



UNIVERSIDADE DE ÉVORA

ESCOLA DE CIÊNCIAS E TECNOLOGIA

DEPARTAMENTO DE MATEMÁTICA

Mestrado em Modelação Estatística e Análise de Dados

Dissertação

**As Projecções Demográficas na (Re)definição de uma rede de Ensino
Superior em Portugal**

Rui Miguel dos Santos Dias

Orientador:

Professora Doutora Maria Filomena Ferreira Mendes,
Universidade de Évora

Professor Doutor Paulo de Jesus Infante dos Santos,
Universidade de Évora

Évora

2012

Mestrado em Modelação Estatística e Análise de Dados

Dissertação

**As Projecções Demográficas na (Re)definição de uma rede de Ensino
Superior em Portugal**

Rui Miguel dos Santos Dias

Orientador:

Professora Doutora Maria Filomena Ferreira Mendes, Universidade de Évora

Professor Doutor Paulo de Jesus Infante dos Santos, Universidade de Évora

Agradecimentos

Em primeiro lugar gostaria de transmitir os meus sinceros agradecimentos pela dedicação e paciência que os meus dois orientadores de dissertação, Professora Doutora Maria Filomena Ferreira Mendes e Professor Doutor Paulo de Jesus Infante dos Santos tiveram ao longo deste ano. Mesmo que em fases distintas os dois souberam manifestar interesse, apoio e disponibilidade, fundamentais para conseguir ultrapassar mais esta fase da minha vida.

Um agradecimento especial também à minha tia Doutora Arminda da Conceição Dias Escária pelo apoio e disponibilidade para rever toda a minha tese.

O meu sincero agradecimento à minha namorada e companheira Natacha por o apoio e motivação incondicional dados ao longo desta viagem que agora acaba. Sem ela não teria conseguido chegar ao fim.

Não quero deixar de agradecer à minha mãe que não tendo estado presente me deu forças para continuar, com a sua memória após a morte e valores transmitidos em vida. Certamente que a deixei orgulhosa.

Finalmente, agradecer a todos aqueles que tornaram possível esta dissertação. Ao longo do projecto foram preciosas as ajudas que obtive. Para eles e elas, que de uma forma ou de outra me ajudaram, aqui fica o agradecimento.

Resumo

A presente dissertação tem como principal objectivo o de dar um contributo, recorrendo a duas metodologias estatísticas (*Lee-Carter* e *Booth-Maindonald-Smith*), para uma projecção mais fiável das taxas de mortalidade em Portugal a médio prazo (25 anos), de modo a dar uma maior consistência às fases seguintes do projecto de investigação em que se insere. Mais especificamente, considerando os dados referentes a Portugal nas bases de dados da *Human Mortality Database*, pretende-se projectar o valor das taxas mortalidade, realizar comparações de resultados usando diferentes metodologias e analisar os ajustamentos dos modelos utilizados. Os resultados deste estudo indicam que existem poucas diferenças entre as projecções dos dois modelos, assim como se demonstra que a utilização dos dois modelos é aceitável, apresentando valores muito interessantes ao nível de erros. A análise de sensibilidade feita aos modelos permite evidenciar a sua robustez.

Palavras-chave: Modelos de Projecção, Mortalidade, método de Lee Carter, método de Booth-Maindonald-Smith, População.

Abstract

Demographic projections in the (Re)definition of a network for the higher education in Portugal

The present dissertation's main objective is to make a contribution, using two models (*Lee-Carter* and *Booth-Maindonald-Smith*), for a more reliable projection of mortality rates in Portugal at a medium term (25 years), so as to give greater consistency to subsequent phases of the research project, which it operates. More specifically, considering the data for Portugal in the databases of the *Human Mortality Database* it is intended to project the value of mortality rates, comparisons of results using different methodologies and analyze the adjustments of the models. The results of this study indicate that there are few differences between the projections of the two models, as well as demonstrated that the use of the two models is acceptable with values very interesting in terms of errors. The sensitivity analysis performed on the models highlights models robustness.

Keywords: Projection Models, Mortality, Lee Carter method, Maindonald Booth-Smith method, Population.

Índice

Agradecimentos.....	I
Resumo.....	III
Abstract	V
Índice.....	VII
Índice de Gráficos	XV
Índice de Tabelas.....	XIX
Índice de Comandos.....	XXI
1. Introdução	1
2. Metodologias.....	6
2.1. Modelo <i>Lee-Carter</i> (MLC)	6
2.2. Modelo <i>Booth-Maindonald-Smith</i> (MBMS).....	8
2.3. Erros associados ao ajustamento dos modelos.....	9
2.4. Erro Quadrático Médio.....	10
3. Software e Linguagem R.....	11
4. Projecções vs Valores Observados.....	13
4.1. População Total.....	15
4.1.1. Projecções.....	15
4.1.2. Erros Associados	17
4.1.3. Projecções dentro ou fora do intervalo de predição	22
4.2. População Feminina.....	23
4.2.1. Projecções.....	23
4.2.2. Erros Associados	25
4.2.3. Projecções dentro ou fora do intervalo de predição	27
4.3. População Masculina.....	28
4.3.1. Projecções.....	28
4.3.2. Erros Associados	30
4.3.3. Projecções dentro ou fora do intervalo de predição	32
4.4. Síntese Final	33
5. Ajustamento dos MLC e MBMS no Software R	35
5.1. Recolha dos parâmetros	37
6. Análise do Ajustamento	39
6.1. Cálculo das Taxas de Mortalidade Ajustadas.....	39

6.2.	Modelo <i>Lee-Carter</i>	39
6.2.1.	Anos 1981-2009.....	39
6.2.1.1.	População Total.....	39
6.2.1.1.1.	Erros Associados.....	39
6.2.1.1.2.	Erro Quadrático Médio.....	41
6.2.1.1.3.	Síntese Final.....	42
6.2.1.2.	População Feminina e Masculina.....	42
6.2.2.	Anos 1990-2009.....	43
6.2.2.1.	População Total.....	43
6.2.2.1.1.	Erros Associados.....	43
6.2.2.1.2.	Erro Quadrático Médio.....	45
6.2.2.1.3.	Síntese Final.....	46
6.2.2.2.	População Feminina e Masculina.....	46
6.2.3.	Anos 2000-2009.....	47
6.2.3.1.	População Total.....	47
6.2.3.1.1.	Erros Associados.....	47
6.2.3.1.2.	Erro Quadrático Médio.....	49
6.2.3.1.3.	Síntese Final.....	50
6.2.3.2.	População Feminina e Masculina.....	50
6.3.	Modelo <i>Booth-Maindonald-Smith</i>	51
6.3.1.	Anos 1981-2009.....	51
6.3.1.1.	População Total.....	51
6.3.1.1.1.	Erros Associados.....	51
6.3.1.1.2.	Erro Quadrático Médio.....	53
6.3.1.1.3.	Síntese Final.....	54
6.3.1.2.	População Feminina e Masculina.....	54
6.3.2.	Anos 1990-2009.....	55
6.3.2.1.	População Total.....	55
6.3.2.1.1.	Erros Associados.....	55
6.3.2.1.2.	Erro Quadrático Médio.....	57
6.3.2.1.3.	Síntese Final.....	57
6.3.2.2.	População Feminina e Masculina.....	57
6.3.3.	Anos 2000-2009.....	58
6.3.3.1.	População Total.....	58

6.3.3.1.1.	Erros Associados	58
6.3.3.1.2.	Erro Quadrático Médio.....	60
6.3.3.1.3.	Síntese Final	60
6.3.3.2.	População Feminina e Masculina.....	61
7.	Projecções	63
7.1.	Ajustamento e Recolha das Projecções do <i>Software R</i>	63
7.2.	Taxas de Mortalidade	65
7.2.1.	Anos 1981-2009	65
7.2.1.1.	População Total.....	66
7.2.1.2.	População Feminina e Masculina.....	68
7.2.2.	Anos 1990-2009	68
7.2.2.1.	População Total.....	68
7.2.2.2.	População Feminina e Masculina.....	71
7.2.3.	Anos 2000-2009	71
7.2.3.1.	População Total.....	72
7.2.3.2.	População Feminina e Masculina.....	74
7.3.	Óbitos	74
7.3.1.	Anos 1981-2009	75
7.3.2.	Anos 1990-2009	77
7.4.	Esperança de Vida.....	79
7.4.1.	Anos 1981-2009	79
7.4.2.	Anos 1990-2009	81
7.5.	Síntese Final	82
8.	Análise de Sensibilidade	83
9.	Conclusões	89
9.1.	Projecções vs Valores Observados.....	89
9.2.	Ajustamento dos modelos	90
9.3.	Projecções	92
9.4.	Análise de Sensibilidade	92
9.5.	Síntese Final	93
10.	Bibliografia	95
11.	Anexos.....	99
	Anexo 1 – EPA nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.	99

Anexo 2 – EPAL nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.	99
Anexo 3 – EPA nos anos de ajustamento entre 1981 e 2009 do MLC para a população feminina.	100
Anexo 4 – EPAL nos anos de ajustamento entre 1981 e 2009 do MLC para a população feminina.	100
Anexo 5 – EPA nos anos de ajustamento entre 1981 e 2009 do MLC para a população masculina.	101
Anexo 6 – EPAL nos anos de ajustamento entre 1981 e 2009 do MLC para a população masculina.	101
Anexo 7 – EPAM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1981 a 2009.	102
Anexo 8 – EPALM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1981 a 2009.	102
Anexo 9 – EPA nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.	103
Anexo 10 – EPAL nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.	103
Anexo 11 – EPA nos anos de ajustamento entre 1990 e 2009 do MLC para a população feminina.	104
Anexo 12 – EPA nos anos de ajustamento entre 1990 e 2009 do MLC para a população masculina.	104
Anexo 13 – EPAM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1990 a 2009.	105
Anexo 14 – EPAL nos anos de ajustamento entre 1990 e 2009 do MLC para a população feminina.	105
Anexo 15 – EPAL nos anos de ajustamento entre 1990 e 2009 do MLC para a população masculina.	106
Anexo 16 – EPALM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1990 a 2009.	106
Anexo 17 – EPA nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.	107
Anexo 18 – EPAL nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.	107
Anexo 19 – EPA nos anos de ajustamento entre 2000 e 2009 do MLC para a população feminina.	108
Anexo 20 – EPA nos anos de ajustamento entre 2000 e 2009 do MLC para a população masculina.	108

Anexo 21 – EPAM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 2000 a 2009.....	109
Anexo 22 – EPAL nos anos de ajustamento entre 2000 e 2009 do MLC para a população feminina.	109
Anexo 23 – EPAL nos anos de ajustamento entre 2000 e 2009 do MLC para a população masculina.	110
Anexo 24 – EPALM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 2000 a 2009.....	110
Anexo 25 – EPA nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.	111
Anexo 26 – EPAL nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.	111
Anexo 27 – EPA nos anos de ajustamento entre 1981 e 2009 do MBMS para a população feminina.	112
Anexo 28 – EPA nos anos de ajustamento entre 1981 e 2009 do MBMS para a população masculina.	112
Anexo 29 – EPAM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1981 a 2009.....	113
Anexo 30 – EPAL nos anos de ajustamento entre 1981 e 2009 do MBMS para a população feminina.	113
Anexo 31 – EPAL nos anos de ajustamento entre 1981 e 2009 do MBMS para a população masculina.	114
Anexo 32 – EPALM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1981 a 2009.....	114
Anexo 33 – EPA nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.	115
Anexo 34 – EPAL nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.	115
Anexo 35 – EPA nos anos de ajustamento entre 1990 e 2009 do MBMS para a população feminina.	116
Anexo 36 – EPA nos anos de ajustamento entre 1990 e 2009 do MBMS para a população masculina.	116
Anexo 37 – EPAM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1990 a 2009.....	117
Anexo 38 – EPAL nos anos de ajustamento entre 1990 e 2009 do MBMS para a população feminina.	117
Anexo 39 – EPAL nos anos de ajustamento entre 1990 e 2009 do MBMS para a população masculina.	118

Anexo 40 – EPALM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1990 a 2009.....	118
Anexo 41 – EPA nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.....	119
Anexo 42 – EPAL nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.....	119
Anexo 43 – EPA nos anos de ajustamento entre 2000 e 2009 do MBMS para a população feminina.....	120
Anexo 44 – EPA nos anos de ajustamento entre 2000 e 2009 do MBMS para a população masculina.....	120
Anexo 45 – EPAM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 2000 a 2009.....	121
Anexo 46 – EPAL nos anos de ajustamento entre 2000 e 2009 do MBMS para a população feminina.....	121
Anexo 47 – EPAL nos anos de ajustamento entre 2000 e 2009 do MBMS para a população masculina.....	122
Anexo 48 – EPALM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 2000 a 2009.....	122
Anexo 49 – Resumo dos erros associados para todos os modelos.....	123
Anexo 50 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1981 a 2009 para população feminina.....	123
Anexo 51 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1981 a 2009 para população masculina.....	124
Anexo 52 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1981 a 2009 para população feminina.....	124
Anexo 53 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1981 a 2009 para população masculina.....	125
Anexo 54 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1990 a 2009 para população feminina.....	125
Anexo 55 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1990 a 2009 para população masculina.....	126
Anexo 56 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1990 a 2009 para população feminina.....	126
Anexo 57 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1990 a 2009 para população masculina.....	127
Anexo 58 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 2000 a 2009 para população feminina.....	127
Anexo 59 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 2000 a 2009 para população masculina.....	128

Anexo 60 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 2000 a 2009 para população feminina.	128
Anexo 61 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 2000 a 2009 para população masculina.	129
Anexo 62 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1990 e 2009 para a população feminina, com intervalos de predição.	129
Anexo 63 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1990 e 2009 para a população masculina, com intervalos de predição.	130

Índice de Gráficos

Gráfico 1 - Curvas de Mortalidade Projectadas pelos MLC para a população total.	15
Gráfico 2 - Curvas de Mortalidade Projectadas pelos MBMS para a população total... 16	16
Gráfico 3 - Curvas de Mortalidade Observadas para a população total.	16
Gráfico 4 – EPA das diversas projecções dos MLC para a população total.	18
Gráfico 5 – EPA das projecções dos MBMS para a população total.	18
Gráfico 6 – EPAM, comparação entre MLC e MBMS para a população total.	19
Gráfico 7 – EPAL das projecções dos MLC para a população total.	20
Gráfico 8 – EPAL das projecções dos MBMS para a população total.	21
Gráfico 9 – EPALM, comparação entre MLC e MBMS para a população total.	21
Gráfico 10 – Curvas de Mortalidade Projectadas pelos MLC para a população feminina.	24
Gráfico 11 – Curvas de Mortalidade Projectadas pelos MBMS para a população feminina.	24
Gráfico 12 – Curvas de Mortalidade Observadas para a população feminina.	25
Gráfico 13 – EPAL das projecções dos MLC para a população feminina.	26
Gráfico 14 – EPAL das projecções dos MBMS para a população feminina.	26
Gráfico 15 – EPALM, comparação entre MLC e MBMS para a população feminina. .	27
Gráfico 16 – Curvas de Mortalidade Projectadas pelos MLC para a população masculina.	29
Gráfico 17 – Curvas de Mortalidade Projectadas pelos MBMS para a população masculina.	29
Gráfico 18 – Curvas de Mortalidade Observadas para a população masculina.	30
Gráfico 19 – EPAL das projecções dos MLC para a população masculina.	31
Gráfico 20 – EPAL das projecções dos MBMS para a população masculina.	31
Gráfico 21 – EPALM, comparação entre MLC e MBMS para a população masculina.	32
Gráfico 22 – EPAM nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.	40
Gráfico 23 – EPALM nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.	41
Gráfico 24 – EPAM nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.	44

Gráfico 25 – EPALM nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.....	45
Gráfico 26 – EPAM nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.....	48
Gráfico 27 – EPALM nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.....	49
Gráfico 28 – EPAM nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.....	52
Gráfico 29 – EPALM nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.....	53
Gráfico 30 – EPAM nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.....	56
Gráfico 31 – EPALM nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.....	56
Gráfico 32 – EPAM nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.....	59
Gráfico 33 – EPALM nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.....	59
Gráfico 34 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1981 a 2009 para população total.....	66
Gráfico 35 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1981 a 2009 para população total.....	67
Gráfico 36 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1990 a 2009 para população total.....	69
Gráfico 37 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1990 a 2009 para população total.....	70
Gráfico 38 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 2000 a 2009 para população total.....	72
Gráfico 39 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 2000 a 2009 para população total.....	73
Gráfico 40 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1981 e 2009 para a população feminina, com intervalos de predição.....	75
Gráfico 41 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1981 e 2009 para a população masculina, com intervalos de predição.....	76

Gráfico 42 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1990 e 2009 para a população feminina, com intervalos de predição.....	77
Gráfico 43 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1990 e 2009 para a população masculina, com intervalos de predição.....	78
Gráfico 44 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1981 e 2009 para a população feminina, com intervalos de predição.....	79
Gráfico 45 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1981 e 2009 para a população masculina, com intervalos de predição.....	80
Gráfico 46 – Erro Percentual Algébrico para o total da população no ano de ajustamento de 2008.	84
Gráfico 47 – Erro Percentual Algébrico das populações totais calculadas para o ano de 2008.	85
Gráfico 48 – EPAL, para as populações com acréscimo de 15% na mortalidade observada para o ano de 2008.....	86
Gráfico 49 – EPAL, para as populações com acréscimo de 100% na mortalidade observada para o ano de 2008.....	87
Gráfico 50 – EPAL, para as populações com acréscimo de 350% na mortalidade observada para o ano de 2008.....	87
Gráfico 51 – Impacto na população de 18 anos de idade dos erros das taxas de mortalidade no ano de 2008.....	88

Índice de Tabelas

Tabela 1 – Média de Percentagem de Valores Dentro do Intervalo nos dois modelos para a população total.	22
Tabela 2 – Média de Percentagem de Valores Dentro do Intervalo nos dois modelos para a população feminina.	28
Tabela 3 – Média de Percentagem de Valores Dentro do Intervalo nos dois modelos para a população masculina.	32
Tabela 4 – Esperança de vida projectada para os diferentes períodos, sexos e modelos.	81

Índice de Comandos

Comando 1 – Ajustamento dos Modelos Lee-Carter e Booth-Maindonald-Smith	13
Comando 2 – Ajustamento do Modelo de Projecções para o LC e BMS	14
Comando 3 – Obtenção das projecções	14
Comando 4 – Carregamento duma base de dados do <i>Human Mortality Database</i>	35
Comando 5 – Ajustamento dos MLC, parte 1	35
Comando 6 – Ajustamento dos MLC, parte 2	36
Comando 7 – Ajustamento dos MBMS.....	36
Comando 8 – Obtenção dos parâmetros do modelo <i>ax , bx</i>	37
Comando 9 – Obtenção do parâmetro do modelo <i>kt</i>	38
Comando 10 - Ajustamento de Projecções para os Modelos Lee-Carter.	63
Comando 11 – Ajustamento de Projecções para os Modelos Booth-Maindonald-Smith.	64
Comando 12 – Obtenção das Taxas de Mortalidade Projectadas.....	64

1. Introdução

Esta dissertação insere-se no plano de trabalhos de uma bolsa de investigação no âmbito do projecto PTDC/CPE-PEC/103727/2008, designado por “Redefinição de uma rede de ensino superior em Portugal: desafios decorrentes da demografia, do crescimento económico e da coesão regional”, co-financiado pela Fundação para a Ciência e a Tecnologia e pelo FEDER através do Programa POCI 2010.

O projecto visa contribuir para a identificação de cenários relativos à dimensão óptima e a distribuição da rede de Instituições de Ensino Superior em Portugal, assumindo, em simultâneo, a consideração de dois objectivos: as pressões com vista a uma utilização eficiente de recursos públicos e a necessidade de melhorar o crescimento económico e a coesão territorial.

Uma rede de Ensino Superior em Portugal deve ter por base a evolução esperada da procura. Essa procura é determinada pelo número de jovens que atingirão os 18 anos, nos próximos anos. Todos os jovens que, no período de tempo compreendido entre 2012 e 2030, estarão em condições de ingressarem no ensino superior já nasceram, pelo que o volume de estudantes dependerá, em primeira instância, das probabilidades de sobrevivência a que estarão sujeitos ano após ano. Em segunda instância, obviamente, serão função das taxas de conclusão versus taxas de abandono do ensino secundário. Finalmente, poder-se-á ainda considerar a influência das migrações. Nesta investigação pretende-se determinar qual o volume da procura, tendo em consideração apenas a questão da mortalidade que se assume como o principal factor da evolução do número de jovens.

Os estudos existentes sobre a procura do Ensino Superior em Portugal, habitualmente, não contemplam cenários de projecção da população, em particular dos estratos em condições de entrar no Ensino Superior, nem estimativas da população futura com idade de entrada. Para além de outras variáveis que condicionam necessariamente o acesso a este grau de ensino (tais como a taxa de conclusão do ensino secundários, a localização geográfica, o perfil socioeconómico, entre outras), a estimação da população com idade de 18 anos, entre os 18 e os 24 anos e os 18 e os 30 anos distingue este projecto conferindo-lhe uma sólida base de partida. Para além disso, este projecto incorpora uma análise de intervalos de confiança para as projecções igualmente inovador neste tipo de estudos, podendo ser elaborados e analisados vários

cenários, no caso específico de mortalidade, desde os mais optimistas aos mais pessimistas.

Em seguimento dos objectivos anteriormente referidos, esta dissertação pretende contribuir, recorrendo a diversas metodologias estatísticas, com a projecção de taxas de mortalidade que possam dar a maior consistência possível às fases seguintes do projecto de investigação em que se insere, onde se procederá à projecção de uma subpopulação e introdução de factores de natureza socioeconómica.

Ao longo dos últimos anos a preocupação da sociedade, dos demógrafos e estatísticos com o envelhecimento da população dos países ocidentais e mais desenvolvidos tem vindo a aumentar. Com o avanço da ciência e medicina a população tem a tendência de cada vez viver mais anos e com a diminuição da natalidade nesses países o envelhecimento da população é uma inevitabilidade.

Talvez por culpa desse aumento de preocupação a comunidade científica tem dedicado especial atenção ao estudo da população e também ao estudo da modelação na área. Nos anos mais recentes surgiram inúmeros modelos e métodos estatísticos com o intuito de estudar, analisar e projectar um fenómeno cuja compreensão poderá ser fundamental para a sustentabilidade da sociedade como a conhecemos.

No final do século XX assistiu-se ao surgimento de inúmeros métodos de previsão da mortalidade em todo o Mundo. Esta tendência crescente verificou-se devido ao declínio muito acentuado da mortalidade, fundamentalmente nos países mais desenvolvidos. Numa primeira abordagem verifica-se que a mortalidade diminuiu drasticamente nas idades mais jovens, mortalidade infantil e juvenil, e também nas gerações mais velhas. É sem dúvida um dado importante, significando um passo em frente no que concerne ao aumento populacional, com menos mortes e maior esperança de vida. Pode no entanto ser um problema e desafio para a Segurança Social de cada país.

Nos últimos 20 anos foram muitos os métodos construídos para modelar e projectar taxas de mortalidade ao longo dos anos e das idades da população, assim como para taxas de natalidade e também para a migração, embora neste último caso com mais dificuldades, visto ser um tema de enorme dificuldade de análise.

No campo da mortalidade existem vários métodos de projecção de taxas de mortalidade. O mais conhecido, e de uma forma geral mais reconhecido pela

comunidade científica, é o modelo *Lee-Carter* (1992). A partir do modelo *Lee-Carter* foram construídas inúmeras variantes, entre as quais mudanças efectuadas no modelo por *Booth, Maindonald e Smith* (2002), no intuito de melhor ajustar a análise aos estudos pretendidos, assim como melhorar o resultado final das projecções realizadas. Vários investigadores têm dedicado muitas horas de trabalho com a publicação de inúmeros artigos para melhorar a análise e modelação do fenómeno de mortalidade nos últimos anos, entre os quais *Ronald Lee* (2000), *Lawrence Carter* (1992), *Timothy Miller* (2000), *Heather Booth* (2001), *John Maindonald* (2002), *Len Smith* (2005), *Shahid Ullah* (2007), *Rob Hyndman* (2007, principal responsável pelo pacote de R utilizado no projecto), *Han Lin Shang* (2011), *Piet de Jong* (2006), *Leonie Tickle* (2006), *Nan Li* (2000), *Shripad Tuljapurkar* (2000), *Carl Boe* (2000), *Coelho* (2001), entre muitos outros.

Neste projecto serão analisados dois modelos de projecção da mortalidade para Portugal, o modelo *Lee-Carter* e o modelo *Booth-Maindonald-Smith*. Os principais objectivos da dissertação são a comparação entre os dois métodos, assim como a verificação se a sua utilização será apropriada para a população portuguesa.

Lee e Carter (1992) referiram que a esperança de vida nos Estados Unidos da América aumentou de 47 para 75 anos entre 1900 e 1988 e que caso essa tendência se continuasse a verificar atingiria a marca dos 100 anos em 2065. A proposta de um novo método de previsão de mortalidade era, portanto, necessária e imperativa para se conseguir conhecer melhor as implicações que terá a mortalidade no futuro ao nível do planeamento de políticas públicas e fiscais.

Em Setembro de 1992, os autores acima mencionados publicam o artigo “*Modeling and Forecasting U. S. Mortality*”, onde apresentam um novo modelo para extrapolação das tendências e dos padrões etários da mortalidade, que ficou conhecido por modelo *Lee-Carter*. Desde então que as aplicações do modelo são várias e em diversos países, nomeadamente *Lee e Nault* (Canadá, 1993), *Lee e Rofman* (Chile, 1994), *Wilmoth* (Japão, 1996), *Fígoli* (Brasil, 1998), *Coelho* (Portugal, 2001), *Booth, Maindonald e Smith* (Austrália, 2001) ou *Tuljapurkar, Li e Boe* (G7, 2000). Tornou-se no método mais consensual e utilizado por todo o Mundo originando diversas variantes para melhor explicar ou melhor se ajustar às necessidades específicas de cada país ou investigador.

Lee e Miller (2000) fazem um balanço do comportamento do modelo *Lee-Carter*, avaliando também as considerações que suscitou ao longo dos quase 10 anos de existência e revendo os resultados obtidos em relação a projecções da mortalidade. Entre as conclusões destacam-se a boa performance das previsões de esperança de vida e padrões etários de mortalidade para os períodos desde a publicação do modelo; o comportamento quase linear da mortalidade; a verificação da hipótese base do modelo de que o vector de coeficiente b_x se mantém fixo no tempo; destaque também para as condições iniciais de previsão do modelo, ao grau de incerteza associado ao modelo e à adequação dos intervalos de probabilidade para a previsão.

Tuljapurkar, Li e Boe (2000) analisaram a mortalidade ao longo de cinco décadas nos sete países mais ricos do mundo (Canadá, França, Alemanha, Itália, Japão, Estados Unidos da América e Reino Unido), concluindo que em todos estes países a mortalidade a cada idade diminuiu exponencialmente a um ritmo praticamente constante (“padrão universal” de declínio da mortalidade).

Booth, Maindonald e Smith (2001) aplicaram o modelo *Lee-Carter* aos dados da população australiana na procura de melhorar a metodologia das projecções de mortalidade na Austrália. Mais especificamente, aumentaram a aplicabilidade do modelo *Lee-Carter*. Em 2002, os mesmos autores apresentaram modificações ao modelo *Lee-Carter*, concluindo que as mesmas aumentaram a esperança de vida quando comparada com o modelo original, assim como reduziram para metade o erro de projecção para os dados australianos (*Booth, Maindonald e Smith, 2002*).

Em relação a Portugal, *Coelho* (2001) aplicou o método a dados de mortalidade do país e comparou-os com valores observados e com previsões até 2049 apresentados pelo INE. As conclusões retiradas são de que de um modo geral o comportamento do modelo *Lee-Carter* é bastante razoável, contudo deixando a nota de que é necessário aprofundar o estudo de alguns aspectos do modelo, nomeadamente o comportamento da mortalidade acima dos 85 anos. As últimas projecções de população residente publicadas pelo INE (2003) para o horizonte de 2050 utilizam o modelo *Lee-Carter* na previsão de mortalidade (*Bravo, 2007*).

A base de dados para a realização desta dissertação é a da *Human Mortality Database* (<http://www.mortality.org>) nos dados referentes a Portugal, de referir que estes dados são oficiais visto estarem a ser fornecidos pelo Instituto Nacional de

Estatística. Também foram retirados alguns dados do próprio sítio da internet do INE que não estariam disponíveis no HMD.

Na primeira secção de análise procedeu-se à comparação das projecções dos modelos com os dados observados já existentes para comprovar a habilidade dos modelos para projectar, através de erros associados às projecções e análise dos intervalos de predição calculados. Posteriormente será analisado o ajustamento de cada modelo também através dos erros associados, assim como de grandezas estatísticas, como o erro quadrático médio. De seguida, compara-se as projecções de cada modelo na procura de diferenças nas taxas de mortalidade, assim como noutras medidas, como na esperança de vida e no número de óbitos. Por fim, realiza-se uma análise de sensibilidade com o intuito de observar o impacto dos erros calculados nas projecções finais, de modo a aferir a qualidade das mesmas. Termina-se com as conclusões gerais do trabalho.

2. Metodologias

Existem bastantes variantes que são em tudo similares ao *Lee Carter* com algumas *nuances*. Uma das mais semelhantes é o modelo *Lee-Carter* sem ajustamento de k_t (LCnone), outras como o método apresentado por *Tuljapurkar, Li e Boe* (2000) que para além de não ajustar o índice do nível geral de mortalidade, restringe o período de ajustamento com um limite inferior no ano de 1950, ou o modelo *Booth-Maindonald-Smith* que determina o critério de linearidade para o índice do nível geral de mortalidade, assim como muitos outros, modelo *Lee-Miller* (2001) ou modelo *Hyndman-Ullah* (2007).

Este trabalho propõe-se comparar o modelo original de *Lee-Carter* com o modelo *Booth-Maindonald-Smith* através da sua aplicação a dados existentes no *Human Mortality Database* para Portugal.

2.1. Modelo *Lee-Carter* (MLC)

É um modelo extrapolativo, ou seja, projecta para o futuro as tendências dos dados históricos e os padrões etários da mortalidade. A partir da informação histórica da mortalidade o modelo estima um índice temporal do nível geral de mortalidade, o qual é modelizado através de uma série temporal (modelos ARIMA) e posteriormente projectado. As projecções para as taxas de mortalidade e esperança de vida são obtidas a partir da projecção estimada do índice temporal da mortalidade.

Como todos os modelos existem vantagens e desvantagens a associar. Uma das principais vantagens destacadas pelos diversos autores é a forma como o modelo combina um modelo demográfico e parcimonioso com um modelo de séries temporais, obtendo assim intervalos probabilísticos para as projecções. Outra é a possibilidade de incorporar no modelo um histórico relativamente longo ou também o facto de permitir que as taxas de mortalidade decresçam de forma exponencial ao longo do tempo, não sendo necessário estabelecer um limite superior arbitrário à esperança de vida. No lado das desvantagens há que relembrar que sendo um modelo extrapolativo partilha os problemas dos modelos semelhantes. A estrutura e evolução verificadas nos dados colocados no modelo podem não se verificar no futuro, podendo existir alterações profundas ao nível estrutural, demográfico ou social que o modelo não contemplará,

como possíveis avanços na medicina, mudanças de contexto socioeconómicas, alterações de estilo de vida ou o aparecimento de novas doenças.

A estrutura do modelo proposto por *Lee e Carter* em 1992 é dada pela função log-bilinear:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{x,t}$$

Onde, $m_{x,t}$ corresponde às taxas de mortalidade observadas na idade x e ano t ; a_x é o padrão etário das taxas de mortalidade logarítmicas ao longo dos anos, ou seja descreve o nível médio das taxas de mortalidade no tempo por idade; k_t é o índice de nível geral da mortalidade no ano t ou seja, descreve as tendências temporais do nível de mortalidade; b_x reflecte as mudanças nas taxas de mortalidade logarítmicas em cada idade x em resposta a alterações no nível geral da mortalidade, é o padrão de desvios de perfil da idade conforme variação do índice de nível global k_t ; $\varepsilon_{x,t}$ é o resíduo na idade x e no ano t , erro aleatório com média zero e variância σ_ε^2 .

Visto o modelo *Lee-Carter* ser sobre-parametrizado, para assegurar a fiabilidade do modelo, *Lee e Carter* (1992) impuseram duas restrições:

$$\sum_{t=1}^n k_t = 0, \quad \sum_{x=x_1}^{x_p} b_x = 1$$

Onde n é o número de anos e p o número de idades existentes no conjunto de dados observados. Isto como forma de garantir a obtenção de solução única para as estimativas dos parâmetros.

O processo de ajustamento do modelo *Lee-Carter* é feito em duas etapas: estimação dos parâmetros e reajustamento do k_t para que o número total de mortes ajustado coincida com o número total de óbitos efectivamente observado.

A estimação dos parâmetros do modelo é efectuada através do método dos mínimos quadrados, ou seja,

$$(\hat{a}_x, \hat{b}_x, \hat{k}_t) = \arg \min \left\{ \sum_{x=x_1}^{x_p} \sum_{t=1}^n (\ln(m_{x,t}) - a_x - b_x k_t)^2 \right\},$$

cuja solução única, atendendo às restrições acima mencionadas, é obtida através da aplicação do método de decomposição em valores singulares (SVD, *Singular Value*

Decomposition) (Bravo, 2007). Método aplicado à matriz dos logaritmos das taxas de mortalidade depois de lhe ter sido subtraída a média temporal do logaritmo das taxas de mortalidade específicas por idade.

Visto que as estimativas dos parâmetros são geradas com base na matriz de logaritmos das taxas de mortalidade e não na matriz de taxas de mortalidade, o número de óbitos obtidos pela aplicação das taxas de mortalidade estimadas à distribuição por idade da população, de um modo geral, não coincide exactamente com o número de óbitos observados. Assim sendo, *Lee e Carter* (1992) propuseram uma nova etapa na estimação de k_t , de forma a que produza exactamente o número total observados para o ano em causa. Assim, a tarefa seria calcular k_t dado que,

$$D_t = \sum_x (N_{x,t} e^{(a_x + b_x k_t)})$$

onde D_t representa o número total de óbitos observados no ano t e $N_{x,t}$ corresponde à população na idade x no ano t .

Para tal reajuste *Lee e Carter* (1992) utilizaram a metodologia de *Box-Jenkins*, seleccionando um modelo ARIMA, ou seja, um passeio aleatório com um termo constante,

$$k_t = k_{t-1} + c + \varepsilon_t$$

Este será um dos aspectos mais importantes do modelo *Lee-Carter*, a possibilidade da obtenção de intervalos de probabilidade para a projecção de k_t , conjuntamente com a previsão pontual do índice de nível de mortalidade para um dado horizonte temporal.

2.2. Modelo *Booth-Maindonald-Smith* (MBMS)

Como uma variante do modelo *Lee Carter*, o método de *Booth-Maindonald-Smith* difere do *Lee-Carter* em três pontos:

- O período de ajustamento é determinado com base num critério estatístico de bondade do ajustamento, com o pressuposto que o resultado da componente principal k_t é linear;

- O ajustamento de k_t passa a ser feito para uma distribuição etária das mortes em vez de para o número total de mortes;
- As taxas “jump-off” são as taxas ajustadas sobre esta metodologia ajustada.

Booth et al. (2002) ajustaram o modelo *Lee-Carter* para dados da Austrália entre 1907-1999 e concluíram que o “padrão universal” (*Tuljapurkar et al.*, 2000) para um declínio da mortalidade constante representado por k_t linear não se manteve ao longo do período de ajustamento. Houve também dificuldades na validação do pressuposto de um b_x constante. Assumindo a linearidade do k_t logo de início, a variante Booth-Maindonald-Smith procura maximizar o modelo geral restringindo o período de ajustamento para maximizar o ajustamento ao pressuposto de linearidade, o que também leva a um melhor validação do pressuposto de b_x constante.

2.3. Erros associados ao ajustamento dos modelos

Uma forma de se comparar o ajustamento dos modelos estudados numa forma quantitativa, e não apenas gráfica, é o cálculo de erros de ajustamento. Optou-se, para verificar a qualidade dos valores projectados ou ajustados, pelo cálculo do Erro Percentual Absoluto Médio (EPAM) e também pelo cálculo do Erro Percentual Algébrico Médio (EPALM). O primeiro porque permite observar qual o erro absoluto, em percentagem, do erro de ajustamento que o modelo cometeu. O segundo porque permite verificar se existe subvalorização ou sobrevalorização por parte do ajustamento de cada modelo.

$$EPAM = \frac{\sum \frac{|\hat{m}_{x,t} - m_{x,t}|}{m_{x,t}}}{n} \times 100$$

$$EPALM = \frac{\sum \frac{(\hat{m}_{x,t} - m_{x,t})}{m_{x,t}}}{n} \times 100$$

onde, $\hat{m}_{x,t}$ corresponde aos valores projectados, $m_{x,t}$ aos valores observados de taxas de mortalidade e n o número de erros do numerador.

De salientar que também se efectua o cálculo do Erro Percentual Absoluto (EPA) e do Erro Percentual Algébrico (EPAL). Tratam-se de dois erros com um cálculo similar aos anteriores com a diferença de, nos casos tratados, ser calculado para cada idade. Este cálculo serve de base para uma análise dos erros em função da idade, tentando observar em que idades os modelos ajustam pior ou melhor. Podendo ou não haver um padrão de ajustamento nesse sentido.

2.4. Erro Quadrático Médio

O erro quadrático médio pode ser usado como uma medida de erro de previsão. Este erro é determinado somando os erros de previsão ao quadrado e dividindo pelo número de erros usados no cálculo.

Assim, é necessário calcular os seguintes valores.

$$SQ = \sum (\hat{m}_{x,t} - m_{x,t})^2$$

Onde,

$m_{x,t} \rightarrow$ valores observados

$\hat{m}_{x,t} \rightarrow$ valores ajustados

Em seguida, basta calcular o erro quadrático médio através da seguinte expressão.

$$EQM = \frac{SQ}{n}$$

Onde,

$n \rightarrow$ número de erros utilizados no numerador

Tendo em conta a fórmula de cálculo do EQM quanto mais o valor se aproximar de zero melhor. No entanto, de realçar que é calculado através de taxas de mortalidade que são valores muito pequenos, há que ter portanto cuidado nas conclusões que se retiram dos valores baixos de EQM.

3. Software e Linguagem R

R é uma linguagem de programação e um ambiente de desenvolvimento integrado para cálculos estatísticos e gráficos. Foi originalmente criada por *Ross Ihaka* e por *Robert Gentleman* no Departamento de Estatística da Universidade de Auckland, Nova Zelândia e tem como principais características o facto de o seu desenvolvimento ser feito à escala mundial com vários colaboradores em todo o mundo e de ser um programa completamente gratuito. O nome da linguagem provém não só dos seus criadores mas também de um jogo figurado com a linguagem S, linguagem essa bastante similar à R (alguns códigos em S funcionam sem alterações em R).

O *software* R tem como principal função a manipulação de dados e modelação estatística, assim como cálculo e representação gráfica.

Neste projecto o principal objectivo no uso da linguagem e *software* R é a modelação estatística. Visto ser uma ferramenta direccionada a dar resposta a este tipo de problemas é o *software* ideal para este trabalho. Nesse sentido fez-se uma pesquisa sobre os pacotes existentes em R sobre demografia, mortalidade, fecundidade e migrações, chegando à conclusão que para os modelos que o projecto pretende utilizar já estão acessíveis *online* e gratuitamente.

O pacote mais importante para este projecto é o pacote criado por *Rob J. Hyndman*, com a colaboração de *Heather Booth*, *Leonie Tickle* e *John Maindonald*, denominado de “*demography package*” com o título “*Forecasting mortality, fertility, migration and population data*”. Este pacote inclui funções para análise demográfica incluindo tábuas, modelos *Lee-Carter*, análise de dados funcional de taxas de mortalidade, fecundidade e migração e também projecção estocástica da população. Publicado inicialmente a 18 de Outubro de 2011 pode ser baixado através da seguinte página,

<http://robjhyndman.com/software/demography/>

Para uma melhor compreensão de como grande parte das funcionalidades do pacote é imperativo que também se baixe o manual de ajuda do pacote, podendo fazê-lo através do seguinte *link*,

<http://cran.r-project.org/web/packages/demography/demography.pdf>

Tendo o pacote e manual ao dispor é apenas uma questão de exploração até que se possa compreender e utilizar da forma mais adequada o *software* e a modelação necessário ao projecto.

Ao longo deste trabalho tentar-se-á disponibilizar toda a informação necessária para realizar os devidos cálculos, modelação ou gráficos através do *software* R, desde o código à interpretação do mesmo. Não entrando, obviamente, numa excessiva centralização do projecto no *software* utilizado, visto não ser esse o objectivo.

4. Projecções vs Valores Observados

A primeira abordagem realizada no âmbito deste projecto aos modelos escolhidos, foi a comparação de projecções obtidas para as taxas de mortalidade através da modelação dos dados para os anos entre 2001 e 2009 com os valores observados dessas mesmas taxas de mortalidade. Pretende-se assim perceber se as projecções realizadas pelos dois modelos são de certa forma realistas e se não se obtém um desvio muito exagerado da realidade verificada, para além de se fazer a comparação entre os dois modelos para observar qual o que menos erro produz nas suas projecções.

Assim sendo, procedeu-se à construção de modelos com um histórico parcialmente semelhante e aumentando a informação no mesmo. Quer isto dizer que as projecções realizadas para 2001 terão menos informação do que as realizadas em 2009, visto que o histórico no modelo será maior em 2009 do que em 2001. O modelo para as projecções de 2001 terá como histórico de dados os valores compreendidos entre 1981 e 2000, o modelo para as projecções de 2002 terá o mesmo histórico mas acrescentando os valores observados em 2001 e assim sucessivamente, até ao modelo construído para as projecções de 2009 que terá como base de dados histórica os anos entre 1981 e 2008. Assim, aumenta-se a fiabilidade de cada um dos modelos, visto estar a acrescentar informação real não se perdendo na comparação de modelos já que apenas se vai comparar projecções e modelos do mesmo ano e não de anos passados ou futuros.

Através do *software R* a construção dos modelos *Lee-Carter* e *Booth-Maindonald-Smith* é bastante simples, assim utilizando o exemplo do comando 1 criaram-se os primeiros modelos.

```
modeloLC1<-lca (portugal, series="total", years=anos81.00)
modeloBMS1<-bms (portugal, series="total", years=anos81.00)
```

Comando 1 – Ajustamento dos Modelos Lee-Carter e Booth-Maindonald-Smith

Como se pode verificar através dos códigos, tem-se um modelo para o modelo *Lee-Carter* e outro para o modelo *Booth-Maindonald-Smith*, utiliza-se a base de dados de Portugal para os anos entre 1981 e 2000 e neste caso específico para a série *total* ou seja para o conjunto da população, não diferenciando homens e mulheres. Para os

modelos seguintes foi apenas uma questão de modificar as séries a utilizar assim como os anos de histórico que se pretendiam. Contudo este comando não nos dará as projecções mas sim os ajustamentos aos parâmetros dos modelos em questão. Será necessário outra modelação no *software* para ser possível retirar as projecções para cada série de dados e para cada ano pretendido. O comando 2 é um exemplo com códigos que constroem a modelação para as projecções.

```
foreLC1 <- forecast(modeloLC1, h=1, level=90)
foreBMS1 <- forecast(modeloBMS1, h=1, level=90)
```

Comando 2 – Ajustamento do Modelo de Projecções para o LC e BMS

Utilizando a função *forecast* pode-se construir através dos modelos já existentes as projecções pretendidas. De realçar que para além de se identificar o modelo subjacente às projecções, também há que referir o número de anos a projectar (parâmetro *h* no código) e também o nível de confiança do intervalo criado para as projecções. Isto porque quando se retiram os valores das projecções para cada ano e para cada idade, as mesmas vêm com um intervalo de predição associado com o nível referido no comando.

Construídos os modelos para ser possível projectar as taxas de mortalidade para os anos entre 2001 e 2009, resta colocá-los de modo a serem tratados e analisados. Para retirar esses valores de projecções do *software* é necessário escrever, por exemplo, o comando 3.

```
write.table (foreLC1$rate, file="LC_2001.txt", sep=" ")
```

Comando 3 – Obtenção das projecções

Neste caso optou-se por retirar os dados para um ficheiro *.txt* de modo a ser mais fácil a conversão, assim como a passagem para folhas de cálculo como o *Microsoft Excel*. Há que referir que este comando extrai os valores das taxas de mortalidade projectadas (*rate*), assim como os intervalos de predição, para o ficheiro que se denominou como *LC_2001* e tendo como separadores caracteres vazios ou seja, espaços.

Com os valores projectados por cada modelo para cada ano e cada género, assim como as projecções para as taxas de mortalidade da população total, podem-se efectuar comparações com os valores de taxas de mortalidade observados na realidade e já registados pelas instituições competentes.

4.1. População Total

4.1.1. Projecções

Tendo sempre como base os valores projectados extraídos da modelação e os valores observados, chega-se aos gráficos 1, 2 e 3 para as taxas de mortalidade, respectivos anos, géneros e modelos:

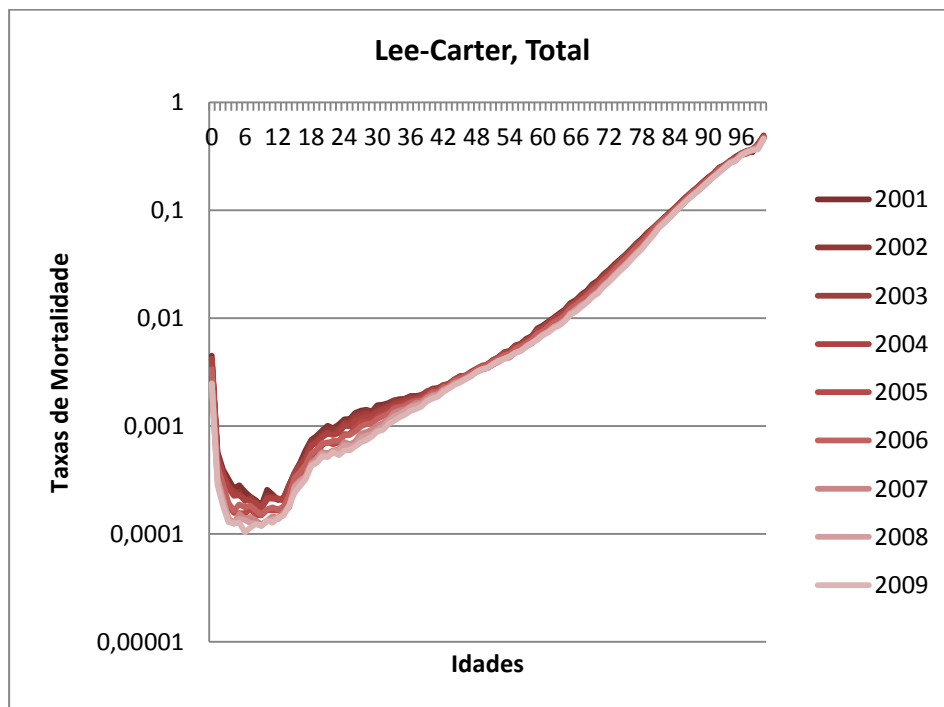


Gráfico 1 - Curvas de Mortalidade Projectadas pelos MLC para a população total.

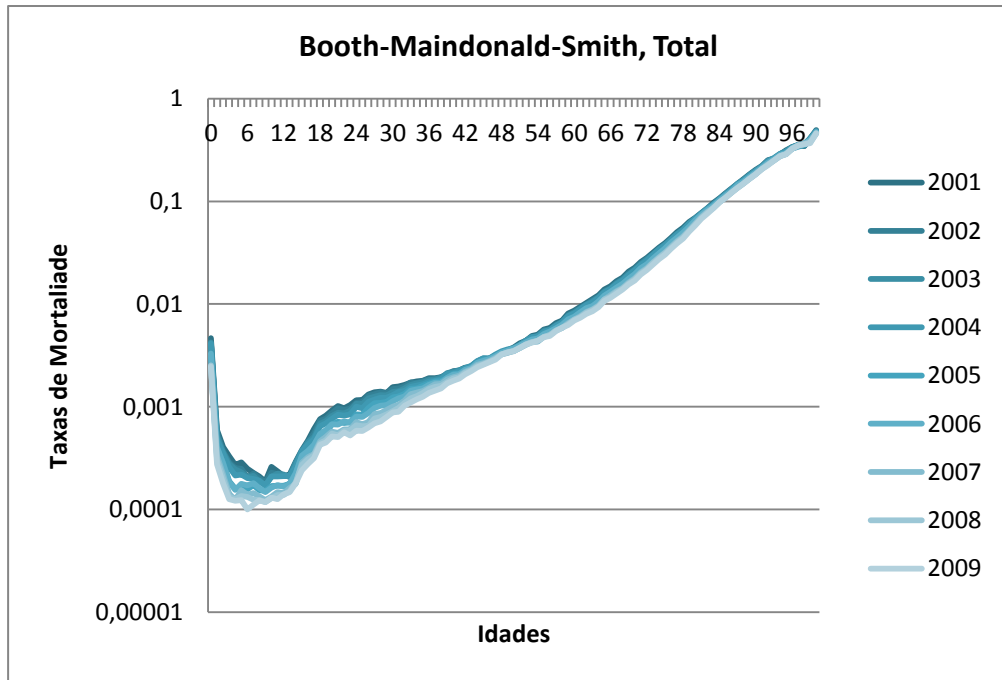


Gráfico 2 - Curvas de Mortalidade Projectadas pelos MBMS para a população total.

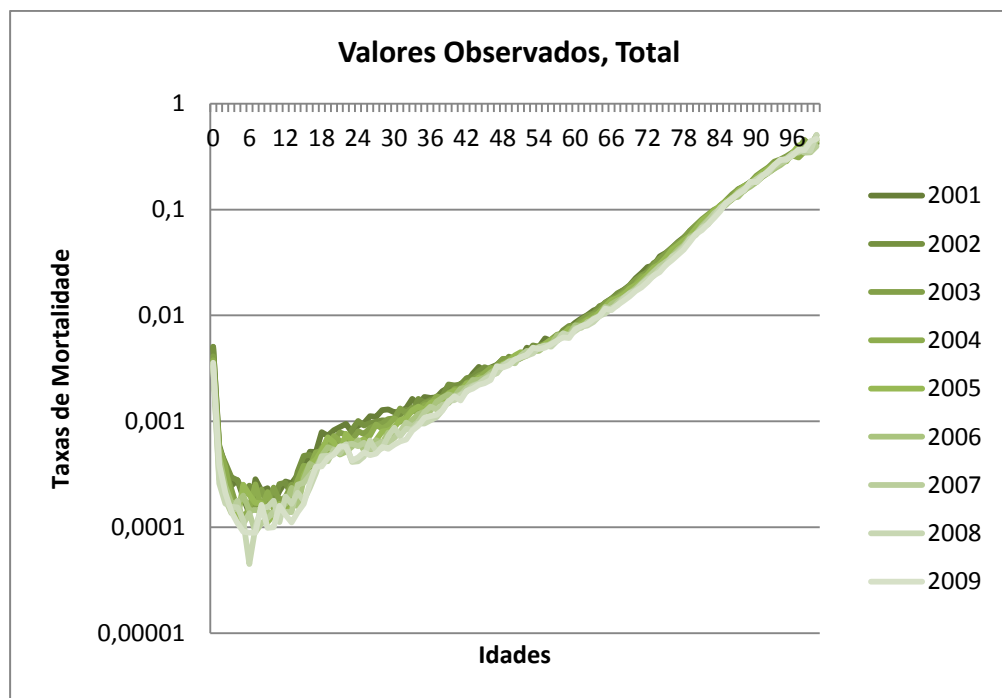


Gráfico 3 - Curvas de Mortalidade Observadas para a população total.

A primeira conclusão que se pode retirar observando gráficos 1, 2 e 3, e aquela que suscita também alguma preocupação é o facto de que os dois modelos entre si não

apresentarem diferenças significativas (gráfico 1 e 2), o que para efeitos de comparação dificulta o processo de escolha final.

Numa segunda análise superficial e apenas com a observação dos gráficos chega-se também à conclusão que o modelo *Lee-Carter* e a sua variante, o modelo *Booth-Maindonald-Smith*, têm um efeito de alisamento nas projecções calculadas, isto é, os gráficos das projecções de cada um dos modelos não apresenta tantos “picos” (gráfico 1 e 2) como os verificados nos valores observados de taxas de mortalidade (gráfico 3). Em termos demográficos acaba por ser o mais desejado, não havendo oscilações excessivas nas taxas de mortalidade quando se passa de idade para idade.

De um modo geral, e mais uma vez salientando o facto desta análise inicial ser apenas uma análise gráfica qualitativa e não quantitativa, para o total da população, a que não faz distinção entre géneros, os dois modelos apresentam uma evolução bastante próxima da real não havendo nenhum indício que haja uma discrepância ao nível de evolução etária e anual, em comparação com os valores observados.

4.1.2. Erros Associados

Para uma melhor análise de comparação e tendo em vista agora uma análise quantitativa e não apenas gráfica, existem vários métodos para o efeito. Uma das possíveis maneiras de verificar a qualidade de um valor projectado ou ajustado é o cálculo do Erro Percentual Absoluto Médio (EPAM) assim como o cálculo do Erro Percentual Algébrico Médio (EPALM). Assim, e observando exclusivamente para o Erro Percentual Absoluto (EPA), pode-se construir um gráfico com esse mesmo erro ao longo das faixas etárias e dos anos de projecção, para os dois modelos em questão.

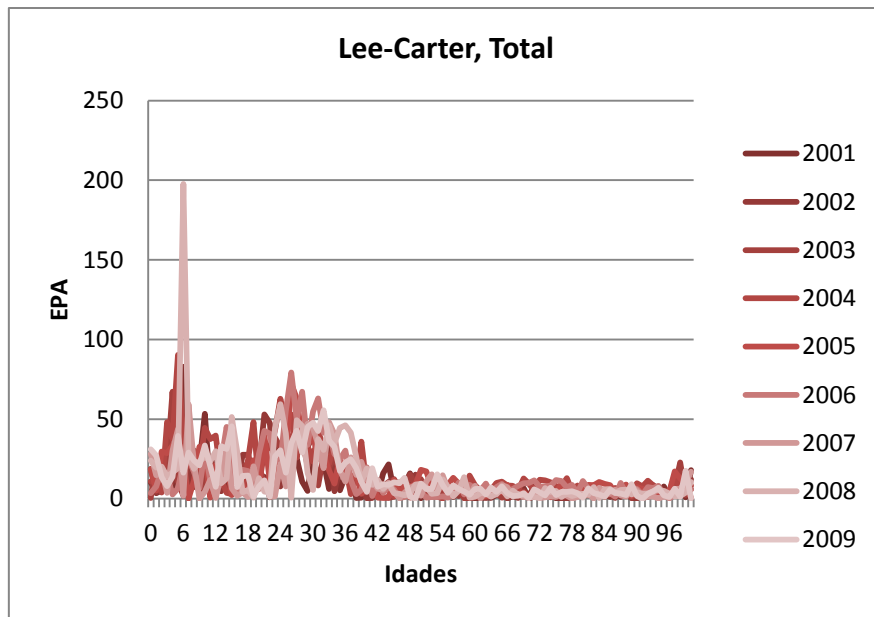


Gráfico 4 – EPA das diversas projecções dos MLC para a população total.

Como se pode observar, pelo gráfico 4, existe uma tendência para um erro superior nas idades mais jovens, até 35 anos, decrescendo substancialmente nas idades superior a 45 anos. Também se regista um pequeno aumento nas idades superiores a 95 anos face às idades anteriores (dos 45 aos 95 anos). E com a excepção para a taxa de mortalidade para os 6 anos do ano de 2008, não existem grandes oscilações entre idades, havendo um EPA a rondar os 50% nas idades até os 35 anos, mas baixando drasticamente depois disso.

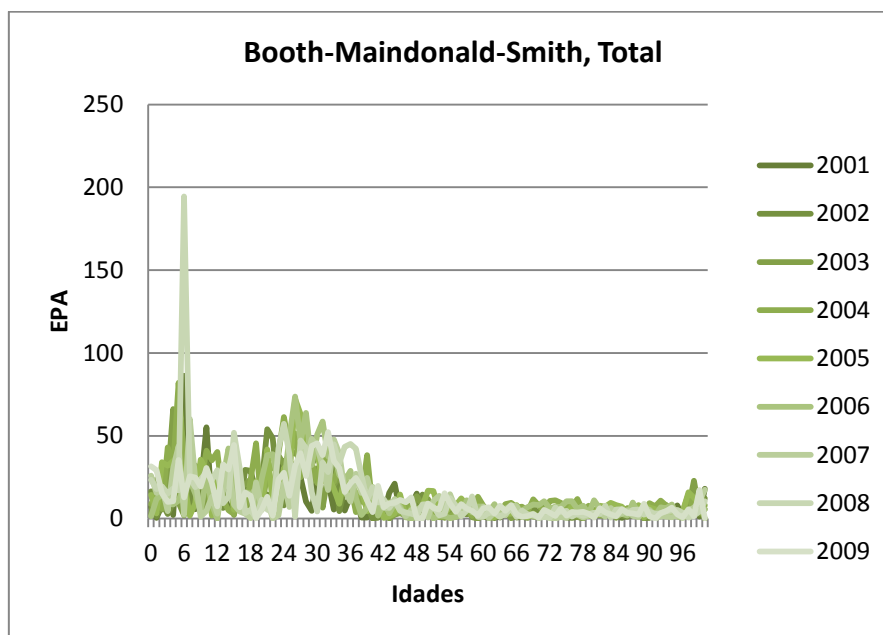


Gráfico 5 – EPA das projecções dos MBMS para a população total.

Através do gráfico 5, pode-se observar que o comportamento do erro é bastante similar no modelo *Booth-Maindonald-Smith*, mais uma vez dificultando a decisão, mas simultaneamente mostrando que as diferenças verificadas nestes dados para estes dois modelos não são grandes, muito pelo contrário. O andamento das curvas ao longo das faixas etárias para a variante é idêntica ao do modelo *Lee-Carter*, incluindo o anormal “pico” no ano de 2008 com a idade de 6 anos. Procurando perceber o que poderia estar associado nesse erro chega-se à conclusão que existe uma taxa de mortalidade observada para essa idade nesse ano muito inferior às idades adjacentes assim como aos anos anteriores e posteriores. Não estando portanto em causa a qualidade da projecção.

Mais uma vez o principal adversário das projecções construídas está nas idades mais jovens, dos 0 aos 35 anos. Obtendo-se valores muito próximos dos 50%, baixando drasticamente a partir dessa idade.

Posto isto, e para uma melhor comparação analítica dos dois modelos, construiu-se um gráfico com o Erro Percentual Absoluto Médio (EPAM) de cada ano projectado em cada modelo.

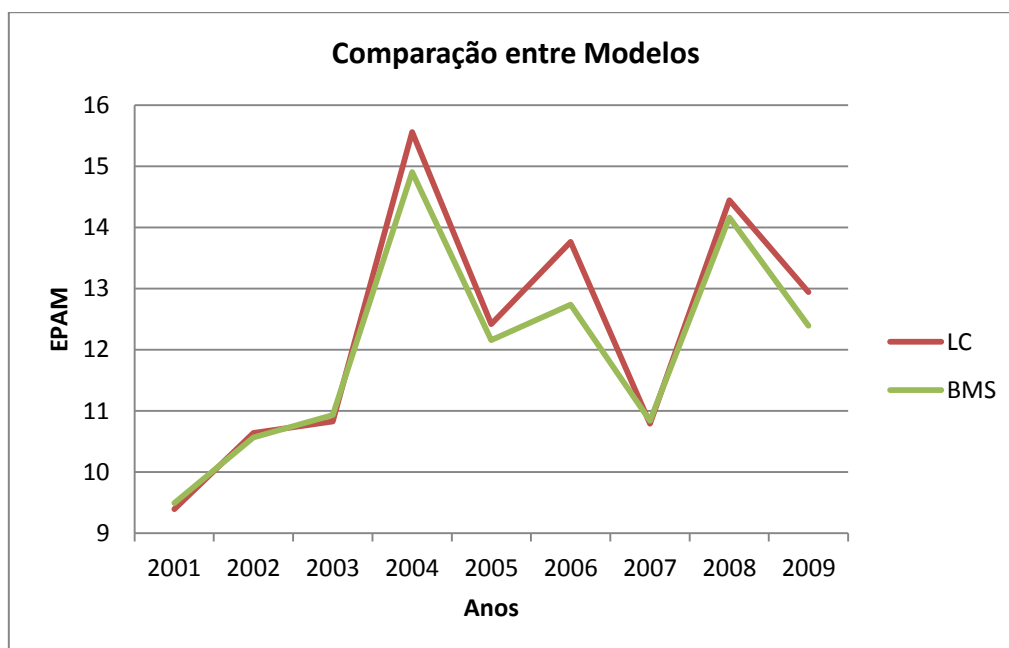


Gráfico 6 – EPAM, comparação entre MLC e MBMS para a população total.

Como se pode verificar pelo gráfico 6 a diferença ao longo dos anos projectados entre cada modelo é bastante pequena. O EPAM situa-se ao longo dos vários anos entre os 8 e os 16%, tendo em conta os valores de EPA anteriormente visto e estes valores pode-se afirmar que para uma projecção de taxas de mortalidade não parece ser uma má

projecção. Destaca-se também o facto do modelo *Lee-Carter* ter valores ligeiramente superiores nalguns anos em comparação com o modelo *Booth-Maindonald-Smith*, no entanto há que sublinhar a proximidade existente entre os dois modelos.

Para analisar mais pormenorizadamente o erro associado à projecção utiliza-se o Erro Percentual Algébrico Médio (EPALM), este erro poderá dar informação sobre a sobrevalorização ou subvalorização que a projecção executa sobre a taxa de mortalidade. Utilizando a mesma metodologia gráfica que a usada para o EPAM, as primeiras observações retractam o erro em cada uma das idades e anos.

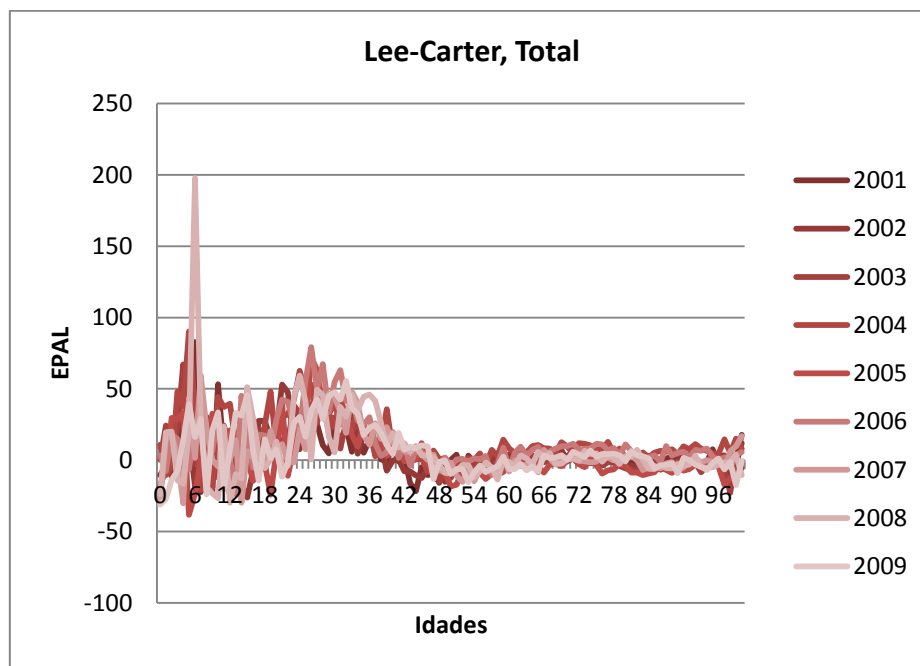


Gráfico 7 – EPAL das projecções dos MLC para a população total.

Numa primeira abordagem ao gráfico 7, pode-se verificar que o comportamento do Erro Percentual Algébrico é em tudo similar aos Erro Percentual Absoluto, com a diferença de haver valores negativos e positivos, podendo assim ser analisada a sobrevalorização ou subvalorização das projecções. De um modo geral pode afirmar-se que nas idades superiores a 40 anos não existe uma constante de sobrevalorização ou subvalorização, mas nas idades mais jovens o mesmo não pode ser afirmado, visto que é perceptível uma constante sobrevalorização (valores positivos) das projecções em relação aos valores das taxas de mortalidade observados.

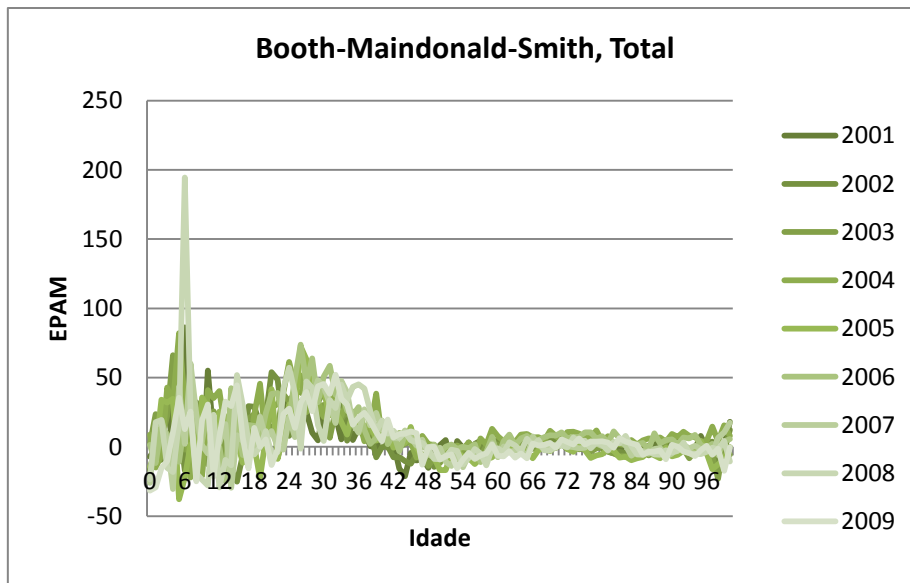


Gráfico 8 – EPAL das projecções dos MBMS para a população total.

De um modo geral, através do gráfico 8, pode-se verificar que o comportamento do Erro Percentual Algébrico no modelo *Booth-Maindonld-Smith* é perfeitamente similar ao verificado no modelo *Lee-Carter*, havendo uma sobrevalorização nas idades mais jovens e em maior escala, e valores em redor do zero a partir dos 35 anos, não havendo a destacar nenhuma sobrevalorização ou subvalorização continuada.

Assim será possível o cálculo do EPALM e respectiva comparação gráfica e analítica, através do gráfico 9.

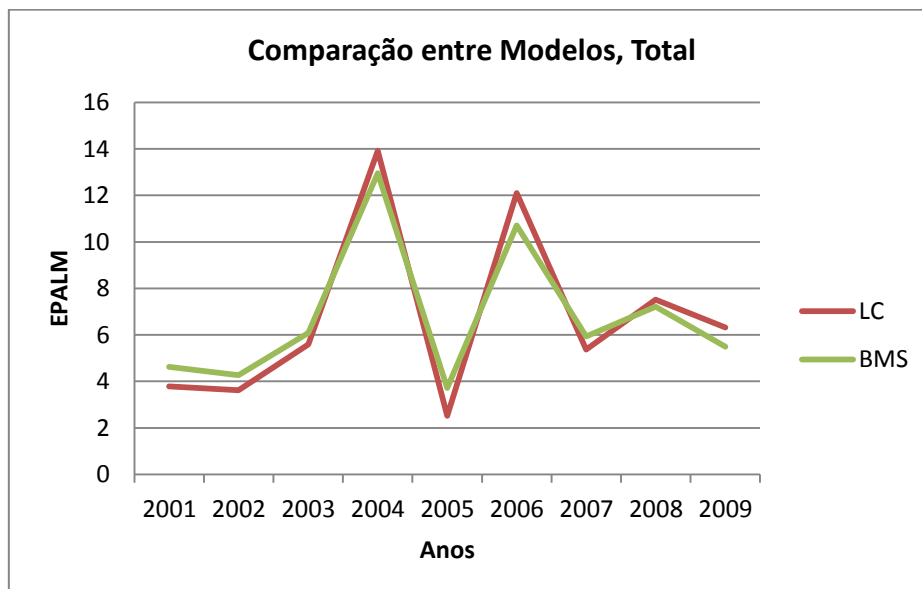


Gráfico 9 – EPALM, comparação entre MLC e MBMS para a população total.

O primeiro destaque que se pode dar é a diminuição do erro quando comparado com o EPAM, obviamente que derivado à sua forma de cálculo seria imperativo esse

facto acontecer. Os valores do erro, desta feita, situam-se entre os 2 e os 14%. Verifica-se também que não existe nenhum domínio ao nível de modelos, não existindo nenhum modelo com um erro sistematicamente superior ao outro, levando a uma conclusão semelhante à registada com o cálculo do EPAM. Ou seja, os modelos não apresentam diferenças muito significativas.

4.1.3. Projecções dentro ou fora do intervalo de predição

Finalmente, e visto que existem intervalos de predição construídos a 90%, é possível verificar se os valores observados nas taxas de mortalidade se situam dentro ou fora desse mesmo intervalo de predição calculado para as projecções. A análise consiste em verificar se para cada idade, a taxa de mortalidade observada se situa dentro do intervalo de predição do modelo. Verifica-se quantas vezes isso sucede em cada ano e em cada modelo. A tabela 1 resume essa análise:

	Média de Percentagem de Valores Dentro do Intervalo
Lee-Carter	45,32
Booth-Maindonald-Smith	37,07

Tabela 1 – Média de Percentagem de Valores Dentro do Intervalo nos dois modelos para a população total.

Assim pode-se verificar que existe um maior número de valores dentro do intervalo de predição no caso do modelo *Lee-Carter* do que no modelo *Booth-Maindonald-Smith*.

Há que referir que os valores encontrados estão longe de ser os melhores, visto que nem a 50% dos valores algum dos modelos chega. Facto que já seria de esperar aquando da análise dos EPAM e EPALM, havendo valores próximos de 50% de erro, seria natural que os valores não se encaixassem nos intervalos de predição calculados para as projecções. Um dos principais motivos para se chegar a estes valores prende-se

pelo nível de confiança exigido, 90%, talvez aumentando o nível de confiança para 95% ou mesmo 99% já se estaria a falar doutro tipo de resultados e conclusões.

Mesmo estando perante resultados parcos de qualidade, ao nível de percentagem de valores dentro do intervalo de predição, há que salientar que o modelo *Lee-Carter* destacou-se em relação ao modelo *Booth-Maindonald-Smith* para os dados estudados, revelando intervalos de predição mais fiáveis.

Uma das principais análises que é necessária fazer neste projecto é o comportamento dos modelos nos diferentes géneros. Ou seja, verificado o comportamento, fiabilidade e analisadas as projecções para o total da população, sem distinção, é necessário fazer o mesmo processo mas neste caso para os géneros em separado. A ideia passa por verificar se existe alguma alteração comportamental de algum dos modelos quando os dados inseridos na modelação não são os do total da população, mas sim a população feminina e a população masculina.

A análise da curva de mortalidade é essencial para se verificar o andamento das taxas de mortalidade quer ao longo das idades quer ao longo dos anos projectados. No entanto, os cálculos dos dois erros poderá tornar-se fastidioso para os sexos. Atendendo ao facto verificado com o total da população, do qual se pôde concluir que apenas com o EPALM seria possível verificar não só a amplitude do erro mas também a sobre ou subvalorização dos valores das projecções, resolveu-se apresentar apenas os gráficos e a comparação para o EPALM, assim como o andamento do erro ao longo das idades.

4.2. População Feminina

4.2.1. Projecções

Começando pelas curvas de mortalidade, pode-se construir os mesmos gráficos que foram criados aquando da avaliação sobre os modelos do total da população.

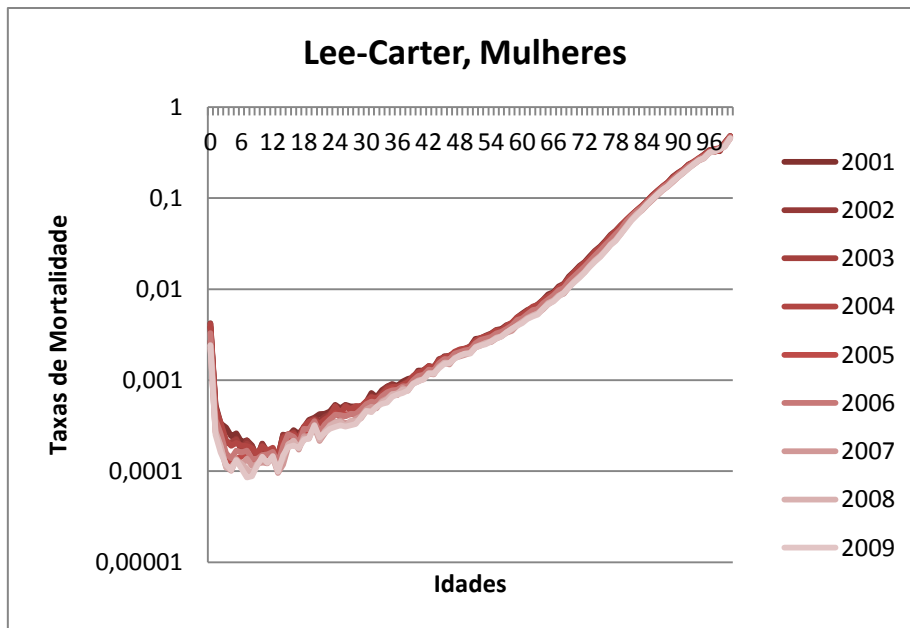


Gráfico 10 – Curvas de Mortalidade Projectadas pelos MLC para a população feminina.

Como se pode comprovar através da comparação entre a curva de mortalidade projectada para a população feminina, pelo modelo *Lee Carter* (gráfico 10), e a curva de mortalidade projectada para a população total, pelo mesmo modelo (gráfico 1), a primeira questão que se destaca é um menor alisamento da curva. O que pode significar uma melhor projecção, visto não ser demais lembrar que os valores observados tendem a ter menor alisamento que os projectados. Por outro lado, há que referir que demograficamente seria vantajoso, as projecções serem as mais alisadas possíveis, dentro dos parâmetros de fiabilidade que pretendemos.

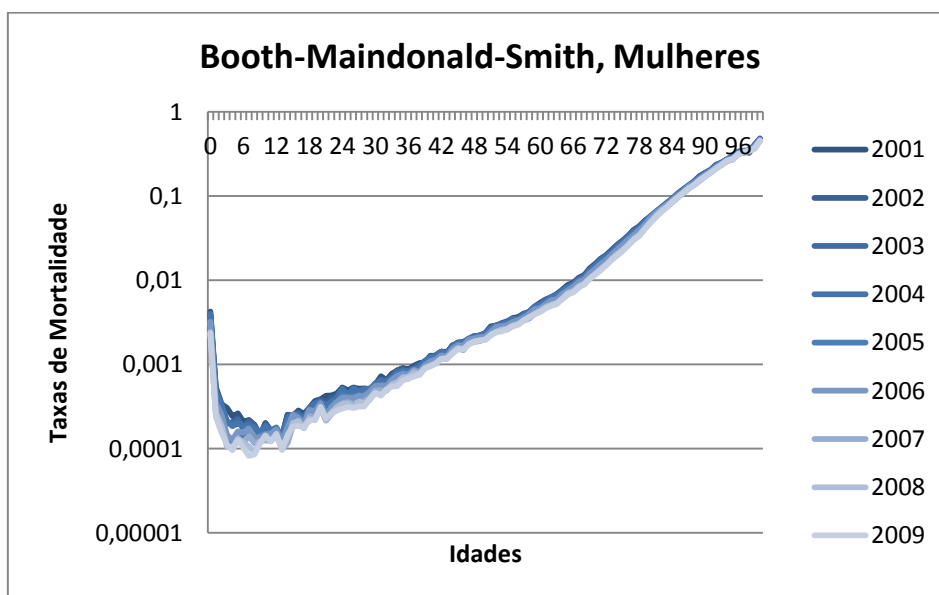


Gráfico 11 – Curvas de Mortalidade Projectadas pelos MBMS para a população feminina.

Observando o gráfico 11, de um modo geral, a curva apresentada pelo modelo *Booth-Maindonald-Smith* é bastante similar à do modelo *Lee-Carter*.

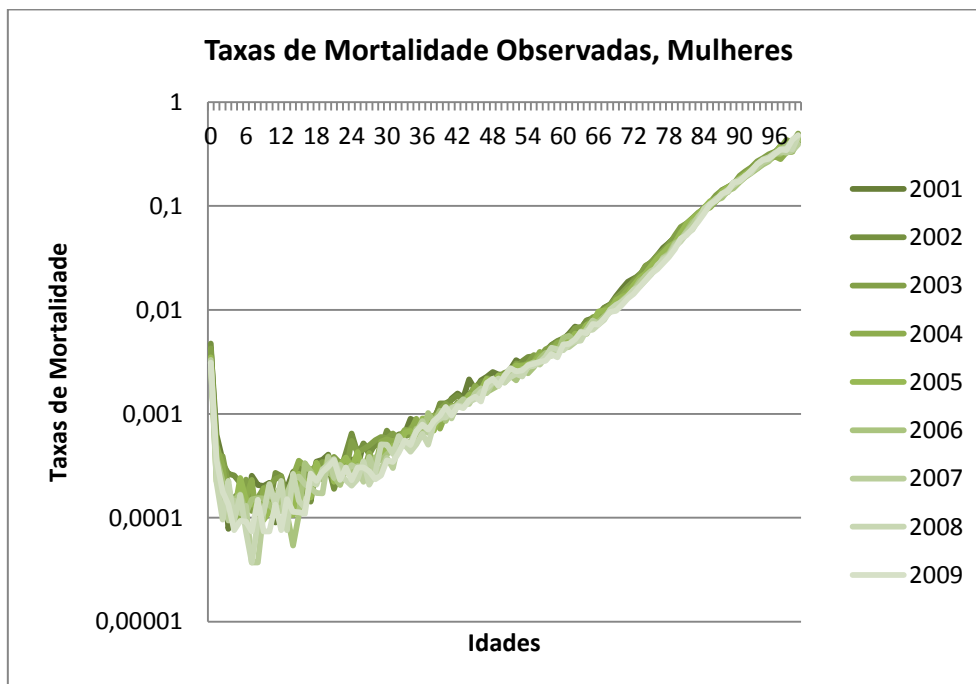


Gráfico 12 – Curvas de Mortalidade Observadas para a população feminina.

A comparação entre as curvas de mortalidade dos modelos construídos e a curva de mortalidade observada para a população feminina (gráfico 12) é bastante parecida com a comparação para a população total. Ou seja, existe um óbvio alisamento nos modelos, “alisando picos” nas taxas de mortalidade e realçando também a melhoria das taxas de mortalidade ao longo dos anos, mais do que os valores observados indicam. Por outras palavras, os modelos construídos projectaram uma diminuição na taxa de mortalidade em quase todas as idades ao longo dos vários anos (a cor no gráfico 10 e 11 vai aclarando de cima para baixo), enquanto no caso das taxas de mortalidade observadas isso não se verifica tão claramente. Não existe portanto uma diferenciação entre anos tão clara como no caso dos dois modelos e respectivas projecções.

4.2.2. Erros Associados

Depois de analisadas as curvas de mortalidade na população feminina para os diferentes cenários (modelos e valores observados), comparam-se os modelos de forma quantitativa através do Erro Percentual Algébrico (EPAL) ao longo das idades e em seguida o Erro Percentual Algébrico Médio (EPALM).

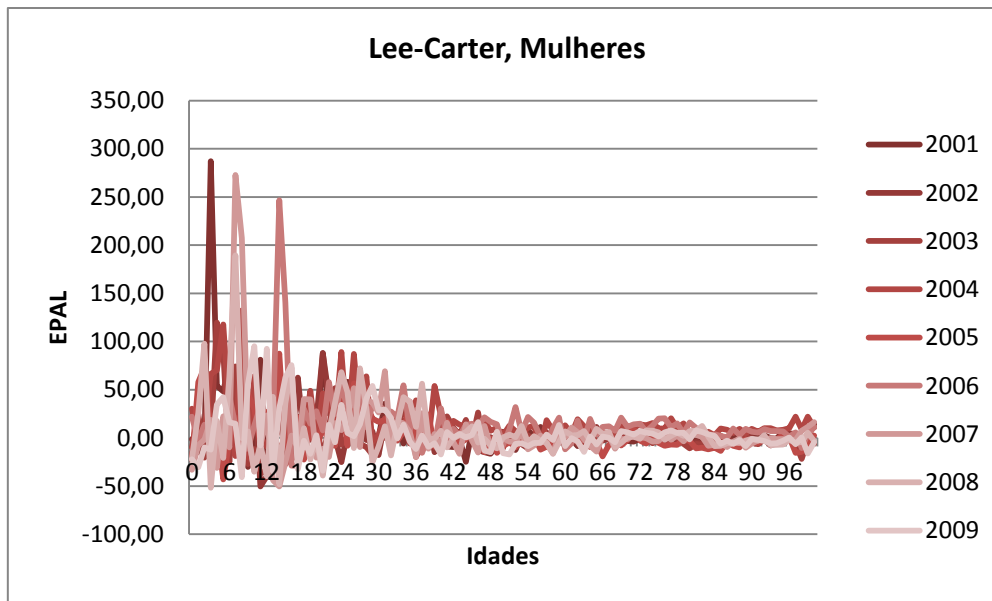


Gráfico 13 – EPAL das projecções dos MLC para a população feminina.

Como se pode verificar através do gráfico 13, para a população feminina existem mais erros bruscos na comparação com os valores de taxas de mortalidade observados, no entanto, e tirando uma ou outra exceção e a existência de mais “picos” no gráfico, o comportamento do erro é semelhante ao comportamento evidenciado para a população total. Ou seja, substancialmente maior erro nas idades mais jovens, até aos 40 anos. E decréscimo acentuado nos 40 a 45 anos, mantendo-se mais ou menos constante para todas as idades seguintes.

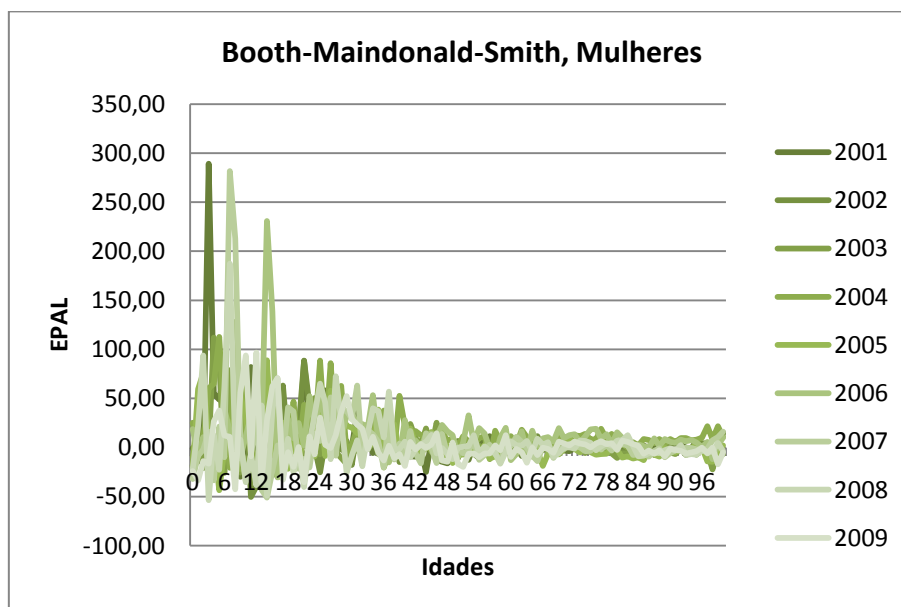


Gráfico 14 – EPAL das projecções dos MBMS para a população feminina.

No caso do modelo *Booth-Maindonald-Smith* (gráfico 14), o comportamento do erro é relativamente similar ao modelo *Lee Carter*.

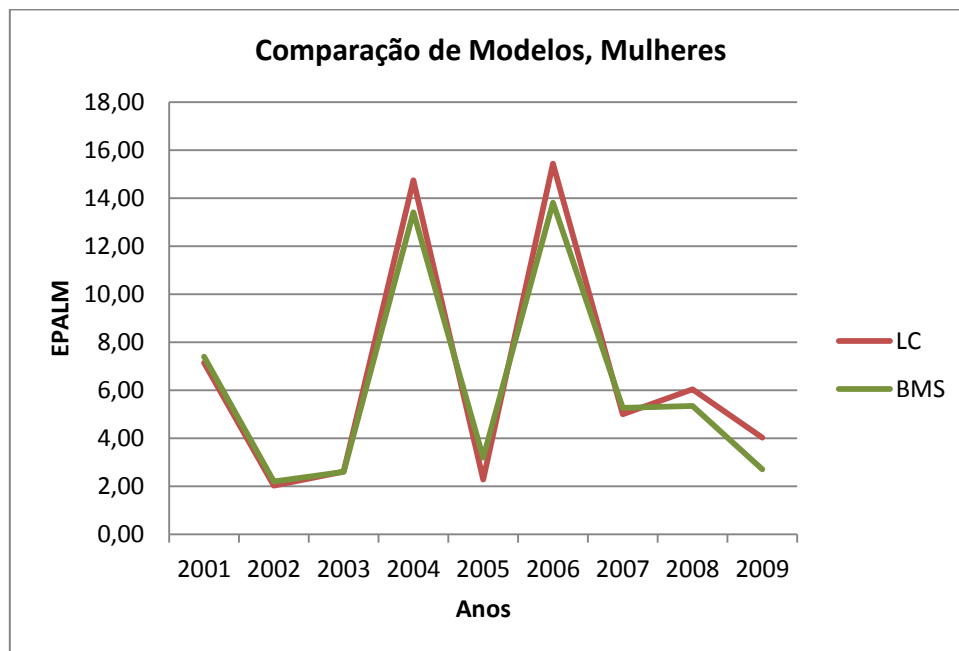


Gráfico 15 – EPALM, comparação entre MLC e MBMS para a população feminina.

Através do gráfico 15 pode-se afirmar que as diferenças entre os dois modelos para a projecção da população feminina são praticamente nulas. Os dois modelos situam os seus erros percentuais algébricos médios entre os 2 e os 16%.

4.2.3. Projecções dentro ou fora do intervalo de predição

Para a análise às projecções para a população feminina estarem completas é necessário verificar se os intervalos de predição se adequam à realidade. Na população total os valores a 90% ficaram bastante aquém do esperado portanto será sempre interessante verificar o que acontece no caso da população feminina para os mesmos 90%. Na tabela 2 resume-se esse estudo.

Média de Percentagem de Valores Dentro do Intervalo	
Lee-Carter	43,67
Booth-Maindonald-Smith	37,18

Tabela 2 – Média de Percentagem de Valores Dentro do Intervalo nos dois modelos para a população feminina.

Os valores dentro do intervalo de predição a 90% mantiveram-se praticamente idênticos quando comparados com o total da população. Mais uma vez refere-se que o facto de se trabalhar com 90% de nível de confiança acaba por poder prejudicar os resultados, não querendo com isto dizer que os valores projectados não têm fiabilidade, até porque os valores dos erros comprovam o contrário. De referir que se continua a ter um modelo *Lee-Carter* mais fiável no que a intervalos de predição diz respeito. Outra possível explicação para haver um tão grande número de valores fora dos intervalos é o facto de se tratar de valores muito pequenos, estando na ordem de grandeza das milésimas, ou mais pequenos, é natural que mesmo um pequeno erro faça o valor observado sair fora do intervalo de predição.

4.3. População Masculina

4.3.1. Projecções

Para a população masculina o procedimento é idêntico às anteriores, com a construção das curvas de mortalidade para verificar o andamento da mesma, assim como o cálculo dos erros e comparação dos modelos finais.

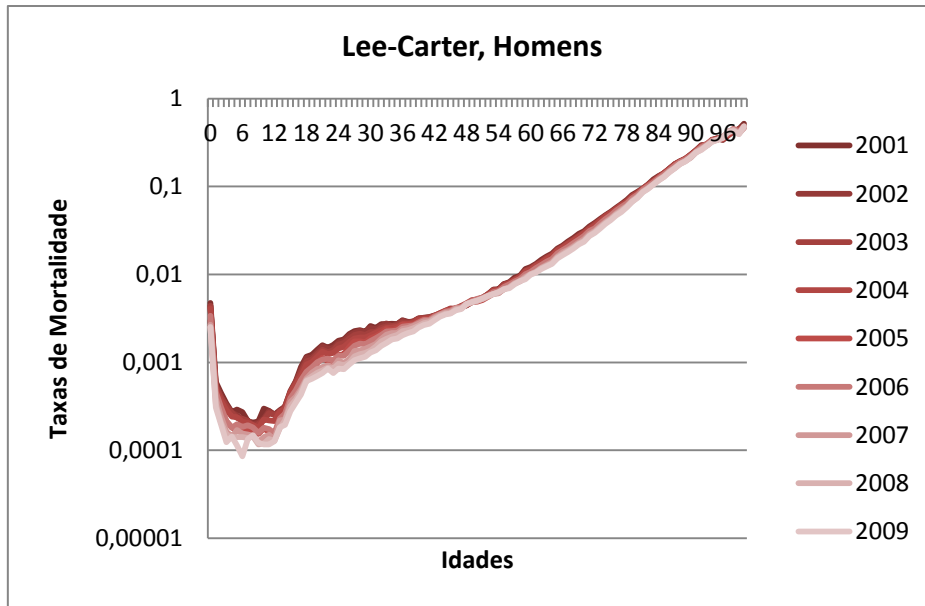


Gráfico 16 – Curvas de Mortalidade Projectadas pelos MLC para a população masculina.

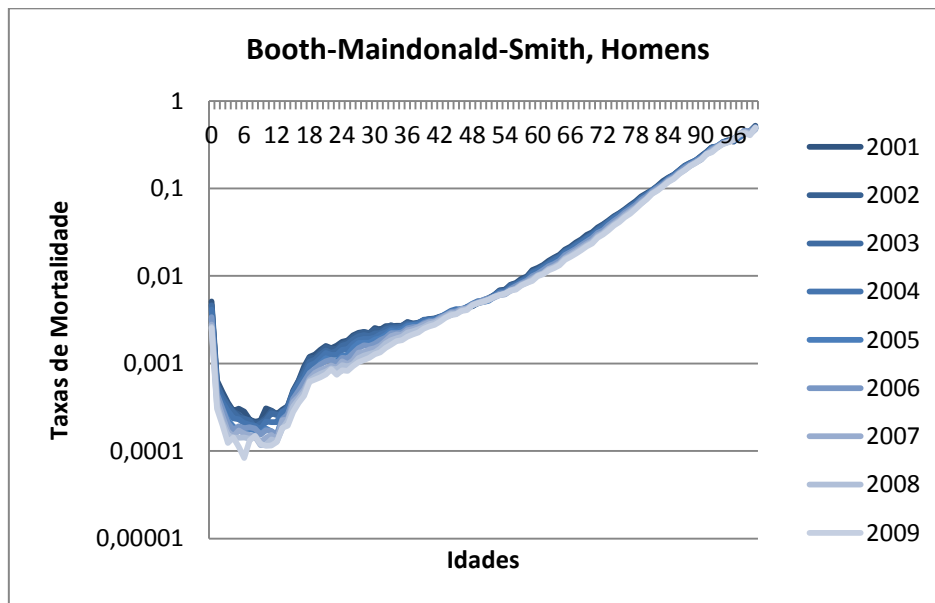


Gráfico 17 – Curvas de Mortalidade Projectadas pelos MBMS para a população masculina.

Mais uma vez as diferenças gráficas apresentadas pelos dois modelos são bastante pequenas, tanto no caso da população total, feminina e agora a masculina. O andamento das duas curvas de mortalidade é semelhante e alisada, não apresenta grandes oscilações. De salientar, apenas como curiosidade demográfica, que a curva de mortalidade projectada para os dois modelos apresenta a “lomba” da mortalidade jovem que está relacionada com os acidentes rodoviários (gráficos 16 e 17).

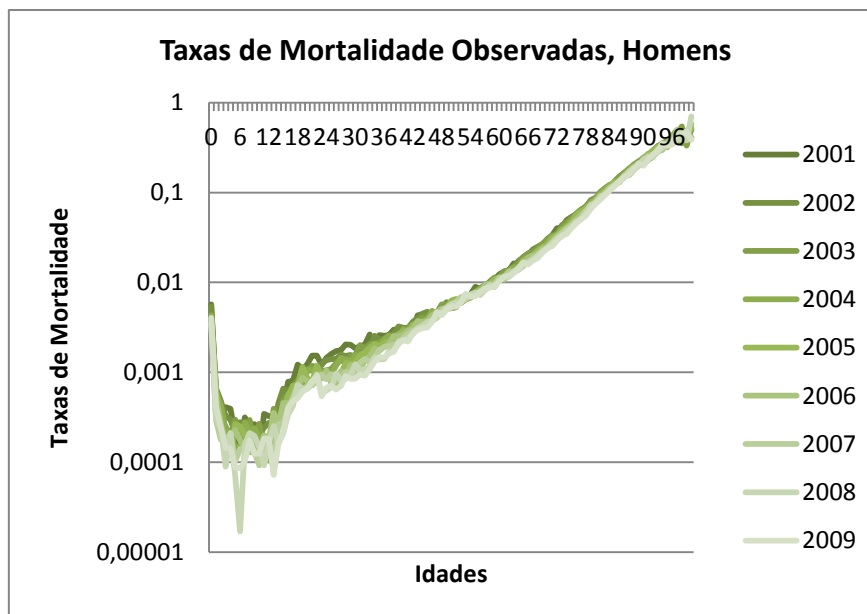


Gráfico 18 – Curvas de Mortalidade Observadas para a população masculina.

Nas taxas de mortalidade observadas (gráfico 18) há acima de tudo que referir um valor bastante anómalo e que acabará por ter algum impacto, embora que relativo, no cálculo dos erros. Trata-se do valor da taxa de mortalidade para os 6 anos de idade no ano de 2008. É um valor claramente mais baixo que o previsto tendo influência não só nos erros que serão calculados em seguida mas também nos erros que anteriormente se calculou para a população total. Não é demais relembrar que havia um valor bastante menor na taxa de mortalidade da população total nesta mesma idade e neste mesmo ano que fazia com que o erro associado a essa projecção chegasse perto dos 200%. Agora percebe-se mais ao pormenor donde virá essa diferença brusca. Por algum motivo, seja um problema da base de dados ou não, o valor de taxa de mortalidade para os 6 anos dos homens em 2008 foi bastante baixo.

Posto isto, verifica-se que mais uma vez os modelos alisam o comportamento das curvas de mortalidade quando comparados graficamente com as taxas de mortalidade observadas.

4.3.2. Erros Associados

Em seguida, calculou-se o Erro Percentual Algébrico (EPAL) para cada idade e ano, assim como o Erro Percentual Algébrico Médio (EPALM) para cada um dos anos projectados.

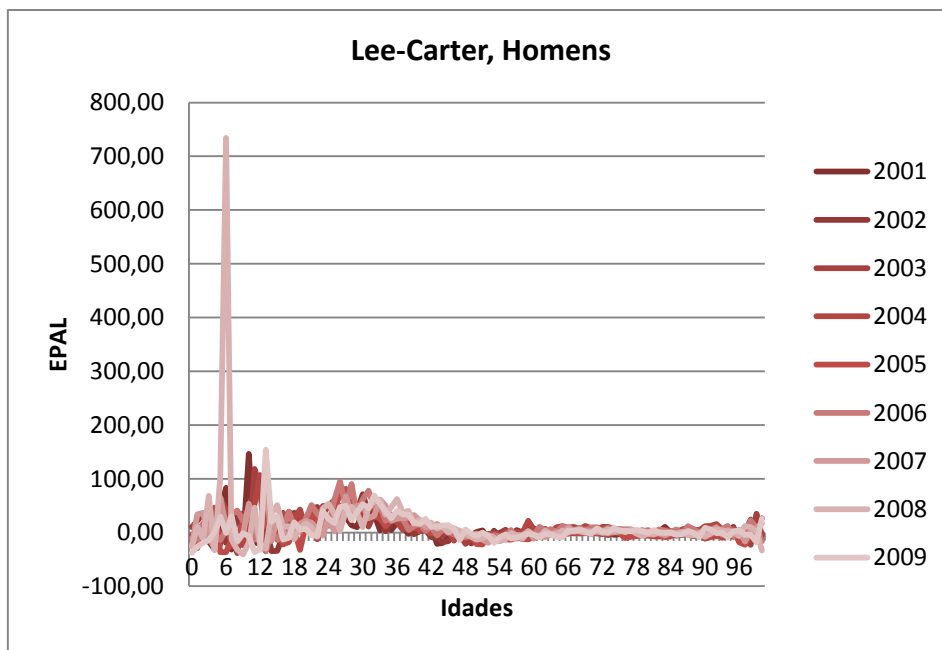


Gráfico 19 – EPAL das projecções dos MLC para a população masculina.

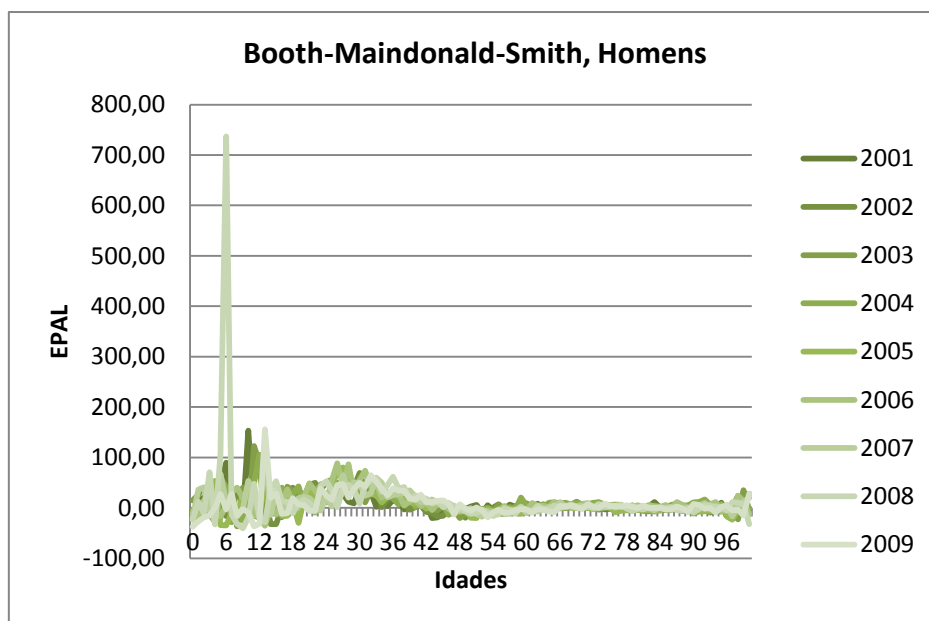


Gráfico 20 – EPAL das projecções dos MBMS para a população masculina.

Como se pode comprovar facilmente pelos gráficos 19 e 20, existe um valor muito acima do esperado e aconselhado. Isto, evidentemente tem a ver com o valor de taxa de mortalidade para a idade de 6 anos no ano de 2008. Mais uma vez se mostra bastante fora do contexto dos outros valores de taxa de mortalidade e com isso aparece um valor de erro acima dos 700%. De realçar que se trata dum valor anormal e que surge nos dois modelos, portanto não pondo em causa a fiabilidade dos mesmos.

Comparando agora os erros dos dois modelos, construiu-se o gráfico 21.

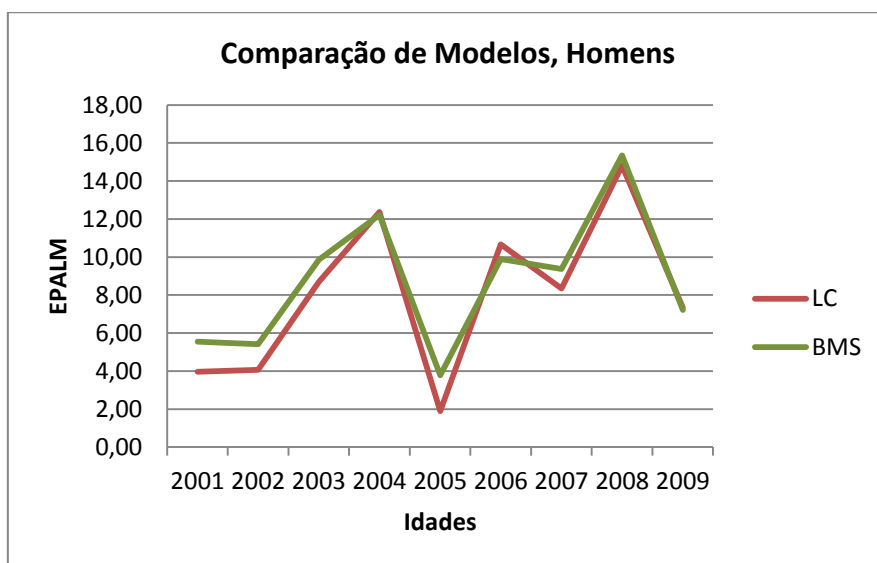


Gráfico 21 – EPALM, comparação entre MLC e MBMS para a população masculina.

Como se pode verificar os dois erros mantêm-se entre valores similares, dos 2% a menos de 16%. Observa-se uma ligeira vantagem para o modelo *Lee-Carter* que apresenta menor erro para as projecções no entanto não será uma diferença muito significativa para dar indicação de qual modelo é o melhor para projecções deste género.

4.3.3. Projecções dentro ou fora do intervalo de predição

Resta então uma análise aos intervalos de predição construídos pelos modelos para as projecções, como se fez para a população feminina, apresenta-se apenas a média de percentagem de valores dentro do intervalo de predição, que dará uma indicação de qual dos modelos terá melhor projecção nos intervalos de confiança.

Média de Percentagem de Valores Dentro do Intervalo	
Lee Carter	37,95
Booth-Maindonald-Smith	29,26

Tabela 3 – Média de Percentagem de Valores Dentro do Intervalo nos dois modelos para a população masculina.

Como se pode observar pelos valores da tabela 3 houve um decréscimo substancial de taxas de mortalidade observadas enquadradas dentro do intervalo de predição quer para o modelo *Lee-Carter* quer para o modelo *Booth-Maindonald-Smith*. De um ponto de vista demográfico é notória a maior dificuldade em projectar intervalos de predição para a população masculina do que para a população feminina. No entanto de destacar que o valor dos Erros Percentuais Algébricos Médios não mostram uma maior dificuldade nas projecções das taxas de mortalidade em relação à população masculina. Como já foi referido, a obtenção de intervalos de predição ao nível de confiança de 90% pode condicionar estes resultados.

4.4.Síntese Final

Resumindo os resultados obtidos é possível concluir-se que a utilização dos dois modelos, quer o modelo *Lee-Carter* quer o modelo *Booth-Maindonald-Smith*, é viável e aceitável de um ponto de vista demográfico.

Observando as curvas de mortalidade para as três bases de dados utilizadas, população total, feminina e masculina, pode-se afirmar que graficamente não existe nenhuma evidência de que a projecção não esteja dentro dos parâmetros aceitáveis nem existe evidência da diferença entre modelos. Para além disso, os dois modelos apresentam projecções com um alisamento positivo e poucas mudanças bruscas de valores de idade para idade, assim como de ano para ano. A comparação das curvas de mortalidade projectadas com as curvas de mortalidade observadas revela que nas taxas observadas existem mais “picos” ao longo da sua curva do que nas curvas projectadas pelos modelos.

Visto tratar-se de projecções existe possibilidade de calcular o erro associado, neste caso optou-se pelos erros percentuais algébricos médios. Em todos os casos estudados surgem valores de erros entre os 2% e os 16%, ou seja, muito abaixo dos 25% que se colocou como patamar de ser relevante, visto tratarem-se de projecções. De referir que as diferenças de erros entre os dois modelos são muito pequenas, não podendo ser factor para a escolha de um dos dois como o melhor. Um dado importante que se pode retirar da análise aos erros das projecções é o facto de o erro ser muito maior nas idades mais jovens (até aos 35 anos) e consideravelmente mais baixo em idades superiores aos 40 anos.

Estes modelos fornecem também uma ferramenta muito útil, os intervalos de predição. Assim, e como mais um factor de escolha, analisou-se o possível enquadramento das taxas de mortalidade observadas dentro dos intervalos de predição correspondentes. O resultado neste particular não foi deveras animador, com os valores das projecções a situarem-se, sempre mais de metade dos mesmos, fora dos intervalos de predição. A conclusão que se pode tirar, e visto que os erros não foram considerados demasiado elevados, é a de que o nível de confiança utilizado, 90%, era demasiado baixo para o género de estatísticas que se está a utilizar.

Como conclusão intermédia deste projecto pode-se afirmar que é positiva a utilização dos modelos *Lee-Carter* e *Booth-Maindonald-Smith* para a projecção de taxas de mortalidade em Portugal. No entanto, este estudo inicial não apresentou diferenças estatísticas suficientes para que uma decisão sobre qual dos dois será o melhor modelo possa ser tomada. Qualquer um dos dois apresenta-se como uma boa solução.

A utilização de diferentes tamanhos de dados históricos utilizados para construção dos modelos pode ser um caminho a explorar para a verificação ou não de diferenças entre os dois modelos. Se neste caso foi utilizado um histórico de 20 anos, poderiam ser utilizados históricos de 30 e 10 anos e tentar perceber se com bases de dados diferentes o comportamento dos modelos se mantém idêntico ou, se por outro lado, se modifica. Assim como projecções para mais do que um ano, nesta análise foi projectado o ano seguinte à base de dados colocada no modelo. No entanto, para projecções com maior amplitude, seja médio ou longo prazo, o comportamento do modelo poderá ser diferente e a comparação dos mesmos pode ser interessante. Os próximos capítulos deste projecto terão em conta estas perguntas.

5. Ajustamento dos MLC e MBMS no Software R

O principal objectivo deste trabalho é a projecção das taxas de mortalidade para Portugal no curto e médio prazo. Assim, e dada a idade de entrada no ensino superior ser, em média, aos 18 anos, optou-se por projectar para o intervalo de 2010 até 2034, ou seja de 25 anos. Para além do intervalo de projecção seleccionado, é necessário também escolher qual o intervalo temporal existente na base de dados que servirá de fundo para as projecções. Tendo em conta que uma das principais componentes do trabalho é a escolha do melhor modelo, faz sentido pensar em vários intervalos temporais e conseqüente decisão sobre qual o melhor para as projecções pretendidas. É de lembrar que o estudo será feito sempre para o total da população e para a população discriminada por sexo.

Utilizando os modelos *Lee-Carter* e *Booth-Maindonald-Smith* para efectuar as projecções é necessário em primeiro lugar ajustar os mesmos, para os dados de Portugal.

Assim, o primeiro passo para o ajustamento dos modelos é o carregamento da base de dados, retirada do *Human Mortality Database*, para o *software* (comando 4).

```
portugal <- read.demogdata("Mx_1x1.txt", "Exposures_1x1.txt",  
type="mortality", label="Portugal")
```

Comando 4 – Carregamento dum base de dados do *Human Mortality Database*

Com a base de dados carregada para o *software*, resta então ajustar os modelos. Para facilitar a compreensão dos mesmos, coloca-se de seguida todos os modelos ajustados e a análise, essa sim, será feita em separado.

```
##Modelos Lee-Carter##  
modeloLCt81<-lca (portugal, series="total", years=anos81.09)  
modeloLCf81<-lca (portugal, series="female", years=anos81.09)  
modeloLCm81<-lca (portugal, series="male", years=anos81.09)
```

Comando 5 – Ajustamento dos MLC, parte 1

```
##Modelo Lee Carter##  
  
modeloLCt90<-lca (portugal, series="total", years=anos90.09)  
modeloLCf90<-lca (portugal, series="female", years=anos90.09)  
modeloLCm90<-lca (portugal, series="male", years=anos90.09)  
  
modeloLCt00<-lca (portugal, series="total", years=anos00.09)  
modeloLCf00<-lca (portugal, series="female", years=anos00.09)  
modeloLCm00<-lca (portugal, series="male", years=anos00.09)
```

Comando 6 – Ajustamento dos MLC, parte 2

Como se pode verificar pelos três conjuntos de ajustamentos, foram ajustados três modelos para cada intervalo temporal (1981-2009, 1990-2009 e 2000-2009), cada um desses modelos será para uma série diferente (total da população, mulheres e homens) (comandos 5 e 6).

O mesmo será realizado para o modelo *Booth-Maindonald-Smith*.

```
##Modelo Booth-Maindonald-Smith##  
  
modeloBMSt81<-bms (portugal, series="total", minperiod=29, years=anos81.09)  
modeloBMSf81<-bms (portugal, series="female", minperiod=29, years=anos81.09)  
modeloBMSm81<-bms (portugal, series="male", minperiod=29, years=anos81.09)  
  
modeloBMSt90<-bms (portugal, series="total", minperiod=20, years=anos90.09)  
modeloBMSf90<-bms (portugal, series="female", minperiod=20, years=anos90.09)  
modeloBMSm90<-bms (portugal, series="male", minperiod=20, years=anos90.09)  
  
modeloBMSt00<-bms (portugal, series="total", minperiod=10, years=anos00.09)  
modeloBMSf00<-bms (portugal, series="female", minperiod=10, years=anos00.09)  
modeloBMSm00<-bms (portugal, series="male", minperiod=10, years=anos00.09)
```

Comando 7 – Ajustamento dos MBMS.

O ajustamento é feito com o comando para o modelo *Booth-Maindonald-Smith*, com a ligeira diferença de ser necessário referir o período mínimo para o intervalo temporal usado, que neste caso coloca-se o número de anos que existe no intervalo temporal utilizado como base de dados (comando 7).

5.1. Recolha dos parâmetros

Com estes comandos (5, 6 e 7) fica-se com os modelos ajustados. O passo seguinte será analisar esse ajustamento com o intuito de perceber se será oportuno efectuar projecções a partir dos modelos ou se, pelo contrário, os modelos apresentam ajustamentos estatisticamente pouco significativos. Essa análise será feita através do cálculo dos erros associados a esses ajustamentos.

Apesar do modelo *Booth-Maindonald-Smith* ser uma variante do modelo *Lee-Carter*, ambos têm a mesma função. Aplicando essa função aos valores ajustados, pode-se retirar as taxas de mortalidade ajustadas para cada ano e idade do intervalo temporal analisado:

$$\ln(\hat{m}_{x,t}) = \hat{a}_x + \hat{b}_x k_t$$

$$\hat{m}_{x,t} = \exp(\hat{a}_x + \hat{b}_x k_t)$$

Assim sendo, resta retirar os parâmetros ajustados (\hat{a}_x, \hat{b}_x) e o parâmetro estimado (k_t) utilizando os seguintes comandos, e será possível calcular as taxas de mortalidade.

Parâmetros \hat{a}_x, \hat{b}_x :

```
write.table(modeloLCt81$ax, file="LCaxt81.txt", sep=" ")
write.table(modeloLCt81$bx, file="LCbxt81.txt", sep=" ")
```

Comando 8 – Obtenção dos parâmetros do modelo \hat{a}_x, \hat{b}_x

Como se pode verificar, o comando 8 trata-se de um exemplo para um dos modelos *Lee-Carter* ajustados anteriormente. Cada código criará um ficheiro *.txt* com os valores do referido parâmetro.

Parâmetro k_t :

```
lct81<-modeloLCt81$fitted  
  
write.table(lct81$y, file="LCt81.txt", sep=" ")
```

Comando 9 – Obtenção do parâmetro do modelo k_t

Neste caso (comando 9) é necessário duas linhas de código do *software* para a criação do ficheiro com a informação pretendida. Mais uma vez será criado um ficheiro *.txt* com os valores do parâmetro pretendido.

Com este procedimento obtiveram-se então os parâmetros para uma análise cuidadosa dos ajustamentos. Contudo, será necessário verificar se os ajustamentos de cada modelo, para cada parcela de população (total, feminina ou masculina), são bem efectuados e se estatisticamente é seguro afirmar-se que se está perante um bom ajustamento.

6. Análise do Ajustamento

6.1. Cálculo das Taxas de Mortalidade Ajustadas

Como já foi referido anteriormente, bastará então calcular as taxas de mortalidade ajustadas para cada ano da base de dados introduzido no modelo e poder-se-á calcular os erros associados do ajustamento de cada modelo. Este cálculo é essencial para a análise do ajustamento e consequentemente essencial para o trabalho pretendido.

Com as taxas de mortalidade ajustadas, de cada ano de base de dados introduzido no modelo calculadas é possível então executar uma análise ao ajustamento de cada modelo. Para uma melhor compreensão desta análise, subdividiu-se a mesma em modelos e em *background* (1981-2009, 1990-2009 e 2000-2009) assim como em séries de população distintas (total, feminina e masculina).

6.2. Modelo *Lee-Carter*

6.2.1. Anos 1981-2009

6.2.1.1. População Total

6.2.1.1.1. Erros Associados

Os dois erros associados que se pretendem analisar são o Erro Percentual Absoluto Médio (EPAM) assim como o Erro Percentual Algébrico Médio (EPALM). No entanto, e como foi o caso numa análise anterior a esta, foram também calculados o Erro Percentual Absoluto (EPA) e Erro Percentual Algébrico (EPAL) para cada ano e idade de *background* associado ao modelo.

Nessa análise das idades em cada ano podem-se verificar algumas particularidades interessantes:

- Nos primeiros anos de ajustamento, o erro é menor. Ou seja, nos anos da década de 80 e início de 90 os erros são bem menores que nos anos posteriores. Verifica-se também que nos anos mais recentes de base de dados existem muito mais “picos” no gráfico. (Anexo 1)

- Outra curiosidade muito interessante é o facto de nos anos da década de 80 e início de 90 os valores serem essencialmente subvalorizados, enquanto nos anos posteriores os valores são essencialmente sobrevalorizados. (Anexo 2)
- Por fim, verifica-se tanto num erro como no outro que nas idades até aos 40 anos são as que apresentam maiores erros de ajustamento e que nas idades superiores a 40 anos o erro acaba sempre por se situar abaixo dos 50%. (Anexo 1 e 2).

Erro Percentual Absoluto Médio

Para melhor interpretação e visualização dos erros, construiu-se o gráfico 22.

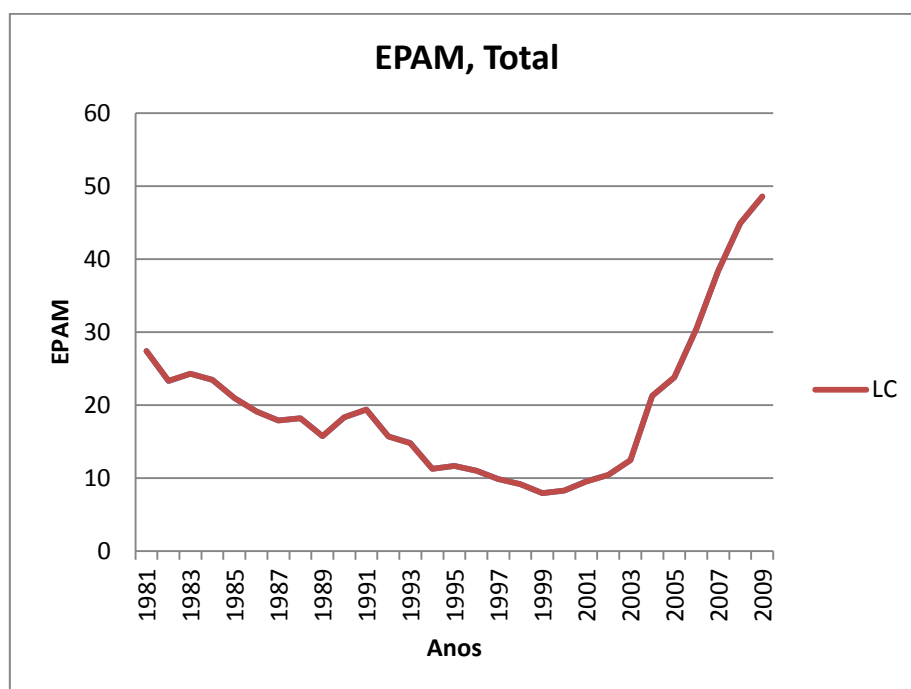


Gráfico 22 – EPAM nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.

Como se pode verificar pelo gráfico, nos cerca de 30 anos de *background* do modelo existe algumas alterações no que o erro diz respeito. Nos primeiros anos de ajustamento do modelo, parece existir um erro um pouco superior ao registado em anos posteriores mas nos últimos anos de ajustamento existe um aumento do erro algo substancial. Repare-se que o erro chega a cair para valores inferiores a 10% no ano 2000 mas aumenta até perto dos 50% no ano de 2009.

De um modo global, o erro percentual absoluto médio situa-se perto dos 20%.

Erro Percentual Algébrico Médio

Mais uma vez para uma melhor interpretação e análise deste erro, construiu-se o gráfico 23.

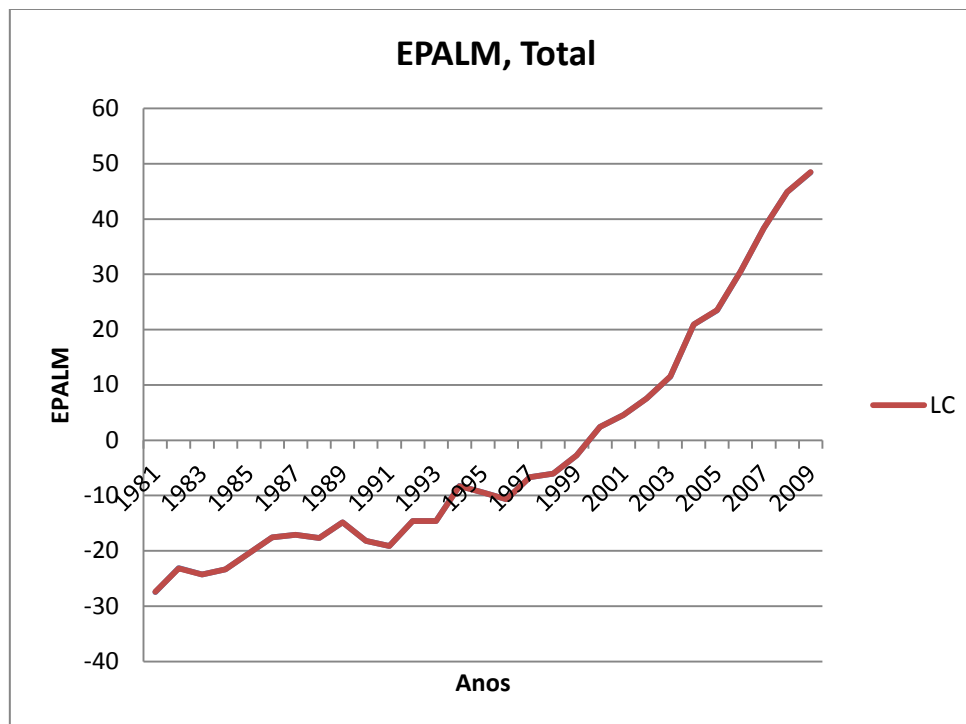


Gráfico 23 – EPALM nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.

É também bastante curioso verificar que o erro do ajustamento neste modelo tem um comportamento como aquele evidenciado no gráfico. É possível observar que existe uma subvalorização das taxas de mortalidade ajustadas até meio dos anos na base de dados introduzidos no modelo e depois uma sobrevalorização para os anos seguintes. De referir que a sobrevalorização é em termos absolutos mais elevada do que a subvalorização.

Globalmente o erro tem um valor na ordem dos -2%.

6.2.1.1.2. Erro Quadrático Médio

Através da fórmula do erro quadrático médio já apresentada anteriormente no capítulo das metodologias foi possível calcular o seguinte valor:

$$EQM = 0,000127$$

O valor do erro quadrático médio encontra-se com uma ordem de grandeza bastante pequena, situando-se muito próximo do zero. É necessário ter em conta que as taxas de mortalidade também têm uma ordem de grandeza pequena, e portanto não se deve retirar muito do seu valor, a não ser a possibilidade de comparação entre os modelos.

6.2.1.1.3. Síntese Final

Fazendo uma análise mais global ao ajustamento do modelo *Lee-Carter* para o total da população e tendo como base de dados introduzida no modelo os valores da população entre 1981 e 2009, pode-se afirmar que o ajustamento não apresenta sinais de ser um mau ajustamento.

Existem de facto alguns valores elevados registados nos erros associados, mas em valores globais não apresentam valores demasiados elevados para o tipo de análise que se está a executar. De salientar também que o erro quadrático médio apresenta um valor próximo de zero, e conseqüentemente aceitável.

6.2.1.2. População Feminina e Masculina

A mesma análise realizada para o total da população é agora feita para a população dividida por género, ou seja, população feminina e população masculina.

Foram calculados os erros associados por idades e anos (Anexos 3 e 4 – População feminina; Anexos 5 e 6 – População masculina), assim como calculados os valores de erros associados para os anos de *background* para cada população (Anexos 7 e 8).

Por fim, foram calculados também os erros globais do modelo. No caso do EPAM, os valores tanto da população feminina tanto da população masculina situam-se perto dos 21%, enquanto no caso do EPALM os valores situam-se entre os -1 e -2% para as duas populações.

De referir que na população masculina existe um valor muito elevado de erro numa idade e apenas num ano. E que de um modo geral nas idades mais novas os erros são maiores que no total da população. No entanto, de uma perspectiva global do erro,

como aliás já foi indicado, o valor não aumenta muito relativamente ao valor obtido no total da população.

De um modo geral, e olhando aos gráficos e quadros como apoio à decisão, pode-se referir que não existe uma diferença substancial entre as duas populações no que se refere a erros associados.

No caso do erro quadrático médio, os valores calculados foram os seguintes:

$$EQM_{feminina} = 0,000134$$

$$EQM_{masculina} = 0,000435$$

Mais uma vez, os valores do erro quadrático médio são pequenos. De notar que o erro para a população masculina é bastante superior.

De um modo geral, não existe muita diferença entre o total da população e as populações diferenciadas por sexo.

6.2.2. Anos 1990-2009

6.2.2.1. População Total

6.2.2.1.1. Erros Associados

Assim como foi efectuado para uma base de dados com os anos compreendidos entre 1981 e 2009, foram calculados os erros associados ao ajustamento do modelo para uma base de dados com um menor número de anos como histórico. Não só os erros globais mas também os erros em cada idade e os erros por ano de ajustamento. Utilizar-se-á a mesma sequência de análise realizada para os anos compreendidos entre 1981 e 2009.

Construídos os gráficos com os erros em cada uma das idades é possível verificar mais uma vez que nas idades inferiores a 40 anos os erros apresentam-se muito superiores aos erros observados nas idades mais elevadas. Também é possível verificar através do EPALM que existe novamente uma subvalorização nos primeiros anos de ajustamento e uma sobrevalorização nos últimos anos de ajustamento. (Anexos 9 e 10)

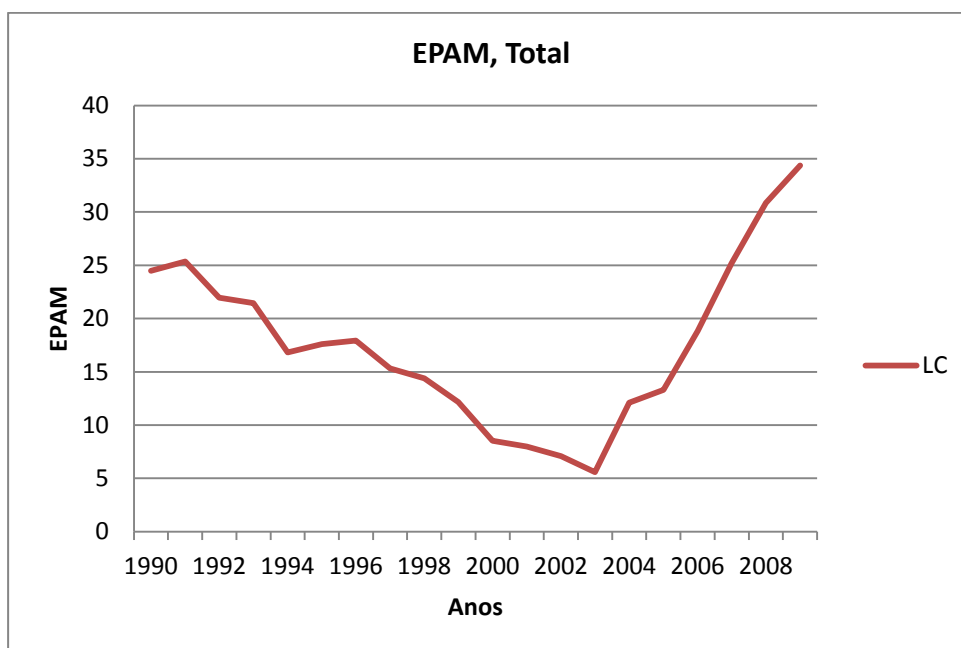
Erro Percentual Absoluto Médio

Gráfico 24 – EPAM nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.

Como se pode verificar pelo gráfico 24, continuam a existir erros absolutos em maior escala no final do período de análise do que no seu início, assim como se observa que o início dos anos 2000 é o período com menores erros. De salientar que, em comparação com o período de 1981-2009, a diferença entre os valores absolutos de erro do início do período com o fim do período diminuiu. Agora, a diferença é apenas de cerca de 10% (25% a 35%).

O valor global do erro percentual absoluto médio é de aproximadamente 17,57%.

A variação da mortalidade é bastante significativa nos últimos anos, em especial nos homens. Logo, quanto maior o histórico, menor o erro.

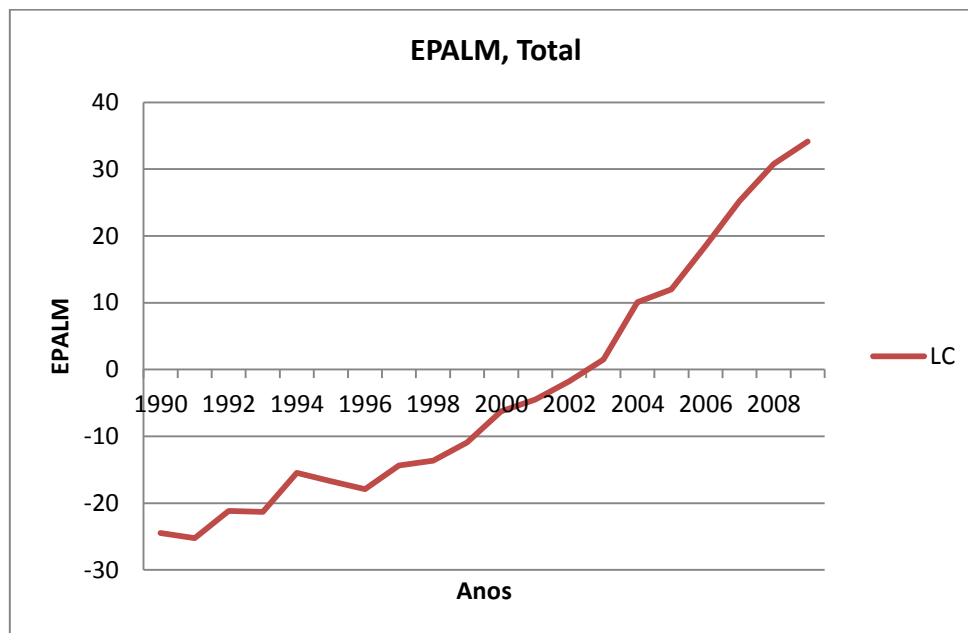
Erro Percentual Algébrico Médio

Gráfico 25 – EPALM nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.

O andamento do erro percentual algébrico é muito idêntico ao do período analisado anteriormente (gráfico 25). Em termos absolutos parece que o erro diminuiu ligeiramente comparativamente ao período anterior. No entanto, continua a existir uma subvalorização nos primeiros anos e uma sobrevalorização nos últimos anos.

O valor global do erro percentual algébrico médio é de aproximadamente $-3,07\%$.

6.2.2.1.2. Erro Quadrático Médio

O valor calculado é de:

$$EQM = 0,000095$$

Mais uma vez, chega-se a um valor bastante próximo de zero.

6.2.2.1.3. Síntese Final

Constata-se que assim como no caso do modelo para o período entre 1981 e 2009, também o modelo com um período de ajustamento entre 1990 e 2009 apresenta valores de erro perfeitamente aceitáveis para este tipo de análises e modelação. Verificam-se realmente alguns “picos” nalgumas idades e nalguns anos específicos mas há que ter em conta que se está a analisar taxas de mortalidade, em que os seus valores mesmo que oscilando pouco, dada a sua ordem de grandeza, podem resultar em erros grandes. Não pondo em causa os valores de erros globais, que se apresentam bastante razoáveis.

Tais erros pontuais não retiraram a habilidade do modelo em ajustar de forma bastante positiva as taxas de mortalidade, facto esse que parece que pode ser destacado pelo valor de erro quadrático médio estar bastante perto do zero.

6.2.2.2. População Feminina e Masculina

De um modo sucinto, e tendo como base de sustentação a metodologia utilizada para as análises anteriores pode-se afirmar que este segundo período de ajustamento dos anos de 1990 a 2009 não apresenta, em termos de análise de ajustamento, grandes diferenças comparativamente ao anterior, no que a população feminina e masculina diz respeito.

Mais uma vez existem valores de erros superiores nas idades mais jovens, havendo tanto na população feminina como na população masculina “picos” acentuados, em determinados anos e idades pontuais. (Anexos 11 e 12) No caso do EPAM pode-se verificar que os valores se situam entre os 5 e os 45%. (Anexo 13) Existe também uma subvalorização nos primeiros anos, assim como uma sobrevalorização dos anos mais recentes de análise. (Anexos 14, 15 e 16)

No que diz respeito a valores globais, o EPAM apresenta valores a rondar os 19% para as duas populações, um pouco mais baixo que o período anterior. O EPALM

por seu lado apresenta valores entre os -2 e os -3%, um pouco maior em valor absoluto que no período anterior.

Quanto ao erro quadrático médio observaram-se os seguintes valores:

$$EQM_{feminina} = 0,000103$$

$$EQM_{masculina} = 0,000214$$

Valores mais uma vez próximos de zero. Continua a existir um erro superior quando se ajusta a população masculina.

Fazendo a comparação entre população total e população diferenciada por sexos, pode-se concluir que continua a não existir diferenças significativas entre as três séries.

6.2.3. Anos 2000-2009

6.2.3.1. População Total

6.2.3.1.1. Erros Associados

Depois de efectuada a análise para dois períodos distintos de 20 e 30 anos, falta fazer a análise para o período mais curto de ajustamento, com 10 anos de informação. Efectua-se o mesmo tipo de análise registada anteriormente, com erros associados ao ajustamento, quer o EPAM como o EPALM.

Da mesma forma, construíram-se os gráficos para os EPA e EPAL ao longo das idades e anos, para que se possa verificar se o andamento dos erros ao longo das idades se mantém igual aos períodos anteriores ou se diferem de alguma maneira. (Anexos 17 e 18) Há a salientar o facto de mais uma vez se manter os maiores erros nas idades mais jovens e de haver uma clara subvalorização nos primeiros anos.

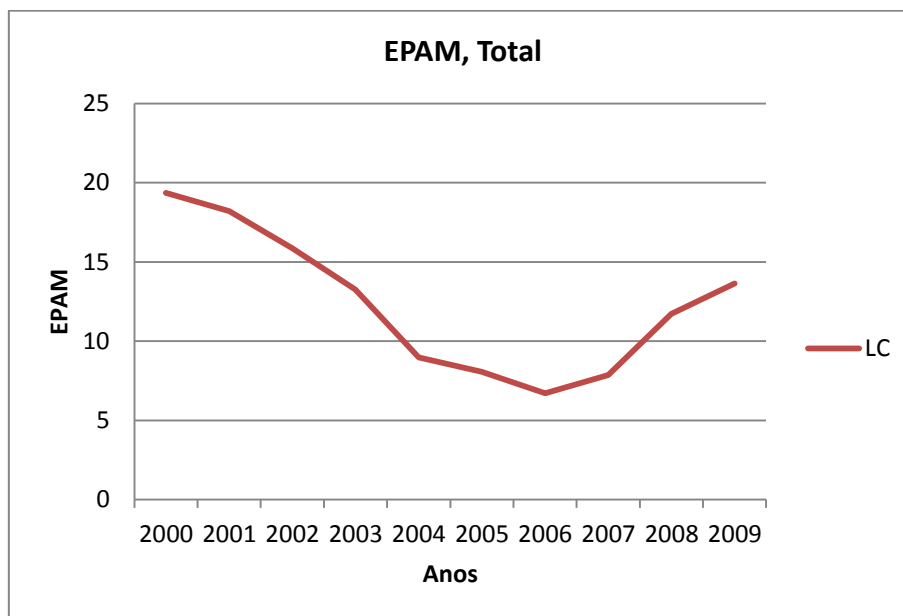
Erro Percentual Absoluto Médio

Gráfico 26 – EPAM nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.

Observando o gráfico 26, encontram-se algumas diferenças em relação aos períodos anteriores. Se nos períodos mais longos de ajustamento os últimos anos do ajustamento tinham sempre erros superiores ao início do ajustamento, aqui isso não se verifica. Os erros no final do ajustamento diminuem consideravelmente em comparação com os períodos anteriores. Continua, no entanto, a haver uns anos pelo meio com menor erro do que no início e no fim do período de ajustamento.

De referir também que, neste caso, os erros no geral são menores que nos períodos anteriores, tendo um EPAM global na ordem dos 12,36%.

Visto que os erros têm vindo a diminuir com os períodos mais curtos é possível afirmar que com períodos mais curtos de informação colocada no ajustamento o erro tem tendência a diminuir.

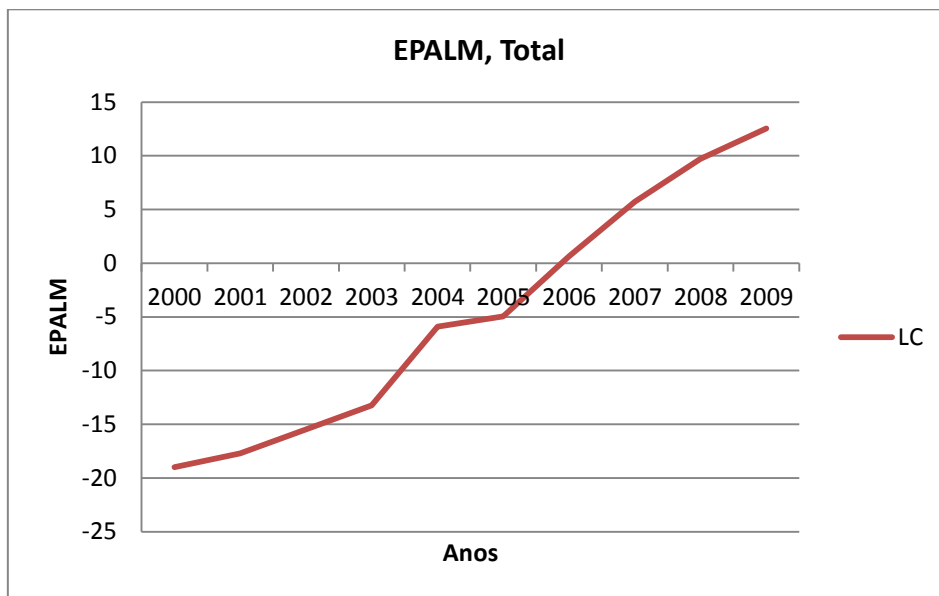
Erro Percentual Algébrico Médio

Gráfico 27 – EPALM nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.

Duas características a reter quando se observa o gráfico do EPALM: a continuada subvalorização dos valores nos primeiros anos de ajustamento e a sobrevalorização dos valores nos últimos anos do ajustamento e a diminuição do erro ao longo do ajustamento. Neste caso, os valores do erro situam-se entre -20% e 15%, encurtando e muito a amplitude de valores se comparada com a dos períodos anteriores (gráfico 27).

O valor de EPALM em termos globais situa-se a rondar os -4,76%.

6.2.3.1.2. Erro Quadrático Médio

O valor calculado é de:

$$EQM = 0,000187$$

Não parece existir uma grande diferença de erros quadráticos médios quando se modifica a base de dados, continua a ser um valor bastante próximo de zero e na mesma ordem de grandeza dos anteriores estudos.

6.2.3.1.3. Síntese Final

A redução do número de anos no ajustamento dos modelos parece indicar uma melhoria ao nível dos erros EPAM, e um ligeiro agravamento ao nível dos EPALM. Enquanto os EPAM melhoram com o redução do número de anos no período de ajustamento, os EPALM tendem a acentuar a subvalorização global do ajustamento.

No entanto, em termos de comportamento continuam a ser bastante similares aos restantes períodos, ou seja, continua a haver subvalorização dos valores de taxas de mortalidade no início do período de ajustamento e sobrevalorização no final do período. E também, por outro lado, continua a haver uma maior dificuldade em ajustar idades mais jovens em comparação com as idades mais avançadas.

No que se refere à habilidade para ajustar, o modelo continua a apresentar um erro quadrático médio próximo de zero.

6.2.3.2. População Feminina e Masculina

Usando como princípio metodológico o utilizado anteriormente verificou-se a existência ou não de diferenças entre a população total e a população com os sexos discriminados, dentro deste período de ajustamento mais curto.

Olhando aos gráficos dos EPA ao longo das idades, pode-se constatar que, tanto nas mulheres como nos homens, os valores de erro não oscilam muito, tirando alguns picos nas idades mais jovens que aliás têm vindo a ser hábito. (Anexo 19 e 20) Na comparação entre as idades pode-se verificar que os homens parecem ter um erro algo inferior ao do verificado na população feminina ao longo de mais anos. (Anexo 21) No que a EPAL diz respeito verificam-se novamente alguns “picos” nas camadas mais jovens, enquanto na comparação entre as duas populações se observa uma similaridade de valores. Voltando novamente a subvalorizar-se nos primeiros anos e sobrevalorizar-se nos últimos anos. (Anexos 22, 23 e 24)

Quanto a valores globais pode-se referir que para o EPAM, os valores situam-se entre os 14 e 15% para as duas populações, feminina e masculina. No caso do EPALM, os valores situam-se a rondar os -4%.

Quanto à análise do erro quadrático médio, chega-se aos seguintes valores:

$$EQM_{feminina} = 0,000059$$

$$EQM_{masculina} = 0,000187$$

Verifica-se uma constante nesta análise, com os erros quadráticos médios sempre superiores na população masculina em relação à população feminina, no entanto com a diferença a ser muito pequena e sempre com valores muito perto do zero.

6.3. Modelo *Booth-Maindonald-Smith*

Este capítulo do trabalho será bastante idêntico ao anterior em termos metodológicos. Neste caso, será para o modelo *Booth-Maindonald-Smith*, visto ser objectivo do trabalho a comparação entre ambos os modelos e tentar aferir qual dos dois o que melhor se adapta aos dados de Portugal.

6.3.1. Anos 1981-2009

6.3.1.1. População Total

6.3.1.1.1. Erros Associados

A primeira análise será feita aos erros de ajustamento do modelo. No que toca a EPA e EPAL ao longo das idades e anos pode-se verificar uma similaridade entre o modelo anterior, o *Lee-Carter*, e este. Ou seja, existe uma predominância pelos valores mais elevados de erros nas camadas mais jovens das idades, assim como nos anos mais recentes uma predisposição para a sobrevalorização, enquanto nos primeiros anos de ajustamento há uma subvalorização de ajustamento. (Anexos 25 e 26)

Erro Percentual Absoluto Médio

No que toca ao EPAM construiu-se o gráfico 28, de modo a registar qual os anos com valores de erros no ajustamento:

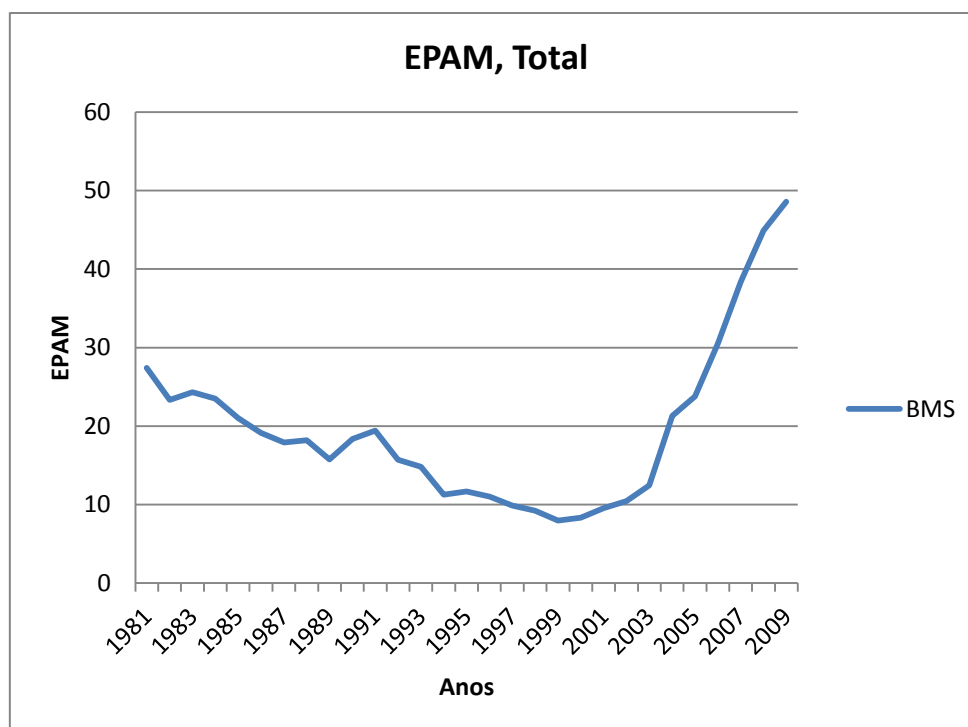


Gráfico 28 – EPAM nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.

No gráfico pode-se verificar a tendência dos valores dos erros para diminuir durante os primeiros 20 anos de base de dados e aumentar consideravelmente nos últimos 10 anos, acabando com um valor de erro superior no final do ajustamento comparativamente ao início.

De um modo global, há que referir que o valor de EPAM do modelo *Booth-Maindonald-Smith*, para a população total, se situa na ordem dos 19,59%.

Comparativamente ao modelo de LC, pode-se afirmar que os dois modelos apresentam erros bastante idênticos, quer em comportamento ao longo dos anos e idades, quer em valor global. (Anexo 49)

Erro Percentual Algébrico Médio

Para avaliar a sobrevalorização e subvalorização dos erros, o EPALM apresentasse como a melhor ferramenta para o fazer. Desse modo, constrói-se um gráfico idêntico mas agora para este erro.

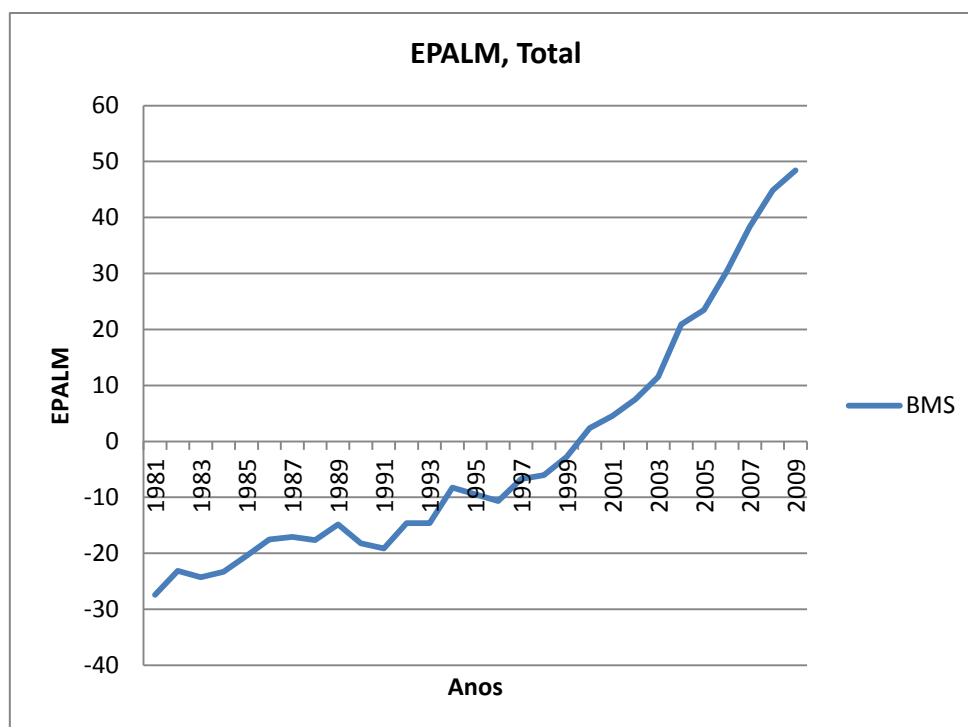


Gráfico 29 – EPALM nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.

Observando o gráfico 29 regista-se uma tendência óbvia e clara para o aumento do valor do erro, em termos reais. Verifica-se novamente a similaridade com o modelo *Lee-Carter*, com o erro a subvalorização nos primeiros anos de ajustamento e a sobrevalorização, em maior escala, nos últimos anos de ajustamento, sendo os meados dos anos 90 a melhor fase de ajustamento.

De um modo global o valor de EPALM para o modelo *Booth-Maindonald-Smith*, para a população total, situa-se a rondar os -2%.

Numa comparação mais directa com o modelo *Lee-Carter*, pode-se afirmar que tantos os valores dos erros de ajustamento como o comportamento dos mesmos ao longo dos anos e idades, para a população total, são bastante idênticos. (Anexo 49)

6.3.1.1.2. Erro Quadrático Médio

O valor de *EQM*, calculado foi de:

$$EQM = 0,000127$$

E apresenta-se como um valor muito similar ao calculado para o modelo *Lee-Carter* nas mesmas circunstâncias que estas, ou seja, mesmo período temporal de ajustamento e mesma população. O que só demonstra a excelente bondade de ajustamento, de qualquer um dos modelos.

6.3.1.1.3. Síntese Final

Analisando apenas o modelo *Booth-Maindonald-Smith*, pode-se concluir que, para o total da população, o ajustamento efectuado é bastante válido. Os valores dos erros apresentados não foram excessivos, tirando alguns erros pontuais facilmente explicados pelas características das taxas de mortalidade, e apresenta um excelente erro quadrático médio.

De um modo geral pode-se afirmar que os valores de erros, quer os EPAM quer os EPALM, são bastante idênticos aos registados na análise ao modelo *Lee-Carter*. O comportamento dos dois modelos ao nível do ajustamento é bastante idêntico para o total da população.

6.3.1.2. População Feminina e Masculina

Como tem sido usual, a primeira análise aos modelos será a análise feita aos erros. Numa primeira abordagem observa-se o comportamento dos EPA ao longo das idades. (Anexo 27 e 28) Mais uma vez, o comportamento nas duas populações é bastante idêntico. Os gráficos podem parecer diferentes mas acaba por ser uma questão de escala, visto que existe um valor na população masculina que foge muito do expectável. Mas quando comparados os dois EPAM, quer para a população feminina quer para a população masculina, pode-se verificar a similaridade entre os mesmos. Não havendo, de facto, nenhuma população com uma predominância em termos de erros de ajustamento em relação à outra. (Anexo 29)

Quanto ao EPAL, a análise é a mesma. O comportamento dos erros do ajustamento comparando ambas as populações, é idêntico, havendo de facto um valor na população masculina que distorce um pouco a percepção disso. (Anexos 30 e 31) Há a salientar, mais uma vez, a constante subvalorização dos modelos nos primeiros anos de ajustamento assim como a sobrevalorização nos últimos anos. (Anexo 32)

Em termos globais, os modelos quer ao nível de EPAM quer ao nível de EPALM, apresentam valores bastante idênticos, a rondar os 21% e os -1%, respectivamente.

Os erros quadráticos médios calculados foram:

$$EQM_{feminina} = 0,000134$$

$$EQM_{masculina} = 0,000435$$

São valores novamente bastante baixos e que indicam uma excelente capacidade de ajustamento dos modelos. Mais uma vez destaque para um valor mais alto quando calculado para a população masculina.

Aferindo agora sobre os dois modelos em comparação, *Lee-Carter* e *Booth-Maindonald-Smith*, pode-se concluir que quer para a população feminina, quer para a população masculina, as similaridades entre ambos são muitas e acentuadas. Não havendo diferenças significativas entre os modelos ao nível do ajustamento. (Anexo 49)

6.3.2. Anos 1990-2009

6.3.2.1. População Total

6.3.2.1.1. Erros Associados

Numa primeira abordagem aos erros, quer o EPA quer o EPAL ao longo dos anos, apresentaram-se idênticos ao período anterior e comparativamente com o modelo *Lee-Carter*, que é a análise que se pretende com este trabalho, o comportamento dos mesmos é similar e sem grandes alterações de registo. (Anexos 33 e 34)

Em seguida serão apresentados os dois erros em termos médios ao longo dos anos para que se possa ter uma ideia do seu comportamento e a ponto de fazer uma comparação obrigatória não só entre períodos de ajustamento mas também entre modelos.

Erro Percentual Absoluto Médio

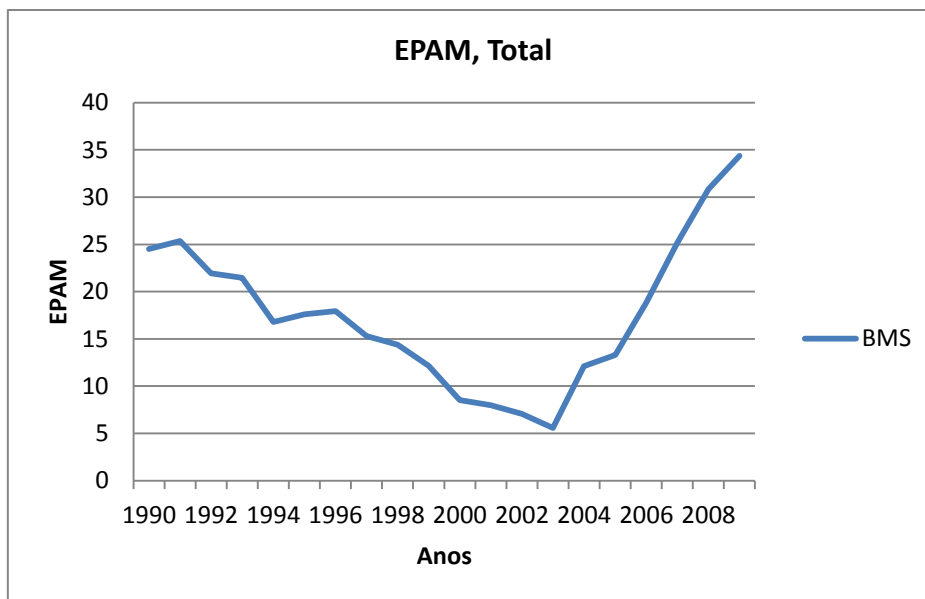


Gráfico 30 – EPAM nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.

Como se pode verificar pelo gráfico 30, o comportamento do EPAM é em tudo idêntico aos anteriores. Diminui nos primeiros anos de ajustamento e aumenta nos posteriores. Em termos globais o valor de erro anda à volta dos 17,56%.

Erro Percentual Algébrico Médio

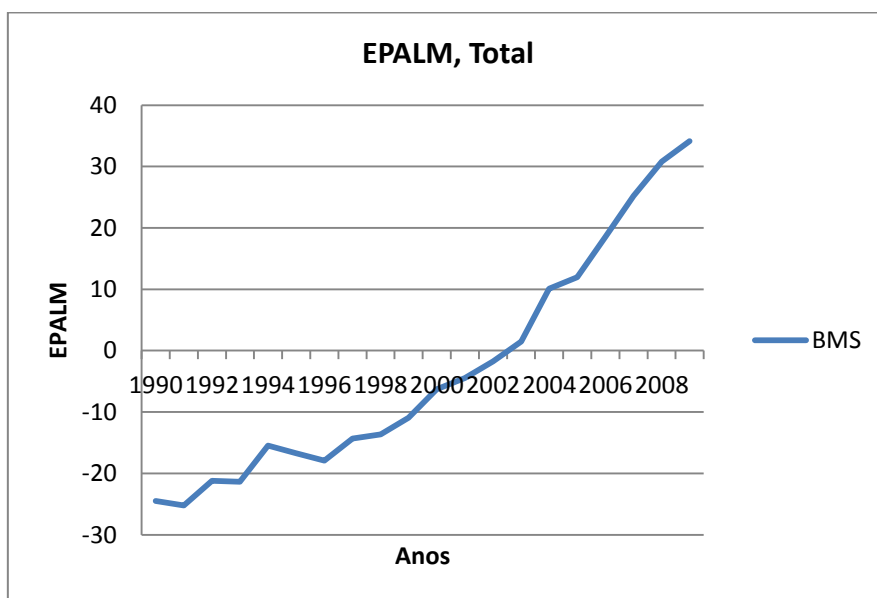


Gráfico 31 – EPALM nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.

O EPALM apresenta um comportamento característico, subvalorizando nos primeiros anos de ajustamento e sobrevalorizando a partir dum determinado ano, mais

ou menos a meio do período (gráfico 31). Este comportamento é em tudo idêntico aos ajustamentos anteriores, quer para períodos diferentes quer para o modelo *Lee-Carter*.

O valor global apresentado por este erro é de cerca de -3%.

6.3.2.1.2. Erro Quadrático Médio

O valor calculado é de:

$$EQM = 0,000095$$

É um valor praticamente idêntico ao calculado para o mesmo período para o modelo *Lee-Carter*. Os dois modelos apresentam valores de erro quadrático médio baixos.

6.3.2.1.3. Síntese Final

Pode-se afirmar que neste período em particular, assim como no período anterior, não existem diferenças significativas entre os modelos *Booth-Maindonald-Smith* e *Lee-Carter*, para o conjunto total da população, quer ao nível de erros associados, quer ao nível de erros quadráticos médios. O comportamento dos erros associados são praticamente idênticos, assim como os seus valores globais.

6.3.2.2. População Feminina e Masculina

Para finalizar a análise a este período de ajustamento, entre os anos de 1990 a 2009, é necessário verificar se a constância existente para os períodos e modelos anteriores se mantém para este período quando aplicado o ajustamento à população feminina e masculina.

No que a erros diz respeito, observando os gráficos tanto para as mulheres como para os homens, não se vislumbra qualquer diferença significativa em comparação com os modelos anteriores. Tanto no EPAM como no EPALM, os comportamentos verificados, assim como os valores globais calculados, mostram que os modelos mantêm uma similaridade quase completa. (Anexos 35 a 40)

Os valores globais de EPAM e EPALM para as duas populações rondam os 19 e os -2%, respectivamente.

Já o erro quadrático médio apresenta os seguintes valores:

$$EQM_{feminina} = 0,000103$$

$$EQM_{masculina} = 0,000214$$

Constata-se, mais uma vez, que os valores da população masculina são superiores e que se situam ambos próximos do zero.

Comparando de uma forma sintética os dois modelos para este período de ajustamento nas três séries de população utilizadas pode-se concluir que os ajustamentos são em tudo semelhantes e não apresentam diferenças significativas, quer ao nível de erros quer ao nível de erro quadrático médio. (Anexo 49)

6.3.3. Anos 2000-2009

6.3.3.1. População Total

6.3.3.1.1. Erros Associados

Finalmente chega-se ao último período em análise realizado para os ajustamentos dos modelos. Os erros registados ao longo das idades e anos não variam consideravelmente quer nos anteriores períodos quer no modelo *Lee-Carter*, muito pelo contrário, os erros continuam a registar poucas diferenças ao nível de população total, apresentando alguns “picos” nas camadas mais jovens e sendo mais regular nas camadas mais idosas. (Anexos 41 e 42)

Erro Percentual Absoluto Médio

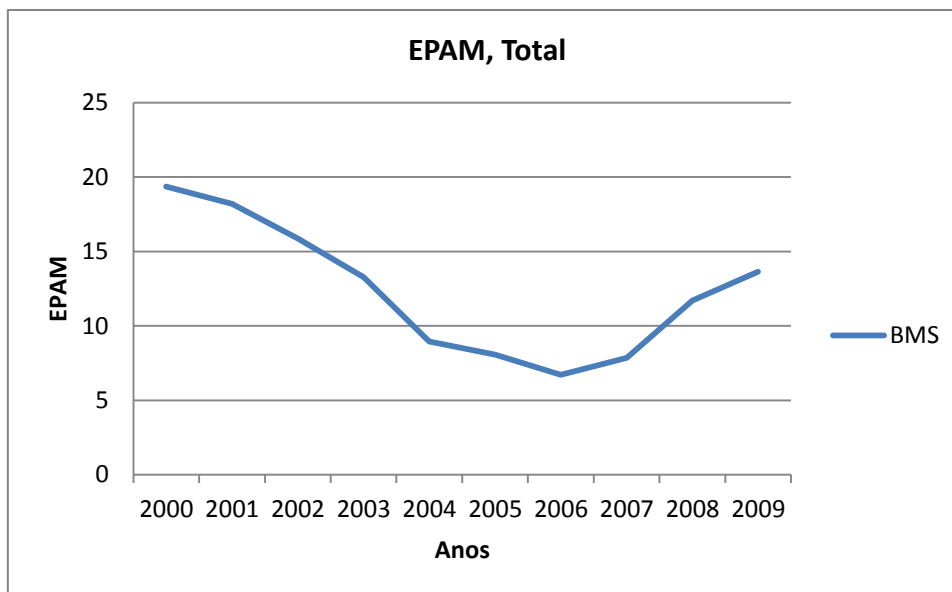


Gráfico 32 – EPAM nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.

Como se pode verificar no gráfico 32, o EPAM continua a apresentar valores mais elevados no início e no fim do período de ajustamento. Neste caso, assim como aconteceu para o modelo *Lee-Carter*, os valores do erro baixam em comparação com os outros períodos.

Em termos globais, o valor de EPAM registado está na ordem dos 12%.

Erro Percentual Algébrico Médio

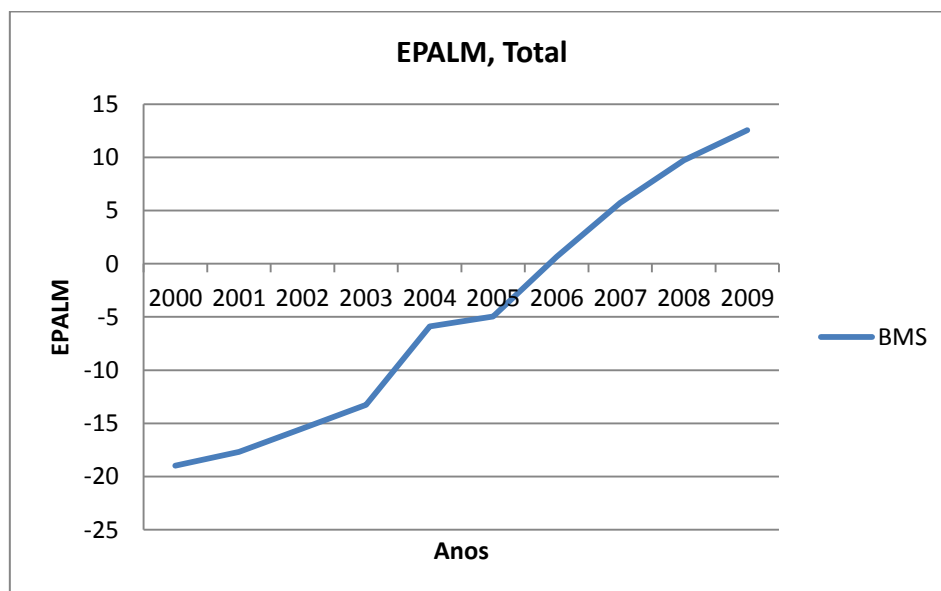


Gráfico 33 – EPALM nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.

Neste caso, o EPALM apresenta um gráfico quase linear (gráfico 33). Apresentando um aumento de erro em valores reais. No entanto, o importante a reter é,

novamente, a subvalorização dos valores das taxas ajustadas nos primeiros anos de ajustamento e uma sobrevalorização nos anos finais de ajustamento.

O valor global deste erro para este modelo e período ronda os -4%.

6.3.3.1.2. Erro Quadrático Médio

O EQM registou um valor de:

$$EQM = 0,000056$$

Como se pode comprovar, volta-se a ter um erro quadrático médio bastante próximo do valor obtido para o mesmo período de tempo mas com o ajustamento de *Lee-Carter*. Sendo novamente bastante próximo do valor zero.

6.3.3.1.3. Síntese Final

De um modo geral pode-se afirmar que os valores dos erros, tanto absolutos como algébricos são bastante similares para os dois modelos em questão (*Booth-Maindonald-Smith* e *Lee-Carter*).

Regista-se também que, como aconteceu para o modelo *Lee-Carter*, os valores dos erros EPAM diminuíram com a diminuição do número de anos colocados no período de ajustamento. Ou seja, quanto menos anos estão no período de ajustamento melhor parece o modelo ajustar. Já no caso dos erros EPALM parece haver um ligeiro agravamento no erro, passando os modelos a subvalorizar um pouco mais os valores ajustados. (Anexo 49)

Também no erro quadrático médio se verifica a mesma situação. Ou seja, os valores são muito idênticos entre modelos usados e os valores melhoram com um menor espaço temporal de ajustamento. (Anexo 49)

6.3.3.2. População Feminina e Masculina

No que respeita à divisão da população total pelos sexos, esta última análise mostra que os valores de erros não variam muito entre modelos. Embora entre períodos haja, como na população total, uma pequena alteração de valores, sendo benéfica no EPAM e em sentido contrário no EPALM, a verdade é que no comportamento dos erros, o efeito das séries de população é diminuto.

Analisando os gráficos de EPA e EPAL ao longo das idades e anos, para as populações, pode-se constatar que, na comparação entre as mesmas o comportamento e os valores observados são em tudo semelhantes aos da população total, assim como os valores do modelo *Booth-Maindonald-Smith* são bastante similares aos apresentados pelo modelo *Lee-Carter*. (Anexos 43 a 48)

No que a erro quadrático médio diz respeito, os valores calculados foram os seguintes:

$$EQM_{feminina} = 0,000059$$

$$EQM_{masculina} = 0,000187$$

Verifica-se mais uma vez que os valores são parecidos aos apresentados, não só pela população total como pelo modelo Lee-Carter para o mesmo período de ajustamento. Continuam a apresentar-se valores próximos de zero e com a população masculina a apresentar um valor superior quando comparado com a população feminina.

7. Projecções

As projecções são de facto o grande objectivo do projecto e serão delas que irão ser tiradas as conclusões mais relevantes do mesmo, pelo que se entende como o capítulo mais importante do estudo.

Neste capítulo ir-se-á proceder a uma divisão dos modelos de forma distinta da efectuada para os ajustamentos. No capítulo de ajustamentos foram divididos em modelos, em seguida em períodos temporais e finalmente entre sexos. Neste caso, será efectuada a comparação entre modelos e portanto dividir-se-á inicialmente por períodos temporais, para poder comparar os dois modelos em situações idênticas.

7.1. Ajustamento e Recolha das Projecções do *Software R*

Antes de se iniciar a análise às projecções é necessário criá-las e retirar-las do *software* utilizado no projecto. Para tal, utilizaram-se várias linhas de código.

```
###Projecções dos Modelos de Lee-Carter###  
foreLct81 <- forecast(modeloLct81, h=25, level=90)  
foreLcf81 <- forecast(modeloLcf81, h=25, level=90)  
foreLcm81 <- forecast(modeloLcm81, h=25, level=90)  
  
foreLct90 <- forecast(modeloLct90, h=25, level=90)  
foreLcf90 <- forecast(modeloLcf90, h=25, level=90)  
foreLcm90 <- forecast(modeloLcm90, h=25, level=90)  
  
foreLct00 <- forecast(modeloLct00, h=25, level=90)  
foreLcf00 <- forecast(modeloLcf00, h=25, level=90)  
foreLcm00 <- forecast(modeloLcm00, h=25, level=90)
```

Comando 10 - Ajustamento de Projecções para os Modelos Lee-Carter.

No comando 10 é possível verificar que é necessário ajustar as projecções aos modelos já criados, utilizando o comando *forecast*, esse ajustamento é fácil. Aliás, exactamente da mesma forma que foi efectuado no capítulo 6, aquando da comparação entre projecções e valores observados já registados. Ou seja, neste comando terá de se

colocar não só o modelo que se está a querer projectar mas também o número de anos a projectar, neste caso 25 anos, e também o nível de confiança do intervalo de predição.

Em seguida fez-se o mesmo procedimento para os modelos *Booth-Maindonald-Smith* (comando 11).

```
###Projecções dos Modelos de Booth-Maindonald-Smith###  
foreBMSt81 <- forecast(modeloBMSt81, h=25, level=90)  
foreBMSf81 <- forecast(modeloBMSf81, h=25, level=90)  
foreBMSm81 <- forecast(modeloBMSm81, h=25, level=90)  
  
foreBMSt90 <- forecast(modeloBMSt90, h=25, level=90)  
foreBMSf90 <- forecast(modeloBMSf90, h=25, level=90)  
foreBMSm90 <- forecast(modeloBMSm90, h=25, level=90)  
  
foreBMSt00 <- forecast(modeloBMSt00, h=25, level=90)  
foreBMSf00 <- forecast(modeloBMSf00, h=25, level=90)  
foreBMSm00 <- forecast(modeloBMSm00, h=25, level=90)
```

Comando 11 – Ajustamento de Projecções para os Modelos Booth-Maindonald-Smith.

Mais uma vez deve-se ter em conta que é necessário colocar os anos que se quer projectar assim como o nível de confiança para o intervalo de predição.

Em seguida, é necessário retirar os valores das projecções das taxas de mortalidade para se poder analisar os mesmos; para tal utilizou-se o comando 12.

```
write.table(foreLCt81$rate, file="fLCt81.txt", sep=" ")
```

Comando 12 – Obtenção das Taxas de Mortalidade Projectadas.

Referir que o comando anterior é apenas um exemplo para um dos modelos projectados. Será necessário recorrer a um comando semelhante para recolher as restantes taxas de mortalidade projectadas e assim proceder à análise das mesmas. Como é facilmente perceptível pelo comando anterior, trata-se de recolher as projecções do modelo que contém o ajustamento do modelo *Lee-Carter* com um período de ajustamento de 1981 até 2009 e para a população total. É criado um ficheiro *.txt* com o nome de *fLCt81* e que tem como separação os espaços. O nome dado para o ficheiro assim como para os modelos fica ao critério de cada um, não havendo praticamente restrições, à excepções de alguns caracteres indesejados.

7.2. Taxas de Mortalidade

Com os valores de taxas de mortalidade projectadas pode-se proceder a uma análise mais rigorosa das projecções calculadas. Se por um lado ficou evidente que o ajustamento dos dois modelos não apresentam em termos estatísticos problemas, é necessário verificar se as projecções efectuadas por cada um deles apresentam qualquer tipo de incoerências ao longo dos anos pretendidos. Para tal análise construíram-se os habituais gráficos em escala logarítmica para as taxas de mortalidade projectadas e procura-se perceber qual o melhor modelo e período de tempo.

7.2.1. Anos 1981-2009

O primeiro período a ser analisado é o mais longo, aquele que numa avaliação de projecções comparadas com valores observados deu o maior EPAM. É o período também que poderá conter mais “erros de evolução”, quer isto dizer, erros que se possam explicar pelo facto de ter havido uma evolução espantosa em termos de medicina e de saúde humana, especialmente no que concerne às taxas de mortalidade infantil e juvenil. O facto de os anos 80 fazerem parte do período de ajustamento pode levar a erros nas projecções que noutros períodos possam não acontecer.

7.2.1.1. População Total

Todo este projecto foi efectuado para as três séries de populações, população total, população feminina e população masculina. No entanto, o destaque será para a população total, não deixando de fazer a análise às restantes populações, se bem que numa forma mais ligeira.

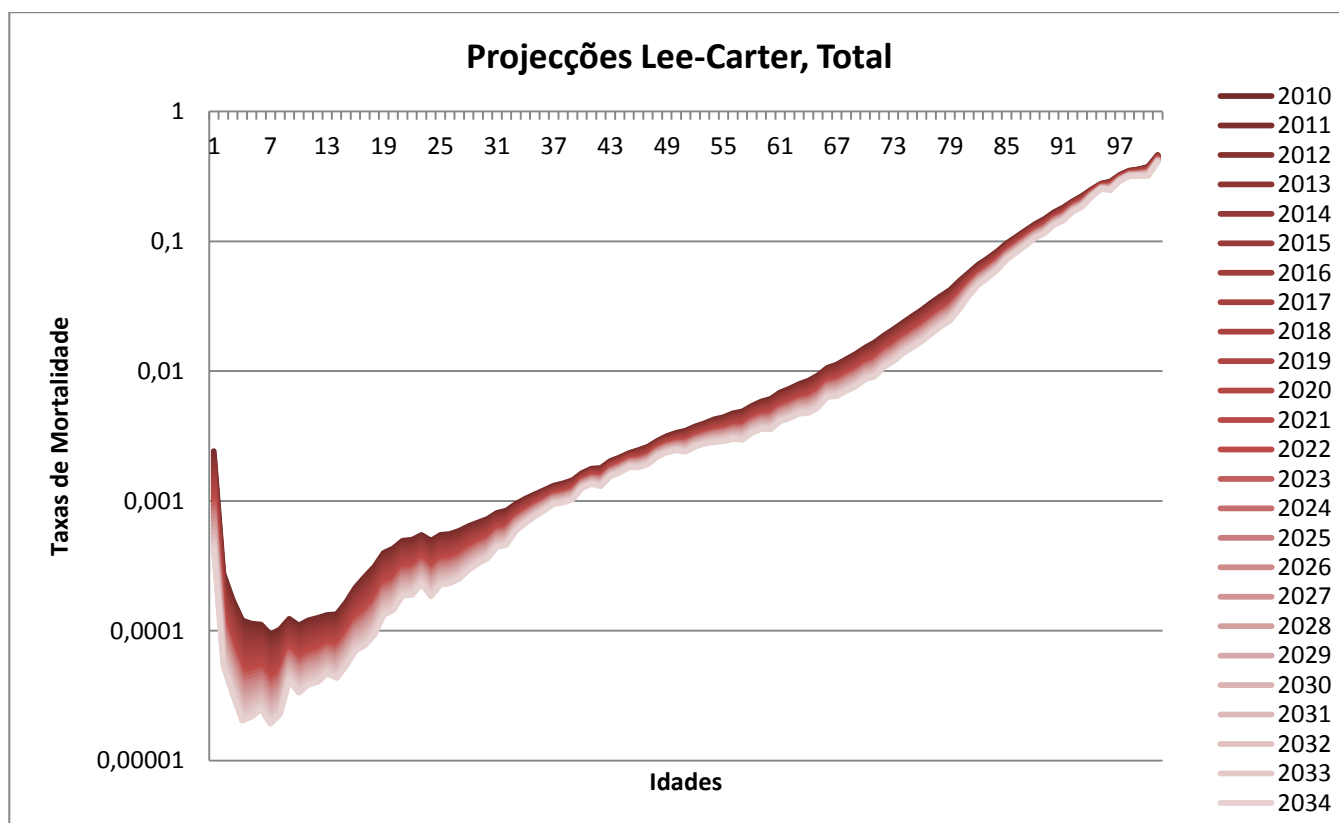
Modelo Lee-Carter

Gráfico 34 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1981 a 2009 para população total.

Como se pode verificar através do gráfico 34 as projecções efectuadas para as taxas de mortalidade, pelo ajustamento dum modelo *Lee-Carter* com um período de ajustamento de 1981 a 2009, não apresentam nenhum aspecto que possa ser apontado como problemático. Verifica-se uma taxa de mortalidade mais elevada no primeiro ano de idade, havendo uma significativa descida nas idades mais jovens. Observa-se também pronunciada a “lomba” da sinistralidade rodoviária que afecta os jovens em Portugal e por fim existe um aumento gradual da taxa de mortalidade ao longo dos anos. De realçar

os poucos ou quase inexistente “picos” no gráfico. O que sublinha o alisamento criado pelo modelo e que no fundo é pretendido para projecções de qualidade.

Enfatizar também o facto de ao longo dos 25 anos de projecções não haver nenhuma mudança brusca nas taxas de mortalidade projectadas, havendo uma melhoria gradual das taxas o que acaba por ser expectável dada a evolução científica e médica dos últimos anos e que se espera continuar a acontecer.

Modelo Booth-Maindonald-Smith

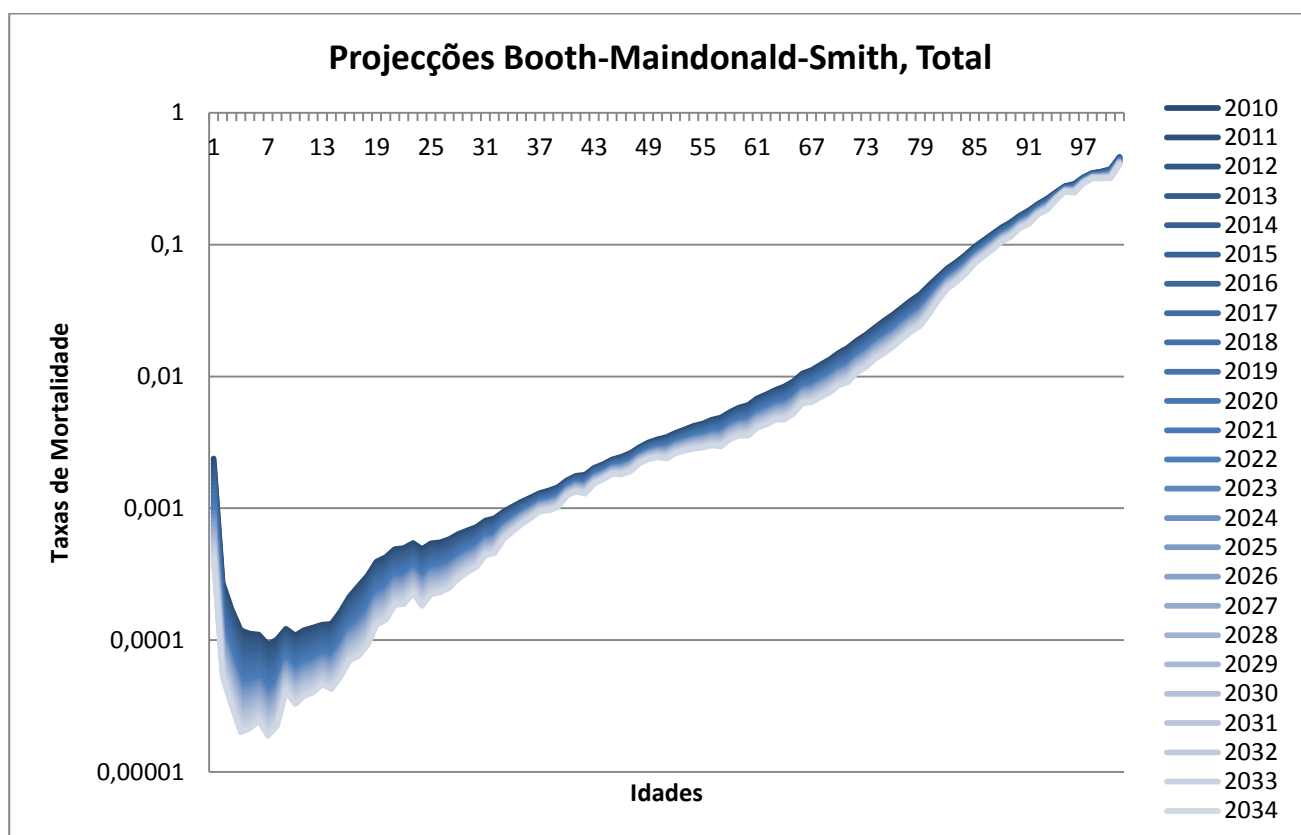


Gráfico 35 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1981 a 2009 para população total.

Observando o gráfico 35 das projecções efectuadas para o mesmo período de ajustamento e para a mesma população, mas para o modelo *Booth-Maindonald-Smith*, verifica-se uma semelhança bastante significativa dos dois modelos. Embora os valores das taxas de mortalidade projectadas não sejam exactamente as mesmas, para este período de ajustamento o gráfico, o andamento e o comportamento, dos dois modelos

são praticamente idênticos. Não se registando diferenças dignas numa avaliação desta natureza.

7.2.1.2. População Feminina e Masculina

Para as populações divididas por sexos, e visto não ser o focus do trabalho, será feita uma análise mais superficial.

De um modo geral, e analisando os gráficos da mesma forma efectuada para a população total, é possível verificar que tanto para a população feminina como para a população masculina, o comportamento dos dois modelos mantém-se similar ao verificado aquando da análise da população total. Ou seja, não existem, neste período de ajustamento, grandes diferenças entre os dois modelos analisados para qualquer uma das populações. (Anexos 50 a 53)

7.2.2. Anos 1990-2009

Este será o período intermédio, de 20 anos. Trata-se de um período de ajustamento que não se ficando apenas com os últimos anos de evolução médico-científica, apresenta também alguma informação histórica que poderá ser relevante para as projecções.

7.2.2.1. População Total

Analisando então os dois modelos ajustados para um segundo período de ajustamento e para a população de ambos os sexos.

Modelo Lee-Carter

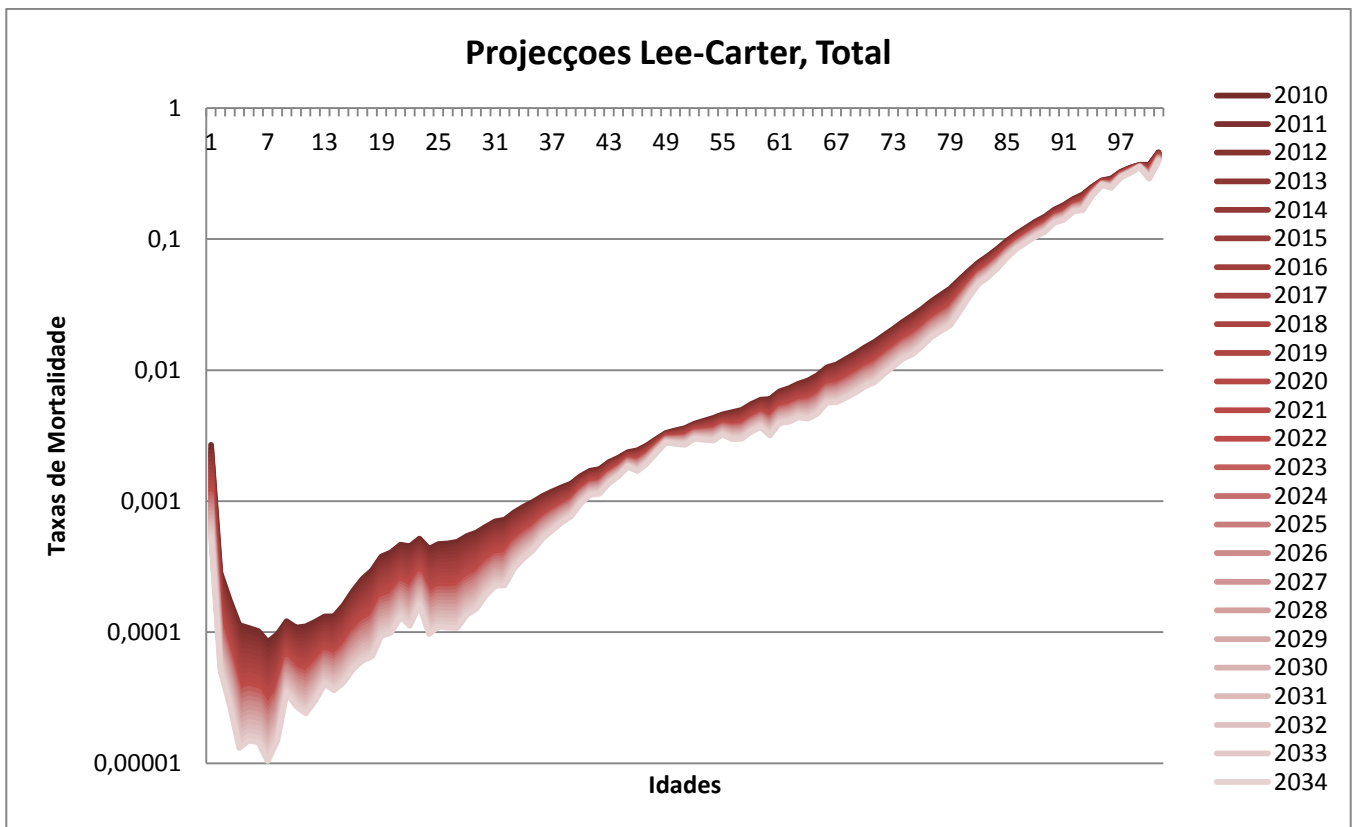


Gráfico 36 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1990 a 2009 para população total.

Mais uma vez se verifica, assim como para o período anterior, um alisamento bastante positivo nas taxas de mortalidade projectadas pelo modelo *Lee-Carter* (gráfico 36). Nota-se uma maior amplitude de valores, quando comparados com o período anterior, mas no geral o comportamento assemelha-se em muito com o período de ajustamento com os anos de 1981 a 2009.

Contudo há que destacar a maior diferença entre ambos e que não sendo absolutamente decisiva é de realçar. Os valores das projecções estão inferiores, parecendo que o gráfico sofreu um pequeno mas visível deslocamento para baixo. Os valores das projecções para os diversos anos continuam a melhorar, mas verifica-se também um afastamento superior. Ou seja, a diferença de taxa de mortalidade no ano de 2010 e de 2034 é superior à diferença apresentada no período anterior em quase todas as idades.

De um modo geral, pode-se afirmar que mais uma vez o modelo *Lee-Carter* apresenta curvas de mortalidade bastante positivas e sem denotar qualquer possível problema ou valor inesperado.

Modelo Booth-Maindonald-Smith

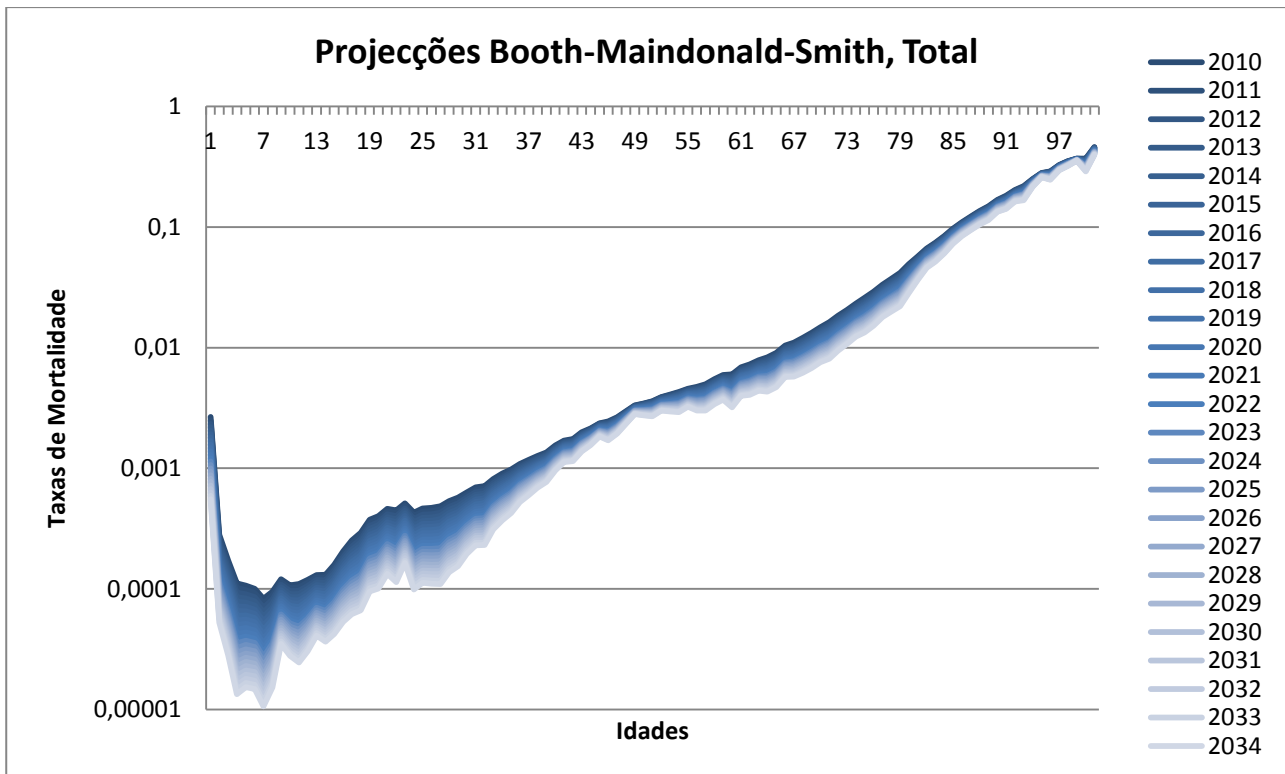


Gráfico 37 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1990 a 2009 para população total.

A análise realizada para as curvas de mortalidade criadas através das projecções do modelo *Booth-Maindonald-Smith* não difere muito das realizadas para o modelo *Lee-Carter*, quando comparados os períodos de ajustamento. Ou seja, as taxas de mortalidade do gráfico 37 parecem ter um comportamento idêntico às do período anterior para o mesmo modelo, ressalvando a maior amplitude entre o primeiro e último ano de projecção. De destacar que, neste caso, o valor das taxas de mortalidade não sofrem grandes diferenças como nos modelos *Lee-Carter* para diferentes períodos, ou seja, o valor das taxas para o período de ajustamento de 1981 a 2009 é bastante similar ao período de ajustamento de 1990 a 2009 para o modelo *Booth-Maindonald-Smith*.

Quanto a comparação, é realizada para os dois modelos, continuam a verificar-se as semelhanças que existiram para os modelos no período anterior.

7.2.2.2. População Feminina e Masculina

A divisão por sexos volta a apresentar curvas de mortalidade bem mais “oscilantes” que as anteriormente observadas com a população total.

Comparando apenas períodos de ajustamento, pode-se verificar que de um modo geral, quer para o modelo *Lee-Carter*, quer para o modelo *Booth-Maindonald-Smith*, as curvas de mortalidade apresentadas são bastantes semelhantes às do período anterior. De realçar que, mais uma vez, no modelo *Lee-Carter* os valores das taxas parecem estar ligeiramente inferiores ao período de ajustamento de 1981 a 2009.

Quando se compara as populações, verifica-se que, especialmente para a população masculina, existem muitas mais oscilações e “picos” nos gráficos do que na população total ou feminina.

Observando agora os dois modelos, verifica-se novamente uma semelhança entre os mesmos. (Anexos 54 a 57)

Não parece haver uma grande diferença para estes dois primeiros períodos, quer ao nível de população quer ao nível de modelos.

7.2.3. Anos 2000-2009

O ajustamento com os anos entre 2000 e 2009 será o terceiro e último período analisado. Por um lado é o período que melhor se adequa em termos evolutivos, pois apresenta os valores de taxas de mortalidade mais recentes, por outro acaba por ser aquele que tem menos informação, sendo o mais curto de todos com apenas 10 anos de ajustamento.

7.2.3.1. População Total

Em seguida serão analisados os dois modelos utilizados para o último período de ajustamento para o total da população.

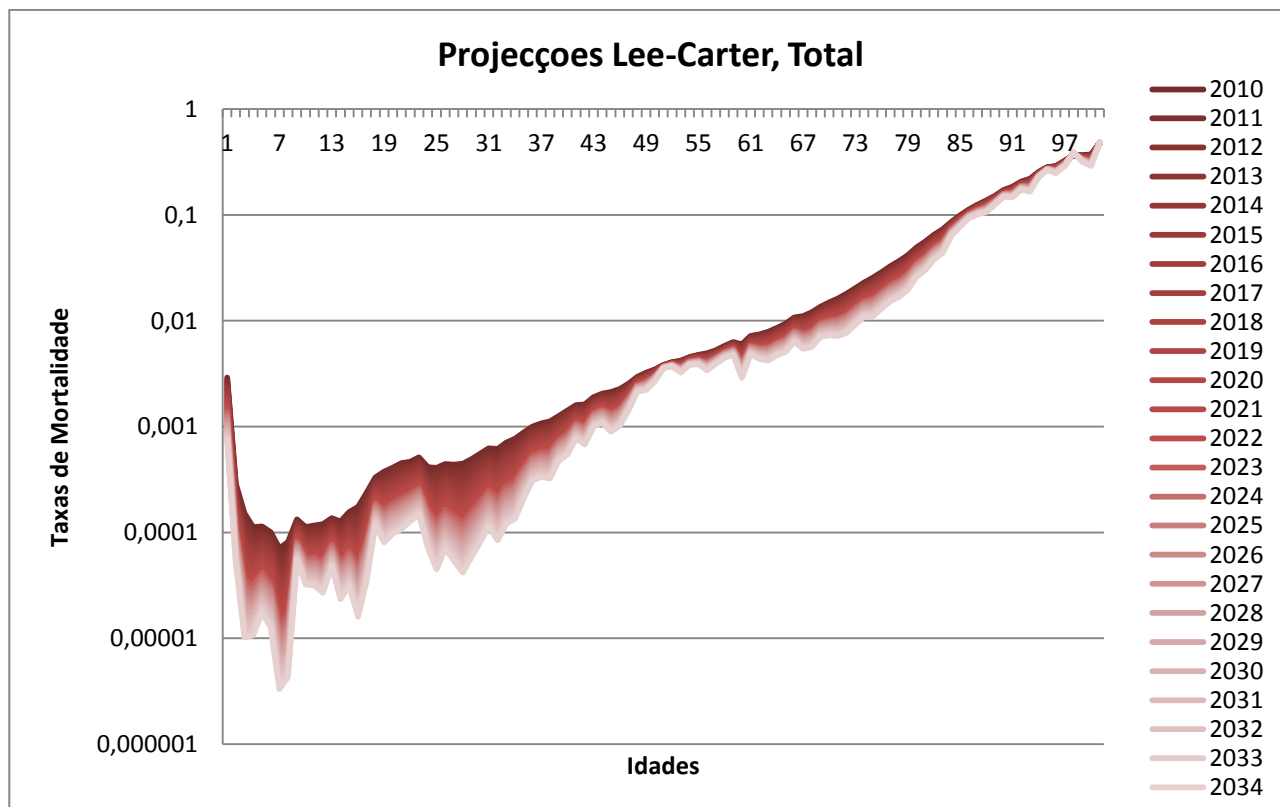
Modelo Lee-Carter

Gráfico 38 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 2000 a 2009 para população total.

As diferenças para com os outros períodos de ajustamento começam a ser evidentes. Verifica-se que para este período de ajustamento o alisamento, que era assinalável para o modelo *Lee-Carter*, deixou de existir. O surgimento de inúmeras oscilações e “picos” no gráfico 38 dá a entender uma maior dificuldade no modelo em projectar com precisão todos os valores da curva de mortalidade.

Realçando principalmente as idades mais jovens como as mais problemáticas, havendo numa idade uma amplitude bastante grande de valores do ano de 2010 para o de 2034 e na idade seguinte uma amplitude bastante baixa de valores, comparativamente à anterior. Também a denominada “lomba” da mortalidade rodoviária aparece bastante mais acentuada nos últimos anos de projecção

comparativamente com os primeiros anos de projecção, o que não será expectável visto que o esperado é uma melhoria das taxas de mortalidade desse aspecto específico. Realçar também o facto de em algumas idades mais velhas a taxa de mortalidade inclusivamente aumentar com os anos, o que é de todo inesperado.

A principal conclusão que se poderá tirar deste gráfico é a de que possivelmente este período não será o mais desejável de se utilizar para o projecto. No entanto há que verificar este comportamento mais indesejável para o modelo de BMS e também se se observa nas populações feminina e masculina.

Modelo de Booth-Maindonald-Smith

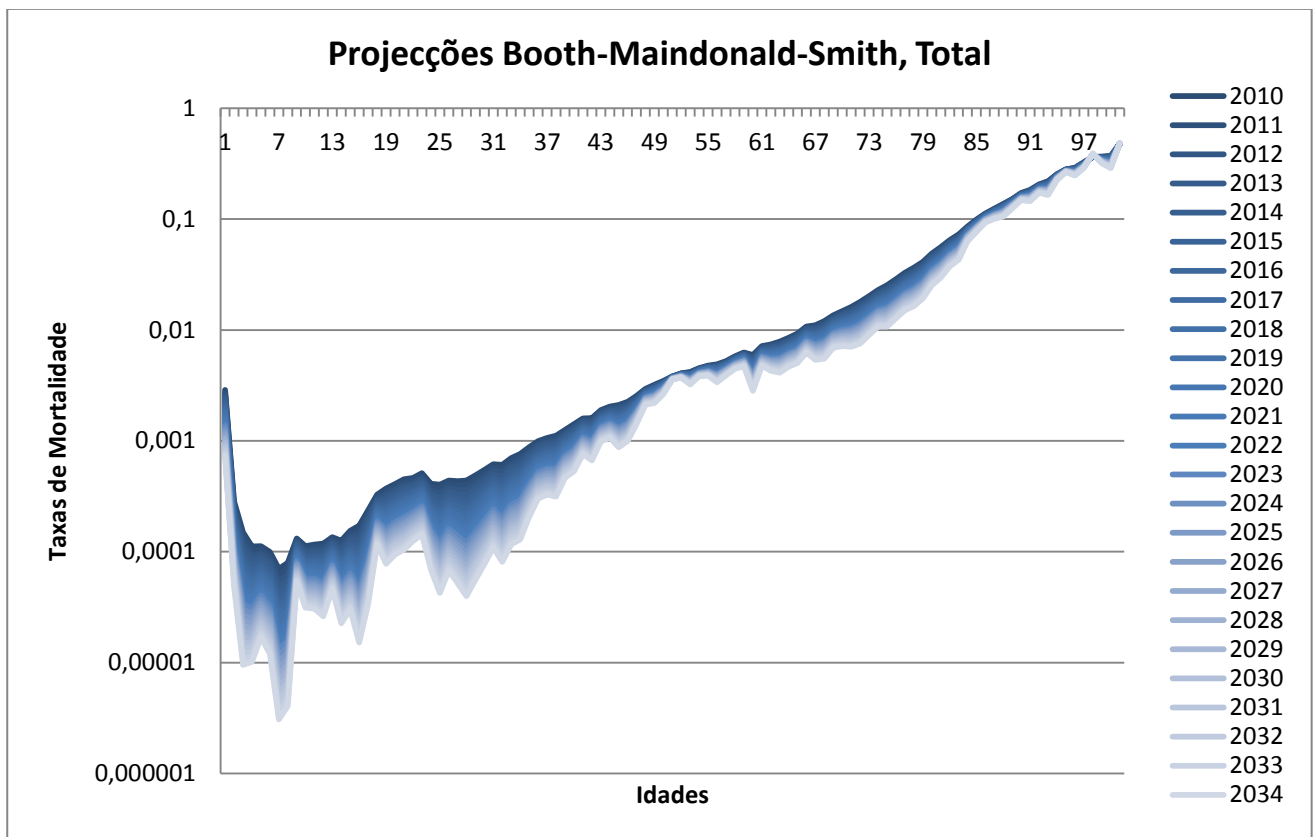


Gráfico 39 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 2000 a 2009 para população total.

De um modo geral o comportamento do modelo *Booth-Maindonald-Smith* para este período é idêntico ao modelo *Lee-Carter*. Tem algumas ligeiras diferenças, mas numa observação mais geral verifica-se que existem demasiados “picos” e oscilações nos valores das curvas de mortalidade, assim como em alguns casos de idades mais

avançadas, a taxa de mortalidade chega a aumentar com o passar dos anos projectados. (gráfico 39) É de facto algo que não será expectável de acontecer e desde logo faz soar o alarme quanto à utilização deste período de ajustamento.

7.2.3.2. População Feminina e Masculina

Fazendo a mesma análise para as duas populações divididas por sexos pode-se constatar a fragilidade dos dois modelos aquando da utilização deste período mais curto de ajustamento. (Anexos 58 a 61)

Na realidade basta observar o comportamento das curvas de mortalidade quer para a população feminina quer para a população masculina para perceber que o uso deste período é não só desaconselhado como indesejado. Nos quatro casos analisados, população feminina e masculina para os dois modelos, é visível a constante subida e descida do gráfico, não se assemelhando em nada com uma curva de mortalidade. No caso da população feminina há pelo menos duas situações em que a taxa de mortalidade aumenta com os anos projectados e em idades consideradas jovens, o que é de todo improvável face à percepção generalizada de que as taxas de mortalidade tendem a diminuir, ainda para mais em idades mais jovens.

De um modo geral, estes resultados para um período mais curto podem ser explicados pelo facto de se estar a projectar para 25 anos com um período de ajustamento de 10 anos. Algo que poderá não ser o mais aconselhado.

7.3. Óbitos

Uma análise mais específica e concreta poderá ajudar na decisão entre qual o melhor modelo e qual o melhor período de ajustamento a utilizar para a população portuguesa. Posto isto, analisou-se qual o impacto que as taxas anteriormente projectadas têm no número de total de óbitos. Deste modo, tentando “humanizar” a avaliação dos modelos ao nível do número de óbitos.

Por outro lado, realizou-se esta análise apenas para as populações femininas e masculinas. Isto porque foram nessas populações que se verificaram as maiores

oscilações e os maiores erros. Também só se realizará para os períodos entre 1981 e 2009 e 1990 e 2009, visto que o último período com menor amplitude de anos mostrou-se muito pouco viável na análise anterior.

Mais uma vez, irá realizar-se a divisão de análises por período, de forma à comparação ser mais simples de se efectuar.

7.3.1. Anos 1981-2009

Para uma melhor compreensão realizaram-se alguns gráficos com o comportamento não só dos diversos modelos como também dos seus intervalos de predição, juntamente com os valores observados nos anos anteriores às projecções.

População Feminina

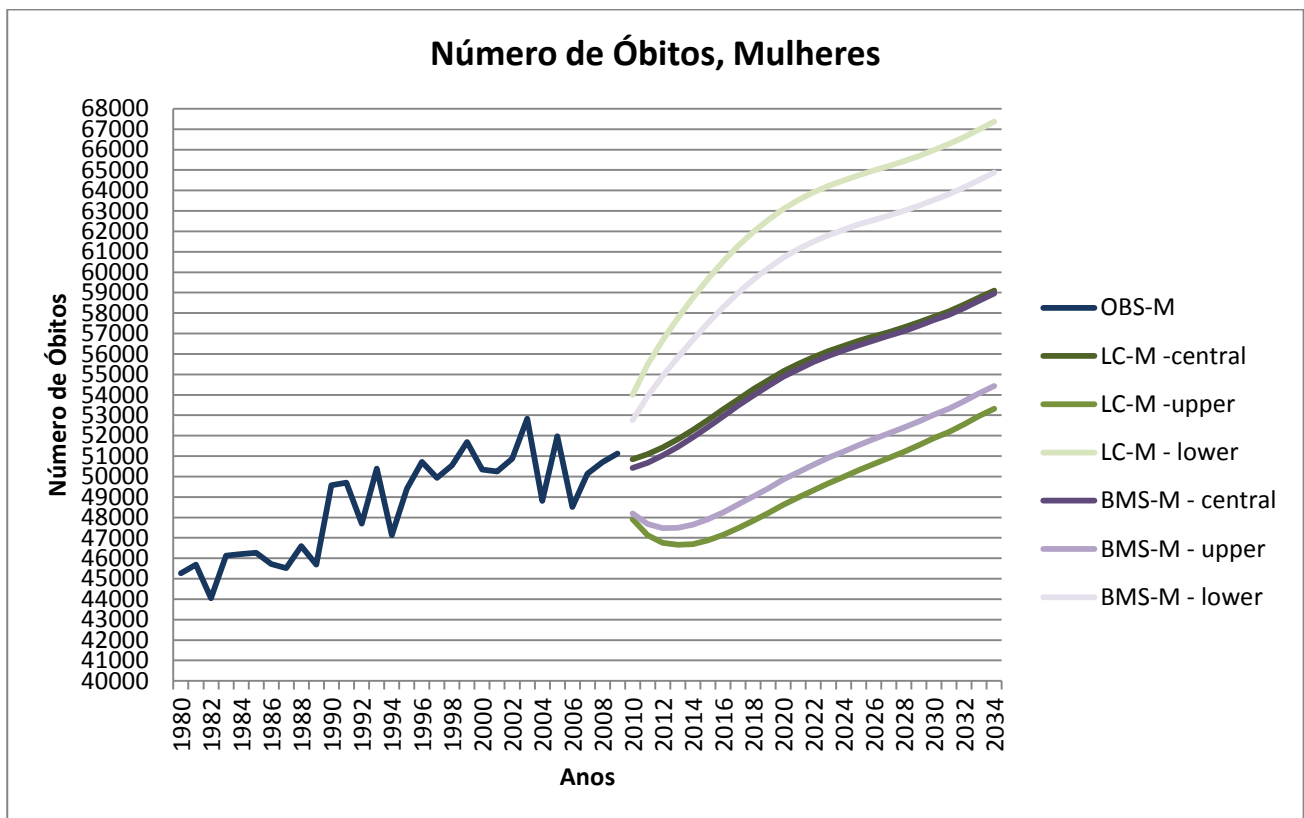


Gráfico 40 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1981 e 2009 para a população feminina, com intervalos de predição.

Como se pode verificar através do gráfico 40, existe de facto um alisamento nas projecções que não coincide com o andamento generalizado dos valores observados

anteriores aos anos projectados. No entanto, pode-se observar que o andamento dos óbitos totais da população feminina tem tendência a crescer e que as diferenças entre os valores projectados para os dois modelos não têm diferença significativa. O único ponto significativo é o facto do intervalo de predição ter uma amplitude bastante grande para os dois modelos.

Salientar que não é de todo inesperado que o número de óbitos aumente dum ponto de vista de valores na ordem de grandeza indicada pelos dois modelos, o que favorece a escolha destes modelos para uma análise à mortalidade em Portugal.

População Masculina

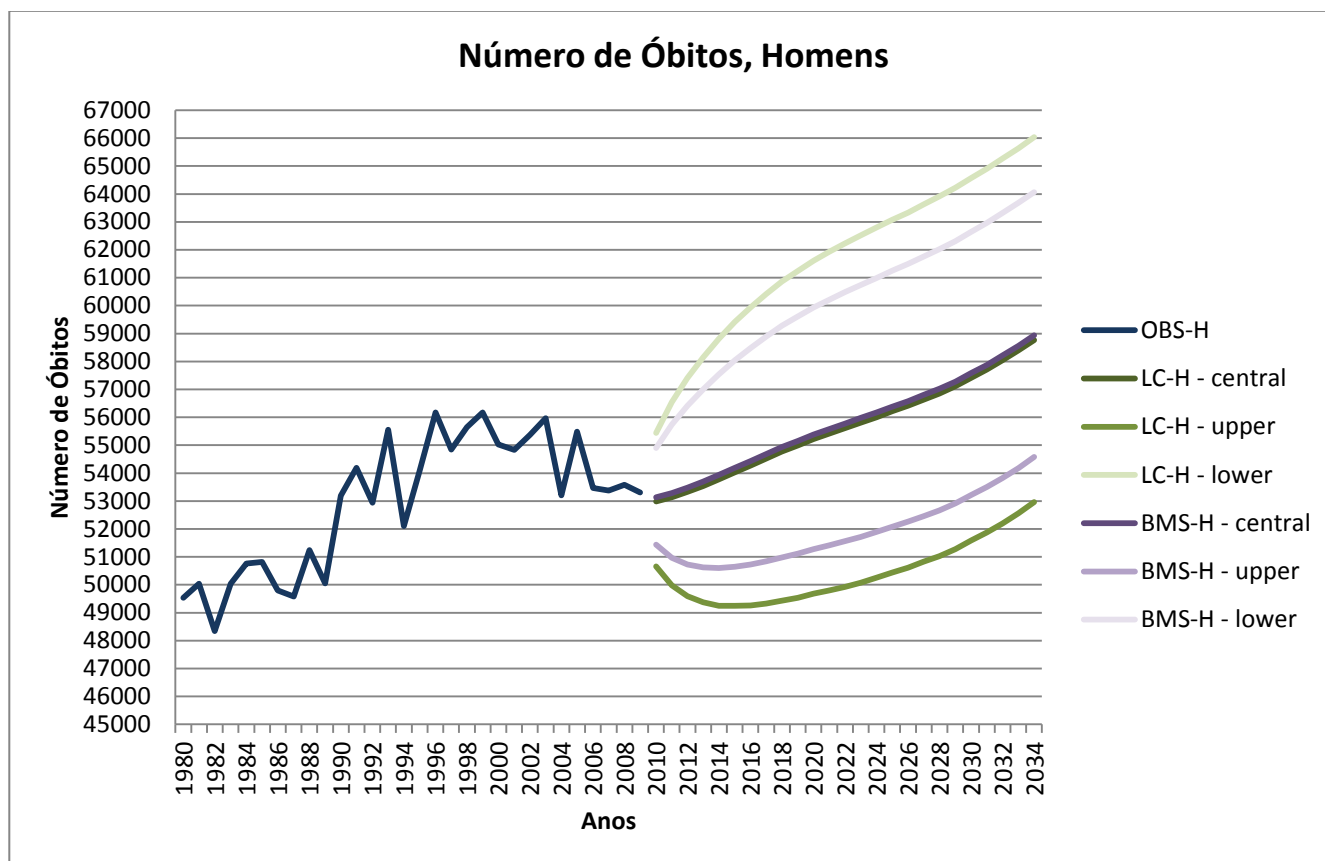


Gráfico 41 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1981 e 2009 para a população masculina, com intervalos de predição.

O número de óbitos na população masculina parece oscilar um pouco mais que no contexto da população feminina (gráfico 41). Se no caso da população feminina havia uma trajectória ascendente bastante clara ao longo dos últimos 30 anos, no caso

dos homens essa trajectória parece ter sido invertida nos últimos 10 anos. No entanto, é previsível que o número de óbitos siga o comportamento verificado pelos dois modelos, ou seja, uma subida a médio prazo. Com o aumento da população mais idosa é também natural que o número de óbitos venha a aumentar.

Nesta análise há duas características que se devem destacar: a primeira sendo a amplitude dos intervalos de predição continuar a ser grande face aos valores apresentados; a segunda, o facto dos valores das projecções dos dois modelos continuarem a ser bastante idênticos nos dois modelos.

7.3.2. Anos 1990-2009

Desta feita serão analisadas as duas populações para o segundo período em estudo.

População Feminina

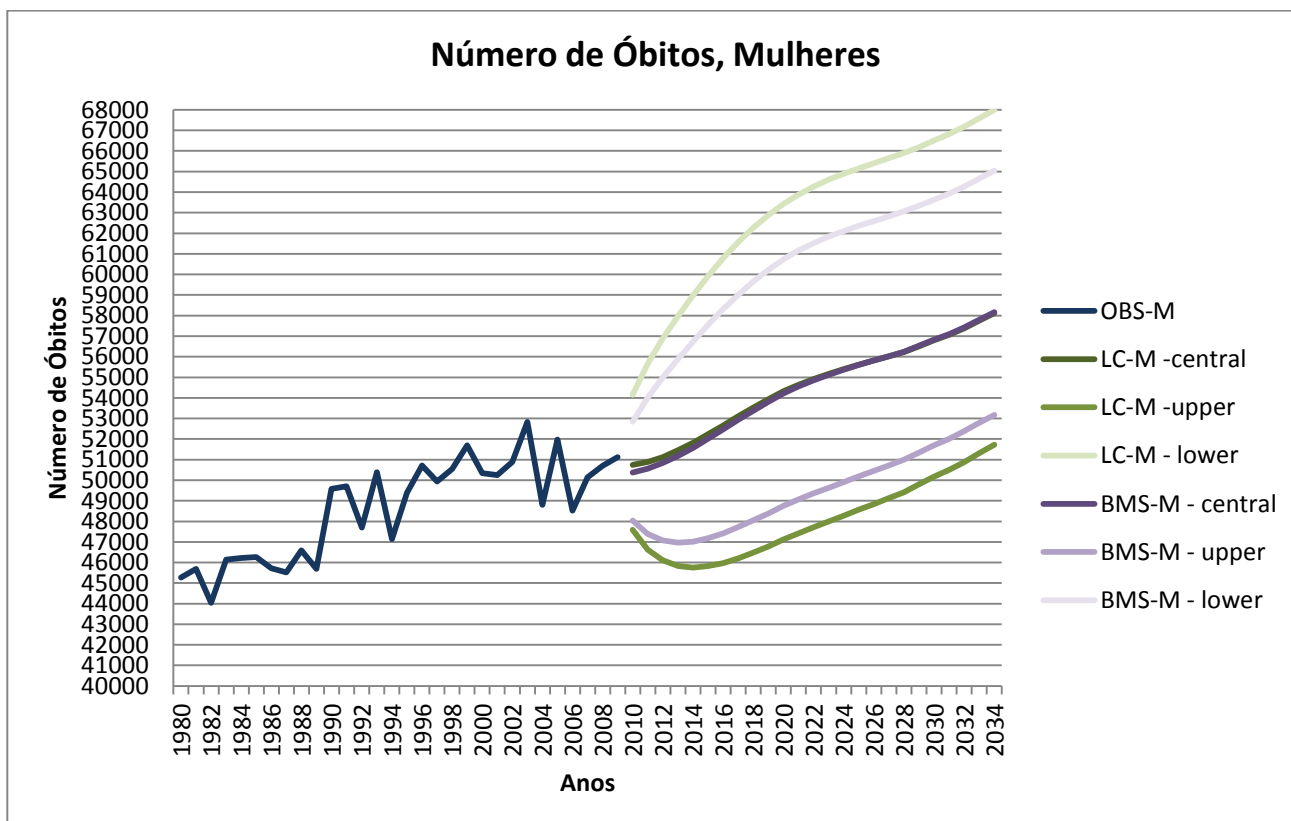


Gráfico 42 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1990 e 2009 para a população feminina, com intervalos de predição.

Neste segundo período em estudo é possível verificar que as diferenças entre os dois gráficos (gráficos 40 e 42) com o número de óbitos para a população feminina são praticamente nulas, principalmente no final das projecções. Quer em termos de andamento quer em termos de diferenças entre modelos. Os dois períodos de ajustamento parecem projectar números de óbitos bastante semelhantes para ambos os modelos, não havendo diferenças na interpretação realizada aquando do período anterior.

População Masculina

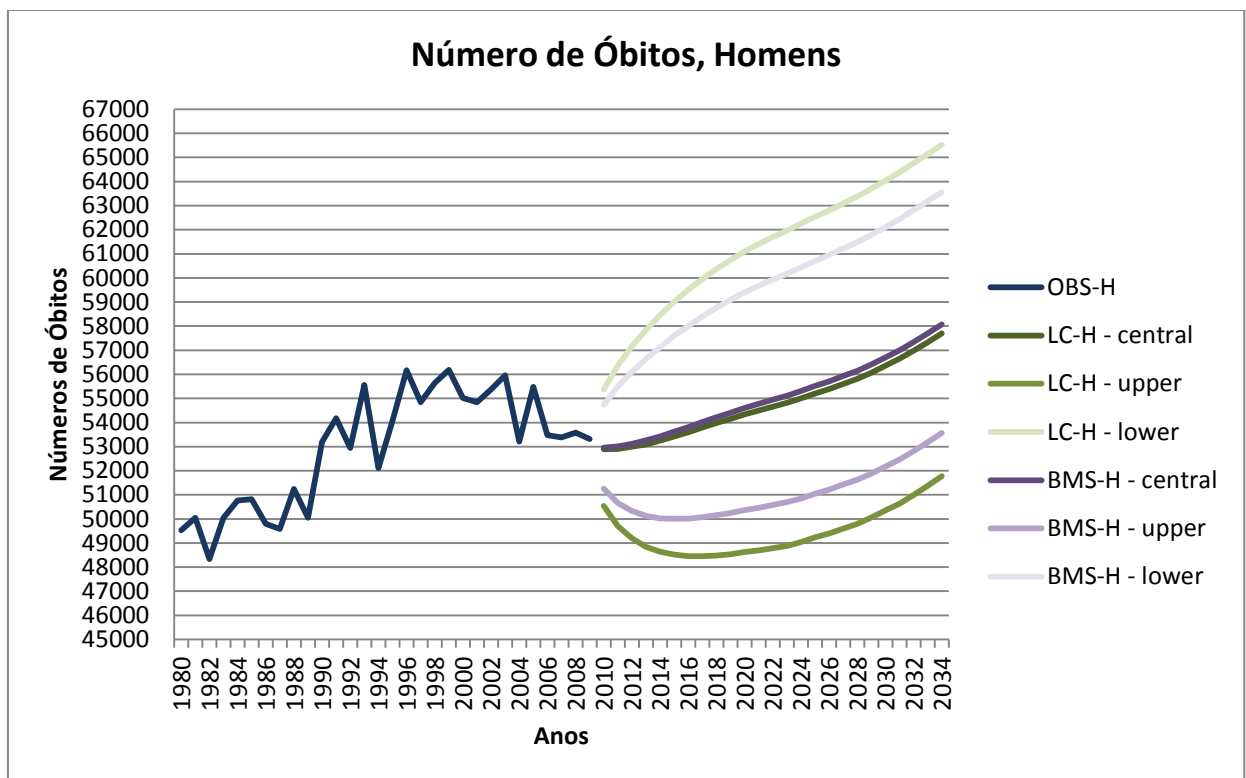


Gráfico 43 – Número de óbitos projectados pelos modelos com período de ajustamento entre 1990 e 2009 para a população masculina, com intervalos de predição.

Do mesmo modo que foi realçado na análise anterior sobre a população feminina, também na população masculina as diferenças entre gráfico 43 e o gráfico do período anterior (gráfico 41) são diminutas. No entanto, há que destacar a diferença entre populações, se na população feminina os valores dos dois gráficos tendem a aproximar-se com o prolongar das projecções, no caso da população masculina tendem a afastar-se, estando muito próximos no início do período de projecção.

Deste modo, pode-se afirmar que o número de óbitos não será um factor de diferenças referente a períodos de ajustamento, bem como nos modelos ajustados. Salientar, no entanto, que confirmam as boas indicações dadas pelos modelos nas anteriores análises.

7.4. Esperança de Vida

Outra análise também mais específica do que as taxas de mortalidade que se pode fazer é o cálculo da esperança de vida. Este cálculo possibilita mais um factor de decisão assim como de informação no que à escolha dos modelos e períodos diz respeito. Mais uma vez, esta análise só será realizada para a população feminina e masculina, assim como para os dois períodos de ajustamento mais longos, de 20 e 30 anos. Os gráficos para além dos valores de projecções e intervalos de predição, terão também os valores observados para os anos anteriores assim como as previsões efectuadas pelo EUROSTAT.

7.4.1. Anos 1981-2009

População Feminina

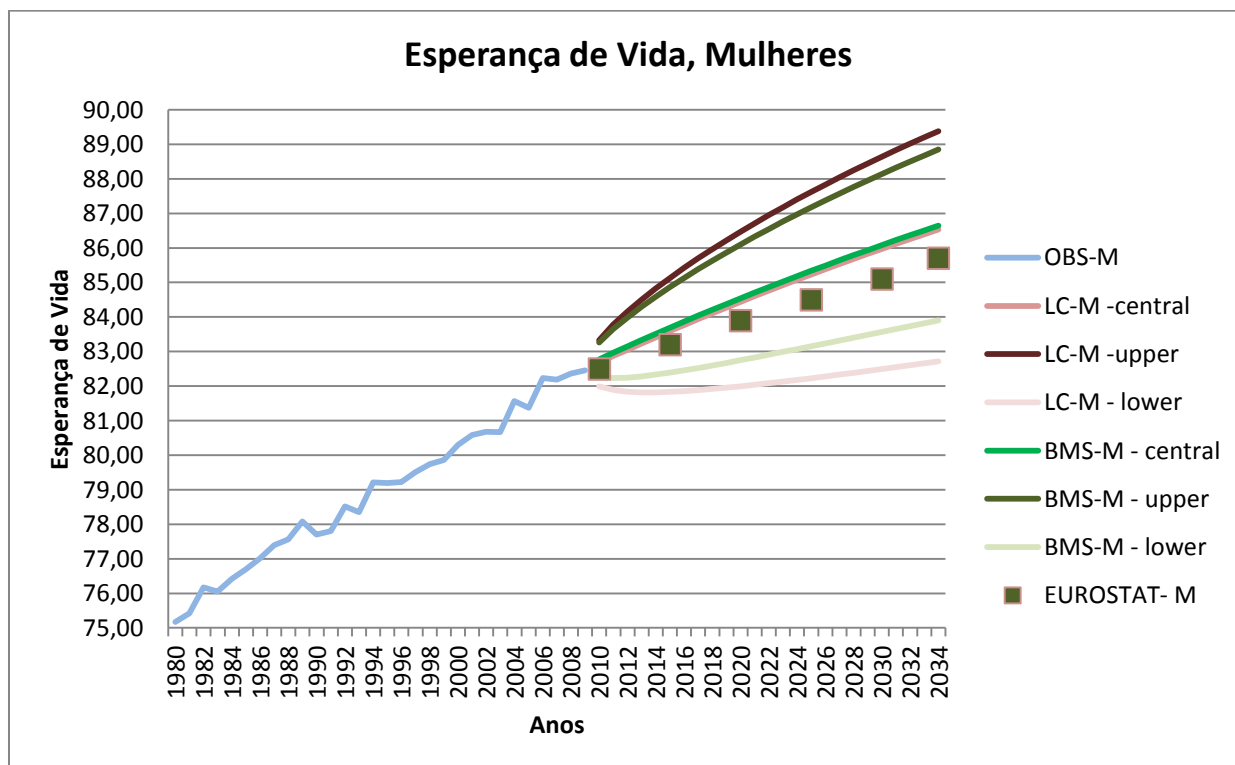


Gráfico 44 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1981 e 2009 para a população feminina, com intervalos de predição.

Um dos aspectos mais relevantes no gráfico 44 é o facto dos valores de esperança de vida serem bastante semelhantes em ambos os modelos. A única diferença que no gráfico é perceptível será o facto da diferença nos intervalos de predição aumentar ao longo dos anos de projecção.

De resto constata-se que as diferenças entre as projecções feitas pelos dois modelos em causa e o EUROSTAT não são muito grandes, ficando nos dois casos a menos de 1 ano de diferença, o que numa projecção a 25 anos não parece ser muito grande.

Relativizando para os valores observados, pode verificar que nos últimos 30 anos se mantém uma curva bastante linear de aumento da esperança de vida, o que nos dois modelos se mantém e com um declive bastante parecido, o que indica uma boa aproximação dos modelos à realidade.

População Masculina

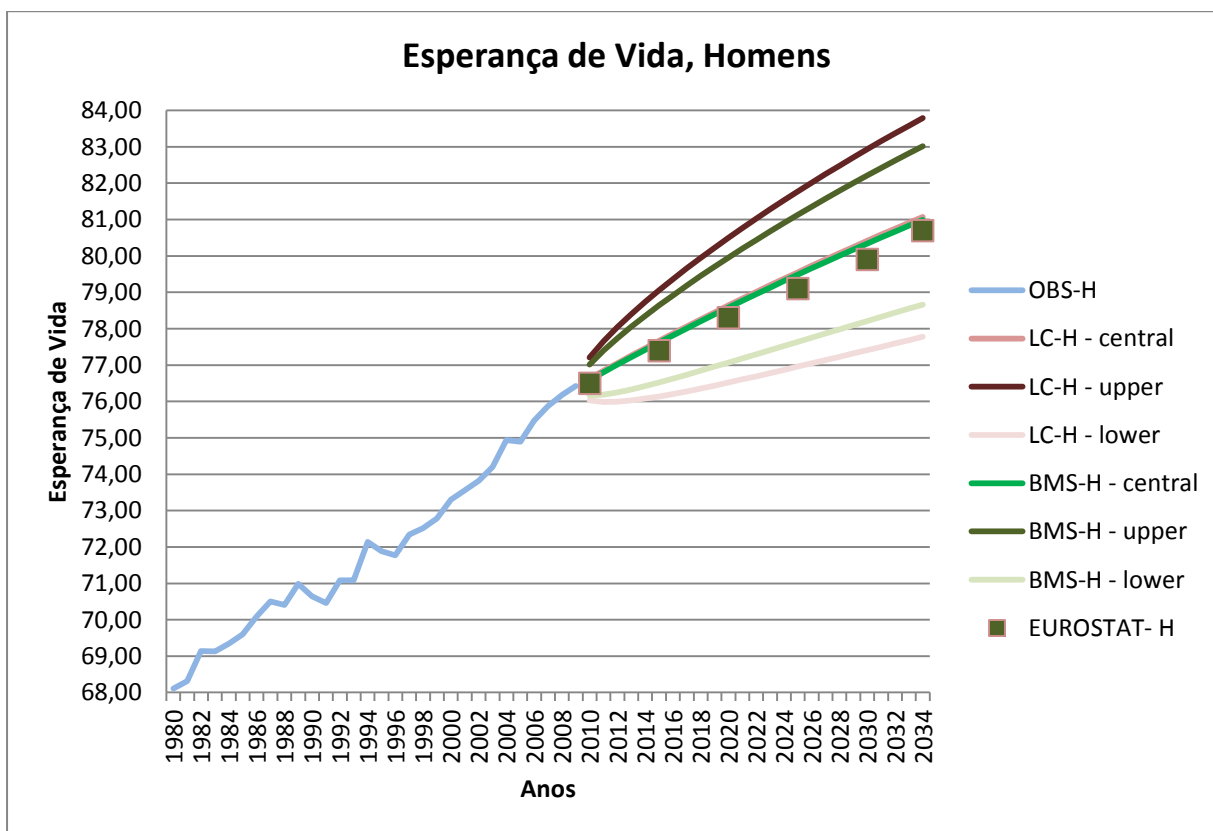


Gráfico 45 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1981 e 2009 para a população masculina, com intervalos de predição.

Observando o gráfico 45 com a esperança de vida para a população masculina nos diversos modelos e para um período de ajustamento de 30 anos, verifica-se que as diferenças entre todos os valores apresentados no gráfico são menores. Os valores do EUROSTAT coincidem praticamente com os valores projectados pelos modelos, ficando a menos de 6 meses ao fim de 25 anos de projecções. E os valores dos limites dos intervalos de predição são menos amplos do que na população feminina.

Mais uma vez, quando comparados com os valores observados nos últimos anos, salienta-se o facto do comportamento dos valores projectados estarem próximos dos expectáveis, aumentando a esperança de vida em alguns anos comparativamente com os últimos valores registados em 2009.

7.4.2. Anos 1990-2009

De um modo geral a análise para o período de ajustamento entre 1990 e 2009 não tem grandes diferenças quando comparado com o período entre 1981 e 2009 que foi descrito anteriormente. (Anexos 62 e 63)

Na tabela 4 pode-se observar os valores de esperança de vida e verificar que são bastante parecidos entre populações idênticas.

	1981 - 2009				1990 - 2009			
	Mulheres		Homens		Mulheres		Homens	
	LC	BMS	LC	BMS	LC	BMS	LC	BMS
2010	82,68	82,77	76,63	76,59	82,72	82,80	76,72	76,70
2015	83,59	83,69	77,67	77,62	83,74	83,81	77,91	77,85
2020	84,44	84,54	78,64	78,58	84,69	84,74	79,00	78,91
2025	85,23	85,34	79,55	79,48	85,56	85,60	80,00	79,89
2030	85,97	86,08	80,41	80,34	86,37	86,40	80,93	80,80
2034	86,53	86,64	81,07	80,99	86,97	87,00	81,63	81,48

Tabela 4 – Esperança de vida projectada para os diferentes períodos, sexos e modelos.

7.5. Síntese Final

Realizando-se três análises distintas, curvas de mortalidade, óbitos e esperança de vida, pôde-se constatar que de um modo geral os dois modelos são bastante semelhantes. Os valores projectados para as taxas de mortalidade são semelhantes e o comportamento ao longo dos 25 anos de projecção também.

No que concerne a períodos de ajustamento viram-se diferenças significativas entre os dois primeiros períodos, com mais anos de ajustamento, em relação ao período de ajustamento mais curto, com vantagem para os períodos mais longos. Mas nos dois casos mais longos globalmente os valores não variam muito.

De realçar que as principais diferenças nos períodos, populações e modelos foram verificadas na análise às curvas de mortalidade. No número de óbitos e nos valores de esperança de vida, essas diferenças foram muito diminutas. Esse facto pode ser explicado por se usar apenas dois períodos de ajustamento nessas análises.

8. Análise de Sensibilidade

A realização de uma análise de sensibilidade revela-se quase sempre muito útil neste tipo de estudos. Neste caso, a análise de sensibilidade procura determinar o efeito de diferentes estimativas, associadas a erros de magnitudes diferentes, no resultado final da projecção. É uma importante forma de calcular a relevância de uma variável sobre o resultado de outra.

Nesta secção, procura-se determinar qual o verdadeiro impacto dos erros das taxas de mortalidade ajustadas nos óbitos da população. Em particular, pretende-se verificar se alguns erros de maior valor, que surgiram nos ajustamentos dos modelos analisados, têm um impacto grande na estimativa da população real, ou se por outro lado são erros com pouco significado no resultado final.

Para esta fase da dissertação analisaram-se os valores de erros associados ao longo dos anos de ajustamento de cada modelo. De um modo geral os maiores valores de erro encontravam-se nos últimos anos de ajustamento e portanto a escolha teria de recair sobre um desses anos. Para além disso verificou-se que havia um valor bastante alto de erro no ano de 2008.

Assim, o que se decidiu foi realizar a análise de sensibilidade para o ano de 2008, especificamente para o período de ajustamento de 1981 a 2009. Ou seja, calcular a população através das taxas de mortalidade ajustadas pelos dois modelos e comparar essa população com a população calculada através das taxas de mortalidade observada, calculando os erros associados. Com isto, o efeito das migrações não terá impacto nos valores desses mesmos erros. De referir que se entendeu ser suficiente a realização desta análise de sensibilidade apenas para a população total e não para a população dividida por sexos, porque o que interessa é a globalidade dos possíveis candidatos ao Ensino Superior.

Para melhor entender o porquê da escolha deste ano de ajustamento, pode observar-se o gráfico 46, para a população total, dos erros associados às taxas de mortalidade ajustadas para o ano de 2008, tanto para o modelo *Lee-Carter* como para o modelo *Booth-Maindonald-Smith*.

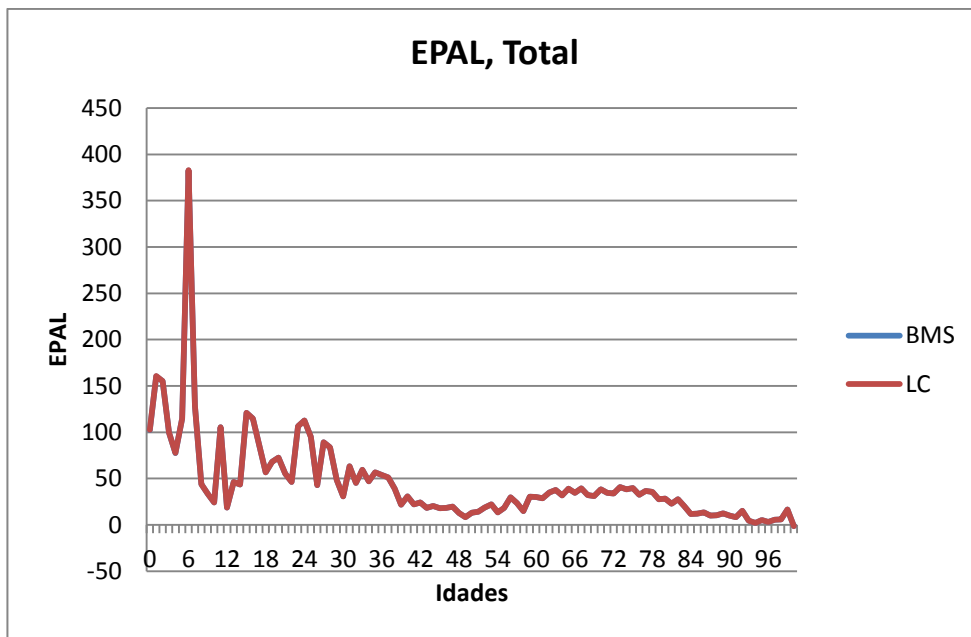


Gráfico 46 – Erro Percentual Algébrico para o total da população no ano de ajustamento de 2008.

Como se pode verificar existem valores na ordem dos 100%, chegando a observar-se um valor acima dos 350%, que justifica opção por este ano. Os erros do modelo *Lee-Carter* e *Booth-Maindonald-Smith* são praticamente idênticos, portanto a análise de sensibilidade recairá apenas num dos modelos, optando-se pelo *Lee-Carter*.

O cálculo da população através das taxas de mortalidade ajustadas pelo modelo é efectuado com o auxílio de tábuas de mortalidade, através das probabilidades de sobrevivência em anos completos. Com essas probabilidades, os nascimentos de 2008 e 2009 e com a população média para o ano de 2008 consegue-se calcular qual a população estimada do ano de 2009.

O mesmo tipo de cálculo é executado para estimar a população através das taxas de mortalidade observadas. Esta população servirá como população base, como se fosse a população observada, caso não existissem migrações. Assim poder-se-á comparar a população calculada com base nas taxas de mortalidade ajustadas no modelo com a população calculada através das taxas de mortalidade observadas.

Calculando o Erro Percentual Algébrico para cada idade do respectivo ano de 2008, ocorrido entre as duas populações anteriormente referidas, chega-se ao gráfico 47.

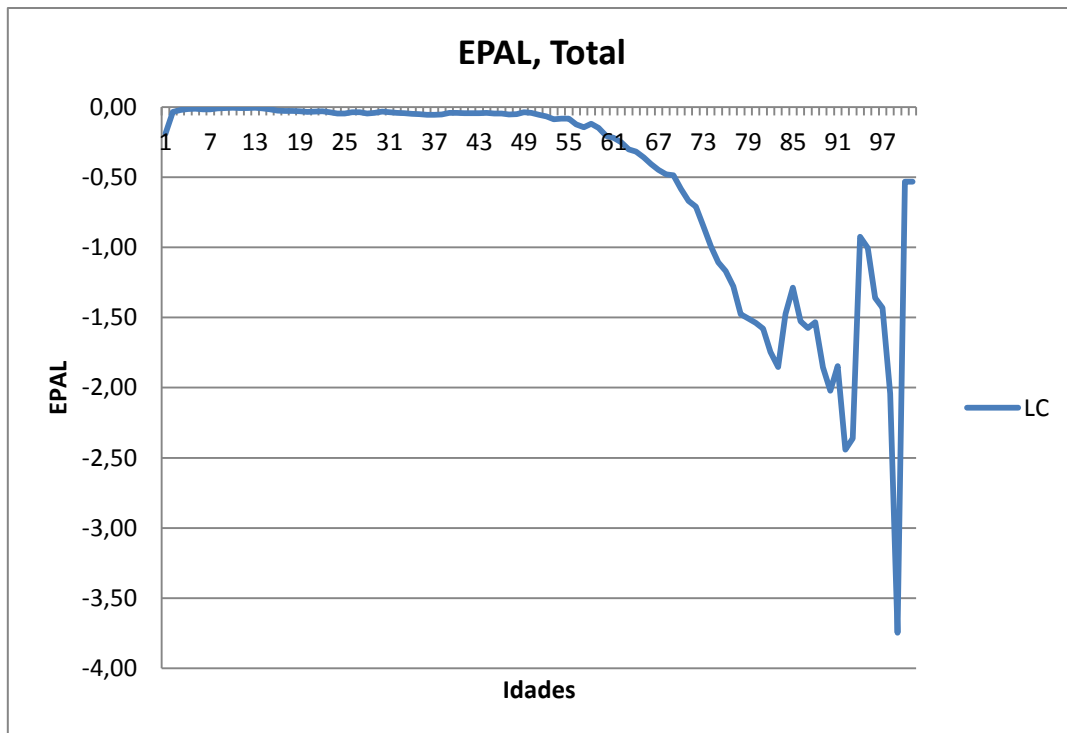


Gráfico 47 – Erro Percentual Algébrico das populações totais calculadas para o ano de 2008.

A característica mais evidente do gráfico é a evidente maior subvalorização da população estimada pelo modelo nas idades mais velhas. Por outras palavras, as taxas de mortalidade ajustadas são sobrevalorizadas o que faz com que a população tende a ser maior do que aquilo que o modelo ajusta.

De um modo global, verifica-se que o valor de Erro Percentual Algébrico Médio para as populações calculadas situa-se perto dos $-0,5\%$, o que é um erro praticamente residual. Mais uma vez, indo de encontro ao expectável. Ou seja, se nos modelos existe uma sobrevalorização das taxas de mortalidade ajustadas, necessariamente terá de haver uma subvalorização na população. Se se ajusta mais mortes do que as ocorridas, ajusta-se consequentemente menos pessoas vivas.

A análise de sensibilidade prossegue através do cálculo de populações através de taxas de mortalidade observadas acrescidas de percentagens de erro (5, 10, 15, 20, ... 350%), para observar qual o impacto na população de possíveis erros nas taxas de mortalidade. Sem perda de generalidade, optou-se por realizar apenas acréscimos nas taxas de mortalidade observadas.

De um modo geral os resultados foram bastante animadores. Veja-se os próximos três gráficos para três percentagens de erros diferentes ao longo de todo o ano.

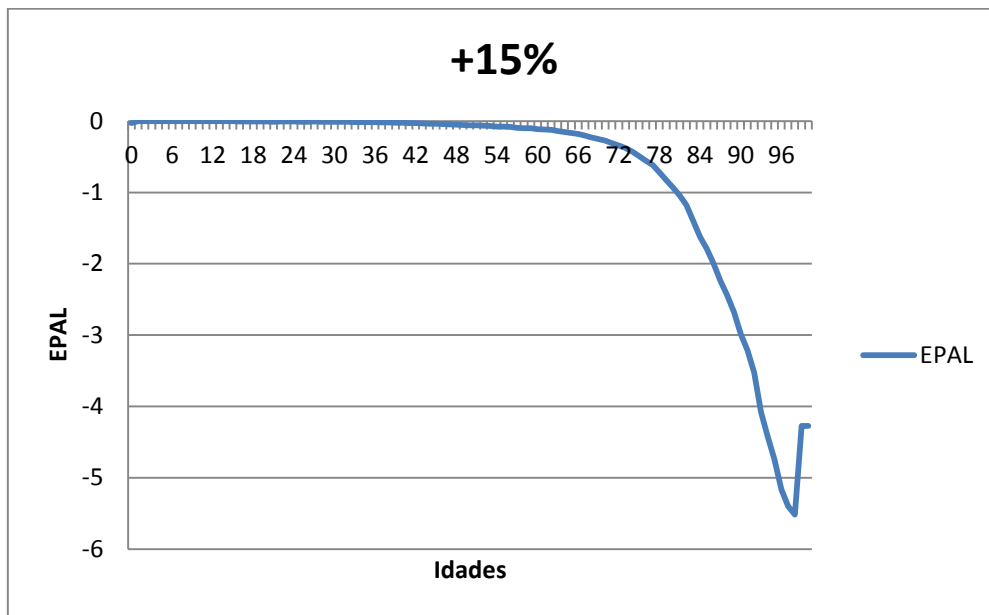


Gráfico 48 – EPAL, para as populações com acréscimo de 15% na mortalidade observada para o ano de 2008.

O gráfico 48 exprime quais seriam os erros associados à população calculada através das taxas de mortalidade observadas com um acréscimo de 15%. Tendo em conta que o Erro Percentual Algébrico Médio se situava sempre em valores inferiores, para praticamente todos os modelos, é interessante verificar que se o erro fosse o mesmo em todas as idades, e de 15%, a população teria no máximo um erro associado na ordem dos 5,5% em termos absolutos. E de constatar que seriam sempre valores abaixo de 1% para idades inferiores a 80 anos.

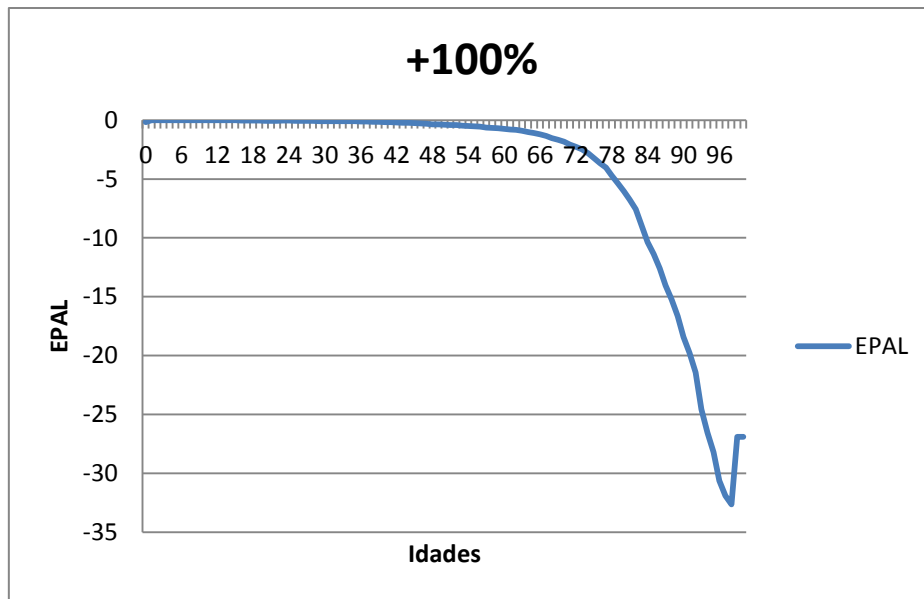


Gráfico 49 – EPAL, para as populações com acréscimo de 100% na mortalidade observada para o ano de 2008.

No gráfico 49 regista-se o comportamento do Erro Percentual Algébrico caso o erro verificado nas taxas de mortalidade ajustadas fosse de 100%, ou seja, que duplicasse o valor das taxas de mortalidade observadas. Pode-se observar que, mais uma vez, só a partir de uma idade mais avançada é que o valor absoluto do EPAL aumenta consideravelmente. Só mesmo para idades superiores a 80% é que o valor do EPAL se situa acima dos 5% (valor absoluto). Ficando sempre abaixo dos 35% de EPAL em valores absolutos.

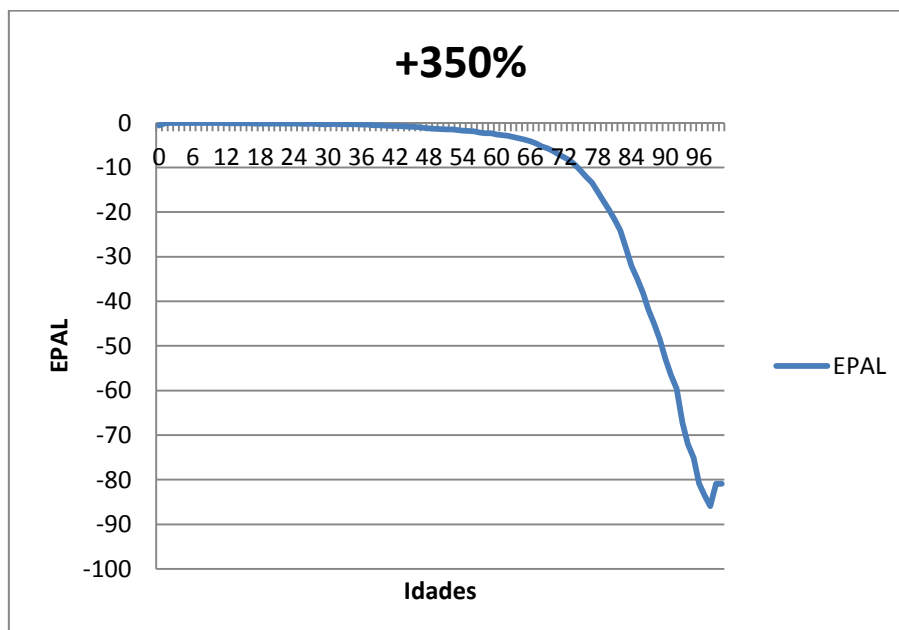


Gráfico 50 – EPAL, para as populações com acréscimo de 350% na mortalidade observada para o ano de 2008.

Tendo em conta que no ano de 2008 foi observado um valor de erro acima dos 350% numa das idades, é com naturalidade que se analisou qual o impacto de um erro dessa magnitude para todas as idades. Pode-se verificar, através do gráfico 50, que mais uma vez, só nas idades mais avançadas é que o EPAL cresce consideravelmente em termos absolutos. Até perto dos 70 anos de idade, o valor de EPAL associado a um erro de 350% nas taxas de mortalidade observadas era sempre menor do que 5%, em valores absolutos. Aumentando bastante nas idades mais altas, podendo atingir cerca de 90%.

Posto isto, e como os erros de maior valor no ano estariam todos nas idades mais jovens, analisando os diferentes gráficos já expostos, é sensato afirmar que os erros nas idades mais jovens pouco influenciam na população final.

Estando este projecto assente numa óptica de analisar principalmente os jovens em idade de entrar para o ensino superior, e tendo a idade dos 18 anos como patamar de extrema importância, resta analisar quais os impactos dos erros da mortalidade nessa idade em concreto na população.

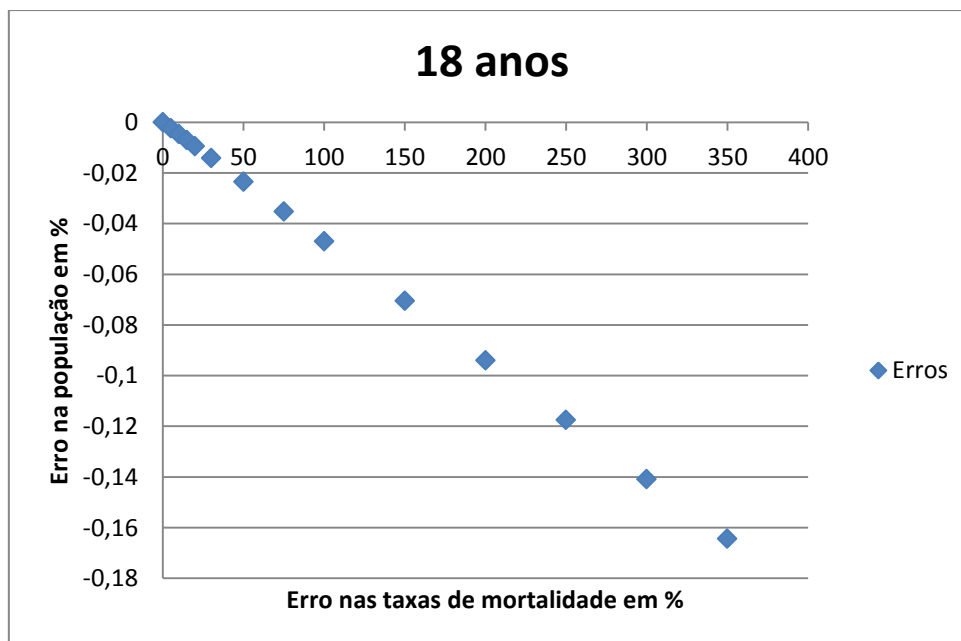


Gráfico 51 – Impacto na população de 18 anos de idade dos erros das taxas de mortalidade no ano de 2008.

Os resultados da análise são de facto encorajadores, isto porque para os 18 anos de idade, mesmo com um erro nas taxas de mortalidade de 350% apenas tem um impacto na ordem dos 0,16% (valor absoluto) na população daquela faixa etária. O que vem dar bastante força ao ajustamento destes modelos para este fim em particular, ou seja, para a estimação de população em idades jovens.

9. Conclusões

Tendo em conta que esta dissertação apresenta várias etapas distintas de trabalho, decidiu-se apresentar as conclusões divididas nessas mesmas etapas, ou seja, nos diversos capítulos de análise que foram apresentados. Desta forma, tenta-se simplificar a leitura e compreensão das conclusões.

9.1. Projecções vs Valores Observados

Numa primeira abordagem aos dois modelos, *Lee-Carter* e *Booth-Maindonald-Smith*, e à projecção de mortalidade em Portugal, registou-se uma semelhança muito grande entre os dois. Destaca-se também o facto de os dois modelos parecerem ser demograficamente viáveis para as projecções pretendidas, visto que não apresentam valores de erros muito significativos.

Quando comparadas as três populações estudadas, população total, feminina e masculina, graficamente não existe nenhuma evidência de que a projecção não esteja dentro dos parâmetros aceitáveis. Os dois modelos apresentam projecções com alisamentos bastante positivos, não apresentando valores extremos. Quando comparadas as curvas de mortalidade projectadas com as curvas de mortalidade observadas verifica-se que nas taxas observadas existem mais “picos” ao longo da curva que nas taxas projectadas por ambos os modelos.

Os erros calculados associados aos ajustamentos apresentam valores considerados aceitáveis para o tipo de análise que se está a realizar. Também neste caso se verifica uma grande semelhança entre modelos, registando-se poucas diferenças entre os valores de erros associados às projecções em ambos. Um dado importante que se pode retirar da análise aos erros é o facto de ser maior nas idades mais jovens (até 35 anos) e consideravelmente mais baixo em idades mais avançadas (acima dos 40 anos).

A utilização de intervalos de predição para a verificação da bondade de ajustamento dos modelos não foi bem sucedida. Os valores observados na sua maioria dos casos situaram-se fora dos intervalos de predição construídos pelos modelos. No entanto, não é motivo suficiente para pôr em causa os resultados dos mesmos, isto porque possivelmente o grau de confiança utilizado deveria ter sido maior, assim como pelo facto de o valor das taxas de mortalidade ser de uma ordem de grandeza bastante

pequena, fazendo com que uma pequena alteração de valor possa fazer com que o valor saia do intervalo de predição.

De um modo geral, neste primeiro capítulo de análise aos modelos de projecções, pode-se concluir que a utilização dos mesmos foi positiva e estatisticamente aceitável. Não parece ter havido qualquer indicação que não seriam projecções de qualidade as realizadas pelos modelos em estudo. No entanto, o estudo não apresentou diferenças relevantes entre os dois modelos, não se podendo concluir quanto à melhor capacidade de projectar de um em relação ao outro. Qualquer um dos dois modelos se apresenta como uma boa solução para o referido.

9.2. Ajustamento dos modelos

A primeira grande conclusão das comparações entre ajustamentos dos modelos efectuadas neste trabalho é a não existência de diferenças relevantes entre os dois. Ao longo de todos os intervalos de tempo utilizados nas bases de dados em estudo, assim como para as populações analisadas, as diferenças entre os dois modelos comparados são muito pequenas, apresentando em todos os casos comportamentos e valores bastante idênticos.

Ao nível das populações as semelhanças verificadas nos modelos continuam a ser vincadas. Ou seja, de um modo geral, para todos os modelos realizados, quando comparadas as populações totais, femininas e masculinas, não existem diferenças relevantes assinaláveis entre elas, dentro de cada período de ajustamento idêntico.

Já no que concerne aos períodos de ajustamento, parece haver diferenças entre os três utilizados. O período de ajustamento com maior número de anos é o que apresenta um nível de erro percentual absoluto médio maior. O nível de erro global apresentado vai aumentando consoante o número de anos ir aumentando, ou pelo menos é o que assim sugerem os valores. No entanto, há que não esquecer que o objectivo do trabalho é a projecção a 25 anos e utilizar bases de dados com apenas 10 anos de intervalo temporal pode ser pouco aconselhado.

De um modo geral, e tendo agora como foco de análise os valores obtidos para os EPAM, EPALM e também o EQM, pode-se afirmar que as conclusões são interessantes e relevantes.

Nos erros absolutos há a destacar que em todos os períodos analisados os valores tendem a decrescer nos primeiros anos de ajustamento e a aumentar nos últimos. Nos dois primeiros períodos em análise esse aumento ultrapassa em muito os valores absolutos dos primeiros anos de ajustamento, enquanto no último período o valor absoluto do erro nos últimos anos de ajustamento não ultrapassa os erros dos primeiros anos. De salientar, mais uma vez, que não existem diferenças, quer ao nível de modelos, quer ao nível de populações.

Quanto aos erros algébricos existe uma curiosa subestimação nos primeiros anos de ajustamento e uma sobrestimação nos últimos anos. Este comportamento é uma constante em todos os modelos e períodos observados, tendo como ponto de viragem, em geral, o centro do período temporal utilizado em cada modelo. Ou seja, em termos médios nos anos centrais de ajustamento não haverá nem subvalorização nem sobrevalorização dos valores das taxas de mortalidade.

O erro quadrático médio apresenta valores baixos, sempre com ordem de grandeza abaixo das milésimas, para qualquer um dos modelos ajustados, período ou população. De referir que o valor dos EQM foi sempre superior nas populações masculinas quando comparados com as populações femininas.

Fazendo uma análise dos erros ao longo das idades conclui-se que tanto num erro como noutro, existe um maior valor de erro para as idades mais jovens (até aos 40 anos) e que nas idades mais avançadas esse mesmo erro tende a diminuir, situando-se quase sempre abaixo dos 50%, não havendo grandes “picos” nos gráficos observados. O que mais uma vez evidencia o facto de nas idades cuja taxa de mortalidade é mais baixa, é mais complicado para os modelos ajustarem-se da melhor maneira. Por contraste, nas idades com uma taxa de mortalidade superior, os modelos têm uma maior facilidade de ajustamento.

Portanto, como já foi referido anteriormente, a principal conclusão do estudo do ajustamento dos modelos é a quase ausência de diferenças entre os dois modelos, *Lee-Carter* e *Booth-Maindonald-Smith*. No entanto, há a destacar o menor valor de erros

associados a períodos mais curtos de ajustamento, assim como as diferenças nos anos de ajustamento, quer ao nível de erros absolutos quer ao nível de erros algébricos.

9.3. Projecções

As principais comparações terão de ser feitas ao nível das projecções. De um modo geral as projecções quer para o modelo LC quer para o modelo BMS são muito semelhantes.

Para as taxas de mortalidade há a destacar as diferenças entre períodos de ajustamento. Há um alisamento muito mais acentuado nas curvas de mortalidade para os períodos de ajustamento mais longos e um menor alisamento principalmente para o período de ajustamento de apenas 10 anos, o que leva a afirmar que possivelmente não será o melhor período de ajustamento para as projecções que se pretendem calcular. Para os dois primeiros períodos de ajustamento, os mais longos, não se vislumbram grandes diferenças entre os modelos.

Na análise das projecções realizadas para os óbitos, as diferenças entre modelos existem apenas de forma relevante nos intervalos de predição. Sendo no caso do modelo de LC bastante maior do que no caso do modelo de BMS. De um modo geral os valores mantêm-se bastante semelhantes para os dois modelos e apresentam comportamentos que não parecem ser de todo improváveis. Nesta análise não existem diferenças entre períodos a destacar.

No cálculo da esperança de vida, a análise é em tudo semelhante à do total de óbitos. Não existem diferenças quer ao nível de valores projectados para os dois modelos quer para os dois períodos, mas apenas diferença nos intervalos de predição. O comportamento da curva da esperança de vida ao longo dos anos projectados parece ir de encontro com o expectável, não havendo motivos para duvidar da bondade das projecções em questão.

9.4. Análise de Sensibilidade

A grande conclusão que se pode retirar da análise de sensibilidade realizada neste projecto é a de que o impacto dos erros nas taxas de mortalidade é maior para as

idades mais avançadas do que para as idades mais jovens. E é-o de uma forma bastante vincada e assinalável.

Verifica-se que mesmo errando nas taxas de mortalidade numa ordem dos 350%, para idades jovens, o erro reflectido na população estimada utilizando essas taxas de mortalidade é sempre menor que 5%. Fortalecendo assim o ajustamento dos dois modelos no âmbito de um projecto que tem como principal população alvo grupos etários jovens.

9.5. Síntese Final

Resumidamente, conclui-se que os dois modelos, de uma forma global, se comportam de maneira semelhante, quer ao nível do ajustamento, quer ao nível da projecção, quando utilizados para a análise de dados referentes à população portuguesa, no que à população total diz respeito, mas igualmente para a população diferenciada por sexo.

Também de referir que é importante um equilíbrio na hora de escolher o período temporal ajustado aos modelos, ou seja, nem muitos anos que tirem veracidade aos números, tendo em conta os avanços científicos na área da saúde, nem poucos anos, tendo em conta a necessária atenção ao número de anos projectados.

De um modo geral, pode-se afirmar que a utilização de qualquer um dos dois modelos é aceitável e bastante positiva, projectando valores de mortalidade bastante fiáveis e consequentemente estimando valores de população para os fins pretendidos muito interessantes.

Observando de um ponto de vista demográfico, pode-se concluir também que o uso de um modelo com um intervalo de confiança mais amplo, contrariamente ao habitual, se torna benéfico pelo facto de dar origem a uma maior amplitude de resultados nos cenários em estudo.

Através deste estudo foi possível não apenas prever que a procura relativa às instituições de Ensino Superior em Portugal tenderá a diminuir nos anos entre 2010 e 2035, como também quantificar essa diminuição.

De um modo geral, pôde-se constatar que a população jovem que mais probabilidade teria de se candidatar ao Ensino Superior português (a população com 18 anos, o grupo dos 18 aos 24 anos e também o grupo dos 18 aos 30 anos), tenderá a diminuir de forma considerável nos próximos anos em todos aqueles grupos etários. A população jovem que está associada à candidatura e frequência do Ensino Superior em Portugal poderá vir a diminuir entre os 20 a 25%, nos próximos 25 anos.

Com os resultados desta análise, a avaliação da procura e a consequente possível redefinição dos pontos da Rede de Ensino Superior em Portugal poderão ser trabalhados de uma forma bastante mais precisa, adequada e sustentável.

Sem acesso a dados dos candidatos dos últimos anos ao Ensino Superior em Portugal é extremamente difícil uma análise mais profunda à população projectada. Caso existissem dados de cada Universidade sobre os candidatos da sua região de origem ou estrato social, por exemplo, seria possível uma análise mais aprofundada sobre regiões de atracção e como redistribuir duma forma mais eficaz os pontos da Rede de Ensino Superior em Portugal.

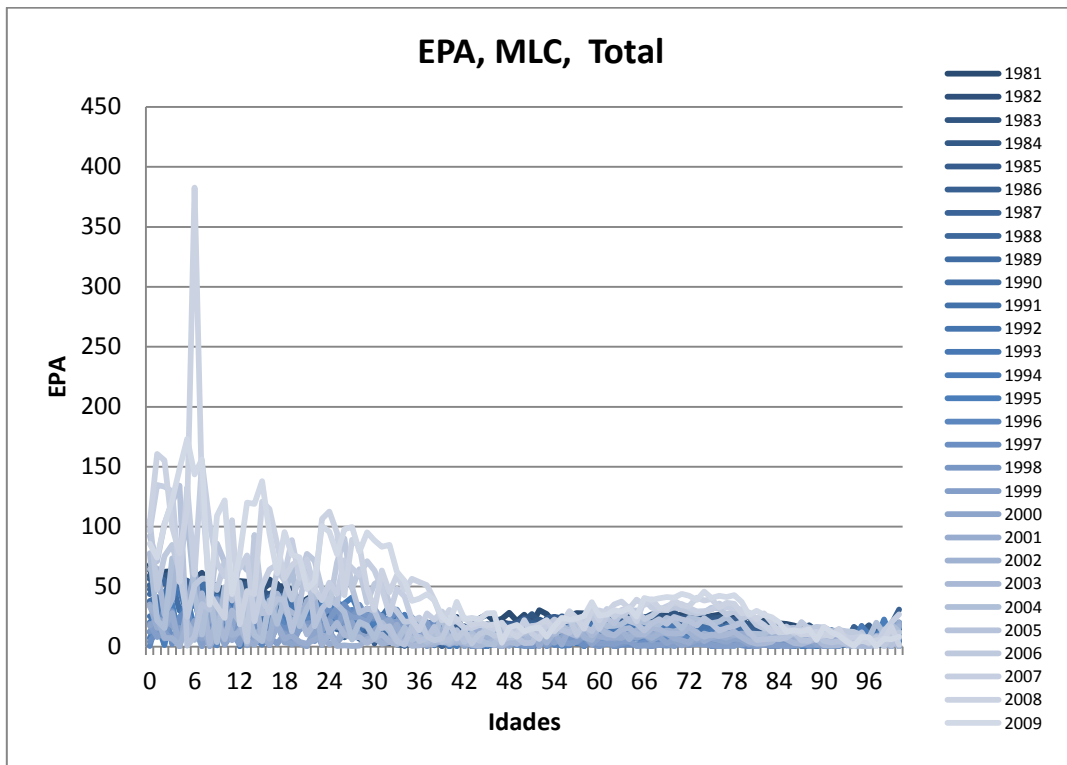
10. Bibliografia

- Booth, Heather, Hyndman, Rob J., Tickle, Leonie e Jong, Piet de (2006), *Lee-Carter mortality forecasting: a multi-country comparasion of variants and extensions*, Demographic Research Journal, Vol. 15, Art. 9: 289-310, Disponível: <http://www.demographic-research.org/Volumes/Vol15/9/>
- Booth, Heather, Maindonald, John e Smith, Len (2001), *Age-time interactions in mortality projection: applying Lee-Carter to Australia*, Working Papers in Demography n° 85. Canberra: Demography and Sociology Program, Australian National University.
- Booth, Heather, Maindonald, John e Smith, Len (2002). *Applying Lee-Carter under conditions of variable mortality decline*. Population Studies, 56 (3): 325–336.
- Booth, Heather, Tickle, Leonie e Smith, Len (2005), *Evaluating of the variants of the Lee-Carter Method of forecasting mortality: A multi-country comparison*, Special Issue, New Zealand Population Review, 31(1): 13-34.
- Box, G., Jenkins, G. M. e Reinsel, G. (2008), *Time Series Analysis: Forecasting & Control*, 4th Edition, Prentice Hall.
- Bravo, J.M.V. (2007), *Tábuas de Mortalidade Contemporâneas e Prospectivas: Modelos Estocásticos, Aplicações Actuariais e Cobertura do Risco de Longevidade*, PhD thesis, Universidade de Évora, Évora.
- Coelho, E. I. F. (2001). “*Método de Lee-Carter para previsão da mortalidade*”, Revista de Estudos Demográficos n° 37 - Instituto Nacional de Estatística, Departamento de Estatísticas Sociais.
- Cowpertwait, Paul S.P. e Metcalfe, Andrew V. (2009), *Introductory Time Series with R*, Springer.
- Fígoli, Moema G. Bueno (1998), *Modelando e projectando a mortalidade no Brasil*, Revista Brasileira de Estudos de População, Vol. 1, n° 1.
- Hyndman, Rob J. e Ulah, M.S. (2007), *Robust forecasting of mortality and fertility rates: A functional data approach*, Computational Statistics & Data Analysis 51(10): 4942-4956.

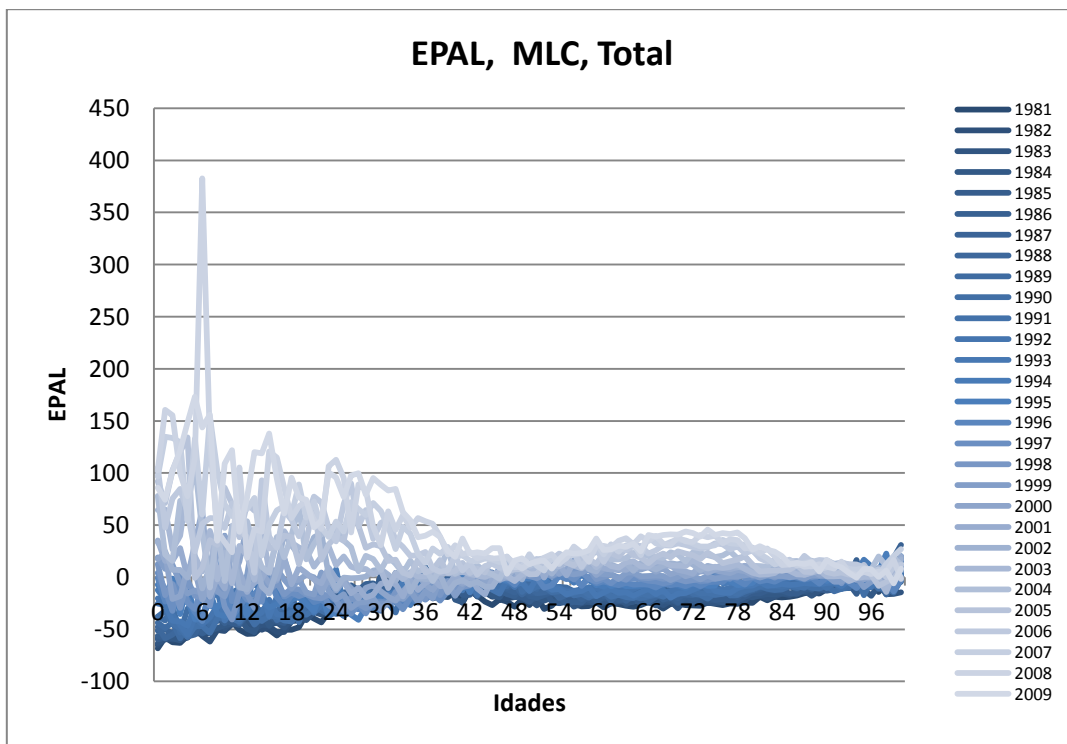
- Hyndman, Rob J., Booth, Heather, Tickle, Leonie e Maindonald, John (2012), *Forecasting mortality, fertility, migration and population data*, [Online], Disponível: <http://robjhyndman.com/software/demography/>
- Lee, Ronald (2000), *The Lee-Carter method for forecasting mortality, with various extensions and applications*, North American Actuarial Journal 4(1): 80-91.
- Lee, Ronald e Carter, Lawrence (1992), *Modeling and Forecasting U.S. Mortality*, Journal of the American Statistical Association, Vol.87, n.º 419, 659-675.
- Lee, Ronald e Miller, Timothy (2000), *Evaluating the Performance of Lee-Carter Mortality Forecasts*, University of California, Berkeley [Online], Disponível: <http://www.demog.berkeley.edu>.
- Lee, Ronald e Miller, Timothy (2001), *Assessing the performance of the Lee-Carter approach to modeling and forecasting mortality*, Demography 38(4): 537-549..
- Lee, Ronald e Nault, F. (2003), *Modeling and forecasting provincial mortality in Canada*, Paper presented at the World Congress of the International Union for the Scientific Study of Population, Montreal.
- Lee, Ronald e Rofman, Rafael (1994), *Modelacion y Proyeccion de la Mortalidad en Chile*, NOTAS 22, nº 59, 182-213. Also available in English from the authors, titled “*Modeling and Forecasting Mortality in Chile*”.
- Li, Nan e Gerland, Patrick (2011), *Modifying the Lee-Carter Method to project mortality changes up to 2100*, Paper presented at the Population Association of America 2011 Annual Meeting - Washington, DC. Session 125: Formal Demography I: Mathematical Models and Methods.
- Li, Nan e Lee, Ronald (2005), *Using the Lee-Carter method to forecast mortality for a group of populations non-divergently*, Paper presented at the XXV International Population Conference, Tours, France, Session 84: Stochastic projection of demographic parameters.
- Moreno, Adriano A. (2011), *Desenvolvimento do segundo e terceiro pilares da Segurança Social: O caso de Cabo Verde*, Tese de Mestrado, Universidade Nova de Lisboa, Lisboa.

- Shang, Han Lin, Booth, Heather e Hyndman, Rob J. (2011), *Point and interval forecasts of mortality rates and life expectancy: A comparasion of then principal component methods*, Demographic Research Journal, Vol. 25, Art. 5: 173-214, Disponível: <http://www.demographic-research.org/Volumes/Vol25/5/>
- Shumway, Robert H. e Stoffer, David S. (2010), *Time Series Analysis and Its Applications: With R examples*, 3rd Edition, Springer.
- Tuljapurkar, S., Li, N. e Boe, C. (2000), *A universal pattern of mortality decline in the G7 countries*, Nature 405: 789-792.
- Wilmoth, J.R. (1996), *Mortality projections for Japan: A comparison of four methods*, In Caselli, G. and Lopez, A.D. (eds) Health and Mortality among Elderly Populations, Oxford: Clarendon Press: 266-287.

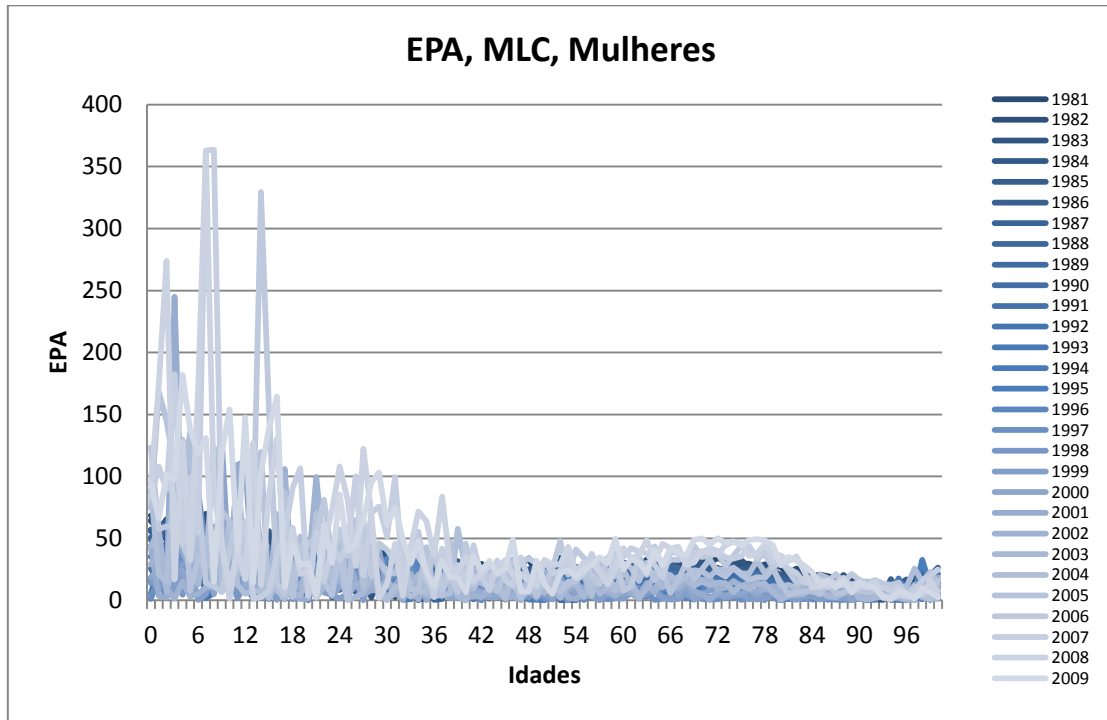
11. Anexos



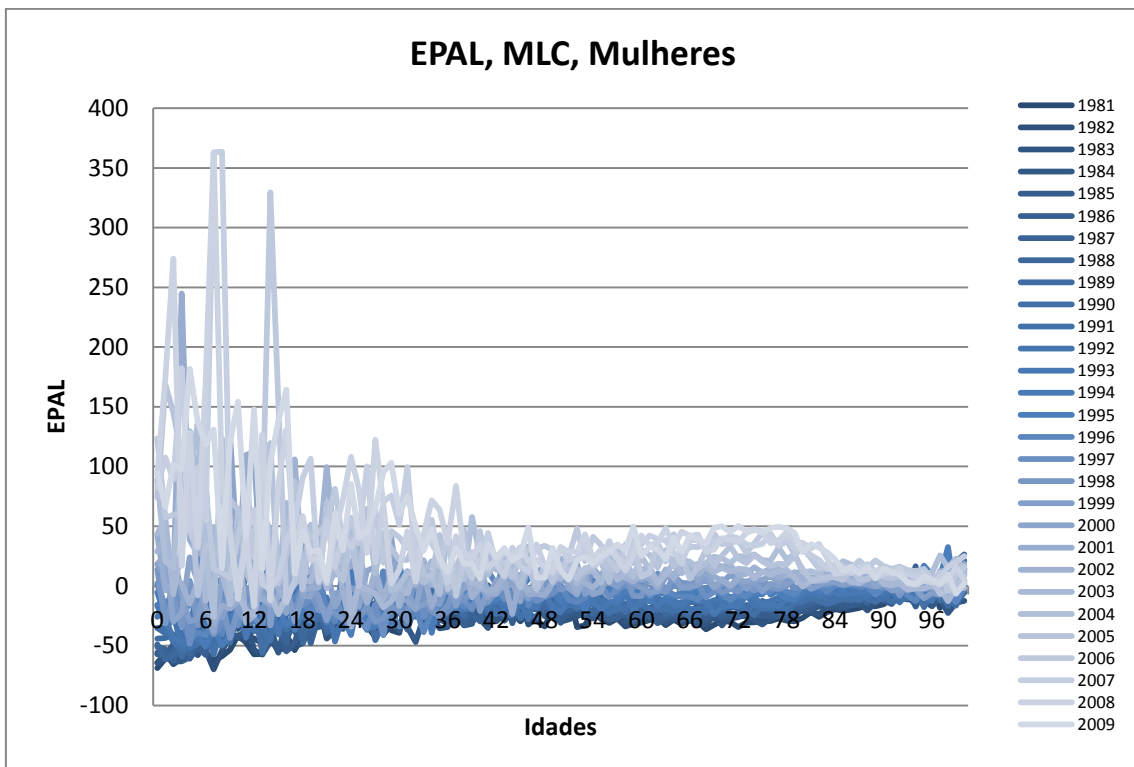
Anexo 1 – EPA nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.



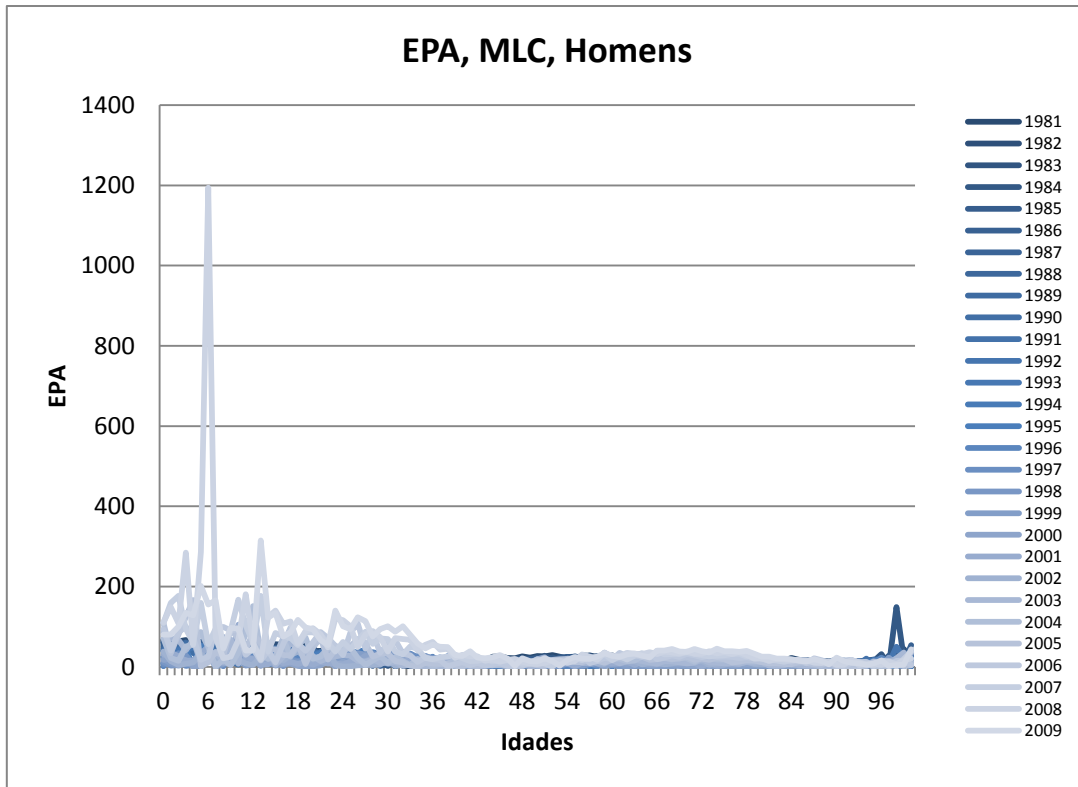
Anexo 2 – EPAL nos anos de ajustamento entre 1981 e 2009 do MLC para a população total.



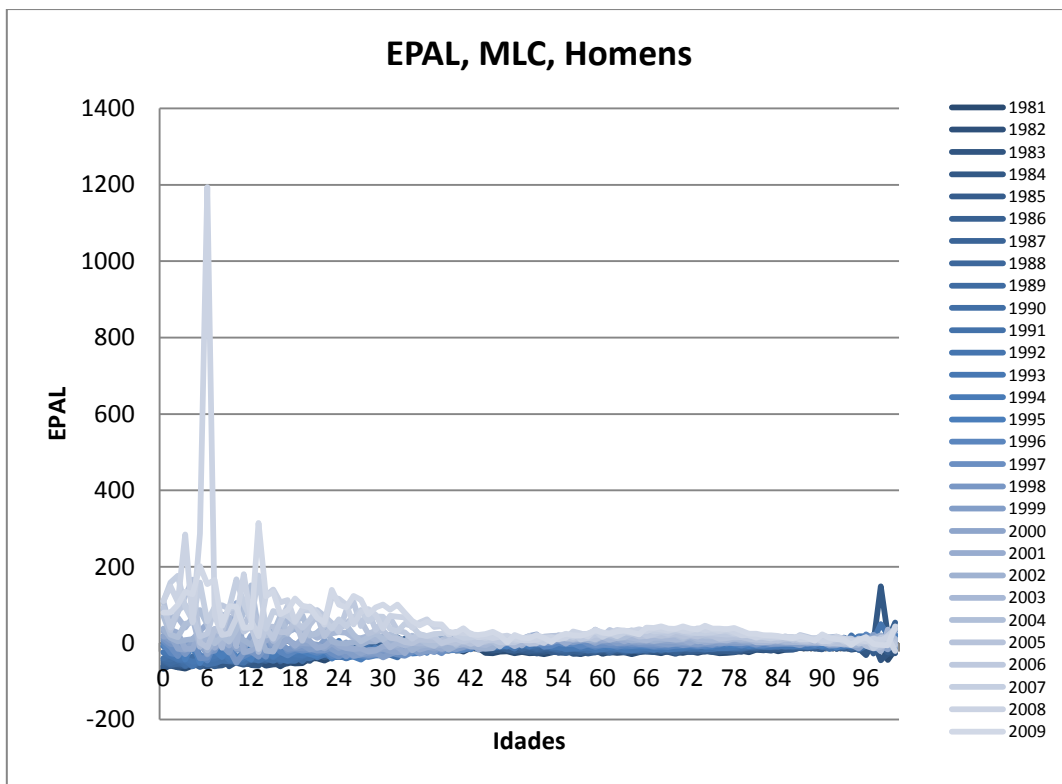
Anexo 3 – EPA nos anos de ajustamento entre 1981 e 2009 do MLC para a população feminina.



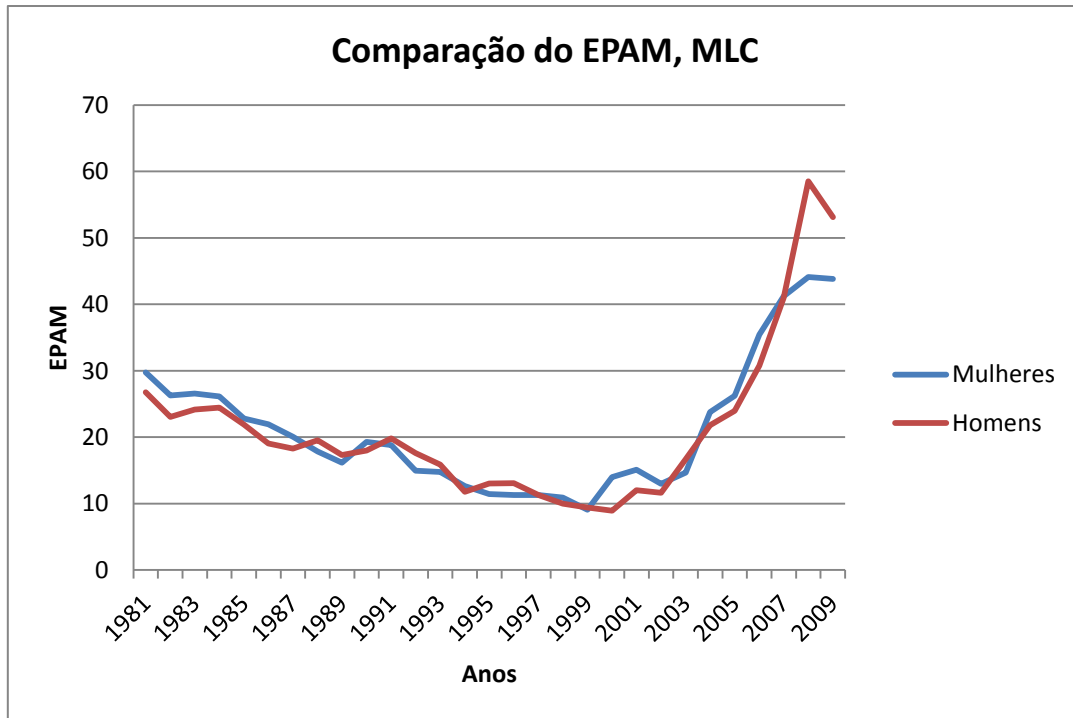
Anexo 4 – EPAL nos anos de ajustamento entre 1981 e 2009 do MLC para a população feminina.



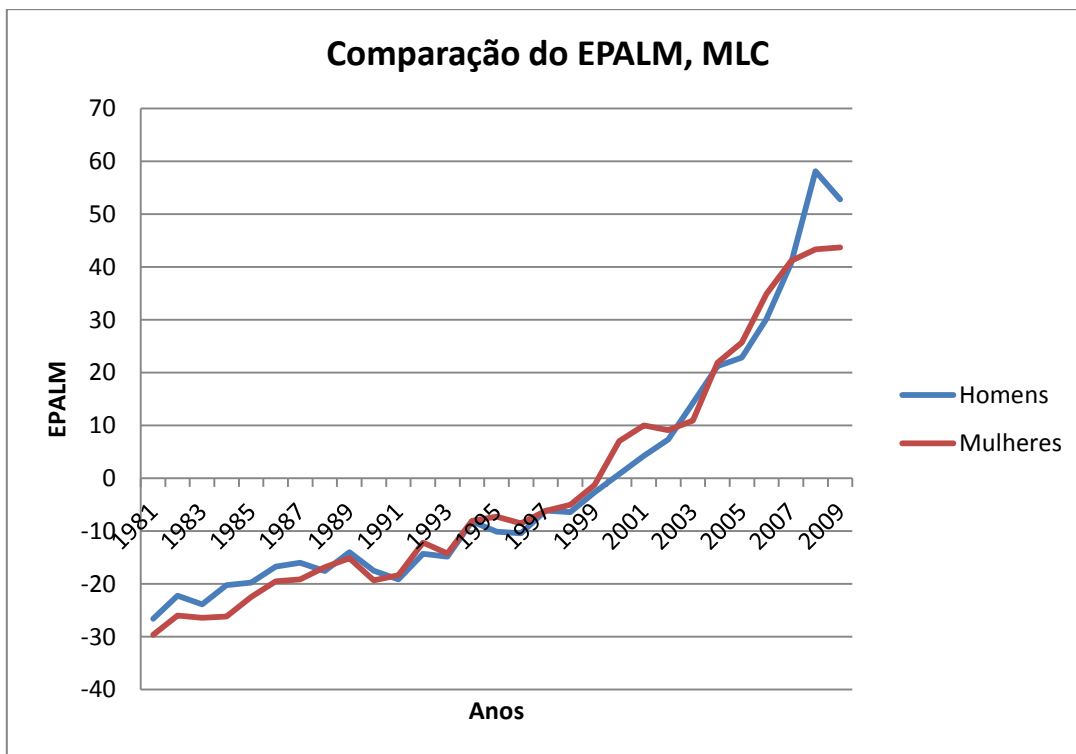
Anexo 5 – EPA nos anos de ajustamento entre 1981 e 2009 do MLC para a população masculina.



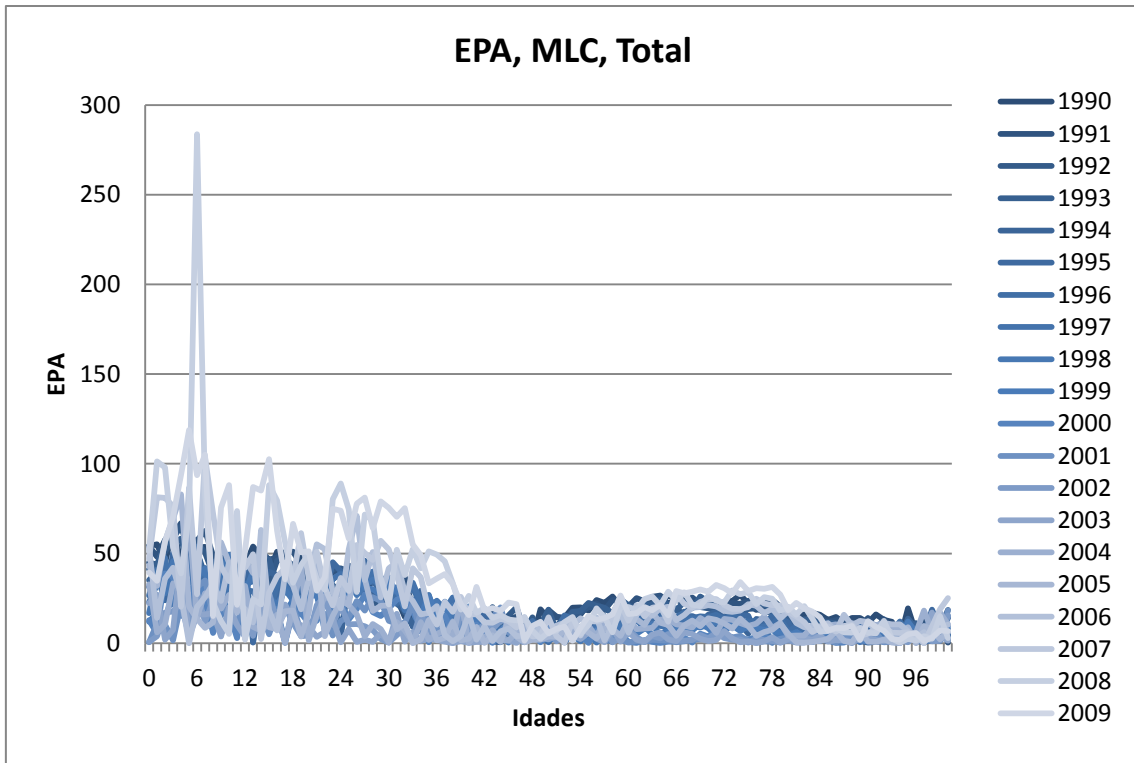
Anexo 6 – EPAL nos anos de ajustamento entre 1981 e 2009 do MLC para a população masculina.



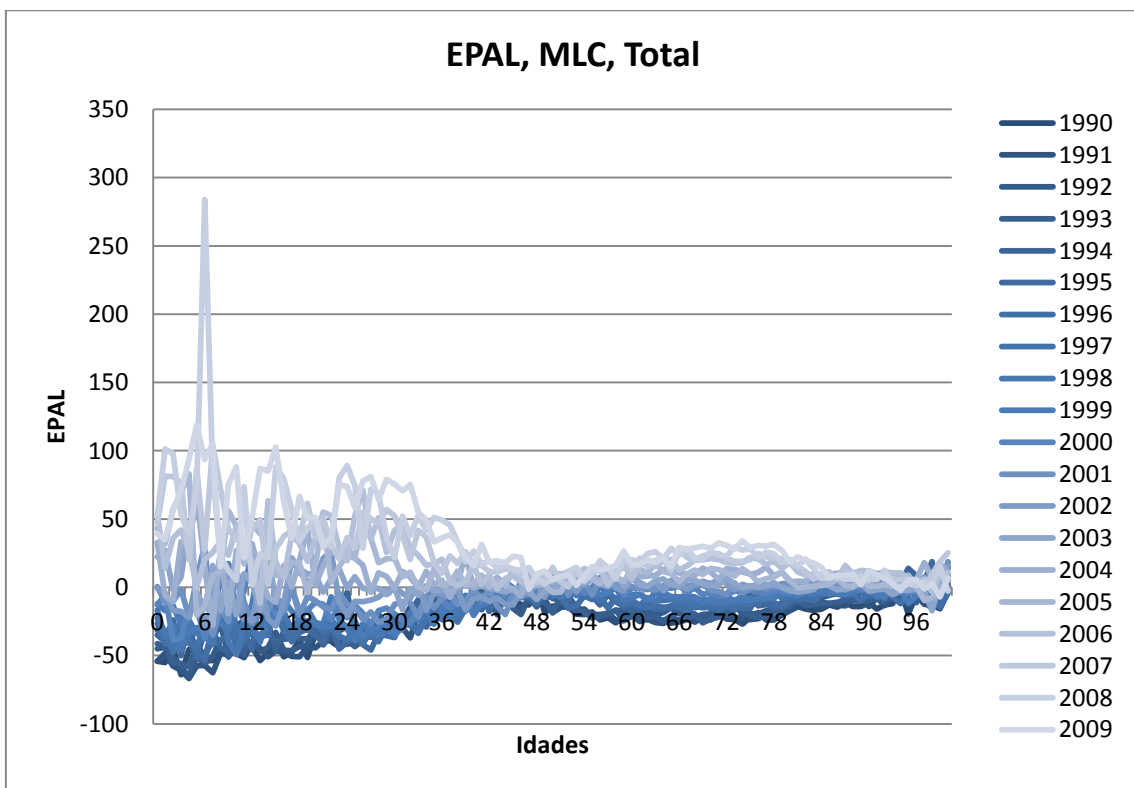
Anexo 7 – EPAM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1981 a 2009.



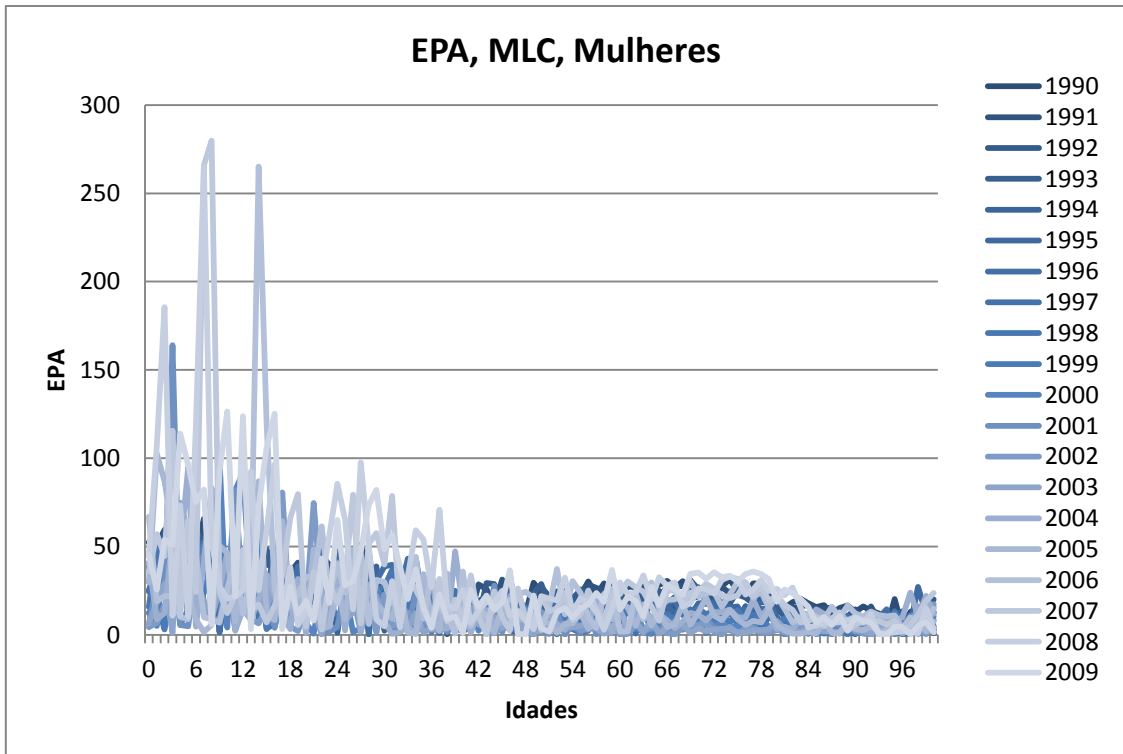
Anexo 8 – EPALM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1981 a 2009.



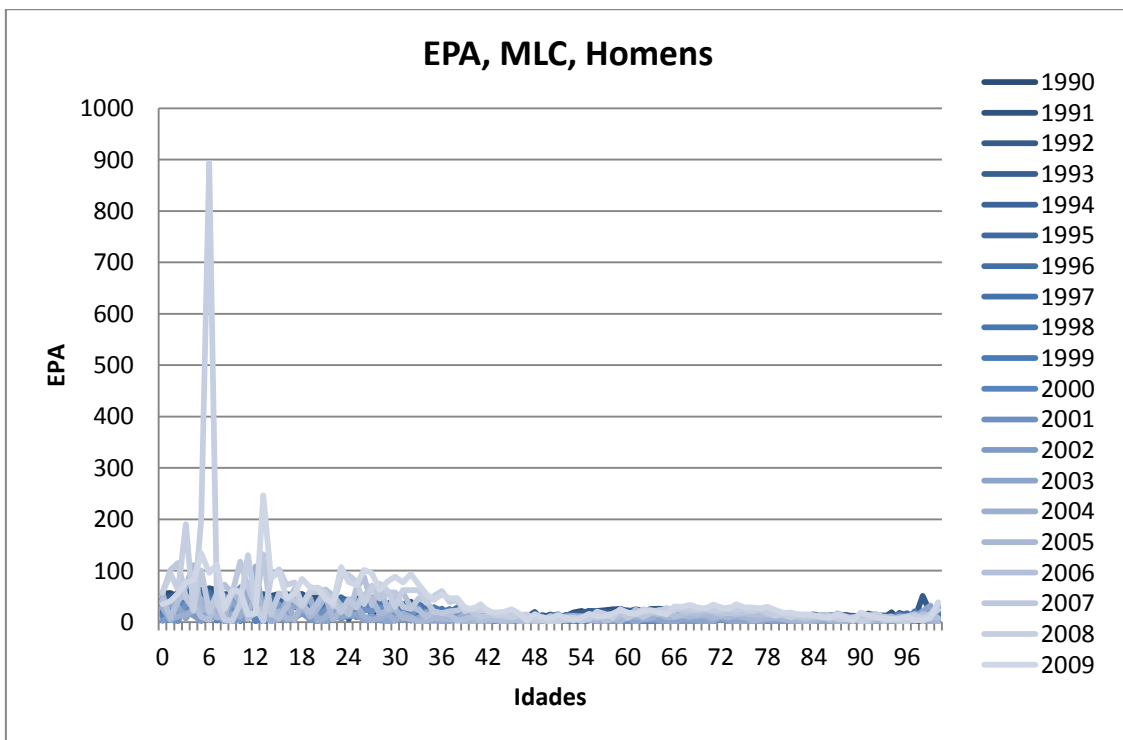
Anexo 9 – EPA nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.



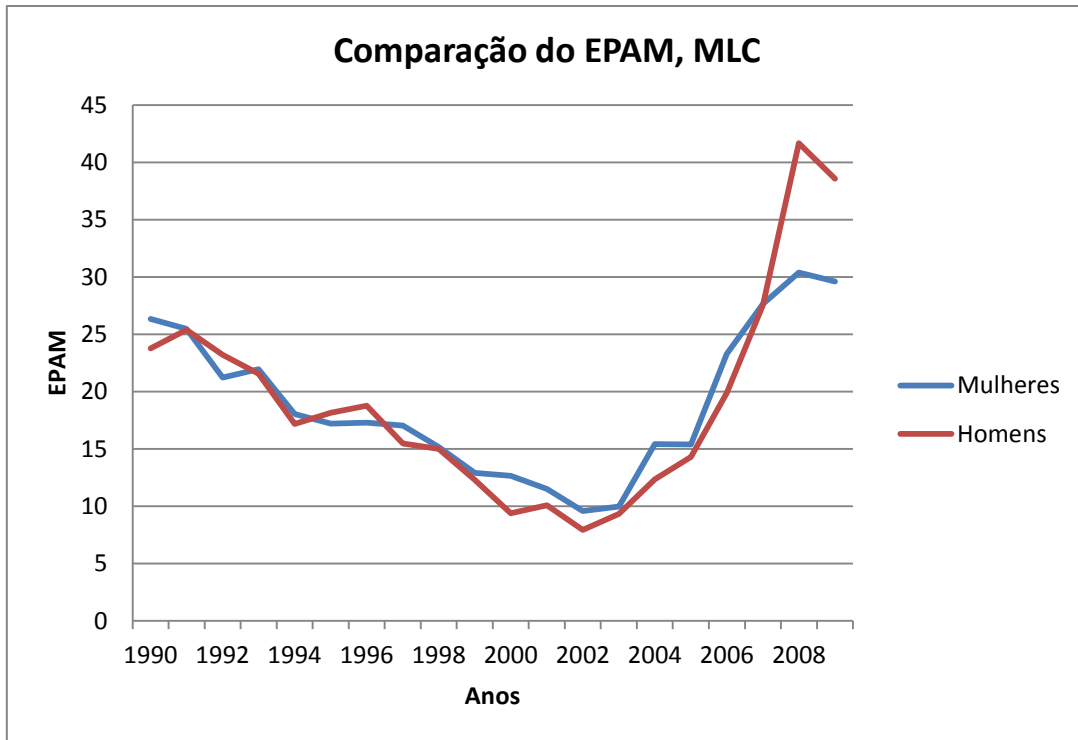
Anexo 10 – EPAL nos anos de ajustamento entre 1990 e 2009 do MLC para a população total.



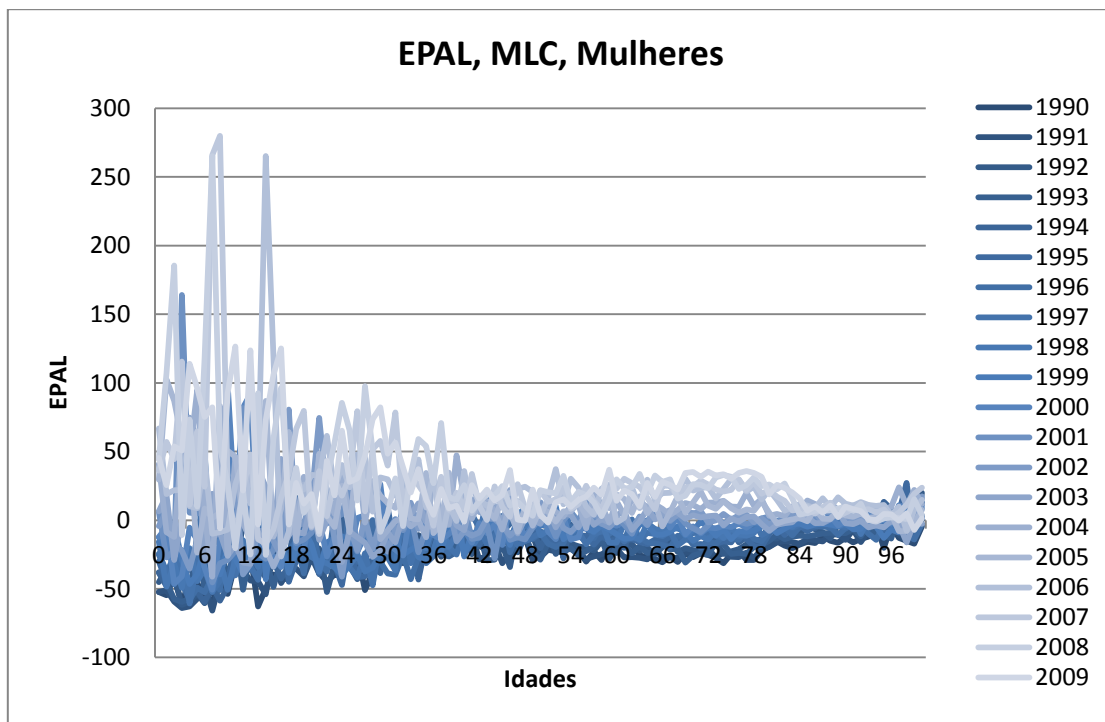
Anexo 11 – EPA nos anos de ajustamento entre 1990 e 2009 do MLC para a população feminina.



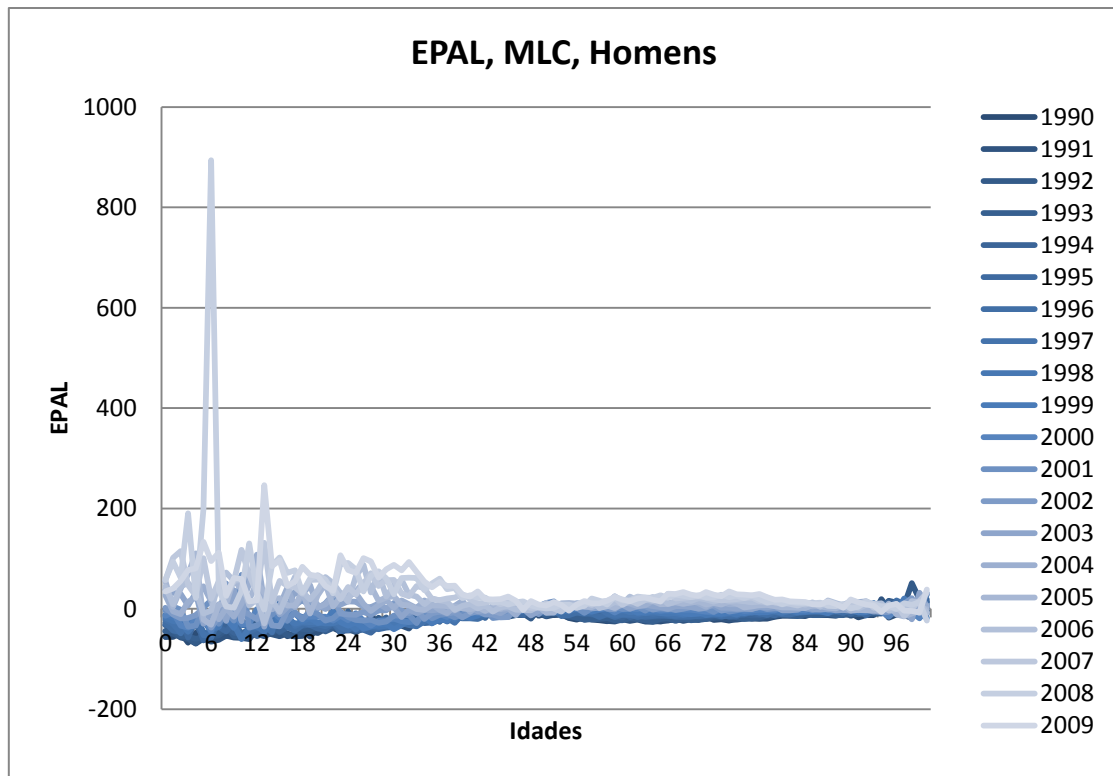
Anexo 12 – EPA nos anos de ajustamento entre 1990 e 2009 do MLC para a população masculina.



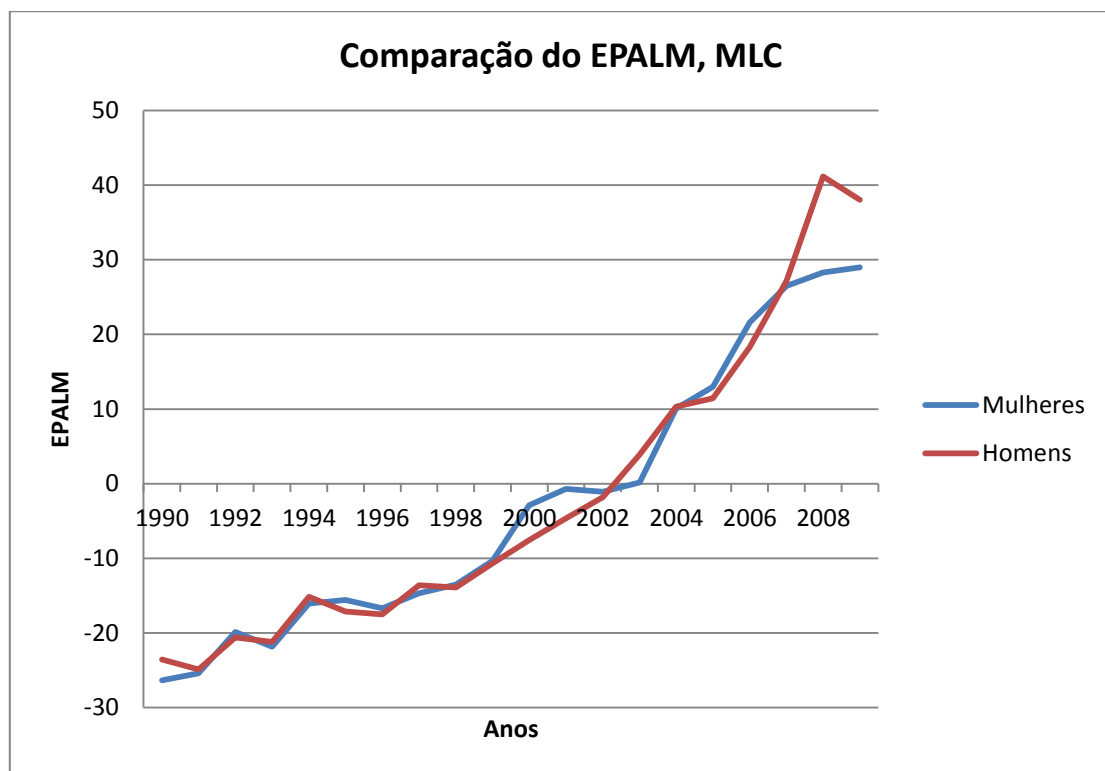
Anexo 13 – EPAM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1990 a 2009.



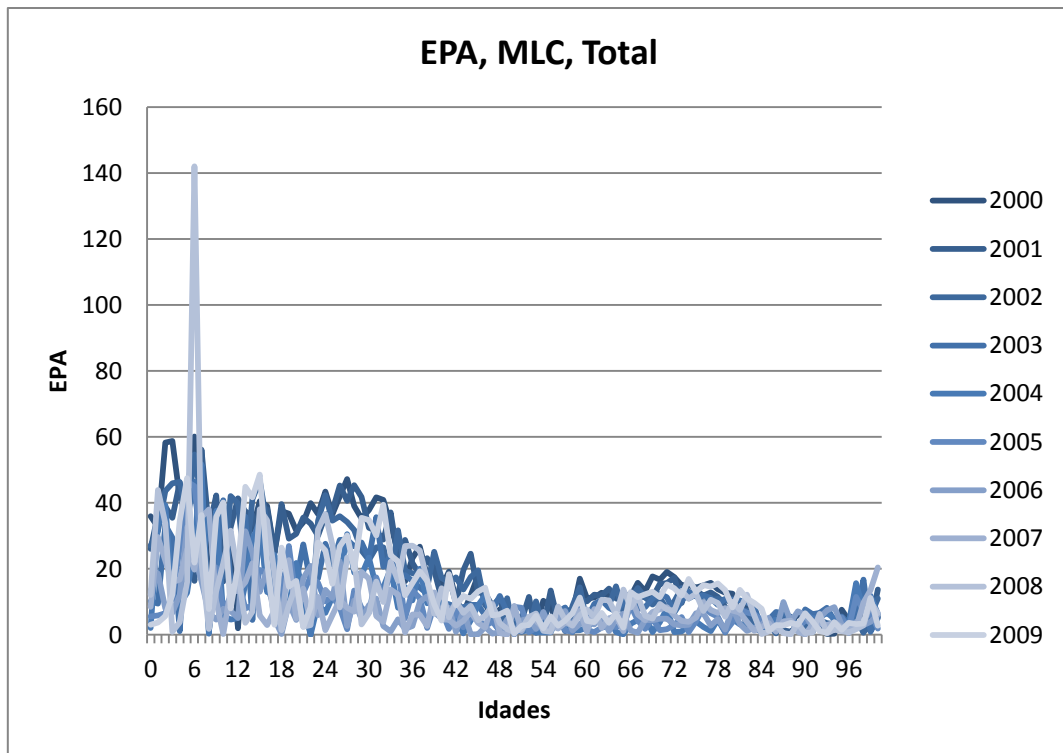
Anexo 14 – EPAL nos anos de ajustamento entre 1990 e 2009 do MLC para a população feminina.



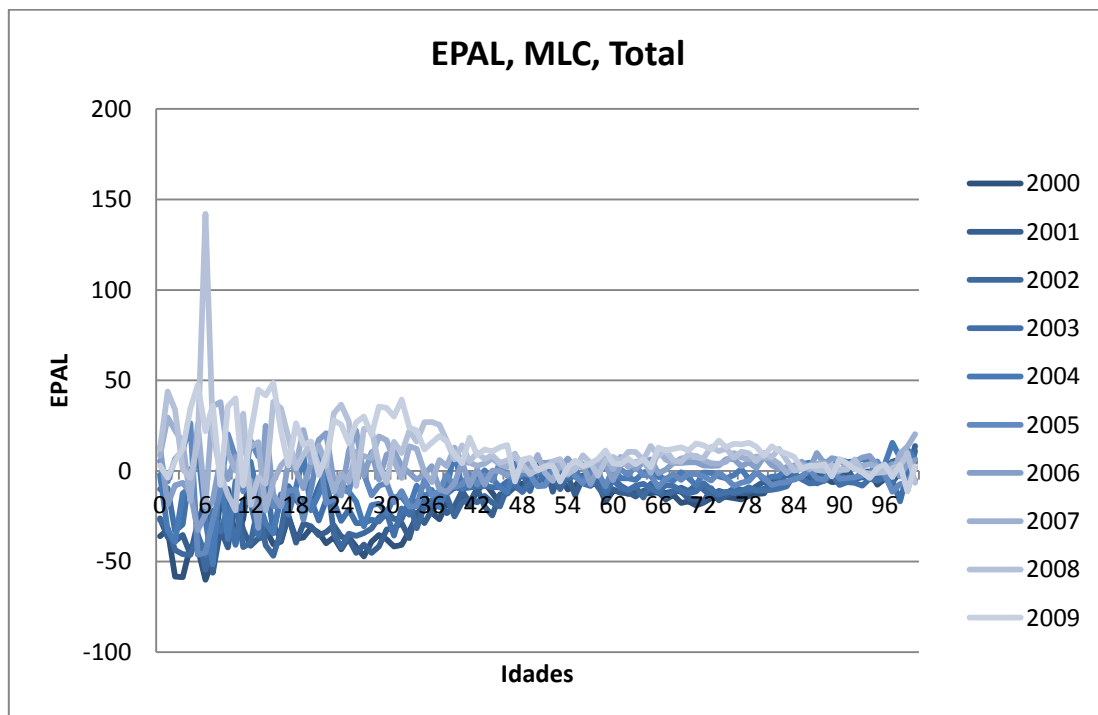
Anexo 15 – EPAL nos anos de ajustamento entre 1990 e 2009 do MLC para a população masculina.



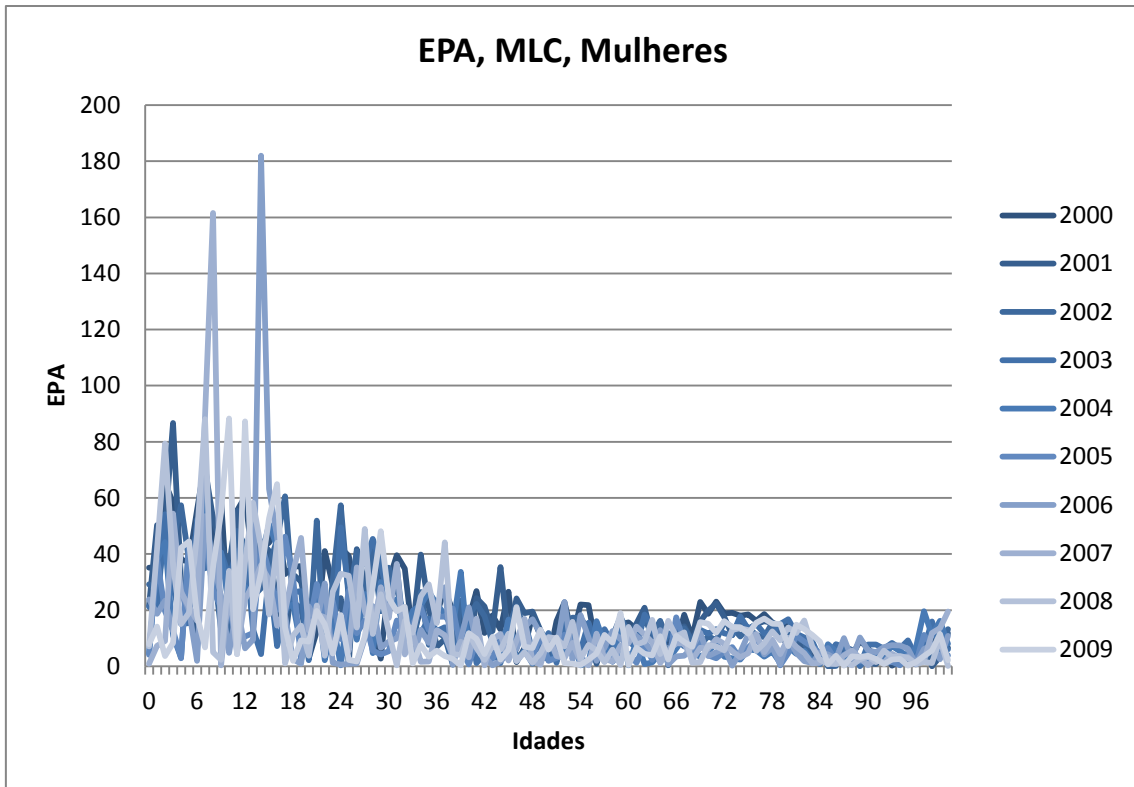
Anexo 16 – EPALM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 1990 a 2009.



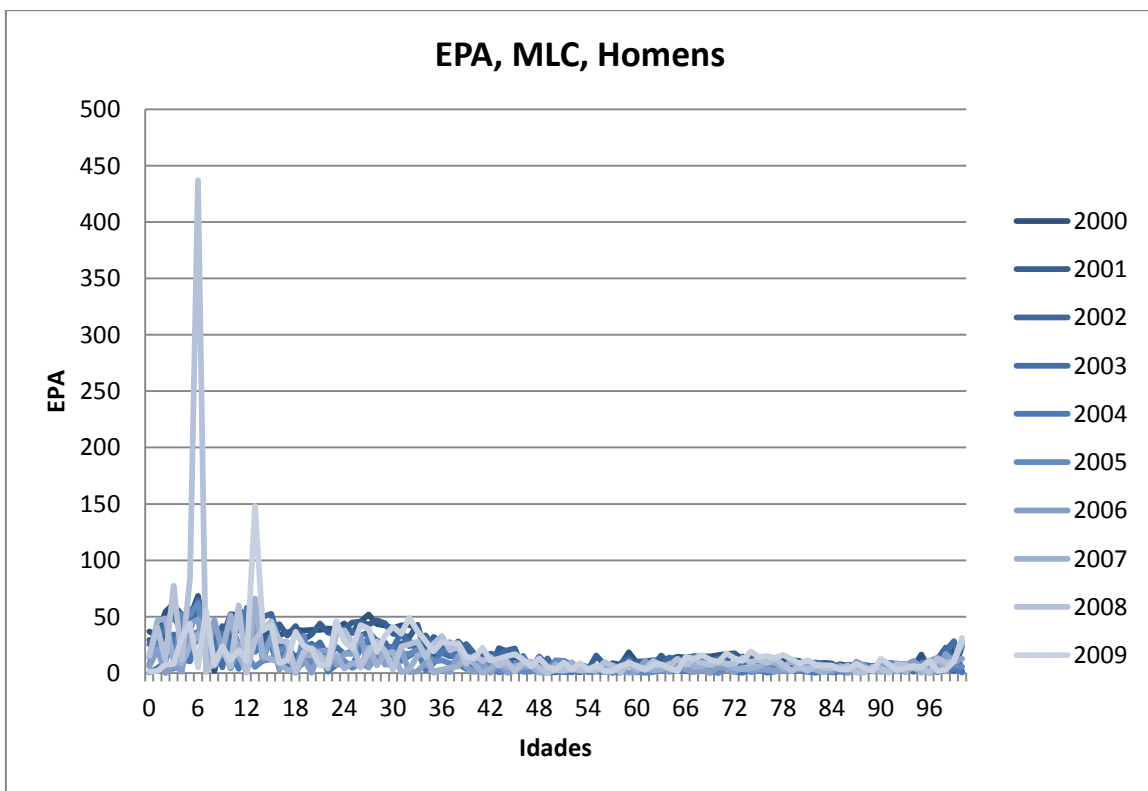
Anexo 17 – EPA nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.



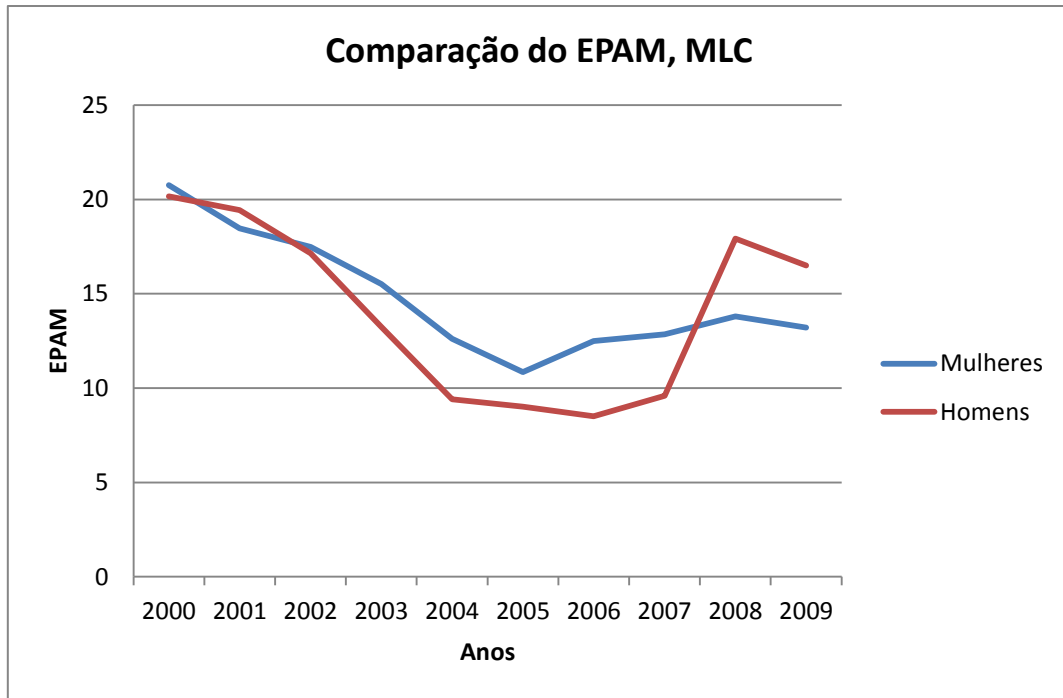
Anexo 18 – EPAL nos anos de ajustamento entre 2000 e 2009 do MLC para a população total.



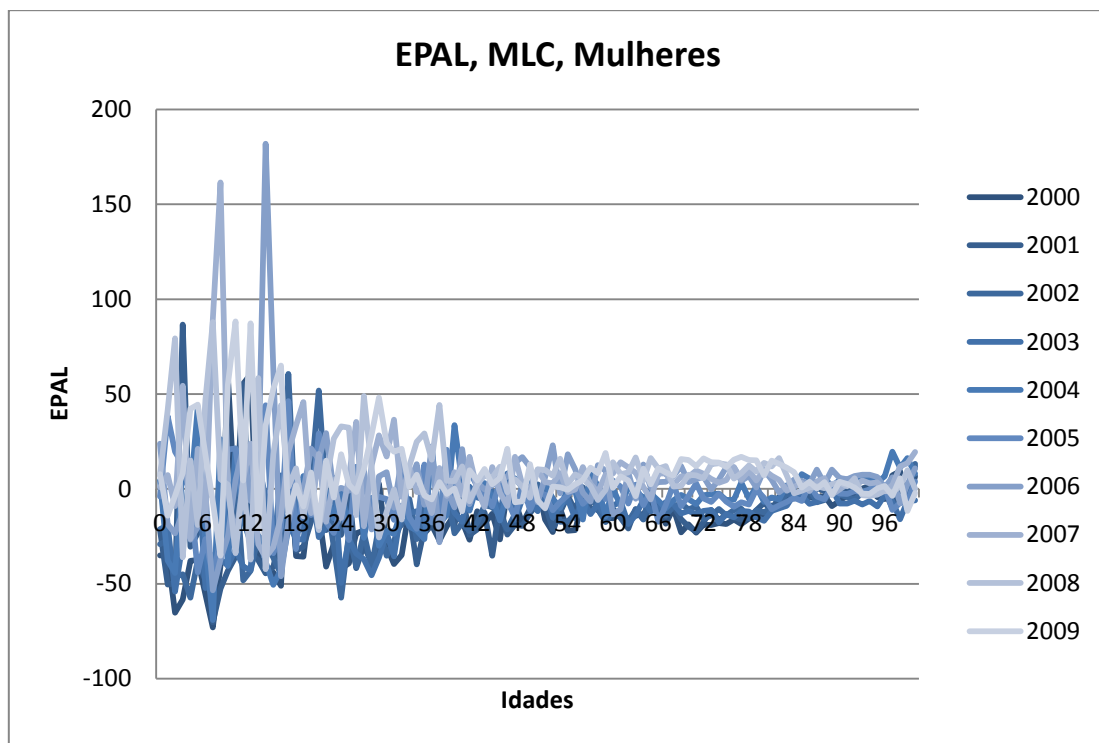
Anexo 19 – EPA nos anos de ajustamento entre 2000 e 2009 do MLC para a população feminina.



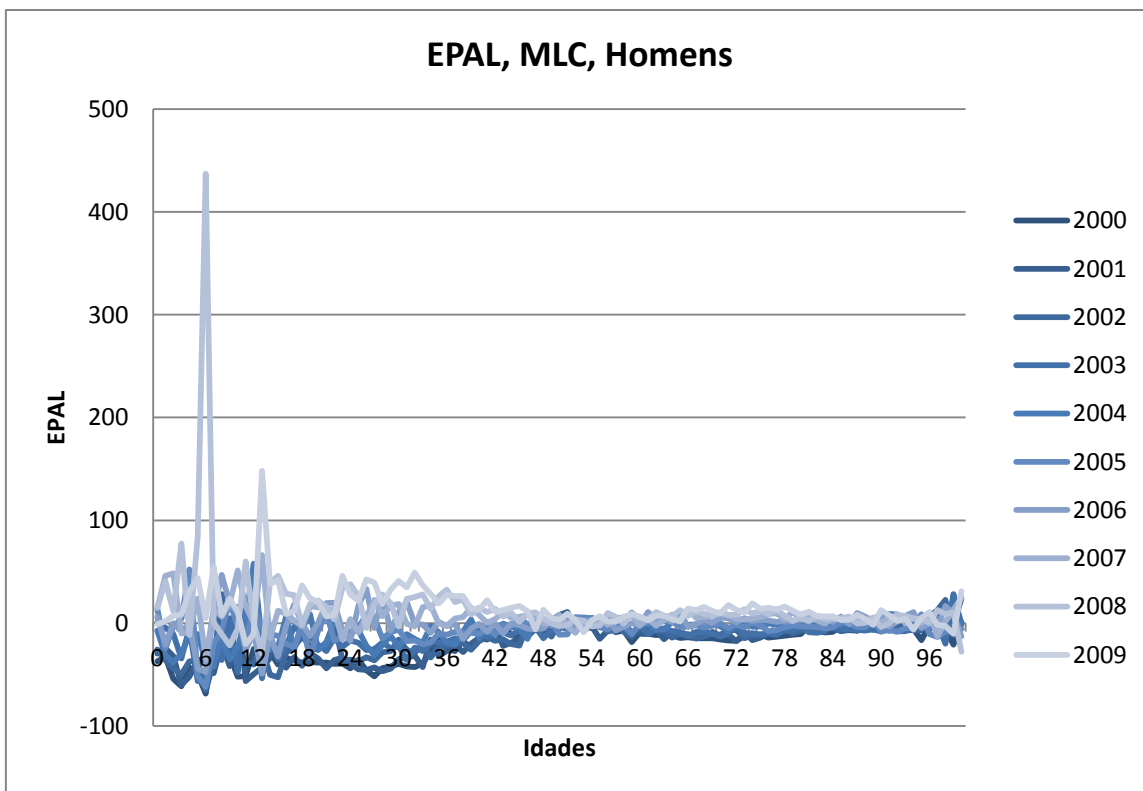
Anexo 20 – EPA nos anos de ajustamento entre 2000 e 2009 do MLC para a população masculina.



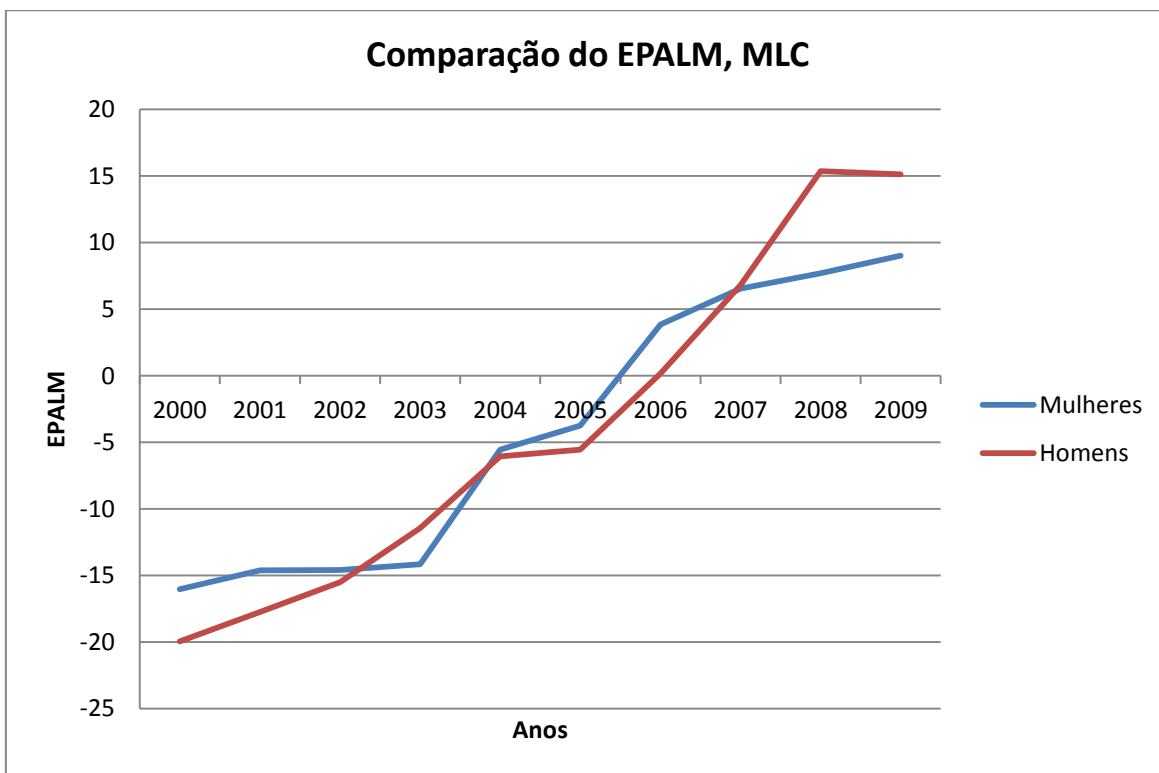
Anexo 21 – EPAM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 2000 a 2009.



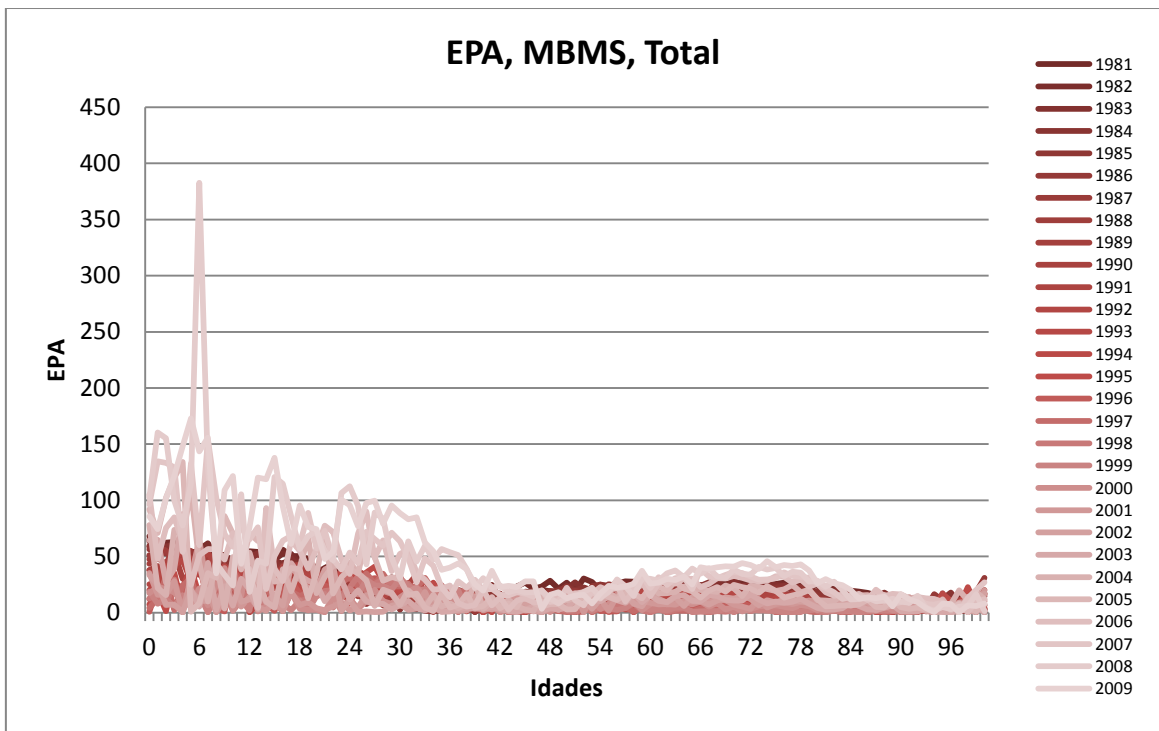
Anexo 22 – EPAL nos anos de ajustamento entre 2000 e 2009 do MLC para a população feminina.



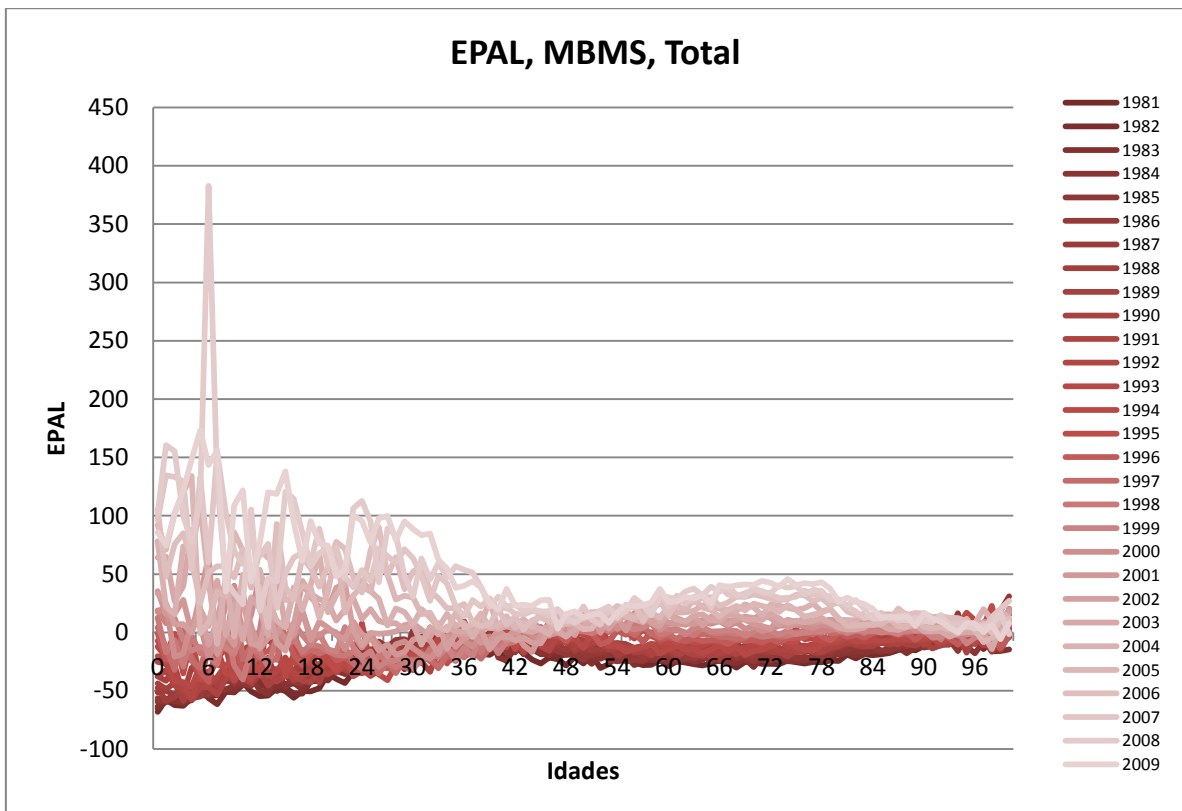
Anexo 23 – EPAL nos anos de ajustamento entre 2000 e 2009 do MLC para a população masculina.



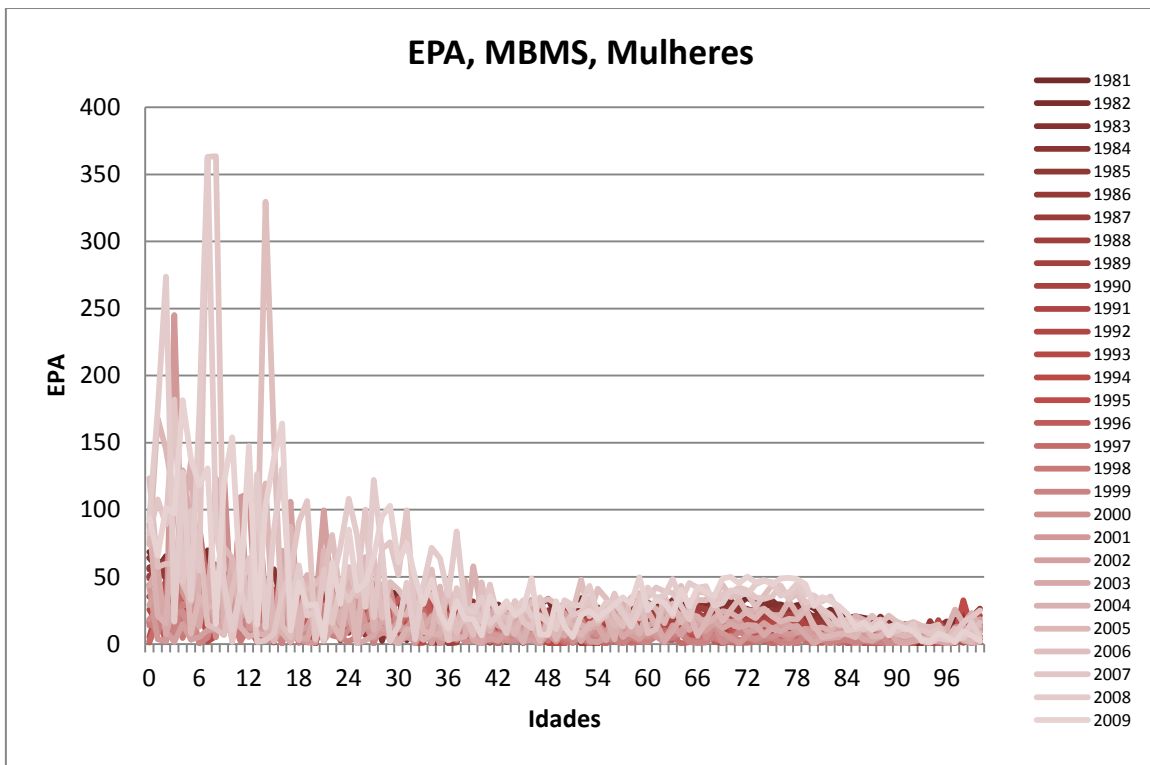
Anexo 24 – EPALM, comparação entre os sexos M e F, do MLC com anos de ajustamento de 2000 a 2009.



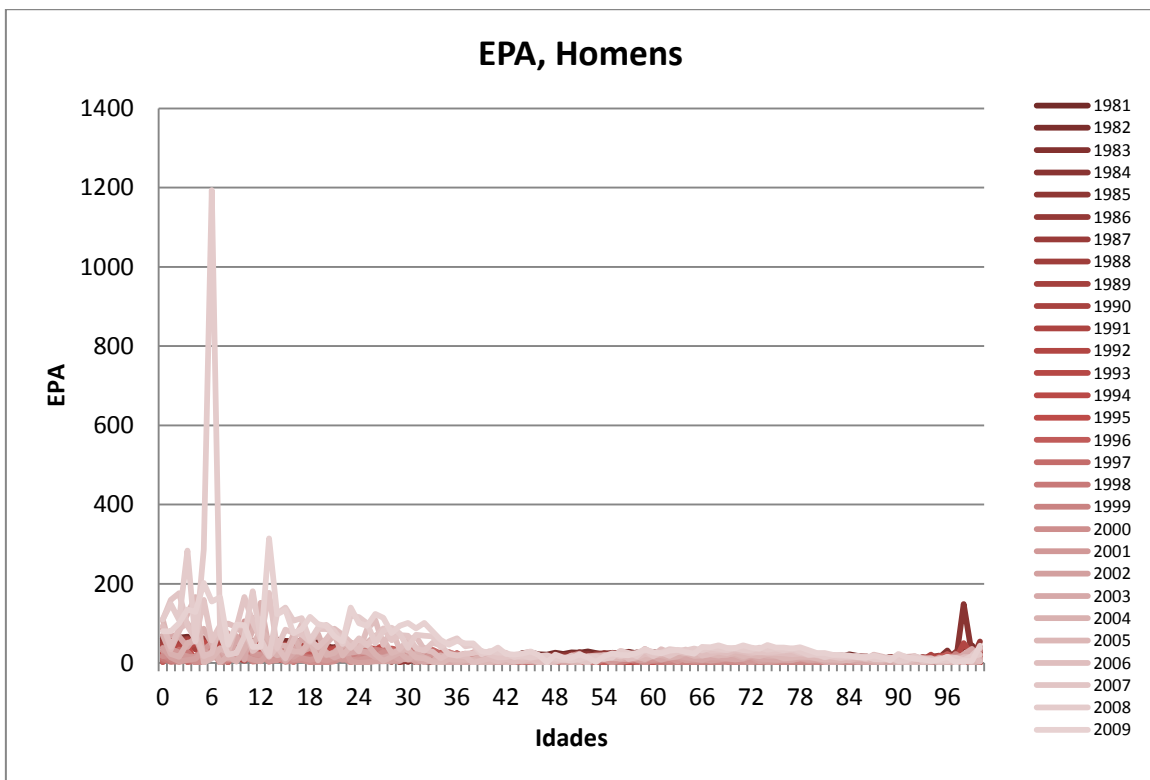
Anexo 25 – EPA nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.



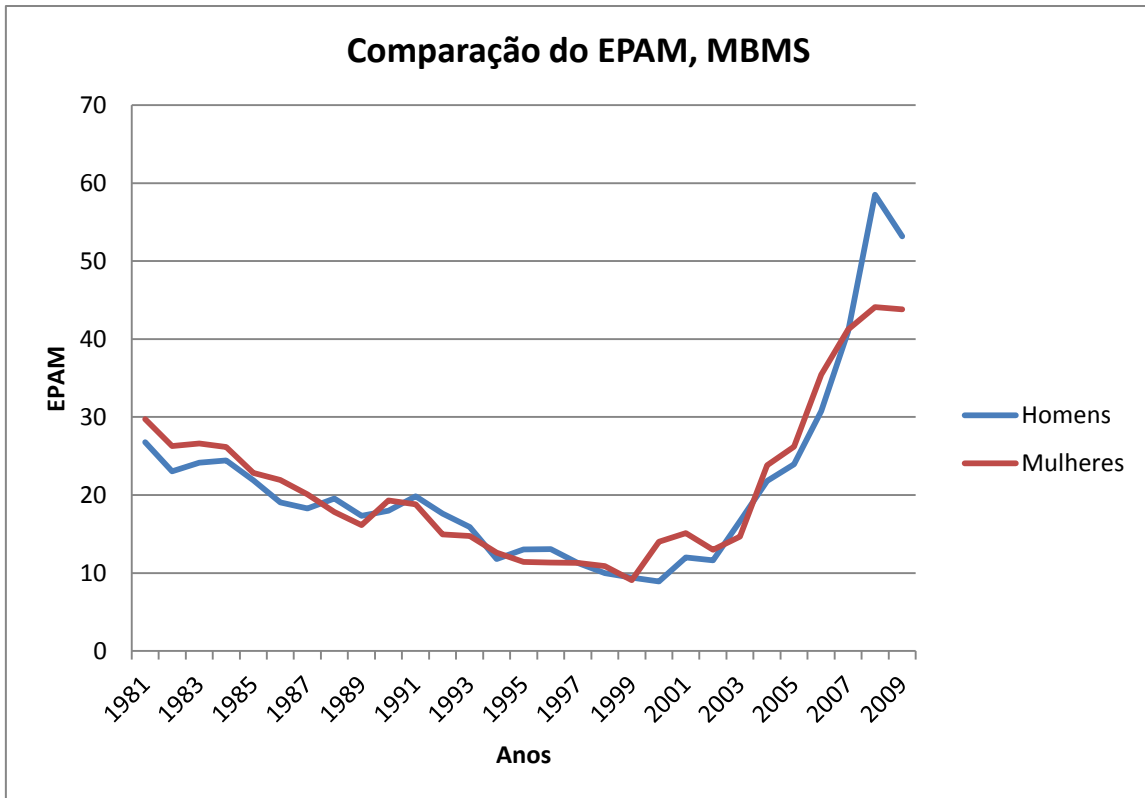
Anexo 26 – EPAL nos anos de ajustamento entre 1981 e 2009 do MBMS para a população total.



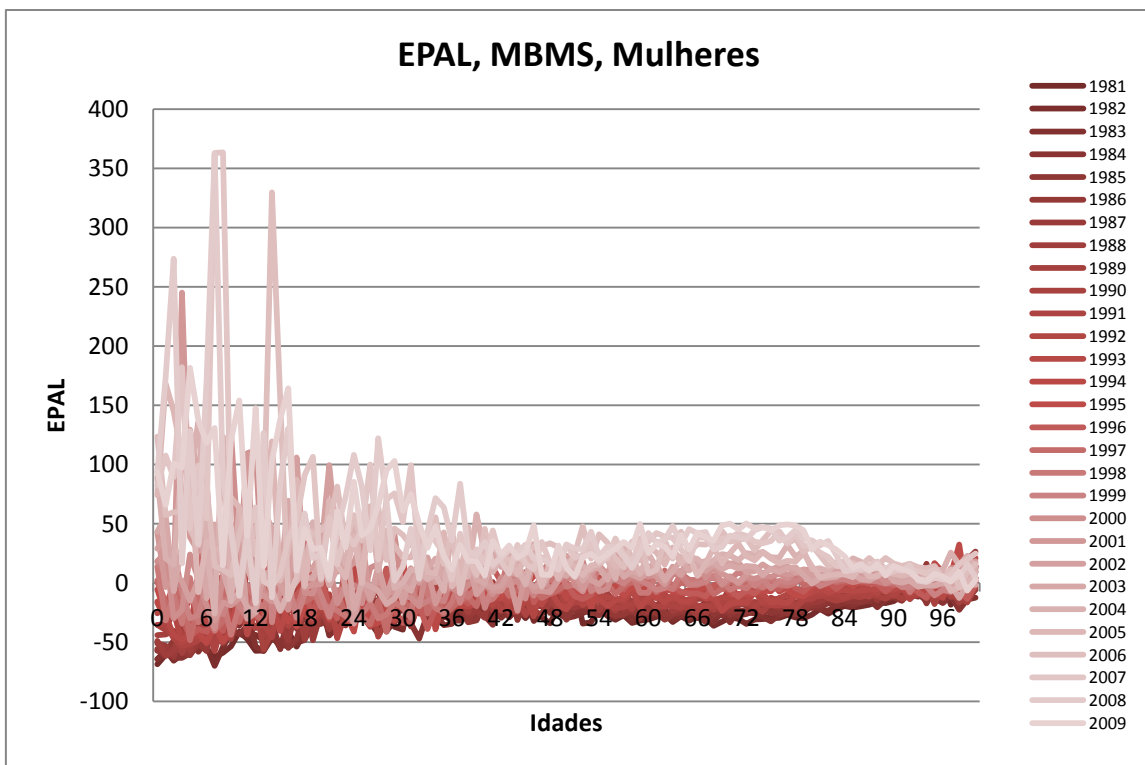
Anexo 27 – EPA nos anos de ajustamento entre 1981 e 2009 do MBMS para a população feminina.



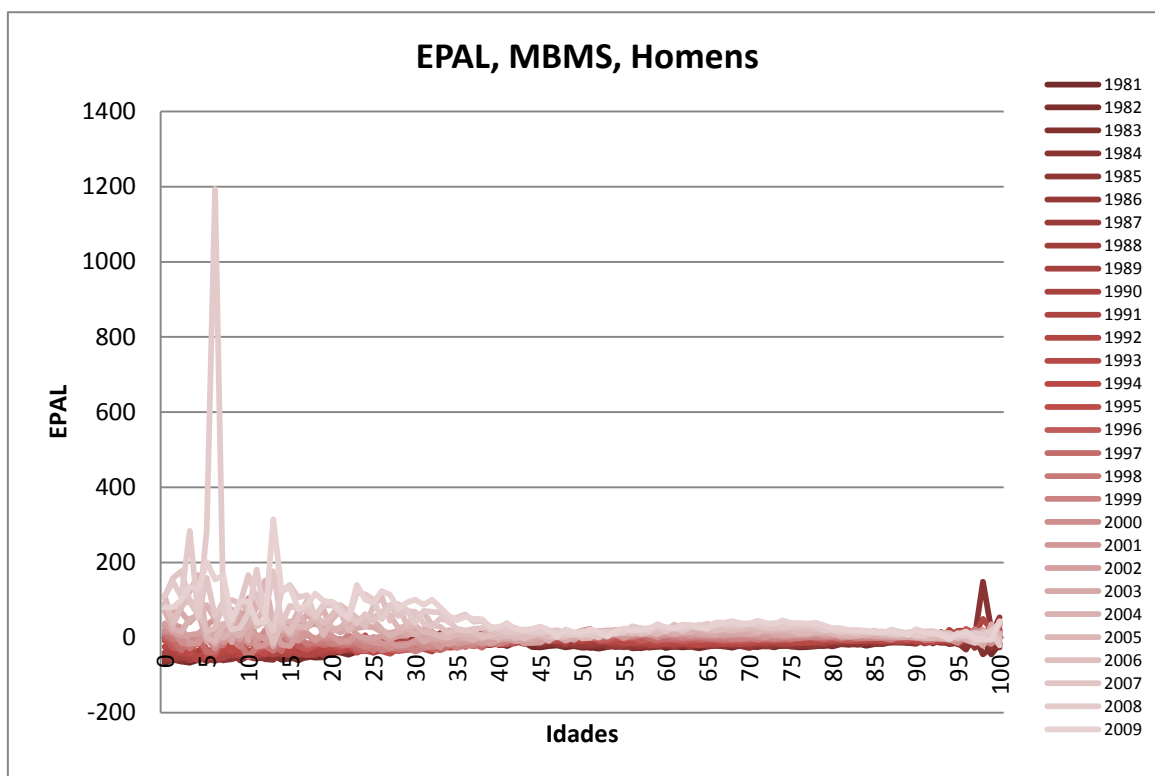
Anexo 28 – EPA nos anos de ajustamento entre 1981 e 2009 do MBMS para a população masculina.



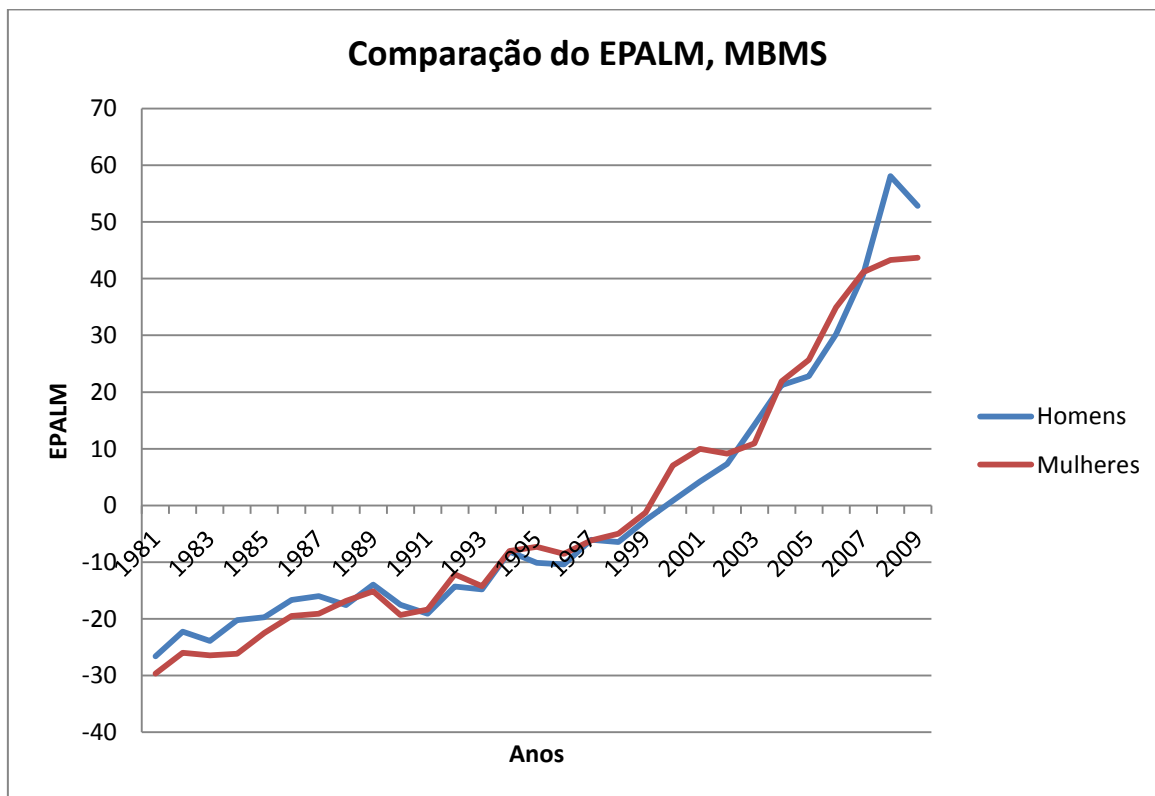
Anexo 29 – EPAM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1981 a 2009.



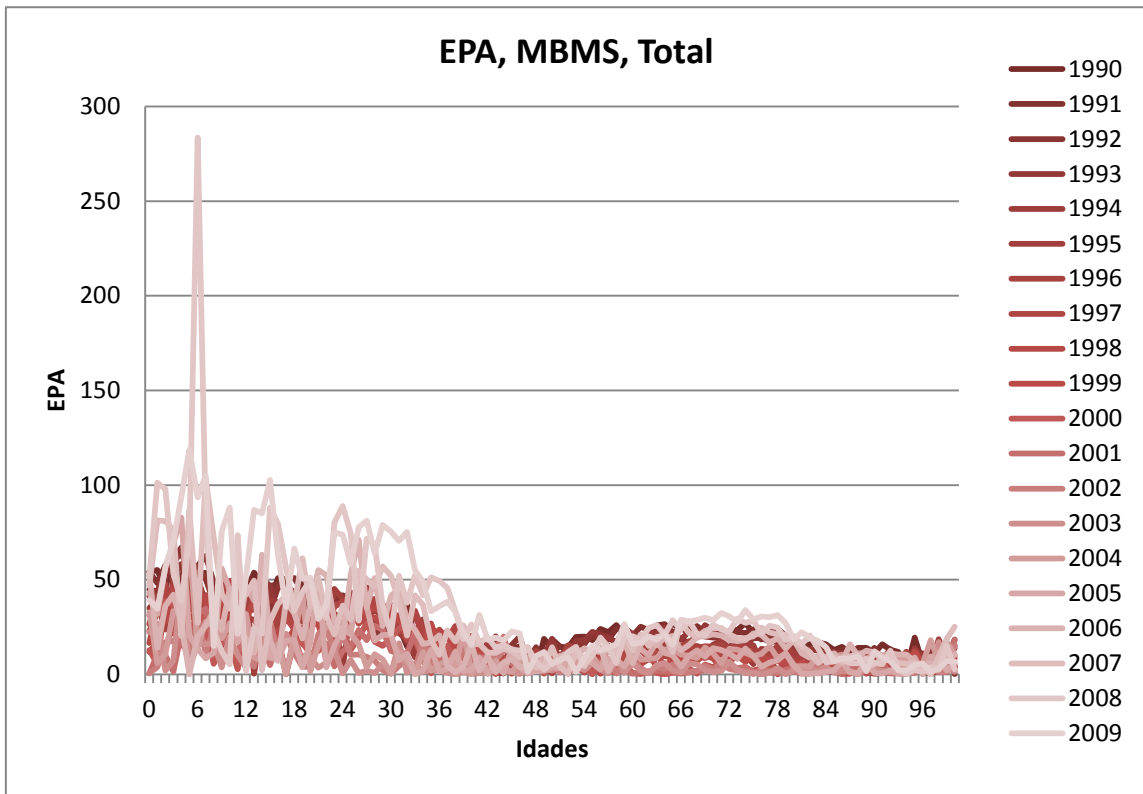
Anexo 30 – EPAL nos anos de ajustamento entre 1981 e 2009 do MBMS para a população feminina.



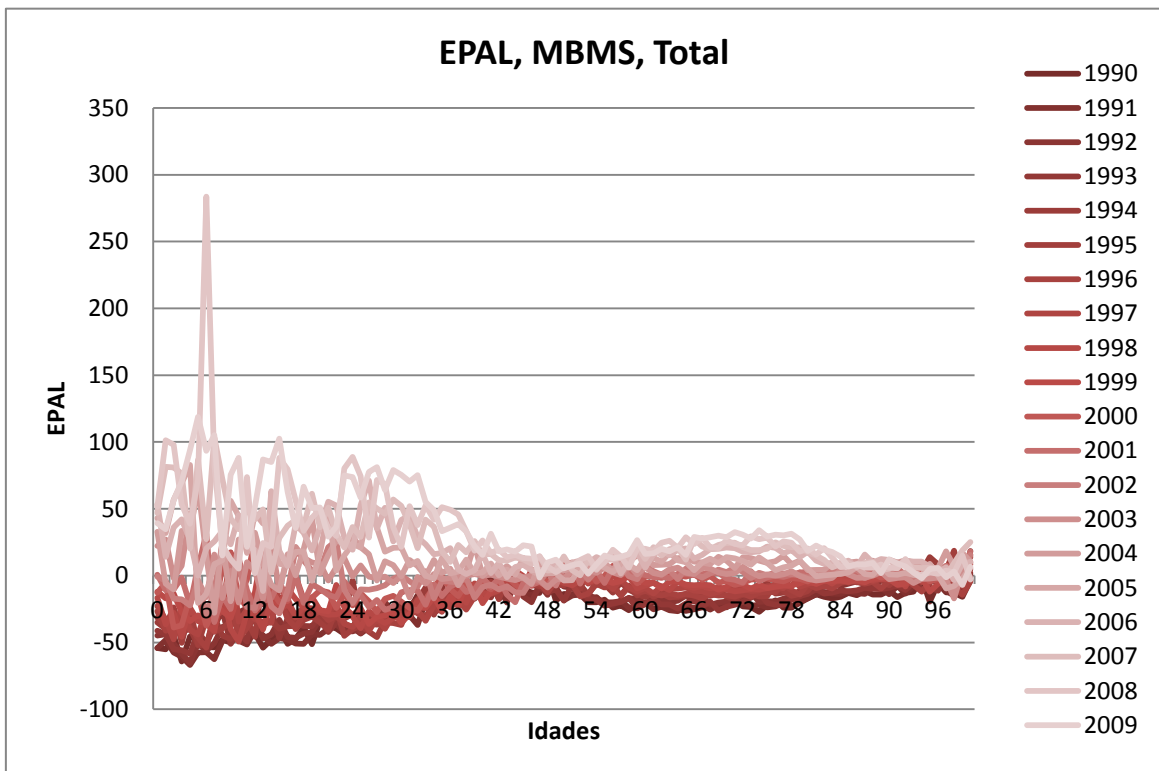
Anexo 31 – EPAL nos anos de ajustamento entre 1981 e 2009 do MBMS para a população masculina.



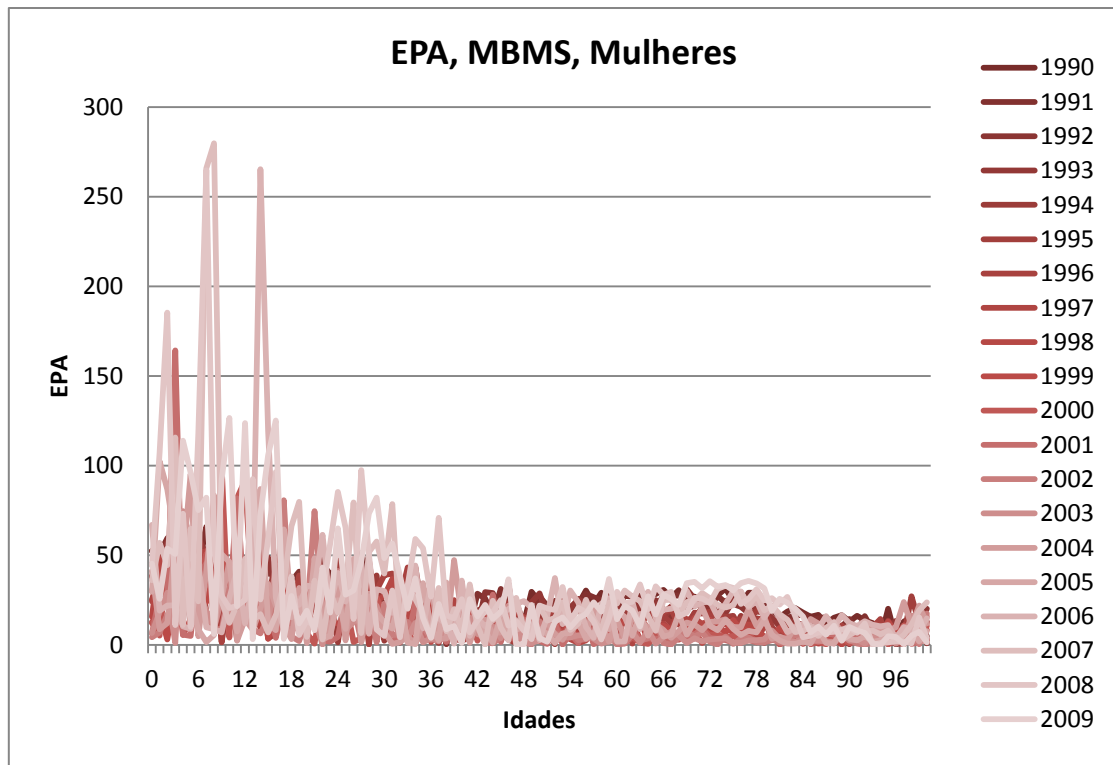
Anexo 32 – EPALM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1981 a 2009.



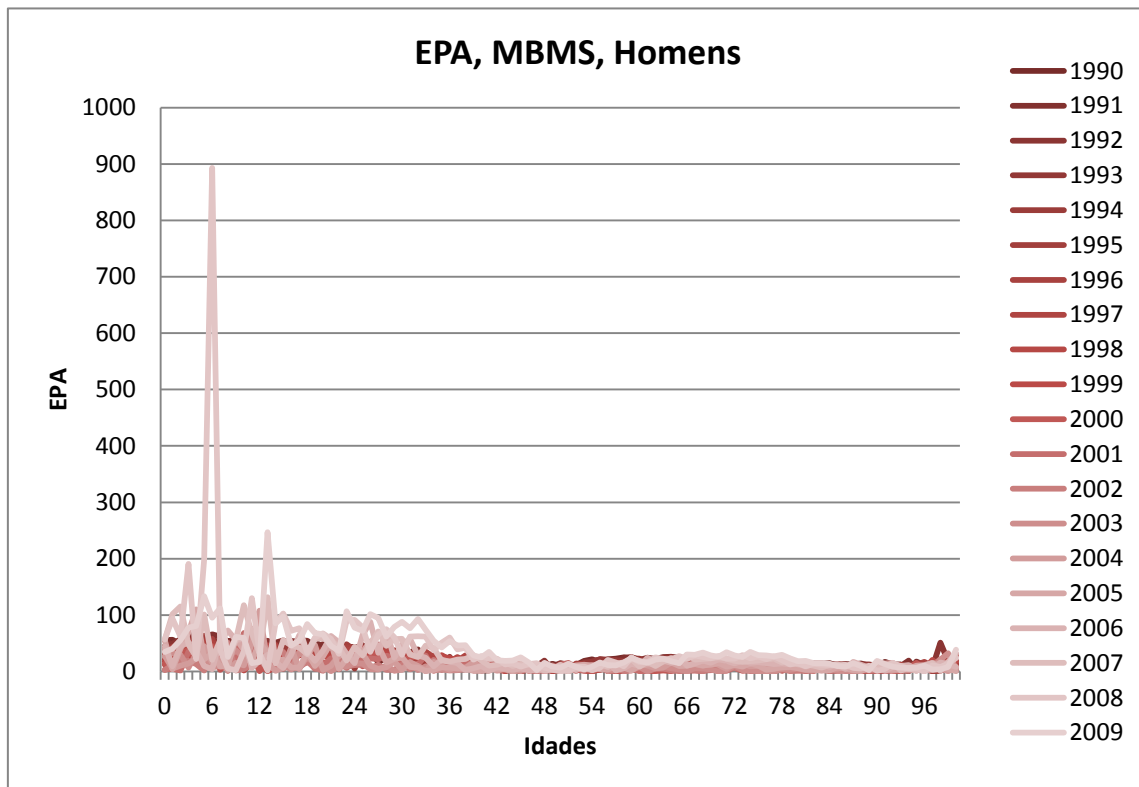
Anexo 33 – EPA nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.



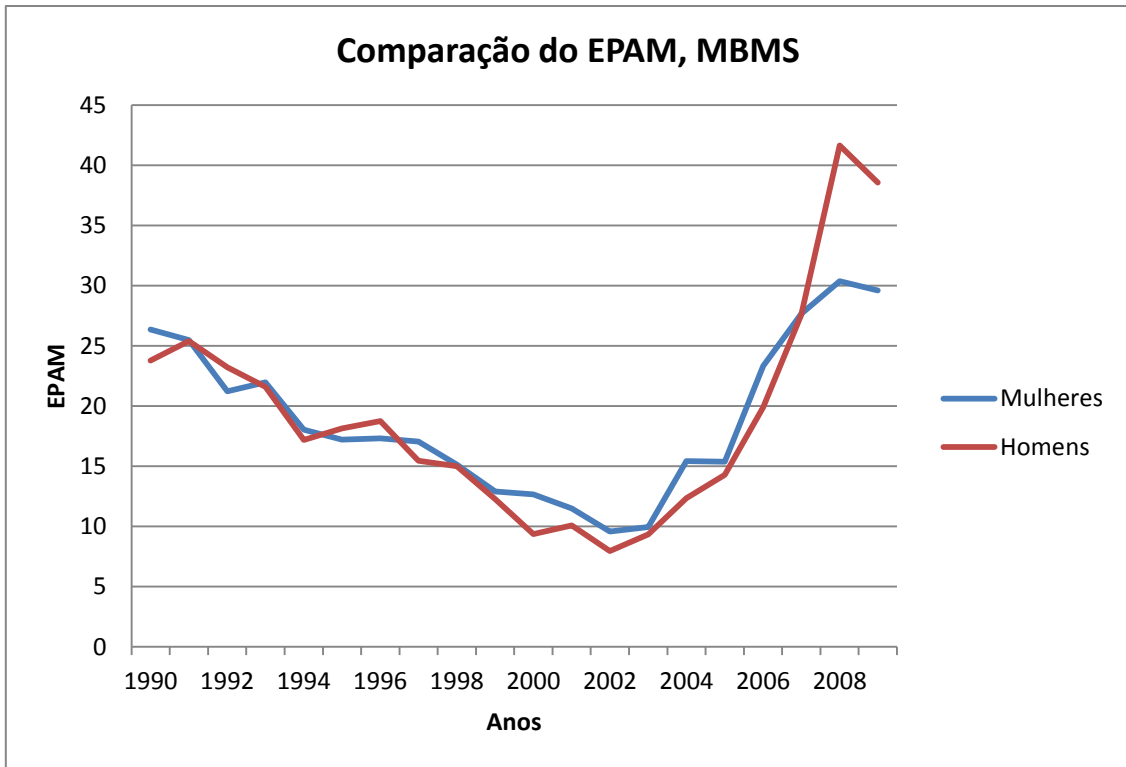
Anexo 34 – EPAL nos anos de ajustamento entre 1990 e 2009 do MBMS para a população total.



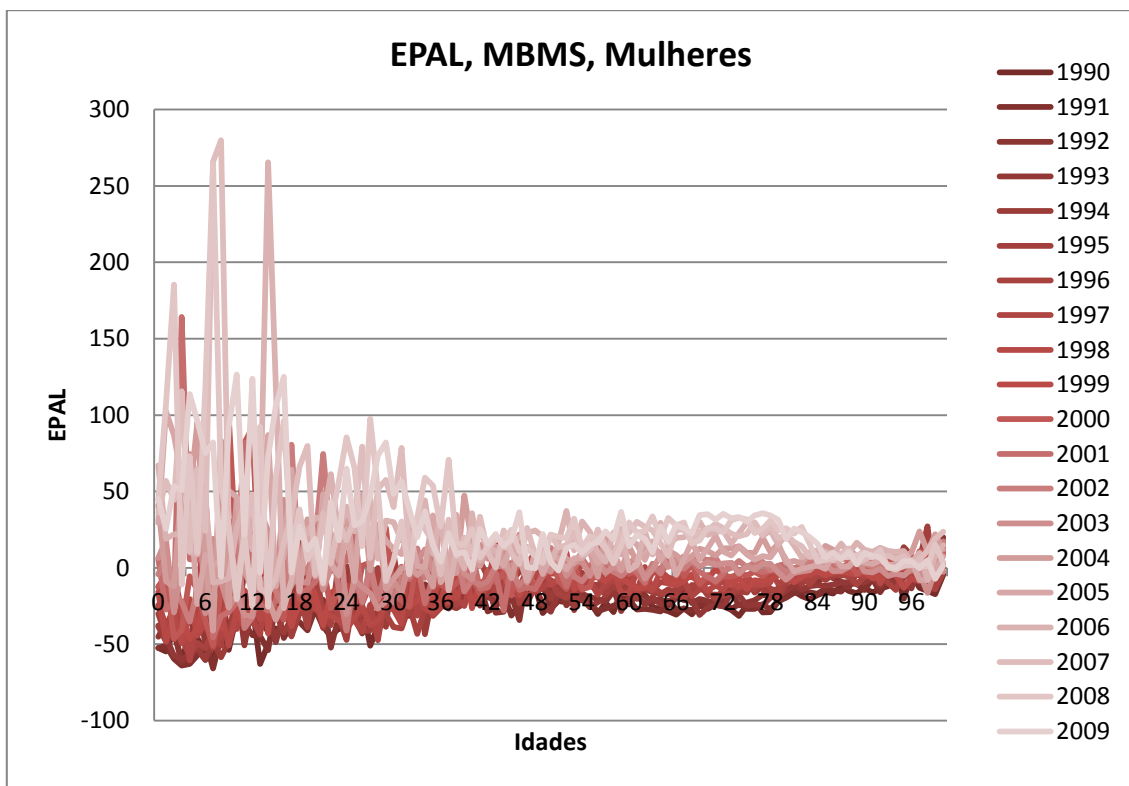
Anexo 35 – EPA nos anos de ajustamento entre 1990 e 2009 do MBMS para a população feminina.



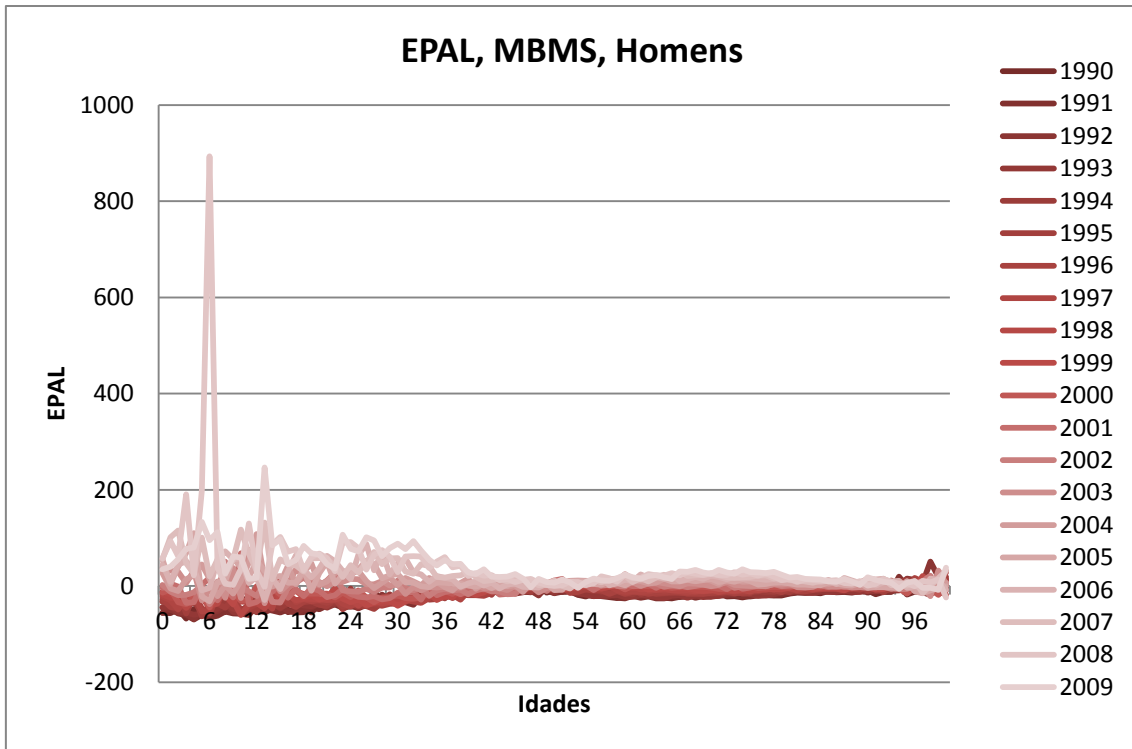
Anexo 36 – EPA nos anos de ajustamento entre 1990 e 2009 do MBMS para a população masculina.



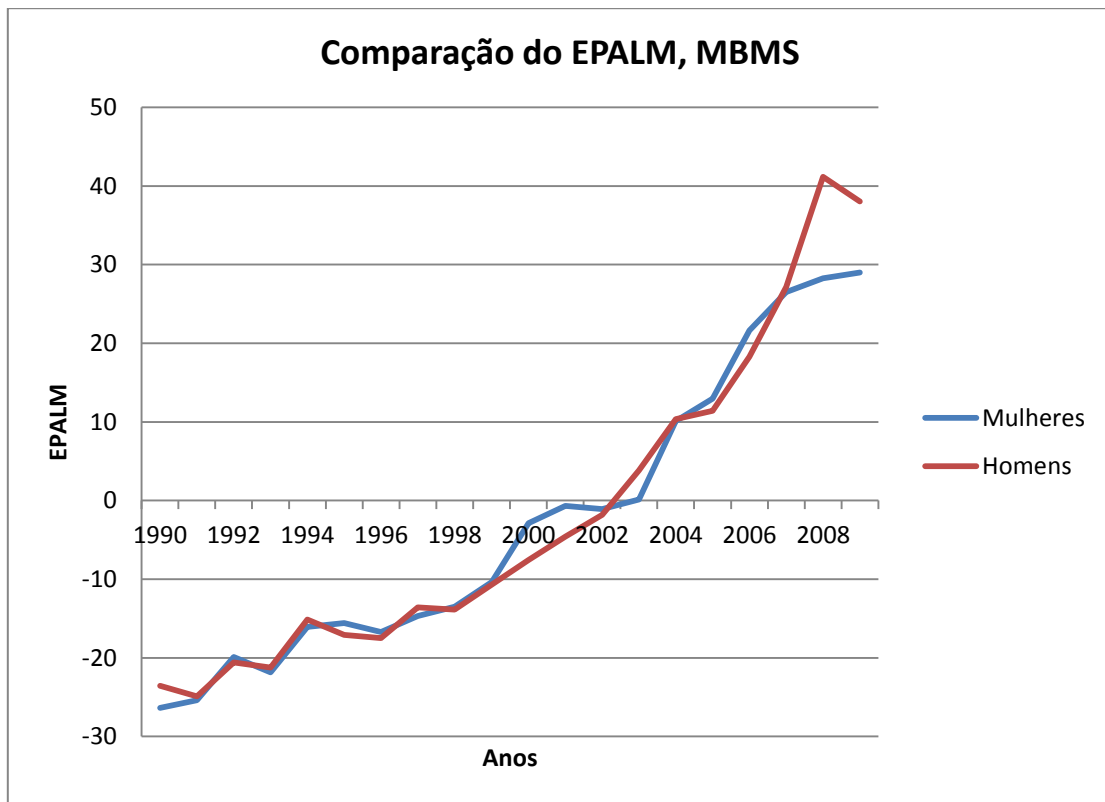
Anexo 37 – EPAM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1990 a 2009.



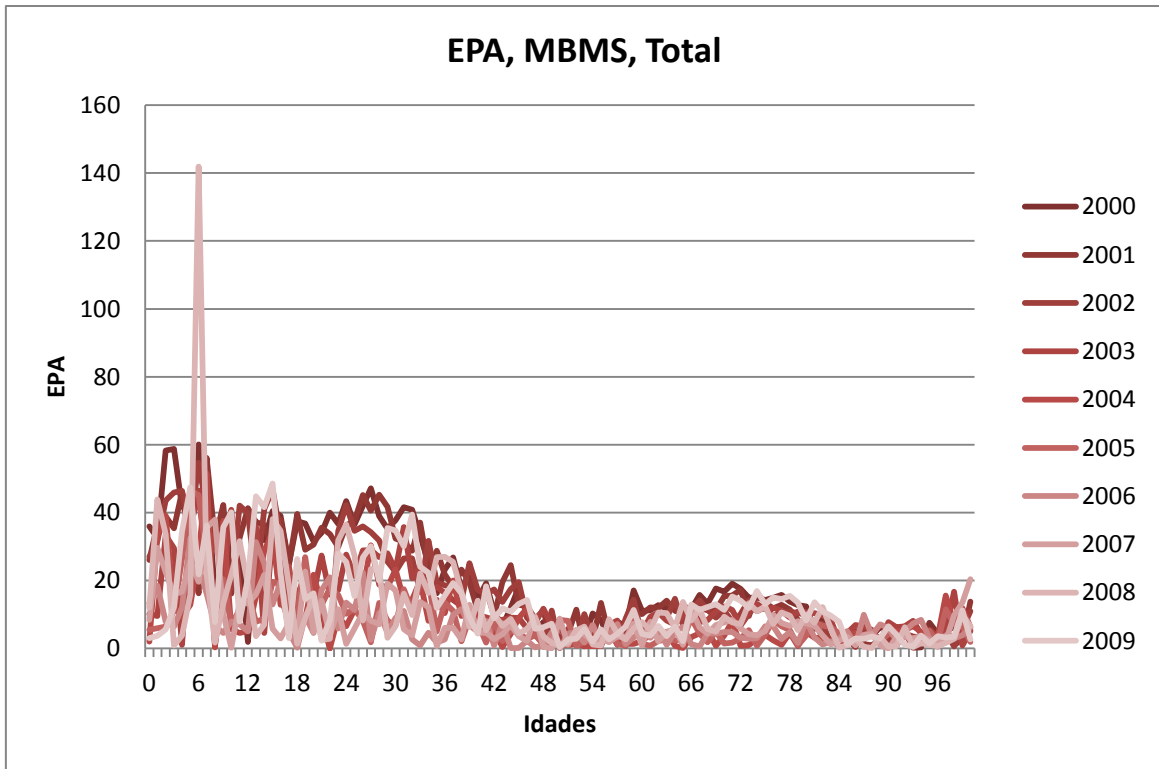
Anexo 38 – EPAL nos anos de ajustamento entre 1990 e 2009 do MBMS para a população feminina.



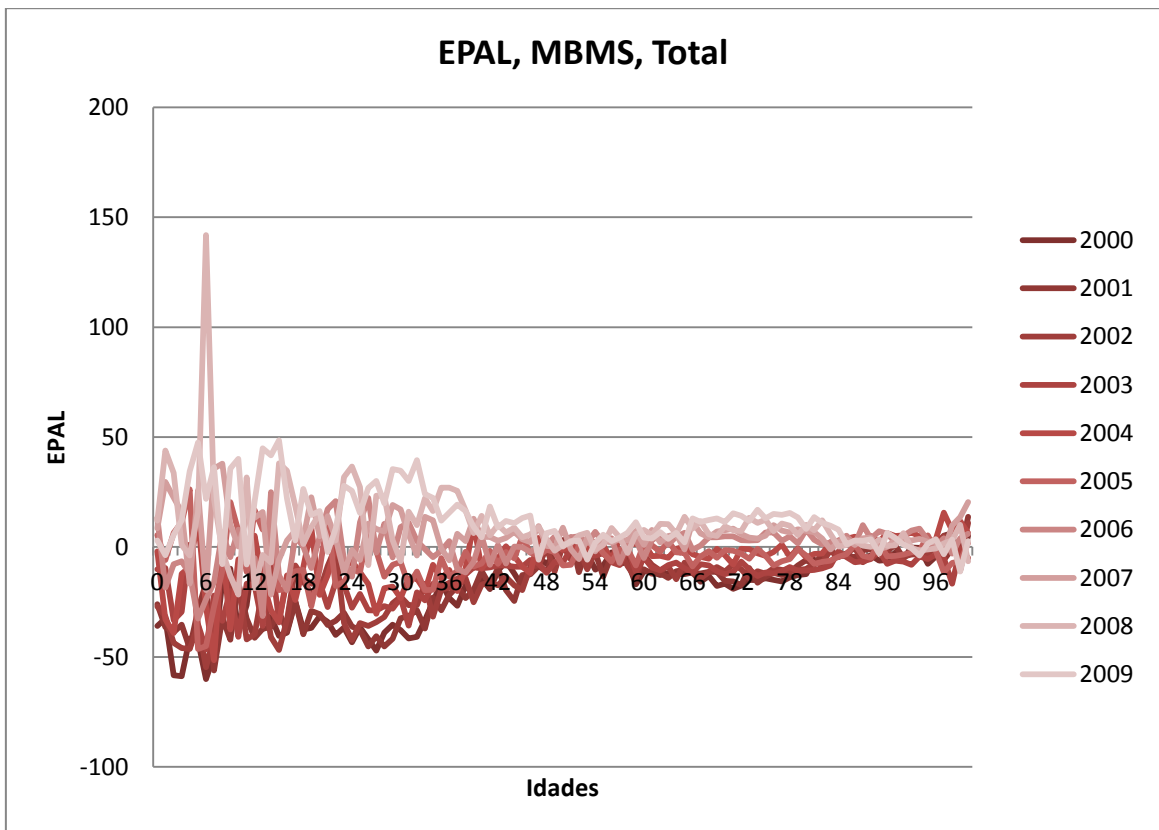
Anexo 39 – EPAL nos anos de ajustamento entre 1990 e 2009 do MBMS para a população masculina.



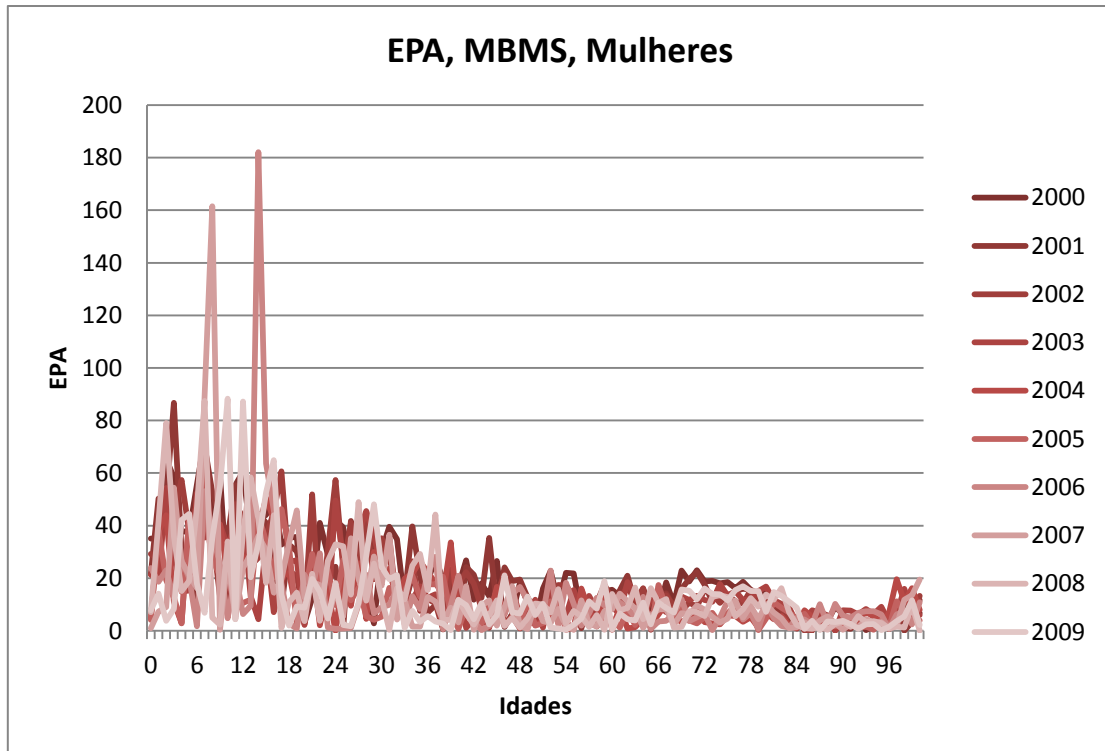
Anexo 40 – EPALM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 1990 a 2009.



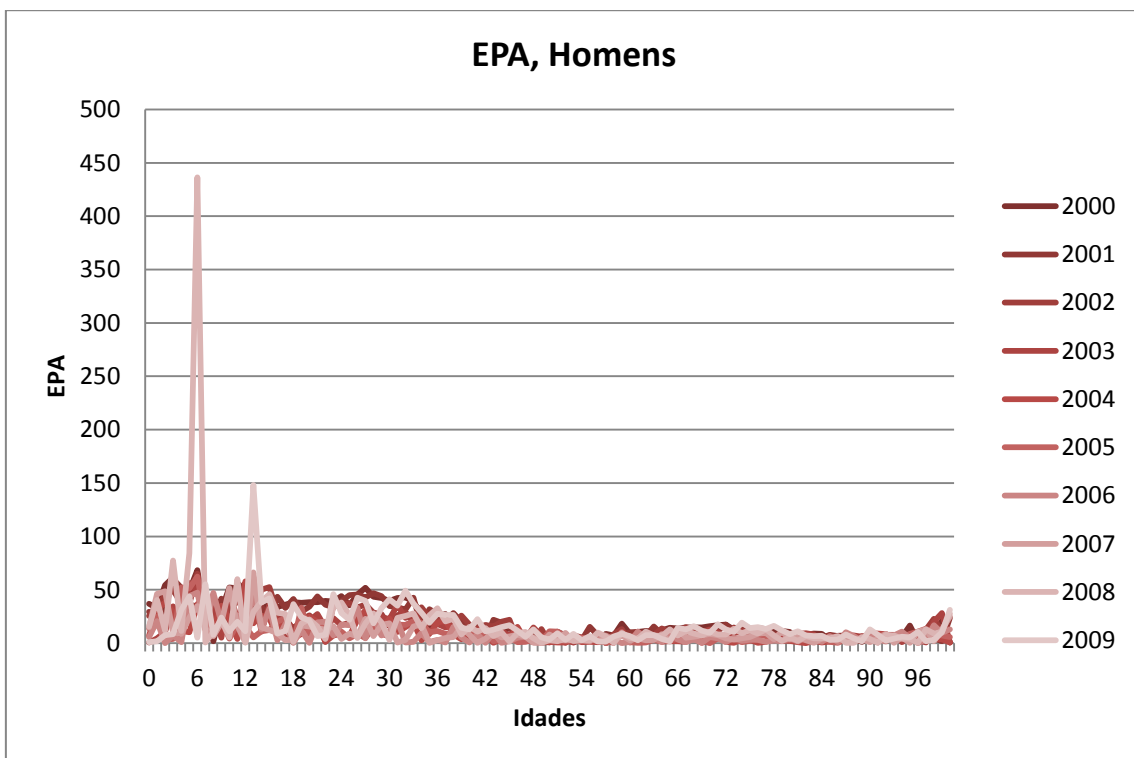
Anexo 41 – EPA nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.



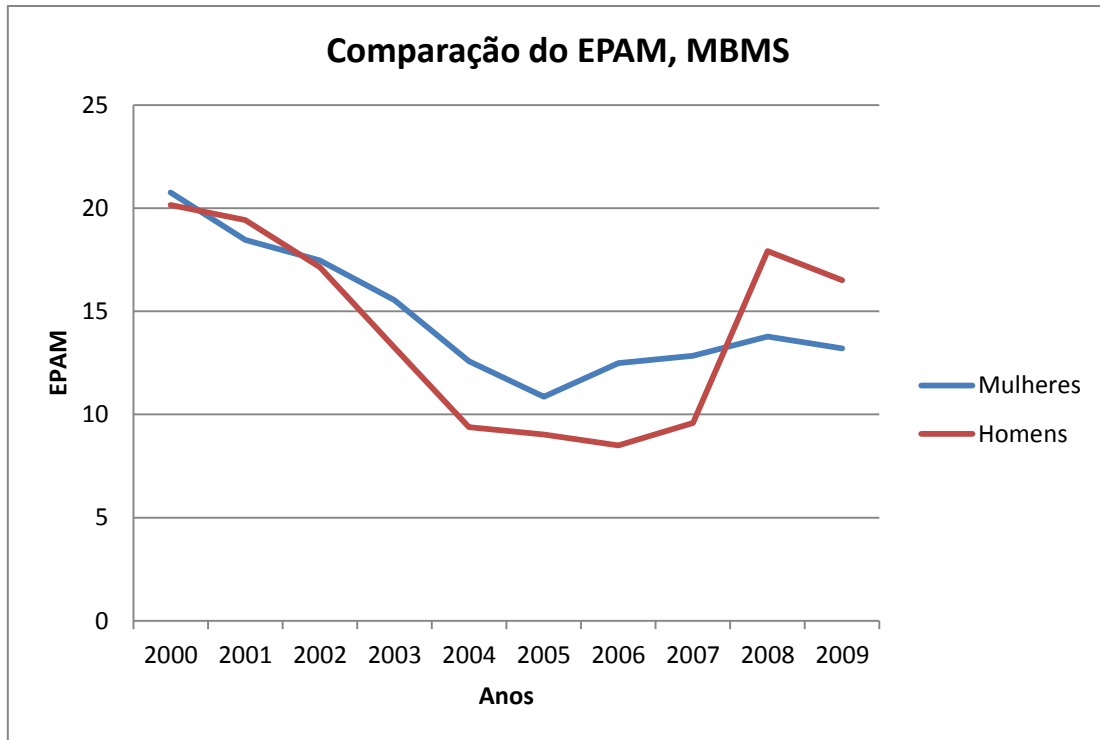
Anexo 42 – EPAL nos anos de ajustamento entre 2000 e 2009 do MBMS para a população total.



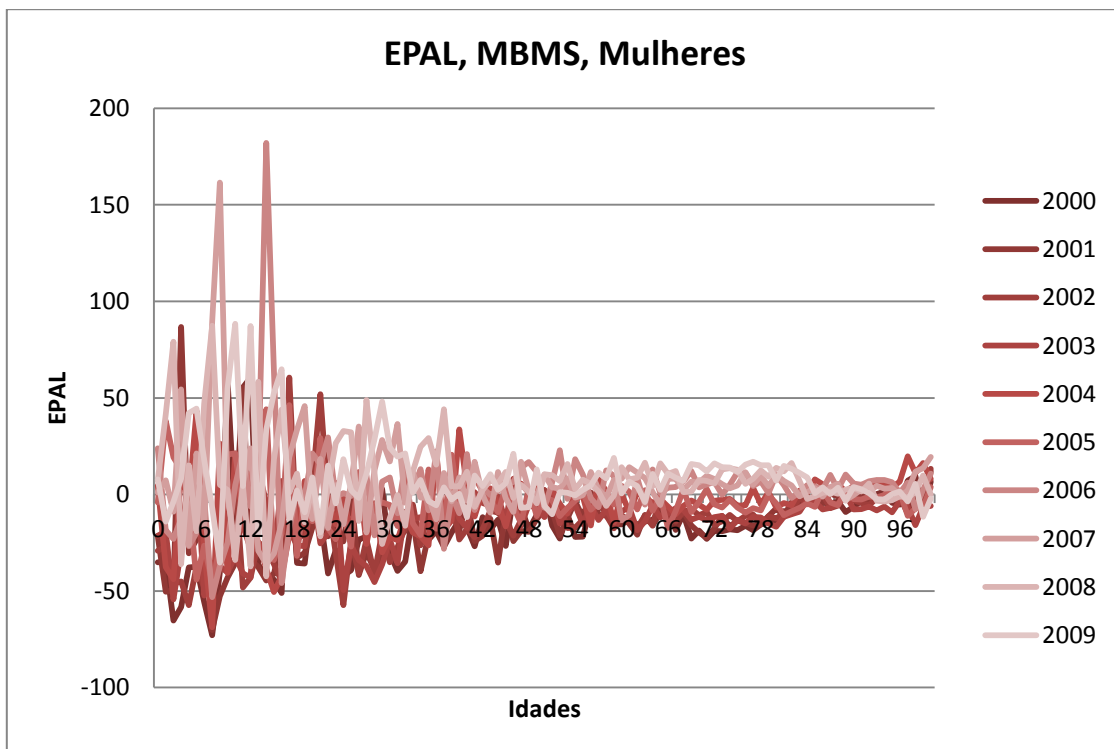
Anexo 43 – EPA nos anos de ajustamento entre 2000 e 2009 do MBMS para a população feminina.



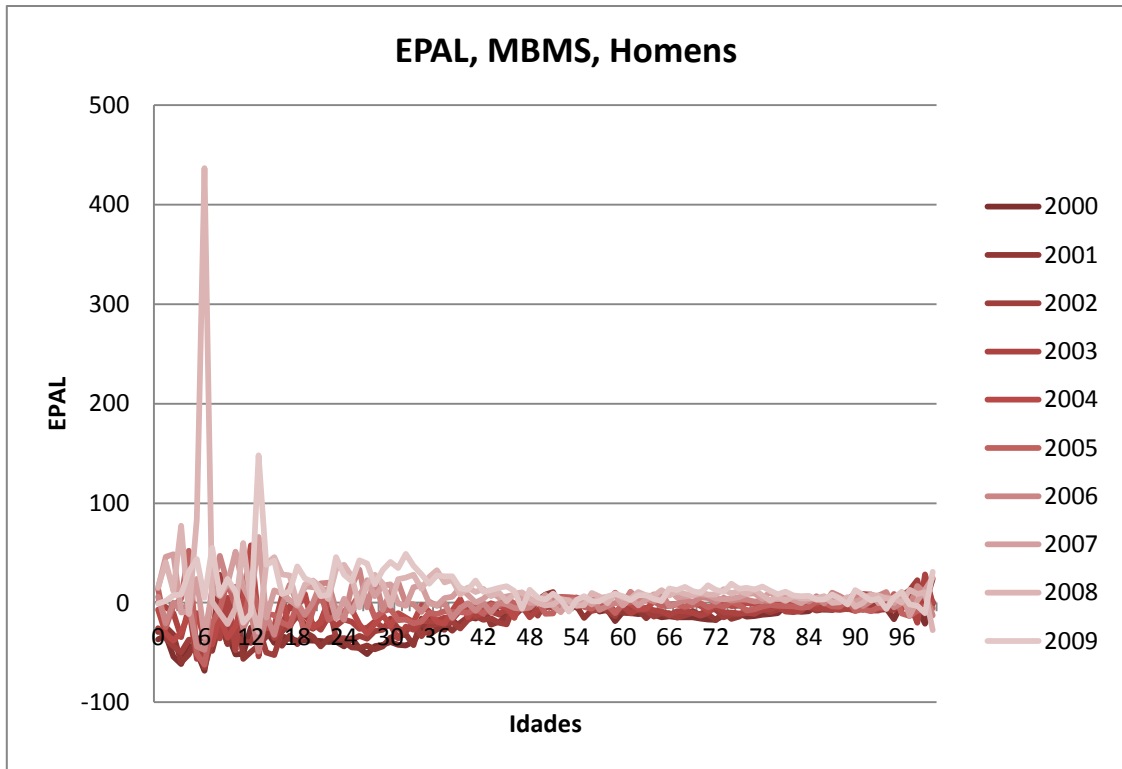
Anexo 44 – EPA nos anos de ajustamento entre 2000 e 2009 do MBMS para a população masculina.



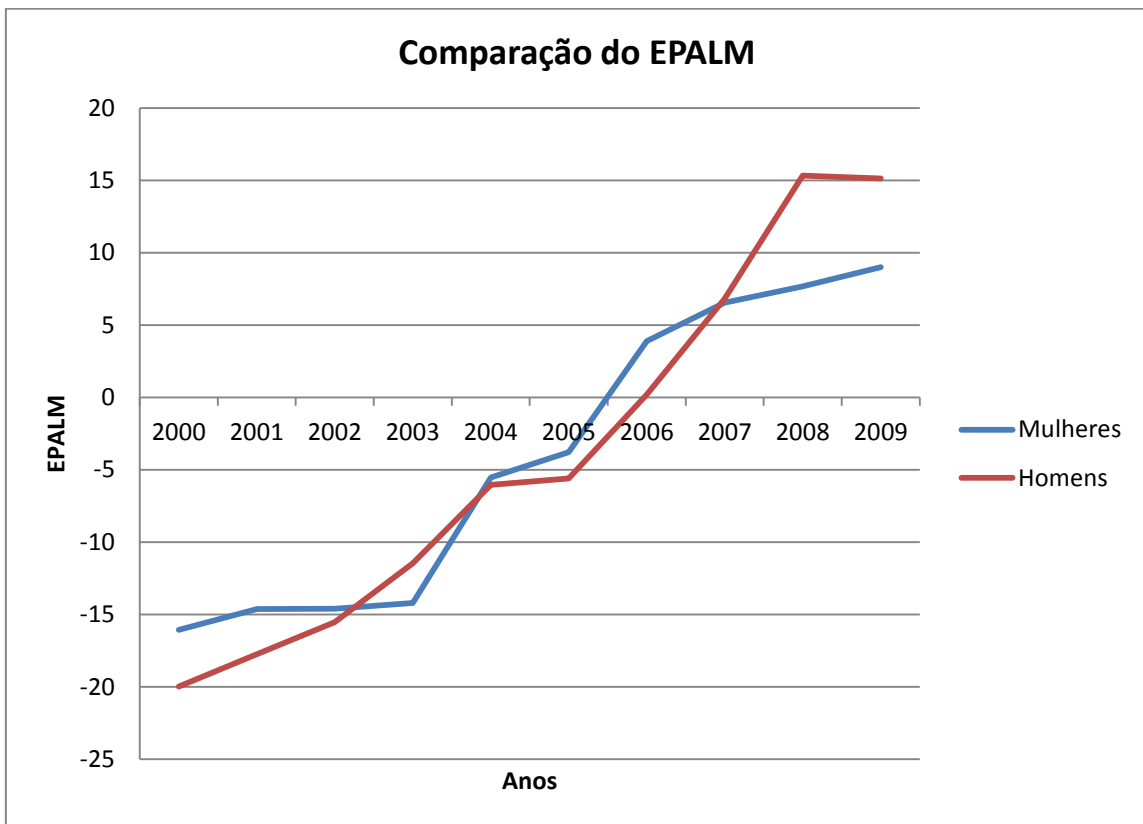
Anexo 45 – EPAM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 2000 a 2009.



Anexo 46 – EPAL nos anos de ajustamento entre 2000 e 2009 do MBMS para a população feminina.



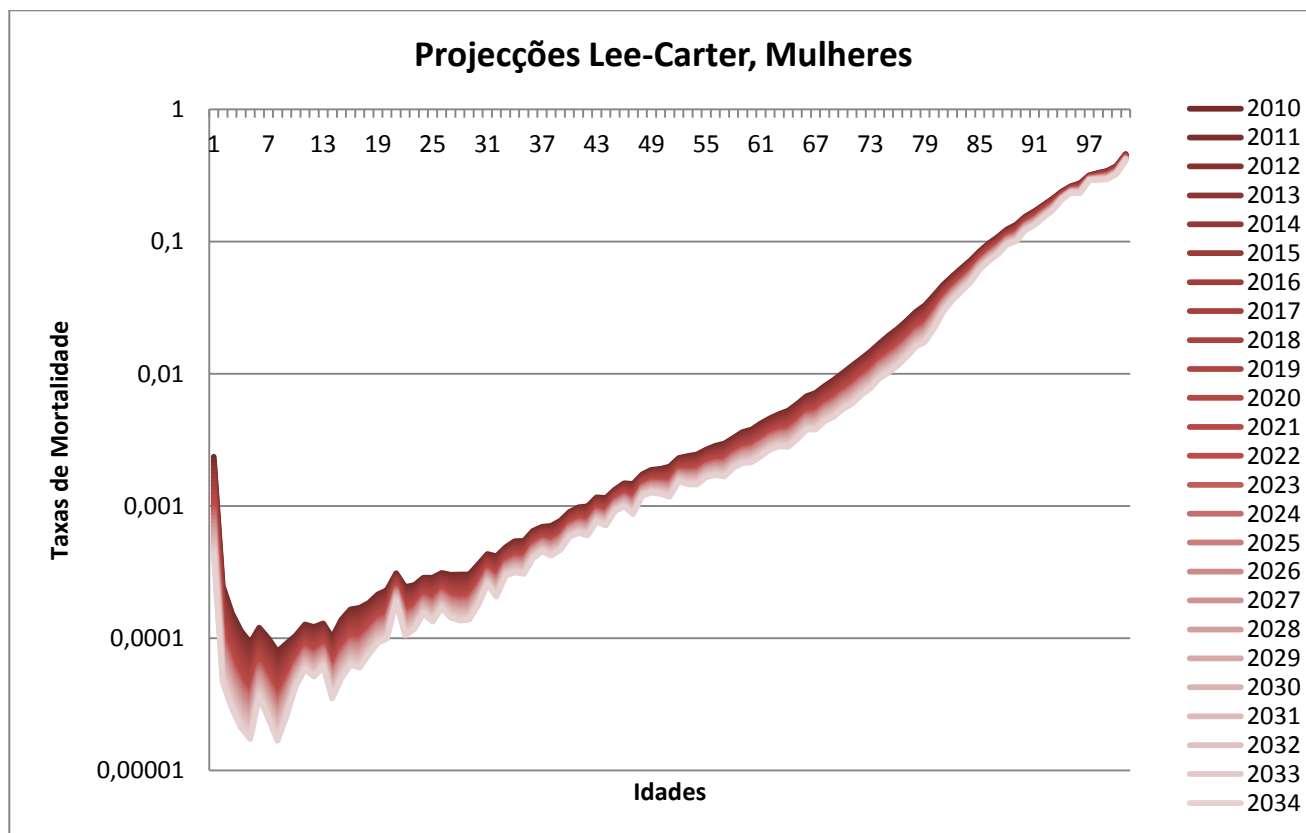
Anexo 47 – EPAL nos anos de ajustamento entre 2000 e 2009 do MBMS para a população masculina.



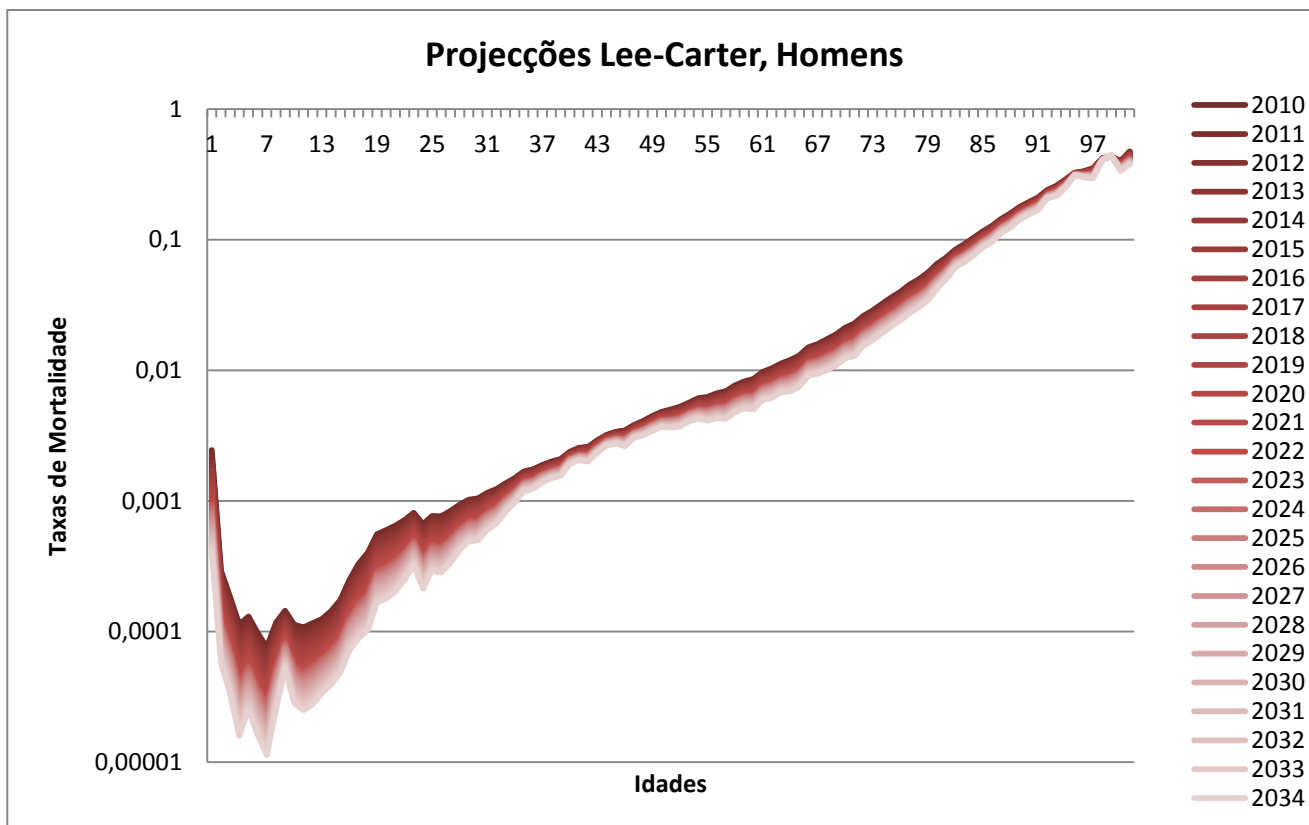
Anexo 48 – EPALM, comparação entre os sexos M e F, do MBMS com anos de ajustamento de 2000 a 2009.

		1981-2009		1990-2009		2000-2009	
		LC	BMS	LC	BMS	LC	BMS
EPALM	Mulheres	-1,8540	-1,8539	-2,8165	-2,8167	-4,1662	-4,1667
	Homens	-1,1695	-1,1689	-2,0946	-2,0944	-3,8866	-3,8865
	Total	-2,1871	-2,1869	-3,0740	-3,0741	-4,7609	-4,7611
EPAM	Mulheres	21,1569	21,1549	18,9036	18,9014	14,8028	14,8017
	Homens	21,1336	21,1305	19,0894	19,0865	14,0924	14,0925
	Total	19,5904	19,5872	17,5672	17,5645	12,3638	12,3628
EQM	Mulheres	0,000134	0,000134	0,000103	0,000103	0,000059	0,000059
	Homens	0,000435	0,000435	0,000214	0,000214	0,000187	0,000187
	Total	0,000127	0,000127	0,000095	0,000095	0,000056	0,000056

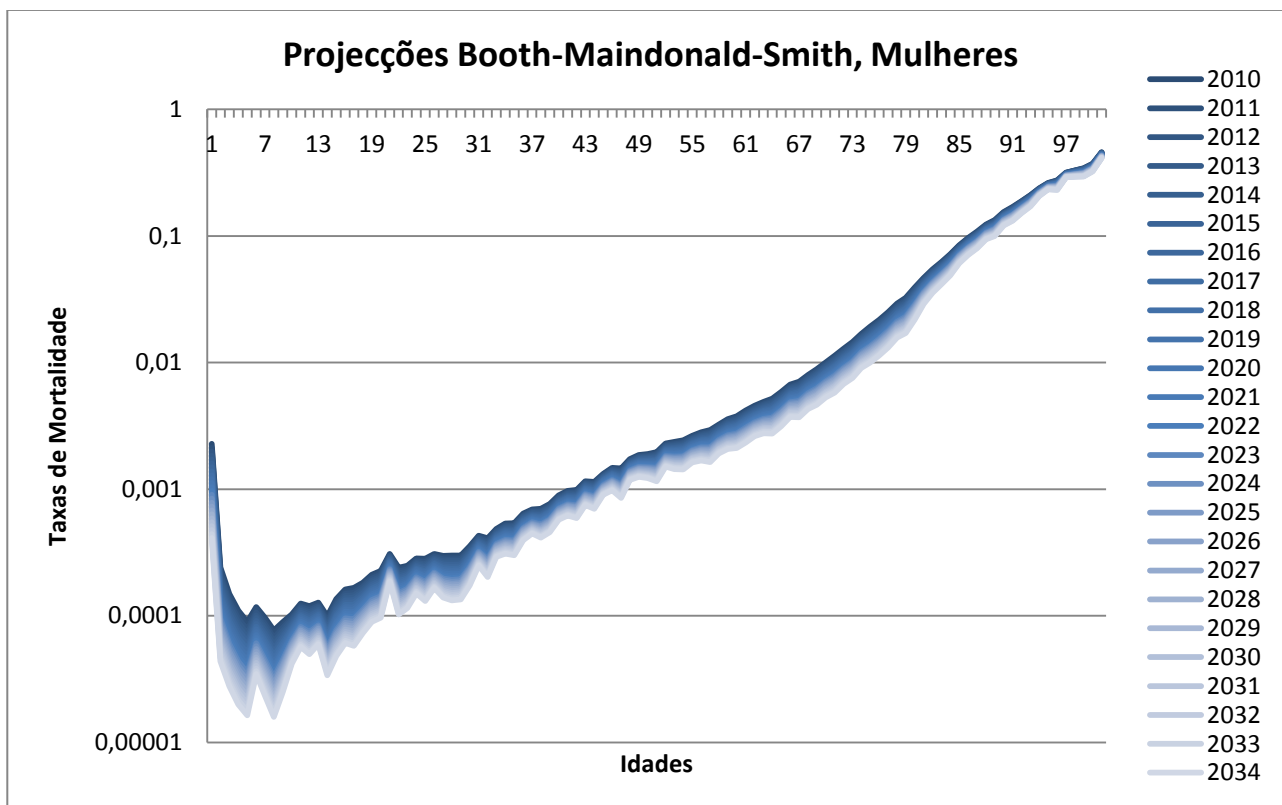
Anexo 49 – Resumo dos erros associados para todos os modelos.



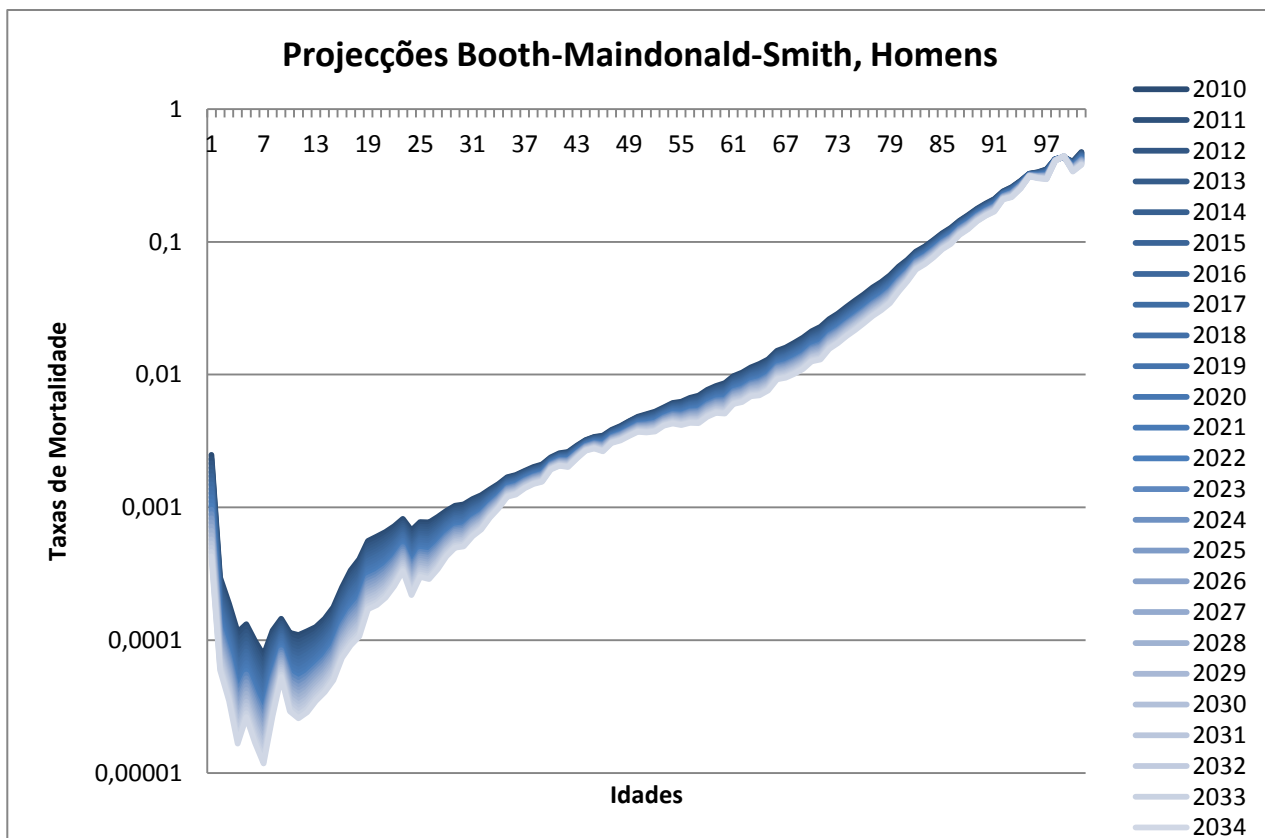
Anexo 50 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1981 a 2009 para população feminina.



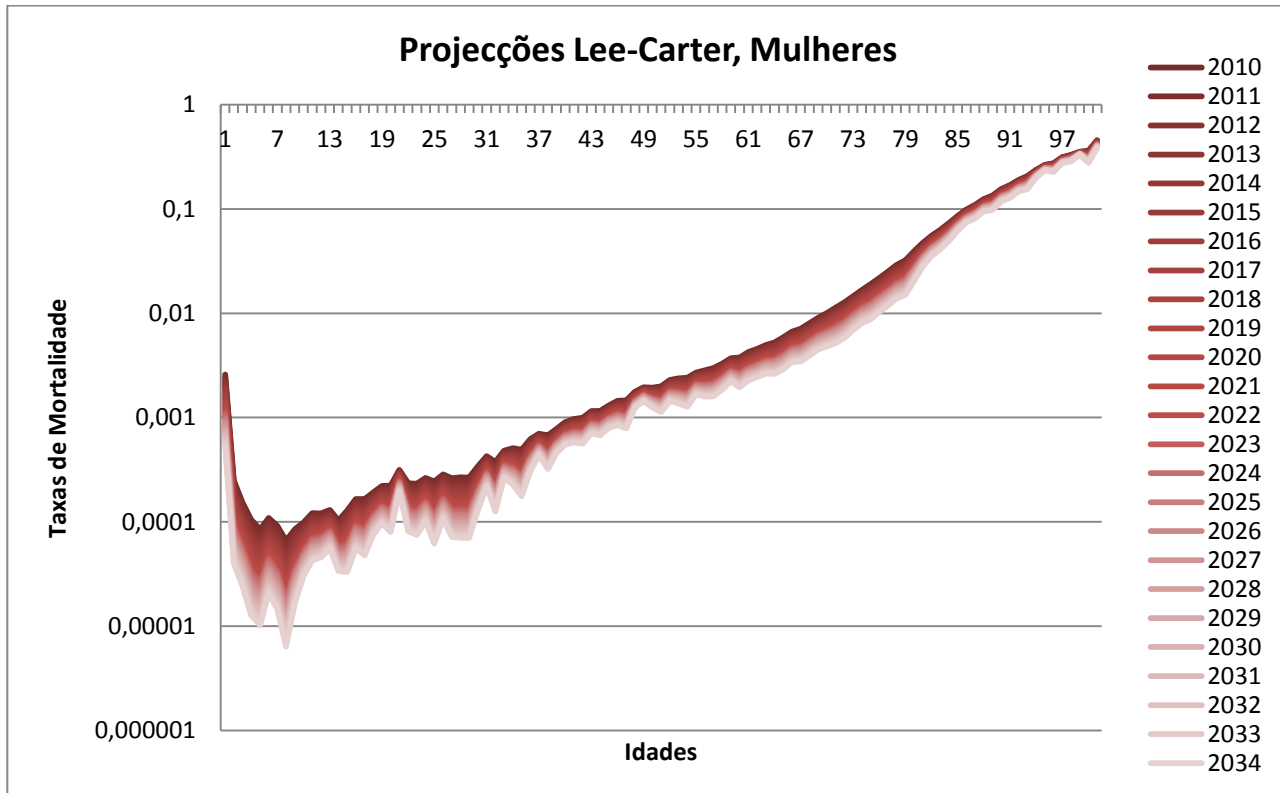
Anexo 51 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1981 a 2009 para população masculina.



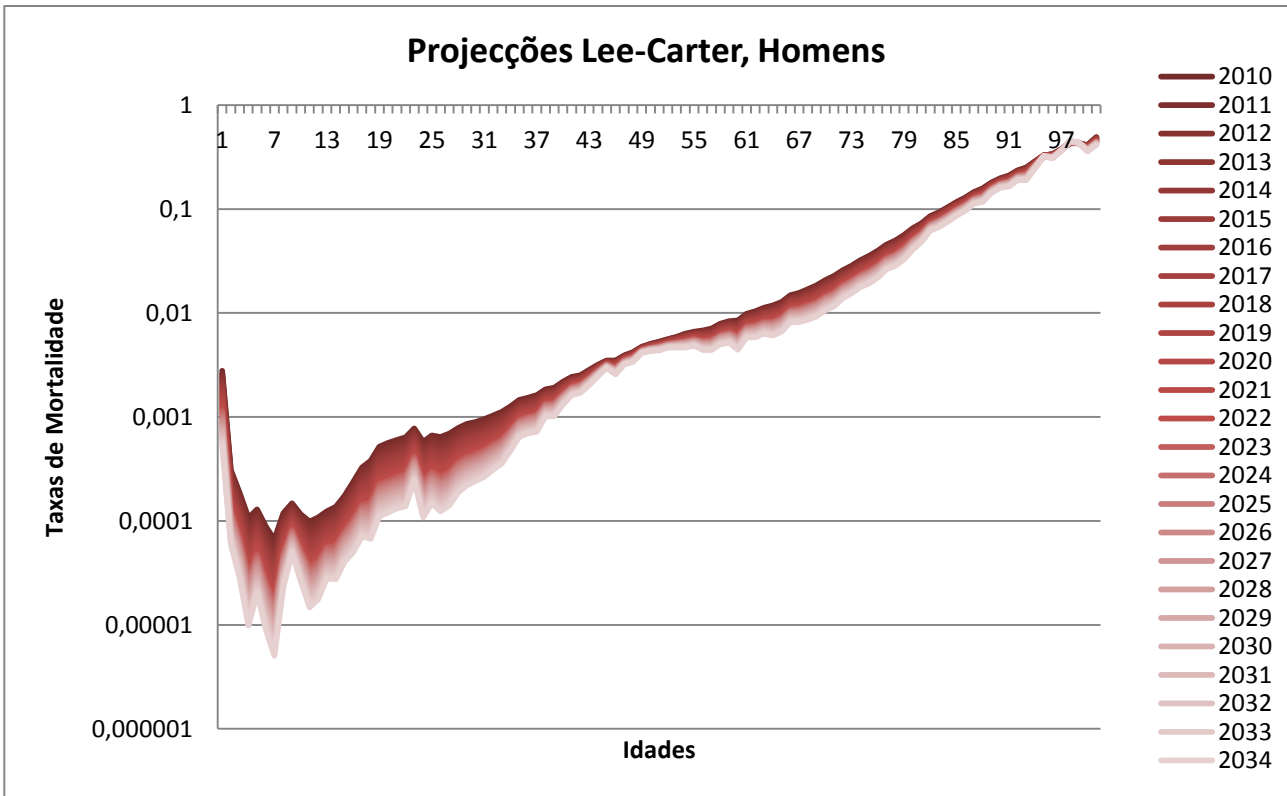
Anexo 52 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1981 a 2009 para população feminina.



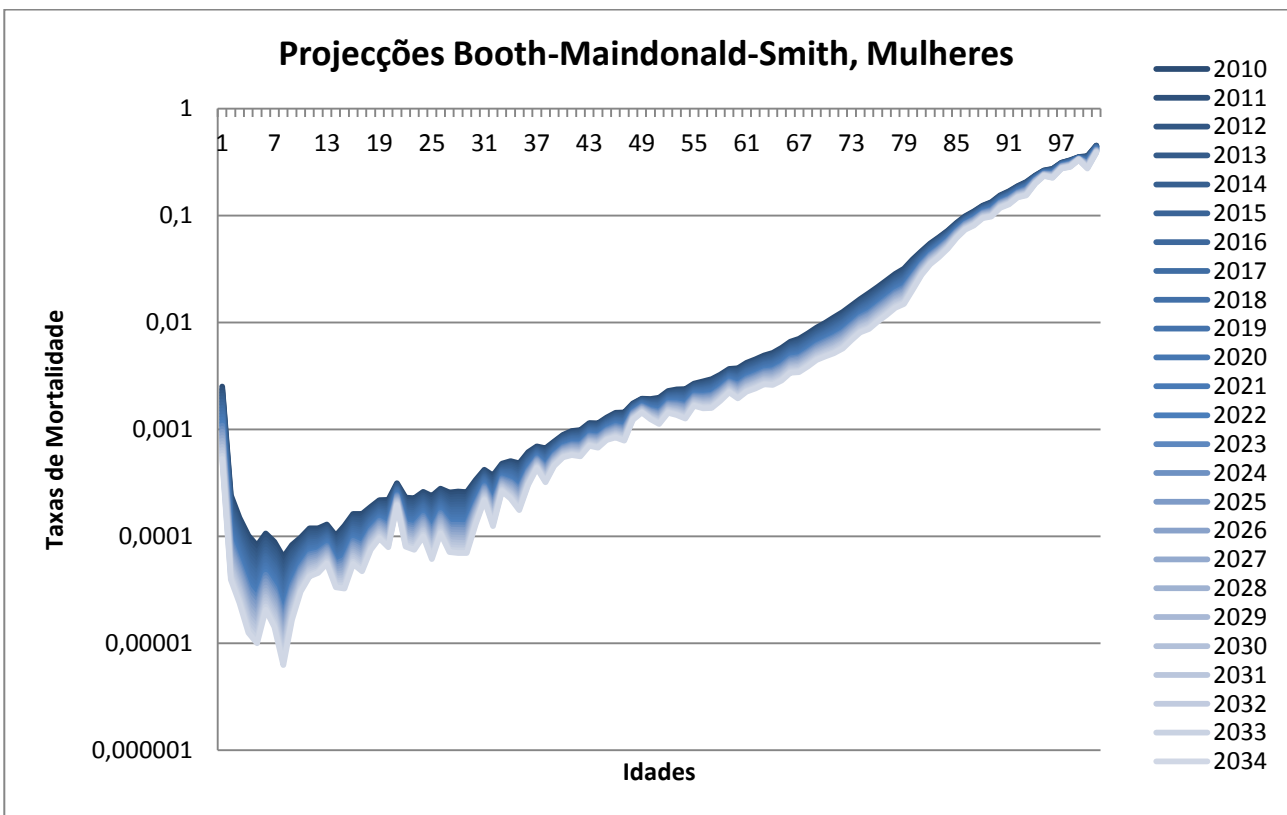
Anexo 53 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1981 a 2009 para população masculina.



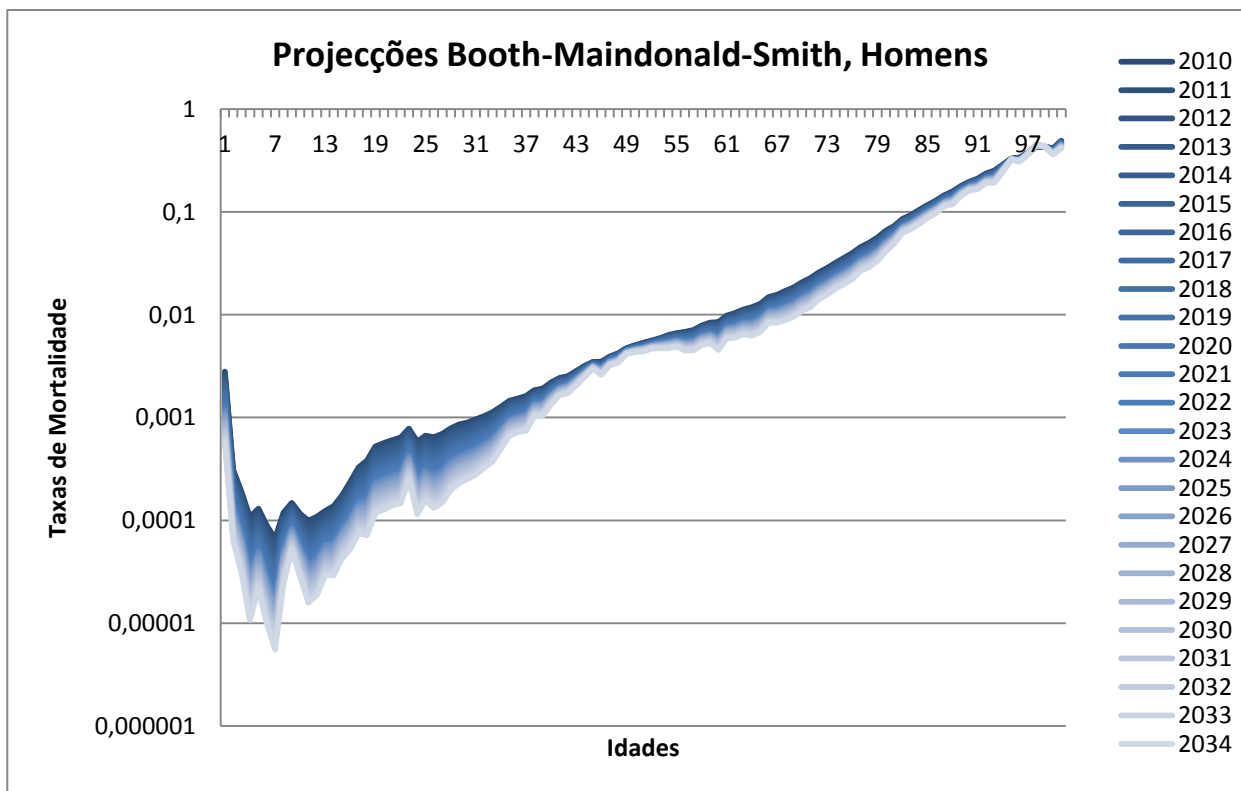
Anexo 54 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1990 a 2009 para população feminina.



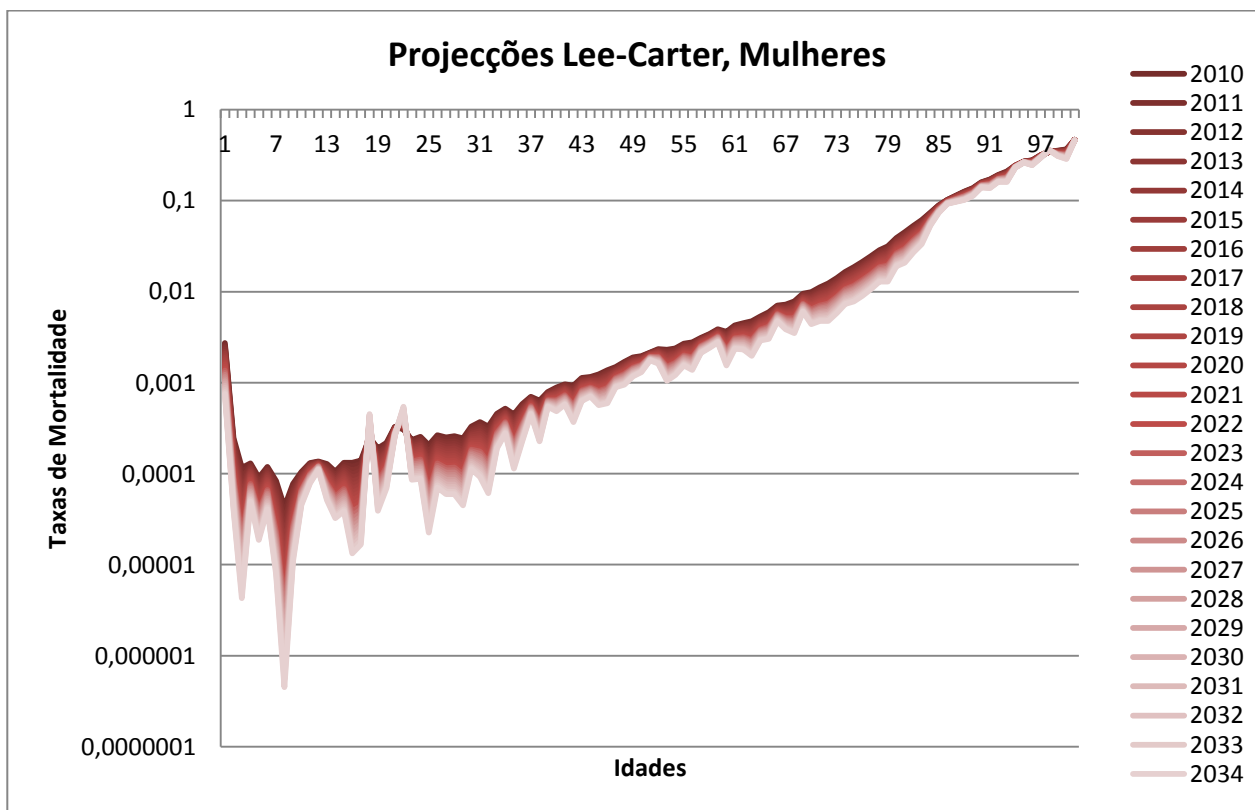
Anexo 55 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 1990 a 2009 para população masculina.



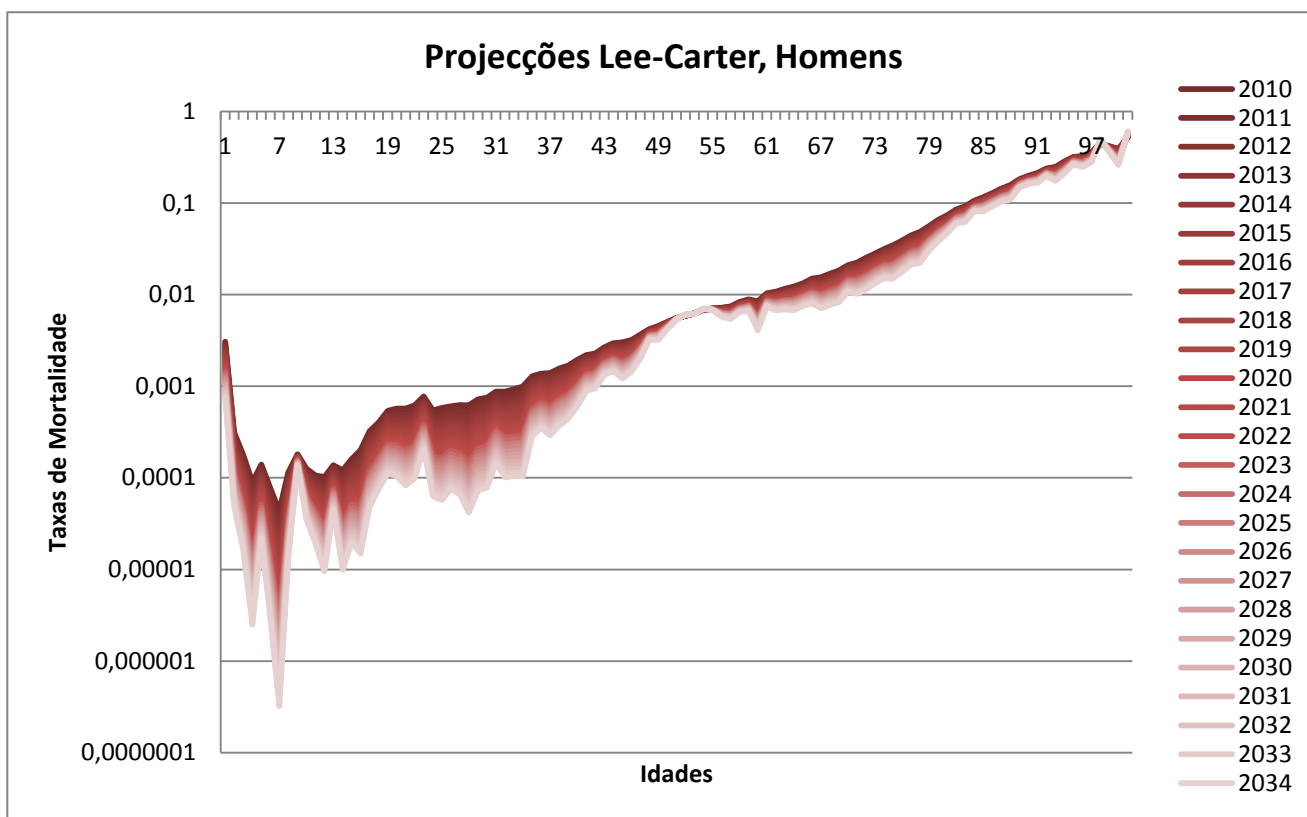
Anexo 56 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1990 a 2009 para população feminina.



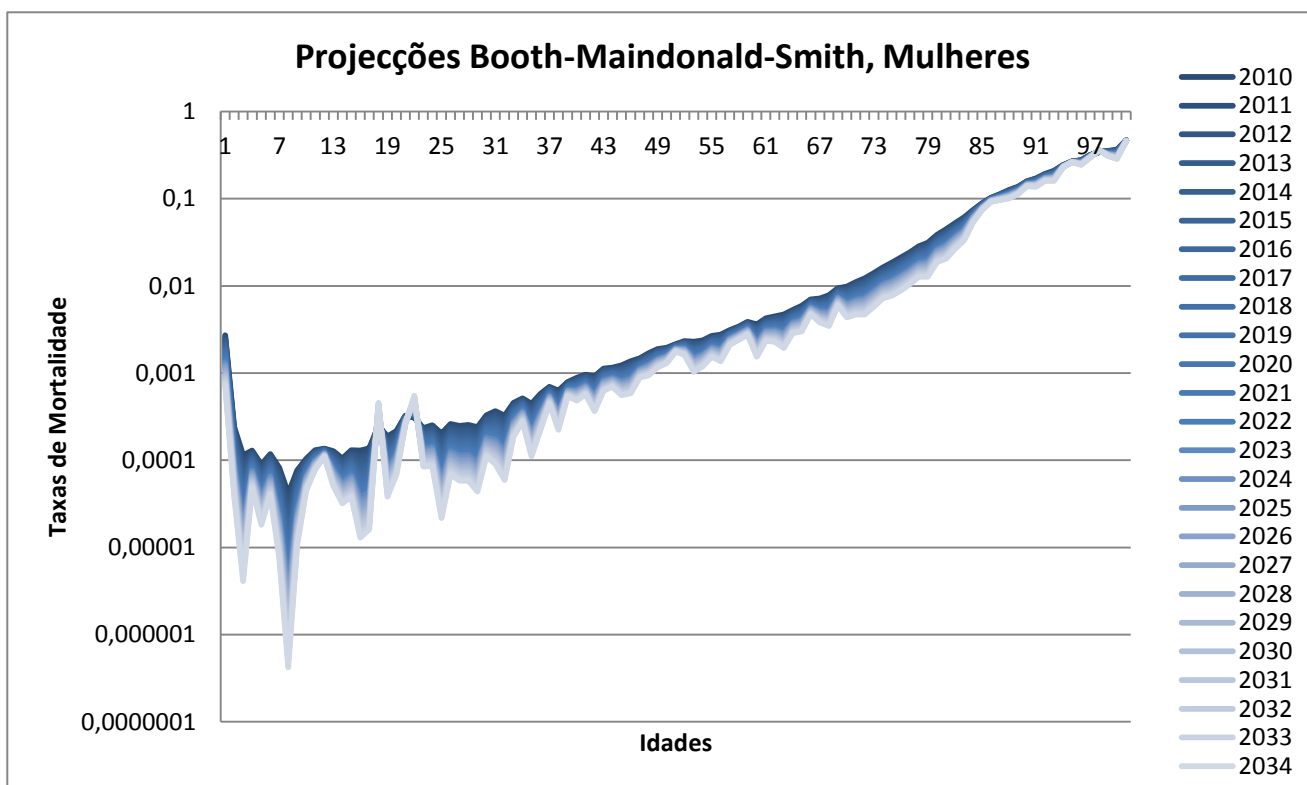
Anexo 57 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 1990 a 2009 para população masculina.



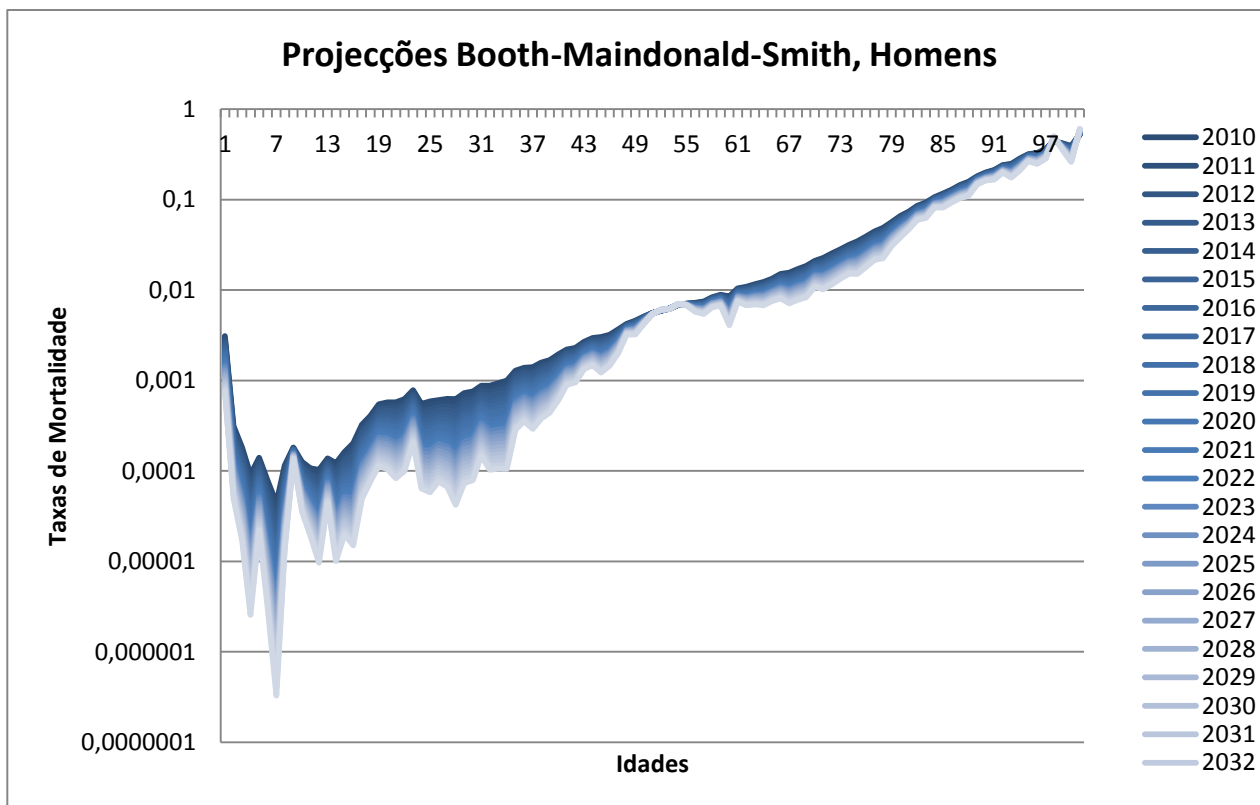
Anexo 58 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 2000 a 2009 para população feminina.



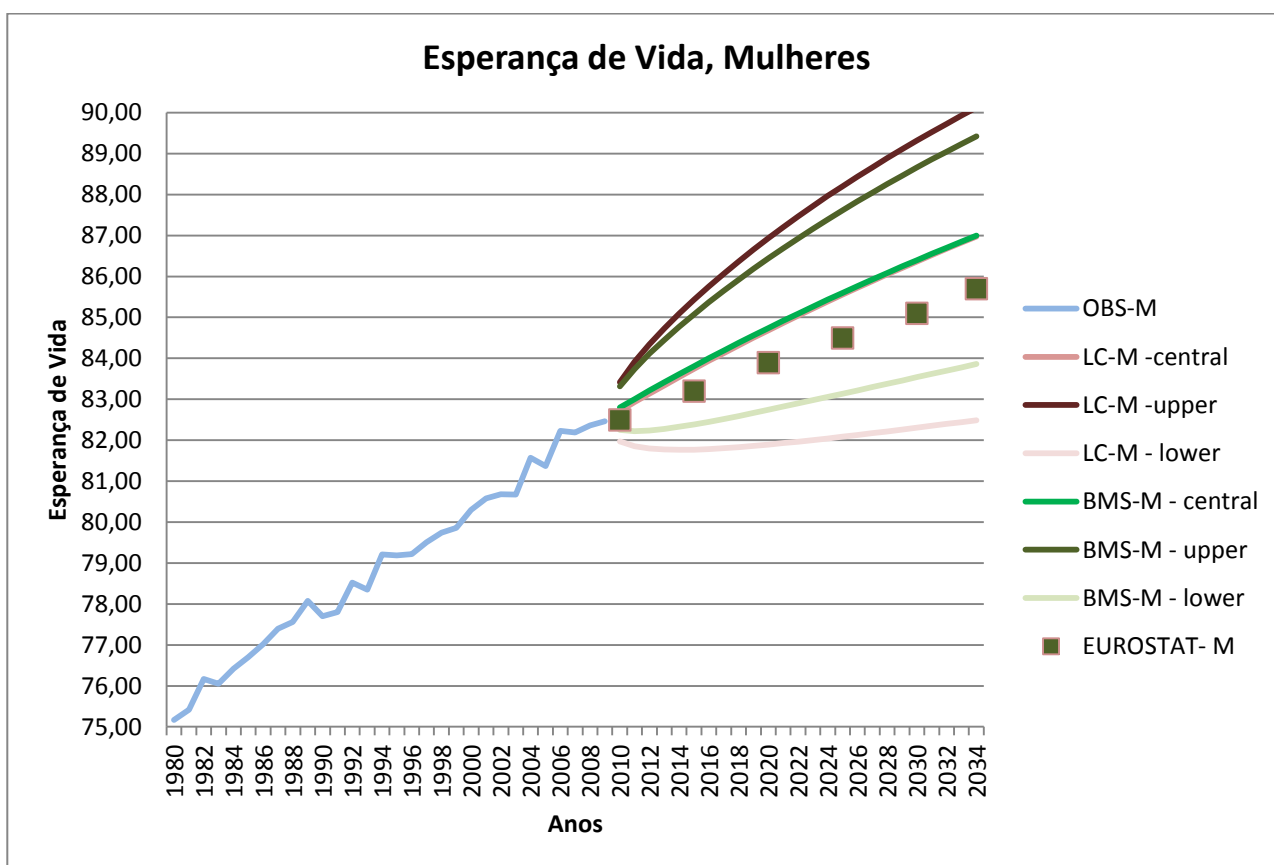
Anexo 59 – Taxas de Mortalidade Projectadas pelo MLC, com ajustamento de 2000 a 2009 para população masculina.



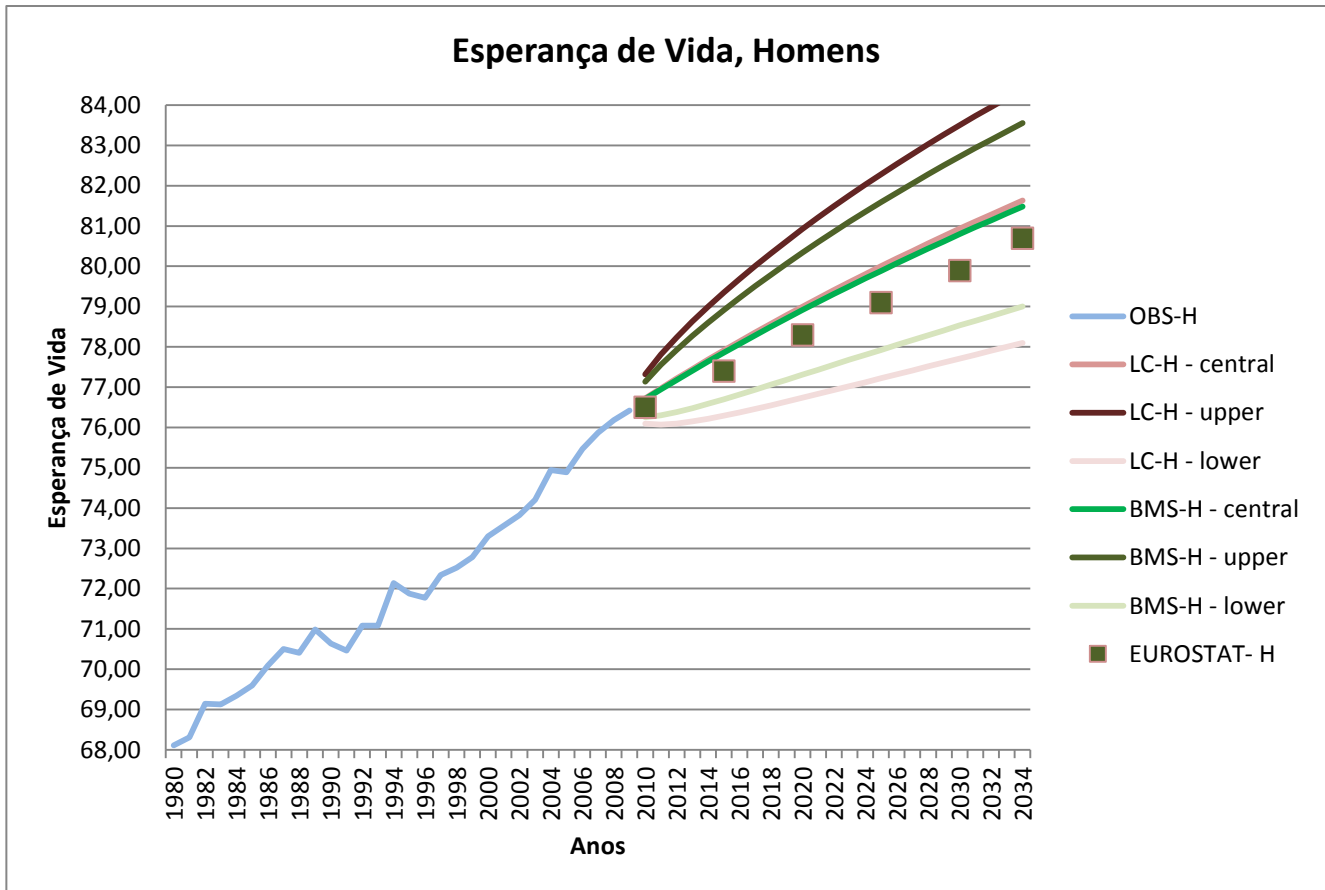
Anexo 60 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 2000 a 2009 para população feminina.



Anexo 61 – Taxas de Mortalidade Projectadas pelo MBMS, com ajustamento de 2000 a 2009 para população masculina.



Anexo 62 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1990 e 2009 para a população feminina, com intervalos de predição.



Anexo 63 – Esperança de vida projectada pelos modelos com período de ajustamento entre 1990 e 2009 para a população masculina, com intervalos de predição.

