ornelas

**Modelos Bayesianos Semi-paramétricos:
uma aplicação a dados de saúde e meio ambiente**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Orientadoras:

Alexandra Mello Schmidt

Mariane Branco Alves

junho de 2006

ORNELAS, CAROLINA PEREIRA

Modelos Bayesianos Semi-paramétricos:
uma aplicação a dados de saúde e meio
ambiente [Rio de Janeiro] 2006

?, ?p. 29,7 cm (IM/UFRJ, M.Sc.,
Estatística, 2006)

Dissertação - Uni-
versidade Federal do Rio
de Janeiro, IM/UFRJ

1.? 2.? 3.? 4.?

I. IM/UFRJ II. Título (série)

Folha de exame

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro como parte dos requisitos necessários à obtenção do título de Mestre em Estatística.

Banca examinadora:

Alexandra M. Schmidt, PhD
Instituto de Matemática - UFRJ

Mariane B. Alves, D. Sc
Instituto de Matemática e Estatística - UERJ

Flávia M. P. Landim, D. Sc
Instituto de Matemática - UFRJ

Antonio Carlos M. Ponce de Leon, PhD
Instituto de Medicina Social - UERJ

Resumo

A classe dos Modelos Aditivos Generalizados (MAG) pode ser vista como uma extensão dos Modelos Lineares Generalizados (MLG), e permite que o preditor linear seja substituído por um preditor aditivo, composto pela soma de funções suaves s . Neste trabalho, em particular, as funções não-paramétricas utilizadas para suavização são *splines*. Com isso, pode-se inserir no preditor variáveis que não influenciam a resposta média de forma linear, criando-se uma estrutura mais flexível para modelagem dos dados.

O procedimento de inferência seguirá os enfoques clássico e bayesiano e, para exemplificar o método, os modelos propostos são ajustados a um conjunto de dados de contagens de óbitos em São Paulo, com observações feitas no período de 01 de janeiro de 1994 a 31 de dezembro de 1997. Pressupõe-se distribuição de Poisson para a contagem de óbitos e a média é explicada em função de variáveis climáticas e poluentes atmosféricos. Uma segunda aplicação é feita utilizando resposta contínua, modelando-se observações de monóxido de carbono (CO) deste mesmo banco de dados. Aqui, as variáveis envolvidas no preditor são temperatura e umidade.

Palavras chave: Modelo aditivo generalizado, spline cúbico.

Abstract

The class of Generalized additive models (GAM) can be seen as an extension of the Generalized linear models (GLM), where the linear predictor is substituted by an additive predictor, composed by a sum of smooth functions s . In particular, in this work the non-parametric functions are *splines*. With this, it can be inserted in the predictor variables that does not influence the mean response in a linear way, creating a more flexible structure for modelling the data.

The inference procedure will follow both the classical and Bayes paradigm and to exemplify the method, the proposed models are applied to a database of daily death counts in São Paulo, registered from January 1st, 1994 until December 31st, 1997. It is assumed a Poisson distribution for the death counts and its mean is described as a function of climatic variables and atmospheric pollutants. A second application is made using continuous response, modelling carbon monoxide (CO) observations of this same database. Here, the variables introduced in the predictor are temperature and humidity.

Key words: Generalized additive models, cubic spline.

Agradecimentos

Agradeço à minha família por compreender minha ausência, e pelo apoio às minhas decisões.

Agradeço também às professoras Alexandra M. Schmidt e Mariane B. Alves pelo incentivo, paciência, oportunidade e confiança.

À amiga Michele Cunha pelas horas de sono perdidas e pelo livro de Medida e Integração que nunca devolvi.

À amiga Raquel Vasconcellos pela companhia nas madrugadas em frente ao computador.

Um agradecimento especial ao amigo Marcus Vinícius. Mais uma vez afirmo o que te escrevi em, se não me engano, setembro de 2004: menino, tua inteligência só não é maior que teu coração!

A todos os companheiros de curso, em especial, Val, Carla, Ana e Lucinha, pela amizade sincera.

A Washington Junger (novamente), pelos conselhos e “puxões de orelha” regados a café.

Ao professor Oswaldo Gonçalves Cruz que por várias vezes solucionou dúvidas sem mesmo me conhecer.

Ao Laboratório de Poluição Atmosférica Experimental da Faculdade de Medicina da USP pela construção do banco de dados de poluentes e ao professor Julio Singer pela disponibilização deste banco de dados em sua página pessoal.

À CAPES, pelo apoio financeiro.

A todos aqueles que de alguma forma participaram desta etapa de minha vida deixo meus mais sinceros votos de felicidades. Obrigada.

Sumário

1	Introdução	1
1.1	Revisão de Literatura	3
1.2	Modelos Lineares	3
1.2.1	Modelos de Defasagens Distribuídas	4
1.3	Modelos Lineares Generalizados	7
2	Modelos Aditivos Generalizados	9
2.1	Splines	9
2.2	Representação por <i>B-splines</i>	11
2.3	Seleção do parâmetro de suavização da spline	12
2.4	Inferência sobre Modelos Aditivos Generalizados	15
2.4.1	Estimação sob o enfoque clássico	16
2.4.2	Estimação sob o enfoque bayesiano	17
3	Modelos propostos	23
3.1	Descrição do banco de dados	24
3.2	Poluentes estudados	26
3.2.1	Monóxido de carbono - CO	28
3.2.2	Material particulado - PM_{10}	28
3.2.3	Dióxido de enxofre - SO_2	28
3.3	Critérios de comparação de modelos	31
3.3.1	Critério de Informação de Akaike - AIC	31
3.3.2	Critério de Informação da Deviance - DIC	32
3.4	Modelos ajustados	33
3.4.1	Aplicação utilizando resposta discreta: Mortalidade em São Paulo	33
3.4.2	Aplicação utilizando resposta contínua: Níveis de CO em São Paulo	58
4	Considerações Finais	72

Referências Bibliográficas	74
Apêndice A: Diferenças divididas	78

Lista de Figuras

1.1	Aproximação polinomial para os coeficientes β_j	5
3.1	Número observado de óbitos por doenças respiratórias em menores de 5 anos no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo. . .	25
3.2	Temperatura mínima observada no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.	25
3.3	Umidade média observada no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.	26
3.4	Nível observado de CO no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.	29
3.5	Nível observado de PM_{10} no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.	30
3.6	Nível observado de SO_2 no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.	30
3.7	Histogramas das amostras das distribuições a posteriori dos coeficientes ρ_0, ρ_1 e ρ_2 para os poluentes PM_{10}, CO e SO_2 , no modelo de defasagens distribuídas com restrições polinomiais de ordem 2.	38
3.8	Evolução temporal dos efeitos estimados associados a aumentos nos níveis de PM_{10}, CO e SO_2 , para os modelos de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4.	42
3.9	Evolução temporal dos efeitos estimados associados a aumentos nos níveis de PM_{10}, CO e SO_2 , para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2.	43

3.10	Evolução temporal e intervalos de confiança / credibilidade dos efeitos estimados associados a aumentos nos níveis de PM_{10} , CO e SO_2 para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2.	44
3.11	Evolução temporal dos riscos relativos estimados associados a aumentos nos níveis de PM_{10} , CO e SO_2 , para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2, seguindo o enfoque bayesiano. . .	45
3.12	Evolução temporal dos efeitos estimados associados ao nível máximo observado dos poluentes PM_{10} , CO e SO_2 , para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2.	46
3.13	Histograma do parâmetro de suavização λ das variáveis umidade e temperatura para os poluentes PM_{10} , CO e SO_2 , no modelo de defasagens distribuídas com restrições polinomiais de ordem 2.	48
3.14	Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável temperatura no modelo contendo o poluente CO como regressora.	49
3.15	Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável umidade no modelo contendo o poluente CO como regressora.	50
3.16	Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável temperatura no modelo contendo o poluente PM_{10} como regressora.	51
3.17	Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável umidade no modelo contendo o poluente PM_{10} como regressora.	52
3.18	Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável temperatura no modelo contendo o poluente SO_2 como regressora.	53
3.19	Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável umidade no modelo contendo o poluente SO_2 como regressora.	54

3.20	Efeito não linear de umidade e temperatura, estimado seguindo o enfoque clássico, no modelo de defasagens distribuídas com restrições polinomiais de ordem 2 para os poluentes PM_{10} , CO e SO_2	56
3.21	Efeito não linear de umidade e temperatura, estimado seguindo o enfoque bayesiano, no modelo de defasagens distribuídas com restrições polinomiais de ordem 2 para os poluentes PM_{10} , CO e SO_2	57
3.22	Histograma do parâmetro de suavização λ da variável temperatura no modelo 9 utilizando $\log(CO)$ e média semanal de $\log(CO)$ como resposta.	65
3.23	Histograma do parâmetro de suavização λ da variável umidade no modelo 13 utilizando CO e média semanal de CO como resposta.	66
3.24	Histograma das amostras das distribuições a posteriori dos coeficientes ϕ para $\log CO$ e a variável temperatura.	67
3.25	Histograma das amostras das distribuições a posteriori dos coeficientes ϕ para CO e a variável umidade.	68
3.26	Efeito não linear da variável temperatura, estimado seguindo os enfoques clássico e bayesiano, no modelo 9 utilizando $\log(CO)$ e média semanal de $\log(CO)$ como resposta.	70
3.27	Efeito não linear da variável umidade no modelo 13, estimado seguindo os enfoques clássico e bayesiano, utilizando CO e média semanal de CO como resposta.	71

Lista de Tabelas

3.1	Estações monitoradoras	27
3.2	Estatísticas descritivas dos poluentes PM_{10} , CO e SO_2	31
3.3	Resultados da estimação dos modelos com média das defasagens 0,1,2 para os poluentes CO , PM_{10} e SO_2	34
3.4	Resultados da estimação dos modelos de defasagens distribuídas por 40 dias, sem restrições para os poluentes CO , PM_{10} e SO_2	36
3.5	Critérios de comparação resultantes da estimação dos modelos de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4 para os poluentes CO , PM_{10} e SO_2	37
3.6	Matrizes de covariância dos coeficientes ρ_0 , ρ_1 e ρ_2 , estimadas seguindo os enfoques clássico e bayesiano para os poluentes PM_{10} , CO e SO_2 , no modelo de defasagens distribuídas com restrições polinomiais de ordem 2.	39
3.7	Coefficientes de efeito global resultantes da estimação do modelo de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4, para os poluentes CO , PM_{10} e SO_2	40
3.8	Estatísticas sumárias do parâmetro de suavização λ das variáveis umidade e temperatura para os modelos tendo como regressoras os poluentes PM_{10} , CO ou SO_2 , com defasagens distribuídas sujeitas a restrições polinomiais de ordem 2.	47
3.9	Representação esquemática dos modelos ajustados aos dados de poluição.	59
3.10	Critérios de comparação de modelos para observações diárias de $\log(CO)$	60
3.11	Critérios de comparação de modelos para observações diárias de CO	61
3.12	Critérios de comparação de modelos para a média semanal de $\log(CO)$	63

3.13 Critérios de comparação de modelos para a média semanal de CO	64
3.14 Média, mediana e intervalo de credibilidade do parâmetro de suavização λ para os modelos 9 e 13 com respostas diárias e semanais.	66

Capítulo 1

Introdução

Neste trabalho será verificado o comportamento de variáveis explicativas que não necessariamente tenham uma influência linear na resposta, e/ou respostas que são afetadas não só pela variável explicativa no dia corrente, mas também no período anterior ou períodos anteriores. Além disso, existe um grande interesse em trabalhar com variáveis resposta com distribuição não gaussiana e, desta forma, se torna necessário trabalhar com uma classe de modelos que suporte esta estrutura.

Embora os modelos lineares generalizados (McCullagh & Nelder, 1989) permitam contornar o problema de não normalidade da variável resposta, estes não suportam que se analise conjuntamente, por exemplo, uma estrutura que contenha covariáveis defasadas no tempo e variáveis explicativas que não impactam a resposta média de forma linear. O uso de modelos com defasagens é justificado pela idéia de que, em muitas situações, o desfecho no dia corrente é afetado não só pelo nível das variáveis explicativas no mesmo dia, mas também do dia anterior ou dias anteriores. Ou seja, a motivação para tais modelos é o reconhecimento de que muitos eventos que afetam a resposta persistem ao longo do tempo.

A idéia de trabalhar com suavizações de variáveis tem suas raízes nas indústrias aeronáutica e de construção de navios. Em Bartels et al (1987) é descrita uma técnica para construir moldes de aviões utilizada pela indústria aérea britânica na 2ª Guerra Mundial, que consistia em passar finas pranchas de madeira por pontos traçados no chão. As pranchas seriam mantidas no lugar em pontos discretos, e entre estes pontos

assumiria formas de energia mínima.

Em estatística a importância do uso de técnicas de suavização de variáveis surge quando, por exemplo, não é possível obter um modelo linear que se ajuste bem aos dados, ou quando existem variáveis explicativas para as quais o interesse está apenas na possível consideração de seus efeitos, mas não na estimativa da magnitude de tais efeitos. Uma forma de introduzir esta particularidade no modelo é trabalhar com uma função da variável em questão - geralmente um polinômio de baixo grau, cujos coeficientes são estimados através de mínimos quadrados (Green e Silverman, 1994). Esta técnica, porém, pode não ser eficiente nos casos em que se tem observações extremas, pois estas podem influenciar várias regiões da curva, e desta forma, o objetivo principal deste estudo será buscar uma estrutura que seja flexível o suficiente para:

- (i) modelar variáveis resposta cuja distribuição pertença à família exponencial;
- (ii) incluir no modelo variáveis que impactam a resposta média de forma não linear;
- (iii) permitir o tratamento de efeitos defasados de covariáveis sobre a resposta.

O procedimento de inferência neste trabalho seguirá os enfoques clássico e bayesiano, e sua organização será feita da seguinte forma:

No capítulo 2 é feita uma apresentação formal das splines, bem como das técnicas de estimação de modelos semi-paramétricos seguindo os enfoques clássico e bayesiano, e são mostradas as formas de se obter amostras da distribuição a posteriori. No capítulo 3 são introduzidos os modelos propostos para ajustar os dados segundo a metodologia apresentada no capítulo anterior e é feita uma breve descrição dos dados, aos quais os modelos propostos são ajustados. Neste capítulo também se encontram descritos os critérios adotados para comparar os modelos estimados e as aplicações, e são listados os resultados alcançados. Finalmente, no capítulo 4 são apresentadas as considerações finais.

1.1 Revisão de Literatura

No contexto do presente trabalho a principal restrição ao uso dos modelos lineares generalizados é o caso em que se tem variáveis explicativas que não influenciam a resposta média de forma linear, e uma alternativa é trabalhar com a classe dos Modelos Aditivos Generalizados (MAG) proposta por Hastie e Tibshirani (1990). Esta classe de modelos permite que o preditor linear seja substituído por um preditor aditivo, composto pela soma de funções f quaisquer. Neste trabalho, em particular, as funções não-paramétricas utilizadas para suavização são *splines*, apresentadas como em Eilers e Marx (1996), que assumem que o efeito f de uma covariável x pode ser aproximado por uma spline polinomial escrita em termos de uma combinação linear de funções base B-spline. Biller e Fahrmeir (1997) apresentaram uma aproximação bayesiana completa para suavizadores como as splines, e Wahba (2000) estudou diversas generalizações das *splines* que são obtidas como solução de problemas variacionais. Fahrmeir e Lang (2001) apresentam uma forma de realizar inferência bayesiana através de simulação de Monte Carlo em modelos aditivos semi-paramétricos e mistos, e Lang e Brezger (2004) introduzem uma generalização deste trabalho, onde mostram uma versão bayesiana para as splines de Eilers e Marx adotando prioris passeio aleatório que servem como prioris de suavização para os coeficientes das B-splines.

Neste capítulo será feita uma breve revisão dos modelos de Hastie e Tibshirani e de como serão tratados os efeitos defasados de covariáveis. Será mostrada também uma formalização das técnicas não-paramétricas utilizadas segundo de Boor (1978).

1.2 Modelos Lineares

Um modelo linear é aquele em que uma variável dependente Y é explicada por um conjunto de variáveis independentes $X_t = X_{1t}, \dots, X_{pt}$ por meio de uma função linear nos parâmetros, da seguinte forma:

$$E(Y_t | \mathbf{X}_t) = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_p X_{pt} \quad (1.1)$$

Em notação matricial, tem-se:

$$E(Y_t | \mathbf{X}_t) = \mathbf{X}_t \beta \quad (1.2)$$

onde:

Y_t é o valor da variável resposta no tempo t ,

$\beta' = (\beta_0, \dots, \beta_p)$ é o vetor de coeficientes,

\mathbf{X}_t é o vetor de dimensão $1 \times (p + 1)$ cujos elementos são dados pela t -ésima observação de p covariáveis independentes, e o primeiro elemento é 1.

1.2.1 Modelos de Defasagens Distribuídas

A idéia de trabalhar com variáveis defasadas no tempo surgiu inicialmente da necessidade de modelar séries econômicas, onde se procurava estudar quando uma determinada medida econômica teria efeito, ou qual seria a durabilidade de algum bem de consumo (Solow, 1960; Chetty, 1971). Esta estrutura de modelo se estendeu naturalmente a outras áreas do conhecimento, sendo utilizada amplamente em ciências sociais e, mais recentemente, na área de saúde pública (Schwartz, 2000; Alves, 2003).

Um modelo de defasagens distribuídas sobre uma variável X apresenta a seguinte forma:

$$E(Y_t | \underline{X}) = \gamma + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_q X_{t-q} \quad (1.3)$$

onde $X_t, X_{t-1}, \dots, X_{t-q}$ representam a covariável de interesse no tempo corrente, no tempo imediatamente anterior, e até q tempos anteriores, e γ é um intercepto. O preditor pode ter, ainda, outros blocos estruturais, representando tendência, sazonalidade, e outras covariáveis.

Apesar de ser justificado pelo fato de a resposta média no tempo corrente ser afetada não só pelo nível de variáveis explicativas no tempo imediato, mas também por seu nível no tempo anterior ou anteriores, o uso de modelos de defasagens distribuídas acarreta alguns problemas como, por exemplo, autocorrelação, visto que é razoável imaginar que observações subsequentes de uma variável sejam dependentes no tempo. Além disso,

acrescentar no modelo defasagens de uma variável implica em aumento do número de parâmetros a estimar, o que pode diminuir a precisão das estimativas e aumentar o custo computacional.

Almon (1965) propõe uma estrutura que vem sendo utilizada em muitos estudos epidemiológicos. Para implementá-la, procede-se do seguinte modo:

Primeiramente, os coeficientes β_j são reescritos como

$$\beta_j = \sum_{k=0}^d \rho_k j^k, \text{ para } j = 0, \dots, q \quad (1.4)$$

onde q é o número máximo de defasagens e d é o grau de um polinômio nas defasagens que busca aproximar, de forma suave, a trajetória descrita pelos coeficientes $\beta_0, \beta_1, \dots, \beta_q$, como ilustra a figura 1.1.

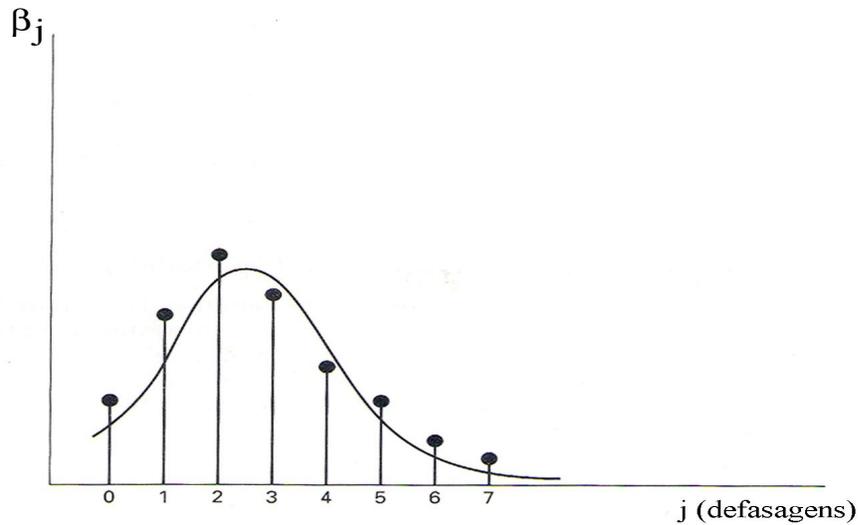


Figura 1.1: Aproximação polinomial para os coeficientes β_j .

Deve-se observar que a imposição de muita restrição sobre os coeficientes β_j (valores baixos de d) pode introduzir tendência no modelo, já que não haverá liberdade suficiente para estimar os coeficientes, e a imposição de pouca restrição (valores altos de d) pode produzir estimativas que serão ruidosas demais para serem informativas.

Tem-se que:

$$\begin{aligned}
\beta_0 &= \rho_0 \\
\beta_1 &= \rho_0 1^0 + \rho_1 1^1 + \rho_2 1^2 + \cdots + \rho_d 1^d \\
\beta_2 &= \rho_0 2^0 + \rho_1 2^1 + \rho_2 2^2 + \cdots + \rho_d 2^d \\
&\vdots \\
\beta_q &= \rho_0 q^0 + \rho_1 q^1 + \rho_2 q^2 + \cdots + \rho_d q^d
\end{aligned} \tag{1.5}$$

Aplicando as restrições (1.4) à estrutura (1.3), o modelo passa a ser escrito como:

$$\begin{aligned}
E(Y_t | X) &= \gamma + \rho_0(X_t) \\
&+ (\rho_0 + \rho_1 + \rho_2 + \cdots + \rho_d)X_{t-1} \\
&+ (\rho_0 2^0 + \rho_1 2^1 + \rho_2 2^2 + \cdots + \rho_d 2^d)X_{t-2} \\
&\vdots \\
&+ (\rho_0 q^0 + \rho_1 q^1 + \rho_2 q^2 + \cdots + \rho_d q^d)X_{t-q}
\end{aligned} \tag{1.6}$$

Coletando os termos que multiplicam cada um dos ρ_i 's e reescrevendo (1.6), tem-se

$$\begin{aligned}
E(Y_t | X) &= \gamma + \rho_0(X_t + X_{t-1} + X_{t-2} + \cdots + X_{t-q}) \\
&+ \rho_1(X_{t-1} + 2X_{t-2} + 3X_{t-3} + \cdots + qX_{t-q}) \\
&+ \rho_2(X_{t-1} + 4X_{t-2} + 9X_{t-3} + \cdots + q^2 X_{t-q}) \\
&\vdots \\
&+ \rho_d(X_{t-1} + 2^d X_{t-2} + 3^d X_{t-3} + \cdots + q^d X_{t-q})
\end{aligned} \tag{1.7}$$

Assim, definindo $d + 1$ novas variáveis $Z_{j,t}$ como somas ponderadas das variáveis \mathbf{X} , tem-se que:

$$\begin{aligned}
Z_{0,t} &= X_t + X_{t-1} + X_{t-2} + \cdots + X_{t-q} \\
Z_{j,t} &= X_{t-1} + 2^j X_{t-2} + 3^j X_{t-3} + \cdots + q^j X_{t-q}, \quad j = 1, 2, \dots, d
\end{aligned} \tag{1.8}$$

e

$$E(Y_t | X) = \gamma + \rho_0 Z_{0,t} + \rho_1 Z_{1,t} + \rho_2 Z_{2,t} + \dots + \rho_d Z_{d,t}. \tag{1.9}$$

Portanto, para trabalhar com um modelo com restrição polinomial, basta reescrevê-lo com parâmetros ρ em termos das variáveis Z obtidas da transformação de X .

Em notação matricial, tem-se:

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_q \end{bmatrix}, \mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1^0 & 1^1 & 1^2 & 1^3 & \cdots & 1^d \\ 2^0 & 2^1 & 2^2 & 2^3 & \cdots & 2^d \\ & & \vdots & & & \\ q^0 & q^1 & q^2 & q^3 & \cdots & q^d \end{pmatrix} \text{ e } \rho = \begin{bmatrix} \rho_0 \\ \rho_1 \\ \vdots \\ \rho_d \end{bmatrix}, \text{ onde } \mathbf{T} \text{ é a matriz}$$

da transformação. Assim, os coeficientes β são escritos como $\beta = \mathbf{T}\rho$, e a matriz de covariâncias dos coeficientes β é dada por $\Sigma_\beta = \mathbf{T}\Sigma_\rho\mathbf{T}'$, onde Σ_ρ é a matriz de covariâncias dos coeficientes ρ .

Ravines, Schmidt e Migon (2006) fazem uma detalhada revisão sobre estimação de modelos de defasagens distribuídas seguindo o enfoque bayesiano, e mostram que estes podem ser escritos como casos particulares de funções de transferência. Tais funções são inseridas em modelos dinâmicos, com o objetivo de avaliar o período e a forma de persistência do efeito das variáveis explicativas.

1.3 Modelos Lineares Generalizados

Uma das principais limitações do modelo de regressão linear reside na hipótese de normalidade dos erros, a qual, condicionalmente às regressoras, implica que a resposta tenha distribuição normal. Torna-se, então, necessária a introdução de uma classe alternativa de modelos, que contemplem respostas não necessariamente gaussianas. Uma possível alternativa, então, para contornar esta limitação é a classe dos Modelos Lineares Generalizados (McCullagh & Nelder, 1989), na qual se pressupõe que a resposta pertença à Família Exponencial.

Diz-se que a distribuição da variável aleatória Y pertence à família exponencial de distribuições se sua função de probabilidades - se Y é discreta - ou sua função densidade de probabilidade - se Y é contínua - pode ser escrita na forma

$$f(y | \chi) = \exp\{a(y)b(\chi) + c(\chi) + d(y)\} \quad (1.10)$$

onde y é uma amostra aleatória de tamanho T da variável Y , χ são parâmetros do modelo, $b(\chi)$ e $c(\chi)$ são funções de χ , $d(y)$ é uma função da amostra e $a(y)$ é uma

estatística suficiente para χ , ou seja, $a(y)$ contém toda a informação proveniente dos dados para a estimação do parâmetro χ . Se $a(y) = y$, diz-se que a distribuição está apresentada na forma canônica, e $b(\chi)$ é chamado parâmetro natural da distribuição. Resultados utilizados para determinar expressões para o valor esperado e variância de $a(y)$ podem ser encontrados em Dobson (1990).

Nos Modelos Lineares Generalizados (MLG) é considerada uma combinação linear dos parâmetros β_1, \dots, β_p , que é igual a alguma função do valor esperado μ_t de Y_t , a variável resposta cuja distribuição pertence à Família Exponencial. Tem-se então:

$$g(\mu_t) = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_p X_{pt} = \mathbf{X}_t \boldsymbol{\beta} = \eta_t \quad (1.11)$$

onde

g é uma função monótona e diferenciável, chamada **função de ligação**;

$$\boldsymbol{\beta}' = (\beta_0, \dots, \beta_p)$$

\mathbf{X}_t é vetor de dimensão $t \times (p + 1)$;

Como se pode ver, a estrutura do preditor de um MLG suporta as defasagens distribuídas sujeitas a restrições polinomiais, apresentadas na seção anterior. Ainda assim, uma das limitações dos modelos lineares generalizados reside no fato de que, em diversas situações, tem-se variáveis explicativas que não impactam a resposta média de forma linear. A hipótese de linearidade nos parâmetros destas variáveis costuma ser muito restritiva e, além disso, no caso das variáveis de confundimento, o foco de atenção não está na estimativa do efeito das mesmas, mas apenas na consideração da possível existência de seus efeitos. Uma solução, portanto, é incorporar tais variáveis no modelo adotado de forma não-paramétrica, o que induz a uma técnica que suporte esta estrutura, e que será apresentada no próximo capítulo.

Capítulo 2

Modelos Aditivos Generalizados

Os modelos lineares generalizados assumem uma forma paramétrica linear para representar a relação entre as covariáveis $X_{1,t}, X_{2,t}, \dots, X_{p,t}$ e a resposta Y_t . Hastie & Tibshirani (1990) propuseram uma classe de modelos, os Modelos Aditivos Generalizados (MAG), onde o preditor linear $\eta = \sum \beta_i X_i$ é substituído pelo preditor aditivo $\sum f_i(X_i)$ - uma soma de funções suaves. Os $f_i(\cdot)$'s são funções não específicas, estimadas por procedimentos iterativos. Alguns exemplos de tais funções são as médias móveis, *splines* e *loess* (Hastie e Tibshirani, 1990; Venables e Ripley, 1994). De forma geral, busca-se funções que suavizem o diagrama de dispersão das variáveis X versus Y . A técnica de estimação pode ser aplicada a qualquer modelo de regressão baseado em verossimilhança, muitos dos quais pertencem à classe dos modelos lineares generalizados. Neste trabalho, em particular, as funções não-paramétricas utilizadas para suavização são *splines*.

2.1 Splines

As *splines* são utilizadas para aproximar uma função por uma série de polinômios definidos em sub-intervalos, cujos extremos são denominados *nós*. É usual forçar os polinômios a unirem-se suavemente nos nós.

De acordo com Hastie e Tibshirani (1990), ao se tentar utilizar como suavizador a curva que minimize $\sum (y_t - f(x_t))^2$, o resultado seria uma curva de interpolação nem um pouco suave, e uma forma de contornar este problema é incluir um termo de penalização

que considere a quantidade de curvatura da função obtida. Isto pode ser feito através de uma medida de suavidade como $\int (f^{(m)}(u))^2 du$, em que $f^{(m)}$ denota a derivada de ordem m .

Eilers e Marx (1996) assumem que uma função de suavização f_j de uma covariável x_j , $j = 1, 2, \dots, p$, pode ser aproximada por uma *spline* polinomial de grau $l = m + 1$ definida em uma série de nós igualmente espaçados. Seja o conjunto de números reais ζ_1, \dots, ζ_n em um intervalo $[a, b]$, com $a \leq \zeta_1 < \zeta_2 < \dots < \zeta_{n-1} < \zeta_n \leq b$.

Busca-se a função f que minimize:

$$S(f_j) = \sum_{t=1}^T (y_t - f_j(\zeta))^2 + \lambda \int_a^b (f^{(m)}(u))^2 du. \quad (2.1)$$

O termo $\int_a^b (f^{(m)}(u))^2 du$ mede a rugosidade da função f , assumindo valor nulo para f linear. λ é um parâmetro desconhecido positivo de suavização: quanto maior, mais suave será f . Nos casos extremos, quando $\lambda = 0$ tem-se a curva de interpolação, e um ajuste linear quando $\lambda \rightarrow \infty$. Sejam $a = \zeta_0$ e $b = \zeta_{n+1}$. Para qualquer valor de λ , a solução de (2.1) será uma *spline* de grau l se as seguintes condições forem satisfeitas:

1. f é um polinômio de grau l em cada sub intervalo $(\zeta_0, \zeta_1), (\zeta_1, \zeta_2), (\zeta_2, \zeta_3), \dots, (\zeta_n, \zeta_{n+1})$;
2. cada dois polinômios por partes¹ se unem no ponto ζ_i , de forma que todas as suas primeiras m derivadas são contínuas em cada ponto ζ_i e, conseqüentemente, em todo o intervalo $[\zeta_0, \zeta_{n+1}]$.

Embora diversas configurações sejam possíveis, uma escolha bastante popular são as *splines* cúbicas ($m = 2$), que consistem em uma coleção de polinômios cúbicos com primeira e segunda derivadas contínuas nos nós (penaliza-se a segunda derivada, que está relacionada à aceleração da curva).

Se a segunda e a terceira derivadas nos pontos ζ_0 e ζ_{n+1} forem iguais a zero, a *spline* cúbica no intervalo $[\zeta_0, \zeta_{n+1}]$ será dita *spline* cúbica natural.

¹Do inglês *piecewise polynomial*

Suponha que f é uma spline cúbica natural com n nós internos $\zeta_1 < \dots < \zeta_n$ e dois nós extremos ζ_0 e ζ_{n+1} . Voltando a (2.1) e reescrevendo $S(f)$ para o caso em que $m = 2$ (*spline* cúbica), tem-se:

$$S(f_j) = \sum_{t=1}^T (y_t - f_j(\zeta))^2 + \lambda \int_{\zeta_0}^{\zeta_{n+1}} (f^{(2)}(u))^2 du. \quad (2.2)$$

A função f_j pode ser representada em termos de uma combinação linear de $n + 2$ bases B de seu espaço - que serão apresentadas na próxima seção - da seguinte forma:

$$f_j(x_j) = \sum_{k=0}^{n+1} \phi_{jk} B_{jk}(x_j), \quad (2.3)$$

onde ϕ_{jk} são coeficientes desconhecidos, n é o número de nós interiores e B_{jk} são as bases.

A existência e unicidade de uma *spline* que minimize (2.2) é garantida pelo teorema de Curry e Schoenberg (de Boor, 1978), e por causa deste teorema Shoenberg chamou as funções B_j de *bases splines* ou *B-splines*. O teorema bem como sua prova podem ser consultados em de Boor (1978).

2.2 Representação por *B-splines*

Considere a sequência não decrescente de nós $\zeta_0, \dots, \zeta_{n+1}$. Segundo de Boor (1978), a k -ésima *B-spline* (normalizada) de ordem $l + 1$ para a sequência de nós ζ é dada por:

$$B_{k,l,\zeta}(x_t) = (\zeta_{k+l} - \zeta_k) [\zeta_k, \dots, \zeta_{k+l}] (\zeta - x_t)_+^l, \forall x \in \mathfrak{R} \quad (2.4)$$

em que $(x)_+ = \max(x, 0)$. A notação “ $(\zeta - x)_+^l$ ” significa que a l -ésima diferença dividida² da função $(\zeta - x)_+^l$ das variáveis ζ e x é considerada como função somente de ζ , fixando-se x . O resultado depende do valor de x escolhido em particular, ou seja, o resultado varia de acordo com x , e daí se obtém a função $B_{k,l,\zeta}$ de x . Desta forma, serão obtidos T vetores de tamanho $n + 2$, cada um com no máximo $l + 1$ componentes não nulas. Para facilitar a notação, será escrito B_k quando l e ζ estiverem subentendidos. Algumas propriedades que motivam o uso de *B-splines* são:

²Ver definição no Apêndice A

(i) $B_k = B_{k,l,\zeta}$ tem suporte compacto, ou seja,

$$B_k(x) = 0 \text{ para } x \notin [\zeta_k, \zeta_{k+l+1}].$$

Para $x \notin [\zeta_k, \zeta_{k+l+1}]$, então $g(\zeta) = (\zeta - x)_+^l$ é um polinômio de grau $\leq l$. Segue que apenas $l + 1$ *B-splines* contém um intervalo $[\zeta_j, \zeta_{j+l}]$ em particular em seu suporte, ou seja, de todas as *B-splines* de ordem $l + 1$ para a sequência de nós ζ , apenas as $l + 1$ *B-splines* $B_{j-l}, B_{j-l+1}, \dots, B_j$ são não nulas no intervalo $[\zeta_j, \zeta_{j+l}]$.

(ii) Tem-se que:

$$\sum_j B_j(x) = \sum_{j=r-l}^{s-l} B_j = 1, \quad \forall \zeta_r < x < \zeta_s.$$

(iii) B_k é positiva em seu suporte, ou seja,

$$B_j(x) > 0 \text{ para } \zeta_k < x < \zeta_{k+l}.$$

De certa forma a escolha dos nós controla a suavidade da curva. Por exemplo, se um nó não é repetido, l condições de continuidade são satisfeitas naquele nó. Uma forma eficiente de calcular os coeficientes das *B-splines* é dada através da relação recursiva:

$$B_{k,l+1}(x) = \left[\frac{x - \zeta_k}{\zeta_{k+l} - \zeta_k} \right] B_{k,l}(x) + \left[\frac{\zeta_{k+l+1} - x}{\zeta_{k+l+1} - \zeta_{k+1}} \right] B_{k+1,l}(x), \quad (2.5)$$

$$\text{com } B_{k,0}(x) = \begin{cases} 1, & \zeta_k < x < \zeta_{k+1} \\ 0, & \text{caso contrário.} \end{cases}$$

Assim, tendo as funções $f_j(x_j)$ reescritas no preditor como em (2.3), os coeficientes ϕ_{jk} são obtidos através de regressão linear usual, utilizando as bases $B_{k,l+1}(x)$ como variáveis explicativas.

2.3 Seleção do parâmetro de suavização da spline

Já foi dito que λ é um parâmetro positivo de suavização que está diretamente ligado ao termo que mede a rugosidade da função f , e uma das dificuldades encontradas na estimação desta função f reside na escolha do parâmetro λ .

Supondo que a solução \hat{f}_λ de (2.2) seja uma *spline* cúbica natural, isto significa que \hat{f}_λ é um polinômio cúbico entre nós sucessivos, e que suas primeira e segunda derivadas são contínuas nestes pontos. Ainda, nos extremos a segunda derivada é igual a zero. Sejam \mathbf{B} a matriz $T \times (n + 2)$ contendo as bases e $\phi' = (\phi_0, \dots, \phi_{n+1})$. Pode-se mostrar (de Boor, 1978) que minimizar (2.2) é equivalente a minimizar o critério de mínimos quadrados penalizado (PSL)

$$PSL(f) = (\mathbf{y} - \mathbf{B}\phi)(\mathbf{y} - \mathbf{B}\phi)' + \lambda f' \Omega f, \quad (2.6)$$

onde $\mathbf{y}' = (y_1, \dots, y_T)$ e $f' = (f(\zeta_0), \dots, f(\zeta_{n+1}))$.

A matriz Ω tem estrutura dada por $\Omega = QR^{-1}Q'$, onde Q uma matriz $n \times (n - 2)$ com entradas q_{ij} , $i = 1, \dots, n$ e $j = 2, \dots, n - 1$, dadas por

$$\begin{aligned} q_{j-1,j} &= h_{j-1}^{-1} \\ q_{j,j} &= -h_{j-1}^{-1} - h_j^{-1} \\ q_{j+1,j} &= h_j^{-1} \end{aligned}$$

para $j = 2, \dots, n - 1$ e $q_{ij} = 0$ para $|i - j| \geq 2$.

A matriz simétrica R tem dimensão $(n - 2) \times (n - 2)$ com elementos r_{ij} , para $i, j \in [2, n - 1]$, dados por:

$$\begin{aligned} r_{i,i} &= \frac{1}{3}(h_{i-1} + h_i), \quad i = 2, \dots, n - 1 \\ r_{i,i+1} &= r_{i+1,i} = \frac{1}{6}h_i, \quad i = 2, \dots, n - 2 \end{aligned} \quad (2.7)$$

e $r_{i,j} = 0$ para $|i - j| \geq 2$, em que $h_i = \zeta_{i+1} - \zeta_i$, $i = 1, \dots, n - 1$.

Um método bastante utilizado para selecionar o parâmetro de suavização e que pode ser escrito em função destas matrizes Q e R é a **validação cruzada** (ou “leaving-out-one”) (Wahba, 2000).

Seja $f_\lambda^{[k]}$ uma função estimada com o parâmetro de suavização λ e deixando de fora a k -ésima observação. $f_\lambda^{[k]}$ minimiza

$$\frac{1}{T} \sum_{t=1, t \neq k}^T (y_t - f(\zeta))^2 + \lambda \int_{\zeta_0}^{\zeta_{n+1}} (f^{(2)}(u))^2 du. \quad (2.8)$$

A função de validação cruzada ordinária, $V_0(\lambda)$, é definida como

$$V_0(\lambda) = \frac{1}{T} \sum_{t=1}^T (y_t - f_\lambda^{[k]}(\zeta_k))^2, \quad (2.9)$$

e a estimativa de λ excluindo a k -ésima observação minimiza $V_0(\lambda)$. É fácil verificar que, para um dos ζ , cada $f(\zeta)$ é uma combinação linear de componentes de $y = (y_1, \dots, y_T)'$, e daí existe uma matriz $A(\lambda)$ satisfazendo

$$\begin{pmatrix} f_\lambda(\zeta_1) \\ \vdots \\ f_\lambda(\zeta_n) \end{pmatrix} = A(\lambda)y,$$

onde $A(\lambda) = (I + \lambda QR^{-1}Q')^{-1}$.

O lema da exclusão de uma observação³ apresenta uma identidade matemática útil, que não será provada aqui, mas é dada por:

$$(y_t - f_\lambda^{[k]}(\zeta_k)) \equiv \frac{(y_t - f_\lambda(\zeta_k))}{(1 - a_{kk}(\lambda))}, \quad (2.10)$$

onde $a_{kk}(\lambda)$ é o k -ésimo elemento da diagonal de $A(\lambda)$. Substituindo (2.10) em (2.9), chega-se a uma forma simplificada para V_0 :

$$\frac{1}{T} \sum_{t=1}^T (y_t - f_\lambda^{[k]}(\zeta_k))^2 \equiv V_0(\lambda) \equiv \frac{1}{T} \sum_{t=1}^T \frac{(y_t - f_\lambda(\zeta_k))^2}{(1 - a_{kk}(\lambda))^2}. \quad (2.11)$$

O termo a direita é mais eficiente computacionalmente que o termo a esquerda. No entanto, a **validação cruzada generalizada**, que será descrita a seguir, é ainda mais eficiente.

Validação Cruzada Generalizada

A validação cruzada generalizada (VCG) é um método de escolha do parâmetro de suavização baseado ainda na idéia de excluir uma observação, mas tem duas vantagens:

³*Leaving-Out-One Lemma* (Craven e Wahba, 1979)

1. é de mais fácil implementação;
2. possui algumas propriedades teóricas que seriam impossíveis de provar para a validação cruzada ordinária.

A função de VCG, $V(\lambda)$, é obtida substituindo $a_{kk}(\lambda)$ em $V_0(\lambda)$ por $\bar{a}_{kk}(\lambda) = \frac{1}{T} \sum_{t=1}^T a_{tt}(\lambda) = \frac{1}{T} \text{tr} A(\lambda)$. A função de VCG é definida por:

$$V(\lambda) = \frac{1}{T} \sum_{t=1}^T \frac{(y_t - f_\lambda(\zeta))^2}{(1 - a_{kk}(\lambda))^2} = \frac{\frac{1}{T} \|(I - A(\lambda))\|^2}{\left[\frac{1}{T} \text{tr}(I - A(\lambda))\right]^2} \quad (2.12)$$

2.4 Inferência sobre Modelos Aditivos Generalizados

Suponha um MAG da seguinte forma:

$$\begin{aligned} Y_t &\sim FE(\mu_t) \\ g(\mu_t) = \eta_t &= \gamma + \varphi_1 C_1 + \varphi_2 C_2 + \cdots + \varphi_l C_l \\ &+ \beta_0 X_t + \beta_1 X_{t-1} + \cdots + \beta_q X_{t-q} \\ &+ s(w_1) + s(w_2) + \cdots + s(w_p), \quad t = 1, \dots, T, \end{aligned} \quad (2.13)$$

com β_j , $j = 1, \dots, q$ aproximado por um polinômio de grau d nas defasagens, com coeficientes ρ_0, \dots, ρ_d como em (1.4), $s(w_j)$ *splines* cúbicas e, portanto parametrizadas por (ϕ'_j, λ_j) , $j = 1, \dots, p$ e C_1, \dots, C_l variáveis de controle inseridas de forma paramétrica. Assim, o número total de regressoras é $m = l + (q + 1) + p$ e o vetor de parâmetros a estimar é dado por

$$\theta = (\gamma, \varphi', \rho_0, \dots, \rho_d, \phi'_1, \dots, \phi'_p, \lambda_1, \dots, \lambda_p) \quad (2.14)$$

e a função de verossimilhança é dada por:

$$L(y_t, \chi_t) = \prod_{t=1}^T p(y_t, \chi_t) \propto \prod_{t=1}^T e^{\{b(\chi_t)a(y_t) + c(\chi_t) + d(y_t)\}}, \quad (2.15)$$

em que χ_t é escrito em termos de θ , como se pode ver em (2.13).

2.4.1 Estimação sob o enfoque clássico

Seja $\mathbf{X} = (X_1, \dots, X_P)$ um vetor aleatório de dimensão P , onde $E[Y_t | \mathbf{X}] = \mu_t$ e $g(\mu_t)$ dado como em (2.13). Seja ainda $\Psi = \gamma + \varphi_1 C_1 + \varphi_2 C_2 + \dots + \varphi_l C_l + \beta_0 X_t + \beta_1 X_{t-1} + \dots + \beta_q X_{t-q}$.

Supondo $s_1(X_1), \dots, s_{j-1}(X_{j-1}), s_{j+1}(X_{j+1}), \dots, s_p(X_p)$ conhecidas, o resíduo parcial pode ser definido como

$$R_j = Y - g^{-1} \left(\Psi + \sum_{k \neq j} s_k(X_k) \right), \quad (2.16)$$

e assim tem-se que $E[R_j | X_j] = s_j(X_j)$.

Observa-se que $s_k(X_k)$ não é conhecido, mas o resíduo parcial fornece uma boa alternativa para a estimação de $\{\hat{s}_j(X_j) | \hat{s}_i(X_i)\}$, $i \neq j$. Este procedimento iterativo é conhecido como algoritmo *backfitting* (Friedman e Stuetzle, 1981), e será brevemente apresentado a seguir.

Algoritmo *backfitting*

O algoritmo *backfitting* ajusta as funções f_j , uma de cada vez, tomando os resíduos (2.16) e suavizando-os contra x_j . Este processo é repetido até a convergência e os termos lineares do modelo (incluindo qualquer termo linear do suavizador) são ajustados por mínimos quadrados. O grau de suavidade será escolhido automaticamente por validação cruzada (Venables e Ripley, 1994). Resumidamente, a estimação clássica semi-paramétrica é feita do seguinte modo:

Primeiramente a parte paramétrica do modelo é estimada por máxima verossimilhança. Dado o preditor linear, é calculado um resíduo parcial associado ao ajuste paramétrico e este resíduo parcial é utilizado como resposta no método da validação cruzada. Voltando à equação (2.12), y_k será o resíduo parcial.

O resultado f_λ que minimiza a função de validação cruzada (2.12) é introduzido como um termo constante na estimação paramétrica e o algoritmo é iterado até que a seqüência de valores da verossimilhança convirja para um critério de parada estabelecido.

2.4.2 Estimação sob o enfoque bayesiano

Sob o enfoque bayesiano, a informação inicial que se tem sobre um parâmetro θ pode ser resumida em $p(\theta)$, a sua *distribuição a priori*.

Admite-se ainda, que com exceção da informação fornecida a priori, toda informação relevante é trazida pela verossimilhança. O que se vai fazer então é combinar a informação inicial à verossimilhança através do **Teorema de Bayes** e deste modo será possível obter a *densidade a posteriori*, que é a densidade atualizada dos parâmetros de interesse. Uma formalização do método de inferência seguindo o enfoque bayesiano será dada seguindo Migon e Gamerman (1999).

Sejam $p(y | \theta)$ e $p(\theta)$ as densidades de $(y | \theta)$ e θ , respectivamente. Tem-se então que:

$$p(\theta | y) = \frac{p(y | \theta)p(\theta)}{p(y)}, \quad (2.17)$$

onde $p(y)$ é a *distribuição preditiva* (ou marginal de y). Antes de observar Y ela é útil para checar a adequação da priori através das predições que ela fornece para Y . Verifica-se que, como ela não depende de θ , funciona como uma constante na determinação de $p(\theta | y)$. Assim, a expressão de $p(\theta | y)$ pode ser mais uma vez reescrita como

$$p(\theta | y) \propto p(y | \theta)p(\theta), \quad (2.18)$$

onde o símbolo \propto denota proporcionalidade. Para recuperar $p(y)$, basta reescrever (2.18) como

$$p(\theta | y) = kp(y | \theta)p(\theta), \quad (2.19)$$

onde $k^{-1} = \int_{\Theta} p(y | \theta)p(\theta)d\theta = E_{\theta}[p(y | \theta)]$.

Em alguns casos a distribuição a posteriori não terá forma fechada, e deverão ser utilizados métodos computacionais para obter amostras desta posteriori. Neste trabalho serão utilizados métodos de Monte Carlo via Cadeias de Markov (MCMC), que são baseados na construção de uma cadeia de Markov que tenha como distribuição limite a densidade da qual se deseja gerar valores. Desta forma, métodos MCMC serão usados para construir uma cadeia que tenha como distribuição estacionária a distribuição a

posteriori de interesse. Depois de constatada a convergência, os valores desta cadeia serão utilizados como amostra da distribuição a posteriori.

Os algoritmos mais utilizados para realizar estas simulações são o *amostrador de Gibbs* e o algoritmo de *Metropolis-Hastings* (Geman, 1997). O amostrador de Gibbs é útil quando não se conhece a densidade conjunta dos parâmetros, mas é possível gerar valores da distribuição condicional completa dos mesmos. O amostrador de Gibbs procede do seguinte modo:

- 1 são dados valores iniciais para os parâmetros de interesse;
- 2 a cada iteração, geram-se valores da distribuição condicional completa, sempre utilizando o último valor atualizado de cada um dos parâmetros;
- 3 repete-se o passo anterior até convergência da cadeia.

Nos casos em que não se consegue gerar valores destas condicionais completas, podem ser inseridos passos de Metropolis-Hastings no passo 2 do amostrador de Gibbs, da seguinte forma:

- 1 gera-se um valor θ^* de uma densidade proposta $q(\theta, \theta^*)$, com θ representando o valor corrente do parâmetro e θ^* o valor proposto gerado de $q(., .)$;
- 2 aceita-se este valor θ^* com probabilidade

$$\min \left\{ \frac{\pi(\theta^*)/q(\theta, \theta^*)}{\pi(\theta)/q(\theta^*, \theta)}, 1 \right\},$$

onde π é a densidade condicional completa de θ .

Seguindo o enfoque bayesiano, o vetor paramétrico dado em (2.14) sofrerá uma pequena alteração, visto que será assumida uma priori de suavização para os coeficientes ϕ em (2.3). Neste caso, serão utilizados passeios aleatórios de segunda ordem como prioris para estes, ou seja, $\phi_{j,k}$ será dado por $2\phi_{j,k-1} - \phi_{j,k-2} + u_{j,k}$, com $u_{j,k} \sim N(0, \tau_j)$, e assim, os parâmetros que devem ser estimados são:

$$\theta = (\gamma, \varphi, \rho_0, \dots, \rho_d, \phi'_1, \dots, \phi'_p, \lambda_1, \dots, \lambda_p, \tau_1^2, \dots, \tau_p^2).$$

Primeiramente toda a informação inicial sobre os parâmetros será resumida em suas densidades a priori:

$$\begin{aligned}
p(\gamma), p(\varphi) &\propto \text{constante} \\
p(\rho_k) &\propto \text{constante}, \quad k = 1, \dots, d \\
\phi_{j,k} &= 2\phi_{j,k-1} - \phi_{j,k-2} + u_{j,k}, \quad u_{j,k} \sim N(0, \tau_j^2), \quad k \geq 2 \\
\phi_{j,0} &\sim N(0, \sigma_{\phi_{j,0}}^2) \\
\phi_{j,1} &\sim N(0, \sigma_{\phi_{j,1}}^2) \\
\tau_j^2 &\sim GI(a_j, b_j),
\end{aligned}$$

e o parâmetro de suavização λ_j será dado por $1/\tau_j^2$. Biller e Fahrmeir (1997) afirmam que em uma aproximação bayesiana empírica τ_j^2 pode ser considerada uma constante desconhecida que pode ser estimada a partir dos dados, e a estimação de τ_j^2 fornece, então, uma escolha do parâmetro de suavização guiada pelos dados. Dadas as prioris, estas serão combinadas à verossimilhança via teorema de Bayes dando origem à densidade atualizada dos parâmetros de interesse

$$\begin{aligned}
p(\theta | y) &\propto L(y, \Theta)p(\gamma)p(\varphi) \prod_{j=0}^d p(\rho_j) \prod_{j=1}^p \prod_{k=3}^n p(\phi_{j,k} | \phi_{j,k-1}, \phi_{j,k-2}, \tau_j^2) p(\phi_{j,0}) \\
&\times p(\phi_{j,1}) p(\tau_j^2).
\end{aligned} \tag{2.20}$$

Se, por exemplo, é assumido que a variável resposta segue uma distribuição *gama*, tem-se que:

$$y_t \sim Ga(\alpha, \beta_t) \Rightarrow f(y_t | \alpha, \beta_t) = \frac{\beta_t^\alpha}{\Gamma(\alpha)} y_t^{\alpha-1} e^{-\beta_t y_t}$$

e daí, sua função de verossimilhança será dada por:

$$L(\mathbf{y} | \alpha, \beta) = f(\mathbf{y} | \alpha, \beta) = \prod_{t=1}^T \frac{\beta_t^\alpha}{\Gamma(\alpha)} y_t^{\alpha-1} e^{-\beta_t y_t}. \tag{2.21}$$

Ainda,

$$g(\mu_t) = g(\alpha/\beta_t) = \log(\alpha/\beta_t) = \gamma + C'\varphi + \sum_{j=1}^p \sum_{k=0}^{n+1} \phi_{j,k} B_{j,k} + \sum_{j=0}^d \rho_j Z_{j,t}$$

e daí,

$$\begin{aligned} \alpha/\beta_t &= \exp\left\{\gamma + C'\varphi + \sum_{j=1}^p \sum_{k=0}^{n+1} \phi_{j,k} B_{j,k} + \sum_{j=0}^d \rho_j Z_{j,t}\right\} \\ \alpha &= \beta_t \exp\left\{\gamma + C'\varphi + \sum_{j=1}^p \sum_{k=0}^{n+1} \phi_{j,k} B_{j,k} + \sum_{j=0}^d \rho_j Z_{j,t}\right\} \end{aligned} \quad (2.22)$$

Substituindo (2.22) em (2.21) e reescrevendo a estrutura da distribuição a posteriori dada em (2.20) chega-se a

$$\begin{aligned} p(\theta | \mathbf{y}) &\propto L(\mathbf{y})p(\gamma)p(\varphi) \prod_{j=0}^d p(\rho_j) \prod_{j=1}^p \prod_{k=2}^{n+1} p(\phi_{j,k} | \phi_{j,k-1}, \phi_{j,k-2}, \tau_j^2) \prod_{j=1}^p p(\phi_{j,0})p(\phi_{j,1})p(\tau_j^2) \\ &\propto \prod_{t=1}^T \frac{\beta_t^\alpha}{\Gamma(\alpha)} y_t^{\alpha-1} e^{-\beta_t y_t} \prod_{j=1}^p (\tau_j^2)^{-a_j-1} \exp\left\{\frac{-b_j}{\tau_j^2}\right\} \\ &\times \prod_{j=1}^p \prod_{k=2}^{n+1} \frac{1}{(\tau_j^2)^{1/2}} \exp\left\{\frac{-1}{2\tau_j^2}(\phi_{j,k} - 2\phi_{j,k-1} + \phi_{j,k-2})^2\right\} \prod_{j=1}^p \prod_{k=0}^1 \frac{1}{(\tau_j^2)^{1/2}} \exp\left\{\frac{-1}{2\tau_j^2}(\phi_{j,k})^2\right\}. \end{aligned} \quad (2.23)$$

A condicional completa de τ_j^2 será dada por:

$$\begin{aligned} p(\tau_j^2 | \cdot) &\propto (\tau_j^2)^{-a_j-1} \exp\left\{\frac{-b_j}{\tau_j^2}\right\} \frac{1}{(\tau_j^2)^{n/2}} \exp\left\{\frac{-1}{2\tau_j^2} \sum_{k=2}^{n+1} (\phi_{j,k} - 2\phi_{j,k-1} + \phi_{j,k-2})^2\right\} \\ &\propto (\tau_j^2)^{-a_j-n/2-1} \exp\left\{\frac{-1}{\tau_j^2} \left[b_j + \frac{1}{2} \sum_{k=2}^{n+1} (\phi_{j,k} - 2\phi_{j,k-1} + \phi_{j,k-2})^2\right]\right\} \end{aligned}$$

Ou seja, $\tau_j^2 \sim GI\left(a_j + \frac{n}{2}, b_j + \frac{1}{2} \sum_{k=2}^{n+1} (\phi_{j,k} - 2\phi_{j,k-1} + \phi_{j,k-2})^2\right)$, e as distribuições dos demais parâmetros não terão forma conhecida sendo necessário, neste caso, utilizar passos de Metropolis-Hastings inseridos no algoritmo de Gibbs para amostrar seus valores.

O esquema de atualização utilizará densidades propostas para o algoritmo Metropolis-Hastings como sugerido por Gamerman (1997). De acordo com este método, um novo

com $U_t = (W_{1,t}, W_{2,t}, \dots, W_{p,t})$. O esquema de amostragem pode ser resumido da seguinte forma:

- **Início do algoritmo**

Calcular a moda a posteriori para $\gamma, \varphi, \rho_0, \dots, \rho_d, \phi'_1, \dots, \phi'_p$ dado um parâmetro de suavização fixo $\lambda_j = 1/\tau_j^2$. A estimativa é utilizada como estado corrente da cadeia $\gamma^c, \varphi^c, \rho_0^c, \dots, \rho_d^c, (\phi'_1)^c, \dots, (\phi'_p)^c, (\tau_i^2)^c$.

- **Atualização de $\Delta = (\gamma, \varphi, \rho_0, \dots, \rho_d)$**

Amostrar um novo valor de cada componente de Δ^p de uma densidade proposta $q(\Delta^c, \Delta^p) = N(m_\Delta, P_\Delta)$, normal multivariada, onde m_Δ e P_Δ são dados como em (2.25). Δ^p é aceito com probabilidade

$$\alpha = \frac{L(y, \dots, \Delta^p) q(\Delta^p, \Delta^c)}{L(y, \dots, \Delta^c) q(\Delta^c, \Delta^p)}.$$

- **Atualização de ϕ'_j**

Um valor proposto para o bloco de parâmetros ϕ'_j é amostrado, para $j = 1, \dots, p$ da densidade normal multivariada $q(\Delta^c, \Delta^p)$. O valor é aceito com probabilidade

$$\alpha = \frac{L(y, \dots, (\phi'_j)^p, (\tau_j^2)^c, \dots, \delta^c) p((\phi'_j)^p | (\tau^2)^c) q((\phi'_j)^p, (\phi'_j)^c)}{L(y, \dots, (\phi'_j)^c, (\tau_j^2)^c, \dots, \delta^c) p((\phi'_j)^c | (\tau^2)^c) q((\phi'_j)^c, (\phi'_j)^p)}.$$

- **Atualização de τ_j^2 e λ_j**

Atualizar a variância amostrando da condicional completa Gama Inversa e fazer $\lambda_j = 1/\tau_j^2$.

Este algoritmo será utilizado para estimar os modelos que serão propostos no próximo capítulo.

Capítulo 3

Modelos propostos

Será dada neste capítulo uma breve descrição dos modelos propostos, bem como do banco de dados utilizado no estudo, e visto que serão utilizadas diferentes estruturas para analisar os dados, serão apresentados também os critérios adotados para compará-las.

Seguindo o enfoque clássico todos os modelos tratados foram ajustados usando o software livre \mathbf{R}^1 , e sob o enfoque bayesiano os mesmos modelos foram ajustados usando o software livre **BayesX**², utilizando 12000 iterações, com período de aquecimento de 2000 iterações, guardadas com saltos de tamanho 10, para evitar autocorrelações altas entre os valores amostrados.

Como já mencionado, os modelos apresentados neste trabalho seguem a seguinte estrutura:

$$\begin{aligned} Y_t &\sim FE(\mu_t) \\ g(\mu_t) = \eta_t &= \gamma + C'\varphi + s(w_1) + \cdots + s(w_p) \\ &+ \beta_0 X_t + \beta_1 X_{t-1} + \cdots + \beta_q X_{t-q}, \quad t = 1, \dots, T, \end{aligned} \tag{3.1}$$

onde FE representa uma distribuição pertencente à família exponencial, γ é um nível comum para todo tempo t , C são variáveis de controle para as quais se assume forma paramétrica, w_1, \dots, w_p são p variáveis suavizadas e X é uma variável cujo impacto cumulativo se deseja avaliar e cujos coeficientes sofrerão restrições polinomiais. Para

¹Disponível para *download* em <http://www.r-project.org/>.

²Disponível para *download* em <http://www.stat.uni-muenchen.de/~lang/bayesx>

resolver (3.1), as funções $s(w)$ são reescritas como em (2.3):

$$\begin{aligned}
 Y_t &\sim FE(\mu_t) \\
 g(\mu_t) = \eta_t &= \gamma + C'\varphi + \sum_{j=1}^p \sum_{k=0}^{n+1} \phi_{j,k} B_{j,k} + \\
 &+ \beta_0 X_t + \beta_1 X_{t-1} + \cdots + \beta_q X_{t-q}, \quad t = 1, \dots, T.
 \end{aligned} \tag{3.2}$$

Os β 's serão reescritos como em (1.4) e o modelo será estimado utilizando as transformações das variáveis X como em (1.9):

$$\begin{aligned}
 Y_t &\sim FE(\mu_t) \\
 g(\mu_t) = \eta_t &= \gamma + C'\varphi + \sum_{j=1}^p \sum_{k=0}^{n+1} \phi_{j,k} B_{j,k} + \\
 &+ \rho_0 Z_{0,t} + \rho_1 Z_{1,t} + \rho_2 Z_{2,t} + \cdots + \rho_d Z_{d,t}, \quad t = 1, \dots, T.
 \end{aligned} \tag{3.3}$$

Ou seja: p variáveis X entrarão de forma suave no modelo, n é o número de nós interiores da *spline* e Z_j , $j = 0, \dots, d$, são as transformações das variáveis X , obtidas ao se aproximar os coeficientes β por um polinômio de grau d nas defasagens.

3.1 Descrição do banco de dados

Os modelos propostos foram ajustados a um conjunto de dados de contagens de óbitos em São Paulo, com observações diárias feitas no período de 01 de janeiro de 1994 a 31 de dezembro de 1997. O banco de dados utilizado foi construído pela equipe do Laboratório de Poluição Atmosférica Experimental da Faculdade de Medicina da USP, e está disponível na página pessoal do professor Julio Singer³. Num primeiro estudo foram considerados como resposta apenas os desfechos relativos a óbitos de menores de 5 anos por doenças respiratórias e a resposta esperada é explicada em função de X_t : material particulado (PM_{10}), monóxido de carbono (CO) ou dióxido de enxofre (SO_2). A série de óbitos de menores de 5 anos por doenças respiratórias é apresentada na figura 3.1. Outras

³Dados disponíveis para *download* em <http://www.ime.usp.br/~jmsinger/>

variáveis consideradas foram temperatura mínima do dia ($^{\circ}$ Celsius) e umidade relativa média do dia (%) e, de forma a controlar outros fatores de confundimento, trabalhou-se com o número de óbitos de crianças menores de 5 anos por causas não respiratórias. As séries de temperatura mínima e umidade média no período de estudo são apresentadas nas figuras 3.2 e 3.3. Percebe-se claramente através da figura 3.2 o comportamento sazonal da temperatura.

Uma segunda aplicação foi feita considerando como resposta observações de monóxido de carbono (CO), e as variáveis envolvidas no preditor são temperatura e umidade.

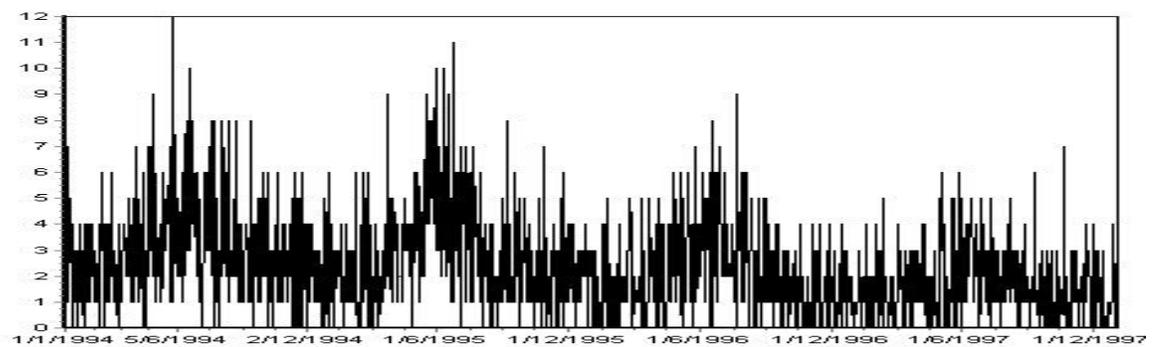


Figura 3.1: Número observado de óbitos por doenças respiratórias em menores de 5 anos no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.

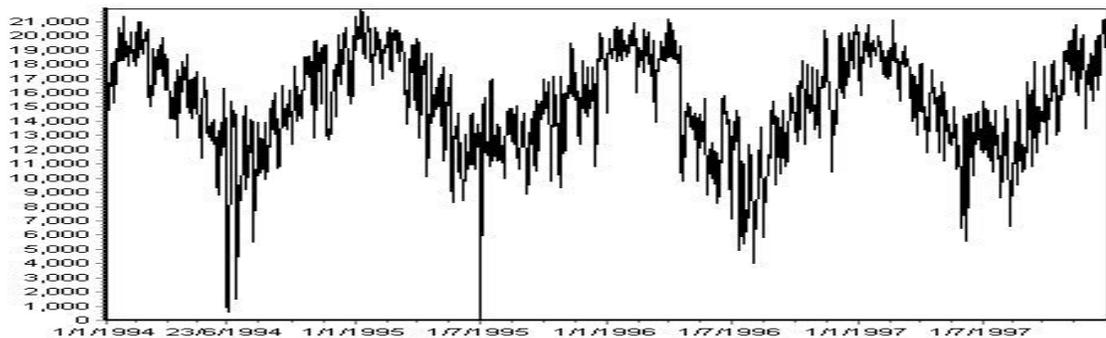


Figura 3.2: Temperatura mínima observada no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.

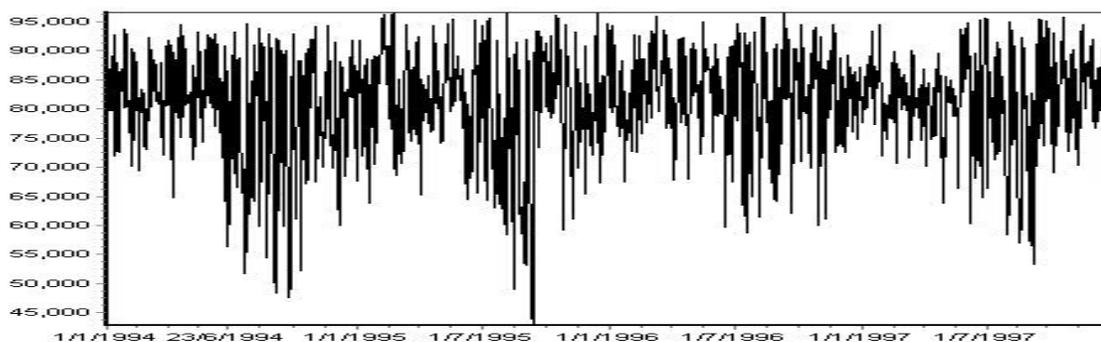


Figura 3.3: Umidade média observada no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.

3.2 Poluentes estudados

As séries dos poluentes consistem em observações diárias feitas em 13 estações monitoradoras. Para o preenchimento de dados omissos foi utilizado o método implementado por Junger (2002), e em uma primeira etapa, cada poluente foi preenchido utilizando a informação das diversas estações monitoradoras para aquele poluente. Posteriormente, foram identificados quais poluentes tinham as maiores correlações, e um novo preenchimento foi feito utilizando estas séries correlacionadas, visto que alguns períodos continuavam incompletos. Após este novo preenchimento, nas séries onde ainda existiam períodos sem observações, foram calculadas médias móveis de 20 dias. A série final de cada poluente utilizada nos modelos é a média das séries preenchidas que, a princípio tinham menos de 50% de dados incompletos, mais as séries com mais de 50% de dados incompletos sem estarem preenchidas.

Em seguida é apresentada uma tabela que contém as 13 estações levadas em consideração no estudo e o “x” indica o local onde cada poluente é medido. Os padrões de qualidade do ar para o Estado de São Paulo (decreto estadual nº 8468 de 08/09/76) são:

$$PM_{10} : 240\mu g/m^3;$$

$$CO : 40000\mu g/m^3 \text{ ou } 35ppm;$$

$$SO_2 : 365\mu g/m^3.$$

Segundo a Companhia de Tecnologia de Saneamento Ambiental - CETESB⁴, estes limites não devem ser excedidos mais de uma vez ao ano.

Tabela 3.1: Estações monitoradoras

Nº da estação	Localização	PM_{10}	SO_2	CO
1	Parque D. Pedro	x	x	x
2	Santana	x	x	
3	Moóca	x	x	x
4	Cambuci	x	x	
5	Ibirapuera	x	x	x
6	N. Senhora do Ó	x	x	
8	Congonhas	x	x	x
9	Lapa	x	x	x
10	Cerqueira César	x	x	x
11	Penha	x	x	
12	Centro	x	x	x
16	Santo Amaro	x	x	x
21	São Miguel Paulista	x	x	

⁴<http://www.cetesb.sp.gov.br/>

3.2.1 Monóxido de carbono - CO

O monóxido de carbono⁵ tem origem, principalmente, na combustão do petróleo e do carvão. O principal dano à saúde causado pelo monóxido de carbono se dá quando este se une à hemoglobina em lugar do oxigênio (formando a carboxiemoglobina), pois, em altas concentrações, prejudica a oxigenação do organismo, causando náuseas, dor de cabeça, diminuição dos reflexos e da acuidade visual. Altos índices do poluente em áreas de fluxo intenso de veículos têm sido apontados como causa adicional de acidentes de trânsito, o que torna o monóxido de carbono um dos principais alvos dos rodízios para controle da poluição já realizados em São Paulo. A série do *CO* no período de estudo é apresentada na figura 3.4.

3.2.2 Material particulado - PM_{10}

O material particulado, presente na fumaça e na poeira, contém pequenas partículas que, quando aspiradas, podem instalar-se no pulmão, provocando danos ao sistema respiratório. Algumas dessas partículas podem conter elementos cancerígenos. Em grande quantidade, o material particulado provoca o agravamento de alergias, asma, bronquite crônica, reduz a resistência às infecções e irritam olhos e garganta. As poeiras em suspensão também potencializam os efeitos dos gases presentes no ar.

A massa de partículas com diâmetro menor que $10\mu m$ (PM_{10}) afeta as pessoas de forma contínua, e contribuem para esta ação um grande número de substâncias químicas, incluindo metais que contenham elementos de transição, sílica, hidrocarbonetos poliaromáticos e agentes biológicos como endotoxinas e glicanos. A série do PM_{10} no período de estudo é apresentada na figura 3.5.

3.2.3 Dióxido de enxofre - SO_2

Um dos poluentes mais comuns e um dos componentes da chuva ácida, o Dióxido de Enxofre é um gás tóxico, incolor em condições normais de temperatura, de sabor ácido,

⁵As informações sobre os efeitos de poluentes foram obtidas via internet, nos endereços [29], [30], [31], [33], [34] e [36].

odor sufocante, e resulta não só da combustão dos derivados do petróleo como a gasolina e o óleo diesel, mas também da atividade dos vulcões e da decomposição de matéria orgânica.

Na indústria, o Dióxido de Enxofre é lançado na atmosfera na queima do petróleo, constituindo cerca de 95% dos compostos de enxofre resultantes da combustão do petróleo. As principais fontes emissoras em refinarias são: caldeiras, unidades de craqueamento e regeneração, e operações de tratamento.

Em baixas concentrações, a tosse é o sintoma mais comum. O contato com a pele provoca irritação, devido à formação de ácido sulfuroso, com o suor.

O Dióxido de Enxofre provoca problemas no sistema respiratório e é causa de bronquites e distúrbios graves, como o enfisema pulmonar. No ar, o dióxido de enxofre pode ser transformado em trióxido de enxofre, que, para as vias respiratórias, é ainda mais irritante que o primeiro. Os vegetais são sensíveis aos óxidos de enxofre: suas folhas amarelecem e, sob altas concentrações de óxido, eles chegam a morrer. A série do SO_2 no período de estudo é apresentada na figura 3.6.

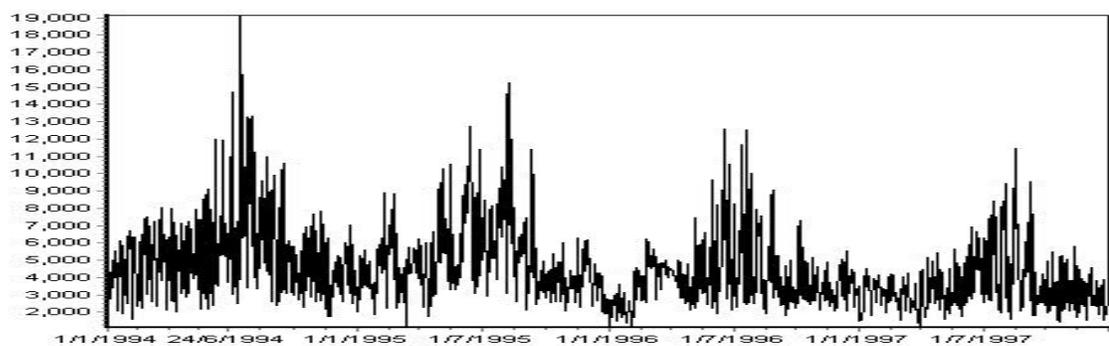


Figura 3.4: Nível observado de CO no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.

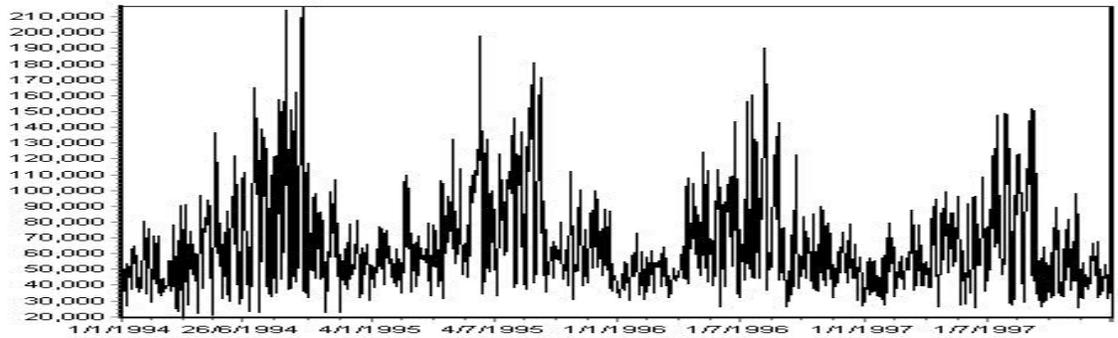


Figura 3.5: Nível observado de PM_{10} no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.

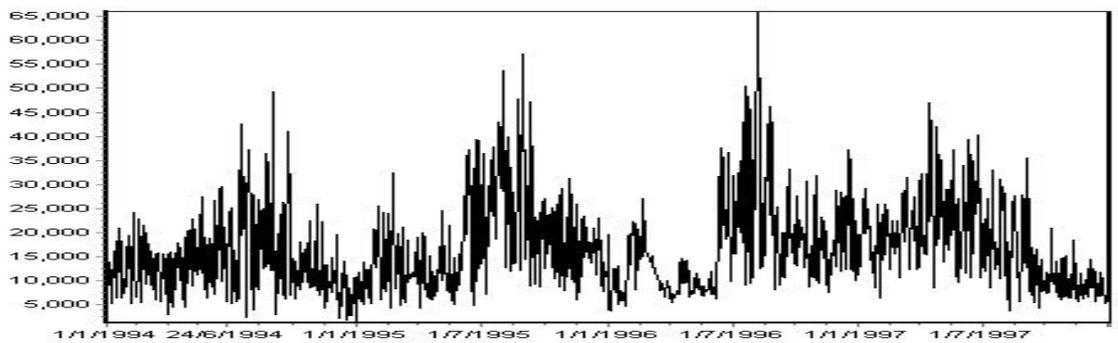


Figura 3.6: Nível observado de SO_2 no período de 01/01/1994 a 31/12/1997, na cidade de São Paulo.

Em seguida é apresentada uma tabela com estatísticas descritivas dos poluentes estudados.

Tabela 3.2: Estatísticas descritivas dos poluentes PM_{10} , CO e SO_2 .

Poluente	média	mediana	desvio padrão	mínimo	máximo	percentil 10	percentil 90
CO	4.52	4.06	2.12	1.12	19.14	2.40	7.30
PM_{10}	65.84	57.26	29.44	19.26	216.56	37.77	106.40
SO_2	16.70	14.70	8.97	1.14	66.03	7.22	28.40

3.3 Critérios de comparação de modelos

Foram utilizados neste trabalho dois diferentes métodos para comparar os modelos estimados: ao se adotar a abordagem clássica para o ajuste, o critério utilizado foi o Critério de Informação de Akaike (AIC). Já para a abordagem bayesiana, utilizou-se o Critério de Informação da Deviance (DIC). Nas seções a seguir serão dadas breves descrições dos critérios.

3.3.1 Critério de Informação de Akaike - AIC

O critério de informação de Akaike - AIC (Akaike, 1974) é uma estatística frequentemente utilizada para a escolha da especificação ótima de uma equação de regressão no caso de alternativas não aninhadas. A melhor especificação é a que produz o menor valor do critério. Por exemplo, o número de defasagens a ser incluídas numa equação com defasagens distribuídas pode ser indicado pela seleção que produz o menor valor do critério de Akaike. O AIC é definido como:

$$AIC = -2 \arg \text{máximo da verossimilhança} + 2p, \quad (3.4)$$

onde p é o número de parâmetros do modelo. Por exemplo, no caso de $y_t \sim Poisson(\mu)$, a log verossimilhança é dada por

$$L(\mu, \mathbf{y}) = \sum_{t=1}^T y_t \log \mu - T\mu + \text{constante}. \quad (3.5)$$

Maximizando (3.5) com relação a μ , tem-se que

$$AIC = -2(\hat{\mu} - p)$$

3.3.2 Critério de Informação da Deviance - DIC

O critério de informação da Deviance (DIC) (Spiegelhalter, Best, Carlin e van der Linde, 2002) pode ser visto como uma generalização do AIC. Ele faz uma comparação entre os logaritmos das verossimilhanças dos modelos saturado e incompleto, procurando estimar se existe diferença significativa entre eles.

Sejam θ^* a média a posteriori de θ no modelo saturado com n parâmetros, θ a média a posteriori no modelo incompleto com $p < T$ parâmetros, $l(\theta^*; y)$ o logaritmo da função de verossimilhança do modelo saturado e $l(\theta; y)$ o logaritmo da função de verossimilhança do modelo incompleto. A função *Deviance* pode, então, ser escrita como:

$$\overline{D(\theta)} = 2[l(\theta^*; y) - l(\theta; y)] \quad (3.6)$$

O número efetivo de parâmetros pD é dado por

$$pD = \overline{D(\theta)} - D(\bar{\theta}), \quad (3.7)$$

e o critério de informação da *deviance*, DIC , é então definido como

$$DIC = \overline{D(\theta)} + pD. \quad (3.8)$$

E assim novamente será considerado o melhor modelo aquele que apresentar menor valor do critério.

3.4 Modelos ajustados

3.4.1 Aplicação utilizando resposta discreta: Mortalidade em São Paulo

Nesta aplicação será modelado o efeito temporal de poluentes sobre óbitos de crianças com idades inferiores a 5 anos, para isto, pressupõe-se um modelo $Poisson(\mu_t)$ para a contagem de óbitos. Escrevendo o modelo como a estrutura dada em (3.3), tem-se que:

$$\begin{aligned} Obitos_t &\sim Poisson(\mu_t) \\ \log(\mu_t) &= \gamma + s(temperatura) + s(umidade) + \varphi NRES5 + \beta_0 POLUENTE_t \\ &+ \beta_1 POLUENTE_{t-1} + \dots + \beta_q POLUENTE_{t-q}. \end{aligned} \quad (3.9)$$

Inicialmente os modelos foram ajustados utilizando apenas as variáveis temperatura e umidade, para diversas combinações de defasagens das mesmas. O melhor ajuste foi obtido considerando-se suavizações da temperatura mínima defasada em dois dias e da umidade média com defasagem de um dia. Para evitar problemas de escala e possibilitar a comparação dos efeitos estimados os poluentes foram padronizados.

Posteriormente foi ajustada uma estrutura semi-paramétrica em que as variáveis temperatura mínima (*tempmin*) e umidade média (*umidmed*) são tratadas de forma não-linear e as demais regressoras (número de óbitos por causas não respiratórias (NRES5) e o poluente estudado) são tratadas de forma linear, paramétrica.

Em uma primeira avaliação considerou-se o modelo utilizando como regressora primeiramente a média dos níveis de cada um dos poluentes no dia corrente e nos dois dias anteriores. Esta é uma forma de avaliar se o efeito deste poluente persiste por mais de um dia. Reescrevendo o modelo tem-se que:

$$\begin{aligned} Obitos_t &\sim Poisson(\mu_t) \\ \log(\mu_t) &= \gamma + \varphi NRES5 + \beta MÉDIA POLUENTE_{0,1,2} \\ &+ s(tempmin) + s(umidmed), \end{aligned} \quad (3.10)$$

onde

$$MÉDIA POLUENTE_{0,1,2} = \frac{1}{3}(POLUENTE_t + POLUENTE_{t-1} + POLUENTE_{t-2}).$$

Em seguida é apresentada a tabela com o resumo dos resultados da estimação do modelo de média móvel de 3 dias, para os poluentes PM_{10} , CO e SO_2 , onde m.p. é média a posteriori, I. Conf. é o Intervalo de Confiança, I. Cred. é o Intervalo de Credibilidade, e pD é o número efetivo de parâmetros. Estes se apresentaram significativos em todos os casos.

Tabela 3.3: Resultados da estimação dos modelos com média das defasagens 0,1,2 para os poluentes CO , PM_{10} e SO_2 .

	Enfoque clássico			Enfoque bayesiano			
	coeficiente	I.Conf.	AIC	m.p.	I.Cred.	pD	DIC
CO	0.19504	0.14712 ; 0.24296	5374.87	0.19860	0.15178 ; 0.24773	14.23	1993.04
PM_{10}	0.13018	0.07937 ; 0.18098	5410.57	0.13311	0.08344 ; 0.18498	13.66	2028.80
SO_2	-0.05670	-0.10164 ; -0.01176	5428.52	-0.05796	-0.10387 ; -0.01529	13.58	2047.25

Deve-se observar que tanto os coeficientes estimados quanto os intervalos de confiança e credibilidade se apresentaram similares. Observa-se também que, ao utilizar o dióxido de enxofre (SO_2) como regressora, este coeficiente se apresentou negativo, o que leva a uma falsa impressão de que um aumento no nível do poluente acarreta uma diminuição no desfecho médio, o que não é razoável. O motivo de este resultado ter sido alcançado está sendo pesquisado.

Visto que estes efeitos foram significativos, é razoável questionar se os mesmos persistem por mais de dois dias. Neste caso, a média utilizada como variável explicativa estaria subestimando o efeito cumulativo dos poluentes e se faz necessário avaliar um modelo no qual estejam presentes mais defasagens do poluente. Foi ajustado em seguida, então, o modelo de defasagens distribuídas por 40 dias, sem que fossem impostas as restrições (1.4) nos coeficientes β . A estrutura apresentada é então reescrita como:

$$\begin{aligned}
 Obitos_t &\sim Poisson(\mu_t) \\
 \log(\mu_t) &= \gamma + \varphi NRES5 \\
 &+ s(temperatura) + s(umidade) \\
 &+ \beta_0 POLUENTE_t + \beta_1 POLUENTE_{t-1} \\
 &+ \dots + \beta_q POLUENTE_{t-40}.
 \end{aligned} \tag{3.11}$$

De acordo com Zanobetti *et al* (2002) e Schwartz (2000), embora o modelo de defasagens distribuídas sem restrições seja muito instável para produzir informação sobre a forma de evolução dos poluentes, ainda assim este fornece uma estimativa não tendenciosa do efeito global dos mesmos. Este efeito é dado pela soma das estimativas dos coeficientes para todas as defasagens, e sob o enfoque bayesiano, dada a amostra destes coeficientes, foi obtida uma amostra dos coeficientes β_i , $i = 0, \dots, 40$. Isto tornou possível calcular a média a posteriori e intervalos de credibilidade para $\sum_{i=0}^{40} \beta_i$, que é um indicador do efeito cumulativo do poluente. Os resultados desta estimação são apresentados na tabela 3.4. Sob o enfoque clássico este indicador foi dado pela soma dos coeficientes β_i , $i = 0, \dots, 40$.

Tabela 3.4: Resultados da estimação dos modelos de defasagens distribuídas por 40 dias, sem restrições para os poluentes CO , PM_{10} e SO_2 .

	Enfoque clássico		Enfoque bayesiano			
	coeficiente	AIC	m.p.	I.Cred.	pD	DIC
CO	0.43102	5273.70	0.45558	0.38995 ; 0.53022	53.83	1889.26
PM_{10}	0.20865	5426.52	0.20243	0.11049 ; 0.29349	53.53	2041.94
SO_2	-0.25233	5353.76	-0.25257	-0.31514 ; -0.19218	52.40	1971.22

Observa-se mais uma vez que o coeficiente associado ao dióxido de enxofre apresentou-se negativo e, ainda, que tanto sob AIC (enfoque clássico) quanto DIC (enfoque bayesiano), a inclusão de mais que três dias no horizonte de impacto dos poluentes produziu, em geral, modelos com critérios de comparação menores que aqueles obtidos com médias móveis de até três dias, como disposto na tabela 3.3. No entanto, apesar de - sob o ponto de vista clássico - o modelo acima fornecer uma estimativa não tendenciosa dos efeitos dos poluentes, as informações trazidas por ele são instáveis, já que a utilização de níveis de poluentes em dias subseqüentes pode introduzir colinearidade, além de sobreparametrizar o modelo, o que pode ser verificado ao observar a coluna de pD . Assim, os parâmetros β sofreram restrições polinomiais, como descrito na Seção 1.2.1 e foram ajustados modelos de defasagens distribuídas com restrições polinomiais de ordens $d = 2, 3$ e 4 . Reescrevendo (3.11) como em (3.3), tem-se:

$$\begin{aligned}
 \log(\mu_t) &= \gamma + \varphi NRES5 \\
 &+ s(temperatura) + s(umidade) \\
 &+ \rho_0 Z_{0,t} + \dots + \rho_d Z_{d,t},
 \end{aligned}$$

com $d = 2, 3$ ou 4 .

O modelo acima foi ajustado seguindo os enfoques clássico e bayesiano. Do ponto de vista clássico, os diferentes ajustes foram comparados usando o critério AIC e, do ponto de vista bayesiano, aplicou-se o critério DIC para comparação de modelos, após convergência do algoritmo MCMC.

A tabela 3.5 exibe AIC , pD e DIC para estes modelos, chegando-se à conclusão de que, tanto segundo o enfoque clássico quanto segundo o enfoque bayesiano, os melhores modelos foram aqueles que consideraram aproximação polinomial de grau 2. Observa-se que houve uma diminuição considerável no número efetivo de parâmetros (pD) estimado, quando estes são comparados aos valores apresentados na tabela 3.4.

Tabela 3.5: Critérios de comparação resultantes da estimação dos modelos de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4 para os poluentes CO , PM_{10} e SO_2 .

	$d = 2$			$d = 3$			$d = 4$		
	AIC	pD	DIC	AIC	pD	DIC	AIC	pD	DIC
CO	5250.00	16.54	1866.30	5251.20	17.86	1867.87	5252.60	18.73	1868.95
PM_{10}	5399.05	16.12	2017.42	5400.38	16.98	2018.17	5400.73	18.07	2018.34
SO_2	5341.86	15.73	1962.10	5343.33	16.41	1962.54	5344.44	17.71	1964.18

Em seguida são apresentados os resultados alcançados a partir do ajuste destes modelos semi-paramétricos. Para facilitar a compreensão, serão apresentados em separado os resultados dos blocos paramétrico e não-paramétrico do modelo.

Análise do bloco paramétrico:

A figura 3.7 mostra os histogramas das amostras das distribuições a posteriori dos coeficientes ρ_0 , ρ_1 e ρ_2 para os poluentes PM_{10} , CO e SO_2 , no modelo de defasagens distribuídas com restrições polinomiais de ordem 2, estimadas seguindo o enfoque bayesiano. Verifica-se que todos se apresentaram centrados na mediana. Dada a amostra destes coeficientes, foi aplicada a transformação (1.4) e, assim, foi obtida uma amostra dos coeficientes β_i , $i = 0, \dots, 40$, tornando possível calcular a média a posteriori e intervalos de credibilidade para $\sum_{i=0}^{40} \beta_i$, que é um indicador do efeito cumulativo do poluente. Sob o enfoque clássico este indicador foi mais uma vez dado pela soma dos coeficientes β_i , $i = 0, \dots, 40$.

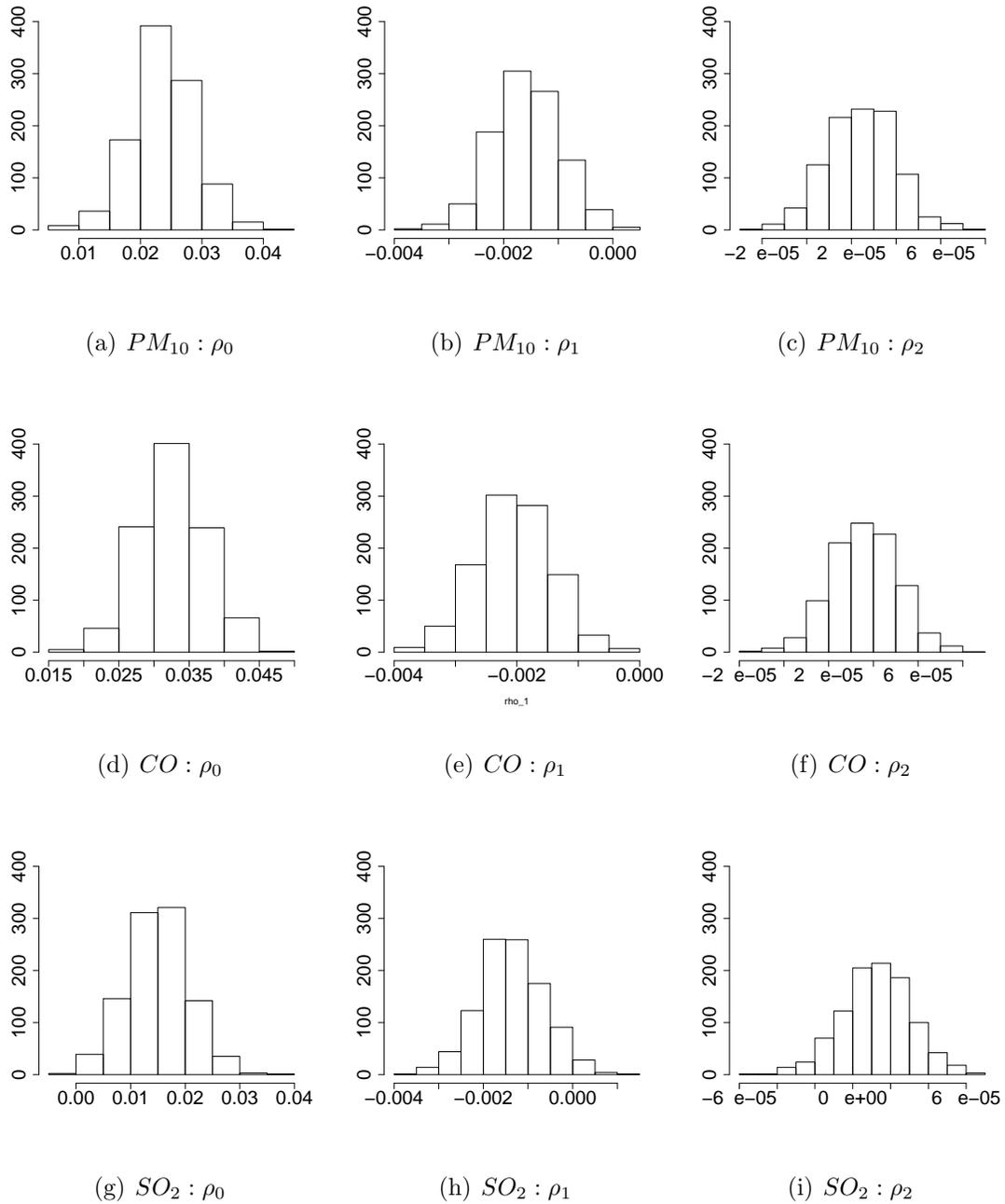


Figura 3.7: Histogramas das amostras das distribuições a posteriori dos coeficientes ρ_0, ρ_1 e ρ_2 para os poluentes PM_{10}, CO e SO_2 , no modelo de defasagens distribuídas com restrições polinomiais de ordem 2.

Em seguida são apresentadas as matrizes de covariâncias dos coeficientes ρ_0 , ρ_1 e ρ_2 , calculadas seguindo os enfoques clássico e bayesiano. Como foi dito, seguindo o enfoque clássico, as matrizes de covariância destes coeficientes ρ são utilizadas para calcular as matrizes de covariância dos coeficientes β , visto que estes são escritos como em (1.4). Assim, tem-se mais uma vez que um indicador do efeito cumulativo do poluente é dado por $\sum_{i=0}^{40} \beta_i$ e intervalos de confiança podem ser calculados assumindo que $\sum_{i=0}^{40} \beta_i \sim N\left(\sum_{i=0}^{40} \hat{\beta}_i, \sum_{i=0}^{40} VAR[\hat{\beta}_i] + 2 \sum \sum_{i < i'} COV[\hat{\beta}_i, \hat{\beta}_{i'}]\right)$.

Tabela 3.6: Matrizes de covariância dos coeficientes ρ_0 , ρ_1 e ρ_2 , estimadas seguindo os enfoques clássico e bayesiano para os poluentes PM_{10} , CO e SO_2 , no modelo de defasagens distribuídas com restrições polinomiais de ordem 2.

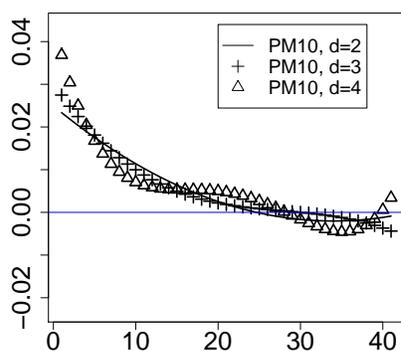
Clássico		ρ_0	ρ_1	ρ_2	Clássico		ρ_0	ρ_1	ρ_2	Clássico		ρ_0	ρ_1	ρ_2
	ρ_0	2.598E-05	-2.931E-06	6.330E-08		ρ_0	2.456E-05	-2.879E-06	6.246E-08		ρ_0	2.598E-05	-2.931E-06	6.330E-08
	ρ_1	-2.931E-06	3.947E-07	-9.334E-09		ρ_1	-2.879E-06	4.029E-07	-9.627E-09		ρ_1	-2.931E-06	3.947E-07	-9.334E-09
	ρ_2	6.330E-08	-9.334E-09	2.317E-10		ρ_2	6.246E-08	-9.627E-09	2.423E-10		ρ_2	6.330E-08	-9.334E-09	2.317E-10
Bayes		ρ_0	ρ_1	ρ_2	Bayes		ρ_0	ρ_1	ρ_2	Bayes		ρ_0	ρ_1	ρ_2
	ρ_0	2.613E-05	-2.922E-06	6.306E-08		ρ_0	2.244E-05	-2.595E-06	5.589E-08		ρ_0	3.174E-05	-3.828E-06	8.551E-08
	ρ_1	-2.922E-06	3.914E-07	-9.226E-09		ρ_1	-2.595E-06	3.682E-07	-8.842E-09		ρ_1	-3.828E-06	5.318E-07	-1.281E-08
	ρ_2	6.306E-08	-9.226E-09	2.282E-10		ρ_2	5.589E-08	-8.842E-09	2.248E-10		ρ_2	8.551E-08	-1.281E-08	3.210E-10

Os efeitos totais estimados no modelo de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4, para os poluentes PM_{10} , CO e SO_2 , dados por $\sum_{i=0}^q \beta_i$, são apresentados na tabela 3.7, juntamente com intervalos de confiança e credibilidade de 95%. O intervalo de confiança é dado por $(\hat{\beta}_i - 1.96DP_{\hat{\beta}_i}; \hat{\beta}_i + 1.96DP_{\hat{\beta}_i})$ e o intervalo de credibilidade é dado pelos quantis 2,5% e 97,5% da distribuição a posteriori de $\sum_{i=0}^{40} \beta_i$. Observa-se que estes efeitos foram significativos em todos os casos. Ainda, verifica-se mais uma vez que os coeficientes relativos ao dióxido de enxofre se apresentaram negativos, e isto está sendo analisado.

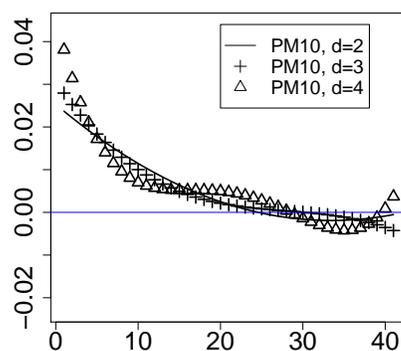
Tabela 3.7: Coeficientes de efeito global resultantes da estimação do modelo de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4, para os poluentes CO , PM_{10} e SO_2 .

Restrição polinomial de ordem 2				
	Enfoque clássico		Enfoque bayesiano	
	coeficiente	I. Conf.	m. p.	I. Cred.
CO	0.45412	0.42009 ; 0.48814	0.46021	0.38752 ; 0.53193
PM_{10}	0.21191	0.17834 ; 0.24549	0.21558	0.13023 ; 0.30272
SO_2	-0.22826	-0.26588 ; -0.19065	-0.23000	-0.29500 ; -0.16754
Restrição polinomial de ordem 3				
	Enfoque clássico		Enfoque bayesiano	
	coeficiente	I. Conf.	m. p.	I. Cred.
CO	0.45588	0.41170 ; 0.50006	0.46094	0.38774 ; 0.53354
PM_{10}	0.21499	0.17301 ; 0.25698	0.21789	0.13306 ; 0.30567
SO_2	-0.22936	-0.27224 ; -0.18649	-0.23021	-0.29093 ; -0.16266
Restrição polinomial de ordem 4				
	Enfoque clássico		Enfoque bayesiano	
	coeficiente	I. Conf.	m. p.	I. Cred.
CO	0.45738	0.40387 ; 0.51088	0.46125	0.38901 ; 0.53455
PM_{10}	0.2253	0.17422 ; 0.27638	0.23378	0.14091 ; 0.31953
SO_2	-0.23185	-0.26946 ; -0.19423	-0.23383	-0.29437 ; -0.17245

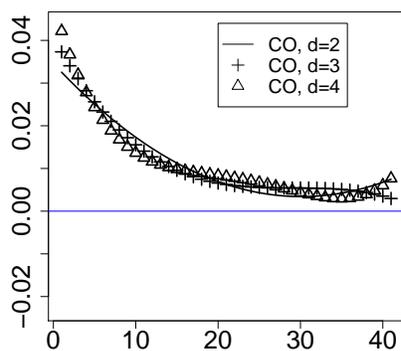
A figura 3.8 exibe a evolução temporal dos efeitos estimados associados a aumentos de 1 desvio-padrão em relação à média nos níveis de PM_{10} , CO e SO_2 , para os modelos de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4, e a figura 3.9 apresenta uma comparação entre os efeitos estimados para cada um dos poluentes no modelo de defasagens distribuídas com restrições polinomiais de ordem 2. Observa-se através dos gráficos que as estimativas se apresentaram bastante similares tanto para o ajuste feito seguindo o enfoque clássico quanto para o enfoque bayesiano. Ainda através dos gráficos pode-se verificar claramente que quanto mais restrição se impõe no polinômio, menos livre será a estimativa dos coeficientes β . Apesar de o impacto dos três poluentes sobre o processo de contagens de óbitos decair gradual e lentamente, o impacto do SO_2 parece se extinguir depois de aproximadamente 15 dias, seguido pelo PM_{10} , que parece persistir até 30 dias. No caso do CO o efeito parece persistir até mais de 40 dias, e para observar melhor estes resultados, na figura 3.10 estes efeitos são mostrados, juntamente com intervalos de confiança / credibilidade de 95%, para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2, que tanto para o enfoque clássico (AIC) quanto para o enfoque bayesiano (DIC) foi o que apresentou valores mais baixos dos critérios de comparação. Observa-se a partir da figura citada que realmente o efeito do SO_2 deixa de ser significativo depois de 15 dias, seguido do PM_{10} , cujo efeito se extingue antes de passar 20 dias. O poluente que parece ter o impacto mais forte e duradouro na resposta é o CO , já que seu efeito só deixa de ser significativo depois de aproximadamente 25 dias.



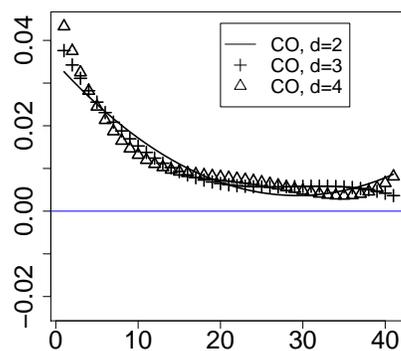
(a) Enfoque clássico



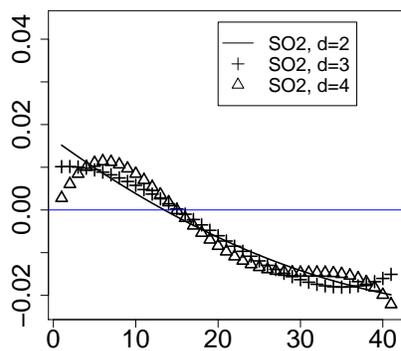
(b) Enfoque bayesiano



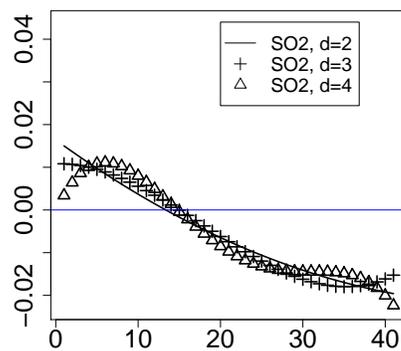
(c) Enfoque clássico



(d) Enfoque bayesiano

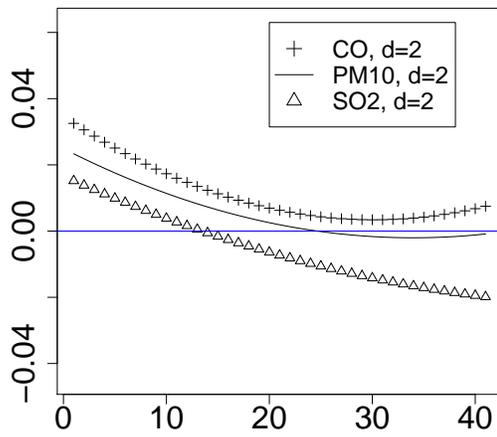


(e) Enfoque clássico

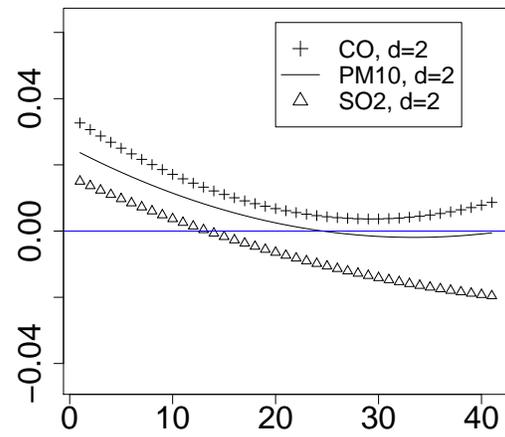


(f) Enfoque bayesiano

Figura 3.8: Evolução temporal dos efeitos estimados associados a aumentos nos níveis de PM_{10} , CO e SO_2 , para os modelos de defasagens distribuídas com restrições polinomiais de ordens 2, 3 e 4.

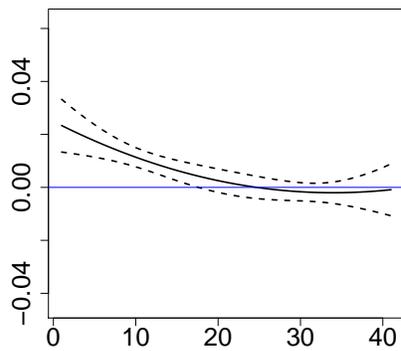


(a) Enfoque clássico

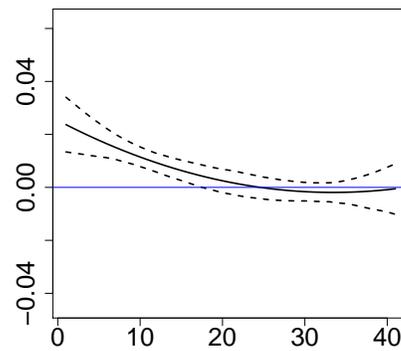


(b) Enfoque bayesiano

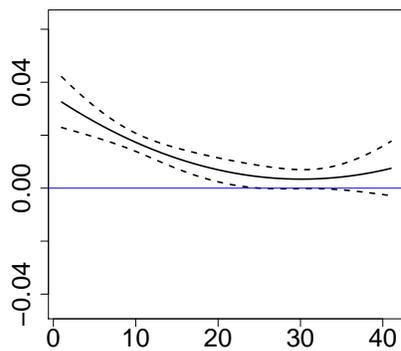
Figura 3.9: Evolução temporal dos efeitos estimados associados a aumentos nos níveis de PM_{10} , CO e SO_2 , para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2.



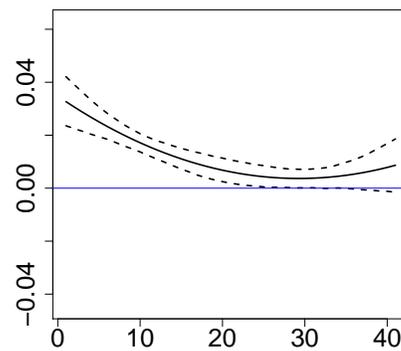
(a) Enfoque clássico



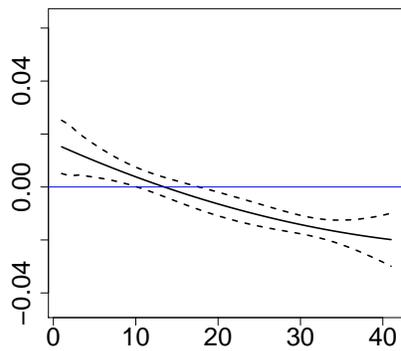
(b) Enfoque bayesiano



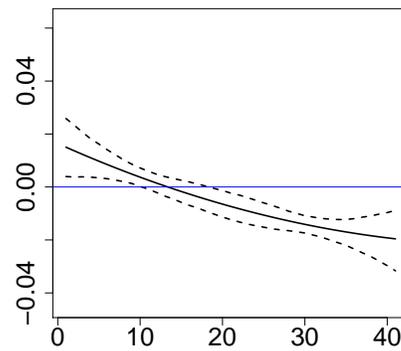
(c) Enfoque clássico



(d) Enfoque bayesiano



(e) Enfoque clássico



(f) Enfoque bayesiano

Figura 3.10: Evolução temporal e intervalos de confiança / credibilidade dos efeitos estimados associados a aumentos nos níveis de PM_{10} , CO e SO_2 para o modelo de defasagens distribuídas com restrições polinômiais de ordem 2.

Em seguida são apresentados gráficos dos riscos relativos estimados sob o enfoque bayesiano para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2. O cálculo destes riscos relativos foi feito tomando toda a amostra dos β 's, e a avaliação destes é importante para verificar o impacto de valores específicos do poluente na resposta.

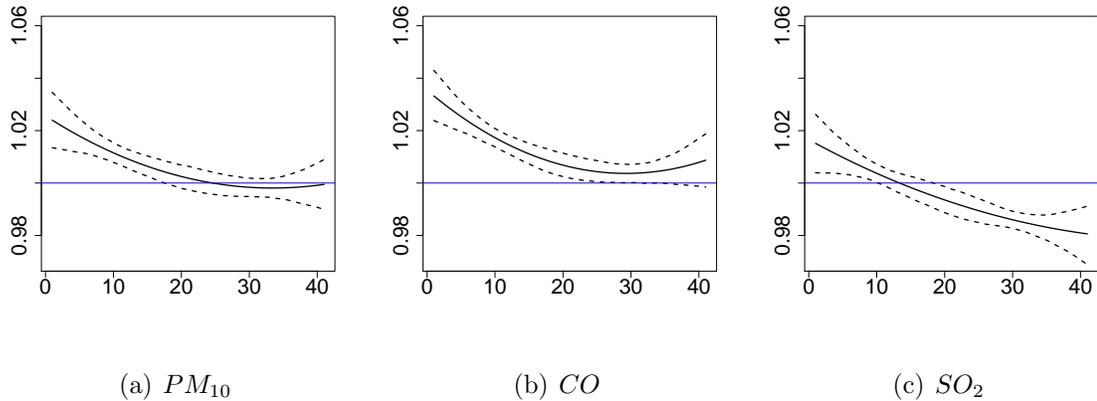
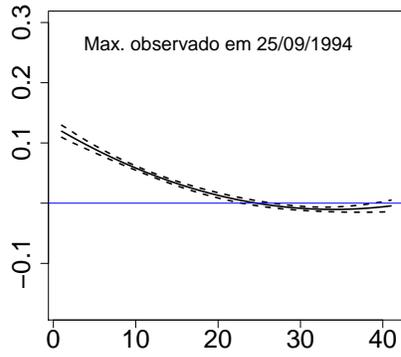
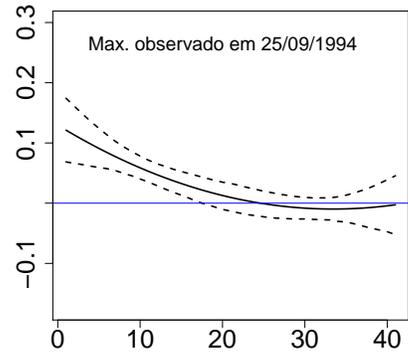


Figura 3.11: Evolução temporal dos riscos relativos estimados associados a aumentos nos níveis de PM_{10} , CO e SO_2 , para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2, seguindo o enfoque bayesiano.

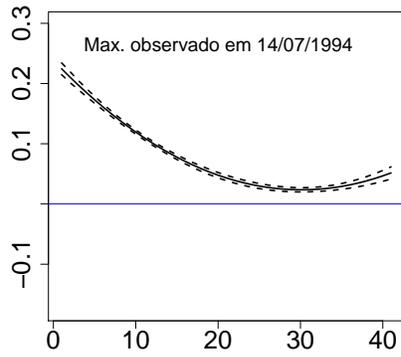
Pode-se observar a partir da figura 3.11 que um aumento de 1 desvio-padrão em relação à média do monóxido de carbono CO acarretaria um aumento imediato de aproximadamente 3,3% no desfecho médio. Nos casos de PM_{10} e SO_2 estes aumentos são de 2,4% e 1,5%, respectivamente. Seria interessante, então, avaliar qual o impacto de valores extremos, como o máximo observado no período de estudo, dos poluentes na resposta. A figura 3.12 mostra a evolução temporal dos efeitos estimados associados ao nível máximo observado de cada um dos poluentes. Verifica-se que o CO , cujo aumento em um desvio-padrão antes causava um aumento de aproximadamente 3,3% na resposta, passou a ter um impacto de mais de 25% nesta resposta. PM_{10} e SO_2 também tiveram seus efeitos aumentados para quase 13% e 9% respectivamente.



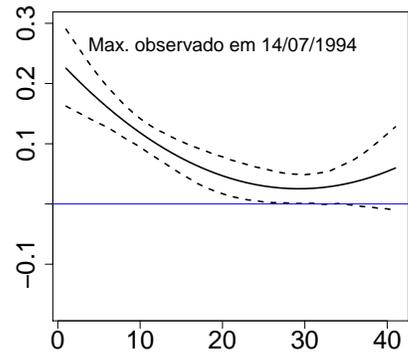
(a) Enfoque clássico



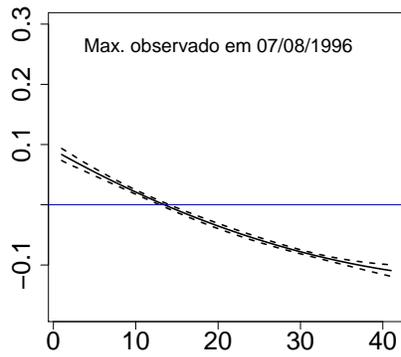
(b) Enfoque bayesiano



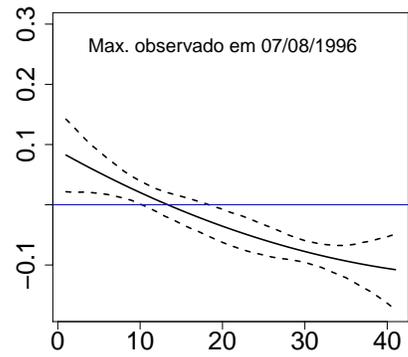
(c) Enfoque clássico



(d) Enfoque bayesiano



(e) Enfoque clássico



(f) Enfoque bayesiano

Figura 3.12: Evolução temporal dos efeitos estimados associados ao nível máximo observado dos poluentes PM_{10} , CO e SO_2 , para o modelo de defasagens distribuídas com restrições polinomiais de ordem 2.

Análise do bloco não-paramétrico:

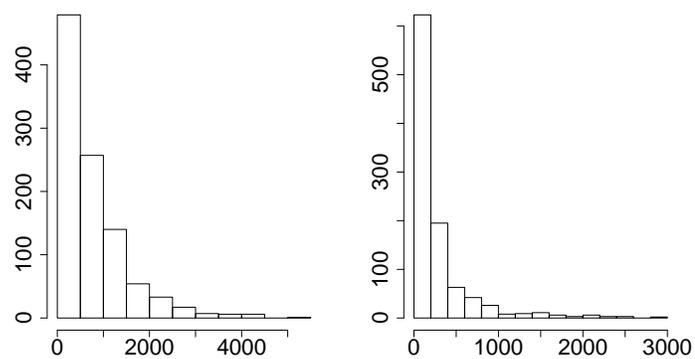
O bloco não-paramétrico do modelo consiste nas suavizações das variáveis temperatura e umidade, e então os resultados apresentados aqui serão referentes aos parâmetros λ e ϕ_k , $k = 0, \dots, n + 1$, onde n é o número de nós interiores da *spline*. Neste estudo foi adotado $n = 20$.

Em seguida é apresentada uma tabela com as estatísticas sumárias dos parâmetros de suavização λ das variáveis temperatura e umidade, estimadas seguindo o enfoque bayesiano. São apresentados também na figura 3.13 os histogramas destes parâmetros.

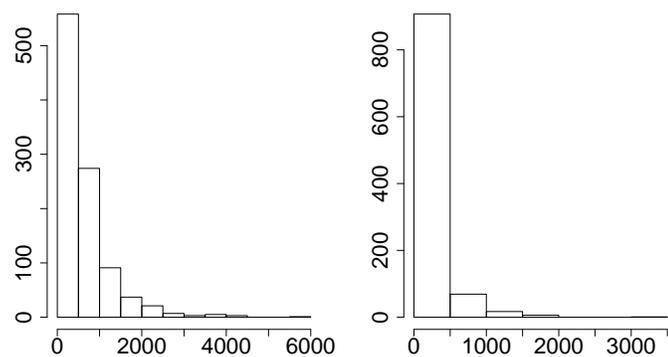
Tabela 3.8: Estatísticas sumárias do parâmetro de suavização λ das variáveis umidade e temperatura para os modelos tendo como regressoras os poluentes PM_{10} , CO ou SO_2 , com defasagens distribuídas sujeitas a restrições polinomiais de ordem 2.

	PM_{10}		CO		SO_2	
	umidade	temperatura	umidade	temperatura	umidade	temperatura
média	767.07	284.77	195.36	606.60	464.19	632.44
mediana	530.92	144.88	104.27	416.12	235.16	475.29
$I.Cred_{.95\%}$	59.2;2800.4	25.0;1572.5	21.7;966.4	48.0;2283.3	28.2;2021.8	64.5;2033.0

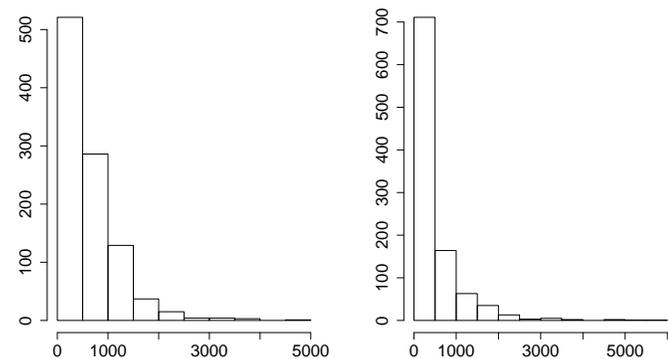
Na sequência, são apresentados nas figuras 3.14 a 3.19 os histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para os poluentes CO , PM_{10} e SO_2 . Gráficos destes coeficientes ϕ são mostrados nas figuras 3.20 e 3.21, juntamente com intervalos de confiança / credibilidade de 95%, que ilustram o efeito não linear de umidade e temperatura, estimados seguindo os enfoques clássico e bayesiano, no modelo de defasagens distribuídas com restrições polinomiais de ordem 2 para os poluentes PM_{10} , CO ou SO_2 .



(a) PM_{10}



(b) CO



(c) SO_2

Figura 3.13: Histograma do parâmetro de suavização λ das variáveis umidade e temperatura para os poluentes PM_{10} , CO e SO_2 , no modelo de defasagens distribuídas com restrições polinomiais de ordem 2.

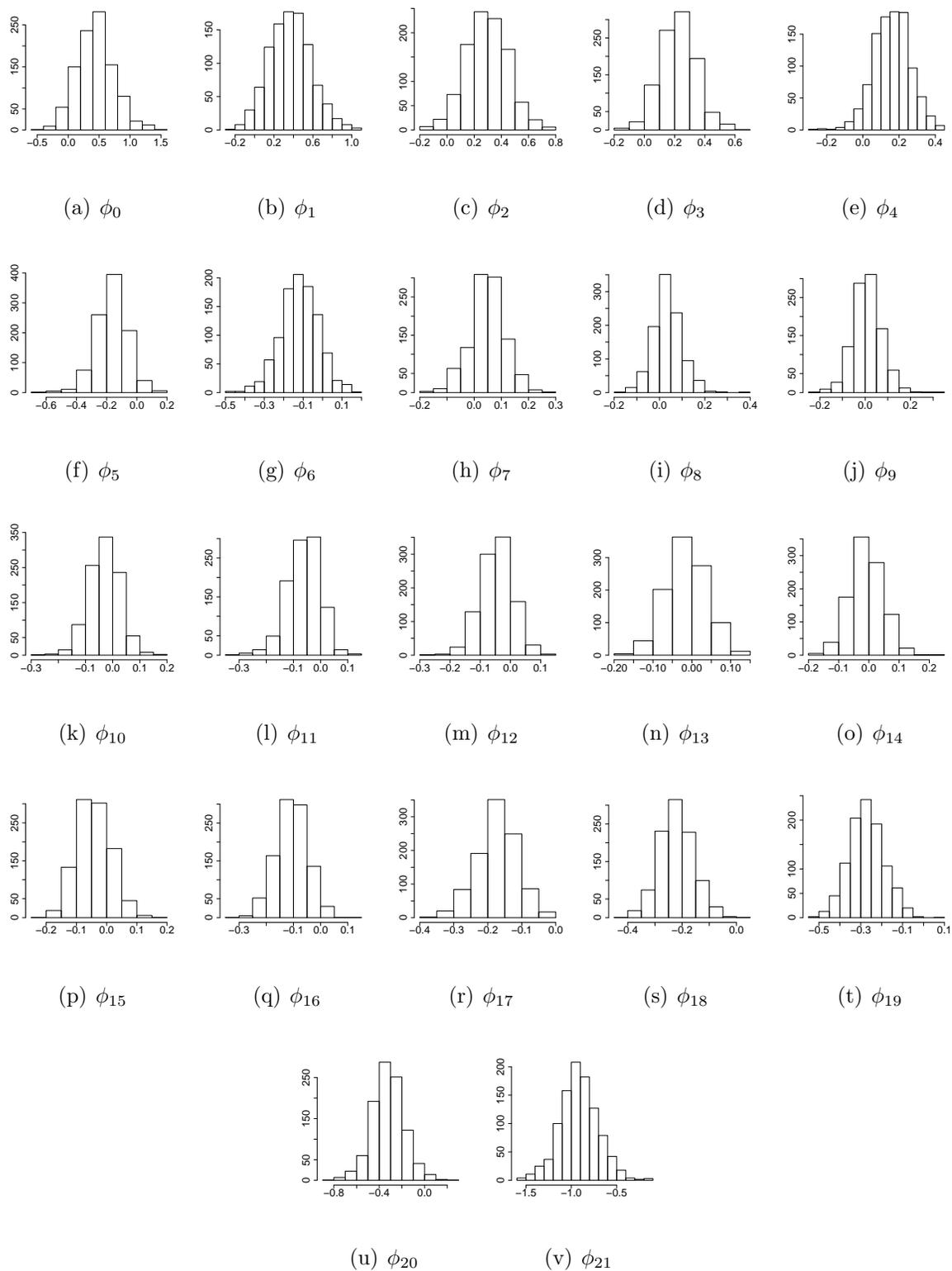


Figura 3.14: Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável temperatura no modelo contendo o poluente CO como regressora.

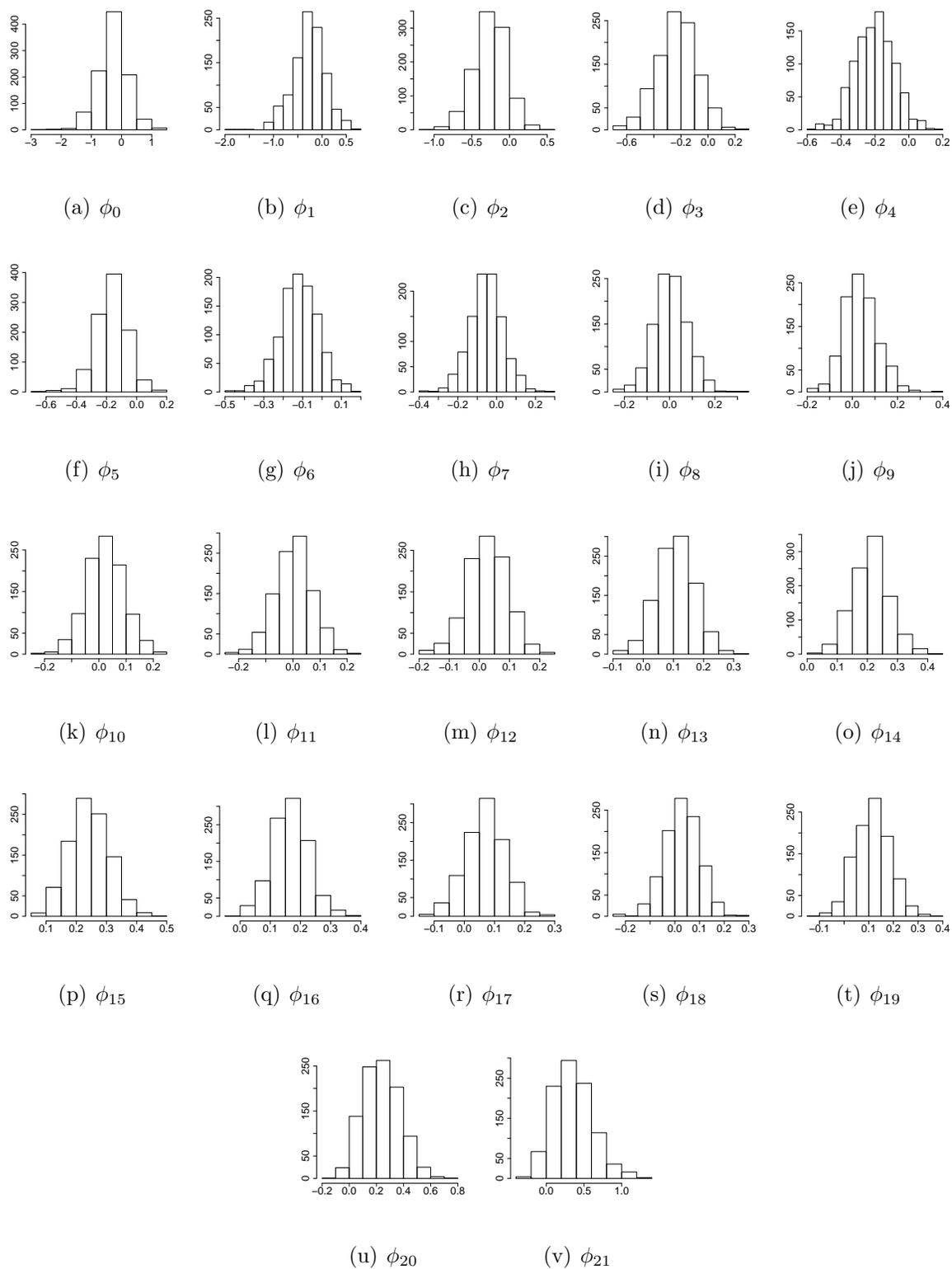


Figura 3.15: Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável umidade no modelo contendo o poluente CO como regressora.

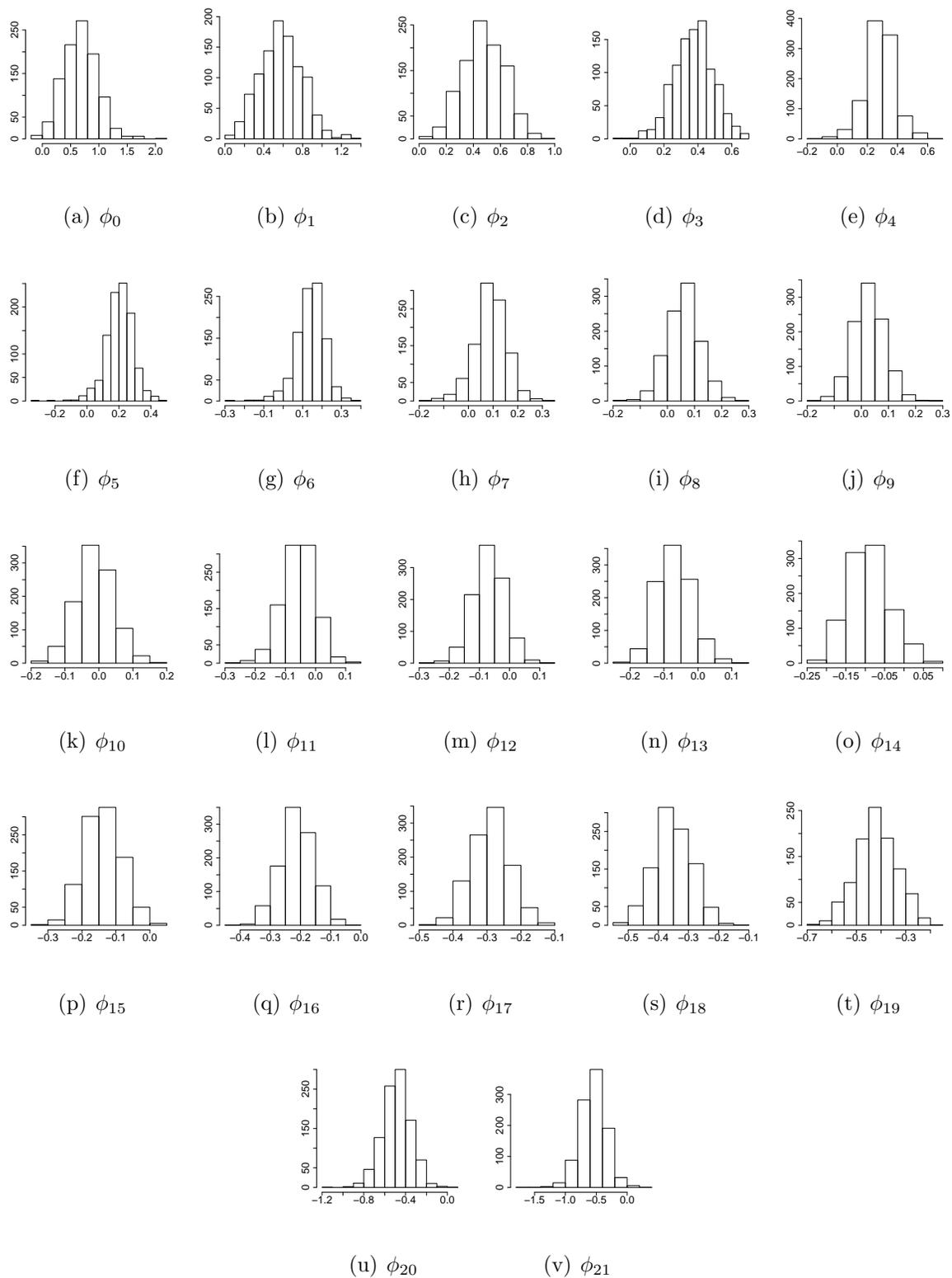


Figura 3.16: Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável temperatura no modelo contendo o poluente PM_{10} como regressora.

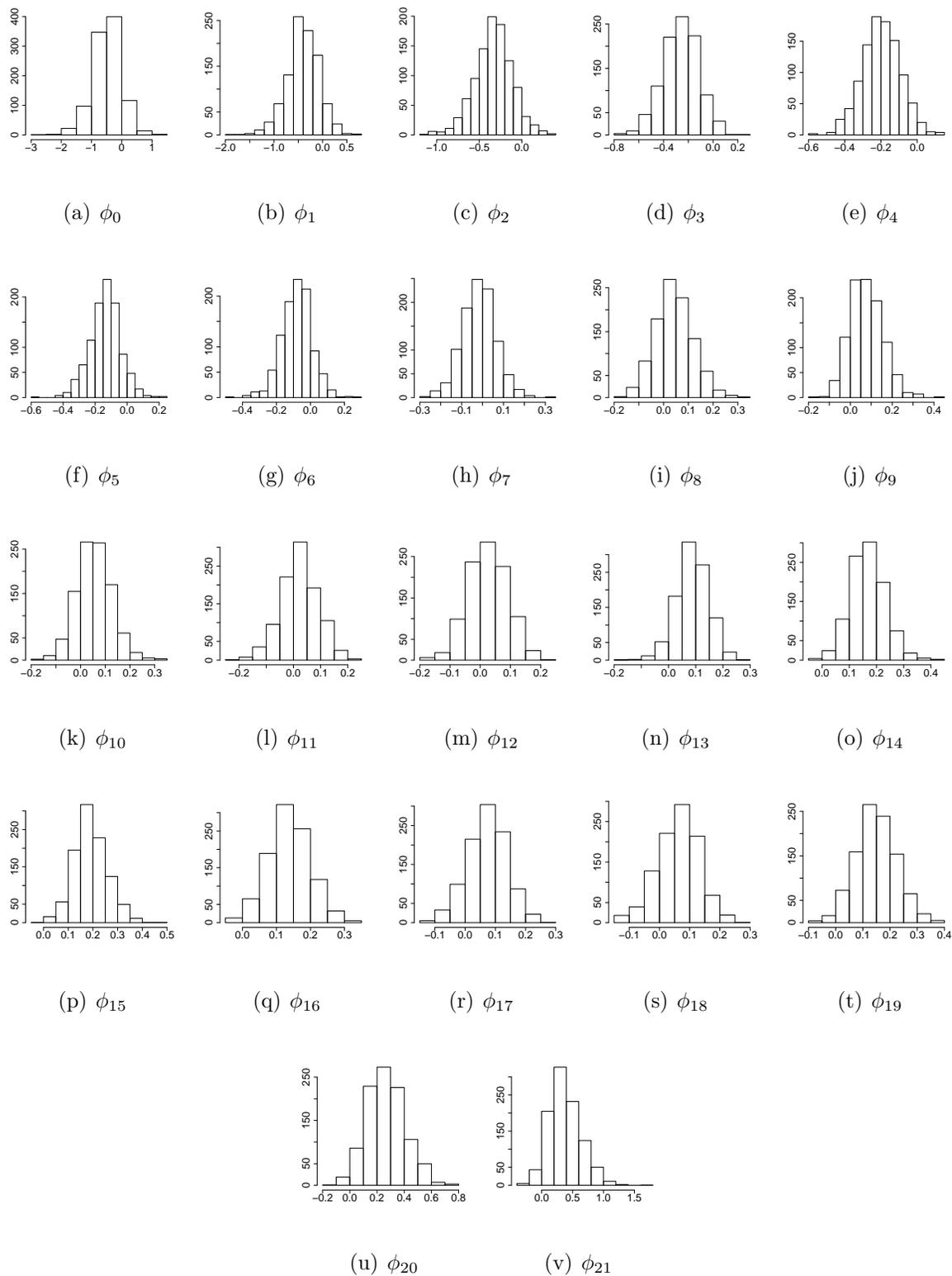


Figura 3.17: Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável umidade no modelo contendo o poluente PM_{10} como regressora.

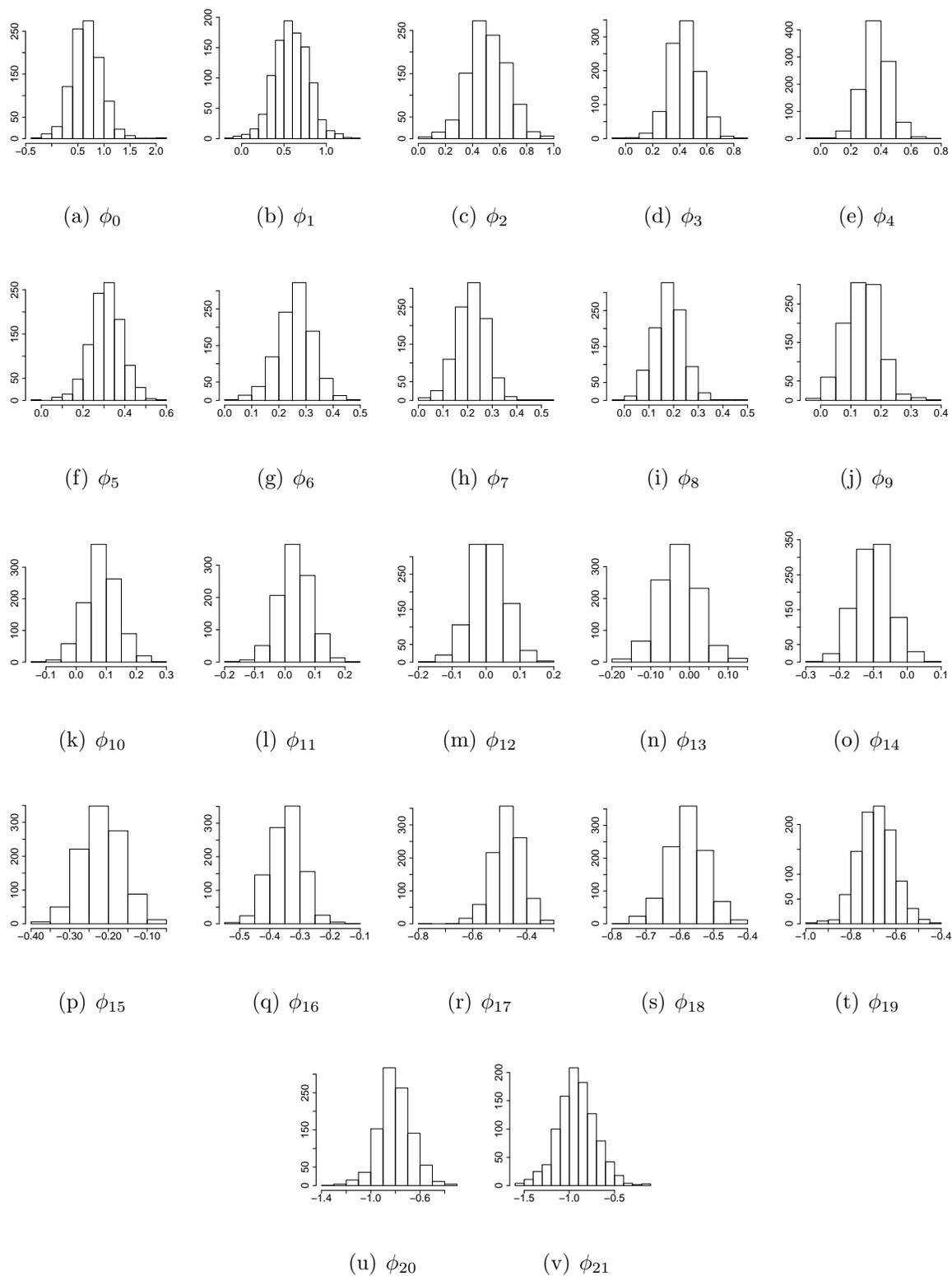


Figura 3.18: Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável temperatura no modelo contendo o poluente SO_2 como regressora.

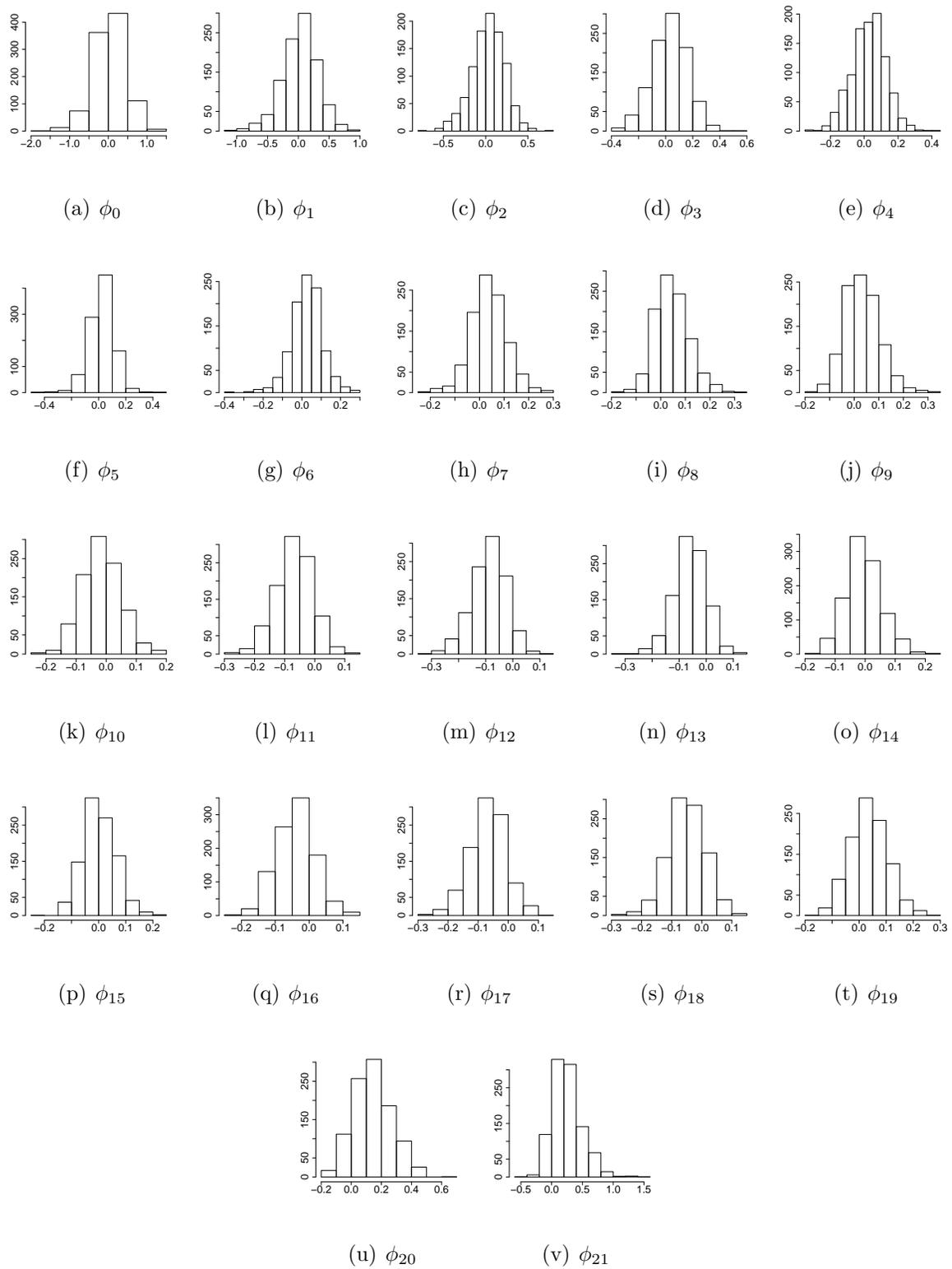
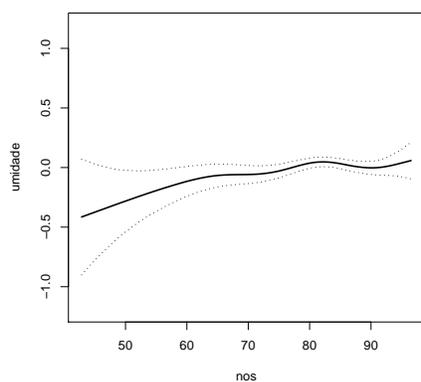
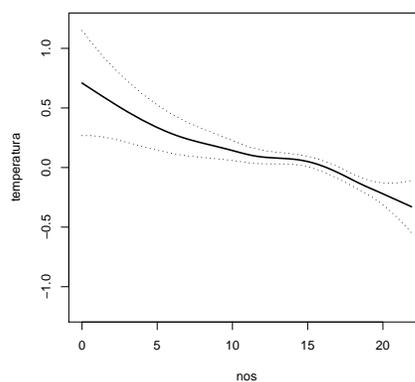


Figura 3.19: Histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para a variável umidade no modelo contendo o poluente SO_2 como regressora.

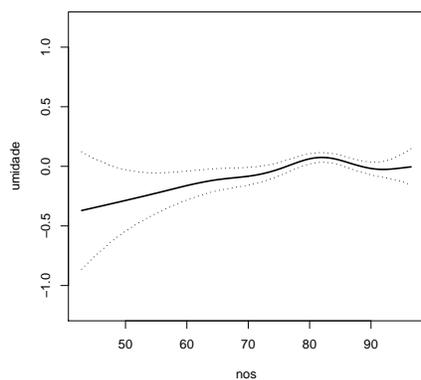
Como foi dito, as figuras 3.20 e 3.21 a seguir apresentam os efeitos não lineares estimados seguindo os enfoques clássico (figura 3.20) e bayesiano (figura 3.21), juntamente com intervalos de confiança / credibilidade de 95%, no modelo de defasagens distribuídas com restrições polinomiais de ordem 2 para os poluentes PM_{10} , CO ou SO_2 . Pode-se observar que estes efeitos estimados se apresentaram bastante similares ao comparar os resultados dos dois enfoques e, ainda, que em geral, a variável temperatura apresentou uma estrutura mais suave que a umidade.



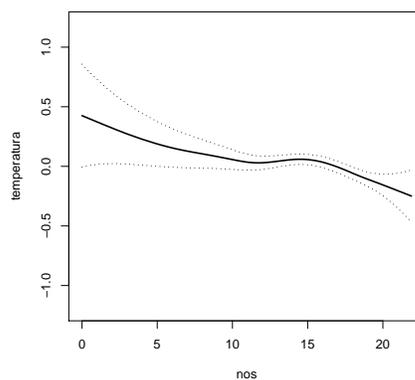
(a) umidade



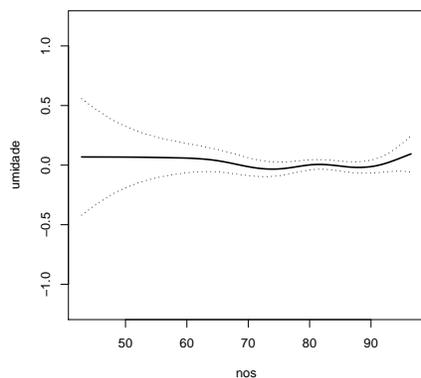
(b) temperatura



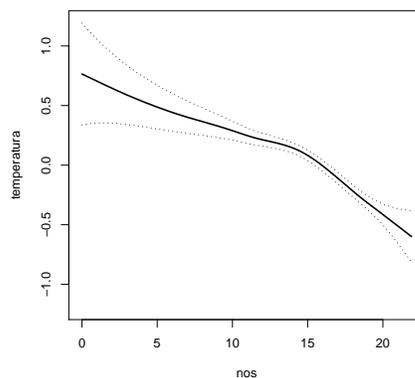
(c) umidade



(d) temperatura

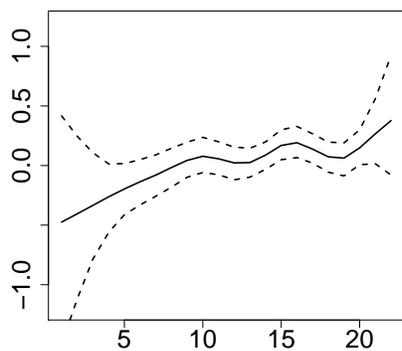


(e) umidade

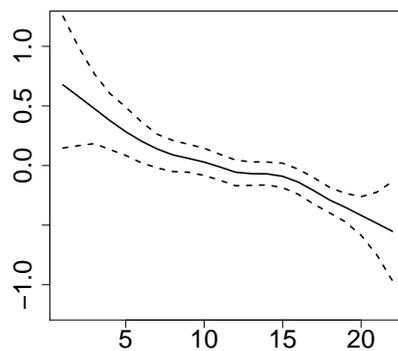


(f) temperatura

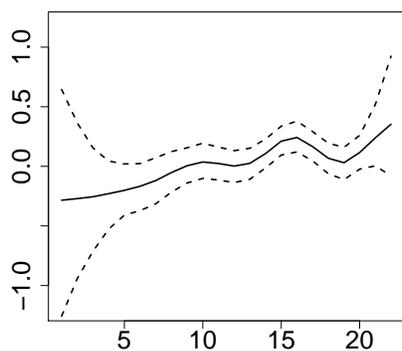
Figura 3.20: Efeito não linear de umidade e temperatura, estimado seguindo o enfoque clássico, no modelo de defasagens distribuídas com restrições polinomiais de ordem 2 para os poluentes PM_{10} , CO e SO_2 .



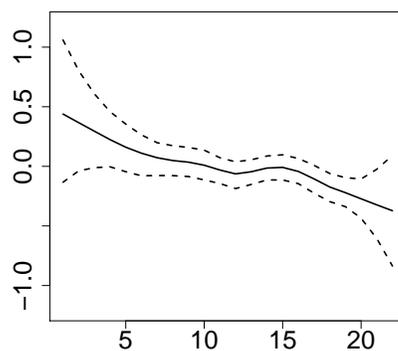
(a) umidade



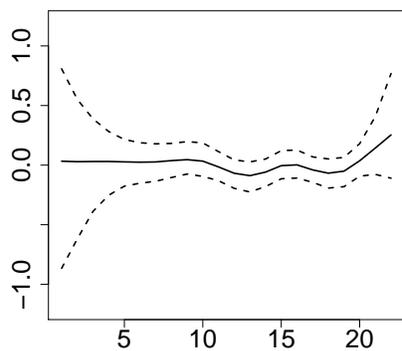
(b) temperatura



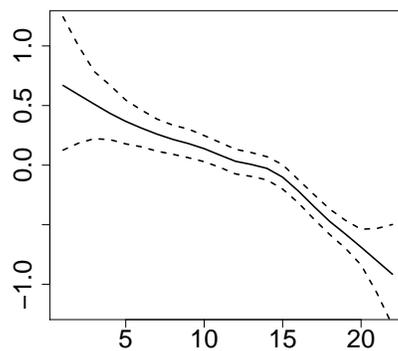
(c) umidade



(d) temperatura



(e) umidade



(f) temperatura

Figura 3.21: Efeito não linear de umidade e temperatura, estimado seguindo o enfoque bayesiano, no modelo de defasagens distribuídas com restrições polinomiais de ordem 2 para os poluentes PM_{10} , CO e SO_2 .

3.4.2 Aplicação utilizando resposta contínua: Níveis de *CO* em São Paulo

Nesta aplicação foram utilizados como resposta os níveis diários e médias semanais do poluente *CO* em São Paulo, pressupondo distribuição $Normal(\mu_t, \sigma^2)$ e $Gama(\alpha, \beta_t)$, e tendo as variáveis umidade média e temperatura mínima como regressoras.

Mais uma vez escrevendo o modelo como a estrutura dada em (3.3), tem-se, para o caso em que foi considerada distribuição $Normal(\mu_t, \sigma^2)$ que:

$$\begin{aligned} \log(Poluenta_t) &\sim Normal(\mu_t, \sigma^2) \\ \eta_t = \mu_t &= \gamma + f(temperatura, umidade) \end{aligned} \quad (3.12)$$

e nos casos em que foi considerada distribuição $Gama(\alpha, \beta)$:

$$\begin{aligned} Poluenta_t &\sim Gama(\alpha, \beta_t) \\ \eta_t = \log\left(\frac{\alpha}{\beta_t}\right) &= \gamma + f(temperatura, umidade), \end{aligned} \quad (3.13)$$

ou seja, a média da variável resposta será função de temperatura e umidade.

Para cada uma das respostas foram testados 16 modelos, que compreendiam combinações das variáveis temperatura mínima e umidade média, ou de suavizações destas mesmas variáveis, e suas defasagens de até três dias. Os modelos pertencem a 4 grupos:

Grupo 1: modelos 1, 2, 3 e 4 as variáveis umidade e temperatura são incluídas de forma paramétrica

Grupo 2: modelos 5, 6, 7 e 8 as variáveis umidade e temperatura são incluídas de forma não-paramétrica

Grupo 3: modelos 9, 10, 11 e 12 a variável umidade é incluída de forma paramétrica e a variável temperatura é incluída de forma não-paramétrica

Grupo 4: modelos 13, 14, 15 e 16 a variável umidade é incluída de forma não-paramétrica e a variável temperatura é incluída de forma paramétrica

Em seguida é apresentada uma tabela com a relação dos 16 modelos ajustados a este banco de dados.

Tabela 3.9: Representação esquemática dos modelos ajustados aos dados de poluição.

Modelos		temperatura _{t-k}				umidade _{t-k}				s(temperatura _{t-k})				s(umidade _{t-k})			
		k				k				k				k			
		0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
Grupo 1	1	x				x											
	2	x	x			x	x										
	3	x	x	x		x	x	x									
	4	x	x	x	x	x	x	x	x								
Grupo 2	5									x				x			
	6									x	x			x	x		
	7									x	x	x		x	x	x	
	8									x	x	x	x	x	x	x	x
Grupo 3	9					x				x							
	10					x	x			x	x						
	11					x	x	x		x	x	x					
	12					x	x	x	x	x	x	x	x				
Grupo 4	13	x												x			
	14	x	x											x	x		
	15	x	x	x										x	x	x	
	16	x	x	x	x									x	x	x	x

As tabelas 3.10 e 3.11 a seguir apresentam os resultados de Critério de Informação de Akaike (AIC) e Critério de Informação da Deviance (DIC) para os 16 modelos ajustados.

Tabela 3.10: Critérios de comparação de modelos para observações diárias de $\log(CO)$.

	$\log(CO) \sim Normal(\mu_t, \sigma^2)$			
	AIC	pD	DIC	<i>deviance</i> ($\hat{\theta}$)
modelo 1	1491.49	5.01	1465.426	1455.416
modelo 2	1447.28	6.89	1466.944	1453.169
modelo 3	1401.76	8.99	1466.834	1448.847
modelo 4	1381.89	10.84	1468.95	1447.261
modelo 5	1450.80	12.64	1470.11	1444.822
modelo 6	1409.05	20.10	1478.247	1438.055
modelo 7	1368.04	28.02	1487.047	1431.014
modelo 8	1355.71	34.75	1493.389	1423.899
modelo 9	1487.15	8.11	1465.410	1449.188
modelo 10	1446.35	13.02	1471.594	1445.563
modelo 11	1405.35	17.54	1477.235	1442.151
modelo 12	1387.65	22.07	1480.244	1436.106
modelo 13	1453.78	9.26	1469.198	1450.688
modelo 14	1409.33	13.68	1468.875	1441.513
modelo 15	1364.15	19.23	1479.798	1441.335
modelo 16	1350.57	24.22	1480.472	1432.023

Tabela 3.11: Critérios de comparação de modelos para observações diárias de CO .

	$CO \sim Gama(\alpha, \beta_t)$			
	AIC	pD	DIC	$deviance(\bar{\theta})$
modelo 1	5649.52	3.73	1351.213	1343.745
modelo 2	5609.09	5.5	1348.656	1337.655
modelo 3	5562.31	7.6	1361.153	1345.961
modelo 4	5543.84	9.62	1360.371	1341.128
modelo 5	5614.74	11.38	1345.421	1322.666
modelo 6	5573.85	21.05	1356.306	1314.197
modelo 7	5531.42	30.85	1379.001	1317.292
modelo 8	5521.16	40.16	1386.52	1306.194
modelo 9	5643.73	7.52	1356.614	1341.575
modelo 10	5606.18	13.19	1359.357	1332.968
modelo 11	5548.88	18.29	1375.388	1338.8
modelo 12	5619.41	23.79	1376.202	1328.616
modelo 13	5563.79	7.88	1339.232	1323.48
modelo 14	5576.14	13.68	1345.050	1317.698
modelo 15	5529.37	20.36	1368.994	1328.274
modelo 16	5516.58	25.81	1372.815	1321.188

Observa-se que entre os modelos testados, o modelo 9 foi aquele que apresentou menor valor do DIC quando se trabalhou com a variável transformada, e o modelo 13, no caso em que se trabalhou com a variável na escala original. No caso do ajuste seguindo o enfoque clássico, tanto para o CO quanto para o $\log(CO)$ o menor valor de AIC foi alcançado pelo modelo 16, que é aquele em que são incluídas defasagens de até 3 dias da variável temperatura e defasagens suavizadas de até 3 dias da variável umidade. Apesar de a utilização destas defasagens ser razoável, no caso do modelo implementado isto pode causar autocorrelação, visto que estas não sofreram restrições.

Uma questão que deve ser levada em consideração é a avaliação da influência do nível de agregação dos dados sobre os resultados. Desta forma, para avaliar tal questão, as observações diárias de CO foram agregadas por semana, e repetiu-se a estimação dos modelos de 1 a 16 apresentados na tabela 3.9 utilizando esta média semanal como variável resposta.

As tabelas 3.12 e 3.13 a seguir apresentam os resultados de Critério de Informação de Akaike (AIC) e Critério de Informação da Deviance (DIC) para os 16 modelos ajustados utilizando os dados agregados como resposta.

Tabela 3.12: Critérios de comparação de modelos para a média semanal de $\log(CO)$.

	$\log(CO) \sim Normal(\mu_t, \sigma^2)$			
	AIC	pD	DIC	<i>deviance</i> ($\bar{\theta}$)
modelo 1	50.87	4.96	213.173	203.250
modelo 2	49.42	7.13	214.564	200.313
modelo 3	52.27	9.02	215.344	197.303
modelo 4	54.52	10.92	218.384	196.541
modelo 5	48.98	10.02	217.018	196.970
modelo 6	53.08	15.37	222.161	191.419
modelo 7	60.16	22.36	230.084	185.355
modelo 8	67.62	26.89	233.35	179.569
modelo 9	52.4	7.29	214.675	200.086
modelo 10	50.59	11.57	218.207	195.059
modelo 11	55.34	15.88	222.249	190.499
modelo 12	58.58	20.24	226.929	186.449
modelo 13	47.37	7.54	215.330	200.254
modelo 14	51.14	10.85	217.897	196.202
modelo 15	56.81	14.58	220.302	191.139
modelo 16	61.53	18.33	225.111	188.451

Tabela 3.13: Critérios de comparação de modelos para a média semanal de CO .

	$CO \sim Gama(\alpha, \beta_t)$			
	AIC	pD	DIC	$deviance(\bar{\theta})$
modelo 1	662.73	3.44	206.712	199.827
modelo 2	662.04	5.44	207.434	196.563
modelo 3	665.2	7.35	209.129	194.422
modelo 4	667.22	9.38	212.543	193.787
modelo 5	665.22	9.74	212.666	193.179
modelo 6	667.57	18.09	222.750	186.561
modelo 7	675.39	28	231.770	175.768
modelo 8	683.79	35.45	239.466	168.564
modelo 9	664.99	6.31	208.440	195.818
modelo 10	661.54	12.22	215.945	191.512
modelo 11	670.24	19.06	222.441	184.312
modelo 12	662.80	24.56	228.088	178.968
modelo 13	666.15	6.49	209.964	196.979
modelo 14	666.72	11.07	214.519	192.375
modelo 15	672.31	16.12	220.121	187.883
modelo 16	677.15	20.35	225.027	184.317

Observa-se através das tabelas 3.10 a 3.13 que tanto para a variável transformada ($\log(CO)$) quanto para a variável na escala original, o ato de agregar as observações (trabalhar com médias semanais) faz com que os modelos com preditor semi-paramétrico deixem de ser os mais significativos com relação a seus equivalentes paramétricos. O que será feito agora é uma comparação dos resultados alcançados ao se trabalhar com observações diárias e médias semanais da variável transformada e da mesma em sua escala original, seguindo os enfoques clássico e bayesiano. Para isto, foram escolhidos os modelos 9 ($\log(CO) \sim Normal(\mu_t, \sigma^2)$) e 13 ($CO \sim Gama(\alpha, \beta_t)$).

Nas figuras 3.22 e 3.23 são apresentados histogramas das amostras das distribuições a posteriori dos parâmetros de suavização λ para os modelos 9 e 13 com respostas diárias e semanais, onde, no modelo 9, o λ está relacionado à variável temperatura e, no modelo 13, o λ está relacionado à variável umidade. A tabela 3.14 apresenta média a posteriori, mediana a posteriori e intervalos de credibilidade destes parâmetros. Pode-se verificar mais uma vez que, tanto para o modelo 9 ($\log(CO) \sim Normal(\mu_t, \sigma^2)$) quanto para o modelo 13 ($CO \sim Gama(\alpha, \beta_t)$), a agregação das observações faz com que a média a posteriori do parâmetro de suavização aumente.

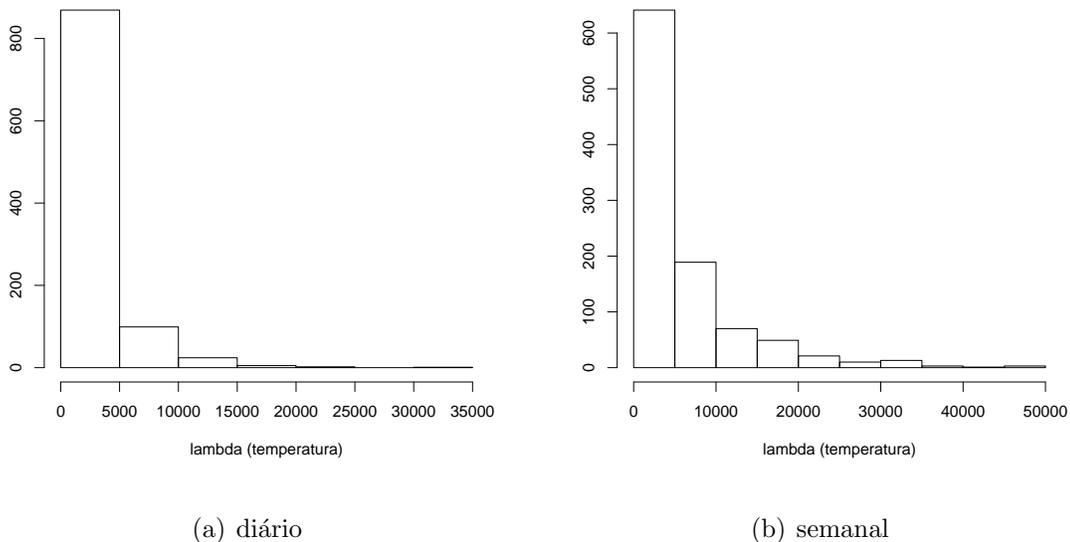


Figura 3.22: Histograma do parâmetro de suavização λ da variável temperatura no modelo 9 utilizando $\log(CO)$ e média semanal de $\log(CO)$ como resposta.

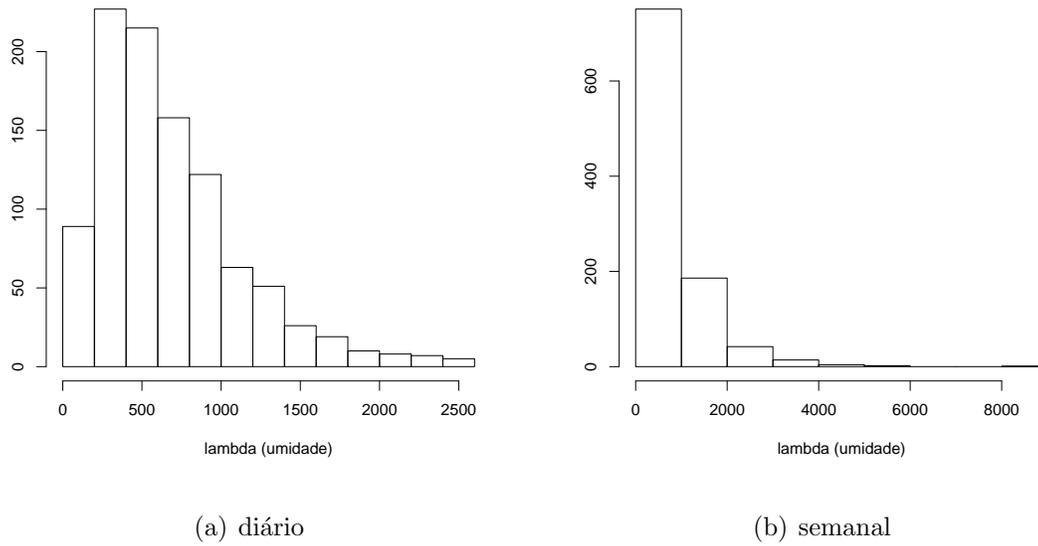


Figura 3.23: Histograma do parâmetro de suavização λ da variável umidade no modelo 13 utilizando CO e média semanal de CO como resposta.

Tabela 3.14: Média, mediana e intervalo de credibilidade do parâmetro de suavização λ para os modelos 9 e 13 com respostas diárias e semanais.

	diário		semanal	
	modelo 9	modelo 13	modelo 9	modelo 13
média	2564.11	678.83	5697.17	765.45
mediana	1581.76	569.83	3053.14	533.78
$IC_{95\%}$	215.65;10750.04	136.11;1866.22	209.75;26535.46	77.98;2843.19

Nas figuras 3.24 e 3.25 são apresentados os histogramas das amostras das distribuições a posteriori dos coeficientes ϕ para os modelos 9 ($\log(CO) \sim Normal(\mu_t, \sigma^2)$) e 13 ($CO \sim Gama(\alpha, \beta_t)$). Gráficos destes coeficientes ϕ , estimados seguindo os enfoques clássico e bayesiano, são mostrados nas figuras 3.26 e 3.27, juntamente com intervalos de confiança / credibilidade de 95%.

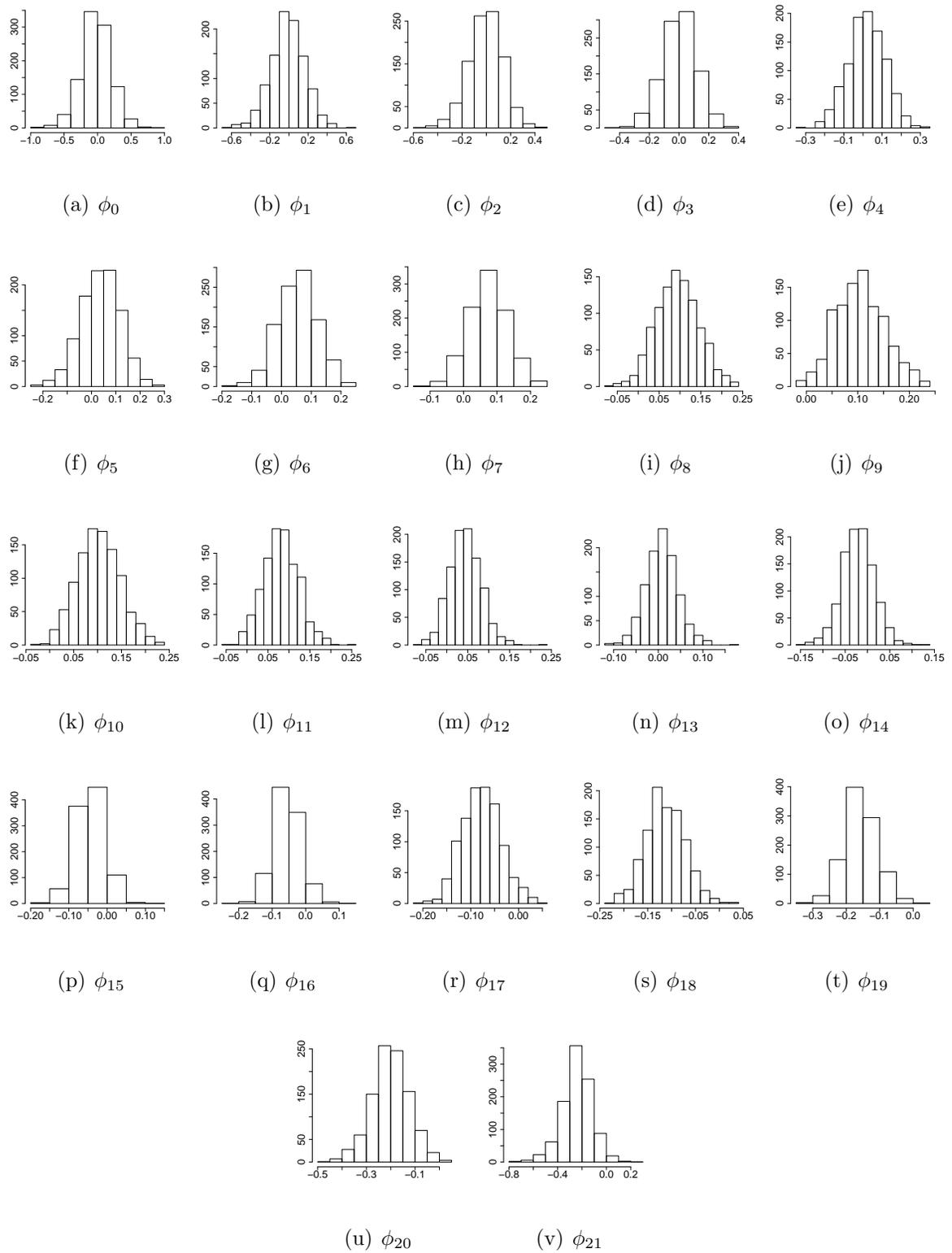


Figura 3.24: Histograma das amostras das distribuições a posteriori dos coeficientes ϕ para $\log CO$ e a variável temperatura.

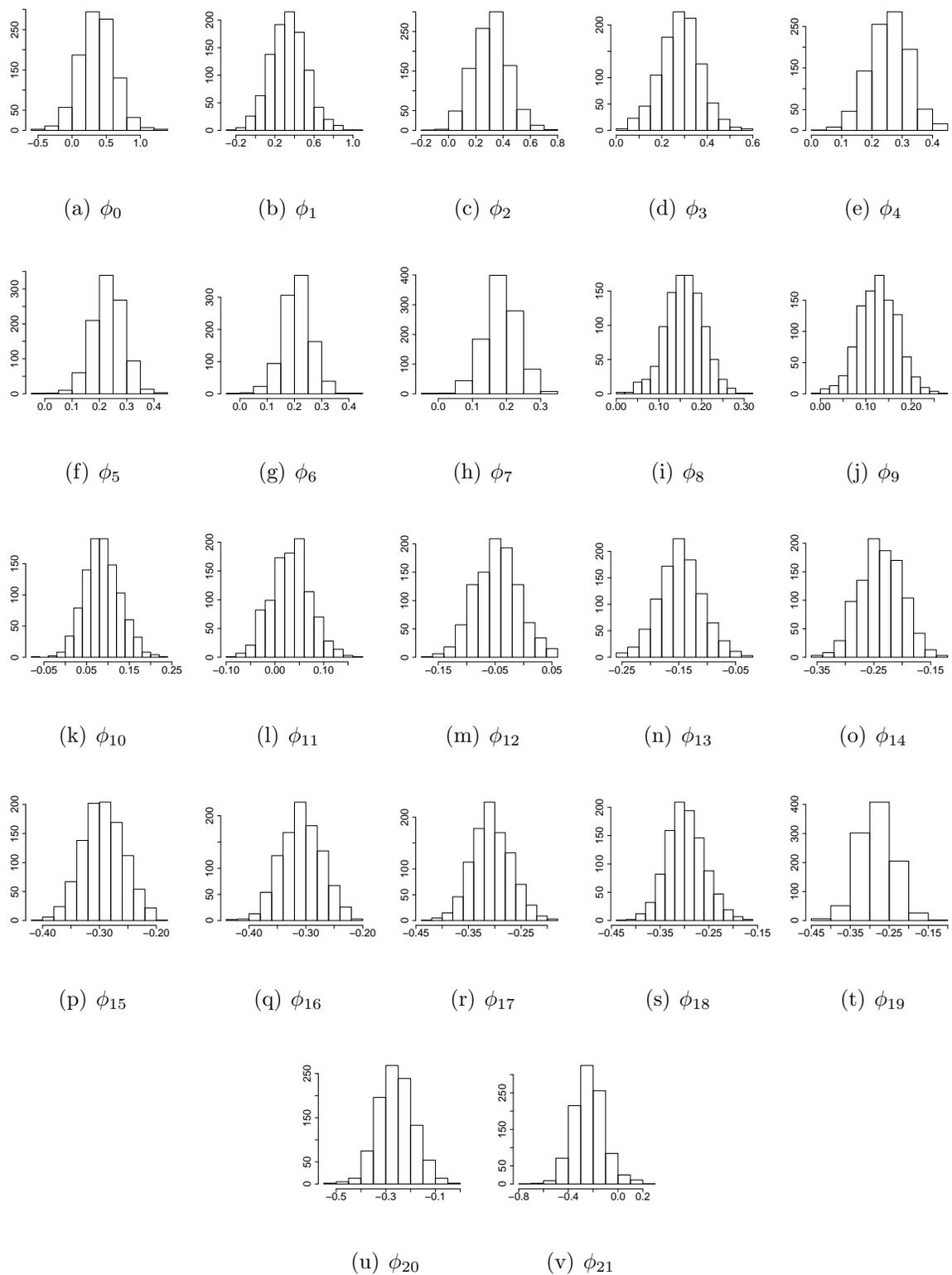
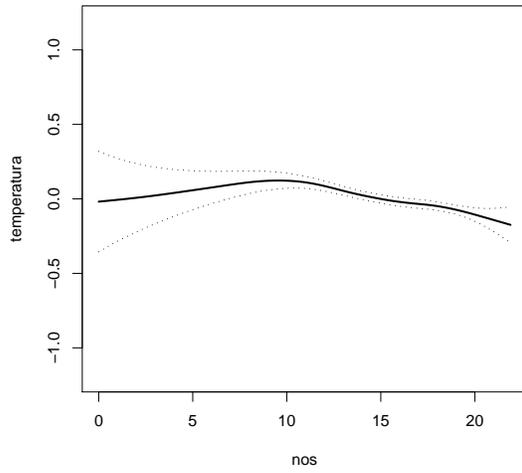
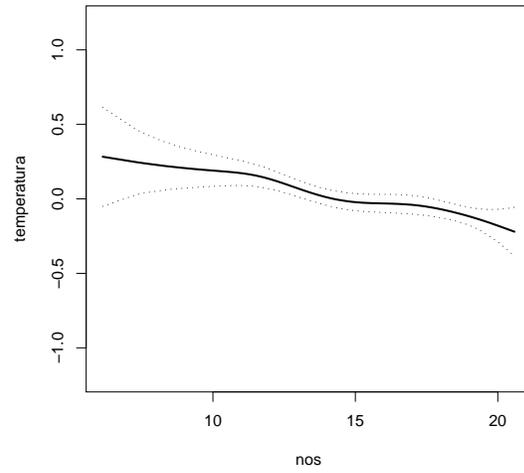


Figura 3.25: Histograma das amostras das distribuições a posteriori dos coeficientes ϕ para CO e a variável unidade.

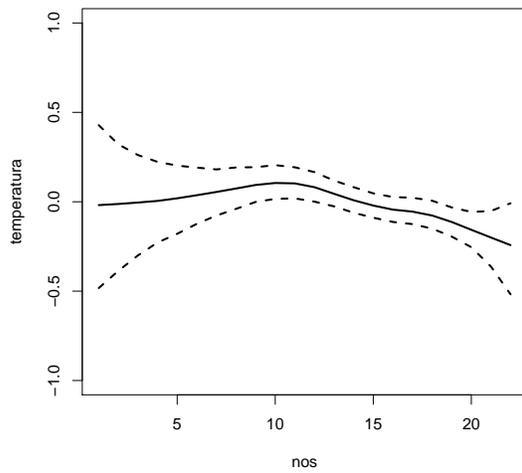
Através das figuras 3.26 e 3.27 a seguir pode-se observar o que foi citado ao comentar a tabela 3.14: a agregação das observações faz com que a média a posteriori do parâmetro de suavização aumente. Além disso, pode-se observar a clara semelhança entre os resultados dos ajustes clássico e bayesiano, chamando atenção para o fato de que, no caso do ajuste clássico, as estimativas se apresentaram mais suaves. No caso especial do modelo 13, utilizando como resposta a média semanal do CO (figura 3.25 (b)), observa-se um comportamento quase linear da função da variável umidade, o que condiz com os resultados apresentados nas tabelas 3.11 e 3.13 que apontavam para o uso de um modelo onde a umidade é incluída de forma paramétrica.



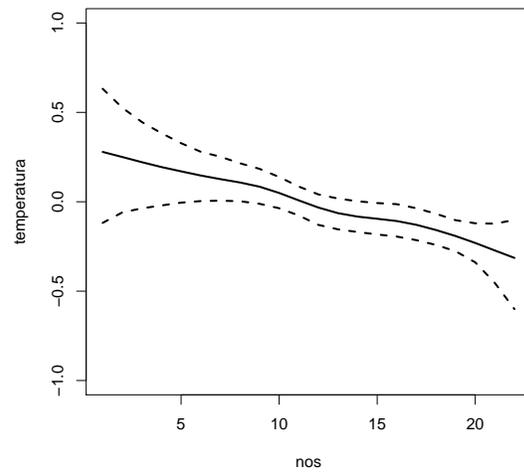
(a) diário (Enfoque clássico)



(b) semanal (Enfoque clássico)

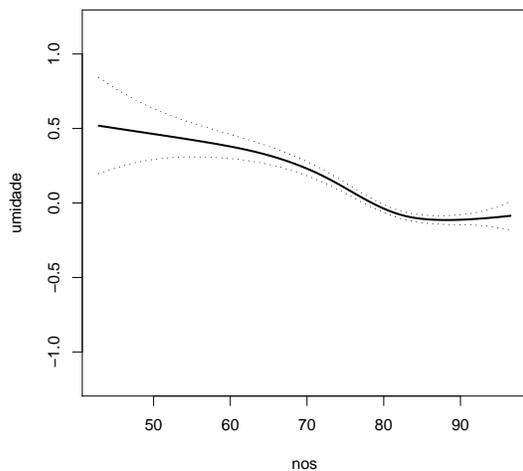


(c) diário (Enfoque bayesiano)

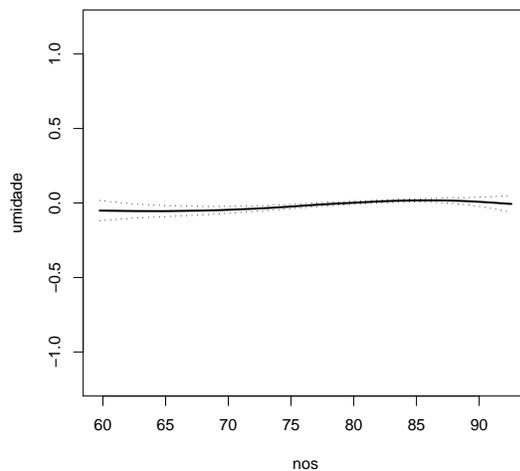


(d) semanal (Enfoque bayesiano)

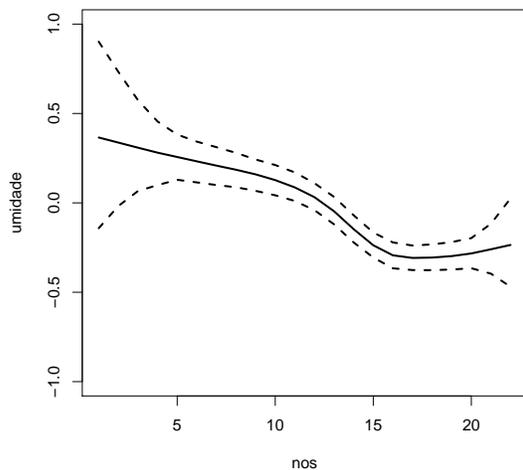
Figura 3.26: Efeito não linear da variável temperatura, estimado seguindo os enfoques clássico e bayesiano, no modelo 9 utilizando $\log(CO)$ e média semanal de $\log(CO)$ como resposta.



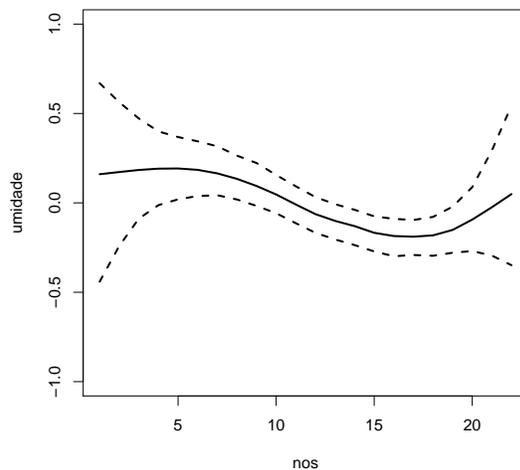
(a) diário (Enfoque clássico)



(b) semanal (Enfoque clássico)



(c) diário (Enfoque bayesiano)



(d) semanal (Enfoque bayesiano)

Figura 3.27: Efeito não linear da variável umidade no modelo 13, estimado seguindo os enfoques clássico e bayesiano, utilizando CO e média semanal de CO como resposta.

Capítulo 4

Considerações Finais

Neste estudo foi analisada uma forma de trabalhar com variáveis que são inseridas no modelo de forma não-paramétrica. Foi visto que a classe dos modelos aditivos generalizados permite trabalhar com diversas respostas com distribuição pertencente à família exponencial e, mais que isso, permite que se inclua no modelo variáveis para as quais não são assumidas hipóteses sobre seu comportamento, ou não se tem definida a forma de impacto sobre o processo estudado. No procedimento de inferência pode-se observar também que existe um paralelo entre as splines e os modelos dinâmicos, visto que a estrutura suave é obtida ao se escolher passeios aleatórios como prioris para os coeficientes das bases destas splines.

Na aplicação feita aos dados de mortalidade em São Paulo verificou-se que a estrutura semi-paramétrica proposta foi bem ajustada, e os modelos de defasagens distribuídas se mostraram bastante eficientes ao estimar a evolução temporal dos efeitos dos poluentes. Foi possível observar também que os efeitos estimados para o poluente CO são maiores e mais duradouros que os estimados para PM_{10} e SO_2 e que, além disso, ao comparar as abordagens clássica e bayesiana, verificou-se que os resultados se mostraram equivalentes em todos os casos. Deve-se salientar que ao utilizar o dióxido de carbono como regressora, os coeficientes / médias a posteriori estimados se apresentaram negativos, e isto está sendo avaliado.

Com relação ao estudo utilizando o monóxido de carbono (resposta contínua), os critérios de comparação divergiram acerca de qual modelo seria o mais indicado segundo

cada um dos enfoques, clássico e bayesiano. Em geral, a estimação clássica apontou para um grupo de modelos que incluísse o maior número testado de defasagens, sendo estas trabalhadas ora de forma paramétrica, ora de forma não-paramétrica, o que poderia introduzir autocorrelação no modelo. Neste caso, uma alternativa é utilizar também aqui a proposta apresentada na seção 1.2.1: adotar restrições polinomiais e verificar se ainda assim estes apresentam os valores mais baixos de AIC. Segundo o enfoque bayesiano, os modelos com menor valor de DIC são aqueles em que a variável explicativa é incluída sem defasagens. Neste caso, as defasagens das variáveis temperatura (modelo 9: $\log(CO) \sim Normal(\mu_t, \sigma^2)$) e umidade (modelo 13: $CO \sim Gama(\alpha, \beta_t)$) admitem suavização quando se trabalha com dados diários, mas um preditor completamente paramétrico é preferido ao modelar a média semanal da resposta.

Referências Bibliográficas

- [1] AKAIKE, H. *A new look at statistical model identification.*, IEEE Transactions on Automatic Control, **AU-19**, 716 - 722, 1974.
- [2] ALMON, S. *The Distributed Lag between Capital Appropriations and Expenditures.*, Econometrica, **33**, 178-196, 1965.
- [3] ALVES, Mariane B. *Funções de Transferência em Modelos Dinâmicos Lineares Generalizados*, Cadernos do IME, Série Estatística, **15**, 2003.
- [4] BARTELS, BEATTY, BARSKY. *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling*, 1987.
- [5] BILLER, C.; FAHRMEIR, L. *Bayesian spline type smoothing in generalized regression models*. Institut für Statistic, Universität München, 1997.
- [6] CHETTY, V. *Estimation of Solow's distributed lag models*. Econometrica, 39, 99 - 117, 1971.
- [7] CRAVEN, P.; WAHBA, G. *Smoothing noisy data with spline functions*. Numer. Math., 31, 377 - 403, 1979.
- [8] de BOOR, C. *A Practical Guide to Splines*. Applied Mathematical Sciences. Springer-Verlag, 1978.
- [9] DOBSON, Annete J. *An Introduction to Generalized Linear Models*. Australia: Chapman & Hall, 1990.

- [10] EILERS, P.H.C.; MARX, B. D. *Flexible smoothing using B-splines and penalized likelihood (with comments and rejoinder)*. *Statistical Science*, 11 (2), 89-121, 1996.
- [11] FAHRMEIR, L.; LANG, S. *Bayesian inference for generalized mixed models based on Markov random field priors*. *Appl. Statist.*, 50, Part 2, 2001-220, 2001.
- [12] FRIEDMAN, J. H.; STUETZLE, W. *Projection pursuit regression*. *J. Amer. Statist. Assoc.*, 76, 817-823, 1981.
- [13] GAMERMAN, D. *Efficient sampling from the posterior distribution in generalized linear models*. *Statistics and Computing*, 7, 57 - 68, 1997.
- [14] GELFAND, A.; GHOSH, S. *Model choice: A minimum posterior predictive loss approach*. *Biometrika*, 85, 1 - 11, 1998.
- [15] GREEN, P. J.; SILVERMAN, B. W. *Nonparametric Regression and Generalized Linear Models: a roughness penalty approach*. Chapman and Hall, London, 1994.
- [16] HASTIE, T.; TIBSHIRANI, R. *Generalized Additive Models*. *Encyclopaedia of Statistical Sciences*, 1990.
- [17] JUNGER, Washington L. *Imputação de dados faltando em séries temporais multivariadas via algoritmo EM*. Projeto final, UERJ. Rio de Janeiro, 2002.
- [18] LANG, S.; BREZGER, A. *Bayesian P-Splines*. *Journal of Computational and Graphical Statistics*, 13, 183-212, 2004.
- [19] McCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. (2ª edição). Chapman & Hall, 1989.
- [20] MIGON, H.; GAMERMAN, D. *Statistical inference: an integrated approach*. Arnold, London UK, 1999.
- [21] RAVINES, R.; SCHMIDT, A. MIGON, H. *Revisiting Distributed Lag Models Through a Bayesian Perspective*. *Applied Stochastic Models in Business and Industry*, Em impressão, 2006.

- [22] SCHWARTZ, Joel. *The Distributed Lag between Air Pollution and Daily Deaths*. Epidemiology, 2000.
- [23] SOLOW, R. *On a family of lag distributions*. Econometrica, 28, 393 - 406, 1960.
- [24] SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P., van der LINDE, A. *Bayesian measures of model complexity and fit (with discussion and rejoinder)*, J. Roy. Statist. Soc., Ser. B, 2002.
- [25] VENABLES, W. N.; RIPLEY, B. D. *Modern Applied Statistics with S-Plus*. Springer-Verlag, 1994.
- [26] WAHBA, G. *Improper Priors, Spline Smoothing and the Problem of Guarding against Model Errors in Regression*, J. R. Stat. Soc. B 40(3), 364-372, 1978.
- [27] ZANOBETTI, A.; SCWARTZ, J.; SAMOLI, E; GRYPARIS, A.; TOULOUMI, G.; ATKINSON, R.; et. al. *The temporal pattern of mortality responses to air pollution: a multicity assessment of mortality displacement*. Epidemiology 13:87-93, 2002.
- [28] WAHBA, G. *Splines in Nonparametric Regression*, Encyclopedia of Environmetrics, 2000.
- [29] http://www.asma-bronquica.com.br/pierre/15POLUICAO_ATMOSFERICA_%20ASMA_URBANA.htm .
- [30] <http://www.estado.com.br/edicao/pano/99/08/10/cid967.html> .
- [31] http://www.feema.rj.gov.br/qualidade_do_ar.htm .
- [32] <http://www.ime.usp.br/jmsinger/>
- [33] <http://www.poluentes.hpg.ig.com.br/mono.html> .
- [34] <http://www.quimica.ufpr.br/ssta/dioxenxof.html> .
- [35] <http://www.r-project.org> .

[36] <http://www.seama.es.gov.br/scripts/sea0501.asp> .

[37] <http://www.stat.uni-muenchen.de/~lang/bayesx> .

Apêndice A

Diferenças divididas

Existem várias formas de se definir diferenças divididas. Será dada aqui a definição segundo de Boor (1978).

Definição: A diferença dividida de l -ésima ordem de uma função f nos pontos $\zeta_k, \dots, \zeta_{k+l}$ é o coeficiente de x^l no polinômio de grau $l + 1$ que coincide com f nestes pontos, e é definida recursivamente por:

$$[\zeta_k, \dots, \zeta_{k+l}]f = \frac{[\zeta_{k+1}, \dots, \zeta_{k+l}]f - [\zeta_k, \dots, \zeta_{k+l-1}]f}{\zeta_{k+l} - \zeta_k}.$$

Por exemplo, para $l = 0, 1$ e 2 tem-se:

$$\begin{aligned} l = 0 & \rightarrow [\zeta_k]f = f(\zeta_k) \\ l = 1 & \rightarrow [\zeta_k, \zeta_{k+1}]f = \frac{f(\zeta_{k+1}) - f(\zeta_k)}{\zeta_{k+1} - \zeta_k} \\ l = 2 & \rightarrow [\zeta_k, \zeta_{k+1}, \zeta_{k+2}]f = \frac{[\zeta_{k+1}, \zeta_{k+2}]f - [\zeta_k, \zeta_{k+1}]f}{\zeta_{k+2} - \zeta_k}. \end{aligned}$$

Serão listadas a seguir duas consequências vindas da definição anterior:

- (i) $[\zeta_k, \dots, \zeta_{k+l}]f$ é uma função simétrica de seus argumentos $\zeta_k, \dots, \zeta_{k+l}$, ou seja, depende somente dos valores $\zeta_k, \dots, \zeta_{k+l}$ e não da ordem em que cada um deles aparece na sequência. Isto fica claro a partir da definição, já que o polinômio de interpolação depende somente dos pontos de interpolação e não da ordem que se escreve estes pontos.

(ii) se f é um polinômio de grau $\leq l$, então $[\zeta_k, \dots, \zeta_{k+l}]f$ é constante como função de $\zeta_k, \dots, \zeta_{k+l}$. Em particular,

$$[\zeta_k, \dots, \zeta_{k+l}]f = 0 \quad \forall f \text{ polinômio de grau } l$$

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)