

MESTRADO

MÉTODOS QUANTITATIVOS PARA A DECISÃO ECONÓMICA E EMPRESARIAL

TRABALHO FINAL DE MESTRADO

RELATÓRIO DE ESTÁGIO

PROJEÇÕES DE MIGRAÇÕES EM PORTUGAL E DETERMINANTES DA SUA EVOLUÇÃO

MANUEL MARIA BARBAS GAIO CARDOTE DA CUNHA

NOVEMBRO - 2020

MESTRADO
MÉTODOS QUANTITATIVOS PARA A DECISÃO ECONÓMICA
E EMPRESARIAL

TRABALHO FINAL DE MESTRADO
RELATÓRIO DE ESTÁGIO

PROJEÇÕES DE MIGRAÇÕES EM PORTUGAL E DETERMINANTES DA
SUA EVOLUÇÃO

MANUEL MARIA BARBAS GAIO CARDOTE DA CUNHA

ORIENTAÇÃO:
PROFESSORA DOUTORA EDVIGES COELHO
PROFESSOR DOUTOR PAULO PARENTE

NOVEMBRO - 2020

Agradecimentos

Gostaria de agradecer ao ISEG pela oportunidade de realizar este estudo e, em particular, aos meus orientadores, a Professora Doutora Edviges Coelho e o Professor Doutor Paulo Parente, pelo acompanhamento e tempo dedicado ao longo da realização deste trabalho.

Ao INE e ao departamento de Estatísticas Demográficas e Sociais pela oportunidade de poder realizar o meu trabalho final de mestrado nesta instituição e também agradecer à Dra. Cláudia Pina por todo o apoio prestado na minha breve passagem pelo INE.

Por fim, quero agradecer à minha família pelo suporte e apoio ao longo do meu percurso académico e aos meus amigos e colegas de mestrado pelo seu apoio nesta aventura.

Resumo

Atualmente, os fluxos migratórios são o fator que determina a dinâmica da população em Portugal, sendo, contudo, o processo demográfico mais volátil. A previsão de fluxos migratórios é uma tarefa multidimensional, caracterizada por elevados níveis de erro, não existindo uma teoria robusta para diferentes fluxos. O presente relatório de estágio tem como objetivo obter projeções para Portugal, entre 2018 e 2038. Para tal, foi aplicado o modelo proposto por Rogers e Castro (1981), bem como uma regressão linear com o intuito de identificar possíveis fatores explicativos. O cálculo dos parâmetros do modelo permite uma representação dos dados e uma melhor interpretação dos mesmos, sendo a sua projeção feita recorrendo a modelos auto-regressivos integrados de médias móveis.

Os resultados alcançados apresentam conclusões positivas para o panorama atual da estrutura etária portuguesa, verificando-se um saldo migratório positivo em alguns dos cenários definidos. Apesar das limitações verificadas, foi também possível aferir o impacto significativo do desemprego no saldo migratório.

Palavras-chave: ARIMA, migração, previsão

Abstract

Nowadays, migratory flows are the factor that determines the population dynamics in Portugal, being, however, the most volatile demographic process. The prediction of migratory flows is a multidimensional task, characterized by high levels of error, and there is no robust theory for different flows. This internship report aims to obtain projections for Portugal, between 2018 and 2038. To this end, the model proposed by Rogers and Castro (1981) was applied, as well as a linear regression in order to identify possible explanatory factors. The calculation of the model parameters allows a representation of the data and a better interpretation of the same, being projected using autoregressive integrated moving average models.

The results achieved show positive conclusions for the current panorama of the Portuguese age structure, with a positive net migration in some of the defined scenarios. Despite the verified limitations, it was also possible to assess the significant impact of unemployment on the net migration.

Keywords: ARIMA, migration, forecast

Índice

Agradecimentos	i
Resumo	ii
Abstract.....	iii
Índice	iv
Índice Figuras	v
Índice Tabelas.....	vi
1. INTRODUÇÃO.....	1
2. REVISÃO DA LITERATURA	3
3. DADOS E METODOLOGIA	8
3.1 Fonte dos dados	8
3.2 Metodologia.....	10
3.2.1 Modelo Rogers-Castro.....	10
3.2.2 Metodologia para aplicação do modelo Rogers-Castro.....	13
3.2.3 ARIMA.....	15
3.2.4 Regressão.....	16
4. RESULTADOS	18
4.1 Projeções em Portugal	19
4.2 Regressão.....	26
5. CONCLUSÃO.....	29
6. REFERÊNCIAS BIBLIOGRÁFICAS	31
7. ANEXOS.....	35

Índice Figuras

Figura 1: Perfil migratório do modelo de 11 parâmetros	11
Figura 2: Perfis migratórios dos quatro modelos R-C	12
Figura 3: Modelo de 13 parâmetros para as Imigrantes Mulheres, em 2013	20
Figura 4: Modelo de 13 parâmetros para as Emigrantes Mulher, em 2018.....	22
Figura 5: Modelo de 13 parâmetros para as Imigrantes Homem, em 2018.....	22
Figura 6: Evolução do Saldo Migratório nos cenários definidos	25
Figura 7: Comparação da estrutura populacional dos imigrantes, em 2018 e 2038.....	25
Figura 8: Evolução do Saldo Migratório entre 1997 e 2018	27
Figura 9: Evolução do saldo migratório e do spread	35
Figura 10: Emigrantes Homem 2007.....	35
Figura 11: Emigrantes Mulher 2012.....	36
Figura 12: Imigrantes Homem 2012.....	36
Figura 13: Emigrantes Homem 2018.....	37
Figura 14: Imigrantes Mulher 2018.....	37
Figura 15: Previsão Emigrantes Homem.....	38
Figura 16: Previsão Emigrantes Mulher	38
Figura 17: Previsão Imigrantes Homem.....	39
Figura 18: Previsão Imigrantes Mulher	39
Figura 19: Comparação da estrutura populacional dos emigrantes, em 2018 e 2038	40

Índice Tabelas

Tabela I: Teste ADF	16
Tabela II: Matriz de Correlação.....	17
Tabela III: Valores dos parâmetros do modelo de 13 parâmetros com maior REQM e respetivo ano.....	21
Tabela IV: Valores dos parâmetros do modelo de 13 parâmetros para o ano de 2018 ..	23
Tabela V: Cenários em 2038 em comparação com 2018	24
Tabela VI: Estimativas dos coeficientes.....	28
Tabela VII: Conjunto de dados.....	40

1. INTRODUÇÃO

Os movimentos migratórios, conjuntamente com os movimentos naturais de nascimento e óbito, são cruciais para o estudo da dinâmica de crescimento de uma população. Portugal tem verificado um envelhecimento demográfico ao longo dos anos, tendo sido, em 2018, o terceiro país da União Europeia com a população mais envelhecida, segundo os dados do índice de envelhecimento do Pordata (2020). Sendo um país de imigração recente, tendo até ao final do século XX sido um país marcado pela emigração, os seus efeitos do seu envelhecimento foram suavizados, uma vez que a população de nacionalidade estrangeira em Portugal é tendencialmente mais jovem (Oliveira e Gomes, 2014).

Comparativamente à fecundidade e à mortalidade, a previsão de migrações é bastante complexa e está dependente de vários fatores, como económicos, políticos e sociais. O propósito do estudo e os dados disponíveis são, também, fatores a ter em conta. O facto de não existir um consenso internacional sobre o método de obtenção dos dados leva a que não haja dados detalhados e que não seja possível uma comparação direta entre países. Acrescida a falta de uma teoria universal, existem razões para procurar solucionar estes problemas. Uma vez solucionados, seria possível aplicar uma mesma teoria a diferentes fluxos e descobrir as suas principais diferenças.

Com o intuito de estimar padrões de migração, sendo estes altamente dependentes entre grupos de idades, Rogers e Castro (1981) desenvolveram um modelo multi-exponencial de taxas de migração por idade, que ficou conhecido como modelo Rogers-Castro (R-C). As taxas de migração tendem a ser mais altas nos jovens adultos e tendem a diminuir com o aumento da idade, podendo alguns padrões verificar um segundo pico no período da reforma. O modelo proposto permite reduzir um grande número de taxas de migração a um pequeno número de parâmetros, permitindo a sua comparação ao longo do tempo e facilitando a sua previsão, recorrendo a modelos de séries temporais.

O presente relatório de estágio tem como objetivo obter projeções dos fluxos migratórios em Portugal por idades e identificar possíveis fatores explicativos. Será aplicado o modelo R-C de modo a obter um conjunto de parâmetros e, à semelhança das projeções realizadas pelo INE, será feita uma previsão com recurso a modelos de séries

temporais, que servirão de limites para os três cenários definidos. Também será aplicada uma regressão para identificar possíveis fatores explicativos, uma vez que, no momento de decisão, as escolhas não se limitam apenas a um país, mas também às expectativas de condições económicas e de outros fatores relevantes no presente e para o futuro.

O estágio realizou-se no Instituto Nacional de Estatística (INE), no Departamento de Estatísticas Demográficas e Sociais, entre 3 de fevereiro e 29 de maio de 2020. O INE, como autoridade estatística central, independente e credível, tem por missão produzir e divulgar, de forma imparcial, informação estatística oficial de qualidade, promovendo a coordenação, a análise, a inovação e a divulgação da atividade estatística nacional, garantindo o armazenamento integrado de dados. Desenvolve processos estatísticos metodologicamente avançados, assegurando a confidencialidade dos cidadãos e entidades, devolvendo à sociedade estatísticas de valor para um melhor conhecimento, investigação e tomada de decisão.

O relatório encontra-se dividido em cinco capítulos. No segundo capítulo, é feita uma breve revisão de literatura; no terceiro, encontram-se descritas as fontes dos dados e a metodologia selecionada; no quarto, apresentam-se os resultados obtidos; o quinto e último capítulo contém as conclusões.

2. REVISÃO DA LITERATURA

Segundo o Glossário sobre Migrações da Organização Internacional para as Migrações (2019), o conceito de emigrante/imigrante pode ser definido da seguinte forma: “da perspectiva da região de partida/chegada, uma pessoa que tenha deixado o seu país ou a sua residência habitual para outro país, sendo o país de destino o seu novo país de residência habitual”. Esta definição foi recentemente adaptada da definição de migrante de longo prazo das Nações Unidas (1998), deixando de fazer referência ao período mínimo de doze meses de permanência. Esta permite assim a emigração/imigração de curto-prazo, definida por um período de três a doze meses.

A Europa encontra-se, maioritariamente, dividida no que diz respeito ao método usado para obtenção de dados, havendo países que mantêm registos populacionais regulares e países que realizam os tradicionais censos populacionais. Os primeiros, como é o caso dos países nórdicos e da Bélgica, utilizam inquéritos com o objetivo de verificar e atualizar os seus registos, enquanto os segundos, onde se inclui Portugal, recorrem a dados sobre nascimentos, mortalidade e migração para atualizar os seus dados. O número de países na Europa a recorrer a registos populacionais regulares está a aumentar, contrariamente ao método dos censos tradicionais, que regista uma diminuição. Os países que investem em registos populacionais frequentes têm dados mais precisos relativos à imigração, enquanto o número de migrações no período compreendido entre dois censos consecutivos, ao ser distribuído de forma proporcional pelos anos intermédios, não permite uma maior precisão destes dados. Esta é uma desvantagem do método por censos populacionais. Relativamente à emigração, não há evidência sobre o método mais preciso. Estas diferenças, juntamente com a escassez de dados em alguns países, dificulta a obtenção de uma base de dados internacional sólida (Jennissen, 2004).

O cálculo de fluxos migratórios pode, de forma útil, recorrer a conceitos utilizados na análise da fecundidade e da mortalidade, que devem ser modificados, quando necessário, tendo em conta as movimentações. A fecundidade e a mortalidade, por apresentarem taxas específicas por idade altamente persistentes, permitiram aos demógrafos representá-las através de expressões matemáticas (Rogers e Castro, 1981). No entanto, sendo os fluxos migratórios o processo demográfico mais imprevisível, não existe uma abordagem uniformizada para os diferentes fluxos (Disney et al., 2015). Métodos determinísticos

para previsão da população não oferecem um sinal adequado da incerteza associada, pelo que as previsões devem recorrer a métodos probabilísticos, uma vez que as suas vantagens superam as limitações dos primeiros e permitem que a sua precisão seja avaliada (Booth, 2006). Os modelos probabilísticos de previsão de população recorrem ao método das componentes por coortes (Keilman et al., 2002).

Entre instituições nacionais de estatística e literatura académica, não existe consenso acerca do modelo de previsão estocástico que produz os melhores resultados. A generalidade das projeções realizadas por instituições estatísticas e organizações internacionais permanece numa abordagem determinística, baseadas em cenários de julgamento (Bijak et al., 2019). Em Portugal, segundo o seu documento metodológico (INE, 2020a), o Instituto Nacional de Estatística (INE) adota modelos funcionais coerentes de séries temporais. Este método foi proposto por Hyndman et al. (2013) e tem como objetivo a análise de componentes principais para dados funcionais, com o intuito da redução da dimensionalidade em análise estatística. Com este método um grande número de variáveis é reduzido a um conjunto de variáveis ortogonais (componentes) e assegura que as previsões para subpopulações que estão relacionadas, mantenham as suas eventuais relações estruturais baseadas na observação histórica. Na última publicação sobre Projeções de População Residente (INE, 2020b), foram definidos quatro cenários de projeção da população, assumindo quatro hipóteses para as migrações: hipótese pessimista, hipótese central, hipótese otimista e hipótese sem migrações.

Rogers e Castro (1981) afirmam que o padrão migratório não é idêntico para ambos os sexos, apresentando os homens maior mobilidade que as mulheres. Contudo, a variabilidade das migrações por sexo é muito menos pronunciada que a variabilidade etária e menos uniforme no tempo e no espaço. O modelo proposto por estes autores, o modelo Rogers-Castro (R-C), permite a modelação das taxas específicas de migração por idade, com a distinção de sexo, de modo a obter um conjunto de parâmetros que permitem uma melhor interpretação dos padrões migratórios. Trata-se de um modelo multi-exponencial, representado por um polinómio geral com 7, 9, 11 ou 13 parâmetros e baseia-se nos perfis de migração verificados nas várias faixas etárias. Possibilita, assim, não só estimar curvas hipotéticas para estudos demográficos quando se está na presença de dados limitados ou inadequados, como também analisar o comportamento das taxas

por idade entre populações distintas ou para períodos diferentes de uma mesma população (Santos e Barbieri, 2019).

Segundo Wilson (2010), os modelos de padrões etários têm-se tornado um instrumento importante, ao permitirem fazer estimativas quando os dados não são totalmente conhecidos ou são suspeitos; a sua graduação quando os padrões etários observados contêm ruído e a redução do número de dados, onde uma grande observação de dados é convertida num número pequeno de parâmetros. Esta redução dos dados permite uma análise comparativa, onde um pequeno número de parâmetros facilita a comparação de padrões etários ao longo do tempo e espaço e, por último, obter projeções de séries temporais dos parâmetros, onde a sua extrapolação permite projetar os perfis migratórios para o futuro. Por outro lado, Bell (2005) constata que, contrariamente aos parâmetros obtidos em modelos de fecundidade e mortalidade, os parâmetros obtidos não descrevem, de forma real, os padrões migratórios das diferentes curvas, pois podem ser semelhantes para os diferentes fluxos. Bernard e Bell (2015) e Ruiz-Santacruz (2019) identificam, também, a escolha arbitrária dos parâmetros iniciais como uma das debilidades da estimação de modelos por meio da otimização, visto que uma variação dos parâmetros iniciais é comum conduzir a diferentes parâmetros ótimos. Para solucionar este problema, é utilizada a mesma semente no início das simulações (o que garante que o resultado é o mesmo independentemente do número de vezes que se faça a otimização) e, posteriormente, é feito um teste de qualidade do ajustamento da curva estimada.

Rogers e Castro (1981) elegeram o estimador do mínimo Qui-quadrado (*minimum chi-squared estimator*) para dar mais peso às faixas etárias com menores taxas de migração. O algoritmo utilizado para minimizar a função objetivo é uma versão modificada do algoritmo de Levenberg-Marquardt. Outros estudos utilizam, também, a estatística do Qui-quadrado como critério de ajustamento, obtendo informação acerca da sua qualidade através da diferença entre os valores observados e ajustados, sendo também definido um critério de seleção da estimativa, o valor de tolerância, baseado nesta diferença. Uma vez que o objetivo primordial do presente relatório é a análise dos parâmetros obtidos e não a avaliação do método de otimização, é utilizado o método dos mínimos quadrados como primeira aproximação e a diferença entre os erros quadráticos médios estimados entre o melhor modelo guardado e o novo modelo estimado como critério de seleção, tal como em Ruiz-Santacruz (2019).

Keilman et al. (2002) reconhece que a generalidade dos métodos de previsão de séries temporais parte da suposição de que é possível inferir um conjunto de informações sobre o comportamento de uma série temporal a partir de observações passadas, sendo este padrão recorrente no tempo. Estes métodos foram desenvolvidos para horizontes curtos e, quando aplicados à previsão de longo prazo, a previsão torna-se constante, gerando elevados níveis de erro. Keilman et al. (2002) identifica a extrapolação de séries temporais como um método de previsão tanto do fluxo total como dos parâmetros do modelo R-C. Assim, os parâmetros extrapolados permitem obter, através dos fluxos previstos, uma previsão do número de emigrantes e imigrantes. Sendo a migração internacional um fenómeno continuamente dinâmico, a sua previsão deve permanecer num horizonte curto, podendo no longo prazo, conjuntamente com os seus intervalos de confiança, tornarem-se irrealistas (Bijak, 2006).

Um método bastante popular para análise de séries temporais é o modelo autoregressivo integrado de médias móveis (autoregressive integrated moving average; ARIMA) (Keilman et al., 2002; Bijak, 2006; Disney et al., 2015). Estes modelos podem ter uma “memória” curta ou longa, dependendo dos seus parâmetros e podem apresentar diferentes propriedades em relação à sua estacionariedade. Uma crítica teórica apontada a estes modelos é que a sua previsão é baseada apenas nos dados passados de um único processo estocástico e pode conduzir a previsões irrealistas (Bijak, 2006). Como alternativa, os modelos econométricos permitem, não só prever fluxos migratórios, como também comprovar teorias económicas com base em dados empíricos e informação covariável (Bijak et al., 2019).

A migração internacional pode variar no tipo, no tempo e entre diferentes países (Jennissen, 2004; Simpson, 2017). A migração com propósito laboral é diferente da migração por motivos familiares, sendo os primeiros geralmente mais novos e com diferentes expectativas de benefícios resultantes da decisão. A semelhança linguística, por exemplo, também é um fator a ter em conta no momento de decisão. Contudo, este fator tem perdido o seu peso, devido à predominância do inglês a nível mundial.

Bruder (2003) afirma que as previsões baseadas apenas no PIB per capita e na taxa de desemprego tendem a sobrestimar os fluxos futuros, devendo estas incluir variáveis socioeconómicas. Deste modo, Castelli (2018) enumera diversos fatores económicos e

não económicos que levam à decisão final de migrar. Estes fatores são divididos em três grupos: macro (independentes das características individuais), micro (características individuais) e meso (elementos relacionados com o indivíduo, mas que não estão sob o seu controlo). Os elementos macro podem ser divididos em políticos, demográficos, económicos, sociais e ambientais. Os elementos micro, por sua vez, estão relacionados, entre outros, com a idade, sexo, etnia, língua e nível escolaridade. Já os elementos meso relacionam-se com os custos de deslocação, tecnologia e o quadro político/legal do país.

Os determinantes económicos estão maioritariamente associados à migração de mão-de-obra, podendo, no entanto, outros tipos de migração ser parcialmente afetados por fatores económicos. Impostos e transferências sociais (como o subsídio de desemprego) surgem, com maior destaque, como fatores de atratividade e não de repulsividade. Em países com impostos progressivos, os trabalhadores de baixas qualificações tendem a ser pouco afetados pela progressividade, sendo, por isso, um fator de atratividade. Países com maiores níveis de impostos tendem a ser menos atrativos.

Segundo Simpson (2017), um maior peso de despesas sociais em percentagem do PIB não atrai necessariamente mais migrantes. A volatilidade de rendimentos, associada a ciclos económicos, incentiva a emigração no país de origem. Um nível elevado de desemprego num país é causador de emigração, enquanto um nível baixo é um fator de atratividade. Para a autora, as diferenças salariais e o nível de desemprego são os principais fatores no momento de decisão.

Jennissen (2004) analisou o impacto de dois tipos de indicadores macroeconómicos: o primeiro tipo relacionado com os indicadores do mercado de trabalho, tais como a taxa de desemprego e a quantidade de capital humano no país; o segundo referente a indicadores de produtividade, como o PIB ou PIB per capita. Esta hipótese é baseada na teoria neoclássica da migração, onde os fluxos ocorrem dos países mais pobres para países mais ricos, na procura de salários mais elevados e melhores ofertas de emprego. Nas várias regressões analisadas, são também adicionadas várias variáveis *dummy* de eventos históricos dos diversos países em análise. Este autor conclui que o PIB per capita e o desemprego são fatores determinantes para o perfil migratório de um país, juntamente com decisões políticas inesperadas ou crises económicas.

Alessandria et al. (2020) analisou na Europa diversos fatores, tendo chegado à conclusão de que o saldo migratório está positivamente correlacionado com a atividade económica, observando-se uma correlação negativa com spreads e desemprego, tal como se verificou durante a crise das dívidas europeias.

De forma a tornarem as suas políticas mais efetivas, os decisores políticos devem entender o impacto dos diversos fatores. A título de exemplo, a existência de estimativas da variação do saldo migratório originadas por flutuações no PIB poderia ser útil à definição de políticas relativas à emigração e imigração. Citando Peixoto (2009), “A demografia pode condicionar as escolhas, mas são os mecanismos sociais que irão desenhar os caminhos futuros”.

3. DADOS E METODOLOGIA

Este capítulo descreve a metodologia selecionada e os procedimentos seguidos no cálculo das projeções da migração portuguesa e de fatores explicativos da mesma. Será descrita a abordagem seguida para a pesquisa empírica, desde as fontes utilizadas, a recolha e o tratamento de dados, a metodologia utilizada e o conjunto de variáveis consideradas adequadas ao caso em estudo. Para este estudo recorreu-se aos *softwares* EViews e R.

3.1 Fonte dos dados

O conceito de emigrante/imigrante escolhido para o presente relatório corresponde ao conceito utilizado pelo INE e ao conceito de migrante de longo-prazo definido no Glossário sobre Migrações da Organização Internacional para as Migrações (2019), referido no capítulo anterior. A periodicidade dos dados é anual.

Os dados referentes à Emigração são obtidos através de operações estatísticas do INE, nomeadamente, as Estimativas da população residente em Portugal e o Inquérito aos Movimentos Migratórios de Saída. As estimativas da população residente são utilizadas para obter os ponderadores para a amostra obtida pelo inquérito. Se o parâmetro a estimar for um total, a expressão do estimador é dada por $\hat{Y} = \sum_{k \in S} w_k y_k$. Se o parâmetro a

estimar for um quociente, tem-se $\hat{R} = \frac{\hat{Y}}{\hat{Z}} = \frac{\sum_{k \in S} w_k y_k}{\sum_{k \in S} w_k z_k}$, onde w_k é o ponderador final associado à unidade k da amostra S , y_k é o valor da característica Y para a unidade k da amostra e z_k é o valor da característica Z para a unidade k da amostra.

O Inquérito aos Movimentos Migratórios de Saída é um inquérito eletrónico por amostragem e funciona como anexo ao Inquérito ao Emprego. Realiza-se no primeiro trimestre de cada ano e tem como alvo de entrevista os indivíduos que fazem ou fizeram parte do agregado familiar a residir no alojamento selecionado e que deixaram o país com intenção de residir no estrangeiro, por um período igual ou superior a três meses, durante o ano anterior ao da recolha da informação, independentemente da sua nacionalidade ou país de nascimento.

Relativamente à Imigração, os seus dados são obtidos por operações estatísticas do INE, nomeadamente, estimativas anuais de população residente e Inquérito ao Emprego; e por procedimentos administrativos, relativos ao Serviço de Estrangeiros e Fronteiras (número de indivíduos que solicitaram autorização de residência em Portugal, por sexo, idade, país de nacionalidade país de nascimento e país de origem) e Ministério dos Negócios Estrangeiros/Direção Geral dos Assuntos Consulares e das Comunidades Portuguesas.

O alisamento tanto para as séries dos emigrantes, como para as dos imigrantes é feito através de médias móveis de vários índices, em função das idades e/ou da nacionalidade (no caso dos imigrantes).

A existência de casos pontuais de migração em idades mais avançadas levou à escolha, para o processo de otimização, do limite de 70 anos no caso da emigração e de 80 anos no caso da imigração. Esta escolha de idades deveu-se também à possibilidade de o modelo R-C representar uma curva pós-reforma.

No que respeita à regressão, a descrição das variáveis utilizadas e a sua respetiva fonte podem ser encontradas na Tabela VII. Também nesta tabela é possível encontrar os sinais esperados para os seus coeficientes, de acordo com a literatura.

3.2 Metodologia

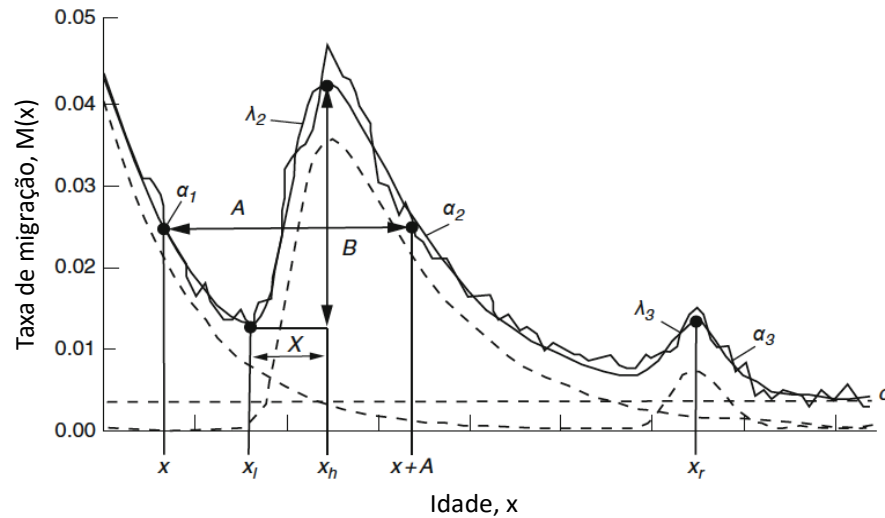
3.2.1 Modelo Rogers-Castro

A constância de comportamentos característicos em cada padrão etário, observados em diversas populações, levou à apresentação do modelo proposto por Rogers e Castro (1981). A intensidade de migração tende a atingir o pico na idade adulta, à medida em que são procuradas oportunidades de emprego e educação, sendo o pico masculino mais tardio que o feminino (Rogers et al., 2010). Por outro lado, a migração torna-se cada vez menos intensa com o avanço da idade, devido a fatores que diminuem a mobilidade, como a estabilidade de carreira e a formação de família. Alguns perfis de idade apresentam um pico secundário, aos 50 e 60 anos. As intensidades migratórias das crianças são semelhantes às dos seus pais, mais frequentes nos primeiros anos de vida e menos frequentes na adolescência.

O modelo R-C permite a representação destes perfis migratórios através da calibração dos parâmetros do modelo. Trata-se de um modelo multi-exponencial que pode ser descrito por um termo constante, ao qual se soma uma curva exponencial negativa no período de infância, uma curva unimodal inclinada para a esquerda no período da força de trabalho, uma curva quase em forma de sino no período da reforma e uma curva exponencial positiva no período pós-reforma. O modelo completo de 13 parâmetros, representativo de todas as curvas descritas, é então composto pelas 5 componentes referidas: curva de infância (N_1), curva da força de trabalho (N_2), curva de reforma (N_3), curva de pós-reforma (N_4) e constante. Refira-se, contudo, que nem todas as componentes são necessárias, podendo o modelo ser composto por 7, 9 ou 11 parâmetros. Efetivamente, de um modo geral, quase todos os perfis de migração requerem a componente de infância, força de trabalho e constante, mas apenas alguns incluem as componentes de reforma e pós-reforma.

A figura 1 apresenta o perfil migratório do modelo de 11 parâmetros e a ilustração dos respetivos parâmetros.

Figura 1: Perfil migratório do modelo de 11 parâmetros



Fonte: Rogers et al. (2010)

Para caracterizar os cronogramas de migração em função de várias medidas cruciais, pode ser avaliado como identificado na figura acima:

- x_l = pico baixo
- x_h = pico alto
- x_r = pico reforma
- X = mudança da força de trabalho
- A = mudanças parentais
- B = salto

O modelo de 13 parâmetros, que acresce à figura 1 a curva ascendente no período pós-reforma, pode ser representado da seguinte forma:

$$\hat{m}(x) = N_1(x) + N_2(x) + N_3(x) + N_4(x) + c \Leftrightarrow$$

$$\Leftrightarrow \hat{m}(x) = a_1 e^{-\alpha_1 x} + a_2 e^{-\alpha_2(x-\mu_2) - e^{-\lambda_2(x-\mu_2)}} + a_3 e^{-\alpha_3(x-\mu_3) - e^{-\lambda_3(x-\mu_3)}} \quad (1)$$

$$+ a_4 e^{\lambda_4 x} + c$$

Onde:

$\hat{m}(x)$ = estimativa da taxa de migração
 x = idade

a_1 = altura da curva da infância
 α_1 = taxa de descida da curva de infância

a_2 = altura da curva da força de trabalho
 λ_2 = taxa de subida da curva da força de trabalho

$\alpha_2 =$ taxa de descida da curva da força de trabalho
 $\mu_2 =$ posição da curva da força de trabalho no eixo etário

$a_3 =$ altura da curva de reforma
 $\lambda_3 =$ taxa de subida da curva de reforma
 $\alpha_3 =$ taxa de descida da curva de reforma
 $\mu_3 =$ posição da curva de reforma no eixo etário

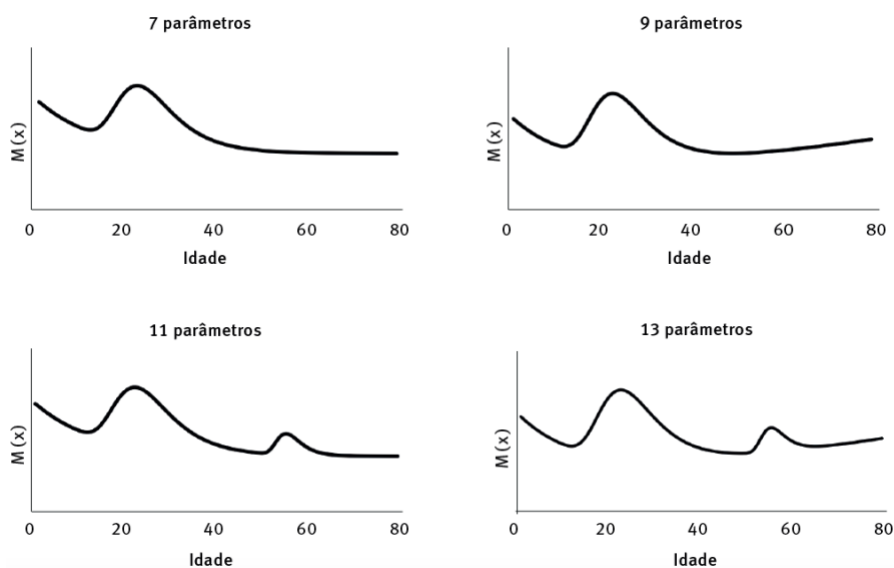
$a_4 =$ altura da curva pós – reforma
 $\lambda_4 =$ taxa de subida da curva pós – reforma

$c =$ constante

O modelo pode ser definido em função do seu padrão e do seu nível. No modelo completo (13 parâmetros), o padrão é definido pelos seguintes parâmetros: $\alpha_1, \alpha_2, \lambda_2, \mu_2, \alpha_3, \lambda_3, \mu_3$ e λ_4 . O nível é determinado pelos restantes parâmetros: a_1, a_2, a_3, a_4 e c .

O modelo de sete parâmetros, correspondente à função padrão, é composto pelas componentes da infância, da força de trabalho e pela componente constante; o modelo com nove parâmetros adiciona a componente ascendente no período pós-reforma. O modelo de onze parâmetros adiciona ao modelo de sete parâmetros o pico no momento da reforma e, no modelo de 13 parâmetros (a função completa), é adicionado ao modelo de sete parâmetros o pico no período da reforma e a componente ascendente no período pós-reforma. Os perfis migratórios podem assim ser descritos pelos quatro modelos, representados na figura 2.

Figura 2: Perfis migratórios dos quatro modelos R-C



Fonte: Santos e Barbieri (2019)

3.2.2 Metodologia para aplicação do modelo Rogers-Castro

Segundo Bernard e Bell (2015), o ajustamento de modelos de padrões migratórios a intensidades de migração por idades observadas requer a utilização de programas de ajustamento não-lineares, devido à sensibilidade da curva ajustada aos valores iniciais dos parâmetros.

Para a aplicação do modelo R-C, foi utilizado o *package* *migraR*, desenvolvido para o *software* R por Ruiz-Santacruz e Garcés (2018). O algoritmo proposto utiliza como *input* a taxa de migração padronizada (soma igual a um) ou a percentagem do número total de migrantes por idades e pretende obter o melhor modelo ajustado, tendo em conta um número máximo de iterações e um valor de critério de paragem escolhidos. Por *default*, mas possível de alteração, são adotadas 100 iterações e, como critério de paragem, um valor de 0,00001 para a diferença entre o Erro Quadrático Médio (EQM) do melhor modelo guardado e do novo modelo estimado.

$$EQM_g = \frac{1}{n} \sum_{i=1}^n (O_i - F_i)^2 \quad (2)$$

Onde g é o modelo guardado, O é o vetor dos valores observados da variável a ser ajustada e F é o vetor dos valores ajustados.

Caso o critério de paragem não se verifique, o algoritmo é executado até atingir o número máximo de iterações definido e será escolhido o modelo guardado, correspondente ao modelo com menor EQM. Ruiz-Santacruz (2019) considera uma diferença (e) de 0,001 um critério de paragem bastante exigente, tendo no seu trabalho necessitado de 40 000 simulações de modo a obter um modelo melhor.

As funções base que levarão ao cálculo do EQM para os modelos de 7, 9, 11 e 13 parâmetros são, respetivamente:

- $\hat{m}(x) = \log (a_1 e^{-\alpha_1 x} + a_2 e^{-\alpha_2(x-\mu_2)} - e^{-\lambda_2(x-\mu_2)} + c) \quad (3)$

$$\bullet \quad \hat{m}(x) = \log (a_1 e^{-\alpha_1 x} + a_2 e^{-\alpha_2(x-\mu_2)-e^{-\lambda_2(x-\mu_2)}} + a_3 e^{\alpha_3 x} + c) \quad (4)$$

$$\bullet \quad \hat{m}(x) = \log (a_1 e^{-\alpha_1 x} + a_2 e^{-\alpha_2(x-\mu_2)-e^{-\lambda_2(x-\mu_2)}} + a_3 e^{-\alpha_3(x-\mu_3)-e^{-\lambda_3(x-\mu_3)}} + c) \quad (5)$$

$$\bullet \quad \hat{m}(x) = \log (a_1 e^{-\alpha_1 x} + a_2 e^{-\alpha_2(x-\mu_2)-e^{-\lambda_2(x-\mu_2)}} + a_3 e^{-\alpha_3(x-\mu_3)-e^{-\lambda_3(x-\mu_3)}} + a_4 e^{\lambda_4 x} + c) \quad (6)$$

Os valores iniciais de cada parâmetro são obtidos através de distribuições uniforme no intervalo (0,100) para o μ_i e no intervalo (0,1) para os restantes. São gerados grupos de n (número máximo de iterações) valores para cada parâmetro, sempre através de uma nova amostra, tornando-os independentes e identicamente distribuídos. Este método permite, assim, obter o melhor modelo de n iterações, através das n curvas geradas.

A otimização, no presente trabalho, será feita através do método dos mínimos quadrados, aplicado ao logaritmo da percentagem do número total, e sem restrições, isto é, os parâmetros estimados não são limitados. Serão utilizadas 5000 simulações e como critério de paragem um valor de 0,00001.

Serão criados três cenários: pessimista, central e otimista, à semelhança das projeções realizadas pelo INE. Para tal, será utilizada uma previsão para o fluxo total de emigrantes e imigrantes, por género, com recurso a modelos ARIMA. O valor médio desta previsão corresponderá ao cenário central e os valores inferior e superior do intervalo de confiança a 80% serão utilizados para o cenário pessimista e otimista, respetivamente. Para a obtenção do perfil migratório em cada ano das projeções, serão utilizados os parâmetros obtidos para o ano de 2018, do melhor modelo R-C identificado.

3.2.3 ARIMA

A previsão do número total de emigrantes e imigrantes, segundo o sexo, será obtida com recurso ao modelo de previsão de séries temporais, ARIMA. A seleção do melhor modelo será feita por um procedimento automatizado.

Uma propriedade importante de uma série temporal é a estacionariedade. Uma série temporal estocástica é considerada estacionária se tem média e variância constantes no tempo e se a covariância entre valores desfasados da série depende apenas da defasagem temporal (Caiado, 2016).

Os modelos autoregressivos e médias móveis, ou modelos ARMA, foram introduzidos por Box e Jenkins (1970), com o intuito de modelar séries temporais. Com a introdução das primeiras diferenças, quando há evidência de não estacionariedade da série, obtém-se uma generalização destes modelos, os modelos autoregressivos e médias móveis integrados (ARIMA). A construção geral de um modelo ARIMA (p,d,q), pode ser representada pela seguinte expressão:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Y_t = (1 - \theta_1 B - \dots - \theta_q B^q) \varepsilon_t \quad (7)$$

Onde:

$(1 - B)^d Y_t$, com $d \geq 1$ é a série estacionária depois de diferenciada d vezes,

o operador B é o operador atraso (operador *backshift* ou operador *lag*),

$\phi_1, \phi_2, \dots, \phi_p$ são os parâmetros autorregressivos e,

$\theta_1, \theta_2, \dots, \theta_p$ os parâmetros médias móveis.

A seleção do modelo será feita automaticamente, segundo o critério Bayesiano de Schwartz (BIC ou SBC), critério alternativo ao critério de Informação de Akaike (AIC). Este critério mede a qualidade de ajuste de um modelo estatístico e também pode ser usado para a seleção de modelos. O seu valor é dado por:

$$BIC = n \ln \hat{\sigma}_\varepsilon^2 + m \ln n \quad (8)$$

Onde $\hat{\sigma}_\varepsilon^2$ é a estimativa da variância dos resíduos, m é o número de parâmetros do modelo e n é o número de observações utilizadas na estimação do modelo. Será selecionado o modelo que apresentar o menor valor BIC.

3.2.4 Regressão

A introdução de uma regressão no presente trabalho tem como principal objetivo analisar os efeitos de fatores económicos e não económicos na evolução do saldo migratório, dado por:

$$\text{Saldo Migratório}_t = \text{Imigrantes}_t - \text{Emigrantes}_t \quad (9)$$

O período de análise começa em 1997 e prolonga-se até 2018, utilizando dados anuais. Para que a inferência estatística não deixe de ser válida e para que se obtenham estimadores consistentes, as variáveis em análise devem ser estacionárias. Uma regressão com variáveis não estacionárias seria uma regressão espúria. Para verificar tal pressuposto recorreu-se ao teste ADF (Dickey-Fuller Aumentado) para averiguar a existência de raízes unitárias, incluindo um termo constante na equação de teste. O teste realizado seleciona automaticamente o número de *lags* da variável dependente a incluir na regressão auxiliar, segundo o critério SIC (Schwarz Info Criterion). De acordo com a seguinte tabela todas as variáveis apresentam raiz unitária ao nível de 5%. Deste modo, a hipótese nula de existência de raiz unitária não é rejeitada e, conseqüentemente, todas as variáveis foram diferenciadas.

Tabela I: Teste ADF

	SM	Pibpercap	Desemp	SalMin	Spread
Nº de <i>lags</i> *	1	0	1	4	2
Estatística	-1,768	-1,430	-2,175	0,336	-1,399
Valor-p	0,385	0,548	0,220	0,973	0,561
Ordem	I(1)	I(1)	I(1)	I(1)	I(1)

* número de *lags* da variável dependente incluídos na regressão auxiliar

De forma a evitar problemas de multicolinearidade quase-perfeita, foi calculada a tabela de correlação com o intuito de auxiliar na compreensão dos coeficientes e interpretações.

Tabela II: Matriz de Correlação

	d(Sm)	d(Pibpercap(-1))	d(Desemp(-1))	d(Salmin(-1))	d(Spread(-1))
d(Sm)	1				
d(Pibpercap(-1))	0,248	1			
d(Desemp(-1))	-0,543	-0,685	1		
d(Salmin(-1))	-0,269	0,063	-0,124	1	
d(Spread(-1))	-0,424	-0,487	0,418	0,307	1

O modelo utilizado para analisar o impacto no saldo migratório é:

$$\Delta SM_t = \beta_0 + \beta_1 \Delta PIB/PopTotal_{t-1} + \beta_2 \Delta Desemp_{t-1} + \beta_3 \Delta SalMin_{t-1} + \beta_4 \Delta Spread_{t-1} + \varepsilon_t \quad (10)$$

Onde a distribuição do vetor de erros (ε), condicional nos regressores (X), é $\varepsilon|X \sim N(0, \sigma^2 I)$.

O modelo foi baseado principalmente nas conclusões de Bruder (2003) e inclui fenómenos analisados em Portugal por Góis (2019).

De acordo com a literatura vigente sobre os determinantes económicos e não económicos, foram compilados os elementos de maior relevo, não tendo sido definidas variáveis de base. A fonte de cada variável encontra-se na Tabela VII (ver Anexos).

Tal como Bruder (2003), todas as variáveis exógenas foram desfasadas um ano, dado que as decisões, baseadas em expectativas e experiências, são formuladas com base no passado.

Ao nível dos fatores macroeconómicos, foram escolhidas as variáveis PIB per capita e número de desempregados. Como visto em Bruder (2003) é esperado um sinal positivo para a variável PIB per capita, isto é, um aumento (diminuição) nesta variável levará a

um aumento (diminuição) do saldo migratório. Em sentido contrário, é esperado um sinal negativo para a variável desemprego, onde um aumento (diminuição), fator originário de emigração, é esperado uma diminuição (aumento), traduzindo-se num aumento da emigração e diminuição da imigração. Em tempos de recessão, como a crise económica e financeira iniciada em 2007/2008, observou-se em Portugal a partir de 2009, um aumento do desemprego, tendo este sido um fator encorajador de emigração e uma barreira à imigração. Consequentemente, observou-se um saldo migratório negativo.

Para determinar o risco internacional, foi considerado o spread da taxa de juro das obrigações a 10 anos entre Portugal e a Alemanha. É esperada uma relação negativa, em que valores mais altos (mais baixos) do spread diminuam (aumentem) o valor da variável dependente, como visto em Alessandria et al. (2020). Taxas de juro mais elevadas representam maiores obrigações do Estado para estabilizar o rácio da dívida, podendo a implementação de certas medidas reduzir as condições económicas. Por isso, a evolução das taxas de juro dos títulos a 10 anos pode-se correlacionar com a evolução das expectativas e, consequentemente, ter impacto na decisão de migrar. A figura 9 (ver Anexos) ilustra a evolução do saldo migratório e do spread.

Espera-se que o coeficiente da variável Salário Mínimo tenha um sinal positivo, onde um aumento (diminuição) originará um aumento (diminuição) do saldo migratório. Como visto em Lange e Gollin (2007), Simpson (2017) e Bruder (2003), segundo a teoria neoclássica da migração, os trabalhadores procuram salários mais altos, geralmente encontrados nos países mais ricos. Níveis baixos de salários são um fator impulsionador de emigração, enquanto que salários mais elevados originam níveis mais altos de imigração e, consequentemente, um aumento do saldo migratório. Como visto em Giulietti (2014), um aumento do Salário Mínimo aumenta a imigração pouco qualificada e não tem efeito na imigração qualificada.

4. RESULTADOS

A análise foi dividida em duas secções: a primeira relacionada com as projeções migratórias em Portugal, baseadas nos três cenários definidos, e a segunda relacionada com a análise de uma regressão linear, identificando possíveis fatores explicativos do

perfil migratório até 2018. Para o efeito foram utilizados os métodos finais descritos no capítulo anterior, bem como analisadas metodologias alternativas.

De seguida, apresentam-se os modelos R-C selecionados e alguns exemplos do seu ajustamento. São também apresentados os fenómenos identificados como explicativos do da evolução do saldo migratório.

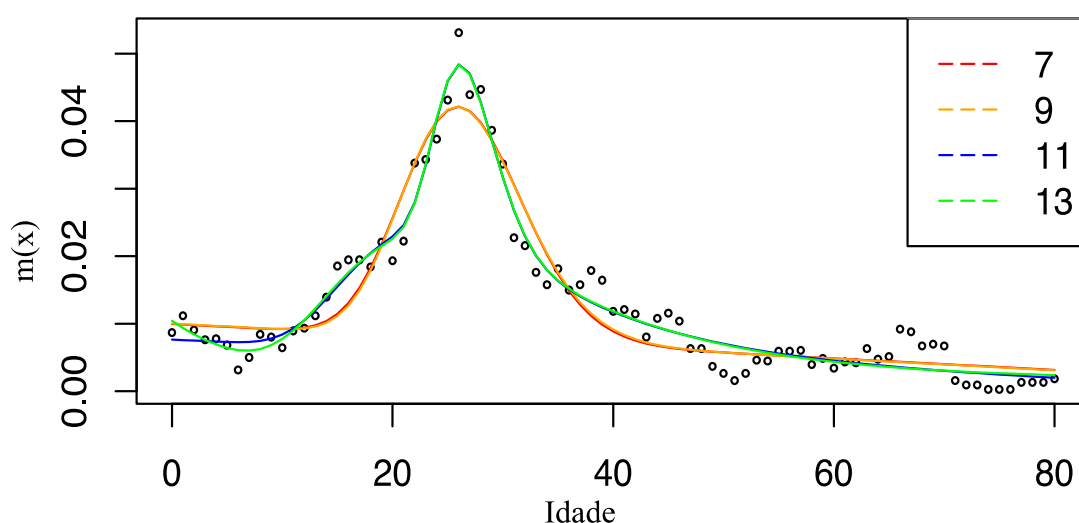
4.1 Projeções em Portugal

O mecanismo mais importante deste subcapítulo é o ajustamento dos dados observados ao modelo R-C. Este ajustamento permite reduzir um grande número de observações num pequeno conjunto de parâmetros, passíveis de interpretação. Espera-se que, com os parâmetros obtidos para o ano de 2018, representativos da migração recente, o perfil migratório das projeções seja realista. Foi também testada a utilização da média dos valores dos parâmetros obtidos em todos os anos observados, contudo, devido às limitações a seguir expostas, não foi possível implementar esta via. Neste método testado, o facto de certos padrões não corresponderem ao esperado, fez com que as médias dos parâmetros fossem irrealistas e representassem curvas exponenciais, em vez dos padrões representados na figura 2.

A falta de informação robusta representa a primeira limitação encontrada. Aos dados obtidos pelo INE, através dos inquéritos realizados, foram aplicadas médias móveis de vários índices e interpolados dos zero aos cem anos. A falta de informação recorrente nas idades mais avançadas levou à escolha dos limites de idade (70 anos no caso da emigração e 80 anos no caso da imigração) para o processo de otimização, evitando a otimização com parâmetros no período de pós-reforma. Esta limitação, segundo Ruiz-Santacruz (2019), faz com que os modelos de 11 e 13 parâmetros sejam geralmente os modelos com melhor qualidade de ajustamento, devido às suas múltiplas curvas exponenciais. De acordo com este autor, a otimização com parâmetros no período de pós-reforma deve-se aos perfis migratórios não encaixarem nos padrões clássicos do modelo (Figura 2), fazendo com que a curva do período pós-reforma se torne crescente ou represente padrões migratórios tardios. Nesta situação, os parâmetros como o μ_2 e μ_3 perdem o significado descrito por Rogers e Castro (1981).

Embora diversos perfis migratórios não correspondam aos padrões esperados, a sua aproximação a um dos padrões é útil para a interpretação da evolução de alguns parâmetros. De seguida, apresenta-se uma figura representativa do ajustamento em 2013 para os imigrantes mulher, correspondente ao valor mais elevado da Raiz do Erro Quadrático Médio (REQM), 0,002403, do modelo de 13 parâmetros, coincidente com o período da crise económica e financeira, onde o perfil migratório não corresponde aos padrões da figura 2.

Figura 3: Modelo de 13 parâmetros para as Imigrantes Mulheres, em 2013



Fonte: Elaboração própria

Nesta figura identificam-se as componentes normalmente observadas, a curva no período da infância e a curva no período laboral, correspondentes ao modelo padrão (sete parâmetros) e não se observam as curvas de reforma e pós-reforma. Na figura observam-se um conjunto de picos no período da vida ativa até ao período da reforma, que o modelo não conseguiu captar. A presença destas múltiplas curvas exponenciais, superiores ao número de curvas representadas no modelo levou a que se esgotassem as típicas curvas descritas e isso originou um pior ajustamento e a um respetivo REQM maior.

A tabela seguinte apresenta o valor dos parâmetros e do R^2 do modelo de 13 parâmetros com maiores valores da REQM e o respetivo ano. À exceção dos emigrantes homem, a que corresponde o ano de 2007, os restantes apresentam valores mais elevados no período da crise financeira e intervenção da Troika em Portugal. Na figura acima

aparece representado o caso das imigrantes mulher, onde não são captados os vários picos observados e onde os parâmetros não correspondem à interpretação clássica. As restantes situações encontram-se representadas nos Anexos.

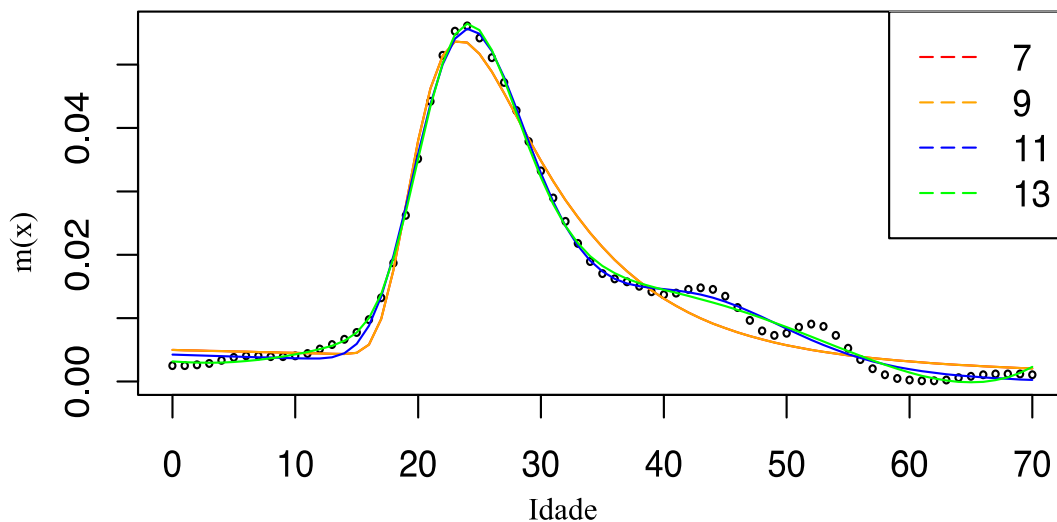
Tabela III: Valores dos parâmetros do modelo de 13 parâmetros com maior REQM e respetivo ano

Parâmetros/Ano	Emigrantes		Imigrantes	
	H	M	H	M
	2007	2012	2012	2013
a_1	0,1274	0,1691	0,1046	0,0088
α_1	0,0304	0,0285	0,0407	0,1331
a_2	0,0391	0,2873	0,0622	0,0484
μ_2	35,4091	63,1973	25,6415	17,8022
α_2	0	0,0403	0,7	0,0689
λ_2	0,3144	0,0206	0,3412	0,1451
a_3	0,0003	0	0,0887	0,011
μ_3	57,6091	47,0881	11,5741	32,5794
α_3	0,393	0,6665	0,0002	0,7
λ_3	0,0602	0,0888	0,1228	0,1989
a_4	0,3216	0,0988	0,4022	0,3017
λ_4	0,0018	0,0129	0	0
c	0,5814	0,6434	0,5068	0,7
REQM	0,00129	0,001623	0,00267	0,002403
R^2	0,21588	0,978	0,94779	0,87264

Uma segunda limitação encontrada é a existência de múltiplas curvas exponenciais em determinados perfis observados, como descrito em cima, o que faz com que se esgotem as curvas exponenciais referidas nos modelos e os parâmetros conseqüentemente não tenham a interpretação desejada. Esta segunda limitação é uma limitação originada pelo modelo, afetando a qualidade da otimização. Esta limitação pode ser verificada tanto no período da crise financeira, como representado na figura 3, como, mais recentemente,

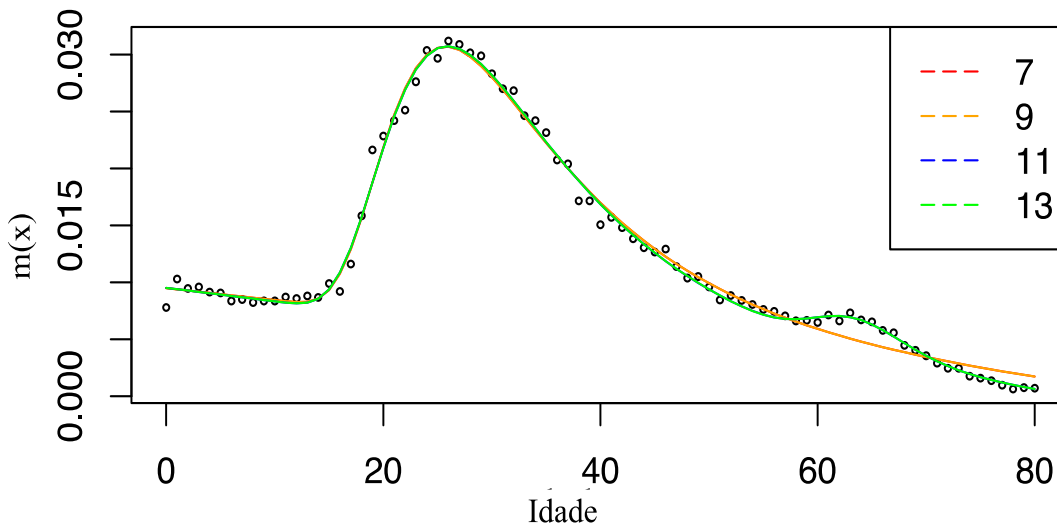
no caso dos emigrantes, como se pode verificar o caso dos emigrantes mulher em 2018 na figura seguinte. O caso das emigrantes homem pode ser encontrado na figura 13 dos Anexos.

Figura 4: Modelo de 13 parâmetros para as Emigrantes Mulher, em 2018



Fonte: Elaboração própria

Figura 5: Modelo de 13 parâmetros para as Imigrantes Homem, em 2018



Fonte: Elaboração própria

Nesta segunda figura, relativa ao ano de 2018 para os imigrantes homem, os padrões são novamente captados, mas a curva no período laboral não capta um fenómeno recente,

a migração estudantil, que pode ser captada pelo modelo de 16 parâmetros, apresentado por Wilson (2010).

A tabela IV, semelhante à tabela anterior, apresenta os parâmetros e o R^2 para os diversos perfis, agora no ano de 2018. Serão estes parâmetros, correspondentes aos perfis migratórios das figuras 4 e 5 e das figuras 13 e 14 que se encontram nos Anexos, que representarão o perfil migratório dos anos projetados.

Tabela IV: Valores dos parâmetros do modelo de 13 parâmetros para o ano de 2018

Parâmetros/Modelo	Emigrantes		Imigrantes	
	H	M	H	M
	13	13	13	13
a_1	0,3807	0,3049	0,7	0,7
α_1	0,0005	0,0022	0,0003	0,0002
a_2	0,0984	0,0054	0,0018	0
μ_2	27,9472	36,4493	81,7524	89,4442
α_2	0,1545	0,5254	0,2883	0,6855
λ_2	0,1187	0,1188	0,0786	0,0823
a_3	0,0492	0,0649	0,0483	0,0508
μ_3	43,3943	72,3538	21,4058	21,3184
α_3	0,0678	0,059	0,0443	0,0752
λ_3	0,1415	0,0271	0,1519	0,2151
a_4	0	0	0	0
λ_4	0,3261	0,12	0,1387	0
c	0,6226	0,6945	0,3085	0,3096
REQM	0,001181	0,000971	0,000616	0,000698
R^2	0,98813	0,99613	0,99347	0,99282

Das Tabelas I e II pode-se concluir que os parâmetros μ_2 e μ_3 obtidos por uma distribuição a priori não informativa, não correspondem à interpretação clássica, definida por Rogers e Castro (1981). O ajustamento aos dados observados gera um REQM baixo, no caso da Tabela II, contudo os parâmetros de localização não correspondem à

interpretação esperada. Poderá esta situação mostrar a pequena influência que estes parâmetros têm na melhor estimativa. Para contornar esta situação poderia ser adicionado um critério de especialista, limitando o intervalo destes parâmetros, podendo igualmente ser geradas curvas que representam um bom ajustamento aos dados observados e a que correspondam um valor baixo da REQM.

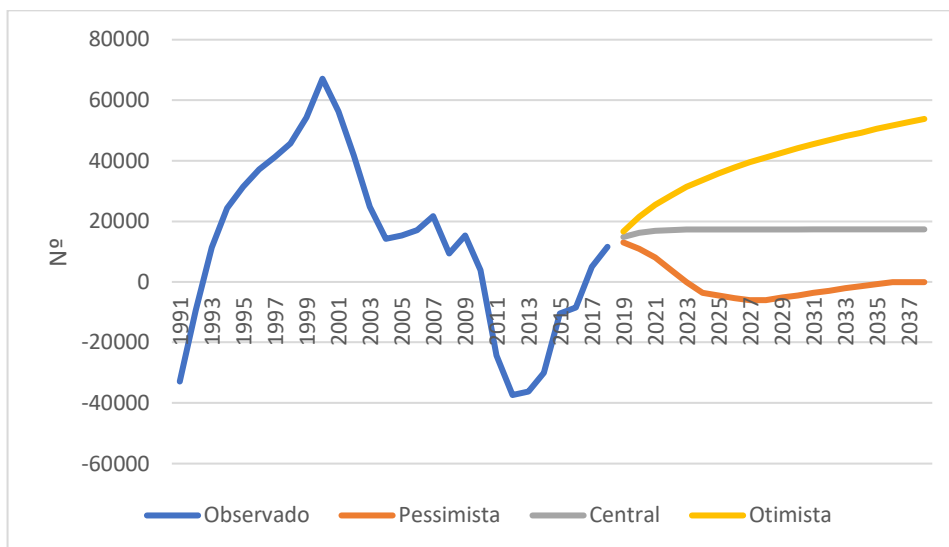
Para as projeções foi aplicada uma metodologia automática de obtenção do melhor modelo ARIMA, isto é, que apresenta o menor valor para o critério BIC. Com estes modelos e o seu respetivo intervalo de confiança a 80% foram desenhados os três cenários definidos. No caso dos emigrantes homem foi selecionado o modelo ARIMA (0,1,0) e para os emigrantes mulher o modelo ARIMA (1,1,0). Para o caso dos imigrantes foi selecionado o modelo ARIMA (1,1,0) para os homens e o modelo ARIMA (1,1,0) para as mulheres. Os valores de previsão para o último ano projetado, 2038, encontram-se na tabela seguinte. No caso do cenário pessimista, tanto para emigrantes como para imigrantes, verificaram-se valores negativos e sendo este um cenário irrealista foi assumido um valor igual a zero nos respetivos anos.

Tabela V: Cenários em 2038 em comparação com 2018

Cenário	Emigrantes		Imigrantes		Saldo Migratório
	H	M	H	M	H+M
2018	23181	8419	20240	22930	11570
Pessimista	0	0	0	0	0
Central	23181	7609	22574	25573	17357
Otimista	47456	31393	69093	63581	53825

Nos cenários central e otimista obtém-se um saldo migratório positivo, traduzindo-se num maior número de imigrantes do que emigrantes. Esta situação, tendo em conta o panorama nacional de uma população envelhecida, constitui uma conclusão positiva. Em contraste, no cenário pessimista o número de emigrantes cresce a um nível superior ao número de imigrantes a partir de 2023, apresentando um saldo migratório negativo. Este fenómeno ocorre até 2037, ano a partir do qual se verifica um saldo migratório nulo. A figura seguinte ilustra a evolução destes cenários.

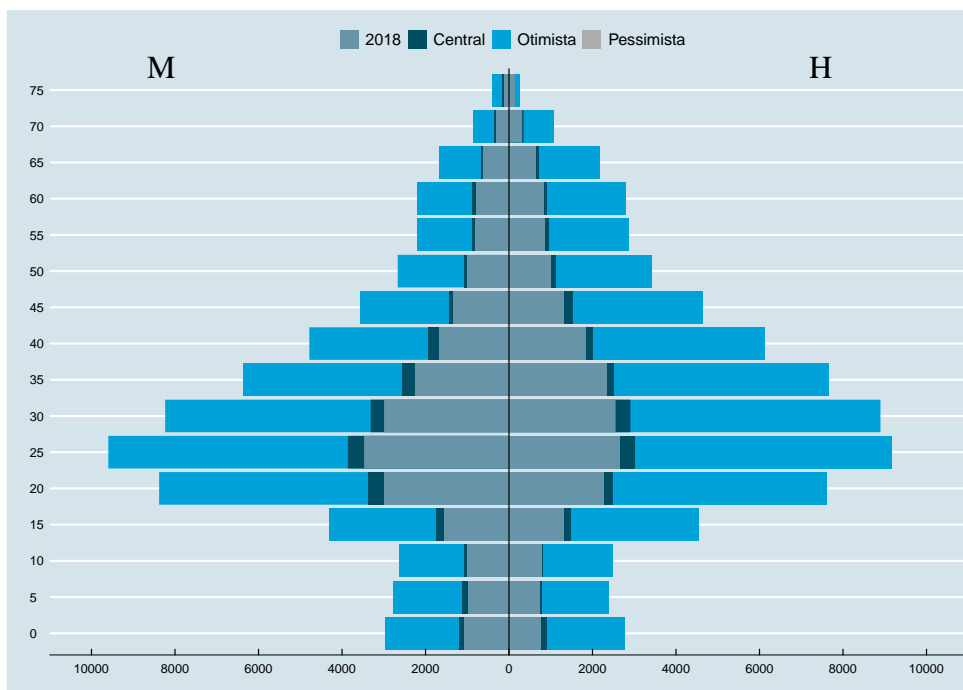
Figura 6: Evolução do Saldo Migratório nos cenários definidos



Fonte: Elaboração própria

Estes valores de previsão, conjuntamente com o perfil migratório obtido para o ano de 2018 permitem construir as projeções até ao ano definido. Na figura seguinte pode-se observar, por grupos quinquenais, a comparação para o caso dos imigrantes, para o ano de 2018 e para os cenários definidos em 2038.

Figura 7: Comparação da estrutura populacional dos imigrantes, em 2018 e 2038



Fonte: Elaboração própria

Nesta figura pode-se observar que apenas o cenário pessimista não supera os valores verificados em 2018. Contudo, há exceção deste cenário, o número de imigrantes foi superior ao número de emigrantes, originando um saldo migratório positivo até ao último ano projetado. Pode-se também observar um padrão em que a imigração feminina é superior à masculina nas faixas etárias mais novas, até aos 45 anos no caso do cenário central e até aos 35 anos no cenário otimista. Assim, conclui-se que é a partir da faixa etária ativa que se começa a verificar uma inversão da tendência, sendo a imigração masculina mais predominante até às últimas faixas etárias analisadas. O cenário pessimista é uma exceção a este fenómeno, uma vez que neste cenário não se verificam movimentos de imigração, sendo uma conclusão pouco plausível de verificar.

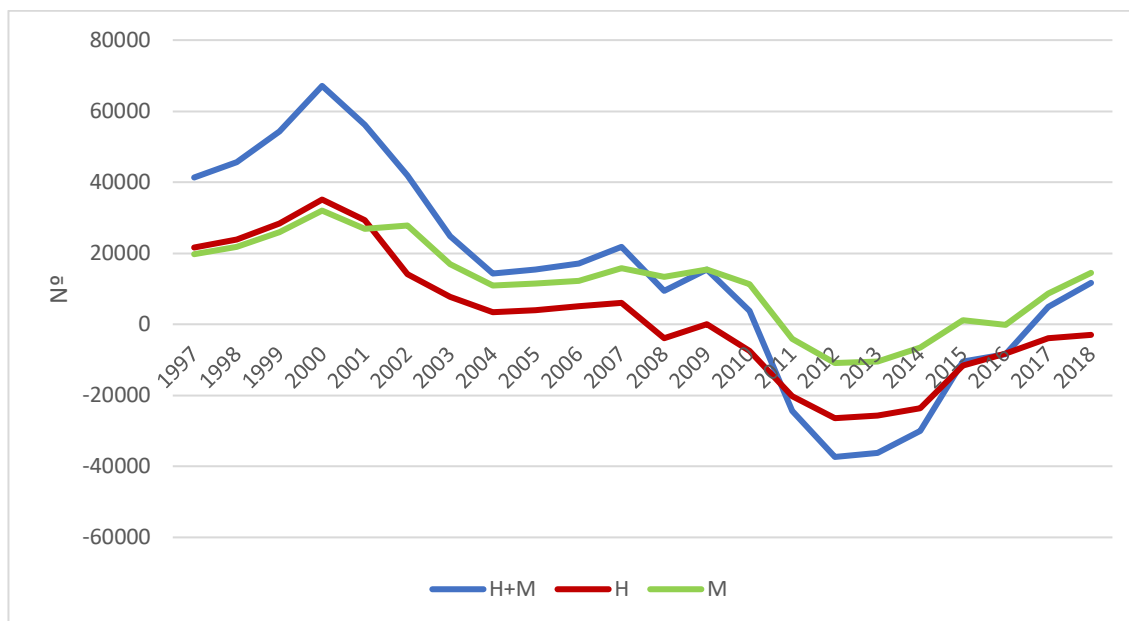
O caso dos emigrantes encontra-se representado na figura 19 dos Anexos. Neste caso observa-se que, para algumas faixas etárias, o cenário central e o cenário otimista apresentam valores superiores aos verificados em 2018. Esta é uma conclusão pouco animadora uma vez que representa um aumento da emigração, originando saldos migratórios mais baixos e, até mesmo nulos no caso do cenário pessimista. Há uma maior predominância de emigração masculina, no entanto, na faixa etária entre os 20 e os 30 anos, observa-se um número elevado de emigrantes mulheres. Tal como verificado no caso dos imigrantes, no cenário pessimista não se verificam movimentos de emigração em nenhuma faixa etária, sendo novamente uma situação pouco plausível de se verificar e, contribui assim, para o menor saldo migratório dos três cenários analisados.

4.2 Regressão

A variável dependente representa o saldo migratório em Portugal entre o ano de 1997 e 2018. Esta variável apresentou valores positivos desde 1997 até 2011, ano em que se verificou um grande crescimento do número de emigrantes e que não foi compensado por um igual crescimento do número de imigrantes. Esta tendência negativa manteve-se até 2017, onde se verificou novamente um saldo migratório positivo e com tendência crescente. A evolução positiva do saldo migratório das mulheres superou o saldo ainda negativo relativo aos homens, o que originou o saldo global positivo. É de destacar uma inversão do padrão em 2002, onde o saldo migratório das mulheres passou a ser superior

ao dos homens. O comportamento desta variável encontra-se representado na figura seguinte:

Figura 8: Evolução do Saldo Migratório entre 1997 e 2018



Fonte: Elaboração própria

Esta análise foi introduzida com o objetivo de identificar fatores explicativos do saldo migratório verificado e que sirvam futuramente como elementos chave na tomada de decisão, com a perceção do conseqüente impacto no padrão migratório.

O modelo proposto não seguiu o caminho desejado, isto é, na significância verificada das suas variáveis. Das variáveis independentes escolhidas algumas tiveram o sinal esperado e apenas uma apresentou significância estatística. Estas não apresentam problemas de multicolineariedade, como se pôde verificar na matriz de correlação apresentada no capítulo anterior. Este facto salienta que nenhuma variável é uma função linear de outra, isto é, não apresentam uma correlação perfeita, que poderia tornar as estimativas imprecisas.

Tabela VI: Estimativas dos coeficientes
(Variável dependente: Saldo Migratório)

	Coeficiente	Estatística t	Valor-p
Δ (Pibpercap(-1))	-7,759	-1,167	0,261
Δ (Desemp(-1))	-0,148	-2,694	0,017**
Δ (Salmin(-1))	-405,702	-1,335	0,202
Δ (Spread(-1))	-1406,715	-0,817	0,427

Nota: os asteriscos ***, **, * indicam significância a 1, 5, 10% respetivamente.

Estes resultados foram obtidos com a estimação do modelo pelo método dos mínimos quadrados (OLS) e deriva de um teste aplicado sobre 20 observações, tendo um R^2 igual a 0,467. A variável relativa ao desemprego foi a única variável a apresentar significância estatística, a um nível de 5%.

O coeficiente obtido da estimativa para a variável “ Δ (Desemp(-1))” corresponde ao sinal esperado. Conforme a literatura revista, o seu sinal negativo levará a que um aumento no número de desempregados origine uma diminuição do saldo migratório, onde há um aumento da emigração em busca de melhores condições e uma diminuição da imigração em busca de oportunidades no nosso país. ¹

Foi realizado o teste Breusch-Godfrey para verificar a autocorrelação dos resíduos e uma vez que a hipótese nula não foi rejeitada pode-se concluir que esta não existe. Foi também assumido que não existe heterocedasticidade, uma vez que, após a realização do teste de White, a hipótese nula da presença de homocedasticidade não foi rejeitada.

¹ Foi testada a introdução de duas variáveis *dummy* representativas de fenómenos verificados em Portugal: as variáveis Troika e Lei da Nacionalidade. A variável Troika refere-se ao período de intervenção da Troika, com o pedido de resgate financeiro, entre 2011 e 2014. Neste período, verificou-se um saldo migratório negativo, originado pelo grande declínio da imigração. A variável Lei da Nacionalidade refere-se à Lei Orgânica n.º 2/2006, de 17 de abril, quarta alteração à Lei n.º 37/81, de 3 de outubro, e que alterou os regimes de atribuição da nacionalidade portuguesa. Segundo Oliveira e Gomes (2014) e Góis (2019), a Lei da Nacionalidade é a medida mais distintiva de inserção social e reconhecida internacionalmente como a melhor política promotora da integração. Esta lei e a sua posterior alteração (Decreto-Lei n.º 237-A/2006, de 14 de dezembro) levaram a que o número de nacionalidades atribuídas passasse de poucos milhares para mais de vinte mil a partir de 2008. Estas variáveis não foram introduzidas devido ao baixo número de observações, no entanto, uma vez que nunca foram utilizadas seria útil a sua análise. De acordo com a literatura revista, seria expectável que a variável Troika tivesse um efeito negativo no saldo migratório, enquanto a Lei da Nacionalidade tivesse um efeito positivo

Nesta análise verificam-se algumas limitações. Primeiramente, o baixo número de observações implicou um baixo número de variáveis explicativas e, conseqüentemente, poderá ter levado a resultados espúrios. Em segundo lugar, a abstração das diferenças entre emigrantes e imigrantes, como a sua idade, nível de riqueza e o seu impacto no crescimento económico. Em terceiro lugar, a inexistência de variáveis relativas ao contexto internacional poderá ter originado coeficientes com sinais contrários, onde seria expectável que a evolução positiva de diversos fatores económicos levasse a um efeito positivo no saldo migratório, pela via da redução da emigração.

5. CONCLUSÃO

De acordo com a estrutura etária da população portuguesa, é imprescindível ao país conseguir projeções sobre as migrações futuras, de modo a estudar o seu impacto na estrutura da população e na economia. Sendo a população migrante mais jovem é de extrema importância que Portugal tome medidas para inverter o cenário atual e consiga rejuvenescer e aumentar a sua população.

O presente trabalho final de mestrado teve como principal objetivo obter vários cenários de projeção dos fluxos migratórios em Portugal e encontrar fatores explicativos para o período já observado, fatores estes que podem ser decisivos para influenciar as migrações futuras. Com o desenvolvimento deste estudo, foi possível munir o INE de um programa implementado em R, que realiza de forma automática, a aplicação do modelo Rogers-Castro aos anos observados, segundo os critérios definidos, e identifica o melhor modelo ARIMA para previsão. Este programa permite, assim, analisar os vários parâmetros dos modelos ajustados e criar cenários fixando o valor dos parâmetros, segundo o julgamento de especialista.

A implementação do modelo Rogers-Castro permitiu contornar o problema da falta de robustez dos dados e, por meio da otimização, ajustá-los às curvas conhecidas como padrão. Os resultados apresentados demonstram que alguns perfis migratórios não correspondem aos padrões estabelecidos, porém foi possível extrair informação importante. No decorrer desta implementação surgiu um outro obstáculo, onde certos parâmetros ao não corresponderem à interpretação clássica, não permitiu a utilização da

média destes como valores para o perfil das projeções. Para contornar esta situação, foi utilizado o perfil obtido pelos parâmetros correspondentes ao ano de 2018.

Com base nos resultados apresentados é possível verificar que os cenários central e otimista permitem obter uma população imigrante em idade ativa superior à verificada em 2018. Estes cenários, até 2038, permitem verificar um saldo migratório positivo, o que é uma conclusão animadora para Portugal. Por seu turno, o cenário pessimista apresenta valores negativos, convergindo para um saldo migratório nulo no último ano projetado.

Para trabalhos futuros poderá ser utilizado outro método de otimização, onde a mesma simulação poderia ser aplicada aos valores iniciais. Poderá também ser possível criar novos cenários onde sejam adicionados valores definidos dos parâmetros aos valores obtidos para a projeção, obtendo diversas estruturas de população para os mesmos valores de projeção.

Refira-se ainda que o INE utiliza médias móveis para interpolação dos dados em falta dos inquéritos eletrónicos realizados, pelo que o estudo e aplicação de métodos alternativos poderiam melhorar os resultados obtidos, ultrapassando os problemas identificados.

À parte do proposto pelo INE, a regressão linear introduzida teve como principal limitação a ausência de um número elevado de observações, prejudicando uma maior precisão das estimativas. Em trabalhos futuros, é de esperar a utilização de uma amostra maior destas variáveis e a introdução de variáveis relativas ao contexto internacional.

O presente trabalho seria melhorado se tanto para a aplicação do modelo R-C como para a regressão, fosse implementado um modelo do tipo gravitacional. Esta análise não foi possível realizar no decorrer do trabalho por falta de disponibilização de dados. Esta permitiria não só chegar aos resultados já obtidos, mas também identificar o tipo e nacionalidade do migrante, isto é, saber quais os motivos e os países de onde provêm os imigrantes e quais os países de destino dos emigrantes.

6. REFERÊNCIAS BIBLIOGRÁFICAS

Alessandria, G.A., Bai, Y., & Deng, M. (2020). Migration and sovereign default risk. *Journal of Monetary Economics*.

Bell, M. (2005) “Towards rigorous cross-national comparison of internal migration: Who collects what?”, paper presented at the 25th International Population Conference, International Union for the Scientific Study of Population (IUSSP), Tours, France, 18-23 July 2005.

Bernard, A., & Bell, M. (2015). Smoothing internal migration age profiles for comparative research. *Demographic Research* 32 (1), 915-948.

Bijak, J. (2006) Forecasting International Migration: Selected Theories, Models, and Methods. Working Paper 4/2006. Central European Forum for Migration Research (CEFMR): Warsaw, Poland.

Bijak, J., Disney, G., Findlay, A.M., Forster, J.J., Smith, P.W.F. & Wiśniowski, A. (2019). Assessing time series models for forecasting international migration: Lessons from the United Kingdom. *Journal of Forecasting* 38 (5), 470-487.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting* 22 (3), 547-581.

Box, G. & Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.

Bruder, J. (2003). East-West Migration in Europe 2004-2015. Thünen- Series of Applied Economic Theory, Working Paper No. 40, University of Rostock.

Caiado, J. (2016). *Métodos de Previsão em Gestão com aplicações em Excel*, 2ª Ed., Lisboa: Edições Sílabo.

Castelli, F. (2018). Drivers of migration: why do people move?. *Journal of Travel Medicine* 25 (1).

Disney, G., Winiowski, A., Forster, J. J., Smith, P. W. F., & Bijak, J. (2015). Evaluation of existing migration forecasting methods and models. Report for the Migration Advisory Committee. ESRC Centre for Population Change, University of Southampton.

Giulietti, C. (2014). Is The Minimum Wage A Pull Factor For Immigrants? *ILR Review* 67, 649-674.

Góis, P. (2019). *Casa Comum. Migração e Desenvolvimento em Portugal. Avançar com as práticas: rumo à inclusão e coesão social*. Lisboa: Caritas Portuguesa.

Hyndman, R. J., Booth, H., & Yasmeen, F. (2013). Coherent mortality forecasting: the product-ratio method with functional time series models. *Demography* 50(1), 261-283.

Instituto Nacional de Estatística (2020a). Documento Metodológico. Projeções de População Residente. Versão 4.0. Lisboa: Instituto Nacional de Estatística.

Instituto Nacional de Estatística (2020b). Projeções de População Residente 2018-2080. Lisboa: Instituto Nacional de Estatística.

International Organization for Migration (2019). Glossary on Migration. International migration law No. 34. Geneva: IOM International Organization for Migration.

Jennissen, R. (2004). *Macro-economic determinants of international migration in Europe*. Amsterdam: Dutch University Press.

Keilman, N., Pham, D. & Hetland, A. (2002). Why Population Forecasts should be Probabilistic - illustrated by the case of Norway. *Demographic Research* 6 (15), 409-454.

Lange, F. & Gollin, D. (2007). Equipping Immigrants: Migration Flows and Capital Movements. IZA Discussion Paper No. 2745. Institute of Labor Economics (IZA). Disponível em SSRN: <https://ssrn.com/abstract=982597>.

Oliveira, C. & Gomes, N. (2014). Monitorizar a integração de Imigrantes em Portugal Relatório Estatístico Decenal, Coleção Imigração em Números do Observatório das Migrações. Lisboa: ACM.

Peixoto, J. (2009). A demografia da população imigrante em Portugal. In Lages, M. F. e Matos, A. T. (coord.), Portugal: Percursos de Interculturalidade (vol. II), pp. 7-47. Lisboa: ACIDI e CEPCEP.

PORDATA (2020). Índice de Envelhecimento [Base de dados], março 2020. Lisboa: PORDATA. Disponível em: <https://www.pordata.pt/Europa/Índice+de+envelhecimento-1609>.

Rogers, A. & Castro, L. J. (1981). Model migration schedules. Research Report 81-30. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A., Little, J. & Raymer, J. (2010). The Indirect Estimation of Migration: Methods for Dealing with Irregular, Inadequate, and Missing Data. Dordrecht, The Netherlands: Springer.

Ruiz-Santacruz, J. & Garcés, J. (2018). migraR. Prototype package for adjusting Rogers and Castro models. Retrieved: 02/03/20

Ruiz-Santacruz, J. (2019). Estimation of migration schedules by simulating the initial values in the parameter optimizations of multi-exponential migration models: an application to intra-Latin American international migration. Papers de demografia 463, 1-77. Bellaterra: Centre d'Estudis Demogràfics.

Santos, R. O. & Barbieri, A. F. (2019). Funções modelo de migração: limites e aplicações. Revista Brasileira De Estudos De População 36, 1-25.

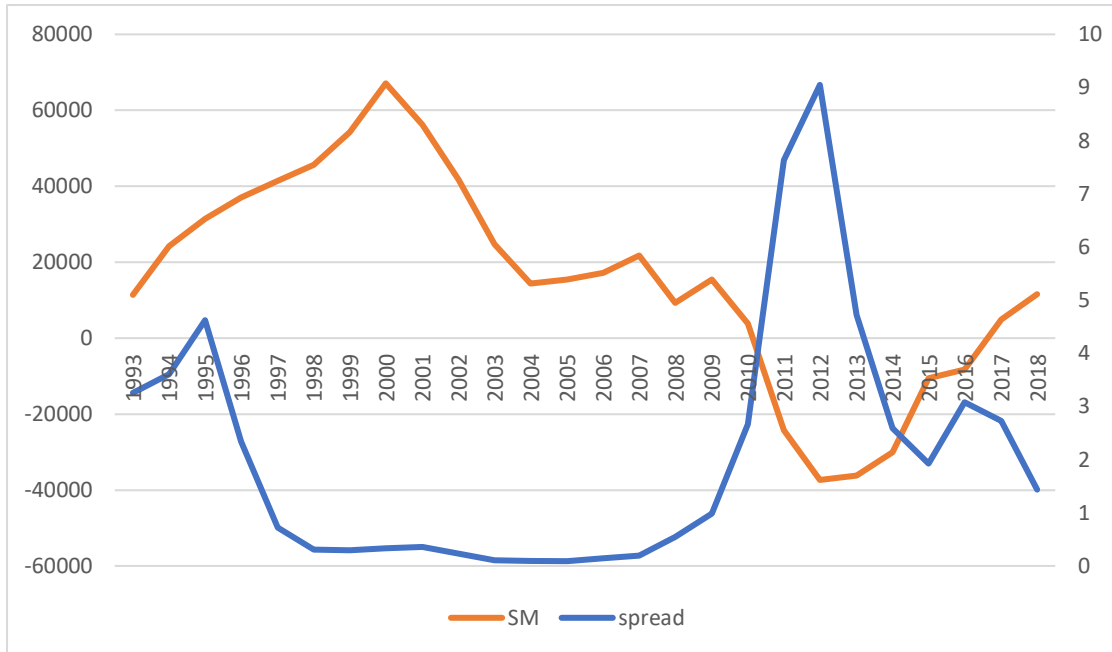
Simpson, N. B. (2017). Demographic and economic determinants of migration: Push and pull factors drive the decision to stay or move. *IZA World of Labor*, 373.

UNDESA Statistics Division. (1998). Recommendations on Statistics of International Migration: Revision 1. *Statistical Papers Series M No. 58*.

Wilson, T. (2010). Model migration schedules incorporating student migration peaks. *Demographic Research* 23 (8), 191-222.

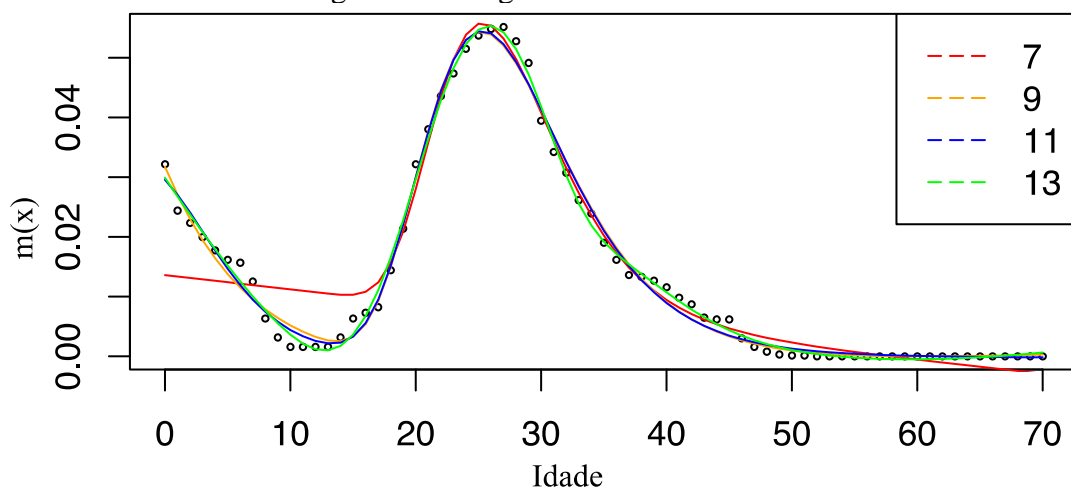
7. ANEXOS

Figura 9: Evolução do saldo migratório e do spread



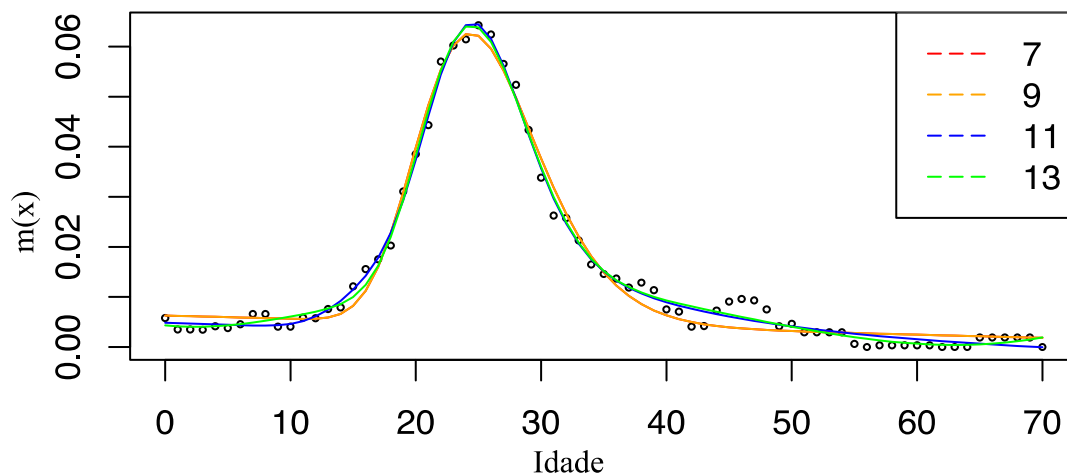
Fonte: Elaboração própria

Figura 10: Emigrantes Homem 2007



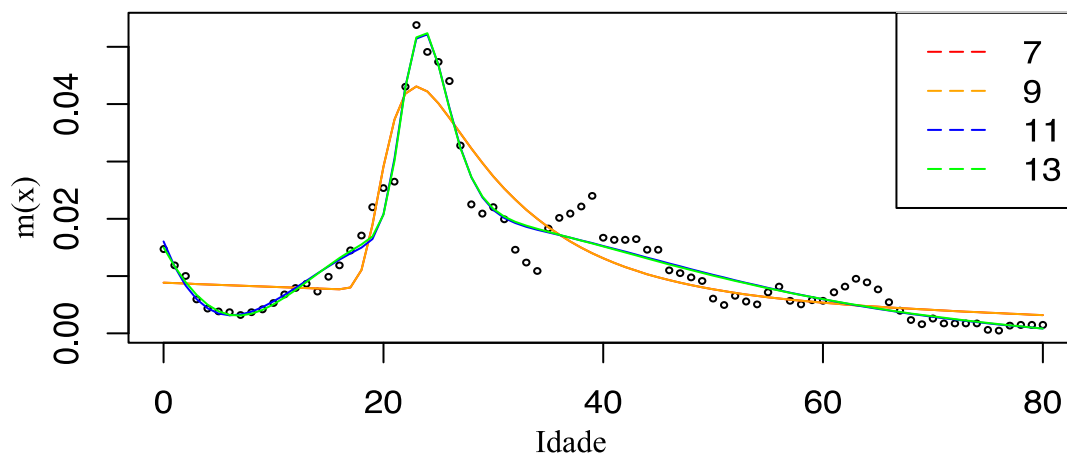
Fonte: Elaboração própria

Figura 11: Emigrantes Mulher 2012



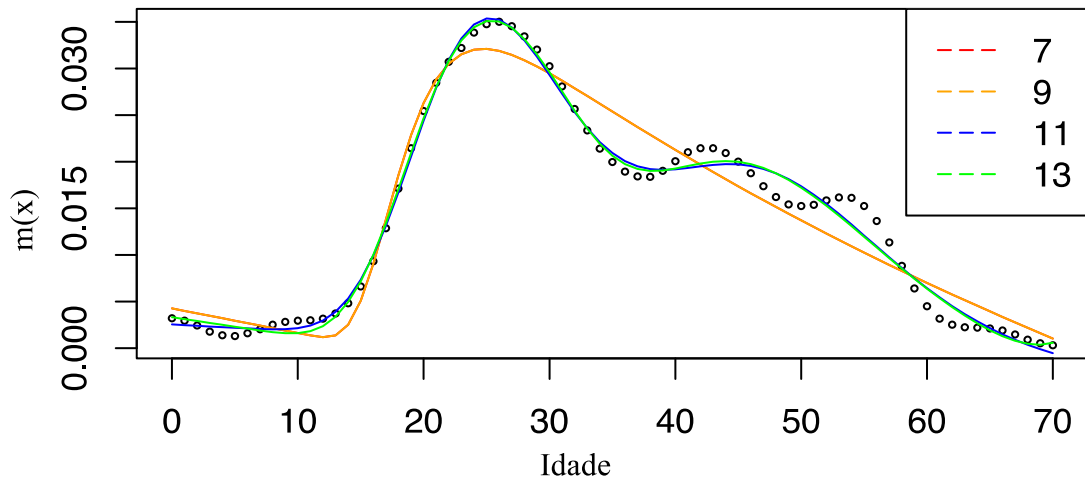
Fonte: Elaboração própria

Figura 12: Imigrantes Homem 2012



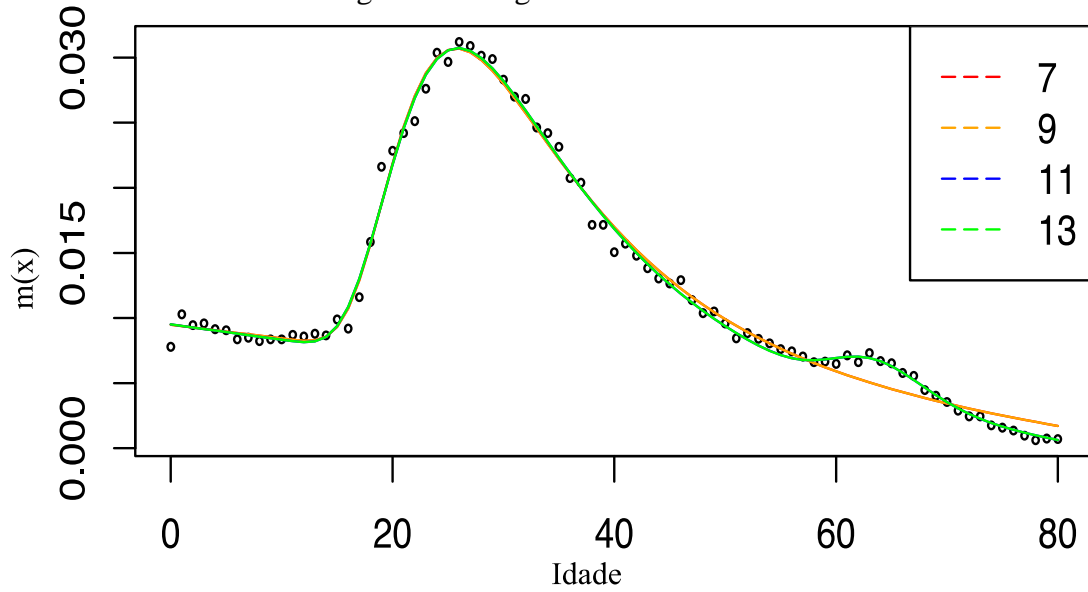
Fonte: Elaboração própria

Figura 13: Emigrantes Homem 2018



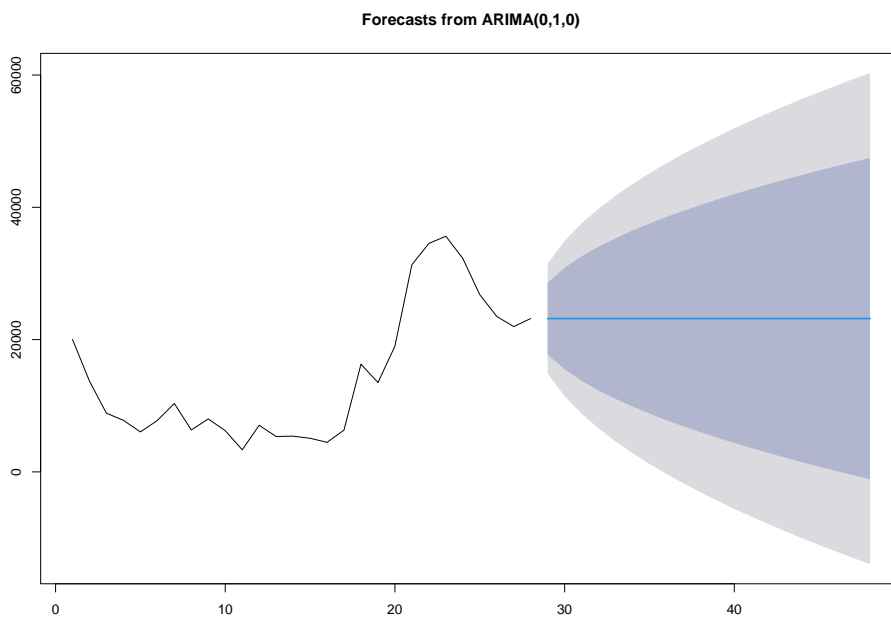
Fonte: Elaboração própria

Figura 14: Imigrantes Mulher 2018



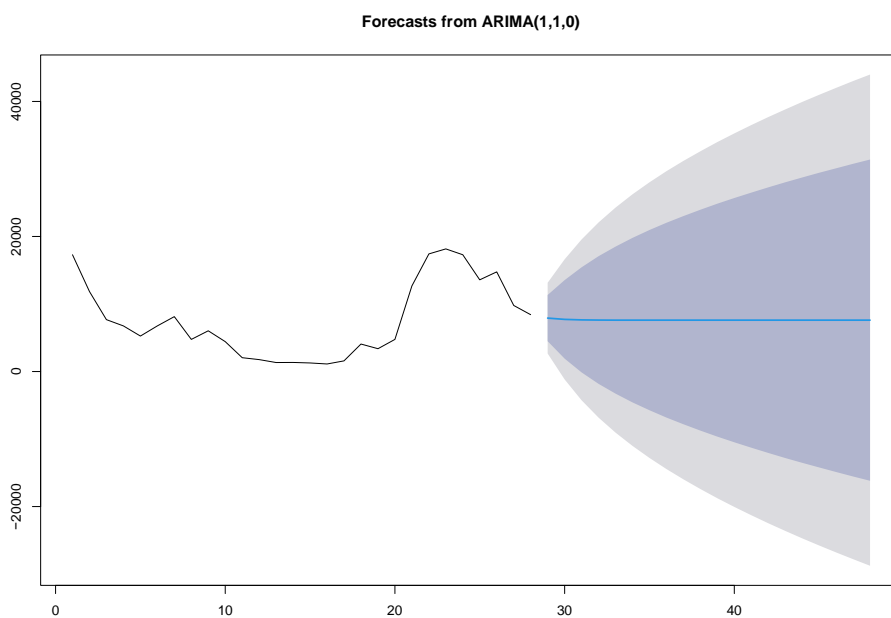
Fonte: Elaboração própria

Figura 15: Previsão Emigrantes Homem



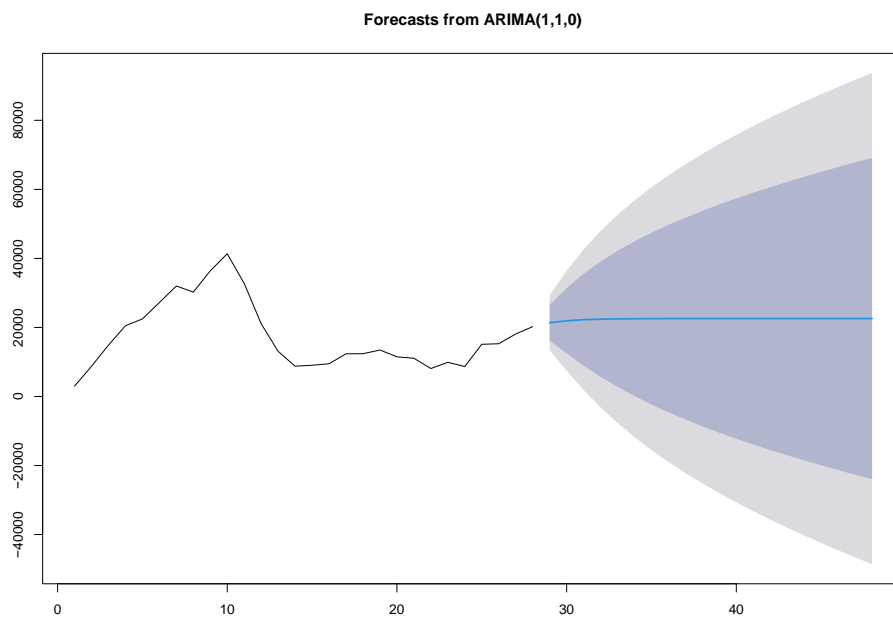
Fonte: Elaboração própria

Figura 16: Previsão Emigrantes Mulher



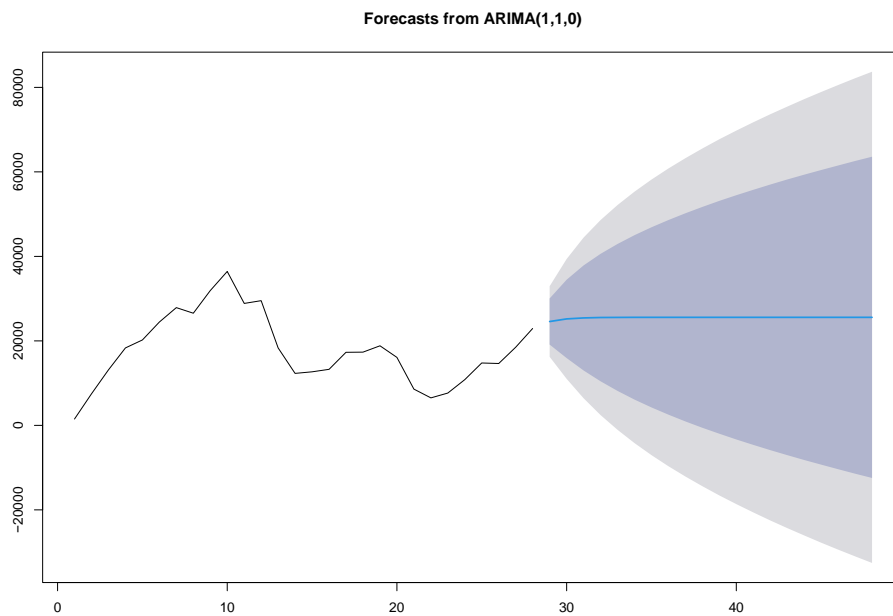
Fonte: Elaboração própria

Figura 17: Previsão Imigrantes Homem



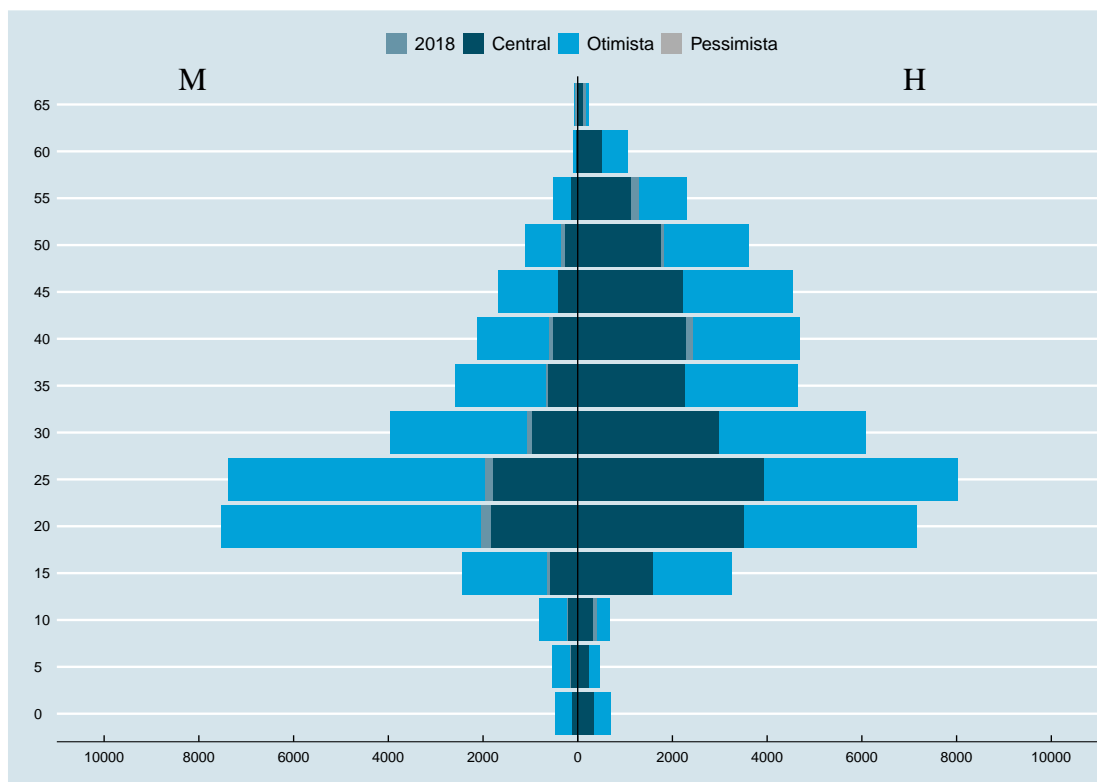
Fonte: Elaboração própria

Figura 18: Previsão Imigrantes Mulher



Fonte: Elaboração própria

Figura 19: Comparação da estrutura populacional dos emigrantes, em 2018 e 2038



Fonte: Elaboração própria

Tabela VII: Conjunto de dados

Variável	Descrição	Fonte	Sinal esperado
SM	Saldo Migratório	INE	
PIB per cap	PIB per capita (PIB no ano civil/População média residente)	PORDATA	(+)
Desemp	Número anual de desempregados	INE	(-)
Sal Min	Salário Mínimo geral mensal	PORDATA	(+)
Spread	Spread da taxa de juro das obrigações a 10 anos portuguesa em relação à alemã	BCE	(-)