



**MARCO HENRIQUE  
VIEIRA MARTO**

**ESTIMAÇÃO CLÁSSICA E BAYESIANA DE  
PARÂMETROS DO MODELO GLOBULUS 2.1**





**MARCO HENRIQUE  
VIEIRA MARTO**

**ESTIMAÇÃO CLÁSSICA E BAYESIANA DE  
PARÂMETROS DO MODELO GLOBULUS 2.1**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Matemática e Aplicações, realizada sob a orientação científica da Doutora Isabel Maria Simões Pereira, Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro.



Dedico este trabalho aos meus Pais e a todos os que me ensinaram a nunca desistir dos sonhos e objectivos sem no mínimo tentar alcançá-los e no máximo concretizá-los.



## **o júri**

presidente

Prof. Doutor Vasile Staicu  
Professor Catedrático do Departamento de Matemática da Universidade de Aveiro

arguente

Prof<sup>a</sup>. Doutora Maria Margarida Branco de Brito Tavares Tomé  
Professora Catedrática do Instituto Superior de Agronomia da Universidade Técnica de Lisboa

orientadora

Prof<sup>a</sup>. Doutora Isabel Maria Simões Pereira  
Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro





## **agradecimentos**

Agradeço o incansável apoio e disponibilidade da Professora Doutora Isabel Pereira (orientadora) e todo o apoio, a disponibilização de dados e opiniões para a realização deste trabalho à Professora Doutora Margarida Tomé. Agradeço também à ALTRI a permissão para poder usar os dados do trabalho.



**palavras-chave**

Estatística Clássica, Estatística Bayesiana, Métodos de Monte Carlo via Cadeias de Markov, Métodos de Regressão Não-Linear.

**resumo**

Este trabalho pretende expôr e contrastar duas perspectivas diferentes de abordar o problema de estimação de parâmetros de modelos Biométricos de altura dominante, área basal e mortalidade para a espécie florestal de *Eucalyptus globulus*, considerando apenas a primeira rotação. As perspectivas são: a metodologia Clássica, já antes utilizada para este estudo e a metodologia Bayesiana, em que este trabalho pretende ser pioneiro em Portugal. Os resultados finais das estimações para cada uma das abordagens são comparáveis, uma vez que se usaram distribuições *a priori* não informativas propostas por Box-Tiao na perspectiva Bayesiana. Obtiveram-se resultados parecidos pelas duas perspectivas, excepto para o caso da estimação dos parâmetros da Mortalidade, em que a metodologia Bayesiana conduziu a um modelo final diferente do obtido pelos métodos clássicos.



**keywords**

Classic Statistics, Bayesian Statistics, Markov Chain Monte Carlo, Non-Linear Regression Methods.

**abstract**

This work pretends to contrast two different perspectives of parameters estimation problem for dominant height, basal area and mortality Biometric Models of *Eucalyptus globules* in first rotation. The two perspectives are: Classical methodology, used before for this kind of study and Bayesian methodology that this work pretends to be the first in Portugal. Final results of parameters estimation for each different perspective are comparable because I used non informative priors proposed by Box-Tiao in Bayesian perspective and were similar for the two perspectives, except for the Mortality parameters because Bayesian methodology drove to a different final model.



ÍNDICE

1	Introdução .....	7
2	Enquadramento florestal do modelo a desenvolver .....	9
2.1	Estrutura do Modelo.....	12
2.2	Compatibilidade entre as Funções de Crescimento e de Inicialização.....	15
2.3	Definição do Modelo Matemático a ajustar .....	20
2.3.1	Altura dominante.....	20
2.3.2	Mortalidade .....	20
2.3.3	Área basal .....	21
2.4	Noções Preliminares de Sistemas de Equações .....	23
2.4.1	Classificações de Sistemas de Equações .....	24
3	Estimação do Modelo com Recurso a Técnicas de Estatística Clássica.....	29
3.1	Métodos de Estimação .....	30
3.1.1	Mínimos Quadrados Não Linear .....	30
3.1.2	Regressão Não Linear Aparentemente Não Relacionada .....	38
3.2	Aplicação ao Modelo: Resultados da Estimação Clássica .....	40
3.2.1	Estudo da Identificabilidade do Sistema de Equações Simultâneas.....	41
3.2.2	Estatísticas Descritivas Sobre os Dados Utilizados para a Estimação .....	41
3.2.3	Estimativas Clássicas dos Parâmetros do Modelo .....	43
4	Estimação do Modelo com Recurso a Técnicas de Estatística Bayesiana .....	55
4.1	Princípios de Estatística Bayesiana.....	55
4.1.1	Introdução .....	55
4.1.2	Distribuições <i>a priori</i> .....	57
4.1.3	Inferência Paramétrica.....	62
4.1.4	Fundamentos de Métodos de Monte Carlo .....	64
4.1.5	Critério DIC para a seleção de modelos .....	71
4.2	Aplicação ao Modelo: Resultados da Estimação Bayesiana.....	73
4.2.1	Desenvolvimento da Metodologia Bayesiana .....	74
4.2.2	Estimativas Bayesianas para o Modelo .....	85
5	Análise de Resultados Clássicos e Bayesianos .....	89
6	Conclusões.....	95

ANEXO I - Evolução em SAS do Modelo de Equações Simultâneas da mortalidade e área basal (metodologia clássica).....	99
ANEXO II - Avaliação da Convergência e Gráficos das Distribuições para os Modelos Bayesianos.....	119
ANEXO III - Exemplos de Macros usadas para VBA-EXCEL.....	125
Bibliografia.....	143



LISTA DE FIGURAS

Figura 1: Estrutura do modelo <i>GLOBULUS 2.1</i> . (de Tomé <i>et al.</i> , 2001).....	14
Figura 2: Ajustamento da função de Lundqvist aos dados da evolução em área basal e altura dominante de uma parcela permanente de eucalipto na zona centro litoral (de Tomé <i>et al.</i> , 2001).....	16
Figura 3: Dados da evolução em área basal e altura dominante de um conjunto de parcelas permanentes de eucalipto na zona centro de Portugal (de Tomé <i>et al.</i> , 2001) .....	16
Figura 4: Distribuições proponentes do modelo da altura dominante.....	76
Figura 5: Distribuições proponentes do modelo da área basal.....	81
Figura 6: Distribuições proponentes do modelo inicial da mortalidade.....	84
Figura 7: Gráfico de perfil das estimativas Clássicas e Bayesianas.....	93



LISTA DE TABELAS

Tabela 1: Variáveis incluídas no modelo <i>GLOBULUS 2.1</i> .....	14
Tabela 2: Estatísticas descritivas dos dados utilizados para a estimação (n = 1477).	42
Tabela 3: Estimativas dos parâmetros do modelo da altura dominante pelo método NLLS.....	43
Tabela 4: Estimativas das covariâncias dos parâmetros da equação da altura dominante.....	44
Tabela 5: Estimativas das correlações dos parâmetros da equação da altura dominante.....	45
Tabela 6: Estimativas dos parâmetros do modelo da mortalidade e da área basal pelo método NSUR.....	47
Tabela 7: Correlações de Pearson entre os pseudo-regressores da equação da mortalidade.....	49
Tabela 8: Correlações de Pearson entre os pseudo-regressores da equação da área basal.....	51
Tabela 9: Estimativas das covariâncias dos resíduos estimados das equações.....	51
Tabela 10: Estimativas das correlações dos resíduos estimados das equações.....	52
Tabela 11: Estimativas das covariâncias dos parâmetros do sistema de equações..	53
Tabela 12: Estimativas das correlações dos parâmetros do sistema de equações...	54
Tabela 13: Estimativas e algumas estatísticas descritivas de localização e escala.....	86
Tabela 14: Estimativas e algumas estatísticas descritivas de localização e escala.....	86
Tabela 15: Estimativas e algumas estatísticas descritivas de localização e escala.....	87
Tabela 16: Estimativas e algumas estatísticas descritivas de localização e escala.....	88
Tabela 17: Estimativas e algumas estatísticas descritivas de localização e escala.....	88
Tabela 18: Estimativas segundo as duas abordagens para a altura dominante.....	89
Tabela 19: Estimativas segundo as duas abordagens para a área basal.....	90
Tabela 20: Estimativas segundo as duas abordagens para a mortalidade.....	92



## CAPITULO 1

### 1 Introdução

O objetivo deste trabalho é comparar duas metodologias distintas para abordar o mesmo problema estatístico, o problema da estimação de parâmetros de modelos florestais, neste caso de eucalipto e para as variáveis: altura dominante, área basal e mortalidade.

As metodologias utilizadas foram a Clássica, já antes bastante desenvolvida e explorada em Portugal, nomeadamente pelo Centro de Estudos Florestais do Instituto Superior de Agronomia da Universidade Técnica de Lisboa (do qual fiz parte como programador por um período superior a 1,5 anos) e a Bayesiana que constitui o grande desafio deste trabalho, nunca antes referenciada em Portugal (para este problema específico), muito distinta da Clássica, embora comparável em algumas situações, nomeadamente quando se utilizam distribuições a priori não informativas e os modelos resultantes contêm exatamente os mesmos parâmetros.

De acordo com o objetivo do trabalho, no capítulo 2 serão definidos os modelos matemáticos, explicam-se as suas origens, assim como se faz uma introdução à teoria dos Sistemas de Equações de Regressão, no capítulo 3 apresenta-se toda a teoria necessária e a estimação dos parâmetros dos modelos segundo a perspectiva Clássica, no capítulo 4 expõe-se toda a teoria necessária e a estimação dos parâmetros dos modelos segundo a perspectiva Bayesiana, para no capítulo 5 se fazer uma análise comparativa dos resultados obtidos pelas duas abordagens e por fim se apresentar alguma reflexão sobre o estudo efectuado nas Conclusões do trabalho.

Nos anexos encontram-se os resultados e a evolução dos modelos estimados pela abordagem Clássica, alguns resultados da estimação importantes da abordagem

Bayesiana que não foram incluídos no texto para o não sobrecarregar e alguns exemplos de programas em “VBA - EXCEL” onde insidiu a maioria do meu intenso esforço computacional neste trabalho.

## CAPITULO 2

### 2 Enquadramento florestal do modelo a desenvolver

Em Portugal têm sido feitos muitos estudos de modelação do crescimento da espécie *Eucalyptus globulus*, mais conhecido correntemente por Eucalipto desde a instalação das primeiras parcelas permanentes (instaladas pela Celbi) em 1971 criando-se também as primeiras tabelas de produção com base nos dados dos primeiros anos de medição destas parcelas permanentes.

Em 1995 a indústria e o ISA/UTL iniciaram um projeto de colaboração tendo como principais objetivos:

- Organizar num formato digital compatível todos os dados disponíveis em Portugal sobre o crescimento do eucalipto;
- Com recurso aos dados, desenvolver um bom modelo de crescimento para as plantações de eucalipto em Portugal, parametrizado para diferentes regiões, podendo ser utilizado para planeamento operacional ao nível do povoamento ou de uma empresa, estratégico a nível regional ou mesmo nacional e no apoio a decisões de política florestal;
- Detectar a necessidade da obtenção de dados adicionais ou da instalação de novos ensaios.

O modelo resultante deste projeto, o modelo *GLOBULUS*, tem sido desde então constantemente melhorado em sucessivas versões, sendo a versão em que inside este estudo designada por *GLOBULUS* 2.1.

Comparativamente com a 1ª versão e os modelos anteriores, o modelo *GLOBULUS 2.1* apresenta as seguintes melhorias:

- Encontra-se baseado na classificação prévia do país em zonas homogêneas com base em variáveis climáticas;
- Para o seu desenvolvimento utilizaram-se dados de todo o país, com uma representatividade razoável das diversas regiões em que o país foi classificado, pelo menos para os povoamentos em 1ª rotação, com exceção da região Norte litoral, na qual apenas se dispõe de dados de um ensaio de compassos e mais um número muito pequeno de observações em plantações industriais. No âmbito do projeto acima referido e para a organização destes dados desenvolveu-se e implementou-se, em "Microsoft ACCESS", a base de dados "sibFlor-Eg";
- Foi feita uma análise da necessidade de parametrização de acordo com a localização do povoamento numa das oito regiões definidas, assim como para os povoamentos em talhadia, testando-se ainda a interação do efeito da talhadia com a localização do povoamento (apenas Norte, contra Sul e interior);
- Desenvolveu-se uma metodologia para a obtenção de modelos compatíveis para a inicialização e predição do crescimento em área basal ( $dap > 5$ ). Por  $dap$  entende-se "diâmetro à altura do peito" (diâmetro do tronco a 1,30m de altura);
- Desenvolveu-se uma metodologia para a obtenção de modelos compatíveis para a inicialização e predição do crescimento em volume do povoamento, por sua vez compatíveis com a área basal e a altura dominante do povoamento;
- Ajustaram-se não só modelos para a predição do volume total ( $dap > 5$ ) com casca, mas também para a predição do correspondente volume total sem casca, assim como de volumes mercantis (com e sem casca) para despontas entre 5 e 10 cm;
- Analisou-se a variabilidade da produtividade dos eucaliptais em Portugal, no todo e em cada uma das regiões definidas. Esta análise é essencial para prever o crescimento potencial de um povoamento de eucalipto de acordo com a sua



localização e com uma classificação subjetiva da estação, relativamente a outras da mesma região, em cinco classes de qualidade: muito alta, alta, média, baixa e muito baixa.

Em relação ao modelo *GLOBULUS 2.0*, o presente modelo apresenta ainda as seguintes alterações/melhorias:

- Realizou-se uma validação adicional dos dados, em consequência da qual se procedeu a algumas correções adicionais;
- Foram acrescentados à base de dados os novos dados provenientes dos ensaios cuja medição é da responsabilidade do ISA/UTL;
- Desenvolveu-se uma redefinição das variáveis indicatrizes das regiões climáticas, de modo a que o valor 1 de cada variável indicatriz corresponda sempre à região de maior produtividade;
- Durante a aplicação do modelo *GLOBULUS 2.0* detectaram-se casos em que os valores observados da altura dominante e especialmente da área basal eram superiores ao valor da assíntota (ver a secção 2.2), facto que levava a estimativas de crescimento decrescentes; por outro lado, a forma do feiche de curvas da altura dominante do modelo *GLOBULUS 2.0* não parecia razoável (as diferentes curvas aproximavam-se excessivamente umas das outras para idades superiores a 20 anos). Assim, optou-se por analisar modelos alternativos que correspondam a valores de assíntotas mais elevados;
- No âmbito do modelo *GLOBULUS 2.1* desenvolveu-se uma metodologia que permite substituir, na função de projecção em área basal, a utilização do número de árvores à plantação / número de varas após monda (consoante se trate de um povoamento em 1ª rotação ou em talhadia) pelo número de árvores vivas em cada idade; por um lado, é difícil avaliar no campo o número de árvores após monda, por outro lado, a utilização do número de árvores vivas em cada idade permitirá introduzir alguma correção à predição do crescimento em povoamentos com bastante mortalidade;
- No modelo *GLOBULUS 2.0* a equação para a simulação da mortalidade era bastante simplificada, não incluindo a influência do índice de qualidade da estação, da

densidade do povoamento ou da região, pelo que se desenvolveu um modelo melhorado;

- Manteve-se o sistema de equações para a predição e projeção de volumes totais com e sem casca e de volumes mercantis que haviam sido ajustados para o modelo GLOBULUS 2.0;
- Desenvolveram-se sub-modelos para a estimação das biomassas totais e por componentes, o que implicou, numa primeira etapa, o ajustamento de equações ao nível da árvore e, numa segunda fase, o ajustamento de um sistema de equações para a predição das biomassas totais e por componentes ao nível do povoamento;
- Com base nas estimativas de biomassa totais e por componentes e nos dados sobre a composição química das diversas componentes de biomassa publicadas, o modelo GLOBULUS 2.1 fornece estimativas de carbono e macro-nutrientes (N, P, K e Ca) acumulados, ao longo do tempo, em cada uma das componentes da biomassa.

## 2.1 Estrutura do Modelo

O modelo GLOBULUS 2.1 é um modelo de povoamento que considera as seguintes variáveis de estado: altura dominante, área basal, volume total e volume mercantil para limites de despona entre 5 e 10 cm (com e sem casca), número de árvores por hectare (ha) e biomassa total e por componentes. O modelo estima também o total de carbono e nutrientes (N, P, K, Ca) acumulados em cada componente. O índice de qualidade da estação, o número de árvores por ha à plantação, o número de varas por ha após a monda de varas, a rotação (alto fuste *versus* talhadia) e a localização numa das 8 regiões definidas são as variáveis de controlo. A rotação foi definida com base numa variável indicatriz da talhadia e a região de acordo com a classificação climática explicada em Tomé et al. (2001). A Tabela 1 apresenta um resumo das variáveis incluídas no modelo.

A estrutura do modelo está representada na Figura 1. Como se pode ver, o modelo considera dois módulos, o módulo de inicialização e o módulo de projeção.

O módulo de projeção é constituído por um conjunto de equações que actualizam o valor das variáveis de estado do instante  $t_1$  para o instante  $t_2$ . No modelo *GLOBULUS 2.1* o módulo de projeção é constituído por funções de crescimento, formuladas como equações às diferenças, e por equações de predição. A maior parte das equações deste módulo, com exceção das equações para a predição dos volumes mercantis e das biomassas total e por componentes, são equações às diferenças, ou seja, equações que estimam o valor de uma variável de estado no instante  $t_2$  em função dos valores das variáveis de controlo e dos valores das variáveis de estado no instante  $t_1$ . Por seu lado, uma equação de predição estima o valor das variáveis de estado num determinado instante  $t$ , função do valor das variáveis de controlo e do valor de outras variáveis de estado no mesmo instante  $t$ . Os volumes mercantis são estimados com base numa equação que estima a proporção do correspondente volume (com ou sem casca) até ao limite de despona, enquanto que as biomassas das diferentes componentes são estimadas com base num sistema de equações alométricas. A projeção, até à idade de corte, de um povoamento que tenha sido objeto de um inventário florestal não necessita de nenhum outro módulo.

No caso de se pretender simular o crescimento de um povoamento no qual não tenha sido realizado nenhum inventário florestal, ou simular o crescimento de um povoamento após o corte (2ª rotação ou superiores) é necessário estimar as condições iniciais para a simulação, função do valor das variáveis de controlo. É esta a função do módulo de inicialização, o qual é constituído apenas por equações de predição.

Variáveis de controlo	Variáveis de estado	
	Variáveis principais	Variáveis secundárias
Ambientais: Índice de qualidade da estação (idade padrão 10) Região climática  Culturais: Densidade à plantação (1ª rotação) Densidade após monda (talhadia) Idade de corte Povoamento: Rotação Idade	Altura dominante Número de árvores ha <sup>-1</sup> Área basal do povoamento Volume total (com e sem casca)	Volumes mercantis (com e sem casca) Biomassa total (parte aérea) Biomassa de raízes Biomassa por componentes: ramos, folhas, madeira, casca Armazenamento de carbono por componentes: ramos, folhas, madeira, casca e raízes

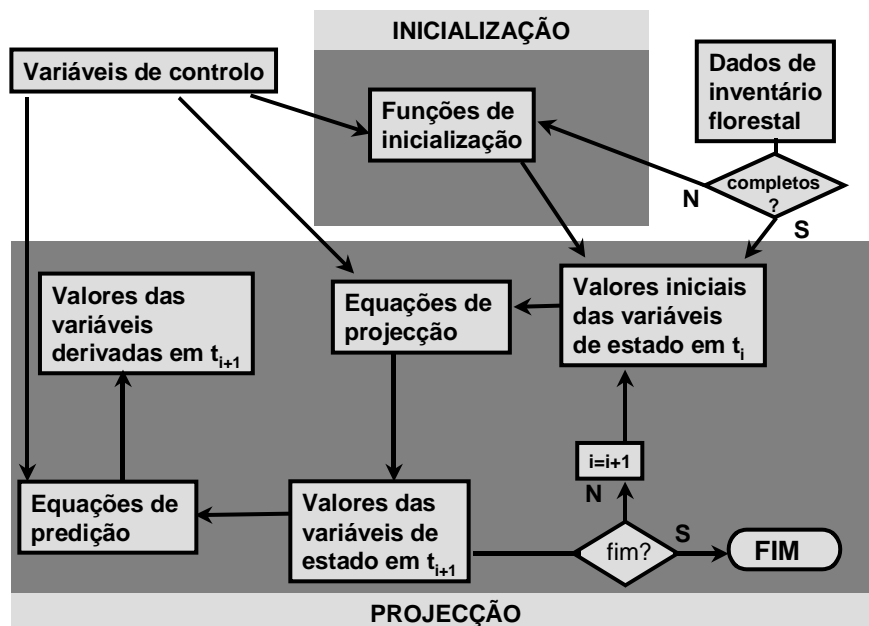


Figura 1: Estrutura do modelo *GLOBULUS 2.1*. (de Tomé *et al.*, 2001)

A compatibilidade entre os dois módulos está assegurada. Ou seja, para os mesmos valores das variáveis de controlo, obtém-se as mesmas estimativas para as variáveis de

estado na idade  $t_2$ , independentemente da idade inicial. Foi desenvolvida, no âmbito do desenvolvimento do modelo *GLOBULUS 2.0*, a metodologia para conseguir esta característica, a qual se considera essencial para garantir uma utilização do modelo na prática. No modelo *GLOBULUS 2.1* desenvolveu-se uma versão melhorada do modelo da área basal então deduzido, o qual permite a expressão da taxa de crescimento em função da densidade do povoamento em vez da densidade de plantação. Esta metodologia é explicada em detalhe no ponto seguinte.

## **2.2 Compatibilidade entre as Funções de Crescimento e de Inicialização**

As funções a utilizar na modelação do crescimento biológico, quer sejam de natureza empírica ou deduzidas a partir de princípios biológicos, devem apresentar uma forma que esteja de acordo com os princípios do crescimento biológico:

- a curva é limitada pela produção zero no início ( $t=0$  ou  $t=t_0$ , consoante a variável em questão toma valores maiores que 0 em  $t=0$  ou mais tarde) e por uma produção máxima finita atingida em idade avançada (existência de assíntota);
- a taxa relativa de crescimento (variação da variável  $x$  por unidade de tempo e de  $x$ ) apresenta um máximo na fase inicial, após o qual é decrescente; na maior parte dos casos, o máximo ocorre suficientemente cedo para podermos utilizar funções decrescentes na sua modelação;
- o declive da curva aumenta com produções crescentes na fase inicial e decresce na fase final (tem um ponto de inflexão).

Como funções de crescimento empíricas têm sido utilizadas diversas, entre as quais a função de Lundqvist:

$$y = Ae^{-kt^{-n}}$$

onde o parâmetro  $A$  representa a assíntota superior do crescimento e o parâmetro  $k$  está relacionado com o declive da curva (taxa de crescimento).

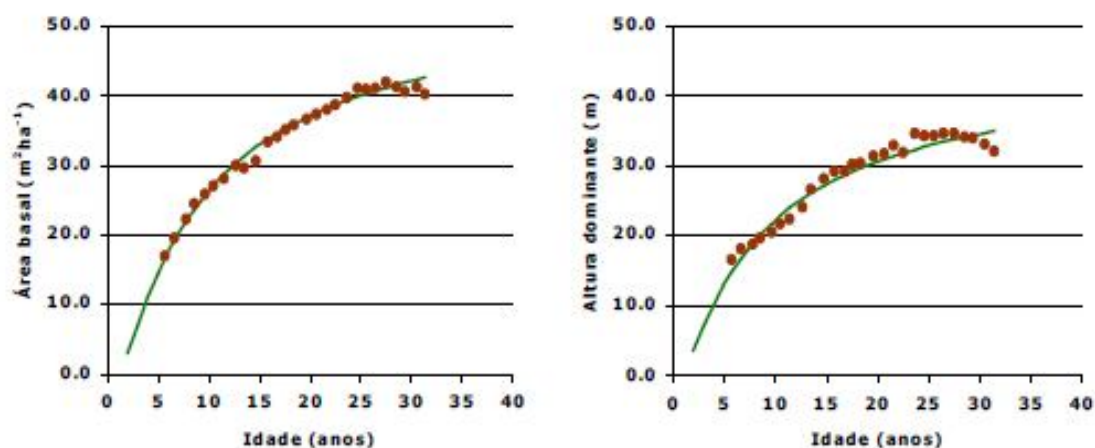


Figura 2: Ajustamento da função de Lundqvist aos dados da evolução em área basal e altura dominante de uma parcela permanente de eucalipto na zona centro litoral (de Tomé *et al.*, 2001).

Mas no desenvolvimento de um modelo de crescimento, o objetivo é modelar, não apenas os dados de uma parcela, mas um conjunto de dados de várias parcelas permanentes.

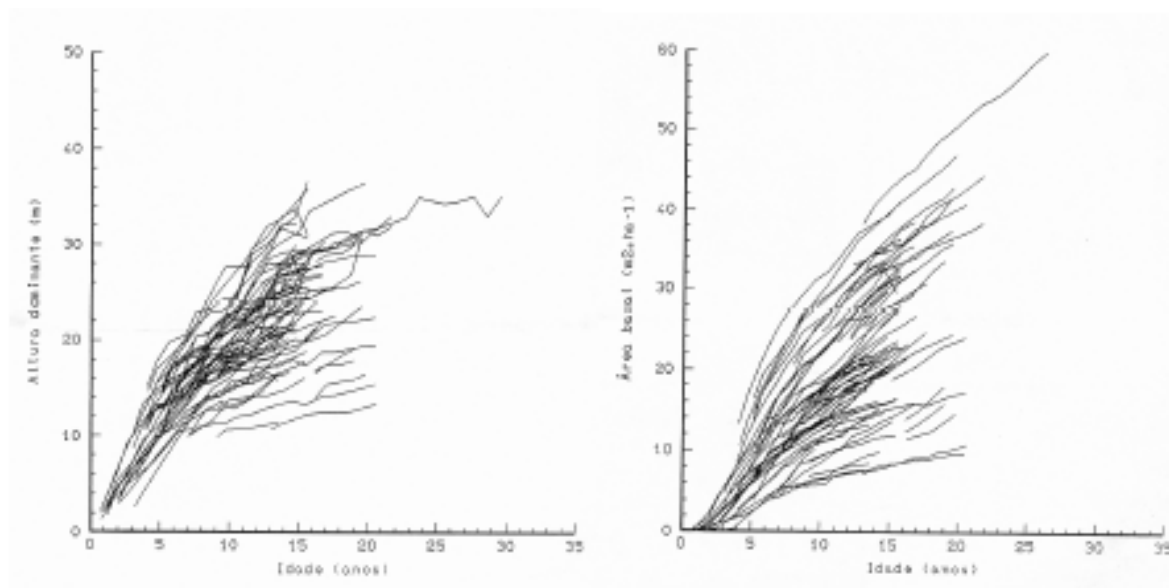


Figura 3: Dados da evolução em área basal e altura dominante de um conjunto de parcelas permanentes de eucalipto na zona centro de Portugal (de Tomé *et al.*, 2001).

Um dos métodos possíveis para modelar simultaneamente o crescimento de um conjunto de parcelas permanentes é através de funções de crescimento formuladas como equações às diferenças. É precisamente esse método que foi usado no modelo GLOBULUS 2.1. A expressão de uma função de crescimento como uma equação às diferenças baseia-se na premissa de que é possível modelar o crescimento com base numa família de curvas correspondente às curvas originadas por essa função de crescimento com todos os parâmetros comuns exceto um, o chamado parâmetro livre. Exemplificando com a função de Lundqvist, temos a seguinte expressão para a função de crescimento:

$$y = Ae^{-kt^{-n}}$$

A partir da função de crescimento nos instantes  $t_1$  e  $t_2$  obtém-se, respectivamente:

$$y_1 = Ae^{-kt_1^{-n}} \Rightarrow \ln(y_1) = \ln(A) - k t_1^{-n} \Rightarrow \ln\left(\frac{A}{y_1}\right) = k t_1^{-n} \Rightarrow k = t_1^n \ln\left(\frac{A}{y_1}\right) \Rightarrow k = \ln\left(\frac{A}{y_1}\right)^{t_1^n}$$

$$y_2 = Ae^{-kt_2^{-n}} \Rightarrow \ln(y_2) = \ln(A) - k t_2^{-n} \Rightarrow \ln\left(\frac{A}{y_2}\right) = k t_2^{-n} \Rightarrow k = t_2^n \ln\left(\frac{A}{y_2}\right) \Rightarrow k = \ln\left(\frac{A}{y_2}\right)^{t_2^n}$$

Igualando agora as duas expressões obtidas para k, vem:

$$-\ln\left(\frac{y_2}{A}\right)^{t_2^n} = -\ln\left(\frac{y_1}{A}\right)^{t_1^n} \Rightarrow \left(\frac{y_2}{A}\right)^{-t_2^n} = \left(\frac{y_1}{A}\right)^{-t_1^n} \Rightarrow y_2 = A\left(\frac{y_1}{A}\right)^{\left(\frac{t_1}{t_2}\right)^n},$$

expressão que corresponde à equação de Lundqvist com o parâmetro k como parâmetro livre, geralmente designada por Lundqvist-k. A família de curvas é definida, para esta equação às diferenças, pelos parâmetros A e n. Um par  $(t_0, y_0)$ , que permite calcular o parâmetro k de acordo com

$$k = \ln\left(\frac{A}{y_0}\right)^{\frac{1}{t_0^n}}$$

é a condição inicial, a qual expressa a forma de uma função da família de curvas definida pela equação às diferenças. Do mesmo modo se podem obter, para cada

função de crescimento, tantas formulações como equações às diferenças quantos os parâmetros dessa função de crescimento.

Os parâmetros  $k$  e  $n$  da função de Lundqvist estão relacionados com a velocidade de crescimento. Sendo o primeiro dado pela condição inicial  $(t_0, y_0)$ , faz todo o sentido que o parâmetro  $n$  esteja relacionado com a densidade do povoamento. Se se assumir uma relação linear entre o parâmetro  $n$  e a densidade do povoamento, o parâmetro  $k$  pode ser expresso do seguinte modo, respectivamente para os instantes  $t_1$  e  $t_2$ :

$$k_{t_1} = \ln \left( \frac{A}{y_1} \right)_{t_1}^{\frac{1}{n_0 + n_n \frac{N_1}{1000}}}$$

$$k_{t_2} = \ln \left( \frac{A}{y_2} \right)_{t_2}^{\frac{1}{n_0 + n_n \frac{N_2}{1000}}}$$

O modelo resultante é:

$$y_2 = A \left( \frac{y_1}{A} \right)_{t_2}^{\frac{t_1}{n_0 + n_n \frac{N_1}{1000}}}$$

A inicialização de uma equação às diferenças para um povoamento particular corresponde à definição da respetiva condição inicial, a qual é essencial para definir a curva da família que corresponde a esse povoamento. Se o povoamento tiver sido objeto de um inventário florestal, a condição inicial é definida com base no resultado deste inventário. Caso contrário, é necessário encontrar uma equação de predição que permita estimar o valor  $y_0$ , em função dos valores das variáveis de controlo e da idade  $t_0$ . Para se garantir a compatibilidade entre esta equação de predição e a equação às diferenças é necessário que a expressão matemática da equação de predição corresponda à função de crescimento que está na base da equação às diferenças. Os parâmetros comuns da equação às diferenças serão também comuns à função de inicialização, sendo o parâmetro livre desta expresso em função dos valores das variáveis de controlo.

Exemplificando com uma das componentes do modelo *GLOBULUS 2.1*, a área basal para os povoamentos de 1ª rotação localizados na região Norte Interior, 6NI. O modelo



selecionado para simular o crescimento em área basal foi a equação Lundqvist-k com a assíntota A expressa em função do índice de qualidade da estação e com o parâmetro de forma n expresso em função do número de árvores por ha à plantação:

$$G_2 = A_g \left( \frac{G_1}{A_g} \right)^{\frac{t_1^{n_{g0} + n_{gn}} \frac{N_1}{1000}}{t_2^{n_{g0} + n_{gn}} \frac{N_2}{1000}}}, \text{ com}$$

$$A_g = 0,1586 Iqe^2$$

$$n_{g0} = 3,893 - 1,0288Iqe$$

$$n_{gn} = 0,1024$$

A função de inicialização foi então obtida por ajustamento da função de Lundqvist na sua forma integral, fixando-se A e n nas expressões acima e estimando o parâmetro k como uma função das variáveis de controlo. Obteve-se então:

$$G = A_g e^{\left( -k_g \frac{1}{t} \right)^{n_{g0} + n_{gn}} \frac{N}{1000}}, \text{ com}$$

$A_g$ ,  $n_{g0}$  e  $n_{gn}$  dados pelas expressões acima e

$$k = -5,5311 + 177,9 \frac{1}{Iqe} + 0,5409 \frac{Npl}{1000} + 16,0146 \frac{100}{Iqe \sqrt{N}}$$

Na metodologia utilizada no ajustamento das funções de crescimento e inicialização compatíveis foi dada prioridade à função de crescimento, ou seja, as expressões obtidas para os parâmetros comuns correspondem à minimização da soma dos quadrados dos resíduos das funções de crescimento. As expressões para os parâmetros livres, presentes nas funções de inicialização, são então estimadas por minimização das respectivas somas de quadrados dos resíduos, mantendo-se as expressões previamente encontradas para os parâmetros comuns. Esta opção foi discutida com os técnicos das empresas de celulose que forneceram dados para a base de dados sibFlor, que a preferiram à alternativa de proceder ao ajustamento simultâneo das funções de inicialização e de projeção.

## 2.3 Definição do Modelo Matemático a ajustar

Matematicamente começa-se por definir o modelo *GLOBULUS 2.1* considerando apenas dados de primeira rotação, não diferenciando o ajustamento por regiões climáticas em Portugal e considerando apenas espaços de tempo unitários.

### 2.3.1 Altura dominante

A variável altura dominante,  $hdom_t$ , representa a altura das árvores dominantes na parcela florestal em metros. A equação foi escolhida com recurso à função Lundqvist-k (ver secção 2.2) com o parâmetro de forma  $n_h$

$$\text{(Equação 1)} \quad hdom_t = A_h \left( \frac{hdom_{t-1}}{A_h} \right)^{\left( \frac{t-1}{t} \right)^{n_h}}, \text{ para } t > 0 \text{ e inteiro,}$$

onde  $hdom_t$  é a variável a explicar (altura dominante) em função de  $hdom_{t-1}$  (variável desfasada) e  $t$  (idade). Os parâmetros a estimar para esta equação são:  $A_h$  e  $n_h$ .

### 2.3.2 Mortalidade

A variável de mortalidade  $N_t$  representa o número de árvores vivas por ha na parcela florestal. Note-se que para esta variável há a necessidade de contemplar uma equação de inicialização e outra de projecção, pois pode-se ter já um povoamento com uma dada idade ( $t > 1$ ) e com dados para as variáveis gerados para o ano anterior (projecção) ou termos de iniciar a sua modelação na idade ( $t$ ) do povoamento (inicialização). Este modelo corresponde ao modelo seleccionado no modelo SOP (Amaro, 1997).

**Equação de Inicialização:**

$$(Equação 2.1) \quad N_t = Npl \exp\left(-\left(am0 + amnp \frac{Npl}{1000} + amQ \frac{Iqe}{10}\right)t\right), \text{ para } t > 0 \text{ e inteiro}$$

Nesta equação  $N_t$  (número de árvores no povoamento) é a variável a explicar em função de  $Npl$ ,  $Iqe$  e  $t$ . Os parâmetros a estimar para esta equação são:  $am0$ ,  $amnp$  e  $amQ$ .

$Npl$  representa o número de árvores por ha à plantação e  $Iqe$  é o índice de qualidade da estação para aquela parcela florestal, pode ser calculado como uma função (inversa) da altura dominante a uma idade fixa do povoamento.

**Equação de Projeção:**

$$(Equação 2.2) \quad N_t = N_{t-1} \exp\left(-\left(am0 + amnp \frac{Npl}{1000} + amQ \frac{Iqe}{10}\right)\right), \text{ para } t > 1 \text{ e inteiro}$$

Nesta equação  $N_t$  (número de árvores no povoamento) é a variável a explicar em função de  $N_{t-1}$  (variável desfasada),  $Npl$  e  $Iqe$ . Os parâmetros a estimar para esta equação são:  $am0$ ,  $amnp$  e  $amQ$ .

$Npl$  e  $Iqe$  têm o significado anteriormente referido.

**2.3.3 Área basal**

A área basal  $G_t$  representa a soma das áreas das árvores da parcela a 1,30 m de altura em metros quadrados por ha. Também para esta variável há a necessidade de especificar uma equação de inicialização e outra de projeção.

**Equação de Inicialização:**

Foi escolhida a função de Lundqvist na sua forma integral com os parâmetros  $A_g$  e  $n_g$  iguais aos da respectiva função de crescimento e com o parâmetro  $k_g$  expresso em função de um conjunto razoável de variáveis de controlo e suas interações.

$$(Equação 3.1) \quad G_t = A_g \exp\left(-k_g \left(\frac{1}{t}\right)^{n_g}\right), \text{ para } t > 0 \text{ e inteiro}$$

e

$$A_g = A_{gQ} Iqe^2$$

$$k_g = k_{g0} + k_{gQ} \frac{1}{Iqe} + k_{gnp} \frac{Npl}{1000} + k_{gf} Fe_t$$

$$n_g = n_{g0} + n_{gQ} \ln(Iqe) + n_{gn} \frac{N_t}{1000}$$

A variável  $Fe$  representa um factor de espaçamento, à semelhança do factor de Wilson (Tomé *et al.* 2001), mas utilizando a altura dominante e a densidade à plantação:

$$Fe_t = \frac{100}{h_{dom_t} \sqrt{Npl}}$$

Nesta equação  $G_t$  (área basal do povoamento) é a variável a explicar em função de  $t$ ,  $Npl$ ,  $Iqe$ ,  $N_t$  e  $Fe_t$ . Os parâmetros a estimar para esta equação são:  $A_{gQ}$ ,  $k_{g0}$ ,  $k_{gQ}$ ,  $k_{gnp}$ ,  $k_{gf}$ ,  $n_{g0}$ ,  $n_{gQ}$  e  $n_{gn}$ .

**Equação de Projeção:**

Selecionou-se a função Lundqvist-k com o parâmetro de forma  $n_g$ .

$$(Equação 3.2) \quad G_t = A_g \left(\frac{G_{t-1}}{A_g}\right)^{\frac{(t-1)^{n_{gt-1}}}{t^{n_{gt}}}}, \text{ para } t > 1 \text{ e inteiro,}$$

onde:

$$A_g = A_{gQ} Iqe^2$$

$$n_{gt} = n_{g0} + n_{gQ} \ln(Iqe) + n_{gn} \frac{N_t}{1000}, \text{ para } t > 0 \text{ e inteiro.}$$

Nesta equação  $G_t$  (área basal do povoamento) é a variável a explicar em função de  $G_{t-1}$  (a variável desfasada),  $t$ ,  $Iqe$  e  $N_t$ . Os parâmetros a estimar para esta equação são:  $A_{gQ}$ ,  $n_{g0}$ ,  $n_{gQ}$  e  $n_{gn}$ .

## 2.4 Noções Preliminares de Sistemas de Equações

O modelo constituído por sistemas de equações interdependentes é um modelo estatístico caracterizado por um sistema de equações que explicitam relações de dependência entre variáveis, podendo assumir simultaneamente interdependência entre os erros das equações. Diz-se linear se traduzir uma relação linear em relação às variáveis e aos parâmetros.

Nestes modelos podem existir vários tipos de equações. Algumas traduzem equações fundamentais entre as variáveis relativas ao fenómeno a ser estudado, e são denominadas equações estruturais, pois estabelecem ligações entre variáveis.

A aleatoriedade do modelo fica estabelecida através da introdução em cada equação estrutural de um termo aditivo aleatório usualmente designado por erro do modelo.

Além dessas variáveis, as equações estruturais contêm também parâmetros desconhecidos, associadas às variáveis, que são chamados parâmetros estruturais do modelo.

As restantes equações do sistema, caso existam, também traduzem ligações entre as variáveis, mas não envolvem termos relativos aos erros aleatórios, nem parâmetros desconhecidos.

No caso geral, em sistemas de equações interdependentes estão incluídas dois tipos de variáveis, as endógenas e as exógenas.

As variáveis endógenas ou dependentes são aquelas cujos valores são determinados pela interação conjunta com outras variáveis dentro do sistema, usualmente surgem no primeiro membro das equações. Por vezes, podem também aparecer como variáveis preditivas no segundo membro das equações do sistema.

As variáveis exógenas ou pré-determinadas, são aquelas cujos valores são determinados exteriormente ao sistema, mas que influenciam os valores das variáveis endógenas. Apenas aparecem no segundo membro das equações e o modelo nada nos diz sobre o modo como estas variáveis são determinadas.

Usualmente, em sistemas de equações interdependentes algumas equações envolvem mais do que uma variável endógena. Para determinar os valores destas variáveis, são necessárias tantas equações quantas variáveis endógenas e tem-se um sistema designado por completo.

### 2.4.1 Classificações de Sistemas de Equações

De acordo com Pindyck e Rubinfeld em 1981, como referenciado em Borders (1989), os sistemas de equações relacionadas podem-se classificar em três categorias: SUR – Equações Aparentemente Não Relacionadas (*Seemingly Unrelated Equations*), equações recursivas e equações simultâneas.

A natureza de cada tipo de sistema de equações é diferente e é necessário perceber a natureza dos diferentes sistemas de equações para os estimarmos.

Um sistema de equações diz-se que está na forma estrutural matricial se tiver a representação:

$$(Equação 2.4.1) \quad BY + \Gamma Z + \varepsilon = 0,$$

onde  $Y$  é a matriz das observações das variáveis endógenas,  $Z$  é a matriz das observações das variáveis exógenas e  $\varepsilon$  é a matriz dos erros aleatórios das equações.

As matrizes  $B$  e  $\Gamma$  contêm os parâmetros estruturais do modelo, correspondendo respectivamente aos coeficientes das variáveis endógenas e aos coeficientes das variáveis exógenas.

A posição dos elementos iguais a zero na matriz  $B$  indica que variáveis endógenas não aparecem nas diferentes equações estruturais. Com o conhecimento da matriz  $B$  podemos distinguir entre os vários tipos de estruturas de sistemas de equações diferentes.

### 1) Sistema de Equações Aparentemente Não Relacionadas (SUR)

Quando a matriz  $B$  é diagonal, i. e., se:

$$B = \begin{bmatrix} B_{11} & 0 & \dots & 0 \\ 0 & B_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & B_{ii} \end{bmatrix}$$

apenas uma variável endógena aparece em cada equação. Isto mostra que as equações não são simultâneas, mas aparentemente relacionadas. Obviamente, não existem relações analíticas entre equações se  $Cov(\varepsilon_i, \varepsilon_j) = 0$  para todos os pares de termos de erro  $i$  e  $j$ , com  $i \neq j$ . Se existe  $Cov(\varepsilon_i, \varepsilon_j) \neq 0$ , para  $i \neq j$ , então as equações são relacionadas. Quando tais correlações entre as equações estão presentes, é desejável usar um procedimento de estimação em três passos discutido em Reed (1986) para este tipo de sistema de equações.

As estimativas SUR são obtidas estimando um conjunto de equações não lineares com restrições impostas entre as equações, mas com matrizes de covariâncias dos resíduos diagonais nas equações. Estas estimativas de parâmetros são usadas para formar um estimador consistente da matriz de covariâncias dos erros, que é depois usada como matriz ponderadora quando o modelo é reestimado para se obterem os novos valores dos parâmetros. Estas estimativas são consistentes e assintoticamente normais, e

dentro de certas condições, assintoticamente mais eficientes que quando as equações são estimadas individualmente.

## 2) Sistema de Equações Recursivas

Ocorre quando B é triangular, i. e. se:

$$B = \begin{bmatrix} B_{11} & 0 & \dots & 0 \\ B_{21} & B_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ B_{G1} & B_{G2} & \dots & B_{GG} \end{bmatrix}$$

O sistema de equações recursivo é um conjunto simultâneo de equações de regressão estimadas uma a uma numa sequência lógica. A condição geral que deve ser encontrada para definir um sistema de equações recursivo é que a matriz B deve ser triangular. Uma condição necessária para ser estimado um sistema recursivo é que seja algebricamente completo no sentido em que o número de equações seja igual ao número de variáveis endógenas. Neste caso a solução para a g-ésima variável endógena envolve apenas as primeiras g equações estruturais. Isto significa que todas as G equações estruturais são necessárias para a solução apenas no caso da última variável endógena  $Y_{Gt}$ . Note-se que a primeira equação estrutural envolve apenas uma variável endógena. Para este tipo de sistema de equações, o método dos mínimos quadrados pode ser usado para obter estimativas dos parâmetros se não existirem correlações entre os erros das diferentes equações no sistema. Numa definição estrita de um sistema recursivo, não existe problema de correlações entre as variáveis endógenas do lado direito e as componentes do erro das variáveis endógenas do lado esquerdo das equações. Contudo, este pressuposto leva-nos ao facto de que as componentes do erro para a equação num sistema recursivo são não relacionadas aos pares. Quando tais correlações existem, não é verdade que as variáveis endógenas do segundo membro são não correlacionadas com as componentes do erro das variáveis endógenas do primeiro membro. Deste modo, quando existem correlações entre as equações num sistema de equações recursivo, o sistema não é recursivo por definição



(Borders, 1989). Neste caso são comumente usados os métodos dos mínimos quadrados em três passos ou SUR.

### 3) Sistema de Equações Simultâneas

Exemplo de um sistema de  $m$  equações simultâneas na forma estrutural:

$$\begin{cases} v_{11}y_{t1} + \dots + v_{m1}y_{tm} + \beta_{11}x_{t1} + \dots + \beta_{k1}x_{tk} + \varepsilon_{t1} = 0 \\ v_{12}y_{t1} + \dots + v_{m2}y_{tm} + \beta_{12}x_{t1} + \dots + \beta_{k2}x_{tk} + \varepsilon_{t2} = 0 \\ \dots \\ v_{1m}y_{t1} + \dots + v_{mm}y_{tm} + \beta_{1m}x_{t1} + \dots + \beta_{km}x_{tk} + \varepsilon_{tm} = 0 \end{cases}$$

Cada equação  $i$ ,  $i = 1, \dots, m$  contém: um erro aleatório  $\varepsilon_{ti}$ ; as variáveis endógenas  $y_{t1}, \dots, y_{tm}$  e os parâmetros estruturais  $\beta_{r2}$ ,  $r = 1, \dots, k$  associados às  $k$  variáveis exógenas  $x_{t1}, \dots, x_{tk}$ .

Se  $B$  não é diagonal nem triangular, é definida uma estrutura integrada ou um sistema de equações simultâneas. Um sistema de equações simultâneas ocorre quando duas ou mais variáveis num sistema são conjuntamente determinadas pelo fenómeno sendo modelado num sistema de equações inter-relacionadas, que mostra que não existe uma ordem sequencial das equações. As relações num sistema destes podem ser complexas e podem consistir em variáveis que aparecem como regressores e variáveis dependentes ao mesmo tempo assim como também correlações entre as componentes dos erros das diferentes equações. Além do mais, podem ser colocadas restrições nos parâmetros tanto dentro como entre as equações. Os métodos mais usados para a estimação deste tipo de sistemas são o SUR e o método dos mínimos quadrados em três passos a fim de se obterem estimadores consistentes e assintoticamente eficientes para os parâmetros.



## CAPITULO 3

### 3 Estimação do Modelo com Recurso a Técnicas de Estatística Clássica

Sabe-se que num sistema de equações simultâneas os estimadores obtidos pelo método dos mínimos quadrados não são consistentes nem eficientes.

Consequentemente, no processo de estimação utilizado tiveram-se em conta os resultados obtidos em estudos anteriores, como Yuancai *et al.*, 2001, em que se obtiveram melhores estimativas (desvios-padrão dos parâmetros mais baixos) pelo método de estimação de Regressão Não Linear Aparentemente Não Relacionada - NSUR (*Nonlinear Seemingly Unrelated Regression*) do que por outros métodos com estimadores com propriedades assintoticamente semelhantes, como o método dos Mínimos Quadrados Não Linear em Três Passos - N3SLS (*Nonlinear Three-Stage Least Squares*).

Assim para a estimação dos parâmetros da equação da altura dominante utilizou-se o método dos Mínimos Quadrados Não Linear (NLLS – *Nonlinear Least Squares*), enquanto os parâmetros das equações da mortalidade e da área basal foram estimadas através do método NSUR aplicado a um sistema de equações simultâneas não linear.

De seguida far-se-á uma exposição de cada um dos métodos, assim como das principais propriedades dos seus estimadores.

### 3.1 Métodos de Estimação

**Definição:** um **modelo de regressão não linear** é um modelo para o qual as equações normais de primeira ordem para a estimação dos mínimos quadrados dos parâmetros são funções não lineares dos parâmetros.

#### 3.1.1 Mínimos Quadrados Não Linear

##### i) A regressão linearizada

O modelo de regressão não linear pode ser representado por:

$$y_i = h(x_i, \beta) + \varepsilon_i, \quad \beta \in \mathfrak{R}^k \quad i = 1, \dots, n$$

onde  $\mathbf{y}$  representa o vector da variável dependente,  $\mathbf{x}$  representa os vectores das variáveis explicativas ou independentes,  $\beta$  representa os coeficientes a estimar da regressão e  $\varepsilon$  o vector dos erros, com valor esperado vector nulo e matriz de covariâncias  $\Sigma = \sigma^2 I$ , segundo as hipóteses do modelo de regressão não linear (Greene, capítulo 9).

Muitos dos resultados que têm sido obtidos para os modelos de regressão não linear são baseados na aproximação linear por séries de Taylor para  $h(x, \beta)$  num valor particular para o vector de parâmetros  $\beta^0$ :

$$h(x, \beta) \approx h(x, \beta^0) + \sum_{i=1}^k \frac{\partial h(x, \beta^0)}{\partial \beta_i^0} (\beta_i - \beta_i^0)$$

Este modelo é o chamado **modelo de regressão linearizado**. Manipulando os termos de outra forma obtém-se:

$$h(x, \beta) \approx \left[ h(x, \beta^0) - \sum_{i=1}^k \beta_i^0 \left( \frac{\partial h(x, \beta^0)}{\partial \beta_i^0} \right) \right] + \sum_{i=1}^k \beta_i \left( \frac{\partial h(x, \beta^0)}{\partial \beta_i^0} \right)$$

Seja  $x_i^0$  a derivada parcial em ordem a  $\beta_i$  calculada no ponto  $\beta^0$ . Para um dado valor de  $\beta^0$ ,  $x_i^0$  é uma função apenas dos dados, não dos parâmetros desconhecidos,

$$x_i^0 = \frac{\partial h(x, \beta^0)}{\partial \beta_i^0}.$$

Temos assim:

$$h(x, \beta) \approx \left[ h^0 - \sum_{i=1}^k x_i^0 \beta_i^0 \right] + \sum_{i=1}^k x_i^0 \beta_i, \quad \text{onde } h^0 = h(x, \beta^0), \text{ podendo ser reescrito}$$

como

$$h(x, \beta) \approx h^0 - x^{0'} \beta^0 + x^{0'} \beta, \text{ o que implica que}$$

$$y \cong h^0 - x^{0'} \beta^0 + x^{0'} \beta + \varepsilon^0.$$

Colocando os termos conhecidos no primeiro membro da equação, obtemos uma equação linear:

$$y - h^0 + x^{0'} \beta^0 = x^{0'} \beta + \varepsilon^0.$$

Designando por  $y^0$  o primeiro membro, vem

$$y^0 = x^{0'} \beta + \varepsilon^0, \quad \text{onde } \varepsilon^0 = \varepsilon + \left[ h(x, \beta) - \left\{ h^0 - \sum_{i=1}^k x_i^0 \beta_i^0 + \sum_{i=1}^k x_i^0 \beta_i \right\} \right].$$

$\varepsilon^0$  incorpora o termo residual original ( $\varepsilon$ ), adicionado do erro da aproximação por série de Taylor de primeira ordem para a verdadeira regressão.

Mediante o conhecimento de  $\beta^0$  (valores iniciais),  $x^0$  e  $y$ , temos  $y^0$  definido, pelo que se torna possível estimar os parâmetros  $\beta$  do modelo linear  $y^0 = x^{0'} \beta + \varepsilon^0$  através do método dos mínimos quadrados linear.

## ii) Propriedades do estimador dos Mínimos Quadrados Não Linear

A função a minimizar para obter as estimativas  $\hat{\beta}$  do método dos mínimos quadrados não linear é dada por

$$S(\beta) = \sum_{i=1}^n [y_i - h(x_i, \beta)]^2 = \sum_{i=1}^n \varepsilon_i^2$$

As condições de primeira ordem para que  $\hat{\beta}$  seja um mínimo são

$$g(\beta) = \frac{\partial S(\beta)}{\partial \beta} = -\sum_{i=1}^n [y_i - h(x_i, \beta)] \frac{\partial h(x_i, \beta)}{\partial \beta} = 0$$

Note-se que na solução,

$$g(\hat{\beta}) = -X_0' \varepsilon = 0,$$

resultado que é similar ao do estimador linear.

O método dos mínimos quadrados continua a ser uma maneira atrativa para estimar os parâmetros. Muitos resultados analíticos têm sido obtidos para o estimador, tais como consistência e normalidade assintótica. Estes resultados podem ser consultados com mais detalhe e rigor por exemplo em Amemiya, 1983. Não poderemos afirmar com toda a certeza que o estimador dos mínimos quadrados não linear é o mais eficiente, exceto em casos de normalidade dos resíduos.

No modelo de regressão clássico, para obter os resultados assintóticos, assume-se que a matriz  $(1/n)X'X$  converge para uma matriz definida positiva  $\mathbf{Q}$ , isto é,

$p \lim \left( \frac{1}{n} \right) X'X = \mathbf{Q}$ . Analogamente, impomos neste caso a mesma condição para os

**pseudo-regressores** (obtidos através das derivadas parciais da função  $h(\cdot)$  em relação a cada um dos parâmetros num dado valor tomado pelo vector dos parâmetros) no modelo linearizado quando eles são calculados nos verdadeiros valores dos

parâmetros. Admitindo que  $\beta^0$  é o vector dos verdadeiros valores dos parâmetros, vem que

$$p \lim \frac{1}{n} X_0' X_0 = p \lim \frac{1}{n} \sum_{i=1}^n \left( \frac{\partial h(x_i, \beta^0)}{\partial \beta^0} \right) \left( \frac{\partial h(x_i, \beta^0)}{\partial \beta^0} \right)' = Q_0$$

, onde  $Q_0$  é uma matriz definida positiva. Para estabelecer a consistência de  $\hat{\beta}$  no modelo linear, é necessário que  $p \lim (1/n) X' \varepsilon = 0$ . Neste caso requeremos o equivalente para os pseudo-regressores:  $p \lim \frac{1}{n} \sum_{i=1}^n x_i^0 \varepsilon_i = 0$ .

Finalmente, a normalidade assintótica pode ser estabelecida dentro de condições gerais (supondo a validade de condições de regularidade, Greene, capítulo 9) se

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n x_i^0 \varepsilon_i \xrightarrow{d} N[0, \sigma^2 Q_0]$$

Dados estes resultados, as propriedades assintóticas do estimador dos mínimos quadrados não linear foram derivadas.

**A consistência de  $\hat{\beta}$  (fraca) fica provada**, se se mostrar que  $p \lim \frac{1}{n} S(\beta)$  é minimizado no verdadeiro valor  $\beta_0$ . A consistência forte é provada mostrando que o mesmo acontece usando a convergência quase certa para a expressão  $\frac{1}{n} S(\beta)$ .

Este método de demonstração é intuitivo, pois parece óbvio que se  $\frac{1}{n} S(\beta)$  está próximo para  $p \lim \frac{1}{n} S(\beta)$  e se o último é minimizado em  $\beta_0$ , então  $\hat{\beta}$  que minimiza o primeiro deve estar próximo de  $\beta_0$ .

Ainda assim necessitamos de três hipóteses para podermos demonstrar a consistência:

**Hipótese 1:** O espaço dos parâmetros  $B$  é compacto (fechado e limitado) e  $\beta_0$  é seu ponto interior;

**Hipótese 2:**  $S(\beta)$  é contínua em  $\beta$ ;

**Hipótese 3:**  $p \lim \frac{1}{n} S(\beta)$  existe, é não estocástico e a sua convergência é não aleatória em  $\beta$ .

O significado da última hipótese é o seguinte. Definindo  $F(\beta) = p \lim \frac{1}{n} S(\beta)$ . Então, dado  $\xi, \delta > 0$ , existe  $n_0$ , independente de  $\beta$ , tal que para todo  $n \geq n_0$  e para todo o  $\beta$ ,  $\text{Prob}\left[ \left| n^{-1} S(\beta) - F(\beta) \right| > \xi \right] < \delta$ .

De seguida far-se-á uma demonstração simplificada e serão mencionadas as hipóteses necessárias.

Das definições do modelo e de  $S(\beta)$ , vem

$$\frac{1}{n} S(\beta) = \frac{1}{n} \sum_{i=1}^n [y_i - h(x_i, \beta)]^2$$

Substituindo  $y_i$  pela sua expressão (com  $\beta = \beta_0$ ) na expressão anterior, obtêm-se a expressão equivalente

$$\frac{1}{n} S(\beta) = \frac{1}{n} \sum_{i=1}^n [h(x_i, \beta_0) + \varepsilon_i - h(x_i, \beta)]^2$$

Arranjando os termos de forma conveniente, obtêm-se

$$\frac{1}{n} S(\beta) = \frac{1}{n} \sum_{i=1}^n [\varepsilon_i + (h(x_i, \beta_0) - h(x_i, \beta))]^2$$

Ou, desenvolvendo o quadrado da adição,

$$\frac{1}{n} S(\beta) = \frac{1}{n} \sum_{i=1}^n \varepsilon_i^2 + \frac{2}{n} \sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta)) \varepsilon_i + \frac{1}{n} \sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta))^2$$

$$\equiv A_1 + A_2 + A_3$$

Em primeiro lugar,  $p \lim A_1 = \sigma^2$ , pela lei dos grandes números.

**Definição - Lei Forte dos Grandes Números:** *Seja  $\{X_k\}_{k \geq 1}$  uma sucessão de variáveis aleatórias i.i.d. distribuídas com  $X$  e defina-se  $S_N = \sum_{i=1}^N X_k$  e  $\bar{X}_N = S_N / N$ . Então, com*



$N \rightarrow \infty$ ,  $\bar{X}_N$  converge quase certamente para  $\mu$  finito se e só se  $E[|X|] < \infty$ , e tem-se  $E[X] = \mu$ .

Em segundo lugar, fixados  $\beta_0$  e  $\beta$ ,  $p \lim A_2 = 0$ , segue da convergência de

$\frac{1}{n} \sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta))^2$  pela desigualdade de Markov:

$$\Pr ob \left[ \left\{ \frac{1}{n} \sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta)) \varepsilon_i \right\}^2 \geq \xi^2 \right] \leq \frac{\sigma^2}{\xi^2 n^2} \sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta))^2$$

ou de uma forma mais explícita

$$\Pr ob \left[ \varepsilon_i^2 \geq \frac{\xi^2 n^2}{\sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta))^2} \right] \leq \frac{\sigma^2}{\xi^2 n^2} \sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta))^2.$$

Uma vez que a convergência de  $A_2$  segue da convergência uniforme do segundo membro da desigualdade anterior, é suficiente assumir que

$\frac{1}{n} \sum_{i=1}^n h(x_i, \beta_1) h(x_i, \beta_2)$  converge uniformemente em  $\beta_1, \beta_2 \in B$ .

Neste momento basta apenas assumir que  $\lim A_3$  é unicamente minimizado em  $\beta_0$ , nomeadamente,

$$\lim \frac{1}{n} \sum_{i=1}^n (h(x_i, \beta_0) - h(x_i, \beta))^2 \neq 0 \text{ se } \beta \neq \beta_0.$$

Somando tudo, o estimador dos mínimos quadrados não linear ( $\hat{\beta}$ ) é consistente se a Hipótese 1 e as duas condições anteriores são satisfeitas. Para mais detalhe sobre a significância e plausibilidade destas hipóteses e condições veja Amemiya, 1983.

Uma estimativa consistente de  $\sigma^2$  pode ser utilizada usando os resíduos estimados (considerando b a estimativa de  $\hat{\beta}$ ):

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n [y_i - h(x_i, b)]^2.$$

A questão de correção dos graus de liberdade por  $(n-k)$  em vez de  $(n)$  não se coloca aqui uma vez que estamos perante resultados assintóticos.

**Teorema:** *Distribuição assintótica do estimador dos Mínimos Quadrados Não Linear.*

Se os pseudo-regressores forem “bem comportados”, isto é, se não violarem as hipóteses definidas para o modelo NLLS (Greene, cap. 9) e se se verificarem as seguintes hipóteses:

- i) O espaço dos parâmetros que contém  $\beta$  é compacto;
- ii) Para um vector  $\beta_0$  pertencente ao espaço dos parâmetros,
 
$$p \lim \frac{1}{n} S(\beta_0) = q(\beta_0);$$
- iii)  $q(\beta_0)$  tem um único mínimo no verdadeiro vector dos parâmetros.

Então:

$$\hat{\beta}^a \sim N \left[ \beta, \frac{\sigma^2}{n} Q_0^{-1} \right] \text{ (assintoticamente), onde}$$

$$Q_0 = p \lim \frac{1}{n} X^{0'} X^0$$

E a estimativa da matriz de covariâncias assintótica é dada por  $\hat{\sigma}^2 (X^{0'} X^0)$ . Uma demonstração rigorosa pode encontrar-se em Jennrich, 1969.

A normalidade assintótica do estimador NLLS, como no caso do estimador da máxima verosimilhança, pode ser derivado da seguinte expansão de Taylor:

$$\frac{\partial S}{\partial \beta} \Big|_{\hat{\beta}} = \frac{\partial S}{\partial \beta} \Big|_{\beta_0} + \frac{\partial^2 S}{\partial \beta \partial \beta'} \Big|_{\beta^*} (\hat{\beta} - \beta_0),$$

onde  $\partial^2 S / (\partial \beta \partial \beta')$  é uma matriz  $k$  por  $k$  das derivadas de segunda ordem e  $\beta^*$  encontra-se entre  $\hat{\beta}$  e  $\beta_0$ . Para se poder escrever a expressão anterior, assume-se que  $h(\cdot)$  é duas vezes continuamente diferenciável em ordem a  $\beta$ . Uma vez que o primeiro membro da expressão anterior é igual a zero (porque  $\hat{\beta}$  minimiza  $S$ ), reescrevendo a igualdade anterior e utilizando algumas manipulações algébricas, obtemos

$$\sqrt{n}(\hat{\beta} - \beta_0) = - \left[ \frac{1}{n} \frac{\partial^2 S}{\partial \beta \partial \beta'} \Big|_{\beta^*} \right]^{-1} \frac{1}{\sqrt{n}} \frac{\partial S}{\partial \beta} \Big|_{\beta_0}.$$

Para provarmos necessitamos de demonstrar que:

- i) O limite em distribuição de  $\frac{1}{\sqrt{n}} \frac{\partial S}{\partial \beta} \Big|_{\beta_0}$  é normal;
- ii)  $\frac{1}{n} \frac{\partial^2 S}{\partial \beta \partial \beta'} \Big|_{\beta^*}$  converge em probabilidade para uma matriz não singular.

Ambas as demonstrações, do ponto i) e do ponto ii) (mais difícil) podem encontrar-se em Amemiya, 1983.

Com estes resultados, inferências e testes de hipóteses podem ser feitos do mesmo modo que para o modelo linear.

Um problema menor prende-se com o facto de que não se garante neste caso que  $R^2$  esteja limitado entre 0 e 1, sendo também denominado de eficiência de modelação.

$$R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Para o cálculo das estimativas ótimas é comum usar-se um processo iterativo com recurso a métodos numéricos, como o método de Gauss-Newton.

### 3.1.2 Regressão Não Linear Aparentemente Não Relacionada

O método SUR (*Seemingly Unrelated Regression*) tem uma grande importância pois assume a existência de correlação contemporânea entre os resíduos de diferentes equações de um sistema de equações simultâneas e tem em conta esse facto no processo de estimação, produzindo estimativas mais eficientes. As mesmas razões são igualmente válidas para sistemas de equações simultâneas não lineares aparentemente não relacionadas (NSUR).

Considerem-se M equações não lineares:

$$\begin{cases} y_1 = h_1(X, \beta) + \varepsilon_1 \\ y_2 = h_2(X, \beta) + \varepsilon_2 \\ \dots \\ y_M = h_M(X, \beta) + \varepsilon_M. \end{cases}$$

Seja  $\varepsilon' = (\varepsilon'_1, \varepsilon'_2, \dots, \varepsilon'_M)$ , assume-se que  $E[\varepsilon\varepsilon'] = \Sigma \otimes I_n$ , onde  $\Sigma$  é uma matriz de covariâncias (M x M) cujo elemento (i,j) é dado por  $\sigma_{ij}$ , onde  $E[\varepsilon_i\varepsilon_j'] = \sigma_{ij}I_n$ . A mesma matriz X e o mesmo vector  $\beta$  aparecem em todas as equações para alargar à possibilidade de que as mesmas variáveis explicativas, e alguns coeficientes, podem ser comuns a mais do que uma equação. Cada equação pode ser uma função não linear diferente de X e  $\beta$ .

O operador  $\otimes$  representa o produto de Kronecker. Relembrando, o produto de Kronecker  $A \otimes B$  entre as matrizes A e B,  $A = [a_{ij}]_{k_1 \times m_1}$  e  $B = [b_{ij}]_{k_2 \times m_2}$ , define-se por

$$A \otimes B = [a_{ij}B]_{(k_1 \times k_2) \times (m_1 \times m_2)}.$$

Este método contempla duas possibilidades de estimação, a **estimação em dois passos** e a **estimação pela máxima verosimilhança**. No âmbito deste trabalho foi utilizada a estimação a dois passos e é esta que vai ser descrita no ponto seguinte.

### i) Estimação em dois passos

Quando  $\Sigma$  é conhecida, um estimador dos mínimos quadrados não linear para  $\beta$  mais eficiente que o obtido aplicando mínimos quadrados não lineares a cada equação individualmente, pode ser obtido minimizando a soma dos quadrados da função:

$$\varepsilon' (\Sigma^{-1} \otimes I_n) \varepsilon .$$

Como  $\Sigma$  é na maioria dos casos desconhecida, uma maneira de proceder é estimar  $\Sigma$  usando os resíduos da aplicação a cada equação separadamente obtidos pelo método dos mínimos quadrados não linear (NLLS), e num segundo passo, determinar por métodos numéricos o valor da estimativa de  $\beta$  que minimiza a expressão algébrica anterior, mas com  $\Sigma$  substituída por uma estimativa  $\hat{\Sigma}$ . Este processo pode ser sumariado através dos seguintes passos:

1. Determinar as estimativas de  $\beta_1, \beta_2, \dots, \beta_M$  que minimizam  $\varepsilon_1' \varepsilon_1, \varepsilon_2' \varepsilon_2, \dots, \varepsilon_M' \varepsilon_M$ , respectivamente;

2. Calcular cada vector de resíduos como  $\hat{e}_i = y_i - h_i(X, b_i), i = 1, 2, \dots, M$ ;

3. Determinar estimativas  $\hat{\sigma}_{ij} = e_i' e_j / n$  e construir a correspondente matriz  $\hat{\Sigma}$ ;

4. Determinar o valor de  $\beta$  que minimiza  $\varepsilon' (\hat{\Sigma}^{-1} \otimes I_n) \varepsilon$ .

Prova-se (Gallant, 1975), que este estimador é consistente e assintoticamente eficiente quando os erros são normalmente distribuídos.

Uma estimativa da matriz de covariâncias assintótica para este estimador, para  $\hat{\beta}$ , é dada por:

$$\hat{\Sigma}_b = \left[ \frac{\partial \varepsilon'}{\partial \beta} \left( \hat{\Sigma}^{-1} \otimes I_n \right) \frac{\partial \varepsilon}{\partial \beta'} \right]^{-1} \Bigg|_{\beta = \hat{\beta}}.$$

Se  $\beta$  de dimensão  $(K \times 1)$ ,  $\frac{\partial \varepsilon'}{\partial \beta}$  é uma matriz  $(K \times Mn)$  dada por:

$$\frac{\partial \varepsilon'}{\partial \beta} = \left[ \frac{\partial h_1}{\partial \beta}, \frac{\partial h_2}{\partial \beta}, \dots, \frac{\partial h_M}{\partial \beta} \right]$$

Se cada equação contiver um conjunto diferente de parâmetros,  $\frac{\partial \varepsilon'}{\partial \beta}$  é uma matriz diagonal por blocos.

A estimação conjunta em dois passos não será mais eficiente que a aplicação feita separadamente a cada equação pelo método NLLS se pelo menos uma das duas condições seguintes ocorrerem: a primeira condição é que  $\Sigma$  é diagonal (não existe correlação contemporânea entre as equações); a segunda condição é que cada função  $h_i(X, \beta_i)$  é a mesma e depende do mesmo conjunto de variáveis explicativas, mas existe um vector  $\beta$  diferente para cada equação.

### 3.2 Aplicação ao Modelo: Resultados da Estimação Clássica

Neste ponto, a primeira coisa a estudar é a identificação das equações do sistema para a mortalidade e área basal, porque só em casos de identificabilidade ou sobreidentificabilidade destas equações é possível estimar este sistema de equações.

De seguida vou proceder à apresentação dos resultados das estatísticas descritivas das variáveis utilizadas para termos uma ideia das distribuições empíricas de todas as variáveis.

Por fim, apresentarei os resultados da estimação do modelo e farei os testes estatísticos para verificar a adequabilidade e bondade do modelo

### **3.2.1 Estudo da Identificabilidade do Sistema de Equações Simultâneas**

Um modelo diz-se identificável quando são identificáveis todos os seus parâmetros, na medida em que podem ser univocamente determinados pela distribuição conjunta das variáveis observadas. Neste caso e atendendo à expressão designada por Equação 2.4.1 do modelo refere-se à identificabilidade dos parâmetros estruturais do modelo contidos nas matrizes  $B$  e  $\Gamma$ .

A identificabilidade em sistemas não lineares é um pouco diferente do que quando se tratam de equações lineares. Amemiya (1988), defende que a não linearidade geralmente ajuda mais do que dificulta a identificabilidade. O número de variáveis exógenas (pré-determinadas) excluídas numa dada equação não precisa ser maior ou igual ao número de parâmetros da mesma equação num modelo não linear, para que haja identificabilidade.

Neste caso, uma condição suficiente para a identificabilidade é que a matriz  $\hat{\Sigma}_b$  (definida na secção 3.1.2) seja não singular.

### **3.2.2 Estatísticas Descritivas Sobre os Dados Utilizados para a Estimação**

Os dados para a estimação foram obtidos de 114 parcelas de Eucalipto em primeira rotação (talhadia), constituindo uma amostra de dimensão 1477 registos. Destes só puderam ser usados 1325 na estimação da equação da altura dominante (devido à não

utilização dos registos em que não se dispunham de dados para a variável desfasada) e só 1436 na estimação do sistema de equações simultâneas da mortalidade e da área basal (devido à não utilização de registos que causavam erros para os cálculos das estimativas pelo “SAS base”).

Na Tabela 2 encontram-se calculadas as principais estatísticas descritivas desta amostra.

**Tabela 2: Estatísticas descritivas dos dados utilizados para a estimação (n = 1477)**

	t	hdom	N	Npl	G	Iqe	Fe
	(anos)	(m)	(arv/ha)	(arv/ha)	(m <sup>2</sup> /ha)		
Min	1,00	0,36	462,66	462,66	0,00	8,48	0,05
Max	32,00	36,40	5000,00	5000,00	52,00	31,18	4,03
Dif. Max	31,00	36,04	4537,34	4537,34	52,00	22,70	3,99
Média	8,37	18,19	1420,81	1405,97	16,31	21,91	0,24
Desvio Padrão	4,88	7,27	821,23	898,43	10,13	3,36	0,41
C.V.(%)	58,32	39,96	57,80	63,90	62,10	15,33	168,5

A média da amostra é calculada do seguinte modo:  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ ; o desvio padrão por

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$
 e o coeficiente de variação através  $CV = \frac{s}{|\bar{x}|}$ .

Realça-se a grande variabilidade das amostras para a generalidade das variáveis exceto para o índice de qualidade da estação e para as alturas dominantes do ano t e do ano t-1 (variável desfasada). A variável que mede o espaçamento entre as árvores, Fe, apresenta mesmo uma variabilidade bastante alta, na ordem dos 168,5%.



### 3.2.3 Estimativas Clássicas dos Parâmetros do Modelo

#### i) Altura dominante

A equação da altura dominante a estimar é dada por

$$hdom_t = A_h \left( \frac{hdom_{t-1}}{A_h} \right)^{\left( \frac{t-1}{t} \right)^{n_h}} + \varepsilon_t, \text{ para } t > 0 \text{ e inteiro,}$$

onde os parâmetros deste modelo foram estimados pelo método NLLS, usando o “Proc Model” do software “SAS base 9.1.3”, obtendo-se os resultados para a estimação que se encontram na tabela 3.

**Tabela 3: Estimativas dos parâmetros do modelo da altura dominante pelo método NLLS.**

Parâmetros	Estimativas	Desvios-Padrão	Estatística-t	Pr >  t
$A_h$	40,44771	1,9072	21,21	<0,0001
$n_h$	0,746454	0,0328	22,76	<0,0001

Como se pode observar pelos resultados da estimação da Tabela 3, rejeitam-se as hipóteses de nulidade de cada um dos parâmetros estimados ( $H_0 : A_h = 0$  e  $H_0 : n_h = 0$ ), com um grau de significância de 5% (a estatística t é testada com uma t-student com  $n - K = 1325 - 2 = 1323$  graus de liberdade), sendo a probabilidade a favor de  $H_0$  nos dois casos inferior a 0,0001. Deste modo, rejeita-se também a hipótese de nulidade conjunta dos parâmetros do modelo ( $H_0 : A_h = n_h = 0$ ), uma vez que as duas variáveis são estatisticamente diferentes de zero (bastava que pelo menos uma fosse para rejeitar a hipótese).

O modelo apresenta uma eficiência da modelação  $R^2 = 0,9539$  (o que representa uma percentagem alta de explicação da variável endógena por todas as variáveis

explicativas), não estando à partida garantido que este esteja para este método (NLLS) entre 0 e 1 como já foi mencionado na descrição do método.

No entanto o modelo apresenta a raiz quadrada do erro quadrático médio =

$$RMSE = \left[ \frac{1}{n-K} \sum_{t=1}^n (hdom_t - \hat{hdom}_t)^2 \right]^{1/2} = 1,3939 \text{ m} = \hat{\sigma}_{hdom}, \text{ um valor bastante baixo}$$

que significa que há um reduzido desvio-padrão dos erros estimados da equação estimada para a variável altura dominante.

Para verificar a existência de colinearidade na regressão calculou-se o coeficiente de correlação entre os dois pseudo-regressores dados pelas seguintes fórmulas:

$$x_1^0 = \frac{\partial hdom_t}{\partial A_h} \Big|_{A_h = \hat{A}_h} = \left( \frac{hdom_{t-1}}{\hat{A}_h} \right)^{((t-1)/t)^{n_h}} + \hat{A}_h \left( \frac{t-1}{t} \right)^{n_h} \left( \frac{hdom_{t-1}}{\hat{A}_h} \right)^{((t-1)/t)^{n_h}-1} \left( -\frac{hdom_{t-1}}{\hat{A}_h^2} \right)$$

$$x_2^0 = \frac{\partial hdom_t}{\partial n_h} \Big|_{n_h = \hat{n}_h} = \hat{A}_h \log \left( \frac{hdom_{t-1}}{\hat{A}_h} \right) \log \left( \frac{t-1}{t} \right) \left( \frac{t-1}{t} \right)^{n_h} \left( \frac{hdom_{t-1}}{\hat{A}_h} \right)^{((t-1)/t)^{n_h}}$$

Obtendo-se um coeficiente de correlação  $\rho = \frac{\sum_{t=1}^n (x_{1t}^0 - \bar{x}_1^0)(x_{2t}^0 - \bar{x}_2^0)}{\sqrt{\left( \sum_{t=1}^n (x_{1t}^0 - \bar{x}_1^0)^2 \right) \left( \sum_{t=1}^n (x_{2t}^0 - \bar{x}_2^0)^2 \right)}} = 0,6960$ . Com

este resultado (uma correlação não muito forte entre os pseudo-regressores), pode concluir-se que não existem indícios de colinearidade.

Em baixo apresentam-se as matrizes de covariâncias e correlações dos parâmetros (Tabelas 4 e 5 respectivamente).

**Tabela 4: Estimativas das covariâncias dos parâmetros da equação da altura dominante.**

Parâmetros	$A_h$	$n_h$
$A_h$	3,637523	-0,056612
$n_h$	-0,056612	0,0010760

Tabela 5: Estimativas das correlações dos parâmetros da equação da altura dominante.

Parâmetros	$A_h$	$n_h$
$A_h$	1,0000	-0,9049
$n_h$	-0,9049	1,0000

Verifica-se e salienta-se a forte correlação entre as estimativas dos dois parâmetros do modelo.

Em resumo, pode-se considerar um bom ajustamento para a altura dominante.

## ii) Modelo da mortalidade e da área basal

Para a inicialização das variáveis da parcela, o sistema de equações simultâneas a estimar é dado por:

$$N_t = Npl \exp\left(-\left(am0 + amnp \frac{Npl}{1000} + amQ \frac{Iqe}{10}\right)t\right) + \varepsilon_{1t}$$

$$G_t = A_g \exp\left(-k_g \left(\frac{1}{t}\right)^{n_g}\right) + \varepsilon_{2t}, \text{ onde:}$$

$$k_g = k_{g0} + k_{gQ} \frac{1}{Iqe} + k_{gnp} \frac{Npl}{1000} + k_{gf} Fe$$

$$n_g = n_{g0} + n_{gQ} \ln(Iqe) + n_{gn} \frac{N_t}{1000}$$

Para a projeção de dados existentes sobre as variáveis da parcela, o sistema de equações simultâneas a estimar é dado por:

$$N_t = N_{t-1} \exp\left(-\left(am0 + amnp \frac{Npl}{1000} + amQ \frac{Iqe}{10}\right)\right) + \varepsilon_{1t}, \text{ para } t > 1 \text{ e inteiro.}$$

$$G_t = A_g \left(\frac{G_{t-1}}{A_g}\right)^{\frac{(t-1)^{n_{gt}-1}}{t^{n_{gt}}}} + \varepsilon_{2t}, \text{ para } t > 1 \text{ e inteiro, onde}$$

$$A_g = A_{gQ} Iqe^2$$

$$n_{gt} = n_{g0} + n_{gQ} \ln(Iqe) + n_{gn} \frac{N_t}{1000}$$

Este sistema de equações simultâneas foi estimado pelo método NSUR, usando o “Proc Model” do software “SAS base 9.1.3”. Verificou-se que os parâmetros  $am0, amQ, k_{g0}, k_{gnp}, k_{gf}$  (veja-se no Anexo I o processo de evolução do modelo) não eram significativamente diferentes de zero com uma significância de 5%, apresentando estatísticas-T em valor absoluto menores que 1,96. Logo as variáveis associadas foram retiradas das equações, passando as equações a estimar a ficar definidas por:

i) Para a inicialização das variáveis da parcela:

$$N_t = Npl \exp\left(-\left(amnp \frac{Npl}{1000}\right)t\right) + \varepsilon_{1t}$$

$$G_t = A_g \exp\left(-k_g \left(\frac{1}{t}\right)^{n_g}\right) + \varepsilon_{2t}, \text{ onde}$$

$$k_g = k_{gQ} \frac{1}{Iqe}$$

$$n_g = n_{g0} + n_{gQ} \ln(Iqe) + n_{gn} \frac{N_t}{1000}$$

ii) Para a projeção das variáveis da parcela:

$$N_t = N_{t-1} \exp\left(-\left(amnp \frac{Npl}{1000}\right)\right) + \varepsilon_{1t}, \text{ para } t > 1 \text{ e inteiro.}$$

$$G_t = A_g \left( \frac{G_{t-1}}{A_g} \right)^{\frac{(t-1)^{n_{gt}-1}}{t^{n_{gt}}}} + \varepsilon_{2t}, \text{ para } t > 1 \text{ e inteiro,}$$

onde:

$$A_g = A_{g0} I q e^2$$

$$n_{gt} = n_{g0} + n_{gQ} \ln(I q e) + n_{gn} \frac{N_t}{1000}$$

A Tabela 6 apresenta as estimativas obtidas para este modelo, os respectivos desvios-padrão, as estatísticas t correspondentes e as probabilidades a favor da hipótese nula de nulidade de cada uma das estimativas.

**Tabela 6: Estimativas dos parâmetros do modelo da mortalidade e da área basal pelo método NSUR.**

Parâmetros	Estimativas	Desvios-Padrão	Estatística-t	Pr >  t
amnp	0,003952	0,000949	4,16	<0,0001
agq	0,12265	0,00658	18,64	<0,0001
ng0	1,938731	0,1429	13,56	<0,0001
ngq	-0,40419	0,0447	-9,05	<0,0001
ngn	0,053529	0,00505	10,60	<0,0001
kgq	136,2317	2,9023	46,94	<0,0001

Verifica-se pela observação dos resultados da estimação da Tabela 6, que se rejeitam as hipóteses de nulidade de cada um dos parâmetros estimados de cada equação, com um grau de significância de 5% (a estatística t é testada com uma t-student com  $n - K = 1436 - 1 = 1435$  graus de liberdade para a equação da mortalidade e  $n - K = 1436 - 5 = 1431$  graus de liberdade para a equação da área basal), sendo a probabilidade a favor de  $H_0$  para todos os parâmetros inferior a 0,0001. Consequentemente, rejeita-se também a hipótese de nulidade conjunta dos parâmetros de cada uma das equações do modelo, uma vez que bastava que pelo menos uma das variáveis fosse estatisticamente diferente de zero.

O modelo apresenta uma eficiência da modelação  $R^2 = 0,9356$  para a equação da mortalidade (pelo que se pode considerar um modelo muito bom, apresenta um valor próximo de 1, traduzindo uma alta percentagem de explicação da variabilidade da variável endógena pelas variáveis exógenas e endógena desfasada) e um coeficiente de determinação  $R^2 = 0,9862$  para a equação da área basal (pelo que à partida se pode considerar uma percentagem elevada de explicação da variabilidade da variável endógena  $G_t$ , pelas variáveis exógenas, pela variável endógena desfasada  $G_{t-1}$ , a variável endógena  $N_t$  e a variável endógena desfasada  $N_{t-1}$ ), não estando garantido à partida que estes coeficientes estejam para este método (NSUR) obrigatoriamente entre 0 e 1.

As equações apresentam as seguintes raízes do erro quadrático médio:

- Equação da mortalidade:  $RMSE_1 = \left[ \frac{1}{n-K} \sum_{t=1}^n (N_t - \hat{N}_t)^2 \right]^{1/2} = 203,5 \text{ arv/ha}$

$= \hat{\sigma}_{N_t}$ . À primeira vista ser considerado um desvio padrão alto estimado para  $\sigma$ , um ajustamento não muito preciso, mas tem de ter-se em atenção os valores altos observados para a variável dependente.

- Equação da área basal:  $RMSE_2 = \left[ \frac{1}{n-K} \sum_{t=1}^n (G_t - \hat{G}_t)^2 \right]^{1/2} = 1,1756 \text{ m}^2/\text{ha}$

$= \hat{\sigma}_{G_t}$ . Um desvio padrão estimado baixo, o que confirma o bom ajustamento.

Pode-se ainda utilizar mais uma medida para avaliar a bondade do ajustamento, a média dos desvios estimados:

Equação da mortalidade:  $m_1 = \frac{1}{n} \sum_{t=1}^n (N_t - \hat{N}_t) = 8,7539 \text{ arv/ha}$

Equação da área basal:  $m_2 = \frac{1}{n} \sum_{t=1}^n (G_t - \hat{G}_t) = 0,2265 \text{ m}^2/\text{ha}$

Resultando um desvio médio muito baixo para a equação da área basal e um mais elevado para a equação da mortalidade (mas bastante bom se tivermos em conta a grandeza dos valores da variável dependente). Tendo em conta os valores amostrais,

os valores dos desvios não são muito elevados, o que revela tratar-se de um bom ajustamento em ambas as equações.

Para verificar a existência de colinearidade na equação da mortalidade calculam-se os coeficientes de correlação entre os seus pseudo-regressores dados pelas seguintes fórmulas (calculadas nos valores ótimos dos parâmetros):

$$x_{1N_t}^0 = \frac{\partial N_t}{\partial amnp} = -\frac{Npl^2}{1000} t \exp\left\{-amnp \frac{Npl}{1000} t\right\}$$

caso estejamos nas condições da Inicialização e

$$x_{1N_t}^0 = \frac{\partial N_t}{\partial amnp} = -\frac{Npl}{1000} N_{t-1} \exp\left\{-amnp \frac{Npl}{1000}\right\}$$

caso estejamos nas condições da projeção.

Neste caso como só há um pseudo-regressor não faz sentido esta análise.

**Tabela 7: Correlações de Pearson entre os pseudo-regressores da equação da mortalidade.**

Pseudo-regressores	$x_1^0$
$x_1^0$	1,00000

Para verificar a existência de colinearidade na equação da área basal calcularam-se os coeficientes de correlação entre os seus pseudo-regressores dados pelas seguintes fórmulas (calculadas nos valores ótimos dos parâmetros):

$$x_{1G_t}^0 = \frac{\partial G_t}{\partial Agq} = iqe^2 \exp\left\{-k_g \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\}\right\}$$

$$x_{2G_t}^0 = \frac{\partial G_t}{\partial kgQ} = -Agq.iqe \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\} \exp\left\{-k_g \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\}\right\}$$

$$x_{3G_t}^0 = \frac{\partial G_t}{\partial ng0} = -Agq.iqe^2 k_g \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\} \exp\left\{-k_g \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\}\right\}$$

$$x_{4G_t}^0 = \frac{\partial G_t}{\partial ngq} = -Agq.iqe^2 k_g \log\left(\frac{1}{t}\right) \log(iqe) \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\} \exp\left\{-k_g \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\}\right\}$$

$$x_{5G_t}^0 = \frac{\partial G_t}{\partial ngn} = -Agq.iqe^2 k_g \frac{N_t}{1000} \log\left(\frac{1}{t}\right) \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\} \exp\left\{-k_g \exp\left\{n_g \log\left(\frac{1}{t}\right)\right\}\right\} ,$$

caso estejamos nas condições da Inicialização e

$$x_{1G_t}^0 = \frac{\partial G_t}{\partial Agq} = iqe^2 \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}} - iqe^2 \exp\{ng_{t-1} \log(t-1) - ng_t \log(t)\} \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}}$$

$$x_{3G_t}^0 = \frac{\partial G_t}{\partial ng0} = -Agq.iqe^2 \log(t)$$

$$\exp\{ng_{t-1} \log(t-1) - ng_t \log(t)\} \log\left(\frac{G_{t-1}}{Agq.iqe^2}\right) \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}} + Agq.iqe^2 \log(t-1)$$

$$\exp\{ng_{t-1} \log(t-1) - ng_t \log(t)\} \log\left(\frac{G_{t-1}}{Agq.iqe^2}\right) \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}}$$

$$x_{4G_t}^0 = \frac{\partial G_t}{\partial ngq} = -Agq.iqe^2 \log(iqe) \log(t)$$

$$\exp\{ng_{t-1} \log(t-1) - ng_t \log(t)\} \log\left(\frac{G_{t-1}}{Agq.iqe^2}\right) \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}} +$$

$$Agq.iqe^2 \log(iqe) \log(t-1)$$

$$\exp\{ng_{t-1} \log(t-1) - ng_t \log(t)\} \log\left(\frac{G_{t-1}}{Agq.iqe^2}\right) \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}}$$

$$x_{5G_t}^0 = \frac{\partial G_t}{\partial ngn} = -Agq.iqe^2 \frac{N_t}{1000} \log(t)$$

$$\exp\{ng_{t-1} \log(t-1) - ng_t \log(t)\} \log\left(\frac{G_{t-1}}{Agq.iqe^2}\right) \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}} +$$

$$Agq.iqe^2 \frac{N_{t-1}}{1000} \log(t-1)$$



$\exp\{ng_{t-1} \log(t-1) - ng_t \log(t)\} \log\left(\frac{G_{t-1}}{Agq.iqe^2}\right) \left(\frac{G_{t-1}}{Agq.iqe^2}\right)^{\frac{(t-1)^{ng_{t-1}}}{t^{ng_t}}}$ , caso estejamos nas condições de projeção.

Na tabela 8 apresentam-se as correlações calculadas recorrendo ao “SAS base” entre os pseudo-regressores da equação da área basal. Verifica-se que há indícios da existência de colinearidade nesta equação, devido à existência de fortes correlações entre alguns pseudo-regressores (assinaladas a sombreado). Tal ocorrência pode explicar-se devido à natureza da equação, que utiliza como regressores a variável  $N_t$  e a mesma variável desfasada.

**Tabela 8: Correlações de Pearson entre os pseudo-regressores da equação da área basal.**

Pseudo-regressores	$X_1^0$	$X_2^0$	$X_3^0$	$X_4^0$	$X_5^0$
$X_1^0$	1,00000				
$X_2^0$	-0,96255	1,00000			
$X_3^0$	-0,74284	1,00000	1,00000		
$X_4^0$	0,96395	-0,98367	-0,76893	1,00000	
$X_5^0$	0,82840	-0,79280	-0,49714	0,76719	1,00000

As tabelas apresentam as matrizes de covariâncias e correlações dos resíduos estimados para as duas equações (Tabela 9 e Tabela 10), necessárias para testar a existência de correlação contemporânea dos erros das equações, que se fará de seguida.

**Tabela 9: Estimativas das covariâncias dos resíduos estimados das equações.**

Equações	$N_t$	$G_t$
$N_t$	41414,46	24,66
$G_t$	24,66	1,38

**Tabela 10: Estimativas das correlações dos resíduos estimados das equações.**

Equações	$N_t$	$G_t$
$N_t$	1	0,103073
$G_t$	0,103073	1

Por construção o método de estimação usado NSUR tem em conta na estimação dos parâmetros a correlação contemporânea dos erros das equações, produzindo também por isso estimativas mais eficientes e consistentes que NLLS e N3SLS. Mas para confirmar se há correlação contemporânea vai fazer-se o teste dos multiplicadores de Lagrange (Griffiths et *al.*, 1993):

$$H_0: \sigma_{21} = 0 \text{ vs } H_1: \sigma_{21} \neq 0.$$

A estatística de teste é dada por  $\lambda_{LM} = n \sum_{i=2}^M \sum_{j=1}^{i-1} r_{ij}^2$ .

$$\lambda_{LM} \overset{a}{\sim} X_{m(m-1)/2}^2$$

onde  $n$  é o número de observações,  $m$  é o número de variáveis endógenas no sistema,  $M$  o número de equações no sistema e  $r_{ij}^2$  é o coeficiente de correlação dos resíduos estimados da  $i$ -ésima equação com os da  $j$ -ésima equação.

Deste modo  $\lambda_{LM} = nr_{21}^2 = 1436(0,103073)^2 = 15,25606$  e o valor crítico de uma Qui-quadrado com  $2 \times 1/2 = 1$  grau de liberdade e com um nível de confiança de 95% é igual a  $3,84 < 15,25606$ , logo rejeita-se a hipótese de ausência de correlação contemporânea entre as equações do sistema e portanto confirma-se a correcta aplicação do método NSUR.

Apresentam-se de seguida as matrizes de Covariâncias e Correlações estimadas dos parâmetros do sistema (Tabelas 11 e 12).

Da análise das tabelas pode salientar-se a forte correlação entre as estimativas dos parâmetros  $ngq$  e  $ng0$  (superior em módulo a 0,75).

**Tabela 11: Estimativas das covariâncias dos parâmetros do sistema de equações.**

Parâmetros	amnp	agq	ng0	ngq	ngn
amnp	$9,0144 \times 10^{-7}$	-0,0000001	-0,0000002	0,0000001	-0,0000001
agq	$-8,301 \times 10^{-8}$	0,0000433	-0,0001472	-0,0000007	-0,0000112
ng0	$-1,903 \times 10^{-7}$	-0,0001472	0,0204276	-0,0062797	-0,0000652
ngq	$1,2132 \times 10^{-7}$	-0,0000007	-0,0062797	0,0019962	0,0000212
ngn	$-6,845 \times 10^{-8}$	-0,0000112	-0,0000652	0,0000212	0,0000255
kgq	$-6,47 \times 10^{-6}$	-0,0087670	0,1096089	-0,0172460	0,0000619

Parâmetros	Kgq
amnp	-0,000006
agq	-0,008767
ng0	0,109609
ngq	-0,017246
ngn	0,000062
kgq	8,423331

Tabela 12: Estimativas das correlações dos parâmetros do sistema de equações.

Parâmetros	amnp	agq	ng0	ngq	ngn
amnp	1,0000				
agq	-0,0133	1,0000			
ng0	-0,0014	-0,1565	1,0000		
ngq	0,0029	-0,0025	-0,9834	1,0000	
ngn	-0,0143	-0,3377	-0,0903	0,0938	1,0000
kgq	-0,0023	-0,4590	0,2642	-0,1330	0,0042

Parâmetros	kgq
amnp	
agq	
ng0	
ngq	
ngn	
kgq	1,0000

Em resumo pode-se considerar um bom ajustamento através de um sistema de equações simultâneas não lineares, apesar da existência de colinearidade nas equações.

## CAPITULO 4

### 4 Estimação do Modelo com Recurso a Técnicas de Estatística Bayesiana

#### 4.1 Princípios de Estatística Bayesiana

##### 4.1.1 Introdução

De acordo com os princípios clássicos, o parâmetro  $\theta$  é tido como desconhecido, mas fixo e é com o objetivo de se estimar o verdadeiro valor de  $\theta$  que se desenvolve a inferência estatística. Considera-se uma amostra  $(X_1, X_2, \dots, X_n)$  da população indexada por  $\theta$  e é com base nos valores observados da amostra  $(x_1, x_2, \dots, x_n) \in \mathfrak{R}^n$  que é baseada toda a informação sobre  $\theta$ .

A abordagem Bayesiana é uma extensão do modelo clássico, que tem raiz numa divergência fundamental: o parâmetro  $\theta \in \Theta$  passa a ser tomado como variável aleatória (ou vector aleatório), não observável; segundo esta abordagem, o que é desconhecido é incerto e toda a incerteza tem de ser quantificada em termos de probabilidade. Admite-se que a informação inicial (antes de realizada a experiência) ou informação *a priori* é demasiado importante para ser ignorada e pode traduzir-se por uma distribuição de probabilidade. Esta informação, que é geralmente subjetiva e reflecte todo o conhecimento que se tem acerca do parâmetro antes de realizada a

experiência, é quantificada através da distribuição *a priori*,  $h(\theta)$ . No caso contínuo,  $h(\theta)d\theta$  exprime o grau de credibilidade que se atribui ao intervalo  $(\theta, \theta + d\theta)$ .

A componente amostral ou experimental é comum às abordagens clássica e bayesiana, assim a família  $\{F_\theta : \theta \in \Theta\}$  também faz parte do modelo bayesiano, para além do domínio dos possíveis valores para  $\theta$ ,  $\Theta \subseteq \mathfrak{R}^k$ . Se considerarmos um qualquer elemento de  $F_\theta$ , seja  $f(x|\theta)$  e a distribuição *a priori*  $h(\theta)$ , o teorema de Bayes para densidades conduz à relação,

$$h(\theta|x) = f(x|\theta)h(\theta) / \int_{\Theta} f(x|\theta)h(\theta)d\theta, \theta \in \Theta,$$

onde  $h(\theta|x)$  é a distribuição *a posteriori* de  $\theta$ . Tendo em conta a observação  $x$ , a informação inicial acerca de  $\theta$  é actualizada passando a ser dada por  $h(\theta|x)$ . A distribuição marginal de  $X$ , também chamada de função preditiva, não depende de  $\theta$ , o seu inverso representa uma constante normalizadora de  $h(\theta|x)$  e é dada por:

$$f(x) = \int_{\Theta} f(x|\theta)h(\theta)d\theta.$$

Assim é frequente escrever-se

$$h(\theta|x) \propto f(x|\theta)h(\theta),$$

onde  $\propto$  representa proporcionalidade entre as expressões e  $f(x|\theta)$  é a função de verosimilhança, isto é:

$$\text{distribuição } a \text{ posteriori} \propto \text{verosimilhança} \times \text{distribuição } a \text{ priori}.$$

A função de verosimilhança tem um papel fundamental na fórmula de Bayes, representa o meio pelo qual se procede à revisão das credibilidades *a priori* com a utilização da informação amostral, ou seja, transforma o conhecimento *a priori* sobre  $\theta$ . Na abordagem bayesiana a distribuição *a posteriori* serve de base a todas as inferências, pois ela incorpora toda a informação sobre o parâmetro

#### 4.1.2 Distribuições *a priori*

A informação *a priori* a incorporar na análise é uma informação possuída por alguém considerado um especialista (perito, *expert*) do problema concreto, investigador, estatístico ou outrem e que pode ser bastante subjetiva. As fontes destes elementos subjetivos são por vezes dados históricos do problema ou de problemas análogos ou factos.

Essa informação *a priori* vai ser reflectida na distribuição *a priori* de  $\theta$ , conhecida antes de se realizar a experiência, representando as credibilidades iniciais sobre  $\theta$ .

A escolha de uma distribuição, que represente a informação que o eliciador tem sobre o parâmetro, pode ser feita em função da conveniência de simplificação da dedução analítica da distribuição *a posteriori* e conseqüentemente na diminuição do esforço computacional. Nesta perspectiva a seleção da distribuição *a priori* adequada que incorpore as crenças existentes sobre  $\theta$  deve ser feita na família conjugada da distribuição  $f(x|\theta)$ .

As distribuições *a priori* conjugadas (da família conjugada) são provenientes da imposição de que a família de distribuições *a priori*  $H$  seja fechada sobre amostragem de qualquer elemento  $F = \{f(x|\theta) : \theta \in \Theta\}$ , isto é que,

$$h(\theta) \in H \Rightarrow h(\theta|x) \propto h(\theta)f(x|\theta) \in H.$$

Nestas condições diz-se também que  $H$  é uma família conjugada de  $F$ .

Muitas das distribuições usadas para representar a informação amostral pertencem à família exponencial (como por exemplo: Binomial, Binomial Negativa, Poisson, Normal, Gama), para a qual sabemos existirem distribuições *a priori* conjugadas.

Quando não existe informação de natureza objetiva nem subjetiva sobre  $\theta$ , ou quando essa informação é escassa, diz-se que se está num estado de ignorância *a priori* ou num estado de conhecimento vago, respectivamente. Nessas condições em que o conhecimento *a priori* é pouco significativo relativamente à informação amostral devem utilizar-se distribuições *a priori* não informativas. Estas distribuições têm um papel relevante para os casos em que: se pretende deduzir as crenças *a posteriori* para quem parte de um conhecimento escasso; se pretende comparar resultados com os da inferência clássica que só usa a informação amostral e averiguar a influência nas inferências da distribuição *a priori* subjetiva que descreve a informação realmente existente, quando confrontada com as que resultam do uso da distribuição *a priori* de referência.

Neste trabalho vamos focar apenas alguns métodos de construção de distribuições *a priori* não informativas. Começamos por apresentar as propostas do método de Bayes-Laplace. Este método baseia-se no “Princípio da Razão Insuficiente” de Bayes e Laplace, segundo o qual na ausência de razão suficiente adota-se a equiprobabilidade. Aqui coloca-se um problema, caso o espaço do parâmetro seja ilimitado, a distribuição é imprópria, i.e.,

$$\int_{\Theta} h(\theta) d\theta = \infty ,$$

o que não causará problemas se a distribuição *a posteriori* for própria. Quando a distribuição *a posteriori* não é própria fica inviabilizado o processo de inferência.

Mas a principal objeção da representação da ignorância por distribuições uniformes é a sua inconsistência no sentido em que se  $\psi = \psi(\theta)$  é uma transformação injetiva de um parâmetro  $\theta$  assumindo uma gama contínua de valores possíveis, as distribuições



uniformes para  $\theta$  e  $\psi$  não são, em regra, probabilisticamente compatíveis. Sendo  $h(\theta)$  uma distribuição *a priori* para  $\theta$ ,

$$h(\psi) = h(\theta(\psi)) \left| \frac{d\theta}{d\psi} \right|$$

é a correspondente distribuição para a reparametrização injetiva  $\psi$ , que não é necessariamente uniforme quando  $h(\theta)$  o é.

A crítica da inconsistência da distribuição uniforme na representação formal da ignorância suscita que esta deva ser invariante sob transformações injetivas.

Entre os procedimentos que asseguram a invariância destaca-se o de Jeffreys e que se baseia no uso da medida de informação de Fisher sobre  $\theta \in \mathfrak{R}$ ,

$$I(\theta) = E \left[ \left( \frac{\partial \log(f(x|\theta))}{\partial \theta} \right)^2 \mid \theta \right] .$$

Para qualquer transformação real injetiva de  $\theta \in \mathfrak{R}$  tem-se

$$I(\psi) = I(\theta(\psi)) \left( \frac{d\theta}{d\psi} \right)^2 ,$$

mostra que a distribuição proposta por Jeffreys para o caso uniparamétrico,

$$h(\theta) \propto [I(\theta)]^{1/2} ,$$

goza da referida propriedade de invariância. A extensão ao caso multiparamétrico faz-se utilizando a raiz quadrada do determinante da matriz  $I(\theta)$ .

$I(\theta)$  é tanto maior quanto maior for a taxa quadrática de variação com  $\theta$  de  $\log[f(x|\theta)]$ , isto é, quanto mais diferenciado pelo modelo estiver  $\theta$  de  $\theta + d\theta$ .

Devem-se considerar mais (menos) plausíveis *a priori* os valores de  $\theta$  com maior (menor)  $I(\theta)$ , isto é, sobre os quais há maior (menor) informação amostral correspondente a considerar, reduzindo tanto quanto possível o efeito da informação *a priori*.

Seguidamente apresentam-se algumas distribuições *a priori* não informativas que são sugeridas consoante o modelo pertence á família de localização, de escala ou de localização-escala:

i) A densidade  $X$  envolve apenas  $\theta$  através do termo  $(x - \theta)$ , ou seja,  $f(x|\theta) = f_0(x - \theta)$ . Nestas ocasiões,  $\theta$  designa-se por parâmetro de localização e uma distribuição *a priori* não informativa para  $\theta$ , invariante sobre transformações de localização (transformações do tipo  $Y = X + c$ , sendo  $c$  uma constante), é a distribuição uniforme no intervalo de variação de  $\theta$ . Caso tenhamos  $\Theta = \Re$  pode utilizar-se  $h(\theta) = 1, \theta \in \Re$ ;

ii) A densidade  $X$  é da forma  $f(x|\theta) = \frac{1}{\theta} f_0\left(\frac{x}{\theta}\right)$ , com  $\theta > 0$ . Nesta situação  $\theta$  designa-se por parâmetro de escala e uma distribuição *a priori* não informativa invariante a transformações de escala (transformações do tipo  $Y = cX$ , sendo  $c$  uma constante) é dada por  $h(\theta) = \frac{1}{\theta}, \theta \in \Re^+ (\Theta = \Re^+)$ ;

iii) A densidade  $X$  combina parâmetros de localização e escala apresentando a forma  $f(x|\theta_1, \theta_2) = \frac{1}{\theta_2} f_0\left(\frac{x - \theta_1}{\theta_2}\right)$ , sendo neste caso  $\theta_1$  ( $\Theta_1 = \Re$ ) o parâmetro de localização e  $\theta_2$  ( $\Theta_2 = \Re^+$ ) o parâmetro de escala. Para este caso admitindo que a ignorância *a priori* é compatível com a independência entre as distribuições dos parâmetros, a distribuição *a priori* conjunta para o vector de parâmetros  $(\theta_1, \theta_2)$  é dada pelo produto das distribuições *a priori*

não informativas para os parâmetros em causa,  $h(\theta_1, \theta_2) = \frac{1}{\theta_2}$ ,  
 $\theta_1 \in \mathfrak{R} \wedge \theta_2 \in \mathfrak{R}^+$  ( $\Theta = \mathfrak{R} \times \mathfrak{R}^+$ ).

O método proposto por Box e Tiao (1973) não reconhece um estado de ignorância absoluta e encara a ignorância das crenças *a priori* unicamente em relação à quantidade de informação da experiência geradora dos dados. A ideia base foi procurar uma reparametrização  $\psi = \psi(\theta)$  do modelo  $\{f(x|\theta) : \theta \in \Theta\}$  para o qual a respectiva verosimilhança fosse apenas trasladada (exata ou, pelo menos, aproximadamente) pelos dados, isto é, tal que:

$$L(\theta | x) \approx g[\psi(\theta) - m(x)] ,$$

onde  $g$  é uma função cuja forma é independente de  $x$  e  $m(x)$  a função que descreve a translação de  $L$  com a variação de  $x$ . Uma vez detectada uma transformação  $\psi$  deste tipo, considera-se para ela essencialmente a verosimilhança normalizada. Assim, as distribuições *a posteriori* de  $\psi$  para diferentes amostras só diferem umas das outras pela localização. Box e Tiao (1973) propuseram para  $\psi$  uma distribuição própria cuja densidade seja sensivelmente constante na gama dos valores apreciáveis de  $L(\psi | x)$ .

$$h(\psi | x) \approx \frac{L(\psi | x)}{\int L(\psi | x)}$$

apresenta uma densidade essencialmente uniforme na gama de valores significativos da verosimilhança, decaindo para zero à medida que se afasta dessa gama, o que justifica a designação de distribuições localmente uniformes.

$$h(\psi) \propto \text{constante},$$

localmente, o que implica que:

$$h(\theta) \propto \left| \frac{d\psi}{d\theta} \right|,$$

aproximadamente proporcional ao Jacobiano da transformação  $\psi(\theta)$ .

Não é demais salientar que, no caso das distribuições *a priori* serem não informativas, as inferências efetuadas a partir da distribuição *a posteriori* são baseadas apenas nos dados – traduzidas pela verosimilhança da amostra - permitindo a comparação destas inferências com as da Estatística Clássica.

### 4.1.3 Inferência Paramétrica

A inferência paramétrica é baseada na distribuição *a posteriori* e inclui a obtenção de estimativas pontuais dos parâmetros, estimativas por intervalos (ou regiões, no caso multiparamétrico) e ainda a formulação de testes de hipóteses. Assim é fundamental analisar as características da distribuição *a posteriori*.

#### i) Estimação pontual

Qualquer medida de localização da distribuição *a posteriori* - valor médio, mediana e moda – podem ser utilizadas como parâmetro. Deve escolher-se a medida sumária da distribuição *a posteriori* que seja mais adequada tendo em conta a forma de  $f(x|\theta)$  e o objetivo do estudo. Por uma questão de simplicidade, desde que a distribuição não apresente assimetrias muito elevadas, é privilegiado o uso da média *a posteriori*.

Daí a importância em saber como calcular valores esperados *a posteriori*. Assim para o cálculo de valores esperados *a posteriori* de uma dada função de  $\theta$ , seja  $g(\theta)$ , temos a seguinte expressão para a média *a posteriori* de  $g(\theta)$ :

$$E_{\theta|x}(g(\theta) | x) = \int_{\Theta} g(\theta)h(\theta | x)d\theta$$

No caso em que  $g(\theta) = \theta$ , ou seja, pretendemos obter  $\hat{\theta} = E_{\theta|x}(\theta | x)$  temos:

$$E_{\theta|x}(\theta | x) = \int_{\Theta} \theta h(\theta | x) d\theta$$

## ii) Estimação por regiões

Para se obter uma análise mais informativa de  $h(\theta | x)$  que a estimação pontual, pode-se recorrer à estimação por regiões do suporte de  $\theta$ ,  $\Theta$ .

**Definição:**  $R(x)$  é uma região de credibilidade  $\gamma$  para  $\theta$  se

$$P[\theta \in R(x) | x] = \int_{R(x)} h(\theta | x) d\theta \geq \gamma.$$

Uma região de credibilidade é definida numericamente, não aleatória e permite uma interpretação probabilística direta e inequívoca.

A fim de se garantir que se escolhe a região de credibilidade que engloba todos os valores de  $\theta$  mais credíveis *a posteriori*, ou seja que se garanta que

$$h(\theta_1 | x) \geq h(\theta_2 | x), \quad \forall \theta_1 \in R(x), \theta_2 \in R(x),$$

temos a seguinte definição.

**Definição:**  $R(x)$  é a região de credibilidade  $\gamma$  com densidade a posteriori máxima (região HPD) se

$$\sup_{\theta \in R(x)} h(\theta | x) \leq c_\gamma \leq \inf_{\theta \in R(x)} h(\theta | x)$$

para algum  $c_\gamma$  tal que  $P[\theta \in R(x) | x] \geq \gamma$ , ou equivalentemente, se

$$R(x) = \{\theta : h(\theta | x) \geq c_\gamma\}$$

com  $c_\gamma > 0$  a maior constante tal que

$$\int_{R(x)} h(\theta | x) \lambda(d\theta) \geq \gamma$$

(relativamente à medida  $\lambda$  em  $\Theta$  em relação à qual a medida a priori apresenta a densidade  $h(\theta)$ ).

### iii) Testes de Hipóteses

No âmbito Bayesiano quando se pretende testar  $H_0: \theta \in \Theta_0$  vs  $H_1: \theta \in \Theta_1 = \Theta - \Theta_0$ , atendendo à interpretação probabilística directa das hipóteses em confronto, basta calcular as respectivas probabilidades *a posteriori* e a decisão de optar por uma delas assenta na sua grandeza relativa, que é traduzida pela quantidade

$$O(H_0, H_1 | x) = P\{H_0 | x\} / P\{H_1 | x\}$$

que se denomina de razão (rácio) de vantagens (chances) *a posteriori* a favor de  $H_0$ .

Sempre que esta quantidade for maior que 1, não há evidências para rejeitar  $H_0$ , caso contrário poderá rejeitar-se  $H_0$ .

Se se pretender medir a influência dos dados correntes  $x$  na alteração da credibilidade relativa de  $H_0$  e  $H_1$  opta-se frequentemente por contrapor a razão *a posteriori* à respectiva razão *a priori*,  $O(H_0, H_1) = P\{H_0\} / P\{H_1\}$ . Ao rácio entre estas duas razões

$$B(x) = O(H_0, H_1 | x) / O(H_0, H_1)$$

chama-se factor de Bayes a favor de  $H_0$ . Portanto um factor de Bayes muito grande (muito pequeno) comparado com a unidade, reflete evidência a favor de  $H_0$  ( $H_1$ ).

#### 4.1.4 Fundamentos de Métodos de Monte Carlo

De acordo com o descrito anteriormente nesta secção, toda a inferência sobre o vector de parâmetros  $\theta$  é feita baseada na distribuição *a posteriori*

$$h(\theta | x) = f(x | \theta)h(\theta) / \int_{\Theta} f(x | \theta)h(\theta)d\theta ,$$

sendo as inferências sumárias obtidas a partir dos valores esperados *a posteriori*

$$E_{\theta|x}[g(\theta)] = \int_{\Theta} g(\theta)h(\theta | x)d\theta$$

para a escolha adequada de  $g(\cdot)$ . Por exemplo se se pretender obter a média ou variância *a posteriori*, considera-se  $g(\theta) = \theta_i$  ou  $g(\theta) = (\theta_i - E_{\theta|x}[\theta_i])^2$ , respetivamente.

Verifica-se que a operação de integração desempenha um papel fulcral em toda a inferência Bayesiana. Porém, por vezes as expressões a integrar não são de cálculo fácil ou não é possível obter expressões explícitas para os integrais em causa. Nestes casos recorre-se normalmente a alguns métodos analíticos, como a aproximação da distribuição *a posteriori* à distribuição normal multivariada, ao método de Laplace de aproximação de integrais ou a métodos de integração numérica (quadratura iterativa).

Uma alternativa apropriada é a recorrência a métodos de Monte Carlo para a aproximação de integrais. Estes métodos são métodos de simulação estocástica, baseando-se na reprodução de valores gerados aleatoriamente de distribuições de probabilidade.

### **i) Monte Carlo Simples**

Sempre que se possa simular uma amostra aleatória  $\theta_1, \theta_2, \dots, \theta_N$  da densidade *a posteriori*  $h(\theta | x)$ , este método aproxima o integral de  $E_{\theta|x}[g(\theta)]$  (como definido anteriormente) pela média empírica

$$\hat{E}_{\theta|x}[g(\theta)] = \frac{1}{N} \sum_{i=1}^N g(\theta_i),$$

que converge quase certamente para  $E_{\theta|x}[g(\theta)]$ , quando  $N \rightarrow \infty$  segundo a Lei Forte dos Grandes Números (mencionada no ponto 3.1.1).

A precisão desta estimativa pode ser medida pelo desvio padrão estimado de Monte Carlo:

$$\frac{1}{\sqrt{N(N-1)}} \left\{ \sum_{i=1}^N \left[ g(\theta_i) - \frac{1}{N} \sum_{i=1}^N g(\theta_i) \right]^2 \right\}^{1/2} .$$

## ii) Monte Carlo via Cadeias de Markov (MCMC)

A simulação de valores com a distribuição  $h(\theta | x)$  nem sempre é tarefa fácil, a utilidade do MCMC está na transformação do problema estático em consideração num problema de natureza dinâmica, através de um processo estocástico temporal preferencialmente fácil de simular e que convirja para a distribuição original. Geralmente é uma cadeia de Markov homogénea cuja distribuição de equilíbrio é a distribuição da qual se pretende simular amostras.

O conhecimento de cadeias de Markov é fulcral para a compreensão do método. Visto que se pretende apenas expôr a teoria necessária para a compreensão deste estudo, a teoria das cadeias de Markov necessária para a compreensão integral do MCMC não será abordada, pode ser vista por exemplo em Paulino *et al.*, 2003.

O algoritmo Metropolis-Hastings serve o propósito de construção de cadeias de Markov com distribuições de equilíbrio específicas. Os denominados MCMC representam a associação deste algoritmo ou algumas das variantes (ver Paulino *et al.*, 2003) para a simulação de distribuições ao método de Monte Carlo para a aproximação de integrais.

O método de amostragem de Gibbs pode ser interpretado como um caso particular do algoritmo Metropolis-Hastings.



### iii) Algoritmo de Gibbs

O algoritmo de Gibbs baseia-se no facto de que se a distribuição conjunta  $h(\theta | x)$  for positiva em  $\Theta_1 \times \Theta_2 \times \dots \times \Theta_k$ , sendo  $\Theta_i$  o suporte da distribuição de  $\theta_i$  para  $i = 1, \dots, k$ , então ela é unicamente determinada pelas distribuições condicionais completas  $h(\theta_i | x, \theta(-i))$ ,  $i = 1, \dots, k$  onde  $\theta = (\theta_1, \dots, \theta_k)$  e  $\theta(-i)$  representa o vector  $\theta$  sem a  $i$ -ésima componente, isto é,  $\theta(-i) = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$ .

Dado um valor inicial arbitrário  $\theta^0 = (\theta_1^{(0)}, \dots, \theta_k^{(0)})$  (estado inicial da cadeia de Markov pertencente ao suporte  $\Theta_1 \times \Theta_2 \times \dots \times \Theta_k$ ), procede-se iterativamente da seguinte forma:

1. Obtém-se  $\theta_1^{(1)}$  de  $h(\theta_1 | x, \theta_2^{(0)}, \dots, \theta_k^{(0)})$ ,  
 obtém-se  $\theta_2^{(1)}$  de  $h(\theta_2 | x, \theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)})$ ,  
 obtém-se  $\theta_3^{(1)}$  de  $h(\theta_3 | x, \theta_1^{(1)}, \theta_2^{(1)}, \theta_4^{(0)}, \dots, \theta_k^{(0)})$ ,  
 ...  
 obtém-se  $\theta_k^{(1)}$  de  $h(\theta_k | x, \theta_1^{(1)}, \dots, \theta_{k-1}^{(1)})$ .

Completa-se assim uma iteração do esquema e uma transição de  $\theta^{(0)}$  para  $\theta^1 = (\theta_1^{(1)}, \dots, \theta_k^{(1)})$ ;

2. O esquema anterior é repetido com  $\theta^1$  anteriormente obtido, como vector inicial, para se obter um novo vector  $\theta^2$  e assim haver uma transição de  $\theta^1$  para  $\theta^2 = (\theta_1^{(2)}, \dots, \theta_k^{(2)})$ ;
3. Itera-se  $t$  vezes este ciclo de geração de observações aleatórias de cada uma das distribuições condicionais, produzindo  $\theta^1, \dots, \theta^t$ .

A sucessão  $\theta^0, \dots, \theta^t, \dots$  é uma realização de uma cadeia de Markov com espaço de estados  $\Theta$  e função de transição (tal como definida em Paulino *et al*, 2003) dada por

$$q(\theta^t, \theta^{t+1}) = \prod_{i=1}^k h(\theta_i^{(t+1)} | x, \theta_j^t, j > i, \theta_j^{(t+1)}, j < i).$$

Resultados apresentados em Paulino *et al* (2003) sobre cadeias de Markov permitem concluir que quando  $t \rightarrow \infty$ ,  $(\theta_1^{(t)}, \dots, \theta_k^{(t)})$  tende em distribuição para um vector aleatório cuja função densidade de probabilidade conjunta é  $h(\theta | x)$ . Particularizando,  $\theta_i^{(t)}$  tende em distribuição para uma quantidade aleatória cuja densidade é  $h(\theta_i | x)$  (densidade marginal *a posteriori* de  $\theta_i$ ) e

$$\frac{1}{t} \sum_{i=1}^t g(\theta^i) \xrightarrow{q.c.} E_{\theta|x}[g(\theta) | x],$$

para qualquer função  $g(\cdot)$ , onde  $E_{\theta|x}[g(\theta) | x]$  representa o valor esperado de  $g(\theta)$  em relação à distribuição *a posteriori*  $h(\theta | x)$ .

Apresenta-se de seguida uma variante deste algoritmo, o algoritmo de Gibbs com o passo de Metropolis, que foi utilizado neste trabalho, visto conhecer-se apenas os núcleos das distribuições (condicionais completas dos parâmetros e *a posteriori*) e estes não sugerirem distribuições conhecidas.

#### iv) Gibbs com o Passo de Metropolis

Este algoritmo resulta da utilização do algoritmo de Metropolis-Hastings dentro do passo de condicionamento do método Gibbs, para se simular de (por exemplo)  $h(\theta_i | \theta^t(-i), x)$  em cada iteração  $t$ .

Assumindo que o processo se encontra na iteração  $t$  e que se pretende obter uma nova observação  $\theta_i^t$  de  $h(\theta_i | \theta^t(-i), x)$ , em que  $\theta^t(-i) = (\theta_1^t, \dots, \theta_{i-1}^t, \theta_{i+1}^{t-1}, \dots, \theta_k^{t-1})$ , esta vai ser obtida usando o algoritmo Metropolis-Hastings para simular de  $h(\theta_i | \theta^t(-i), x)$ , construindo-se uma subcadeia de comprimento  $T$ .

Seja  $\pi(x) = h(x | \theta^t(-i))$ , em que  $x$  aqui representa um valor admissível de  $\theta_i$  e  $q(x, y)$  a distribuição "proposta" (envelope ou ainda a função de transição apropriada).

De seguida enumeram-se os passos para a aplicação do algoritmo.

1. Escolhe-se um  $x$  como estado inicial da (sub) cadeia;
2. Obtem-se um  $y$  aleatório de  $q(x, y)$ , onde  $x$  é o estado actual da subcadeia;
3. Aceita-se  $y$  como novo estado da subcadeia quando  $Min\left(\frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1\right) \geq z$ , onde  $z$  é um valor aleatório de uma distribuição Uniforme(0;1). Caso contrário permanece em  $x$ .

A probabilidade de aceitação é dada por

$$\alpha(x, y) = Min\left(\frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1\right) I(\pi(x)q(x, y) > 0) + I(\pi(x)q(x, y) = 0), \text{ onde } I(.) \text{ é}$$

a função indicatriz.

A quantidade  $\theta_i^t$  é então o valor da T-ésima iterada na subcadeia de Metropolis-Hastings. O método foi proposto inicialmente por Muller, 1991 e o valor de T seria um valor determinado em que a subcadeia converge para a distribuição de equilíbrio  $\pi(x)$ . Estudos teóricos e práticos mostram que basta considerar  $T = 1$ .

Caso na primeira iteração o  $y$  seja aceite, então a cadeia passa de  $\theta_i^t, \dots, \theta_{i-1}^t, \theta_i^{t-1}, \theta_{i+1}^{t-1}, \dots, \theta_k^{t-1}$  para  $\theta_i^t, \dots, \theta_{i-1}^t, y, \theta_{i+1}^{t-1}, \dots, \theta_k^{t-1}$ , caso contrário permanece no mesmo estado, passando-se à geração de uma nova componente  $\theta_{i+1}^t$ .

Para a eficácia deste algoritmo é de extraordinária importância uma escolha apropriada para a distribuição envelope  $q(x, y)$ , que deve ser fácil de simular e garantir uma convergência rápida da subcadeia para a distribuição alvo  $\pi(x) = h(x | \theta^t(-i))$  (ter uma taxa de aceitação tanto maior quanto possível). Geralmente são usadas como distribuições proponentes (envelopes) as distribuições Normal, t-Student ou Gama.

Na prática, o que o algoritmo faz é gerar valores com probabilidades de aceitação altas para os quais a distribuição *a posteriori* tem valores altos e probabilidades de aceitação baixas associadas a valores baixos da distribuição *a posteriori*. Se  $y$  é um valor gerado com probabilidade *a posteriori* alta, o algoritmo tende a ficar neste valor.

Um valor candidato é aceite com uma probabilidade alta se se encontra numa região de probabilidade mais alta que o actual enquanto se se encontrar numa região de probabilidade mais baixa tende a ser rejeitado com probabilidade alta. Além disso a fim de poder gerar uma aproximação boa da distribuição, a cadeia tende frequentemente, mas não sempre, a mover-se para zonas de probabilidades *a posteriori* elevadas.

O número total de iterações  $T$  a gerar depende dos seguintes aspectos fundamentais: a velocidade de convergência da cadeia para o estado de equilíbrio, a correlação entre os sucessivos valores gerados e a precisão pretendida para as estimativas dos parâmetros.

É necessário considerar um número  $m$  de iterações que vão ser descartadas, estas são consideradas como fazendo parte do período de aquecimento, não pertencem portanto à distribuição de equilíbrio. Há que ter em conta também, pela maneira como as cadeias são geradas da possibilidade de correlação entre estados (observações) muito próximos, por isso consideram-se estados espaçados de um valor  $h$  (espaçamento de amostragem) após o período de aquecimento. A precisão pretendida pode ser monitorizada pela dimensão da amostra final ( $N$ ).

O método de observação gráfica da evolução das médias ergódicas poderá servir para uma adequada escolha de  $m$  e  $N$  e conseqüentemente  $T$ . O número total de iterações  $T$  necessárias para obter uma amostra de dimensão  $N$ , é  $T = m + h \times N$ . Para a inferência será usado apenas o subconjunto extraído da cadeia, constituído pelos vectores  $\theta^{(m+h)}, \theta^{(m+2h)}, \dots, \theta^{(m+N \times h)}$ .

#### **v) Avaliação da convergência**

Um dos riscos da aplicação destes métodos é o de se considerar que há convergência antes da distribuição de equilíbrio ter sido atingida.

Existem alguns métodos (ver Paulino *et al.*, 2003), uns visuais, como a monitorização das médias ergódicas das cadeias e outros analíticos que implicam cálculos, como o

diagnóstico de Geweke, 1992 (que se apresenta de seguida) que permitem averiguar se as cadeias em questão não convergem.

Seja  $g(\theta)$  a função do parâmetro cujo valor esperado se pretende estimar, a sua trajetória  $g^1, g^2, \dots, g^N$ , construída a partir de  $g^t = g(\theta^t)$ , define uma sucessão cronológica.

Observando a sucessão ao longo de um número  $N$  suficientemente longo de iterações, calculando-se  $g_a = \sum g(\theta^t) / n_a$  à custa das  $n_a$  primeiras iterações e  $g_b = \sum g(\theta^t) / n_b$  à custa das  $n_b$  últimas iterações. A cadeia é estacionária se a média da primeira parte for semelhante à média da última parte da cadeia. Sendo fixos  $n_a / N$  e  $n_b / N$  e  $N \rightarrow \infty$ , pode mostrar-se (Geweke, 1992) que

$$\frac{(g_a - g_b)}{\sqrt{s_a^2 / n_a + s_b^2 / n_b}} \rightarrow N(0,1),$$

onde  $s_a^2$  e  $s_b^2$  são estimativas independentes das variâncias assintóticas de  $g_a$  e  $g_b$ .

Assim dado um valor calculado desta estatística, rejeita-se a não convergência ( $H_0$ ) com um grau de significância de 5% quando o valor-p desta estatística for maior que 5%. Este método encontra-se implementado entre outros softwares no "R" através do código "BOA" (Smith, 2007).

#### 4.1.5 Critério DIC para a seleção de modelos

Para a escolha dos parâmetros relevantes para os modelos, que se querem o mais parcimoniosos possível, recorreu-se neste trabalho ao critério de verosimilhança penalizada DIC (*Deviance Information Criteria*).

A ideia proposta por Akaike, 1973, baseia-se numa relação simples entre a discrepância de Kullback-Liebler e o máximo da log-verosimilhança.

Esta discrepância (ou distância) de Kullback-Liebler entre os modelos  $f$  e  $g$  é definida, para funções contínuas, como o integral (múltiplo), sempre que exista

$$I(f, g) = \int_x f(x) \log \left( \frac{f(x)}{g(x)} \right) dx .$$

É utilizada a notação  $I(f, g)$  pois esta discrepância quantifica a “informação” perdida quando  $g$  é usada para aproximar  $f$ . Encontrar o modelo que origina a menor perda de informação é equivalente a encontrar a função  $g$  que minimiza  $I(f, g)$ .

A discrepância de Kullback-Liebler entre modelos constitui a base teórica, preconizada por Akaike para a seleção de modelos, surgindo (entre outros) o critério DIC por Spiegelhalter *et al.* em 2002.

Na seleção bayesiana, Dempster, 1974 sugere que se utilize a distribuição *a posteriori* da *deviance*. Seguindo esta sugestão, Spiegelhalter *et al.* (2002) propõem o critério de informação da *deviance* para a seleção bayesiana de modelos, definindo a *deviance* Bayesiana do modelo  $M_i$ , com o vector de parâmetros  $\theta$  como sendo

$$D_i(\theta) = -2 \log \left( \frac{f(x | \theta, M_i)}{h(x)} \right),$$

onde  $h(x)$  é uma função apenas dos dados que não tem impacto na escolha do modelo e  $f(x | \theta, M_i)$  é a função densidade de probabilidade do modelo  $M_i$  com vector de parâmetros  $\theta$ . Sendo o objetivo de utilização a seleção de modelos, as funções  $h(x)$  têm de ser idênticas, de modo a não ter influência na escolha do modelo. Nestas condições considera-se neste trabalho  $h(x) = 1$ .

Spiegelhalter *et al.*, (2002) sugerem como medida de adequabilidade do modelo o valor esperado *a posteriori* de  $D_i(\theta)$  e como penalização  $p_{D_i}$  associada à complexidade do modelo a diferença entre este valor esperado e o valor de  $D_i(\theta_i)$  calculado no valor esperado *a posteriori* de  $\theta_i$ . Define-se então:

$$p_{D_i} = E_{(\theta_i | x, M_i)} [D_i(\theta_i)] - D_i(E_{(\theta_i | x, M_i)} [\theta_i]) \text{ e}$$

$$DIC_i = E_{(\theta_i | x, M_i)} [D_i(\theta_i)] + p_{D_i} = 2 \times E_{(\theta_i | x, M_i)} [D_i(\theta_i)] - D_i(E_{(\theta_i | x, M_i)} [\theta_i]).$$

Quando se pretende comparar dois modelos,  $M_1$  e  $M_2$ , pode usar-se a diferença

$$\Delta DIC(x) = DIC_{M_1}(x) - DIC_{M_2}(x),$$

onde  $\Delta DIC$  representa a alteração do valor do DIC do primeiro para o segundo modelo. Se  $\Delta DIC < 0$ , selecciona-se o primeiro modelo; caso contrário selecciona-se o segundo.

Os valores DIC não têm significado intrínseco, têm apenas significado as diferenças dos valores dos DIC entre modelos.

Para o presente trabalho o cálculo dos DIC não é transcendente, obtendo-se uma estimativa de  $E_{(\theta_i|x,M_i)}[D_i(\theta_i)]$ , monitorizando  $D_i(\theta_i)$  e considerando a média amostral dos valores simulados de  $D_i(\theta_i)$ . Após estes cálculos, basta determinar  $D_i(E_{(\theta_i|x,M_i)}[\theta_i])$  por estimativa directa, monitorizando  $\theta_i$  por forma a calcular a sua média amostral ( $\hat{\theta}$ ) e obter  $D_i(\hat{\theta})$ .

## 4.2 Aplicação ao Modelo: Resultados da Estimação Bayesiana

Os modelos bayesianos da altura dominante, área basal e mortalidade foram estimados individualmente, devido à complexidade que implicava a estimação do sistema de equações da área basal e mortalidade, aplicando-se neste caso o princípio da parcimónia.

Para a estimação dos parâmetros dos modelos individuais de cada uma das três variáveis biométricas, recorreu-se à metodologia bayesiana descrita anteriormente, utilizando distribuições *a priori* não informativas propostas por Box e Tiao, de acordo com Koop (2003) que são uniformes para  $\log(\sigma^2)$  e para os restantes parâmetros de cada uma das equações. Devido ao facto de conhecermos apenas os núcleos das distribuições que resultam da aplicação da metodologia Bayesiana e esses não sugerirem nenhuma distribuição conhecida, houve necessidade de se aplicarem Métodos de Monte Carlo via Cadeias de Markov para a estimação dos parâmetros,

aplicando-se a este caso o Método de Metropolis within Gibbs, i.e. Gibbs com o passo de Metropolis, tal como sugere Koop (2003).

## 4.2.1 Desenvolvimento da Metodologia Bayesiana

### i) Modelo de altura dominante

Seja

$$hdom_i = A_h \left( \frac{hdom_{i-1}}{A_h} \right)^{\binom{t-1}{t}^{n_h}} + \varepsilon_i \quad (i = 1, 2, \dots, n) ,$$

como anteriormente definida, em que  $\varepsilon_i \sim N(0, \sigma^2)$ , iid, sendo  $\beta' = [A_h \quad n_h] \in \mathfrak{R}^2$ ,  $\sigma^2 \in \mathfrak{R}^+$  e  $X = [hdom_{i-1} \quad t_i \quad t_{i-1}]$  a matriz dos valores das variáveis independentes.

Por simplificação faz-se:

$$f_i(x_i, \beta) = A_h \left( \frac{hdom_{i-1}}{A_h} \right)^{\binom{t_i-1}{t_i}^{n_h}} .$$

Como

$$hdom_i | x_{i-1}, \beta, \sigma^2 \sim N(f_i(x_i, \beta), \sigma^2) \quad (i = 1, 2, \dots, n),$$

com  $x_{i-1}$  conhecido, tem-se a seguinte função de verosimilhança da amostra:

$$f_{hdom_1, hdom_2, \dots, hdom_n | \beta, \sigma^2, hdom_0}(hdom_1, hdom_2, \dots, hdom_n) = \prod_{i=1}^n \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{1}{2\sigma^2} [hdom_i - f_i(x_i, \beta)]^2\right\}$$

ou

$$f_{hdom_1, hdom_2, \dots, hdom_n | \beta, \sigma^2, hdom_0}(hdom_1, hdom_2, \dots, hdom_n) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n [hdom_i - f_i(x_i, \beta)]^2\right\}$$

Tomando para a distribuição *a priori*, a distribuição não informativa proposta por Box e Tiao:

$$p(\beta, \sigma^2) \propto \frac{1}{\sigma^2}$$



tem-se a distribuição *a posteriori*:

$$[4.2.1] \quad p(\beta, \sigma^2 | hdom) \propto \frac{1}{(2\pi\sigma^2)^{n/2+1}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n [hdom_i - f_i(x_i, \beta)]^2\right\}$$

Note-se que esta distribuição não é conhecida, como tal ter-se-ão de calcular as distribuições condicionais completas para se poder utilizar o método de Monte Carlo via cadeias de Markov. Assim:

$$[4.2.2] \quad p(A_h | n_h, \sigma^2, hdom) \propto \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \left[ hdom_i - A_h \left( \frac{hdom_{i-1}}{A_h} \right)^{\binom{t_{i-1}}{t_i}} \right]^{n_h} \right\}^2$$

$$[4.2.3] \quad p(n_h | A_h, \sigma^2, hdom) \propto \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \left[ hdom_i - A_h \left( \frac{hdom_{i-1}}{A_h} \right)^{\binom{t_{i-1}}{t_i}} \right]^{n_h} \right\}^2$$

$$[4.2.4] \quad p(\sigma^2 | \beta, hdom) \propto \frac{1}{(\sigma^2)^{n/2+1}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n [hdom_i - f_i(x_i, \beta)]^2\right\}$$

Estas distribuições condicionais completas também não são identificadas e não têm uma forma fechada. Pelo que para simular uma amostra de valores destas distribuições usou-se o algoritmo de Gibbs com o passo de Metropolis.

Como distribuições proponentes (ou envelopes) para estas distribuições condicionais completas ([4.2.2] a [4.2.4]), de acordo com a sua representação gráfica a menos da constante de normalização (para este modelo), foram utilizadas as distribuições que se apresentam na Figura 4.

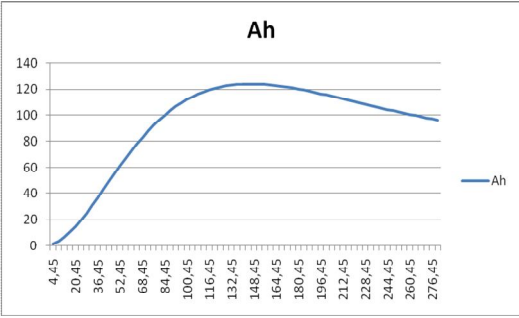
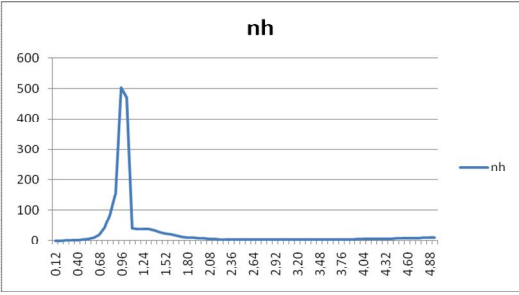
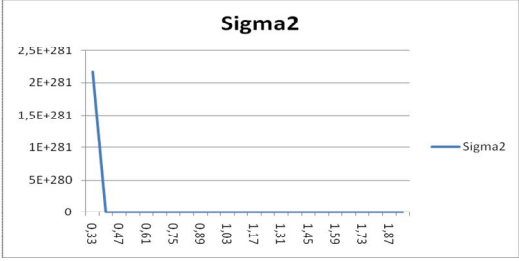
Distribuição Condicional Completa	Distribuição Proponente	Gráfico da Distribuição Empírica a menos de uma constante de normalização
[4.2.2]	Gama(30;2)	
[4.2.3]	Normal(1;1)	
[4.2.4]	Gama(1;2)	

Figura 4: Distribuições proponentes do modelo da altura dominante

ii) Modelo de área basal

Seja o modelo

$$G_i = \left[ Ag_i \exp\left\{-kg_i \left(\frac{1}{t_i}\right)^{ng_i}\right\}\right] I(i = i_1) + \left[ Ag_i \left(\frac{G_{i-1}}{Ag_i}\right)^{\left(\frac{(t-1)^{ng_{i-1}}}{t^{ng_i}}\right)}\right] (1 - I(i = i_1)) + \varepsilon_i,$$

como definido anteriormente, onde  $I(\cdot)$  representa a função indicatriz,  $i_1$  o primeiro registo de idade de uma parcela e sejam

$$Ag_i = Ag_1 l q e_i^2$$

$$ng_i = ng_0 + ngq \ln(Iqe_i) + ngn \frac{N_i}{1000}$$

$$ng_{i-1} = ng_0 + ngq \ln(Iqe_i) + ngn \frac{N_{i-1}}{1000}$$

$$kg_i = kgq \frac{1}{Iqe_i}$$

$$\varepsilon_i \sim N(0, \sigma^2) \text{ (iid) } (i = 1, 2, \dots, n).$$

Denotando por  $\beta' = [Agq \ ng0 \ ngq \ ngn \ kgq] \in \mathfrak{R}^5$ , tem de se estimar o vector dos parâmetros  $(\beta, \sigma^2)$ , com  $\sigma^2 \in \mathfrak{R}^+$ . Seja  $X = [t_i \ G_{i-1} \ Iqe_i \ N_i \ N_{i-1} \ 1]$ , a matriz dos valores das variáveis independentes (pré-determinadas) e considere-se o modelo definido equivalentemente por:

$$f_{1i}(x_i, \beta) = Ag_i \exp\left\{-kg_i \left(\frac{1}{t_i}\right)^{ng_i}\right\}$$

$$f_{2i}(x_i, \beta) = Ag_i \left(\frac{G_{i-1}}{Ag_i}\right)^{\left(\frac{(t-1)^{ng_{i-1}}}{t^{ng_i}}\right)}.$$

Como

$$G_i | x_{i-1}, \beta, \sigma^2 \sim N(f_{1i}(x_i, \beta)I(i = i_1) + f_{2i}(x_i, \beta)(1 - I(i = i_1)), \sigma^2) \quad (i = 1, 2, \dots, n),$$

com  $x_{i-1}$  conhecido e assumindo:

$m = n^\circ$  de parcelas existentes na amostra;

$n = n^\circ$  total de registos existentes na amostra;

$n_j = n^\circ$  total de registos de cada parcela, tais que  $n = \sum_{j=1}^m n_j$ ;

$t_{n_j}$  = idade do último registo da parcela  $j$ ,

tem-se a seguinte função de verosimilhança da amostra:

$$f_{G_1, G_2, \dots, G_n | \beta, \sigma^2, x}(G_1, G_2, \dots, G_n) =$$

$$\begin{aligned}
 &= \prod_{j=1}^m \left\{ \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{(G_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2}{2\sigma^2} \right\} \prod_{i>i_1}^{t_{n_j}} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp \left\{ -\frac{(G_{ij} - f_{2ij}(x_{ij}, \beta))^2}{2\sigma^2} \right\} \right\} \\
 &\propto \frac{1}{(\sigma^2)^{n/2}} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (G_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}
 \end{aligned}$$

Tomando para distribuição *a priori*, a distribuição não informativa proposta por Box e Tiao:

$$p(\beta, \sigma^2) \propto \frac{1}{\sigma^2}$$

tem-se a distribuição *a posteriori*:

$$[4.2.5] \quad p(\beta, \sigma^2 | G) \propto$$

$$(\sigma^2)^{-\frac{n}{2}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (G_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

Note-se que esta distribuição não é conhecida e não apresenta uma forma fechada pelo que não se conseguem obter os valores médios *a posteriori*, consequentemente irá aplicar-se a metodologia MCMC, através do cálculo das distribuições condicionais completas.

$$[4.2.6] \quad p(Agq | ng_0, ngq, ngn, kgq, \sigma^2, G) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (G_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

$$[4.2.7] \quad p(ng_0 | Agq, ngq, ngn, kgq, \sigma^2, G) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (G_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

$$[4.2.8] \quad p(ngq | Agq, ng_0, ngn, kgq, \sigma^2, G) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (G_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

$$[4.2.9] \quad p(\text{ngn} \mid \text{Agq}, \text{ng}_0, \text{ngq}, \text{kgq}, \sigma^2, G) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i,j} - f_{1i,j}(x_{i,j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (G_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

$$[4.2.10] \quad p(\text{kgq} \mid \text{Agq}, \text{ng}_0, \text{ngq}, \text{ngn}, \sigma^2, G) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i,j} - f_{1i,j}(x_{i,j}, \beta))^2 \right] \right\}$$

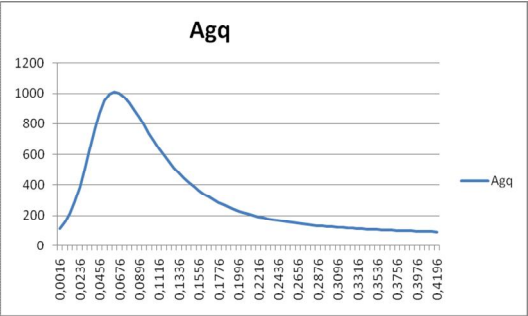
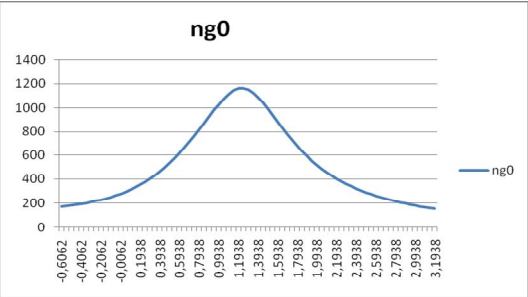
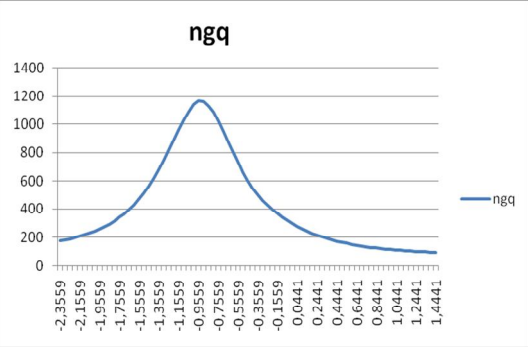
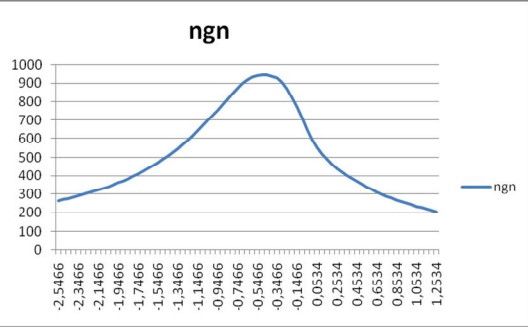
$$[4.2.11] \quad p(\sigma^2 \mid \text{Agq}, \text{ng}_0, \text{ngq}, \text{ngn}, \text{kgq}, G) \propto$$

$$(\sigma^2)^{\frac{n}{2}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (G_{i,j} - f_{1i,j}(x_{i,j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (G_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

Atendendo às razões expostas anteriormente para se obter uma amostra aleatória dos valores dos parâmetros modelados pela distribuição *a posteriori*, foi usado o algoritmo de Gibbs com o passo de Metropolis.

Como distribuições proponentes (ou envelopes) para estas distribuições condicionais completas ([4.2.6] a [4.2.11]), de acordo com a sua representação gráfica a menos da constante de normalização (para este modelo), foram utilizadas as distribuições que se apresentam na Figura 5.

Note-se que neste modelo optou-se por pré-definir o modelo de acordo com o modelo final segundo a abordagem clássica (utilizaram-se os mesmos parâmetros) de forma a torná-lo mais parcimonioso em termos de parâmetros e variáveis.

Distribuição Condicional Completa	Distribuição Proponente	Gráfico da Distribuição Empírica a menos de uma constante de normalização
[4.2.6]	Gama(0,25;2)	 <p>Gráfico da distribuição empírica <b>Agq</b>. O eixo horizontal representa o valor da variável, variando de 0,0016 a 0,4196. O eixo vertical representa a densidade, variando de 0 a 1200. A curva mostra uma distribuição assimétrica à direita, com um pico próximo a 0,0456 e uma cauda longa que se estende até 0,4196.</p>
[4.2.7]	Normal(1,24;1,2)	 <p>Gráfico da distribuição empírica <b>ng0</b>. O eixo horizontal representa o valor da variável, variando de -0,6062 a 3,1938. O eixo vertical representa a densidade, variando de 0 a 1400. A curva é simétrica e em forma de sino, com um pico centralizado em 1,24.</p>
[4.2.8]	Normal(-1;1,2)	 <p>Gráfico da distribuição empírica <b>ngq</b>. O eixo horizontal representa o valor da variável, variando de -2,3559 a 1,4441. O eixo vertical representa a densidade, variando de 0 a 1400. A curva é simétrica e em forma de sino, com um pico centralizado em -1.</p>
[4.2.9]	Normal(-0,55;2)	 <p>Gráfico da distribuição empírica <b>ngn</b>. O eixo horizontal representa o valor da variável, variando de -2,5466 a 1,2534. O eixo vertical representa a densidade, variando de 0 a 1000. A curva é simétrica e em forma de sino, com um pico centralizado em -0,55.</p>

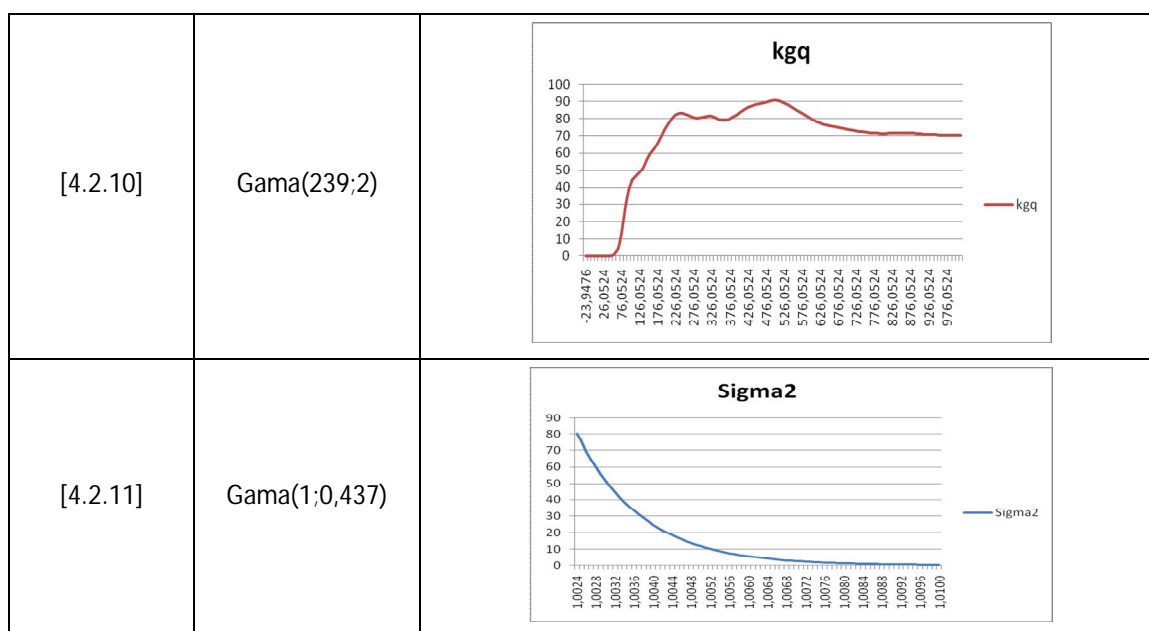


Figura 5: Distribuições proponentes do modelo da área basal

### iii) Modelo de mortalidade

Seja

$$N_i = \left[ Npl_i \exp\left\{-\left(am0 + amnp \frac{Npl_i}{1000} + amq \frac{Iqe_i}{10}\right)t_i\right\} \right] I(i = i_1) +$$

$$\left[ N_{i-1} \exp\left\{-\left(am0 + amnp \frac{Npl_i}{1000} + amq \frac{Iqe_i}{10}\right)\right\} \right] (1 - I(i = i_1)) + \varepsilon_i,$$

como anteriormente definida, onde  $I(\cdot)$  representa a função indicatriz,  $i_1$  o primeiro registo de idade de uma parcela e  $\varepsilon_i \sim N(0, \sigma^2)$ , iid, com  $i = 1, 2, \dots, n$ .

Por uma questão de simplificação na notação, seja

$$\beta' = [am0 \quad amnp \quad amq] \in \mathfrak{R}^3, \quad \sigma^2 \in \mathfrak{R}^+,$$

a matriz dos valores das variáveis independentes (pré-determinadas)

$$X = [t_i \quad Npl_i \quad N_{i-1} \quad 1 \quad Iqe_i]$$

De uma forma simplificada e equivalente considere-se o modelo

$$f_{1i}(x_i, amnp) = Npl_i \exp\left\{-\left(am0 + amnp \frac{Npl_i}{1000} + amq \frac{Iqe_i}{10}\right)t_i\right\}$$

$$f_{2i}(x_i, amnp) = N_{i-1} \exp\left\{-\left(am0 + amnp \frac{Npl_i}{1000} + amq \frac{Iqe_i}{10}\right)\right\}.$$

Como

$$N_i | x_{i-1}, \beta, \sigma^2 \sim N(f_{1i}(x_i, \beta)I(i = i_1) + f_{2i}(x_i, \beta)(1 - I(i = i_1)), \sigma^2) \quad (i = 1, 2, \dots, n),$$

com  $x_{i-1}$  conhecido e assumindo:

$m = n^\circ$  de parcelas existentes na amostra;

$n = n^\circ$  total de registos existentes na amostra;

$n_j = n^\circ$  total de registos de cada parcela, tais que  $n = \sum_{j=1}^m n_j$  ;

$t_{n_j} =$  idade do último registo da parcela  $j$ ,

tem-se a seguinte função de verosimilhança da amostra:

$$\begin{aligned} f_{N_1, N_2, \dots, N_n | \beta, \sigma^2, x}(N_1, N_2, \dots, N_n) &= \\ &= \prod_{j=1}^m \left\{ \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(N_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2}{2\sigma^2}\right\} \right\}_{i>i_1}^{t_{n_j}} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left\{-\frac{(N_{ij} - f_{2ij}(x_{ij}, \beta))^2}{2\sigma^2}\right\} \Bigg\} \\ &\propto \frac{1}{(\sigma^2)^{n/2}} \exp\left\{-\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (N_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (N_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\} \end{aligned}$$

Tomando para a distribuição *a priori*, a distribuição não informativa proposta por Box e Tiao

$$p(\beta, \sigma^2) \propto \frac{1}{\sigma^2}$$

tem-se a distribuição *a posteriori*:

$$[4.2.12] \quad p(\beta, \sigma^2 | N) \propto$$

$$\left(\sigma^2\right)^{\frac{n}{2}-1} \exp\left\{-\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (N_{i_j} - f_{1i_j}(x_{i_j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (N_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$



Consequentemente as condicionais completas que serão utilizadas para simular pelo algoritmo de Gibbs com o passo de Metropolis valores aleatórios para os parâmetros da distribuição *a posteriori*, são dadas por:

$$[4.2.13] \quad p(am0 \mid amnp, amq, \sigma^2, N) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (N_{i,j} - f_{1i,j}(x_{i,j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (N_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

$$[4.2.14] \quad p(amnp \mid am0, amq, \sigma^2, N) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (N_{i,j} - f_{1i,j}(x_{i,j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (N_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

$$[4.2.15] \quad p(amq \mid am0, amnp, \sigma^2, N) \propto$$

$$\exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (N_{i,j} - f_{1i,j}(x_{i,j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (N_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

$$[4.2.16] \quad p(\sigma^2 \mid am0, amnp, amq, N) \propto$$

$$(\sigma^2)^{\frac{n}{2}-1} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \sum_{j=1}^m (N_{i,j} - f_{1i,j}(x_{i,j}, \beta))^2 + \sum_{j=1}^m \sum_{i>i_1}^{t_{n_j}} (N_{ij} - f_{2ij}(x_{ij}, \beta))^2 \right] \right\}$$

Como distribuições proponentes (ou envelopes) para estas distribuições condicionais completas ([4.2.13] a [4.2.16]), de acordo com a sua representação gráfica a menos da constante de normalização (para este modelo inicial), foram utilizadas as distribuições que se apresentam na Figura 6.

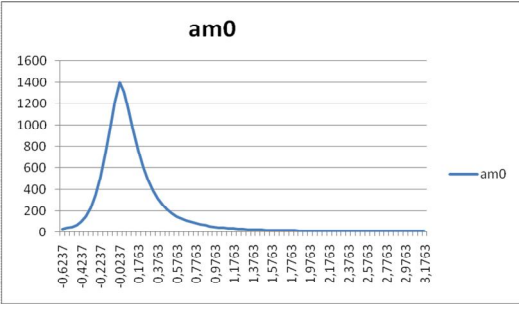
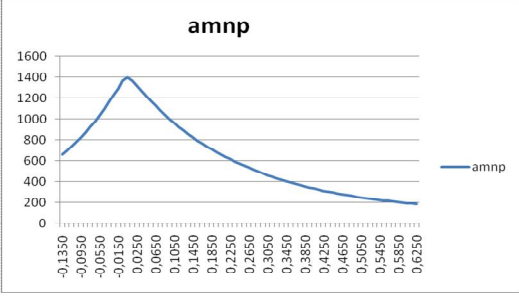
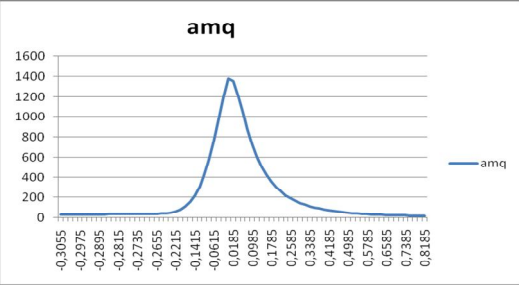
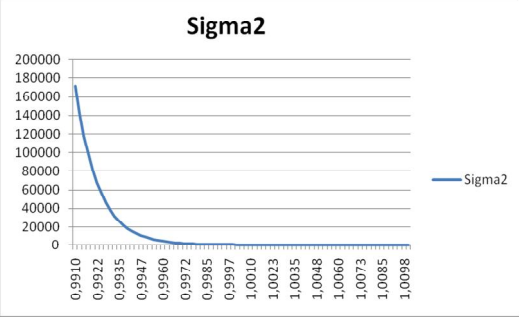
Distribuição Condicional Completa	Distribuição Proponente	Gráfico da Distribuição Empírica a menos de uma constante de normalização
[4.2.13]	Normal (-0,02372;4)	
[4.2.14]	Normal (0,005;4)	
[4.2.15]	Normal (-0,001;4)	
[4.2.16]	Gama(1;0,432)	

Figura 6: Distribuições proponentes do modelo Inicial da mortalidade

## 4.2.2 Estimativas Bayesianas para o Modelo

Nesta secção descreve-se toda metodologia seguida com a finalidade de escolher os melhores modelos bayesianos para a altura dominante, área basal e mortalidade, tendo em conta a informação amostral, a escolha das distribuições *a priori* e a escolha das distribuições proponentes (já descrito em parte nos pontos anteriores).

Comecei por estimar os modelos iniciais (descritos no ponto 4.2.1.), tendo implementado o método de Gibbs com passo de Metropolis para cada um dos três modelos, gerando três cadeias de Markov com 30000 réplicas para cada parâmetro de cada modelo respetivo.

Neste processo recorreu-se a Macros em “Excel-VBA” (ver Anexo III) programadas adequadamente com o algoritmo Gibbs com passo de Metropolis e de acordo com as distribuições apresentadas na secção 4.2.1.

Das 30000 réplicas para cada parâmetro, selecionaram-se apenas as últimas 4000, começando na réplica nº 10005 (verifica-se uma estabilização das médias ergódicas a partir deste estado) e com a finalidade de eliminar a ocorrência de correlação (dependência) entre réplicas, considere apenas uma observação de 5 em 5. No Anexo II encontram-se os gráficos das evoluções das médias ergódicas das 30000 réplicas geradas para cada parâmetro de cada modelo para estudar a convergência da cadeia, as estatísticas teste de Geweke assim como os gráficos das respectivas distribuições para os modelos finais para as 4000 réplicas selecionadas.

Nas Tabelas 13, 14 e 15 apresentam-se as estimativas dos parâmetros (médias amostrais), os primeiros, segundos e terceiros quartis e os desvios-padrão amostrais para os modelos considerados e para cada uma das variáveis biométricas em estudo. Todas estas estatísticas permitem-nos fazer a seleção natural de variáveis via critério DIC, com o objetivo de escolher o melhor modelo, segundo a perspectiva bayesiana. Esta metodologia funciona relativamente bem quando o número de covariáveis (neste caso parâmetros estimados de cada modelo) é moderado. O objetivo consiste em selecionar os modelos que minimizem o DIC.

## i) Modelo de altura dominante

Tabela 13: Estimativas e algumas estatísticas descritivas de localização e escala.

Parâmetros	Média Amostral	$Q_{0,025}$	$Q_{0,5}$	$Q_{0,975}$	Desvio-Padrão Amostral
ah	59,7085	38,0851	59,2864	85,7652	12,6540
nh	0,8891	-1,0392	0,7065	3,0330	1,0247
$\sigma^2$	2,9848	0,0842	2,1790	9,9516	2,7026

## ii) Modelo de área basal

Tabela 14: Estimativas e algumas estatísticas descritivas de localização e escala.

Parâmetros	Média Amostral	$Q_{0,025}$	$Q_{0,5}$	$Q_{0,975}$	Desvio-Padrão Amostral
agq	1,7766	0,0078	1,1473	6,9863	1,9376
kgq	476,6805	416,7483	476,3520	540,2616	31,6947
ng0	1,2174	-1,4045	1,2363	3,8795	1,3639
ngn	-0,6444	-5,0411	-0,5401	3,8041	2,3110
ngq	-1,0523	-3,7235	-1,0148	1,6257	1,3864
$\sigma^2$	0,6500	0,0219	0,4889	2,1046	0,5685

## iii) Modelo de mortalidade (M0)

Tabela 15: Estimativas e algumas estatísticas descritivas de localização e escala.

Parâmetros	Média	$Q_{0,025}$	$Q_{0,5}$	$Q_{0,975}$	Desvio-Padrão
	Amostral				Amostral
am0	-0,0259	-9,2302	0,0051	9,1073	4,7891
amnp	-0,0476	-8,7133	0,0674	8,9838	4,6484
amq	-0,1240	-9,2200	0,0253	9,0780	4,7491
$\sigma^2$	0,6414	0,0194	0,4810	2,0917	0,5646

Analisando os valores das estatísticas de localização e escala destas tabelas verifica-se que apenas no modelo da mortalidade existem duas covariáveis suscetíveis de poderem ser consideradas nulas e portanto retiradas do modelo, são elas: am0 e amnp. Os outros modelos não têm nenhuma covariável que possa ser suscetível de assumir o valor zero.

Apresentam-se nas Tabelas 16 e 17 as estatísticas descritivas (tal como se fez anteriormente) dos dois modelos concorrentes com o modelo inicial para a mortalidade, um sem am0 e outro sem amnp; pois esta análise é mais coerente quando se retira apenas uma covariável e se analisa o resultado (variação positiva ou negativa do DIC para o modelo de partida). Note-se que para estes modelos concorrentes da mortalidade houve necessidade de se ajustar os valores dos parâmetros para as distribuições proponentes (envelopes) para cada covariável a ter em conta.

**Modelo de mortalidade sem am0 (M1)****Tabela 16: Estimativas e algumas estatísticas descritivas de localização e escala.**

Parâmetros	Média	$Q_{0,025}$	$Q_{0,5}$	$Q_{0,975}$	Desvio-Padrão
	Amostral				Amostral
amnp	-0,1199	-8,9675	0,0783	8,8825	4,6757
amq	-0,1052	-8,8874	0,0270	8,8260	4,6906
$\sigma^2$	0,6347	0,0189	0,4571	2,1446	0,5884

**Modelo de mortalidade sem amnp (M2)****Tabela 17: Estimativas e algumas estatísticas descritivas de localização e escala.**

Parâmetros	Média	$Q_{0,025}$	$Q_{0,5}$	$Q_{0,975}$	Desvio-Padrão
	Amostral				Amostral
am0	-0,3073	-9,0106	-0,0802	8,7366	4,6200
amq	-0,1452	-8,7295	-0,0445	8,7206	4,6761
$\sigma^2$	0,6522	0,0195	0,4799	2,1727	0,5934

Após os cálculos dos DIC (programados em VBA com as diferentes 4000 réplicas para cada covariável tida em conta no modelo – ver anexo III), verificou-se que:  $DIC(M0) = 9,9 \times 10^{303}$ ;  $DIC(M1) = 2,7 \times 10^{304}$  e  $DIC(M2) = 1,1 \times 10^{249}$ . Logo conclui-se que o modelo final para a mortalidade é o M2, pois neste parece não existir nenhuma covariável suscetível de poder assumir o valor zero (analisando as suas estatísticas descritivas).

## CAPITULO 5

### 5 Análise de Resultados Clássicos e Bayesianos

A modelação Bayesiana para os casos das variáveis em estudo (altura dominante, área basal e mortalidade) para o modelo *GLOBULUS 2.1*, como deparamos, apresenta metodologias bastante distintas das Clássicas de abordar este problema de estimação.

Os resultados para alguns modelos em que coincidem as variáveis e parâmetros estimados, são comparáveis, uma vez que na abordagem Bayesiana utilizaram-se distribuições *a priori* não informativas propostas por Box-Tiao.

Apresentam-se novamente os resultados finais dos parâmetros por modelo, com a finalidade de comparar os seus valores obtidos na estimação para cada uma das abordagens.

#### i) Altura dominante

**Tabela 18: Estimativas segundo as duas abordagens para a altura dominante.**

Parâmetros	Estimativas Clássicas	Estimativas Bayesianas
ah	40,4477	59,7085
nh	0,7465	0,8891

Para a altura dominante verificou-se o mesmo sinal dos parâmetros, sendo os obtidos pela abordagem bayesiana mais elevados.

Apresenta-se a **Equação 1** segundo as duas abordagens:

- **Clássica:** 
$$hdom_t = 40,4477 \left( \frac{hdom_{t-1}}{40,4477} \right)^{\left( \frac{t-1}{t} \right)^{0,7465}}, \text{ para } t > 0 \text{ e inteiro;}$$

- **Bayesiana:** 
$$hdom_t = 59,7085 \left( \frac{hdom_{t-1}}{59,7085} \right)^{\left( \frac{t-1}{t} \right)^{0,8891}}, \text{ para } t > 0 \text{ e inteiro.}$$

## ii) Área basal

**Tabela 19: Estimativas segundo as duas abordagens para a área basal.**

Parâmetros	Estimativas	Estimativas
	Clássicas	Bayesianas
agq	0,1227	1,7766
kgq	136,2317	476,6805
ng0	1,9387	1,2174
ngn	0,0535	-0,6444
ngq	-0,4042	-1,0523

No caso da área basal, em primeiro lugar há que ter em atenção que para o caso da abordagem Clássica se recorreu a um sistema de equações simultâneas e este modelo foi estimado conjuntamente com o modelo da mortalidade, no caso Bayesiano estimamos isoladamente cada um dos modelos.

Segundo a Tabela 19, verificamos de um modo geral que para os parâmetros cujos valores estimados são maiores, atingem valores mais elevados na abordagem Bayesiana, enquanto que para os parâmetros cujos valores estimados são mais baixos, são comparativamente mais baixos para a abordagem Bayesiana. O único parâmetro que desobedece a esta constatação é agq.



Resultam os seguintes modelos de inicialização (Equação 3.1) e projeção (Equação 3.2) segundo as abordagens:

- **Clássica:**

**Equação 3.1**  $G_t = A_g \exp\left(-k_g \left(\frac{1}{t}\right)^{n_g}\right)$ , para  $t > 0$  e inteiro,

com

$$A_g = 0,1227Iqe^2$$

$$k_g = 136,2317 \frac{1}{Iqe}$$

$$n_g = 1,9387 - 0,4042 \ln(Iqe) + 0,0535 \frac{N_t}{1000}$$

**Equação 3.2**  $G_t = A_g \left(\frac{G_{t-1}}{A_g}\right)^{\frac{(t-1)^{n_{gt-1}}}{t^{n_{gt}}}}$ , para  $t > 1$  e inteiro,

com

$$A_g = 0,1227 Iqe^2$$

$$n_{gt} = 1,9387 - 0,4042 \ln(Iqe) + 0,0535 \frac{N_t}{1000}, \text{ para } t > 0 \text{ e inteiro;}$$

- **Bayesiana:**

**Equação 3.1**  $G_t = A_g \exp\left(-k_g \left(\frac{1}{t}\right)^{n_g}\right)$ , para  $t > 0$  e inteiro,

com

$$A_g = 1,7766Iqe^2$$

$$k_g = 476,6805 \frac{1}{Iqe}$$

$$n_g = 1,2174 - 1,0523 \ln(Iqe) - 0,6444 \frac{N_t}{1000}$$

**Equação 3.2**  $G_t = A_g \left( \frac{G_{t-1}}{A_g} \right)^{\frac{(t-1)^{n_{gt}-1}}{t^{n_{gt}}}}$ , para  $t > 1$  e inteiro,

com

$$A_g = 1,7766 Iqe^2$$

$$n_{gt} = 1,2174 - 1,0523 \ln(Iqe) - 0,6444 \frac{N_t}{1000}, \text{ para } t > 0 \text{ e inteiro.}$$

### iii) Mortalidade

**Tabela 20: Estimativas segundo as duas abordagens para a mortalidade.**

Parâmetros	Estimativas	Estimativas
	Clássicas	Bayesianas
am0	-	-0,3073
amq	-	-0,1452
amnp	0,0040	-

No caso do modelo da mortalidade, como se pode verificar na tabela 20, não há possibilidade de comparação das estimativas das duas abordagens, uma vez que nos modelos finais resultantes das respetivas metodologias de ambas as abordagens não resultaram parâmetros comuns.

Segundo as duas abordagens apresentam-se as equações finais, para inicialização (Equação 2.1) e projeção (Equação 2.2):

- **Clássica:**

**Equação 2.1**  $N_t = Npl \exp\left(-\left(0,0040 \frac{Npl}{1000}\right)t\right)$ , para  $t > 0$  e inteiro

**Equação 2.2**  $N_t = N_{t-1} \exp\left(-\left(0,0040 \frac{Npl}{1000}\right)\right)$ , para  $t > 1$  e inteiro;

- **Bayesiana:**

**Equação 2.1**  $N_t = Npl \exp\left(-\left(-0,3073 - 0,1452 \frac{Iqe}{10}\right)t\right)$ , para  $t > 0$  e inteiro

**Equação 2.2**  $N_t = N_{t-1} \exp\left(-\left(-0,3073 - 0,1452 \frac{Iqe}{10}\right)\right)$ , para  $t > 1$  e inteiro.

Todos estes comentários acerca da comparação entre as estimativas das duas abordagens, podem ser confirmados pela observação da Figura 7.

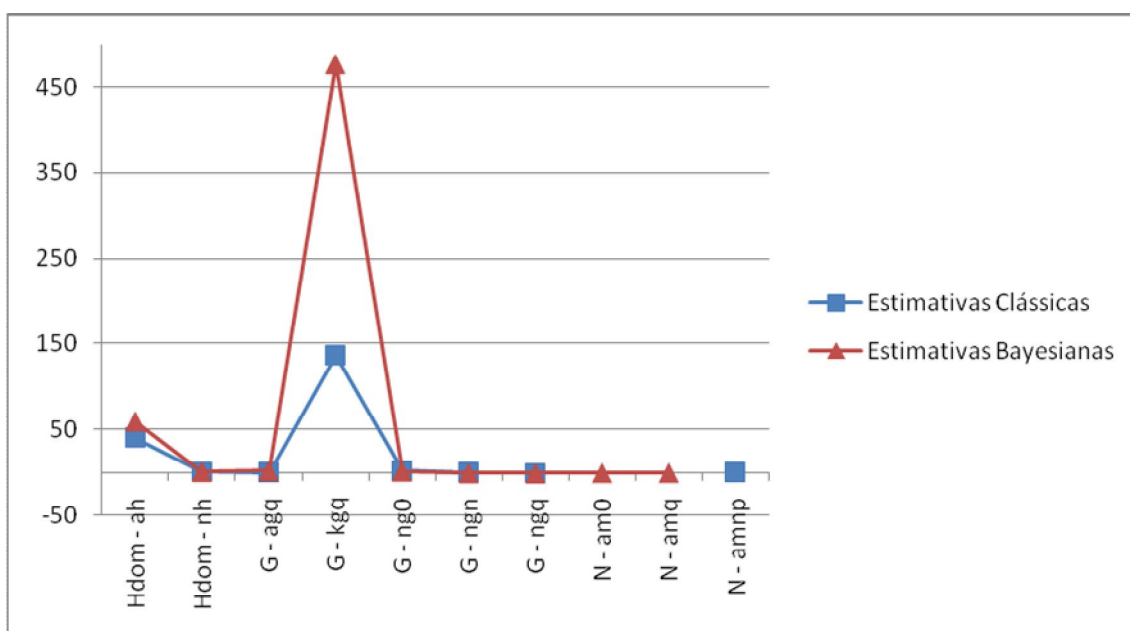


Figura 7: Gráfico de Perfil das estimativas Clássicas e Bayesianas



## CAPITULO 6

### 6 Conclusões

O trabalho foi levado a bom porto muito pela persistência do autor e da Orientadora, pois no que concerne à abordagem Bayesiana existe pouquíssima literatura e trabalhos de aplicação em dados florestais. Para isso foram usados motores de busca da internet e constatei a quase inexistência nas bases de dados da Biblioteca do Instituto Superior de Agronomia de resultados relacionados com o tema em concreto.

Foi um trabalho muito exigente em termos computacionais, pois a programação de todas as macros para a geração de amostragens e a obtenção dessas amostras para a estimação dos parâmetros Bayesianos não foram tarefas fáceis, mas sim árduas e muito morosas e com a agravante de praticamente todo o trabalho ter sido feito enquanto colaborava como analista/programador em empresas prestadoras de serviços informáticos (consultadoria).

Conclui-se para os casos em estudo, reforçando as análises do capítulo anterior, que quando comparáveis os resultados da Metodologia Bayesiana são próximos dos obtidos pela Metodologia Clássica. Uma questão que se coloca é a de saber qual destas metodologias é a melhor em termos preditivos, mas isso sugeria para um outro trabalho de mestrado ou quiçá de doutoramento.

Com este trabalho contribuí para alargar os horizontes na forma de estimar parâmetros de equações de variáveis florestais com a introdução da metodologia Bayesiana e abri o caminho para quem queira tirar partido desta para novos desenvolvimentos e aplicações em Portugal. Espero ter construído uma alavanca para

a implementação da metodologia Bayesiana para a estimação de parâmetros de modelos biométricos de variáveis florestais.

Como novos desafios de investigação na abordagem Bayesiana deste trabalho propunha melhorar os seguintes pontos:

- fazer toda a programação e estimação necessária recorrendo às grandes potencialidades do software "R", nomeadamente fazer uso de outros algoritmos nele implementados;
- estimar os modelos da área basal e mortalidade conjuntamente num sistema de equações simultâneas;
- aplicar esta abordagem bayesiana aos modelos mais recentes de variáveis biométricas florestais.

## ANEXOS





## ANEXO 1

## Evolução em SAS do Modelo de Equações Simultâneas da mortalidade e área Basal (metodologia clássica)

## i) Resultados Estimação – Modelo Inicial

```

The MODEL Procedure

Model Summary

Model Variables      2
Parameters           11
Equations             2
Number of Statements 10

Model Variables      Npov G
Parameters(Value)  am0(0) amnp(0.1) anq(0) agq(0.5) ng0(0.1) ngq(0.1)
                   ngn(0.1) kg0(0) kgq(0.1) kgnp(0) kgf(0)
Equations           Npov G

The 2 Equations to Estimate

Npov = F(am0, amnp, anq)
G     = F(agq, ng0, ngq, ngn, kg0, kgq, kgnp, kgf)

NOTE: At SUR Iteration 4 CONVERGE=0.001 Criteria Met.
WARNING: At SUR Iteration 4 a total of 82 execution errors occurred for 41 observations.

```

The MODEL Procedure  
SUR Estimation Summary

Data Set Options

DATA=        MODELOS.DADOS1  
OUTEST=      MODELOS.DADOS1\_  
OUTS=        MODELOS.S

Minimization Summary

Parameters Estimated        11  
Method                      Marquardt  
Iterations                    4

Final Convergence Criteria

R                            0.000559  
PPC (kgF)                   0.048088  
RPC (kgF)                   0.09409  
Object                      1.227E-6  
Trace(S)                    41366.29  
Objective Value             1.977392  
Lambda                      1E-7

Observations Processed

Read                        1477  
Solved                      1477  
Used                        1436  
Missing                     41

The MODEL Procedure

Nonlinear SUR Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	Label
Npov	3	1433	59275907	41364.9	203.4	0.9358	0.9357	Npov
G	8	1428	1977.2	1.3846	1.1767	0.9862	0.9861	G

Nonlinear SUR Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr >  t
am0	-0.01771	0.0160	-1.11	0.2692
amp	0.005814	0.00210	2.77	0.0057
amq	0.005658	0.00785	0.72	0.4711
aq	0.122322	0.00682	17.94	<.0001
ng0	1.641947	0.2477	6.63	<.0001
ngq	-0.30557	0.0780	-3.92	<.0001
ngn	0.044078	0.00596	7.39	<.0001
kg0	1.653961	1.0183	1.62	0.1046
kgq	100.8081	24.4692	4.12	<.0001
kgnp	-0.17022	0.1171	-1.45	0.1462
kgf	0.487847	0.8049	0.61	0.5446

Number of Observations

Statistics for System

Used	1436	Objective	1.9774
Missing	41	Objective*N	2840

ii) Estimação do Modelo sem o parâmetro *kgf*

## The MODEL Procedure

## Model Summary

Model Variables	2
Parameters	10
Equations	2
Number of Statements	10

```

Model Variables:  Npov G
Parameters(Value): an0(0) amnp(0.1) anq(0) agq(0.5) ng0(0.1)
                  ngq(0.1) ngn(0.1) kg0(0) kgq(0.1) kgnp(0)
Equations:        Npov G

```

## The 2 Equations to Estimate

$$\begin{aligned}
 Npov &= F(an0, amnp, anq) \\
 G &= F(agq, ng0, ngq, ngn, kg0, kgq, kgnp)
 \end{aligned}$$

NOTE: At SUR Iteration 3 CONVERGE=0.001 Criteria Met.

WARNING: At SUR Iteration 3 a total of 82 execution errors occurred for 41 observations.

The MODEL Procedure  
SUR Estimation Summary

Data Set Options

DATA=        MODELOS.DADOS1  
OUTEST=      MODELOS.DADOS1\_  
OUTS=        MODELOS.S

Minimization Summary

Parameters Estimated        10  
Method                      Marquardt  
Iterations                    3

Final Convergence Criteria

R                            0.00029  
PPC(kgrp)                  0.004298  
RPC(kg0)                    0.023504  
Object                      1.924E-6  
Trace(S)                    41364  
Objective Value             1.971251  
Lambda                      1E-7

Observations Processed

Read                        1477  
Solved                      1477  
Used                         1436  
Missing                     41

The MODEL Procedure

Nonlinear SUR Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	Label
npov	3	1433	59272629	41362.6	203.4	0.9358	0.9357	Npov
i	7	1429	1975.6	1.3825	1.1758	0.9862	0.9862	G

Nonlinear SUR Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr >  t
am0	-0.01832	0.0160	-1.14	0.2530
amnp	0.005996	0.00210	2.85	0.0044
amq	0.005713	0.00785	0.73	0.4667
agq	0.120303	0.00643	18.72	<.0001
ng0	1.686266	0.2449	6.88	<.0001
ngq	-0.31803	0.0775	-4.10	<.0001
ngn	0.047475	0.00603	7.88	<.0001
kg0	1.525311	1.0330	1.48	0.1400
kgq	107.2223	23.5580	4.55	<.0001
kgnp	-0.18457	0.1014	-1.82	0.0690

Number of Observations

Statistics for System

Used	1436	Objective	1.9713
Missing	41	Objective*N	2831

iii) Estimação do Modelo sem os parâmetros  $kgf$  e  $amq$ 

## The MODEL Procedure

## Model Summary

Model Variables	2
Parameters	9
Equations	2
Number of Statements	10

```

Model Variables:  Npov G
Parameters(Value)  am0(0)  amp(0.1)  agq(0.5)  ng0(0.1)  ngq(0.1)
                   ngn(0.1)  kg0(0)   kgq(0.1)  kgnp(0)
Equations:        Npov G

```

## The 2 Equations to Estimate

$$\begin{aligned}
 Npov &= F(am0, amp) \\
 G &= F(agq, ng0, ngq, ngn, kg0, kgq, kgnp)
 \end{aligned}$$

NOTE: At SUR Iteration 3 CONVERGE=0.001 Criteria Met.

WARNING: At SUR Iteration 3 a total of 82 execution errors occurred for 41 observations.

The MODEL Procedure  
SUR Estimation Summary

Data Set Options

DATA=        MODELOS.DADOS1  
OUTEST=      MODELOS.DADOS1\_  
OUTS=        MODELOS.S

Minimization Summary

Parameters Estimated	9
Method	Marquardt
Iterations	3

Final Convergence Criteria

R	0.000288
PPC(kgD)	0.004235
RPC(kgD)	0.023114
Object	1.902E-6
Trace(S)	41351.16
Objective Value	1.971925
Lambda	1E-7

Observations Processed

Read	1477
Solved	1477
Used	1436
Missing	41



The MODEL Procedure

Nonlinear SUR Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	Label
Npov	2	1434	59295576	41349.8	203.3	0.9357	0.9357	Npov
G	7	1429	1975.5	1.3825	1.1758	0.9862	0.9862	G

Nonlinear SUR Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr >  t
am0	-0.00715	0.00468	-1.53	0.1270
amp	0.006529	0.00196	3.32	0.0009
agq	0.120268	0.00642	18.73	<.0001
ng0	1.68191	0.2449	6.87	<.0001
ngq	-0.3166	0.0775	-4.08	<.0001
ngn	0.047572	0.00603	7.89	<.0001
kg0	1.524327	1.0334	1.48	0.1404
kgq	107.2399	23.5681	4.55	<.0001
kgnp	-0.18397	0.1015	-1.81	0.0701

Number of Observations

Statistics for System

Used	1436	Objective	1.9719
Missing	41	Objective*N	2832

## iv) Estimação do Modelo sem os parâmetros kgf, amq e kg0

## The MODEL Procedure

## Model Summary

Model Variables	2
Parameters	9
Equations	2
Number of Statements	10

Model Variables	Npov G
Parameters(Value)	am0(0) amnp(0.1) agq(0.5) ng0(0.1) ngq(0.1) ngn(0.1) kg0 kgq(0.1) kgnp(0)
Equations	Npov G

## The 2 Equations to Estimate

$$Npov = F(am0, amnp)$$

$$G = F(agq, ng0, ngq, ngn, kgq, kgnp)$$

NOTE: At SUR Iteration 2 CONVERGE=0.001 Criteria Met.

WARNING: At SUR Iteration 2 a total of 82 execution errors occurred for 41 observations.

The MODEL Procedure  
SUR Estimation Summary

Data Set Options

DATA=        MODELOS.DADOS1  
OUTEST=      MODELOS.DADOS1\_  
OUTS=        MODELOS.S

Minimization Summary

Parameters Estimated	8
Method	Marquardt
Iterations	2

Final Convergence Criteria

R	0.000609
PPC(kgrp)	0.010496
RPC(kgrp)	0.044925
Object	0.000027
Trace(S)	41350.46
Objective Value	1.97257
Lambda	1E-7

Observations Processed

Read	1477
Solved	1477
Used	1436
Missing	41

The MODEL Procedure

Nonlinear SUR Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	Label
Npov	2	1434	59294577	41349.1	203.3	0.9357	0.9357	Npov
G	6	1430	1976.3	1.3820	1.1756	0.9862	0.9862	G

Nonlinear SUR Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr >  t
am0	-0.00732	0.00468	-1.56	0.1179
amp	0.006613	0.00197	3.36	0.0008
agq	0.120173	0.00635	18.92	<.0001
ng0	1.958866	0.1455	13.46	<.0001
ngq	-0.40597	0.0453	-8.95	<.0001
ngn	0.050022	0.00599	8.35	<.0001
kgq	140.8363	4.5001	31.30	<.0001
kgnp	-0.14923	0.1017	-1.47	0.1426

Number of Observations

Statistics for System

Used	1436	Objective	1.9726
Missing	41	Objective*N	2833

## v) Estimação do Modelo sem os parâmetros kgf, amq, kg0 e kgnp

## The MODEL Procedure

## Model Summary

Model Variables	2
Parameters	9
Equations	2
Number of Statements	10

Model Variables	Npov G
Parameters(Value)	am0(0) amnp(0.1) agq(0.5) ng0(0.1) ngq(0.1) ngn(0.1) kg0 kgq(0.1) kgnp
Equations	Npov G

## The 2 Equations to Estimate

$$\begin{aligned} \text{Npov} &= F(\text{am0}, \text{amnp}) \\ \text{G} &= F(\text{agq}, \text{ng0}, \text{ngq}, \text{ngn}, \text{kgq}) \end{aligned}$$

NOTE: At SUR Iteration 2 CONVERGE=0.001 Criteria Met.

WARNING: At SUR Iteration 2 a total of 82 execution errors occurred for 41 observations.

The MODEL Procedure  
SUR Estimation Summary

Data Set Options

DATA=        MODELOS.DADOS1  
OUTEST=      MODELOS.DADOS1\_  
OUTS=        MODELOS.S

Minimization Summary

Parameters Estimated	7
Method	Marquardt
Iterations	2

Final Convergence Criteria

R	0.000199
PPC(ngn)	0.000811
RPC(ngn)	0.0079
Object	0.000017
Trace(S)	41349.91
Objective Value	1.97186
Lambda	1E-7

Observations Processed

Read	1477
Solved	1477
Used	1436
Missing	41

The MODEL Procedure

Nonlinear SUR Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	Label
Npov	2	1434	59293792	41348.5	203.3	0.9357	0.9357	Npov
G	5	1431	1977.1	1.3816	1.1754	0.9862	0.9862	G

Nonlinear SUR Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr >  t
am0	-0.00748	0.00468	-1.60	0.1101
amp	0.006689	0.00197	3.40	0.0007
agq	0.122142	0.00652	18.73	<.0001
ng0	1.943242	0.1434	13.55	<.0001
ngq	-0.40549	0.0448	-9.05	<.0001
ngn	0.054708	0.00508	10.77	<.0001
kgq	136.2608	2.9123	46.79	<.0001

Number of Observations

Statistics for System

Used	1436	Objective	1.9719
Missing	41	Objective*N	2832

vi) Estimação do Modelo sem os parâmetros  $kgf$ ,  $amq$ ,  $kg0$ ,  $kgnp$  e  $am0$  (modelo final)

## The MODEL Procedure

## Model Summary

Model Variables	2
Parameters	9
Equations	2
Number of Statements	10

```

Model Variables:  Npov G
Parameters(Value)  am0 amnp(0.1) agq(0.5) ng0(0.1) ngq(0.1) ngn(0.1) kg0 kgq(0.1) kgnp
Equations:        Npov G

```

## The 2 Equations to Estimate

```

Npov = F(amnp)
G = F(agq, ng0, ngq, ngn, kgq)

```

NOTE: At SUR Iteration 2 CONVERGE=0.001 Criteria Met.

WARNING: At SUR Iteration 2 a total of 82 execution errors occurred for 41 observations.



The MODEL Procedure  
SUR Estimation Summary

Data Set Options

DATA=        MODELOS.DADOS1  
OUTEST=      MODELOS.DADOS1\_  
OUTS=        MODELOS.S

Minimization Summary

Parameters Estimated	6
Method	Marquardt
Iterations	2

Final Convergence Criteria

R	0.00023
RPC(ngn)	0.00096
RPC(ngn)	0.000603
Object	0.00002
Trace(S)	41415.84
Objective Value	1.971991
Lambda	1E-7

Observations Processed

Read	1477
Solved	1477
Used	1436
Missing	41

The MODEL Procedure

Nonlinear SUR Summary of Residual Errors

Equation	DF Model	DF Error	SSE	MSE	Root MSE	R-Square	Adj R-Sq	Label
Npov	1	1435	59429749	41414.5	203.5	0.9356	0.9356	Npov
G	5	1431	1977.8	1.3821	1.1756	0.9862	0.9862	G

Nonlinear SUR Parameter Estimates

Parameter	Estimate	Approx Std Err	t Value	Approx Pr >  t
amp	0.003952	0.000949	4.16	<.0001
agq	0.12265	0.00658	18.64	<.0001
ngD	1.938731	0.1429	13.56	<.0001
ngq	-0.40419	0.0447	-9.05	<.0001
ngn	0.053529	0.00505	10.60	<.0001
kgq	136.2317	2.9023	46.94	<.0001

Number of Observations

Statistics for System

Used	1436	Objective	1.9720
Missing	41	Objective*N	2832

The MODEL Procedure

Covariances of Parameter Estimates

	amp	agq	ng0	ngq	ngn	kgq
amp	9.0144E-7	-.0000001	-.0000002	0.0000001	-.0000001	-0.000006
agq	-8.301E-8	0.0000433	-.0001472	-.0000007	-.0000112	-0.008767
ng0	-1.903E-7	-.0001472	0.0204276	-.0062797	-.0000652	0.109609
ngq	1.2132E-7	-.0000007	-.0062797	0.0019962	0.0000212	-0.017246
ngn	-6.845E-8	-.0000112	-.0000652	0.0000212	0.0000255	0.000062
kgq	-6.47E-6	-.0087670	0.1096089	-.0172460	0.0000619	8.423331

Correlations of Parameter Estimates

	amp	agq	ng0	ngq	ngn	kgq
amp	1.0000	-0.0133	-0.0014	0.0029	-0.0143	-0.0023
agq	-0.0133	1.0000	-0.1565	-0.0025	-0.3377	-0.4590
ng0	-0.0014	-0.1565	1.0000	-0.9834	-0.0903	0.2642
ngq	0.0029	-0.0025	-0.9834	1.0000	0.0938	-0.1330
ngn	-0.0143	-0.3377	-0.0903	0.0938	1.0000	0.0042
kgq	-0.0023	-0.4590	0.2642	-0.1330	0.0042	1.0000

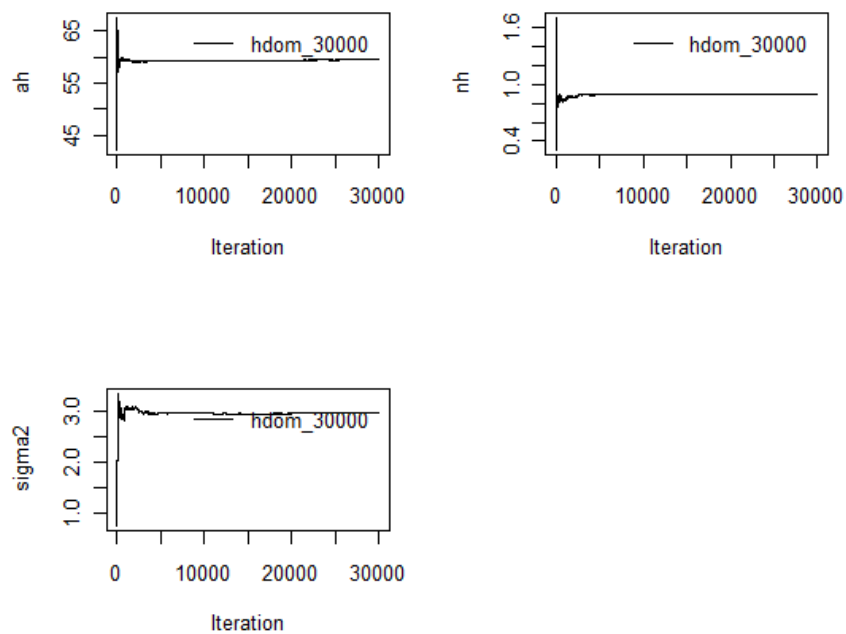


## ANEXO II

### Avaliação da Convergência e Gráficos das Distribuições para os Modelos Bayesianos

#### i) Modelo da altura dominante

#### Sampler Running Mean



GEWEKE CONVERGENCE DIAGNOSTIC:

=====

Fraction in first window = 0.1

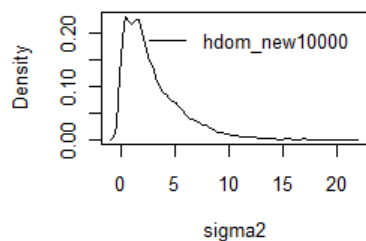
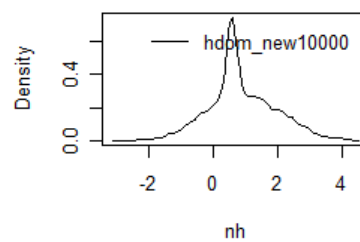
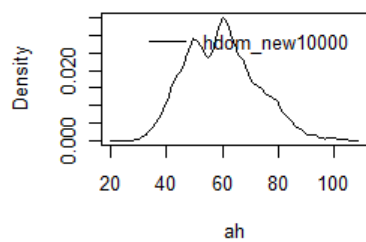
Fraction in last window = 0.5

Chain: Hdom\_new10000

-----

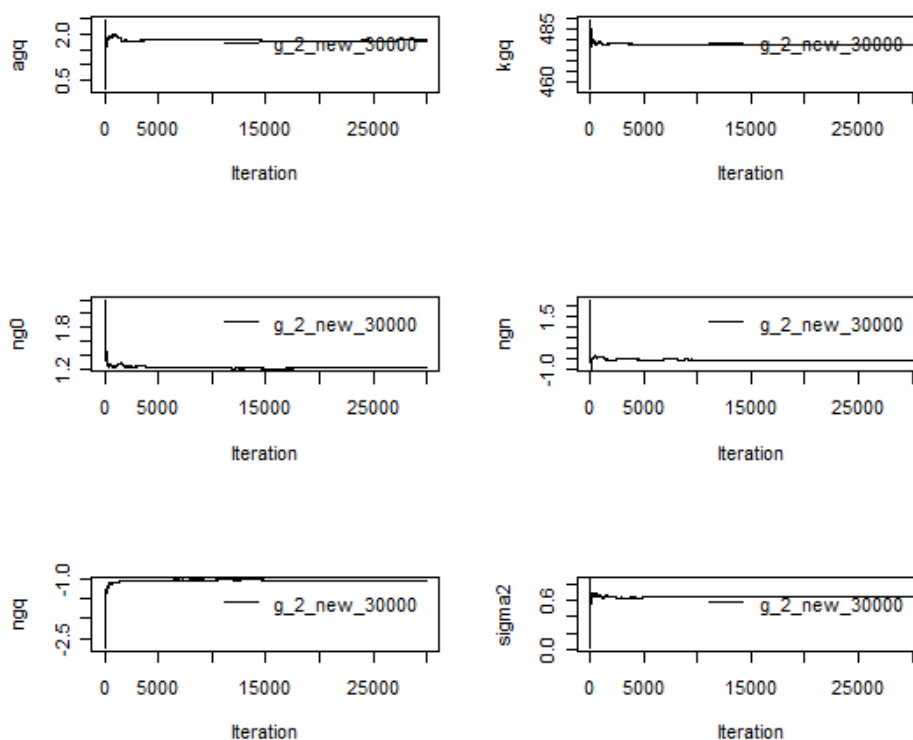
	ah	nh	sigma2
Z-Score	-0.2819171	-1.2766205	-0.5851038
p-value	0.7780071	0.2017363	0.5584779

### Estimated Posterior Density



ii) Modelo da Area Basal

## Sampler Running Mean



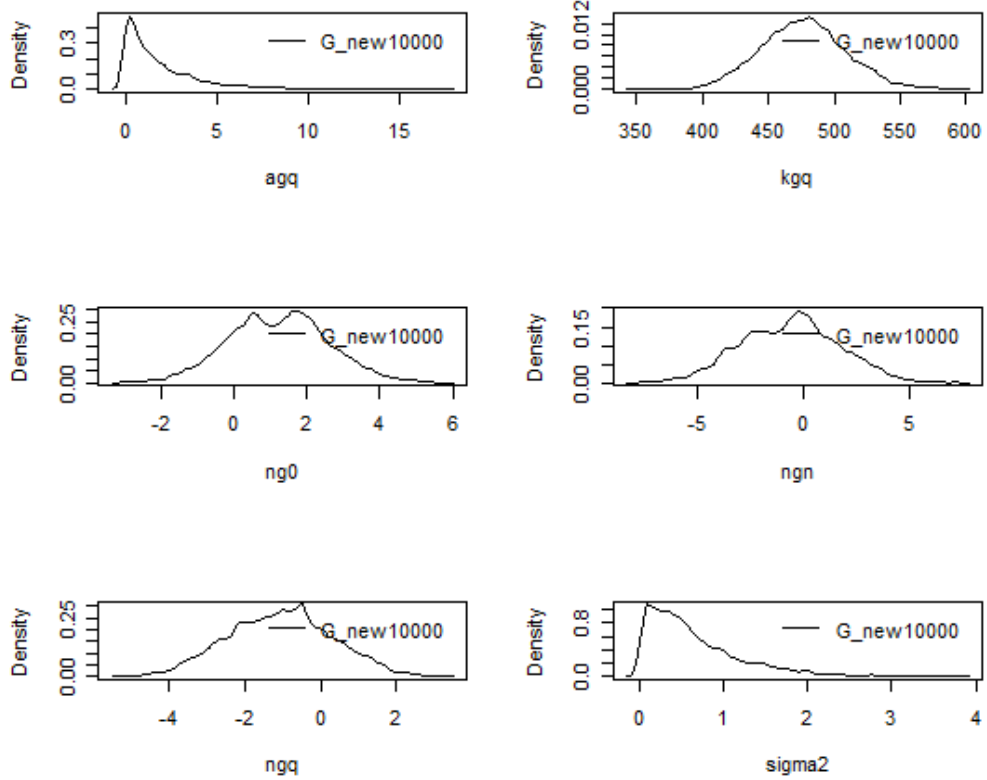
GEWEKE CONVERGENCE DIAGNOSTIC:

Fraction in first window = 0.1  
 Fraction in last window = 0.5

Chain: G\_new10000

	agq	kgq	ng0	ngn	ngq	sigma2
Z-Score	1.6367789	-0.8565673	-0.3238401	-0.3822976	0.8062041	-0.4130256
p-value	0.1016767	0.3916841	0.7460591	0.7022406	0.4201252	0.6795878

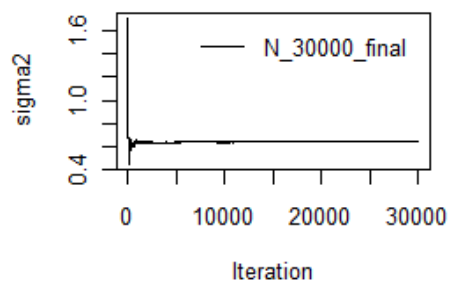
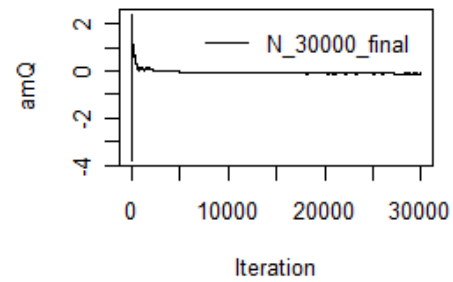
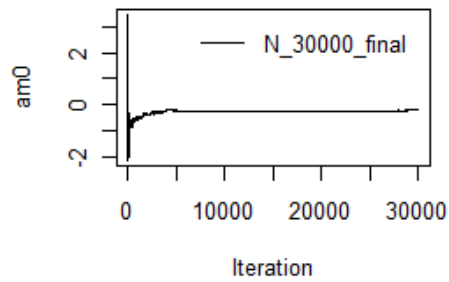
## Estimated Posterior Density





## iii) Modelo da mortalidade

## Sampler Running Mean



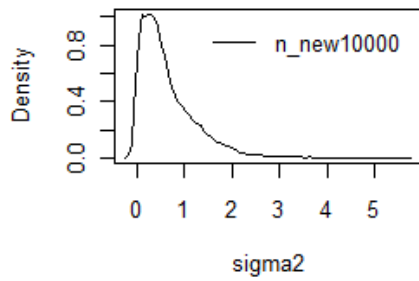
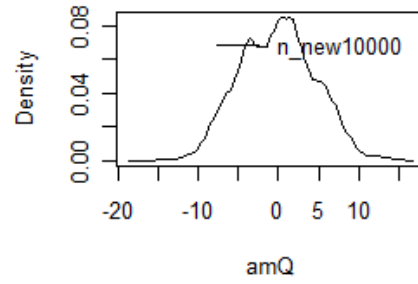
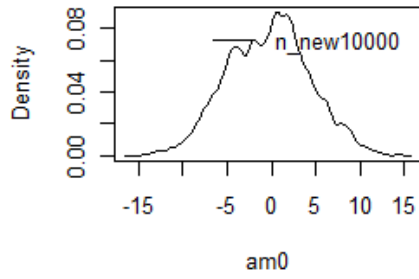
## GEWEKE CONVERGENCE DIAGNOSTIC:

=====  
 Fraction in first window = 0.1  
 Fraction in last window = 0.5

Chain:  $N_{new10000}$   
 -----

	$am_0$	$am_Q$	$\sigma^2$
Z-Score	-0.7349880	-0.1339209	0.7561331
p-value	0.4623468	0.8934651	0.4495694

## Estimated Posterior Density



## ANEXO III

### Exemplos de Macros usadas para VBA-EXCEL

#### i) Cálculo das Réplicas pelo Método de Gibbs com o passo de Metropolis

```
Sub Macro1()  
  
    If Sheets.Count < 2 Then  
        Dim wks As Worksheet  
        For Each wks In Worksheets  
            Sheets.Add after:=wks  
        Next wks  
    End If  
  
    Worksheets(1).Name = "Dados"  
    Worksheets(2).Name = "Resultados"  
  
    Dim avdata(5) As Variant  
  
    avdata(0) = "agq"  
    avdata(1) = "ng0"  
    avdata(2) = "ngq"  
    avdata(3) = "ngn"
```

avdata(4) = "kgq"

avdata(5) = "sigma2"

Application.UseSystemSeparators = True

Dim range1 As Excel.Range

Set range1 = Sheets.Item(2).Range("\$A\$2:\$F\$65000")

Dim vector() As Variant

ReDim vector(range1.Rows.Count - 1, range1.Columns.Count - 1)

Dim l As Integer

l = 3

Dim i, k As Integer

Dim data\_inicio, data\_fim As Date

data\_inicio = Now()

Dim r, r2, r3, r4, r5, r6, agq, ng0, ngq, ngn, kgq, sigma2, agq0, ng00, ngq0, ngn0, kgq0, sigma20, e2\_ant, e2, N As Double

agq = CDbI(Sheets.Item(1).Range("A1").Offset(1, 0).Text)

ng0 = CDbI(Sheets.Item(1).Range("A1").Offset(1, 1).Text)

ngq = CDbI(Sheets.Item(1).Range("A1").Offset(1, 2).Text)

ngn = CDbI(Sheets.Item(1).Range("A1").Offset(1, 3).Text)

kgq = CDbI(Sheets.Item(1).Range("A1").Offset(1, 4).Text)

sigma2 = CDbI(Sheets.Item(1).Range("A1").Offset(1, 5).Text)

agq0 = agq

ng00 = ng0

ngq0 = ngq

ngn0 = ngn

kgq0 = kgq

sigma20 = sigma2

N = 0

While CStr(Sheets.Item(1).Range("A1").Offset(l, 0).Text) <> ""

N = N + 1

l = l + 1

Wend

'.....

For k = 3 To 5002 Step 1

'agq

agq0 = agq

While agq0 = agq

e2 = 0

e2\_ant = 0

I = 3

r = WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))

While CStr(Sheets.Item(1).Range("A1").Offset(l, 0).Text) <> ""

If CStr(Sheets.Item(1).Range("A1").Offset(l, 2).Text) <> "" Then

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGproj(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, ng0, ngq, ngn) ^ 2

e2 = e2 + (EpsilonGproj(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), WorksheetFunction.GammaInv(r, 0.25, 2), ng0, ngq, ngn) ^ 2

Else

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGpl(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, kgq, ng0, ngq, ngn) ^ 2

e2 = e2 + (EpsilonGpl(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l,  
 4).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text),  
 WorksheetFunction.GammaInv(r, 0.25, 2), kgq, ng0, ngq, ngn) ^ 2

End If

I = I + 1

Wend

If Not Err.Number <> 0 Then

If (Alpha(WorksheetFunction.GammaDist(agq, 0.25, 2, False) /  
WorksheetFunction.GammaDist(WorksheetFunction.GammaInv(r, 0.25, 2), 0.25, 2, False) \* WorksheetFunction.Max(10  
^ (-200), ProbP(sigma2, e2)) / WorksheetFunction.Max(10 ^ (-200), ProbP(sigma2, e2\_ant))) >=  
WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))) Then

agq = WorksheetFunction.GammaInv(r, 0.25, 2)

End If

End If

Wend

'ng0

ng00 = ng0

While ng00 = ng0

e2 = 0

e2\_ant = 0

l = 3

r2 = WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))

While CStr(Sheets.Item(1).Range("A1").Offset(l, 0).Text) <> ""

If CStr(Sheets.Item(1).Range("A1").Offset(l, 2).Text) <> "" Then

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGproj(CDb(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDb(Sheets.Item(1).Range("A1").Offset(l, 4).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDb(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, ng0, ngq, ngn)) ^ 2

e2 = e2 + (EpsilonGproj(CDb(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDb(Sheets.Item(1).Range("A1").Offset(l, 4).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDb(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, WorksheetFunction.NormInv(r2, 1.24, 1.2), ngq, ngn)) ^ 2

Else

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGpl(CDb(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 4).Text), CDb(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDb(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, kgq, ng0, ngq, ngn)) ^ 2

e2 = e2 + (EpsilonGpl(CDb(Sheets.Item(1).Range("A1").Offset(l, 0).Text), CDb(Sheets.Item(1).Range("A1").Offset(l,  
 4).Text), CDb(Sheets.Item(1).Range("A1").Offset(l, 3).Text), CDb(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq,  
 kgq, WorksheetFunction.NormInv(r2, 1.24, 1.2), ngq, ngn)) ^ 2

End If

l = l + 1

Wend

If Not Err.Number <> 0 Then

If (Alpha(WorksheetFunction.NormDist(ng0, 1.24, 1.2, False) /  
 WorksheetFunction.NormDist(WorksheetFunction.NormInv(r2, 1.24, 1.2), 1.24, 1.2, False) \*



## ESTIMAÇÃO CLÁSSICA E BAYESIANA DE PARÂMETROS DO MODELO *GLOBULUS 2.1*

---

```
WorksheetFunction.Max(10 ^ (-200), ProbP(sigma2, e2)) / WorksheetFunction.Max(10 ^ (-200), ProbP(sigma2,  
e2_ant))) >= WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))) Then
```

```
ng0 = WorksheetFunction.NormInv(r2, 1.24, 1.2)
```

```
End If
```

```
End If
```

```
Wend
```

```
'ngq
```

```
ngq0 = ngq
```

```
While ngq0 = ngq
```

```
e2 = 0
```

```
e2_ant = 0
```

```
l = 3
```

```
r3 = WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))
```

```
While CStr(Sheets.Item(1).Range("A1").Offset(l, 0).Text) <> ""
```

```
If CStr(Sheets.Item(1).Range("A1").Offset(l, 2).Text) <> "" Then
```

```
On Error Resume Next
```

```
e2_ant = e2_ant + (EpsilonGproj(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, ng0, ngq, ngn)) ^ 2
```

```
e2 = e2 + (EpsilonGproj(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, ng0, WorksheetFunction.NormInv(r3, -1, 1.2), ngn)) ^ 2
```

Else

On Error Resume Next

```
e2_ant = e2_ant + (EpsilonGpl(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),
CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, kgq, ng0, ngq, ngn)) ^ 2
```

```
e2 = e2 + (EpsilonGpl(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l,
4).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq,
kgq, ng0, WorksheetFunction.NormInv(r3, -1, 1.2), ngn)) ^ 2
```

End If

l = l + 1

Wend

If Not Err.Number <> 0 Then

```
If (Alpha(WorksheetFunction.NormDist(ngq, -1, 1.2, False) /
WorksheetFunction.NormDist(WorksheetFunction.NormInv(r3, -1, 1.2), -1, 1.2, False) * WorksheetFunction.Max(10 ^ (-
200), ProbP(sigma2, e2)) / WorksheetFunction.Max(10 ^ (-200), ProbP(sigma2, e2_ant))) >=
WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))) Then
```

```
ngq = WorksheetFunction.NormInv(r3, -1, 1.2)
```

End If

End If

Wend

'ngn

ngn0 = ngn

While ngn0 = ngn

e2 = 0

e2\_ant = 0

l = 3

r4 = WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))

While CStr(Sheets.Item(1).Range("A1").Offset(l, 0).Text) <> ""

If CStr(Sheets.Item(1).Range("A1").Offset(l, 2).Text) <> "" Then

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGproj(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, ngn0, ngq, ngn) ^ 2

e2 = e2 + (EpsilonGproj(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text),

## ESTIMAÇÃO CLÁSSICA E BAYESIANA DE PARÂMETROS DO MODELO *GLOBULUS 2.1*

---

```
Cdbl(Sheets.Item(1).Range("A1").Offset(l, 5).Text), Cdbl(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
Cdbl(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, ng0, ngq, WorksheetFunction.NormInv(r4, -0.55, 2)) ^ 2
```

```
Else
```

```
On Error Resume Next
```

```
e2_ant = e2_ant + (EpsilonGpl(Cdbl(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
Cdbl(Sheets.Item(1).Range("A1").Offset(l, 4).Text), Cdbl(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
Cdbl(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, kgq, ng0, ngq, ngn) ^ 2
```

```
e2 = e2 + (EpsilonGpl(Cdbl(Sheets.Item(1).Range("A1").Offset(l, 0).Text), Cdbl(Sheets.Item(1).Range("A1").Offset(l,  
4).Text), Cdbl(Sheets.Item(1).Range("A1").Offset(l, 3).Text), Cdbl(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq,  
kgq, ng0, ngq, WorksheetFunction.NormInv(r4, -0.55, 2)) ^ 2
```

```
End If
```

```
l = l + 1
```

```
Wend
```

```
If Not Err.Number <> 0 Then
```

```
If (Alpha(WorksheetFunction.NormDist(ngn, -0.55, 2, False) /  
WorksheetFunction.NormDist(WorksheetFunction.NormInv(r4, -0.55, 2), -0.55, 2, False) * WorksheetFunction.Max(10  
^ (-200), ProbP(sigma2, e2)) / WorksheetFunction.Max(10 ^ (-200), ProbP(sigma2, e2_ant))) >=  
WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))) Then
```

```
ngn = WorksheetFunction.NormInv(r4, -0.55, 2)
```

```
End If
```

```
End If
```

Wend

'kgq

kgq0 = kgq

While kgq0 = kgq

e2 = 0

e2\_ant = 0

l = 3

r5 = WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))

While CStr(Sheets.Item(1).Range("A1").Offset(l, 0).Text) <> ""

If CStr(Sheets.Item(1).Range("A1").Offset(l, 2).Text) <> "" Then

e2\_ant = e2\_ant + 0

e2 = e2 + 0

Else

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGpl(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
CDbl(Sheets.Item(1).Range("A1").Offset(l, 4).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
CDbl(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, kgq, ng0, ngq, ngn)) ^ 2

e2 = e2 + (EpsilonGpl(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l,  
4).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq,  
WorksheetFunction.GammaInv(r5, 478 / 2, 2), ng0, ngq, ngn)) ^ 2

End If

I = I + 1

Wend

If Not Err.Number <> 0 Then

If (Alpha(WorksheetFunction.GammaDist(kgq, 478 / 2, 2, False) /  
WorksheetFunction.GammaDist(WorksheetFunction.GammaInv(r5, 478 / 2, 2), 478 / 2, 2, False) \*  
WorksheetFunction.Max(10 ^ (-200), ProbP(sigma2, e2)) / WorksheetFunction.Max(10 ^ (-200), ProbP(sigma2,  
e2\_ant))) >= WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))) Then

kgq = WorksheetFunction.GammaInv(r5, 478 / 2, 2)

End If

End If

Wend

'sigma2

sigma20 = sigma2

While sigma20 = sigma2

e2\_ant = 0

I = 3

r6 = WorksheetFunction.Max(10 ^ (-10), (WorksheetFunction.RandBetween(0, 1000000) / 1000000))

While CStr(Sheets.Item(1).Range("A1").Offset(l, 0).Text) <> ""

If CStr(Sheets.Item(1).Range("A1").Offset(l, 2).Text) <> "" Then

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGproj(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 5).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, ng0, ngq, ngn) ^ 2

Else

On Error Resume Next

e2\_ant = e2\_ant + (EpsilonGpl(CDbI(Sheets.Item(1).Range("A1").Offset(l, 0).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 4).Text), CDbI(Sheets.Item(1).Range("A1").Offset(l, 3).Text),  
 CDbI(Sheets.Item(1).Range("A1").Offset(l, 1).Text), agq, kgq, ng0, ngq, ngn) ^ 2

End If

I = I + 1

Wend

If Not Err.Number <> 0 Then

```
If (Alpha(WorksheetFunction.GammaDist(sigma2, 1, 0.437, False) /  
WorksheetFunction.GammaDist(WorksheetFunction.GammaInv(r6, 1, 0.437), 1, 0.437, False) *  
WorksheetFunction.Max(10 ^ (-200), ProbPS(WorksheetFunction.GammaInv(r6, 1, 0.437), e2_ant, N)) /  
WorksheetFunction.Max(10 ^ (-200), ProbPS(sigma2, e2_ant, N))) >= WorksheetFunction.Max(10 ^ (-10),  
(WorksheetFunction.RandBetween(0, 1000000) / 1000000))) Then
```

```
sigma2 = WorksheetFunction.GammaInv(r6, 1, 0.437)
```

```
End If
```

```
End If
```

```
Wend
```

```
vector(k - 3, 0) = CDbI(agq)
```

```
vector(k - 3, 1) = CDbI(ng0)
```

```
vector(k - 3, 2) = CDbI(ngq)
```

```
vector(k - 3, 3) = CDbI(ngn)
```

```
vector(k - 3, 4) = CDbI(kgq)
```

```
vector(k - 3, 5) = CDbI(sigma2)
```

```
Next
```

```
Sheets.Item(2).Range("$A$1:$F$1").Value = avdata
```

```
range1.Value = vector
```

```
data_fim = Now()
```

```
MsgBox ("Data Inicio = " & CStr(data_inicio) & " || Data Fim = " & CStr(data_fim))
```

```
End Sub
```



## ii) Cálculo do DIC (Deviance)

```
Sub Macro1()
```

```
    If Sheets.Count < 2 Then
```

```
        Dim wks As Worksheet
```

```
        For Each wks In Worksheets
```

```
            Sheets.Add after:=wks
```

```
        Next wks
```

```
    End If
```

```
    Worksheets(1).Name = "Dados"
```

```
    Worksheets(2).Name = "Resultados"
```

```
    Dim avdata(0) As Variant
```

```
    avdata(0) = "DIC"
```

```
    Application.UseSystemSeparators = True
```

```
    Dim range1 As Excel.Range
```

```
    Set range1 = Sheets.Item(2).Range("$A$2:$A$2")
```

```
    Dim vector() As Variant
```

```
    ReDim vector(range1.Rows.Count - 1, range1.Columns.Count - 1)
```

```
    Dim l As Integer
```

```
    Dim i, k As Integer
```

```
    Dim e2, N As Double
```

N = 0

While CStr(Sheets.Item(1).Range("H1").Offset(l, 0).Text) <> ""

N = N + 1

l = l + 1

Wend

l = 1

e2 = 0

While CStr(Sheets.Item(1).Range("H1").Offset(l, 0).Text) <> ""

k = 1

e2 = 0

While CStr(Sheets.Item(1).Range("A1").Offset(k, 0).Text) <> ""

If CStr(Sheets.Item(1).Range("A1").Offset(k, 1).Text) <> "" Then

On Error Resume Next

e2 = e2 + (EpsilonN\_1(CDbI(Sheets.Item(1).Range("A1").Offset(k, 0).Text), CDbI(Sheets.Item(1).Range("A1").Offset(k, 1).Text),  
CDbI(Sheets.Item(1).Range("A1").Offset(k, 2).Text), CDbI(Sheets.Item(1).Range("A1").Offset(k, 3).Text),

Cdbl(Sheets.Item(1).Range("H1").Offset(l, 0).Text), Cdbl(Sheets.Item(1).Range("H1").Offset(l, 1).Text),  
Cdbl(Sheets.Item(1).Range("H1").Offset(l, 2).Text))) ^ 2

Else

e2 = e2 + (EpsilonNpl(Cdbl(Sheets.Item(1).Range("A1").Offset(k, 0).Text), Cdbl(Sheets.Item(1).Range("A1").Offset(k, 2).Text),  
Cdbl(Sheets.Item(1).Range("A1").Offset(k, 3).Text), Cdbl(Sheets.Item(1).Range("A1").Offset(k, 4).Text),  
Cdbl(Sheets.Item(1).Range("H1").Offset(l, 0).Text), Cdbl(Sheets.Item(1).Range("H1").Offset(l, 1).Text),  
Cdbl(Sheets.Item(1).Range("H1").Offset(l, 2).Text))) ^ 2

End If

k = k + 1

Wend

If l <= 4000 Then

vector(0, 0) = vector(0, 0) - 4 \* (-Log((2 \* WorksheetFunction.Pi \* Cdbl(Sheets.Item(1).Range("H1").Offset(l, 3).Text)) ^ ((k - 1)  
/ 2)) - 1 / (2 \* Cdbl(Sheets.Item(1).Range("H1").Offset(l, 3).Text)) \* e2) / (N - 1)

Else

vector(0, 0) = vector(0, 0) + 2 \* (-Log((2 \* WorksheetFunction.Pi \* Cdbl(Sheets.Item(1).Range("H1").Offset(l, 3).Text)) ^ ((k - 1)  
/ 2)) - 1 / (2 \* Cdbl(Sheets.Item(1).Range("H1").Offset(l, 3).Text)) \* e2)

End If

l = l + 1

Wend

Sheets.Item(2).Range("\$A\$1:\$A\$1").Value = avdata

range1.Value = vector

End Sub

## Bibliografia

- 1) Akaike, H. (1973). *Information Theory and an Extension of the Maximum Likelihood Principle*. In B. Petrov and F. Csake, eds., Second International Symposium on Information Theory. Akademiai Kiado. Budapest.
- 2) Amaro, Ana P. N. (1997). *Modelação do crescimento de povoamentos de Eucalyptus globulus Labill de 1ª rotação em Portugal*. Tese de Doutoramento em Engenharia de Sistemas. IST/UTL. Lisboa.
- 3) Amemiya, T. (1983). *Handbook of Econometrics: Non-Linear Regression Models (Chapter 6)*. North-Holland Publishing Company. Stanford.
- 4) Borders, B. E. (1989). *System of Equations in forest stand modelling*. For. Sci., 35: 548- 558.
- 5) Casella, G. e Berger, R. L. (2002). *Statistical Inference, 2<sup>nd</sup>Ed.* . Duxbury Advanced Series. USA.
- 6) Dempster, A.P. (1974). *The direct use of likelihood in significance testing*. In Proc. Conf. Foundational Questions in Statistical Inference (eds. O. Barndorff-Nielsen. P. Blaesild and G. Sihon),335-352.
- 7) Erdman, D. , Little, M. (1994). *Nonlinear Regression Analysis and Nonlinear Simulation Models*. SAS Institute Inc. , Cary, NC.
- 8) Gallant, A. (1975). *Seemingly Unrelated Nonlinear Regressions*. Journal of Econometrics 3 (1975) 35-50. North Carolina State University, Raleigh, NC 27607, U.S.A.
- 9) Gamerman, D. and Lopes, F. H. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, 2<sup>nd</sup>Ed.* . Chapman & Hall/CRC. USA.
- 10) Geweke, J. (1992). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (with discussion)*. In Bayesian Statistics 4, (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds), 169-193. Oxford University Press. Oxford.

- 11) Gilks, W. R. , Richardson, S. and Spiegelhalter, D. J. (eds) (1996). *Markov Chain Monte Carlo in Practice*. Chapman and Hall, New York.
- 12) Green, E. J. , Roesch Jr., F. A. , Smith, A. F. M. and Strawderman, W. E. (1994). *Bayesian Estimation for Three-Parameter Weibull Distribution with Tree Diameter Data*. *Biometrics* 50, 254-269. USA.
- 13) Green, E. J. and Strawderman, W. E. (1996). *A Bayesian growth and yield model for slash pine plantations*. *Journal of Applied Statistics*, Vol. 23, Nos 2&3, 285-299. Oxford.
- 14) Greene, W. H. (2002). *Econometric Analysis, 5<sup>th</sup>Ed.* . Prentice Hall. New Jersey.
- 15) Griffiths, W. E. , Hill, R. C. and Judge, G. G. (1993). *Learning and Practicing of Econometrics*, Wiley. New York.
- 16) Griffiths, W. E. (2001). *Bayesian Inference in the Seemingly Unrelated Regressions Model*. Economics Department, University of Melbourne Vic. 3010 Australia.
- 17) Jennrich, R. I. (1969). *The Asymptotic Properties of Nonlinear Least Squares Estimators*. *Annals of Statistics*, 2, pp. 633–643.
- 18) Judge, G. G. , Hill, R. C. , Griffiths, W. E. , Lütkepohl, H. and Lee, T.-C. (1988). *Introduction to The Theory and Practice of Econometrics, 2<sup>nd</sup>Ed.* . Wiley. New York.
- 19) Koop, G. (2003). *Bayesian Econometrics*. Wiley. London.
- 20) Muller, P. (1991). *Metropolis based posterior integration schemes*. Technical Report, Statistics Department, Pardue University.
- 21) Murteira, B. (1988). *Estatística: Inferência e Decisão*. INCM. Lisboa.
- 22) Murteira, B. (1990). *Probabilidades e Estatística, 2<sup>a</sup>Ed. Rev. ,vol. I e II.* McGraw-Hill. Lisboa.
- 23) Paulino, D. C. , Turkman, M. A. e Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian. Lisboa.

- 24) Pedrosa, A. C. e Gama, S. M. (2004). *Introdução Computacional à Probabilidade e Estatística*. Porto Editora. Porto.
- 25) Pindyck, Robert S. and Rubinfeld, Daniel L. (1981). *Econometric models and economic forecasts*. McGraw-Hill. New York.
- 26) Reed, D. D. (1986). *Estimation procedures for analytically related growth and yield models*. In Annual Meeting of the Society of American Foresters held at Birmingham, AL: 89-92.
- 27) Smith, B. J. (2007). *boa: An R Package for MCMC Output Convergence Assessment and Posterior Inference*. Journal of Statistical Software, Volume 21, Issue 11.
- 28) Spiegelhalter, D. J. , Best, N. G. , Carlin, B. P. e van der Linde, A. (2002). *Bayesian measures of model complexity and fit (with discussion)*. J. Royal Statist. Soc., B, 64, 583-639.
- 29) Tomé, M. , Ribeiro, F. , Soares, P. (2001). *O modelo GLOBULUS 2.1. Relatórios Técnico-Científicos do GIMREF nº 1/2001*. Centro de Estudos Florestais, Instituto Superior de Agronomia, Lisboa.
- 30) Velosa, S. F. , Pestana, D. D. (2008). *Introdução à Probabilidade e à Estatística, 4ªEd.*. Fundação Calouste Gulbenkian. Lisboa.
- 31) Wakefield, J. C. , Smith, A. F. M. , Racine-Poon, A. , Gelfand, A. E. (1994). *Bayesian Analysis of Linear and Non-Linear Population Models by using the Gibbs Sampler*. App. Statist. (1994) 43, No 1, pp 201-221.
- 32) Wooldridge, J. M. (2005). *Introductory Econometrics: A Modern Approach*. South Western College Publishing. USA.
- 33) Yuancai, L. , Marques, C. P. , Bento, J. M. (2001). *Simultaneous modelling of stand volume yield, dominant height and basal area growth models*. Journal of Forest Science, 47, 2001 (7): 285-293. Praha.
- 34) Zellner, A. (1993). *Bayesian Method of Moments (BMOM) Analysis of Mean and Regression Models*. University of Chicago.