



**Universidade de Brasília
Faculdade de Tecnologia**

**Agregação de modelos adaptativa para previsão
de produção em reservatórios**

Thiago Tomás de Paula

PROJETO FINAL DE CURSO
ENGENHARIA DE CONTROLE E AUTOMAÇÃO

Brasília
2023

P324a Paula, Thiago Tomás de.
Agregação de modelos adaptativa para previsão de produção em reservatórios / Thiago Tomás de Paula; orientador: Eugênio Libório Feitosa Fortaleza; coorientador: José Oniram de Aquino Limaverde Filho. -- Brasília, 2023.
91 p.

Projeto Final de Curso (Engenharia de Controle e Automação)
-- Universidade de Brasília, 2023.

1. Previsão a curto prazo. 2. Método de Agregação. 3. Filtro Adaptativo. 4. Reservatório. I. Fortaleza, Eugênio Libório Feitosa, orient. II. Filho, José Oniram de Aquino Limaverde, coorient. III. Título

**Universidade de Brasília
Faculdade de Tecnologia**

**Agregação de modelos adaptativa para previsão
de produção em reservatórios**

Thiago Tomás de Paula

Projeto Final de Curso submetido como requisito
parcial para obtenção do grau de Engenheiro de
Controle e Automação

Orientador: Prof. Dr. Eugênio Libório Feitosa Fortaleza
Coorientador: Dr. José Oniram de Aquino Limaverde Filho

Brasília
2023

**Universidade de Brasília
Faculdade de Tecnologia**

**Agregação de modelos adaptativa para previsão de
produção em reservatórios**

Thiago Tomás de Paula

Projeto Final de Curso submetido como requisito
parcial para obtenção do grau de Engenheiro de
Controle e Automação

Trabalho aprovado. Brasília, 12 de dezembro de 2023:

Prof. Dr. Eugênio Libório Feitosa
Fotaleza, UnB/FT/ENM
Orientador

Dr. José Oniram de Aquino Limaverde
Filho, UnB/FT/ENM
Coorientador

Dr. William Humberto Cuéllar Sánchez
Examinador interno

Prof. Dr. Manoel Porfírio Cordão Neto,
UnB/FT/ENC
Examinador externo

Brasília
2023

Este trabalho é dedicado aos professores Luiz Antônio Rosa Braz e Paulo Luiz da Silva Ramos, que me ensinaram matemática.

Agradecimentos

Ao meu coorientador, José Oniram, pelo esforço e carinho demonstrados durante as revisões do manuscrito, e pela competência em delinear seu desenvolvimento. Ao meu orientador, Eugênio Fortaleza, por me mostrar o que um engenheiro pode ser.

“A educação tem raízes amargas, mas os seus frutos são doces.”
(Aristóteles)

Resumo

Estratégias de previsão de produção de reservatórios que gerem resultados precisos de maneira eficaz são motivo de extensa pesquisa na literatura, geralmente financiada pela indústria. Nesta, destacam-se duas abordagens: o ajuste de histórico, e a agregação de modelos. No primeiro método, uma malha 3D que imita o reservatório real tem parâmetros como porosidade e permeabilidade calibrados através de comparações entre simulações de produção e observações da produção do reservatório físico. Feita a calibração, simulações de produção da malha gerarão estimativas precisas da produção do reservatório real. Trata-se de um método pouco abstrato, embora extremamente exigente em complexidade computacional. Por outro lado, a agregação de modelos é um método numérico que baseia suas previsões em um grupo de especialistas e no erro acumulado durante previsões passadas. Os especialistas consistem em modelos sintéticos do reservatório que para cada passo de tempo fornecem uma sugestão da produção; através de uma regra de agregação, as sugestões são ponderadas com respeito ao erro acumulado, gerando a estimativa. Em comparação ao ajuste de histórico, trata-se de um método mais abstrato, mas com aplicabilidade flexível além de reservatórios e exigência mínima de capacidade computacional. Todavia, os métodos de agregação de modelos sedimentados na literatura não são capazes de considerar um fenômeno típico em reservatórios: o ruído de medição. Este trabalho compara, através de simulações numéricas, a precisão de 3 métodos de agregação bem estabelecidos na literatura, a saber, o Ridge, o Lasso e a Média Ponderada Exponencial, contra um quarto e novo método, a Agregação de Modelos Adaptativa, que usa filtro ativo inspirado no trabalho de [Fortaleza, Gomes et al. \(2022\)](#) para tornar as previsões mais robustas contra ruído de medição, em particular as previsões à curto prazo (1 ano). O Filtro Adaptativo é um mecanismo matemático simples e de fácil implementação industrial que busca melhorar sensibilidade intrínseca da Média Móvel àqueles especialistas que contribuem menos ao erro acumulado. Após apresentar os detalhes e funcionamento de cada método de agregação, a análise comparativa dos erros relativos mostrou que o método proposto forneceu as melhores previsões, que se tornam tão melhores, i.e., menos imprecisos, quanto melhor ajustados forem seus parâmetros. Em suma, a Agregação de Modelos Adaptativa foi validada como método de agregação com previsões precisas e velocidade computacional relativamente alta.

Palavras-chave: Previsão a curto prazo. Método de Agregação. Filtro Adaptativo. Reservatório.

Abstract

Strategies for predicting reservoir production that generate accurate results effectively are the subject of extensive research in the literature, often funded by the industry. Two approaches stand out: history matching and model aggregation. In the first method, a 3D grid resembling the actual reservoir has parameters like porosity and permeability calibrated by comparing production simulations with observations from the physical reservoir. Once calibrated, production simulations from the grid will generate precise estimates of the actual reservoir's production. This method is less abstract but highly demanding computationally. On the other hand, model aggregation is a numerical method that bases its predictions on a group of experts and the accumulated error during past forecasts. The experts consist of synthetic models of the reservoir that provide a production suggestion for each time step; through an aggregation rule, these suggestions are weighted concerning the accumulated error, generating the estimate. Compared to history matching, this method is more abstract but has flexible applicability beyond reservoirs and minimal computational requirements. This work compares, through numerical simulations, the accuracy of three well-established aggregation methods in the literature, namely Ridge, Lasso, and Exponentially Weighted Average, against a fourth and new method, Adaptive Model Aggregation. The latter employs an active filter inspired by the work of [Fortaleza, Gomes et al. \(2022\)](#) to make predictions more robust against measurement noise. The Adaptive Filter is a simple mathematical mechanism with easy industrial implementation aiming to enhance the intrinsic sensitivity of the Moving Average to those experts contributing less to the accumulated error. After presenting the details and operation of each aggregation method, the comparative analysis of relative errors showed that the proposed method provided the best predictions, improving as its parameters were better adjusted. In summary, Adaptive Model Aggregation was validated as an aggregation method with accurate predictions and relatively high computational speed.

Keywords: Short-term prediction. Aggregation Method. Adaptive Filter. Reservoir.

Lista de ilustrações

Figura 1.1 – Ilustração a nível qualitativo de uma partida no jogo da estratégia de agregação.	23
Figura 2.2 – Ilustração do jogo da estratégia de agregação a nível quantitativo.	30
Figura 2.3 – Visão geral da estratégia de agregação proposta.	32
Figura 2.4 – Etapas da atualização do hiperparâmetro em tempo real para o poço p e passo t	40
Figura 2.5 – Casco convexo com $p = 2$	45
Figura 3.6 – Vista oblíqua da realização 01 do reservatório sintético OLYMPUS.	48
Figura 3.7 – Vista superior da realização 01 do reservatório sintético OLYMPUS.	49
Figura 3.8 – Algumas realizações do reservatório sintético OLYMPUS.	49
Figura 3.9 – (a) Malha 3D da permeabilidade ao longo do eixo x do modelo EGG base. (b) Distribuição espacial dos poços produtores e injetores do EGG.	51
Figura 3.10–Realizações 01 a 10 do reservatório sintético EGG.	51
Figura 4.11–Faixas de previsão determinadas para o (a) EGG e (b) OLYMPUS.	54
Figura 4.12–Sequência de operações para se obter o erro percentual de um poço.	55
Figura 4.13–Previsões do óleo feitas pela AMA do 1º ao 5º ano do EGG	56
Figura 4.14–Comparação de previsão de produção de óleo do 1º ao 5º ano do EGG.	56
Figura 4.15–Comparação de previsão de produção de água do 1º ao 5º ano do EGG.	57
Figura 4.16–Realizações do OLYMPUS e VPLs associados.	59
Figura 4.17–Previsões do óleo feitas por todos os métodos para 5 anos de vida do OLYMPUS.	60
Figura 4.18–Previsões da água feitas por todos os métodos para 5 anos de vida do OLYMPUS.	61
Figura A.19–Previsões para o óleo do PROD-2 no OLYMPUS A	80
Figura A.20–Previsões para a água do PROD-6 no OLYMPUS B	80
Figura A.21–Previsões para a água do PROD-2 no OLYMPUS F	81
Figura A.22–Previsões para a água do PROD-3 no OLYMPUS F	81

Lista de tabelas

Tabela 1.1 – Principais aspectos de alguns trabalhos de Agregação de Modelos encontrados na literatura	26
Tabela 2.2 – Comparação estrutural de diferentes funções objetivo.	34
Tabela 3.3 – Tipos e intervalos de distribuição de propriedades das fácies nos modelos do OLYMPUS.	47
Tabela 3.4 – Propriedades dos fluidos nos modelos do OLYMPUS.	47
Tabela 3.5 – Algumas propriedades de geometria e do fluido nos modelos do EGG.	52
Tabela 3.6 – Comparação entre EGG e o OLYMPUS.	52
Tabela 4.7 – Erros na previsão de produção do 1° ao 5° ano do EGG.	58
Tabela 4.8 – Erros da AMA no EGG para diferentes ajustes em v_{OPR}/v_{WPR}	58
Tabela 4.9 – Erros de previsão para o OLYMPUS A (AMA e MPE).	63
Tabela 4.10–Erros de previsão para o OLYMPUS A (Lasso e Ridge).	63
Tabela 4.11–Erros de previsão para o OLYMPUS B (AMA e MPE)	64
Tabela 4.12–Erros de previsão para o OLYMPUS B (Lasso e Ridge).	64
Tabela 4.13–Erros de previsão para o OLYMPUS C (AMA e MPE).	65
Tabela 4.14–Erros de previsão para o OLYMPUS C (Lasso e Ridge).	65
Tabela 4.15–Erros de previsão para o OLYMPUS D (AMA e MPE).	66
Tabela 4.16–Erros de previsão para o OLYMPUS D (Lasso e Ridge).	66
Tabela 4.17–Erros de previsão para o OLYMPUS E (AMA e MPE).	67
Tabela 4.18–Erros de previsão para o OLYMPUS E (Lasso e Ridge).	67
Tabela 4.19–Erros de previsão para o OLYMPUS F (AMA e MPE).	68
Tabela 4.20–Erros de previsão para o OLYMPUS F (Lasso e Ridge).	68
Tabela B.21–Erros de previsão para o OLYMPUS A (AMA e MPE).	82
Tabela B.22–Erros de previsão para o OLYMPUS A (Lasso e Ridge).	83
Tabela B.23–Erros de previsão para o OLYMPUS B (AMA e MPE)	84
Tabela B.24–Erros de previsão para o OLYMPUS B (Lasso e Ridge).	84
Tabela B.25–Erros de previsão para o OLYMPUS C (AMA e MPE).	85
Tabela B.26–Erros de previsão para o OLYMPUS C (Lasso e Ridge).	85
Tabela B.27–Erros de previsão para o OLYMPUS D (AMA e MPE).	86
Tabela B.28–Erros de previsão para o OLYMPUS D (Lasso e Ridge).	86
Tabela B.29–Erros de previsão para o OLYMPUS E (AMA e MPE).	87
Tabela B.30–Erros de previsão para o OLYMPUS E (Lasso e Ridge).	87
Tabela B.31–Erros de previsão para o OLYMPUS F (AMA e MPE).	88
Tabela B.32–Erros de previsão para o OLYMPUS F (Lasso e Ridge).	88
Tabela C.33–Erros de previsão da AMA para o OLYMPUS A: 45:8 e 1:1.	89

Tabela C.34–Erros de previsão da AMA para o OLYMPUS B: 45:8 e 1:1	89
Tabela C.35–Erros de previsão da AMA para o OLYMPUS C: 45:8 e 1:1.	90
Tabela C.36–Erros de previsão da AMA para o OLYMPUS D: 45:8 e 1:1.	90
Tabela C.37–Erros de previsão da AMA para o OLYMPUS E: 45:8 e 1:1.	91
Tabela C.38–Erros de previsão da AMA para o OLYMPUS F: 45:8 e 1:1.	91

Lista de abreviaturas e siglas

AMA	Agregação de Modelos Adaptativa	26
AS	Análise de Sensibilidade	17
ASIPP	Abordagem de Sistemas Integrados aplicada à Produção de Petróleo	46
EDF	Electricité de France	24
EPC	Expansão Polinomial de Caos	20
FDA	Função Distribuição Acumulada	59
FKA	Filtro de Kalman por Agregação	16
FMI	Fundo Monetário Internacional	24
GE	Gradiente Exponenciado	23
IA	Inteligência Artificial	20
KG	Kriging	20
LB	Líquido-Bruto	47
MI	Modelo Intermediário	19
MII	Modelo Intermediário Inteligente	20
MIT	Modelo Intermediário Tradicional	20
MME	Média Móvel Exponencial	27
MMF	Modelo Multi-Fidelidade	20
MOR	Modelo de Ordem Reduzida	20
MPE	Média Ponderada Exponencial	23
MQO	Mínimos Quadrados Ordinário	24
MQR	Mínimos Quadrados Recursivo	24
PFP	Pressão de Fundo de Poço	24
PID	Proporcional-Integral-Derivativo	36
RP	Regressão Polinomial	20
RVS	Regressão do Vetor Suporte	20
TPA	Taxa de Produção de Água	24
TPO	Taxa de Produção de Óleo	24

Sumário

1	Introdução	15
1.1	Contextualização	15
1.2	Ajuste de histórico	16
1.2.1	Modelo Susstituto ou Intermediário	19
1.3	Agregação de Modelos	22
1.4	Proposta do trabalho	26
1.5	Objetivo geral e objetivos específicos	27
1.6	Organização do trabalho	27
2	Regressão em tempo real	29
2.1	Agregação de Modelos Adaptativa	31
2.1.1	Integração dos algoritmos	32
2.1.2	Filtro Adaptativo	34
2.2	Alguns métodos de agregação da literatura	38
2.2.1	Ridge	41
2.2.2	Lasso	43
2.2.3	Média Ponderada Exponencial	44
3	Reservatórios sintéticos	46
3.1	OLYMPUS	46
3.1.1	Caracterização	46
3.1.2	Realizações	48
3.2	EGG	49
3.2.1	Caracterização	49
3.2.2	Realizações	50
4	Resultados e análise	53
4.1	Resultados referentes ao EGG	55
4.2	Resultados referentes ao OLYMPUS	59
5	Conclusões e trabalhos futuros	69
5.1	Conclusões	69
5.2	Trabalhos futuros	69
	Referências	71

Anexos	79
Anexo A Comportamento anômalo no OLYMPUS	80
Anexo B Erros relativos no OLYMPUS com limiar 1	82
Anexo C Resultados da AMA para diferentes razões v_{OPR}/v_{WPR}	89

1 Introdução

1.1 Contextualização

Nas últimas décadas, o investimento em pesquisa e desenvolvimento por parte da indústria petrolífera concentrou-se em aprofundar a compreensão das características estruturais e geológicas dos mais diversos tipos de reservatórios, bem como propor estratégias de otimização da produção que proporcionem retorno econômico que viabilize a operação, como detalham [Khor, Elkamel e Shah \(2017\)](#) e [Mirzaei-Paiaman, Santos e Schiozer \(2021\)](#). Essas estratégias frequentemente requerem conhecimento da produção futura do reservatório. Naturalmente, esse conhecimento do futuro não é possível de ser obtido no presente, e os melhores esforços da pesquisa estão em aprimorar algoritmos e tecnologias que estimem os valores de interesse. Esta estimativa é de maneira geral feita por simulações numéricas ou físicas.

O potencial econômico do uso de simulações físicas de reservatórios tem sido explorado na indústria petrolífera de maneira histórica ([OLIVER; CHEN, 2011](#)). De maneira geral, primeiro aproxima-se o reservatório de interesse a uma ou múltiplas malhas tridimensionais através de estudos geológicos e sísmicos, cujos componentes, denominadas de células, armazenam informações locais relativas a porosidade, pressão, concentração de petróleo, água e gás, viscosidade relativa, presença de poço, etc; em seguida, utilizam-se equações parciais da dinâmica dos fluidos para atualizar tais informações iterativamente a cada passo de tempo. Os resultados ao final da simulação correspondem a perfis espaciais e temporais da produção, os quais podem ser interpretados por engenheiros de reservatório a fim de propor uma estratégia de produção ou otimizar uma estratégia existente, e.g., a introdução de um novo poço injetor, ou o fechamento temporário de um produtor.

Dentro desse contexto, é essencial construir malhas suficientemente próximas ao reservatório em estudo. Para isso, uma das estratégias mais bem consolidadas na literatura calibra os parâmetros locais da malha (pelo menos aqueles passíveis de ajuste) de acordo com o comportamento passado observado do reservatório. Denominada de Ajuste de Histórico (do inglês *History Matching*), esta técnica é popular na literatura de previsão de produção de reservatórios. É escrita a seguir uma breve revisão de trabalhos da literatura relacionados ao ajuste de histórico de reservatórios. Findada a discussão, a próxima seção começa a detalhar a técnica de ajuste como tal.

[Sayyoub \(1981\)](#) e [Ghoniem et al. \(1984\)](#) utilizaram a técnica de programação de metas (do inglês *goal programming*) e o método de Programação Linear de Mínimos Quadrados, respectivamente, em seus estudos para melhorar a correspondência de dados históricos.

Conforme declarado por [Oliver e Chen \(2011\)](#), no início dos anos 2000, grande parte dos esforços na correspondência de histórico está associada à aplicação do Filtro de Kalman por Agregação (FKA, do inglês *Ensemble Kalman Filter*) ([LORENTZEN et al., 2001](#)), cujo desenvolvimento fora revisado anteriormente por [Aanonsen et al. \(2009\)](#). [Emerick e Reynolds \(2012\)](#) introduziram o suavizador de agregado com múltiplas assimilações de dados para auxiliar na correspondência de histórico, fornecendo tanto uma correspondência de dados aprimorada quanto uma melhor quantificação de incerteza em comparação com o FKA. Uma versão adaptativa desse suavizador é proposta por [Le, Emerick e Reynolds \(2016\)](#) para evitar valores extremos nos campos de propriedades de rochas em problemas complexos.

Uma abordagem alternativa para aprimorar a correspondência de histórico em reservatórios de petróleo envolve a aplicação de algoritmos bioinspirados ou algoritmos evolutivos. [Riazi et al. \(2016\)](#) discutiu o uso de um processo automático de correspondência de histórico de reservatório usando tanto a Otimização por Enxame de Partículas quanto o Algoritmo Competitivo Imperialista. Da mesma forma, [Chakra e Saraf \(2016\)](#) empregou um Algoritmo Genético Adaptativo no contexto da caracterização de reservatórios. Finalmente, vale comentar que, mais recentemente, alguns algoritmos de aprendizado de máquina também podem ser usados para melhorar a correspondência de histórico em reservatórios de petróleo, como o fazem [Razak e Jafarpour \(2020\)](#), [Lacerda, Emerick e Pires \(2021\)](#) e [Alguliyev et al. \(2022\)](#).

1.2 Ajuste de histórico

[Oliver e Chen \(2011\)](#) e [Ma \(2019\)](#) apresentam o ajuste de histórico como um tipo de problema de inversão no qual medidas observadas do comportamento passado do reservatório são utilizadas para estimar as variáveis no modelo do reservatório que causaram tal comportamento. Quando bem sucedido, o modelo tende a representar o reservatório de maneira fiel dentro do intervalo onde foi ajustado, o que o torna potencial candidato a fornecer uma previsão da produção. Por sua vez, esta previsão pode abalizar as futuras estratégias de produção que agirão sobre o reservatório em análise. Eventualmente, para acelerar a simulação, pode-se introduzir no processo uma aproximação do modelo.

De modo geral, o ajuste de histórico é constituído por quatro etapas consecutivas: parametrização, definição da função objetivo, maximização da função objetivo, quantificação de incerteza. A seguir detalha-se cada etapa, e, feito isso, apresenta-se a técnica do Modelo Substituto, cuja popularidade muito cresceu nas últimas décadas devido ao alívio de carga computacional que contribui ao ajuste.

Parametrização. Em preparação para simulação, cada célula do modelo deve ser imbuída por diversos parâmetros, como permeabilidades vertical e horizontal, e porosidade. Existem ainda outras propriedades do reservatório que não são necessariamente entradas da simulação, mas auxiliam a estimação da distribuição de, por exemplo, permeabilidade

pelo reservatório: pode-se citar barreiras de xisto e a posição de canais altamente permeáveis. Dada a complexidade do reservatório e o elevado número de células que lhe é atribuído, essa etapa busca reparametrizar, ou projetar, o modelo para um subespaço de ordem reduzida em relação ao espaço de todos os parâmetros. Uma reparametrização razoável irá reduzir o número de variáveis sem deixar de levar em conta aquelas propriedades às quais o reservatório reage de maneira relativamente intensa. À sistematização dessa seleção de parâmetros, dá-se o nome de Análise de Sensibilidade (AS). Embora nem todo ajuste realize a projeção a um subespaço menor, a AS é uma ferramenta essencial para lidar com grandes reservatórios.

Definição da função objetivo. Estabelecidas as variáveis sobre os quais o ajuste de histórico atuará, define-se agora a relação pela qual a qualidade do ajuste será aferida. O modelo é iterativamente ajustado enquanto a qualidade não chegar à um patamar desejado mediante um critério pré-estabelecido. Nesta etapa, define-se este critério em função do erro entre as medidas observadas no reservatório e a saída do modelo numérico que buscam aproximar essas observações. A saída é gerada com base no vetor de variáveis do modelo, que varia conforme o modelo é ajustado. Ademais, diz-se função objetivo porque é sobre ela que serão feitas as estratégias de otimização para minimizar esse erro. No caso particular onde o modelo lida com uma quantidade de dados maior que a quantidade de variáveis, a função objetivo se reduz a uma norma da diferença entre o vetor de observações e o vetor de saída do modelo. No caso geral, além do vetor diferença entre as observações e a saída, a função objetivo dependerá também da diferença entre as variáveis do modelo atual e do anterior. Esses termos são ponderados por matrizes de covariância do ruído nas medições e dos parâmetros do modelo, respectivamente (OLIVER; CHEN, 2011, seção 3.2).

Minimização da função objetivo. Seja na forma mais simples ou na mais complexa, a função objetivo apresenta tanto mínimos locais quanto mínimos globais. A forma pela qual uma estratégia de ajuste de histórico atinge um mínimo, e se esse mínimo é local ou global, é foco de estudo de diversos trabalhos até mesmo fora do contexto de reservatório. Bahrami, Moghaddam e Lesley A James (2022) e Piyang Liu, Kai Zhang e Yao (2023) classificam os algoritmos de otimização em dois grupos: os baseados em gradiente, de abordagem determinística, e os que não são baseados em gradiente, de abordagem estocástica. Nos métodos de gradiente, destacam-se os algoritmos do tipo Newton, o método da maior queda (*steepest descent*), e o método do gradiente conjugado. Todos os algoritmos de gradiente necessitam do cálculo do gradiente da função objetivo (ou pelo menos uma aproximação deste), sendo que os do tipo Newton requerem ainda a Hessiana da função. Métodos nesse grupo são eficientes, mas aderem apenas a mínimos locais e tendem a apresentar elevado custo computacional quando aumenta-se o tamanho do modelo. Por outro lado, algoritmos estocásticos são capazes de encontrar o mínimo global, e, embora não compartilhem uma metodologia comum, tendem a apresentar convergência mais lenta. Nesta classe destacam-se os algoritmos Monte Carlo da Corrente de Markov, de busca heurística, e de assimilação de

dados, como suavizador de agregado (*ensemble smoother*) e FKA. [Bahrami, Moghaddam e Lesley A James \(2022\)](#) apontam e detalham outros algoritmos padrões para a obtenção de mínimos globais, como recozimento simulado, e evolução diferencial. Note que o FKA pode ser adaptado para os algoritmos baseados em gradiente, como mostram [Oliver e Chen \(2011, equação 34\)](#).

Quantificação de incerteza. Feita a minimização da função objetivo, o modelo ajustado resultante não será necessariamente leal ao reservatório real no que tange o comportamento futuro deste. Esta etapa busca resolver ou atenuar o problema de sobreajuste (do inglês, *overfitting*) inerente. Para tanto, são refeitos os passos anteriores diversas vezes, a fim de obter múltiplas malhas que satisfazem a minimização da função objetivo, porém distintas entre si. Deste conjunto, é escolhido um seleto grupo de modelos, os quais serão utilizados para estimar o erro nas previsões feitas pelas malhas.

Elencadas os passos gerais num ajuste de histórico, é útil ver como eles são feitos na prática. [Ning Liu e Oliver \(2005\)](#) propuseram a realização do ajuste de taxas de produção e injeção dos poços produtores e injetores (respectivamente) não através da modificação de propriedades físicas das células, como esperado num ajuste mais ortodoxo, mas sim através do posicionamento geológico/geométrico das fácies do reservatório. Na ocasião da publicação do artigo, ferramentas tradicionais para ajuste de histórico automático (i. e., sem a necessidade de um engenheiro de petróleo ou qualificado similar modificar os parâmetros durante a etapa de minimização da função objetivo) de reservatórios com fronteiras de fácies desconhecidas não eram comuns, dados o alto custo computacional exigido e a fraca diferenciabilidade entre tipos de fácies. Por ser fácil de introduzir em um simulador de reservatório, possuir custo computacional relativamente baixo (uma iteração de simulação por realização, gradiente calculado implicitamente) e utilizar matriz de correlação entre variáveis do modelo, a qual aproxima-se à uma diferenciabilidade entre fácies, o FKA foi a técnica escolhida para atualizar a distribuição das fácies. Por outro lado, a fácies enquanto variável não é nem diferenciável nem gaussiana, e portanto, como o FKA assume entradas gaussianas, o filtro não seria a ferramenta apropriada para o ajuste. Todavia, o estudo deixou claro, em análise dos ajustes obtidos sobre dois modelos distintos de reservatórios, que a escolha astuta das variáveis de estado do FKA é capaz de gerar resultados satisfatórios, e. g., redução média de 16% no erro médio quadrático das taxas antes e depois do ajuste. Os campos gaussianos que geravam os mapas de fácies foram tomadas como variáveis de estado do filtro. Em suma, o problema de ajuste de fácies pode aproveitar as vantagens do FKA sem sofrer com suas desvantagens, desde que implementado com alguma sofisticação e conhecimento técnico do problema.

Apesar da popularidade de métodos de ajuste de histórico e seus reconhecidos benefícios na gerência de reservatório, o tamanho da complexidade computacional permanece um ponto de preocupação. Perceba que no trabalho de [Ning Liu e Oliver \(2005\)](#), a etapa de

quantificação e análise de incerteza fora muito mais privilegiada que, por exemplo, a etapa de apresentação da função objetivo, que na verdade não fora dada explicitamente. Poderia-se perguntar como o custo computacional do mesmo problema poderia ser reduzido caso não fosse utilizado o FKA, ou, de maneira mais incisiva, que tecnologia moderna seria capaz de cumprir a função daquelas “ferramentas tradicionais de ajuste de histórico” indisponíveis na época do artigo.

Uma primeira abordagem para mitigar essa complexidade é utilizar modelos simplificados de reservatórios, como exemplificado em [Heidari et al. \(2013\)](#). No entanto, esses modelos simplificados frequentemente carecem de realismo e não são implementados na prática. Uma estratégia alternativa para lidar com os custos computacionais em reservatórios de petróleo reais envolve a criação de modelos substitutos (também conhecidos como modelos de proxy) que, embora ainda substanciais em tamanho, são menos intrincados, como demonstrado no trabalho de [Bahrami, Moghaddam e Lesley A James \(2022\)](#). A estratégia do Modelo Substituto é explorada a seguir.

1.2.1 Modelo Substituto ou Intermediário

As etapas relacionadas ao ajuste de histórico estão naturalmente atreladas a altas cargas computacionais. Graças à evolução dos processadores de ponta, o método numérico denominado de Modelo Substituto (*surrogate*) ou Intermediário (*proxy*) vem sendo utilizado na literatura como uma alternativa para atenuar os custos computacionais em aplicações de diferentes áreas da engenharia.

Dentro do contexto de engenharia de reservatórios ([BAHRAMI; MOGHADDAM; JAMES, L. A., 2022](#)), a ideia central consiste em introduzir um modelo reduzido em paralelo ao modelo principal do reservatório alvo. Também conhecido como modelo intermediário (MI), o modelo reduzido consiste em uma sofisticada tabela de interpolações não lineares que aproximam uma malha de alta fidelidade, isto é, uma malha com grande número de células. Todo MI é construído com o auxílio de uma malha fiel. A grande vantagem desta técnica é que, uma vez acabada a construção do MI, são precisos apenas poucos segundos para executá-lo, não obstante alguma queda na qualidade da previsão ([BAHRAMI; MOGHADDAM; JAMES, L. A., 2022](#)).

Pondo de maneira genérica, a relação malha-modelo se resume no seguinte conjunto de ações: uma pequena amostra dos parâmetros de entrada é escolhida; a malha fiel é executada dentro do espaço definido pela amostra, gerando saídas; finalmente, o MI ajusta esses dados. A princípio, o intermediário será válido apenas no intervalo da amostragem escolhida. Este conjunto de ações esconde, entretanto, a complexidade do processo. Em particular, quais parâmetros devem ser tomados, e como eles devem ser amostrados, são dois problemas distintos e não triviais.

O primeiro é decidido através do objetivo do MI. Por exemplo, taxas de injeção e de produção, viscosidade da água, entre outras grandezas, podem ser os parâmetros se o objetivo é otimizar o corte de água (do inglês, *water cut*) durante a injeção contínua de água (do inglês, *waterflooding*).

Já em relação ao segundo, tem-se que, em síntese, a amostragem lida com o problema de maximizar a quantidade de informação com o menor número de pontos (ou de execuções da malha fiel). Normalmente, as estratégias de amostragem, cujos detalhamentos acabam se enquadrando fora do escopo do presente manuscrito, são classificadas em dois grupos: estacionária, aquelas que realizam a amostragem de um conjunto imutável de pontos; e sequencial, aquelas que realizam a amostragem incrementando o conjunto inicial de pontos, caso o MI não tenha gerado resultados precisos o suficiente com o conjunto anterior. Mais importante, um MI também é passível de classificação. [Bahrami, Moghaddam e Lesley A James \(2022\)](#) buscam estabelecer um sistema de classificação completo e coeso, até então ausente na literatura, na forma de quatro divisões, conforme a maneira pela qual o MI ajusta a saída da malha: Modelo de Ordem Reduzida (MOR), Modelo Multi-Fidelidade (MMF), Modelo Intermediário Tradicional (MIT), e Modelo Intermediário Inteligente (MII).

No MOR, a malha de alta fidelidade é projetada num sistema com poucas dimensões e características similares com menores graus de liberdade. Para tanto, é feita uma AS anterior à projeção de quais parâmetros devem ser mantidos e quais devem ser omitidos após a projeção. Note que o processo aqui é paralelo à etapa de parametrização no ajuste de histórico. O MFM por sua vez é uma simplificação física da malha fiel, através, por exemplo, da diminuição na resolução da discretização da malha. [Wilson e Durlinsky \(2012\)](#) foram mais extremos, substituindo reservatório com duas porosidades e permeabilidades a um de única porosidade e permeabilidade.

O MIT e MII, no lugar de lidar diretamente com a dinâmica subjacente da malha, usam as suas saídas para treinar ou validar modelos de inteligência artificial. Os modelos de IA mais comuns neste contexto são a Regressão Polinomial (RP), Kriging (KG) ou regressão por processo de Gauss, Regressão do Vetor Suporte (RVS), e Expansão Polinomial de Caos (EPC). Uma vez que estas classes são construídas mais rápida e facilmente que o MOR e MFM, elas se tornam as preferidas nos trabalhos de engenharia de reservatório. A diferença entre o MIT e o MII é que, enquanto o primeiro supre o modelo de Inteligência Artificial (IA) diretamente após a execução da malha fiel, o último treina seu modelo com um conjunto de dados formado pela combinação de entradas e saídas de múltiplas realizações da malha. Dessa forma, o MII se torna apto a encontrar padrões de comportamento que não são visíveis ao MIT, gerando resultados mais precisos a longo tempo.

A princípio, o MII é superior ao MIT, e esses dois têm melhor custo benefício que o MOR e MFM. Recentemente, por exemplo, [Zhen Zhang et al. \(2023\)](#) propõem uma nova metodologia para ajuste de histórico usando MII no contexto de engenharia de reservatórios.

A grosso modo, o ajuste consistia no seguinte: primeiro, parâmetros são identificados e fornecidos ao MII; o MII é construído usando estratégia estática de amostragem; os resultados do intermediário são balizados por simulações de modelos de alta fidelidade, de forma que o ajuste continua se não houver discrepância significativa, e o MI é refeito se houver; os resultados do intermediário são balizados por dados reais, de forma que o ajuste conclui caso não haja discrepância significativa, e o MI é refeito se houver, dessa vez com parâmetros de incerteza alterados. Foram considerados 100 passos de tempo e 11 parâmetros de entrada, como viscosidade do óleo, saturações de água e óleo, e curvas de permeabilidade relativa. A fim de otimizar o ajuste, AS é realizada pervasivamente durante a construção do MI. Destacam-se aquelas sobre o conjunto de treino (tudo o mais constante, amostras foram variadas entre 100 e 500) e a arquitetura interna do modelo, ou seja, por quantas e quais etapas principais passa o vetor de entrada para gerar o vetor de saída (quatro arquiteturas foram consideradas, com 4 a 6 etapas). Fora do MI, foi feita AS, por exemplo, em métodos estocásticos que processam a entrada e saída da malha fiel. A metodologia de ajuste proposto economizou 98% do tempo de computação necessário a ajustes que utilizavam MII já estabelecidos na literatura, tais como EPC, Kriging e RVS.

Na última década, como alternativas para reduzir os custos computacionais associados aos métodos previsão via ajuste de histórico, novas abordagens que tratam do problema de previsão de reservatórios sem atualização de modelo têm surgido na literatura relacionada, e.g., os trabalhos de [Satija e Caers \(2015\)](#), [Scheidt, Renard e Caers \(2015\)](#), [Satija, Scheidt et al. \(2017\)](#) e [Sun e Durlinsky \(2017\)](#). Mais precisamente, um conjunto de modelos de reservatórios é usado para estabelecer uma relação estatística entre as variáveis de histórico e de previsão. A relação estimada é então usada para obter uma previsão estatística com base em dados reais de produção observados. Isso permite quantificar a incerteza posterior sobre a variável de previsão sem inversão explícita ou ajuste de histórico.

Métodos de agregação de modelos também vem se mostrando úteis para fornecer previsões de produção com base na análise do histórico de produção sem atualização do modelo. Para isso, o sistema físico que produz um evento de interesse é aproximado por múltiplos modelos sintéticos, que serão ponderados para gerar a previsão do evento, aqui denominada previsão agregada. É lógico que a previsão será tão melhor quanto melhor for a ponderação, o que leva à inspeção dos dados dos modelos em relação às observações passadas do evento, dando um quê de ajuste de histórico ao processo. Todavia, em comparação com o Ajuste, a Agregação frequentemente carece de uma interpretação física, mas experimenta uma carga computacional muito menor, e já foi aplicada em várias áreas além de reservatórios de petróleo. Na próxima seção e capítulo essa nova técnica é detalhadamente explicada.

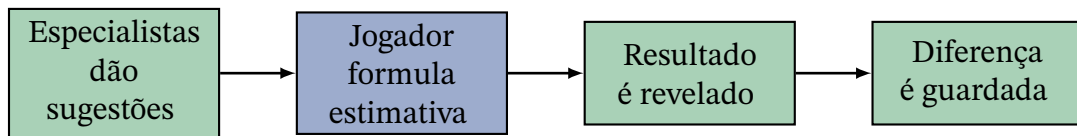
1.3 Agregação de Modelos

O método de agregação é uma maneira de se prever acontecimentos futuros de um sistema tendo em mãos tanto observações passadas do acontecimento quanto sugestões de previsão, ditas especialistas (do inglês *experts*) na literatura (CESA-BIANCHI; FREUND et al., 1997). Os especialistas são entidades que sugerem valores para o acontecimento a ser previsto, mas cujos mecanismos de funcionamento são a priori desconhecidos do ponto de vista matemático. A nível de implementação um conjunto de especialistas são formados, por exemplo, ao variar um ou mais parâmetros de um modelo sintético do sistema. Dessa forma, na prática, não é necessariamente verdade que os especialistas são desconhecidos, mas como deles serão observados apenas as saídas durante a execução do método de agregação, às vezes são ditos caixas pretas. Devaine et al. (2013) em particular usam dois conjuntos de especialistas, sendo o primeiro (seção 4) caixa preta e o segundo (seção 5) construído de maneira transparente. Vale comentar ainda que, no contexto de reservatórios, os especialistas são mais comumente chamados de realizações.

A proposta de previsão de acontecimentos futuros baseada no método de agregação fica bem exemplificada num contexto de jogo iterativo entre o ambiente ou sistema e o jogador. A cada rodada, ou passo de tempo, um jogador recebe as sugestões dos especialistas, e é desafiado pelo ambiente a prever o acontecimento na forma de um resultado numérico. O jogador formula sua previsão, dita previsão agregada, com base nas sugestões dadas e nos resultados e informações de rodadas passadas, e, em seguida, o ambiente revela o resultado. Se o jogo está na primeira rodada, o jogador baseia-se apenas nos especialistas. O jogador guarda a diferença entre o esperado e o ocorrido para uso nas rodadas futuras. A rodada presente termina e a próxima se inicia, continuando o jogo indefinidamente.

Naturalmente, pode-se impor que a partida termine com a previsão do jogador, de modo que a próxima se inicie com a quantificação do erro. A determinação da divisão adequada entre partidas se dará a nível de implementação, sendo ambos as interpretações suficientes para uma explicação qualitativa. A forma como o jogador usa as informações passadas junto com as sugestões presentes é conhecido na literatura de problemas de regressão em tempo real, e é o que distingue, em essência, os métodos de agregação entre si (CESA-BIANCHI, 1999; KIVINEN; WARMUTH, 1997). Esses problemas são respondidos por uma regra de agregação, que realizará alguma forma de regressão sobre as revelações do sistema com os dados dos especialistas para gerar a previsão do jogador. A Figura 1.1 apresenta diagrama que resume o protocolo da partida descrita anteriormente.

Figura 1.1 – Ilustração a nível qualitativo de uma partida no jogo da estratégia de agregação.



Fonte: Autor.

A Agregação pode ser aplicada a qualquer problema de previsão cujo algoritmo de regressão em tempo real esteja bem definido e sistema admita algum número de especialistas. Esta flexibilidade de aplicação é uma característica intrínseca marcante da estratégia de agregação, e é razoável vê-la como vantagem sobre o ajuste de histórico, que é sensível à mudança de reservatório sendo analisado. Por exemplo, simplificações da malha motivadas por uma AS levam em conta particularidades do respectivo reservatório, e por isso deve ser refeita no ajuste de um novo reservatório. Não menos importante, o problema central da Agregação, isto é, como usar e casar de maneira eficaz os especialistas e as informações de partidas passadas, é muito mais fácil de encarar que as não-linearidades e inversões do ajuste de histórico. Para mais informações sobre o método de agregação, convida-se o leitor a buscar [Cesa-Bianchi e Lugosi \(2006\)](#). A seguir, apresenta-se alguns trabalhos da literatura que aplicaram a estratégia de agregação em diversos contextos de engenharia. Essa apresentação é resumida na [Tabela 1.1](#)

[Mallet, Stoltz e Mauricette \(2009\)](#) comparam métodos de agregação para obter previsões dos perfis (i) de concentração de ozônio, de hora em hora, e (ii) do valor de pico da concentração de ozônio, dia a dia, referentes à atmosfera de pontos na Europa ocidental. Mais especificamente, as simulações são feitas sobre o verão de 2001. Nesse trabalho, utilizam-se principalmente duas regras de agregação distintas, o Ridge e o Gradiente Exponenciado, GE (do inglês *Exponentiated Gradient*), assim como variações sobre as mesmas, chamadas Janelamento e Desconto. Em relação ao protocolo da agregação, a primeira variação consiste em considerar apenas os dados mais recentes dos especialistas, enquanto que o segundo busca diminuir a importância das amostras antigas dos especialistas multiplicando suas saídas por um valor inversamente proporcional ao valor do passo de tempo. Os autores usam dois grupos de especialistas, criados com base no sistema de modelagem Polyphemus, principalmente um com 48 e outro com 51, uma expansão do agregado principal, a fim de avaliar a qualidade das previsões obtidas a partir do método de agregação proposto. Ao final do trabalho, o Ridge com desconto se demonstrou superior aos demais métodos, isto é, apresentou o menor erro RMS acumulado ao longo dos diferentes casos de teste.

[Devaine et al. \(2013\)](#) propuseram utilizar a Média Ponderada Exponencial (MPE), assim como outras variações de sua implementação, para prever o consumo de energia elétrica da Eslováquia e da França via Agregação. O método foi aplicado com referência a um conjunto de dados para cada país, fornecidos pelo órgão provedor de energia elétrica da

França, Electricité de France (EDF). Em particular, o conjunto francês foi fragmentado em duas partes. Um período dito conjunto de estimação, usado para projetar os especialistas, que por sua vez vão fornecer as previsões para o segundo período. Este, dito o período de validação, corresponde aos 320 dias considerados a efeito de agregação. Os resultados das simulações demonstraram que, para as mais bem sucedidas versões da MPE, existe um aumento de 5% (conjunto eslovaco) a 15% (conjunto francês) na precisão das previsões em relação à melhor combinação constante dos especialistas.

[Amat, Michalski e Stoltz \(2018\)](#) usam o método de agregação para prever as taxas de intercâmbio de moedas fortes no período de 1973-2014. Os dados foram fornecidos pelo Fundo Monetário Internacional (FMI). As previsões são feitas mês a mês, e as regras de agregação utilizadas foram a MPE e o Ridge Sequencial, ambas com Desconto. Apesar de serem notadamente semelhante em espírito e objetivo ao Desconto em [Mallet, Stoltz e Mauricette \(2009\)](#), as fórmulas dos fatores de desconto são distintas neste artigo. Ambas regras de agregação são validadas contra outras da literatura, a saber, Mínimos Quadrados Ordinário (MQO) e Mínimos Quadrados Recursivo (MQR), sob diferentes condições de simulação. Ao todo, a MPE com desconto e o Ridge Sequencial com Desconto mostraram-se superiores às demais formas de previsão.

Outra vantagem relevante do método de agregação de modelos em comparação ao ajuste de histórico, mesmo com o uso de MI, é a drástica queda de carga computacional, uma vez que são necessários aos algoritmos de regressão apenas os valores dados pelos especialistas e um algoritmo para usá-los na formação de uma estimativa no lugar de uma malha com múltiplas informações de cada célula. Por exemplo, as simulações mais lentas feitas para geração de resultados de manuscrito duraram cerca de 96 horas, enquanto [Zhen Zhang et al. \(2023\)](#) reportam tempo necessário superior a 150 horas, que por sua vez representa uma redução de 98% de métodos de ajuste comuns. O ganho de velocidade é especialmente útil para trabalhos de reservatório, como observado nos trabalhos a seguir.

[Deswarte et al. \(2019\)](#) observam os resultados das previsões de 3 algoritmos de agregação quanto à pressão de fundo de poço (PFP), à taxa de produção de água (TPA) e à taxa de produção de óleo (TPO) do reservatório sintético Brugge. A saber, os algoritmos são a MPE, o Ridge, e o Lasso. Numa primeira etapa, cada algoritmo gerou previsões sobre o reservatório em estudo indo de passo de tempo em passo de tempo, um a um, até chegar ao fim da vida útil do Brugge. Feito isso, usa-se o algoritmo não para prever no instante $t - 1$ o sistema em t , mas em $t + k$, com k possivelmente grande. Essa mudança de paradigma acarreta em complicações na implementação, a saber, na existência de um período de treino para decidir um subconjunto dos especialistas a ser utilizado no período de previsão e num desvio padrão sobre as medidas do reservatório em estudo. Para esse caso, foi mostrado ainda como gerar um envelope limitando as possíveis previsões. Para cada método de agregação, mostraram-se histogramas da raiz do erro médio quadrático entre a medição e a previsão

de quatro medidas: PFP nos produtores e injetores, e as produções de óleo e de água nos produtores. Os mesmos histogramas foram feitos para o modelo que sozinho fazia a melhor previsão das medidas citadas, assim como a combinação de modelos que juntos faziam a melhor previsão. Analisando os histogramas, o Lasso performou melhor (apresentou menor erro) que os dois métodos estáticos acima em cada histograma, o Ridge performou pior apenas nas produções de água e o MPE, apenas nas de pressões dos injetores.

Fortaleza, Filho e Munerato (2022) e Oliveira et al. (2022) propõem estratégias de agregação de modelos de previsão a curto (1 mês) e médio prazo (6 meses) da produção de óleo e gás. As estratégias consistem numa média ponderada dos dados das realizações, cujos pesos são obtidos através da chamada função objetivo. A diferença central entre os artigos está em como essa função é definida: enquanto a no segundo leva em conta apenas o erro médio acumulado pelas realizações, a no primeiro leva em conta também o erro acumulado da realização individual. Ambos os trabalhos validam a sua estratégia individual através de simulações numéricas sobre o reservatório sintético OLYMPUS, e ambos os trabalhos apresentaram resultados satisfatórios: em particular, Oliveira et al. (2022) mostram que o erro médio percentual absoluto da previsão em relação à produção observada de diversas realizações de referência não chega a 3%.

Majoritariamente, os 6 trabalhos apresentados acima fornecem abordagens sequenciais, onde a previsão agregada é feita para o próximo passo de tempo apenas. Nesse caso, considera-se neste manuscrito que a previsão é de curto prazo. Por outro lado, métodos que realizem previsão para além do próximo passo de tempo serão ditos de médio ou longo prazo, conforme a localização da previsão dentro do conjunto de dados disponível. Quando mais próximo ao fim, diz-se longo prazo; caso contrário, médio. Deswarte et al. (2019, seção 5) realizam previsões para os últimos três anos e meio de vida do Brugge, e por isso é considerado o único método de agregação a longo prazo daqueles citados aqui. A Tabela 1.1 exibe de maneira resumida as características principais dos trabalhos acima no que se refere à Agregação.

Tabela 1.1 – Principais aspectos de alguns trabalhos de Agregação de Modelos encontrados na literatura

Trabalho da literatura	Características				
	Objeto de previsão	Conjunto de especialistas	Algoritmos de Agregação	Prazo da previsão	Tempo de amostragem
Mallet, Stoltz e Mauricette (2009)	Concentração de ozônio	Criados via Polyphemus	Ridge, GE	Curto	1h
Devaine et al. (2013)	Consumo de energia	Com base no EDF	MPE	Curto	1 dia (SK) 30 min (FR)
Amat, Michalski e Stoltz (2018)	Taxas de câmbio	Com base no FMI	MPE, Ridge, MQO, MQR	Curto	1 mês
Deswarte et al. (2019)	TPO, TPA e PFP	Reservatório Brugge	MPE, Ridge e Lasso	Curto e longo	30 dias
Fortaleza, Filho e Munerato (2022)	TPO, TPA	Reservatório OLYMPUS	Baseado em gradiente	Curto e médio	30 dias
Oliveira et al. (2022)	TPO, TPA	Reservatório OLYMPUS	Baseado em gradiente	Curto e médio	30 dias

Fonte: Autor.

Nesse contexto de estratégia de agregação de modelos, os trabalhos mencionados anteriormente, sejam sobre reservatórios ou outros sistemas, não consideram meios de mitigar o impacto do ruído associado às medições fornecidas pelos sensores no processo de previsão. Portanto, dado as características particulares das plataformas de petróleo, vislumbra-se que desenvolver estratégias que mitiguem essa influência proporcionaria melhores previsões de produção, o que auxiliaria os engenheiros em determinar as melhores estratégias de otimização ao longo da vida útil do reservatório. Na próxima seção, a proposta deste trabalho que aborda esse aprimoramento a ser explorado na estratégia de agregação de modelos será melhor discutida.

1.4 Proposta do trabalho

Este manuscrito procura validar um novo método de agregação de modelos, nomeado Agregação de Modelos Adaptativa (AMA). A AMA baseia-se na integração da estratégia de agregação de modelos descritas em [Fortaleza, Filho e Munerato \(2022\)](#) e [Oliveira et al. \(2022\)](#). Nessa versão integrada, deve-se gerar, a cada passo de tempo, uma matriz de pesos dos poços produtores. Nela, cada coluna consiste no vetor de pesos de um dado produtor, cujos elementos são os pesos estipulados pelas realizações (especialistas) para aquele poço. Tal matriz é obtida ao minimizar, para cada poço, uma função custo que leva em conta tanto o erro acumulado de cada realização como o erro médio acumulado pelas realizações. Por consequência, espera-se que o erro da previsão agregada seja inferior ao simplesmente utilizar pesos iguais para as realizações disponíveis.

Obtida a matriz de pesos, a AMA utiliza uma abordagem baseada na Média Móvel Exponencial (MME) para reduzir o impacto do ruído de medição nos dados observados do reservatório. Para isso, a cada passo de tempo, filtram-se os elementos da matriz de pesos através da MME, cujo nível de suavização é variável e depende do quanto a previsão fornecida pelo modelo mais recente é estatisticamente relevante em relação ao ruído de medição conhecido.

1.5 Objetivo geral e objetivos específicos

O objetivo principal deste manuscrito centra-se em desenvolver uma metodologia de agregação de modelos para um reservatório a fim de gerar previsões de produção a partir de dados sujeitos a ruído de medição. Para atingir esse objetivo geral, propõem-se alcançar os seguintes objetivos específicos:

1. Apresentar o método de agregação como alternativa ao ajuste de histórico.
2. Apresentar matematicamente os métodos de agregação escolhidos da literatura para a comparação, assim como o da minimização com filtro adaptativo.
3. Apresentar os reservatórios sintéticos sobre os quais se debruçarão os métodos de agregação.
4. Partindo de simulações numéricas, validar por análise comparativa de erros acumulados o método de agregação AMA como superior ante os outros métodos escolhidos.

1.6 Organização do trabalho

Aproximadamente, cada capítulo representa um passo da lista na [seção 1.5](#). Neste capítulo, o método de agregação de modelos foi apresentado como solução frente à necessidade de previsão de produção de reservatório, em detrimento do ajuste de histórico.

No [Capítulo 2](#), essa estratégia é detalhada em duas frentes, primeiro qualitativamente, abordando a lógica e protocolo do método, e depois matematicamente, estabelecendo o conceito de regressão linear e dando norte à notação que será usada pelo resto do texto. Primeiro, é estruturada a forma e simbologia comum que irão seguir os métodos de agregação quando aplicados à previsão de reservatório, para na [seção 2.2](#) e [seção 2.1](#) serem apresentadas as particularidades de cada método.

O [Capítulo 3](#) detalha os reservatórios sintéticos, EGG e OLYMPUS, sobre os quais serão feitas as simulações numéricas de cada método de agregação. Os resultados associados são apresentados e discutidos no [Capítulo 4](#).

Finalmente, o [Capítulo 5](#) faz um arremate total dos pontos mais importantes observados durante o desenvolvimento da proposta conforme apresentado no restante do documento, finalizando com estipulações de ideias para trabalhos futuros com base neste trabalho atual.

2 Regressão em tempo real

Omitiu-se até aqui aspectos quantitativos da estratégia de agregação, ou seja, a regra de agregação e como é feita a regressão dentro da regra. Os aspectos quantitativos, i.e., matemáticos, começarão a ser considerados agora, em duas etapas. Primeiro, será estruturada uma versão mais simples da regressão, que estima uma única variável a cada passo de tempo, para depois disso ampliar a estrutura numa versão da regressão capaz de estimar múltiplas variáveis a cada passo de tempo, visando melhor se encaixar no contexto de reservatórios.

Ainda no contexto do jogo descrito no capítulo anterior, tem-se que o ambiente gera uma sequência $(y_i \in \mathcal{D})_{i \in \mathbb{N}}$ de elementos que devem ser estimados pelo jogador através do auxílio de N especialistas e das informações de partidas passadas. \mathcal{D} é um subconjunto de \mathbb{R} . Denote por $m_{t,i}$, $1 \leq i \leq N$, a previsão fornecida pelo especialista j para o passo de tempo $t \geq 0$, e sejam $w_{t,i}$ números associados a $m_{t,i}$ através de alguma metodologia conhecida.

Para compor $\hat{y}_t \in \mathcal{D}$, a previsão de y_t , o jogador usará o vetor de estimativas $\mathbf{m}_t = (m_{t,1}, m_{t,2}, \dots, m_{t,N})^T$ assim como o vetor $\mathbf{w}_t = (w_{t,1}, w_{t,2}, \dots, w_{t,N})^T$. A regressão é dita linear se \hat{y}_t é obtida através da [Equação 2.1](#),

$$\hat{y}_t = \mathbf{w}_t \cdot \mathbf{m}_t = \sum_{i=1}^N w_{t,i} m_{t,i} \quad (2.1)$$

e é dita não-linear caso contrário. É comum na literatura se referir a $w_{t,i}$ por peso, uma vez que, frequentemente, $0 \leq w_{t,i} \leq 1$ e $\sum_i w_{t,i} = 1$. Alternativamente, pesos com tais propriedades são ditos convexos. Neste manuscrito serão consideradas apenas regressões lineares, sendo uma delas convexa.

Apesar de cada método de agregação ter um regra própria para relacionar $w_{t,i}$ a $m_{t,i}$, a maioria segue um padrão bem definido: ao ser determinado \hat{y}_t , o ambiente revela y_t , e o jogador registra a discrepância entre \hat{y}_t e y_t através de uma função erro $L : \mathcal{D}^2 \rightarrow \mathbb{R}_{\geq 0}$. Uma função erro é qualquer função contínua não-negativa tal que $L(x, y) = 0 \iff x = y$. $L(y_t, \hat{y}_t)$ será usado pelo jogador na próxima partida para atualizar os pesos de cada modelo. Além disso, o objetivo do jogador será minimizar o erro acumulado durante a duração do jogo, ou seja, $\sum_{\tau=0}^t L(y_\tau, \hat{y}_\tau)$.

Contudo, em contextos práticos como reservatórios, é preciso estimar múltiplas variáveis a cada passo de tempo, ao invés de apenas uma medida observada y_t . Por exemplo, observe que a previsão de produção do reservatório como um todo advém da soma das previsões de produção de cada poço. Dessa forma, a preocupação do momento é generalizar a notação acima para comportar um vetor de estimativas.

Seja P o número de variáveis que procura-se prever, denotadas por $y_{t,p}$, $1 \leq p \leq P$, e

$\hat{y}_{t,p}$ as estimativas associadas. Como antes, o índice t indentifica o passo de tempo. A ideia aqui é realizar a versão simples do jogo, onde y_t era um escalar, para cada elemento do vetor $\mathbf{y}_t = (y_{t,1}, y_{t,2}, \dots, y_{t,p})^T$. Assim, para cada p , há um vetor de estimativas disponibilizado pelo especialistas $\mathbf{m}_{t,p} = (m_{t,1,p}, m_{t,2,p}, \dots, m_{t,N,p})^T$, e um vetor de pesos associado $\mathbf{w}_{t,p} = (w_{t,1,p}, w_{t,2,p}, \dots, w_{t,N,p})^T$, que resultam em $\hat{y}_{t,p}$ através do produto interno na [Equação 2.2](#):

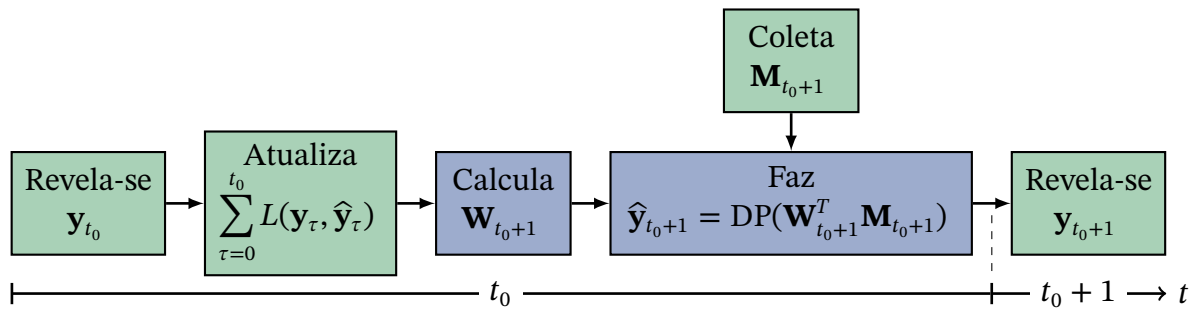
$$\hat{y}_{t,p} = \mathbf{w}_{t,p} \cdot \mathbf{m}_{t,p} = \sum_{i=1}^N w_{t,i,p} m_{t,i,p} \quad (2.2)$$

Ao todo, o vetor de previsões $\hat{\mathbf{y}}_t = (\hat{y}_{t,1}, \hat{y}_{t,2}, \dots, \hat{y}_{t,p})^T$ é obtido através da multiplicação entre a matriz de estimativas $\mathbf{M}_t = (\mathbf{m}_{t,1}, \mathbf{m}_{t,2}, \dots, \mathbf{m}_{t,p})$ e a matriz de pesos $\mathbf{W}_t = (\mathbf{w}_{t,1}, \mathbf{w}_{t,2}, \dots, \mathbf{w}_{t,p})$:

$$\hat{\mathbf{y}}_t = \text{DP}(\mathbf{W}_t^T \mathbf{M}_t) \quad (2.3)$$

onde $\text{DP} : \mathbb{R}^{P \times P} \rightarrow \mathbb{R}^{P \times 1}$ extrai a diagonal principal do argumento na forma de um vetor. Por exemplo, $\text{DP}(\mathbb{I}_{2 \times 2}) = (1, 1)^T$, sendo \mathbb{I} a matriz identidade. Quanto à função perda, a adaptação necessária aqui é simplesmente pôr \mathcal{D} como subconjunto de $\mathbb{R}^{P \times 1}$. As propriedades de $L : \mathcal{D}^2 \rightarrow \mathbb{R}_{\geq 0}$ se preservam, e o objetivo do jogador, ou algoritmo, é minimizar $\sum_t L(\mathbf{y}_t, \hat{\mathbf{y}}_t)$. A [Figura 2.2](#) resume o discutido até aqui por meio de um diagrama. O instante genérico $t = t_0$ começa com a observação dos novos dados do sistema, e termina com a estipulação da previsão agregada, passando pela atualização do acúmulo de erros, pelo cálculo da matriz de pesos, e pela coleta de sugestões dos especialistas. O próximo passo se inicia com a observação do sistema, reiniciando o processo executado no passo anterior.

Figura 2.2 – Ilustração do jogo da estratégia de agregação a nível quantitativo.



Fonte: Autor.

Na literatura, a matriz de pesos é às vezes chamada de modelo. O modelo gerado no método de agregação pode então alimentar um processo de otimização da produção, uma vez que guarda informações sobre quais realizações, e qual combinação destas, tenderam a melhor representar o reservatório alvo. Naturalmente, tal processo de otimização está além da comparação de estratégias de previsão em si, mas é essencial ter em mente o lugar da previsão dentro do objetivo maior no qual ela é efetivamente inserida: o aumento da produção do reservatório. A seguir, apresentam-se as maneiras como cada algoritmo da

agregação - AMA, MPE, Ridge e Lasso - calcula a matriz de pesos \mathbf{W}_t , bem como definem e utilizam a função de perda. A exposição começa com o algoritmo central da proposta deste manuscrito, a AMA.

A notação seguirá o mais perto possível aquela que foi apresentada acima mediante as adaptações ao contexto de reservatório de petróleo: N representará o número de realizações ou modelos, e P o número de poços em cada realização; $y_{t,p}$ representará alguma taxa de produção observada no poço p da referência, como óleo ou água, as quais a literatura simboliza frequentemente por q_{OPR} e q_{WPR} ; $m_{t,i,p}$ representará a previsão dessas taxas pela realização i no poço p , e pode vir a ser escrita como $q_{\text{OPR}}(t, i, p)$, por exemplo; finalmente, para evitar confusão entre o passo de tempo t em análise e a variável de um laço no tempo, esta última será representada por τ .

2.1 Agregação de Modelos Adaptativa

O mais importante objetivo deste manuscrito é validar uma nova estratégia de agregação, a AMA. Diferente das outras estratégias, aqui procura-se modelar, e atenuar, a interferência do ruído de medição ao qual os dados observados estão sujeitos no cálculo da matriz de pesos. Consequentemente, a metodologia proposta tem uma estrutura única em comparação com outras da literatura uma vez que, além de requerer uma etapa para o cálculo da matriz de pesos, existe uma segunda etapa subsequente para modificar, isto é, filtrar, a matriz em questão. Indique por $\widetilde{\mathbf{W}}_t$ a matriz de pesos \mathbf{W}_t filtrada. O filtro em si é ajustado a cada passo de tempo, comparando uma previsão agregada para o passo presente com o observado do ambiente. Essa previsão particular ao filtro, denotada por \hat{y}_t^+ , é feita não por \mathbf{W}_t , cujos efeitos do ruído de medição ainda não foram mitigados, mas por $\widetilde{\mathbf{W}}_{t-1}$. Mais precisamente, \hat{y}_t^+ é usado para definir um parâmetro da MME.

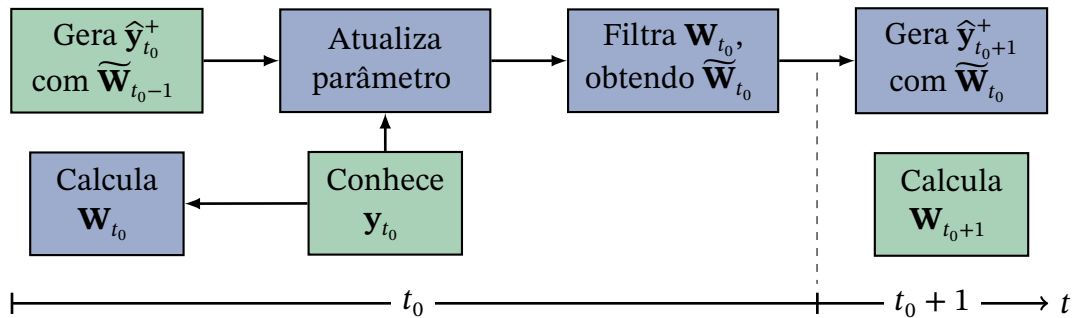
As etapas se relacionam através da seguinte sequência de operações, ilustrada na [Figura 2.3](#). Indicando por $t = t_0$ o passo de tempo atual, usa-se $\widetilde{\mathbf{W}}_{t_0-1}$ para obter as previsões agregadas $\hat{y}_{t_0}^+$ de óleo e água; em seguida, dado \mathbf{y}_{t_0} , o parâmetro da MME é ajustado; na sequência, calcula-se \mathbf{W}_{t_0} através de minimização de função objetivo; por fim, usa-se a MME para encontrar $\widetilde{\mathbf{W}}_{t_0}$. No próximo passo de tempo, $t = t_0 + 1$, $\widetilde{\mathbf{W}}_{t_0}$ gerará $\hat{y}_{t_0+1}^+$, reiniciando o processo.

Evidentemente, as matrizes de pesos filtradas obtidas ao longo da simulação podem ser armazenadas para realizar as previsões agregadas \hat{y}_t^+ via, por exemplo, a [Equação 2.3](#). Dito isso, é importante deixar claro algumas particularidades desse método. Primeiro, assume-se que o passo de tempo presente já possui as informações de produção daquele passo de tempo, ou seja, que pelo menos \mathbf{y}_t já é conhecido em t , o que não se assume em outros métodos da literatura. Segundo, \hat{y}_t^+ é obtida via a [Equação 2.4](#), uma adaptação de (2.3). Terceiro, a minimização da função objetivo compreende dois pontos de tempo, o presente $t = t_0$ e o

passado $t = t_0 - 1$, necessitando dos dados observados \mathbf{y}_t em ambos.

$$\hat{\mathbf{y}}_t^+ = \text{DP}(\widetilde{\mathbf{W}}_{t-1}^T \mathbf{M}_t). \quad (2.4)$$

Figura 2.3 – Visão geral da estratégia de agregação proposta.



Fonte: Autor.

O entendimento pleno desta estratégia de agregação pede a compreensão de suas duas partes principais, (i) a minimização que resulta em \mathbf{W}_t e (ii) a filtragem que resulta em $\widetilde{\mathbf{W}}_t$. A próxima subseção se preocupa em explicitar a minimização, e a seguinte, a filtragem.

Notadamente, a metodologia para se obter a matriz de pesos antes da filtragem é tida como uma versão integrada das regressões detalhadas em [Fortaleza, Filho e Munerato \(2022\)](#) e [Oliveira et al. \(2022\)](#), assim como superior. As elaborações subsequentes irão deixar claro o porquê.

2.1.1 Integração dos algoritmos

A regressão proposta obtém a matriz de pesos minimizando uma função objetivo de poço em poço, de forma a encontrar um vetor de pesos por minimização. Essa metodologia para obter a matriz também é a empregada pelo Lasso e Ridge, sendo o diferencial a formulação da função objetivo.

A função foi projetada para levar em consideração conjunta, dado um poço p , tanto o erro individual de cada realização em relação à produção observada de p , quanto o erro médio entre uma dada combinação das realizações em relação à produção observada de p . Aqui, a produção de p leva em conta tanto a vazão de água quanto a vazão de óleo, através de uma ponderação de cada vazão pelos seus custos. Essa ponderação permite prever o reservatório por uma única matriz de pesos, em vez de precisar gerar uma matriz para prever cada tipo de produção, como é o caso da MPE, Ridge e Lasso. Para iniciar o detalhamento matemático da proposta, veja a [Equação 2.5](#), que mostra a minimização da função objetivo

que gera $\mathbf{w}_{t,p}$:

$$\mathbf{w}_{t,p} \in \underset{(\omega_1, \dots, \omega_N)^T = \boldsymbol{\omega} \in \mathbb{R}^{N \times 1}}{\operatorname{argmin}} \left\{ c_1 \sum_{\tau=t-1}^t \left(\sum_{i=1}^N \omega_i \operatorname{EI}(\tau, i, p) \right) + c_2 \sum_{\tau=t-1}^t \operatorname{EM}(\tau, \boldsymbol{\omega}, p) \right\} \quad (2.5)$$

com c_1, c_2 reais positivos fixos conhecidos. Note que a minimização se dá em dois pontos, o passo de tempo presente t e o passado $t - 1$, conferindo um quê de Janelamento ao algoritmo. [Oliveira et al. \(2022\)](#) usam uma versão particular dessa minimização, com $c_1 = 0$ e $c_2 = 1$. Por outro lado, [Fortaleza, Filho e Munerato \(2022\)](#) consideram ambas as constantes positivas. Considera-se a minimização em (2.5) um avanço em relação àquelas nos trabalhos citados visto que o novo método refina a função objetivo sem contudo introduzir elementos supérfluos.

Resta definir as funções EI e EM. Primeiramente, tem-se que $\operatorname{EI}(\tau, i, p)$ modela o erro individual da realização i , em relação ao poço p , através da [Equação 2.6](#):

$$\operatorname{EI}(\tau, i, p) = [\nu_{\text{OPR}} \Delta q_{\text{OPR}}(\tau, i, p)]^2 + [\nu_{\text{WPR}} \Delta q_{\text{WPR}}(\tau, i, p)]^2 \quad (2.6)$$

onde

$$\Delta q_{\text{OPR}}(\tau, i, p) = q_{\text{OPR}}(\tau, i, p) - q_{\text{OPR}_{\text{OD}}}(\tau, p) \quad (2.7)$$

$$\Delta q_{\text{WPR}}(\tau, i, p) = q_{\text{WPR}}(\tau, i, p) - q_{\text{WPR}_{\text{OD}}}(\tau, p) \quad (2.8)$$

são as diferenças da taxa de produção de óleo e de água, respectivamente, do poço p na realização i em relação à produção observada desse mesmo poço. Já $\operatorname{EM}(\tau, \boldsymbol{\omega}, p)$ captura através da [Equação 2.9](#) o erro médio das realizações, em relação ao poço p , quando ponderadas por vetor de pesos $\boldsymbol{\omega}$:

$$\operatorname{EM}(\tau, \boldsymbol{\omega}, p) = [\nu_{\text{OPR}} (q_{\text{OPR}_{\boldsymbol{\omega}}}(\tau, \boldsymbol{\omega}, p) - q_{\text{OPR}_{\text{OD}}}(\tau, p))]^2 + [\nu_{\text{WPR}} (q_{\text{WPR}_{\boldsymbol{\omega}}}(\tau, \boldsymbol{\omega}, p) - q_{\text{WPR}_{\text{OD}}}(\tau, p))]^2 \quad (2.9)$$

com

$$q_{\text{OPR}_{\boldsymbol{\omega}}}(\tau, \boldsymbol{\omega}, p) = \sum_{i=1}^N \omega_i q_{\text{OPR}}(\tau, i, p) \quad (2.10)$$

$$q_{\text{WPR}_{\boldsymbol{\omega}}}(\tau, \boldsymbol{\omega}, p) = \sum_{i=1}^N \omega_i q_{\text{WPR}}(\tau, i, p) \quad (2.11)$$

termos iguais às produções previstas de óleo e água de um dado poço p quando os especialistas são ponderados por $\boldsymbol{\omega}$ e $\nu_{\text{OPR}}, \nu_{\text{WPR}}$ constantes positivas que modelam a importância econômica na produção de óleo e água, respectivamente. Por exemplo, se $\nu_{\text{OPR}} = 45$ e $\nu_{\text{WPR}} = 5$, o ajuste da minimização dará 9 vezes mais importância ao erro relativo ao óleo em detrimento da água. Uma vez que a técnica utilizada para realizar a minimização em (2.5) fora o algoritmo da região de confiança de Byrd-Omojokun ([LALEE; NOCEDAL; PLANTEGA, 1998](#)),

que necessita do gradiente da função objetivo, as funções erro EM e EI usam quadrados das diferenças ao invés de, por exemplo, valores absolutos. Ainda sobre as funções de erro, é importante notar que não é difícil incluir nestas outras produções do reservatório, desde que haja um custo para essa taxa. De fato, a função objetivo como apresentada é equivalente à que considera taxas de óleo, água e gás com custo de produção de óleo nulo, i. e., $\nu_{\text{GPR}} = 0$. Tomou-se a liberdade de desconsiderar o gás na função objetivo porque os reservatórios de validação EGG e OLYMPUS não produzem gás.

Fortaleza, Filho e Munerato (2022) e Oliveira et al. (2022) também consideram apenas as taxas de óleo e água, mas no primeiro artigo as funções que modelam os erros médio e individual independem do poço, uma vez que lá consideram-se $\sum_p EI(\tau, i, p)$ no lugar de apenas EI, e $\sum_p EM(\tau, \omega, p)$ no lugar de apenas EM. Em outras palavras, os erros a serem considerados na função objetivo foram acumulados por todo o reservatório, de poço em poço. Como consequência, em relação à notação deste manuscrito, é como se \mathbf{W}_t nesse trabalho tivesse colunas idênticas. Estabelecida ciência dessa particularidade, a Tabela 2.2 resume as principais diferenças entre as três funções objetivo.

Tabela 2.2 – Comparação estrutural de diferentes funções objetivo.

Função objetivo	Características				
	c_1	c_2	Erro Individual	Erro Médio	Resultado
Fortaleza, Filho e Munerato (2022)	> 0	> 0	Soma sobre p	Soma sobre p	Colunas da matriz de pesos
Oliveira et al. (2022)	$= 0$	$= 1$	—	Particular a p	Coluna da matriz de pesos
Equação 2.5	> 0	> 0	Particular a p	Particular a p	Coluna da matriz de pesos

Fonte: Autor.

Como comentado no início desta subseção, \mathbf{W}_t é obtido iterando a minimização em (2.5) para $1 \leq p \leq P$. Encontrado \mathbf{W}_t , resta saber como é aplicado o filtro para encontrar $\widetilde{\mathbf{W}}_t$, o que é resolvido a seguir.

2.1.2 Filtro Adaptativo

Assim como a compreensão plena da AMA passa pelo entendimento da etapa do cálculo da matriz de pesos via minimização e da etapa de filtragem dessa matriz, a compreensão plena da filtragem passa pelo entendimento de três pontos-chave: (i) a anatomia, vantagens e desvantagens da média móvel; (ii) a forma teórica de operação do filtro, conforme estabelecida em Fortaleza, Gomes et al. (2022), e; (iii) a forma como o filtro é implementado na prática, dentro e fora do método de agregação em proposta. Cada item citado é abordado a seguir de maneira segmentada.

Média móvel. A MME foi introduzida por [Brown \(1956\)](#) como um suavizador exponencial voltado para previsão e controle de inventários, mas com aplicações além dessa área, como, a saber, controle de qualidade ([HUNTER, 1986](#)). Brown Dado um sinal discreto no tempo $y(t)$, Brown propõe uma equação para prever seu valor médio ou esperado, e outra para prever seu valor real. Visto que em geral as aplicações da MME assumem a cada passo de tempo erros independentes e de média nula sobre y , é a equação que Brown deu para a média que veio a ser reconhecida como a MME, embora na ocasião do artigo sua intenção era validar a segunda equação. Com isso dito, a fórmula da MME é apresentada a seguir.

Sejam $\hat{y}(t)$ a previsão (da média) de $y(t)$ gerada pela MME, e $0 < \alpha < 1$ fixo. A previsão de $y(t + 1)$ é dada pela ponderação entre $y(t)$ e o erro de $\hat{y}(t)$, isto é,

$$\hat{y}(t + 1) = \hat{y}(t) + \alpha(y(t) - \hat{y}(t)) \quad (2.12)$$

$$= \alpha y(t) + (1 - \alpha)\hat{y}(t) \quad (2.13)$$

As equações (2.12) e (2.13), embora equivalentes matematicamente, destacam diferentes interpretações da média. A primeira, como comentado, expõe a média como uma ponderação entre a última observação do evento y e o erro de previsão a ele associado, enquanto a segunda como a combinação convexa entre o último dado observado e a previsão feita sobre ele.

Em ambas as interpretações, o parâmetro α tem o papel óbvio de calibrar a importância das parcelas, mas talvez sua ação mais importante seja a de calibrar a importância de valores passados de y , dando memória à previsão. De fato, aplicar (2.13) recursivamente resulta em (assumindo $\hat{y}(0) = 0$)

$$\hat{y}(t + 1) = \sum_{\tau=0}^t \alpha(1 - \alpha)^\tau y(t - \tau) \quad (2.14)$$

de forma que as medidas passadas até o presente são consideradas para estimação futura via ponderação cujos fatores são exponenciais que dependem de α . Em particular, valores recentes, próximos de t , são privilegiados em detrimento dos valores antigos próximos de zero. Essa atenuação é algo não muito diferente das técnicas de Janelamento e Desconto visto nos trabalhos de [Mallet, Stoltz e Mauricette \(2009\)](#) e [Amat, Michalski e Stoltz \(2018\)](#) no capítulo 1, mas aqui ela é intrínseca à equação, e não uma perturbação externa.

Uma vez que α controla o período cujas observações contribuem efetivamente para a nova previsão, é necessário olhar para o contexto da aplicação para decidir o valor ótimo, ou suficientemente próximo do ótimo, de α . Como explica [Brown \(1956\)](#), α pequeno é adequado para reagir apenas às mudanças de longo prazo, e ruim para reagir a mudanças de curto prazo, situação exatamente simétrica ao caso em que α é grande, que prevê melhor mudanças de curto prazo em detrimento das de longo prazo.

Em resumo, a MME é uma maneira computacionalmente eficiente de se obter previsões de uma medida, especialmente se já se tem em mente escopo da previsão em mente, i.e., curto, médio ou longo prazo. No que segue, a equação da média será ligeiramente modificada para servir como uma filtragem de um sinal discreto de entrada. Mais importantemente, o parâmetro α será atualizado no tempo com o intuito de melhor considerar as observações passadas.

Filtro adaptativo teorizado. Até aqui, a MME foi vista como um método de previsão cujo parâmetro α é estático. Entretanto, se ao invés de $y(t + 1)$ no lado esquerdo de (2.13) houvesse $y(t)$, a equação poderia ser interpretado como um filtro passa baixas, atenuando as altas frequências de y .

Assim, para α bem escolhido, a MME atenuaria ruído ao mesmo tempo em que possuiria boa resposta ao degrau, fatores que, em conjunto com a sua simples complexidade computacional, a tornam um interessante candidato a filtro, como propuseram [Fortaleza, Gomes et al. \(2022\)](#). Para resolver o problema de ter α “bem escolhido”, os autores introduziram um algoritmo para atualização do parâmetro α , que será agora calculado a cada passo de tempo de maneira adaptativa. Em outras palavras, α se torna uma função no tempo $\alpha = \alpha(t)$. Ao todo, o filtro proposto tem o intuito de ser inserido em malha fechada cujo sistema de controle é do tipo Proporcional-Integral-Derivativo (PID), o mais corrente em aplicações práticas de controle. Cabe ver seu funcionamento. Denotando por $y_I(t)$ o sinal contínuo de entrada do filtro no instante t , e $y_O(t)$ sua saída, define-se

$$y_O(t) = y_O(t - 1) + \alpha(t)[y_I(t) - y_O(t - 1)] \quad (2.15)$$

uma adaptação de (2.13).

A ideia de como definir ou atualizar o coeficiente α tem embrião no artigo de [J. O. A. Límaverde Filho E. L. F. Fortaleza e Campos \(2022\)](#). No contexto de controle em malha fechada, propõe-se uma variação adaptativa do filtro de Kalman para melhor lidar com a covariância do ruído no processo, um parâmetro conhecido a priori no filtro de Kalman original ([KALMAN, 1960](#)). Uma vez que a covariância do ruído está relacionada à covariância do vetor de medições, os autores estimam a primeira covariância pela segunda, e obtém a segunda por uma MME matematicamente equivalente a da [Equação 2.15](#). Entretanto, o parâmetro desse artigo é definido de maneira estática no tempo, em dependência com o período de amostragem e o polo mais rápido do sistema em malha fechada.

Por sua vez, [Fortaleza, Gomes et al. \(2022\)](#) assumem que y_I é sujeito a um ruído de medição gaussiano de média nula $n \sim \mathcal{N}(0, \sigma_n)$ tal que $y_I(t) = y(t) + n$, onde y é o sinal observado na planta. Através de várias hipóteses estatísticas feitas sobre y_O e y_I , os autores concluem que

$$\alpha_t = \operatorname{erf} \left(\frac{|y_O(t - 1) - y_I(t)|}{2\sigma_n\sqrt{2}} \right) \quad (2.16)$$

onde $\text{erf} : \mathbb{R} \rightarrow [-1, 1]$ é a função de erro de Gauss:

$$\text{erf}(z) = \frac{1}{\sqrt{\pi}} \int_{-z}^z e^{-x^2} dx. \quad (2.17)$$

Filtro adaptativo aplicado. Gomes, A. Limaverde Filho et al. (2023) e Gomes, Longhi et al. (2023, não publicado) usam a MME na Equação 2.15 em aplicações práticas de ponte rolante, e corte de naphtha leve, respectivamente. Porém, a aplicação dessa fórmula para a média não é imediata no contexto de agregação de modelos, uma vez que algumas das hipóteses tomadas no contexto de malha fechada não mais se aplicam, e além disso, a agregação de lida com múltiplas variáveis simultaneamente (pense por exemplo no vetor de previsões agregadas). Essa adaptação da MME original numa situação vetorial/matricial acarretará mudanças também na fórmula adaptativa de α .

Primeiro, discute-se os efeitos sobre a MME, para depois tratar do seu coeficiente. Ao transladar esse filtro do contexto de controle em malha fechada à nova estratégia de agregação de modelos, a MME trabalha não mais com escalares, e, mais importantemente, agirá não sobre as previsões agregadas em si, mas sim sobre os pesos das realizações. Esclarecendo a segunda afirmação, o vetor de pesos $\mathbf{w}_{t,p}$ obtido na otimização da subseção anterior é filtrado através da relação abaixo, baseada na Equação 2.15.

$$\tilde{\mathbf{w}}_{t,p} = \alpha_{t,p} \mathbf{w}_{t,p} + (1 - \alpha_{t,p}) \tilde{\mathbf{w}}_{t-1,p} \quad (2.18)$$

Note que agora pode-se falar de um vetor $P \times 1$ de parâmetros α , denotado por $\alpha_t = (\alpha_{t,1}, \dots, \alpha_{t,2}, \alpha_{t,p})^T$, sendo cada entrada associada a um poço do reservatório. Por sua vez, $\alpha_{t,p}$, $1 \leq p \leq P$ é dado por

$$\alpha_{t,p} = \text{erf} \left(u_{\text{OPR}} \frac{|\Delta q_{\text{OPR}\tilde{\mathbf{w}}_{t-1,p}}|}{f_{\sigma} \sigma_{\% \text{OPR}} \sqrt{2}} + u_{\text{WPR}} \frac{|\Delta q_{\text{WPR}\tilde{\mathbf{w}}_{t-1,p}}|}{f_{\sigma} \sigma_{\% \text{WPR}} \sqrt{2}} \right) \quad (2.19)$$

onde

$$\Delta q_{\text{OPR}\tilde{\mathbf{w}}_{t-1,p}} = q_{\text{OPR}_{\text{OD}}}(t, p) - q_{\text{OPR}_{\omega}}(t, \omega, p) \Big|_{\omega=\tilde{\mathbf{w}}_{t-1,p}} \quad (2.20)$$

$$\Delta q_{\text{WPR}\tilde{\mathbf{w}}_{t-1,p}} = q_{\text{WPR}_{\text{OD}}}(t, p) - q_{\text{WPR}_{\omega}}(t, \omega, p) \Big|_{\omega=\tilde{\mathbf{w}}_{t-1,p}} \quad (2.21)$$

introduzem o erro entre a previsão feita por $\tilde{\mathbf{w}}_{t-1,p}$ e o observado $y_{t,p}$ para as produções de óleo e água, vide equações (2.10) e (2.11). Os símbolos $\sigma_{\% \text{OPR}}$ e $\sigma_{\% \text{WPR}}$ indicam os desvios padrões na taxa de produção de óleo e água, nessa ordem, na medida observada de um dado poço em um dado passo de tempo. Esses desvios estão relacionados com uma porcentagem de ruído de medição assumida fixa. Por exemplo, se no passo de tempo t a produção de óleo observada é $y_{t,p} = 100$, e a porcentagem de ruído é 5%, então $\sigma_{\% \text{OPR}} = 5$. Idem para a água.

É claro que (2.19) é uma adaptação de (2.16). As principais diferenças nesta versão são duas. Primeiro, uma vez que as hipóteses estatísticas feitas em situação de controle em malha

fechada não são todas válidas, o denominador em (2.16) não advém uma dedução matemática. Assim, o fator $f_\sigma > 0$ deve ser determinado empiricamente, de maneira a melhorar a performance da AMA. Na prática, observou-se que os resultados com $f_\sigma = 1, 2$ ou 4 tiveram pouca diferença entre si, embora $f_\sigma = 1$ tenha performado melhor. A segunda diferença é que $\alpha_{t,p}$, bem como a minimização em (2.5), leva em conta simultaneamente erros nas previsões de óleo e água, que são ponderados de acordo com sua relevância econômica. Em relação ao diagrama na Figura 2.3, $\alpha_{t,p}$ varia de maneira proporcional à diferença $|y_{t,p} - y_{t,p}^+|$, o erro cometido pela previsão agregada quando o modelo não é atualizado. Logicamente, esse erro não pode ser desconsiderado no contexto de reservatório uma vez que a produção só é possível em regime transiente e o tempo de amostragem não é pequeno. O papel da erf aqui é atualizar $\alpha_{t,p}$ de acordo com a relevância estatística desse erro em regime transiente, diferindo do uso em (2.16), cujo tempo de amostragem é mínimo e o regime estacionário é rapidamente obtido.

Note que, para o passo de tempo inicial (e. g., $t = 0$), a Equação 2.18 não pode ser seguida, de modo que, nesse caso, $\widetilde{\mathbf{W}}_t = \mathbf{W}_t$. Novamente, assim como na etapa de minimização, é possível incluir a produção de gás na determinação de α , mas essa produção é ignorada uma vez que os reservatórios de validação adotados neste manuscrito não produzem gás.

Dissecada a AMA em seus dois componentes principais, e etapa de filtragem e da estimação de modelo, a seção subsequente trata de apresentar os métodos da literatura contra os quais a AMA será comparada.

2.2 Alguns métodos de agregação da literatura

A MPE, o Lasso e o Ridge, assim como métodos de agregação em geral, usam um ou dois parâmetros dentro de suas regras de agregação. Também chamados de hiperparâmetros, esses números são reais positivos ou não-negativos introduzidos durante o cálculo dos pesos a fim de diminuir o acúmulo de erro das rodadas da agregação. Matematicamente, os pesos são funções explícitas dos hiperparâmetros. Existe na literatura extenso material dedicado ao estudo em como obter valores ótimos para os hiperparâmetros, isto é, constantes cujos pesos associados minimizem o erro acumulado ao longo da agregação, e. g. Cesa-Bianchi e Lugosi (2006) e Vovk (2001). Como mostram Mallet, Stoltz e Mauricette (2009) e Devaine et al. (2013), na prática os valores teoricamente ótimos não tendem a fornecer as melhores previsões. Além disso, conforme apontado por Deswarte et al. (2019), um método de agregação pode não ter a garantia de um parâmetro ótimo, como é o caso do Lasso, e, mesmo quando tenha, pode não ser possível de calcular a priori, como é o caso do Ridge e MPE.

Desconsiderada forma analítica pela qual definir o parâmetro no cálculo dos pesos, resta a técnica de tentativa e erro. Isso pode ser feito manualmente, pelo engenheiro ou

programador responsável, ou automática e dinamicamente, por algoritmo integrante à regressão.

No método manual, o objetivo é definir após uma bateria de testes um valor constante para o hiperparâmetro. O hiperparâmetro resultante não depende nem do poço em análise, nem do passo de tempo atual. De maneira geral, esse valor é encontrado definindo uma lista $\mathbf{h} = (h_1, h_2, \dots, h_J)^T$, $h_j > 0$, $1 < j < J$, de candidatos e executando simulações para cada h_j . Aquele que possuir as previsões associadas mais precisas será tomado como o valor do parâmetro a ser adotado durante a regressão em tempo real. Logicamente, \mathbf{h} também é definida por meio de simulações e testes sobre o reservatório, em geral através de refinações de $\mathbb{R}_{>0}$ para subconjuntos cada vez menores. Obteve-se c_1 e c_2 da seção 2.1 dessa forma. Apesar de sua estrutura simples e direta, infelizmente o método manual tanto é sensível ao reservatório quanto introduz uma etapa *offline* que potencialmente possa comprometer economicamente diversas aplicações industriais..

A maneira automática visa resolver o problema de otimização do parâmetro de forma *online* e insensível ao reservatório. Isso pode ser feito de duas maneiras, a depender da existência ou não existência de um valor ótimo para o parâmetro dada a fórmula dos pesos associados. Auer, Cesa-Bianchi e Gentile (2002) e Cesa-Bianchi, Mansour e Stoltz (2007) apresentam métodos de atualização de parâmetro em tempo real projetados para se aproximar do parâmetro ideal, mas como já comentado, na prática esses valores não tendem a resultar em boas performances. Por outro lado, Devaine et al. (2013) e Deswarte et al. (2019) adaptam a estratégia do método manual ao laço da regressão (ver Figura 2.2), através de minimização de função objetivo erro, a ser explicada a seguir.

Sejam t o passo de tempo presente, e p o poço para o qual deseja-se obter a previsão agregada do próximo passo de tempo, $\hat{y}_{t+1,p}$. Indique por $h_{t+1,p}^* \in \mathbf{h}$ o hiperparâmetro ótimo. Calcula-se para cada h_j em \mathbf{h} o erro acumulado

$$e_{t,p}(h_j) = \sum_{\tau=1}^t L(\hat{\gamma}_{\tau,p}(h_j), y_{\tau,p}) \quad (2.22)$$

e toma-se $h_{t+1,p}^*$ como o candidato responsável pela menor entrada no vetor de erros $\mathbf{e}_{t,p} = (e_{t,p}(h_1), \dots, e_{t,p}(h_J))^T$. Aqui, $\hat{\gamma}_{\tau,p}(h_j)$ é a previsão agregada do poço p e instante τ quando usando o candidato h_j . Matematicamente,

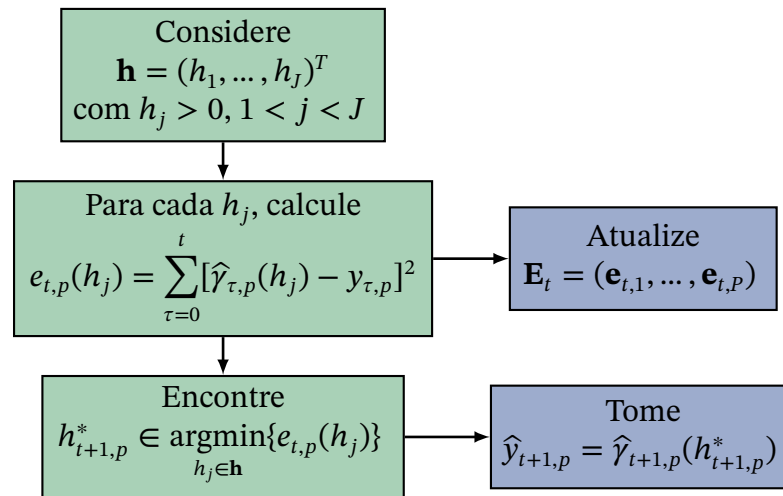
$$\hat{\gamma}_{\tau,p}(h_j) = \mathbf{w}_{\tau,p}(h_j) \cdot \mathbf{m}_{\tau,p} \quad (2.23)$$

de forma que $\hat{y}_{t+1,p} = \hat{\gamma}_{t+1,p}(h_{t+1,p}^*)$. Note que $\hat{\gamma}$ também depende implicitamente do algoritmo de agregação em análise. Neste manuscrito, assim como no trabalho de Deswarte et al. (2019), toma-se

$$L(\hat{\gamma}_{\tau,p}(h_j), y_{\tau,p}) = [\hat{\gamma}_{\tau,p}(h_j) - y_{\tau,p}]^2 \quad (2.24)$$

para a MPE, Ridge e Lasso. A Figura 2.4 resume a técnica de determinação de $h_{t+1,p}^*$ abordada acima num diagrama.

Figura 2.4 – Etapas da atualização do hiperparâmetro em tempo real para o poço p e passo t .



Fonte: Autor.

A matriz \mathbf{E} guarda os vetores de perda ao longo dos passos de tempo, tornando possível usar a relação abaixo, o que acelera a simulação.

$$e_{t,p}(h_j) = e_{t-1,p}(h_j) + [\hat{y}_{t,p}(h_j) - y_{t,p}]^2 \quad (2.25)$$

Mesmo nesta versão automática existem dois problemas com o vetor de candidatos \mathbf{h} : primeiro estabelecer J e segundo, relação entre as entradas h_j . Devaine et al. (2013) usaram a seguinte estratégia para obter bons hiperparâmetros de uma versão modificada da MPE com dois parâmetros, $\eta_t > 0$ e $0 \leq \alpha_t \leq 1$: defina uma lista de candidatos inicial \mathbf{h}_η para η_{t+1} , e execute a agregação do passo de tempo presente t . Se o η escolhido foi o maior valor R do vetor, acrescente a ele as entradas $2R$, $4R$ e $8R$; se o menor r , acrescente $r/2$, $r/4$ e $r/8$. Feito o ajuste, rode uma simulação para cada $\alpha \in \mathbf{h}_\alpha$, com $\mathbf{h}_\alpha = (0 \ 0,005 \ 0,01 \ 0,05 \ 0,1 \ 0,2 \ 0,5 \ 1)^T$. O hiperparâmetro α_t será o valor testado que acumulou o menor erro. Admite-se que as potências de 2 na expansão de \mathbf{h}_η foram impostas, após alguns testes, de maneira essencialmente arbitrária, e que desde que o vetor cubra intervalo suficientemente grande, bons parâmetros serão encontrados.

Deswarte et al. (2019) afirmam ainda que a ordem de grandeza dos candidatos é mais influente sobre \hat{y} que seus valores exatos, de forma que \mathbf{h} nesse trabalho tem valores distribuídos exponencialmente. A abordagem desses autores é portanto similar à estratégia de atualização de α_t em Devaine et al. (2013), embora melhor sistematizada, e será a adotada neste manuscrito. Detalhes das definições de \mathbf{h} para os hiperparâmetros da MPE, Ridge e Lasso serão abordados no Capítulo 4.

As próximas subseções explicitam uma a uma a fórmula dos pesos em função do hiperparâmetro ótimo, $\mathbf{w}_{t+1,p} = \mathbf{w}_{t+1,p}(h_{t+1,p}^*)$ do Ridge, Lasso e MPE, nessa ordem. Ademais, são discutidas também as vantagens e desvantagens de cada método, assim como é dada

a contextualização de cada um. Findadas essas subseções, inicia-se no próximo capítulo a apresentação dos reservatórios de validação.

2.2.1 Ridge

Motivação. Arthur Edwin Hoerl Jr. (1921-1994[†]), enquanto engenheiro mecânico e estatístico, se preocupou em analisar problemas de regressão linear. A caracterização matemática do problema é dada a seguir, já no contexto de reservatório. Em relação à linguagem de agregação de modelos e à notação introduzida no início desta seção, primeiro assume-se que para o vetor de registros de um poço p , $\mathbf{y}_p = (y_{1,p}, y_{2,p}, \dots, y_{t,p})^T$, existe vetor de pesos $\mathbf{w} = (w_1, \dots, w_N)^T$ para o qual $\mathbf{y}_p = \mathbf{M}_p \mathbf{w} + \boldsymbol{\epsilon}$, sendo $\boldsymbol{\epsilon}$ vetor de discrepância com média nula $(0 \dots 0)^T$ e variância $\sigma \mathbf{I}$, e $\mathbf{M}_p = (\mathbf{m}_{1,p}, \mathbf{m}_{2,p}, \dots, \mathbf{m}_{t,p})^T$ a matriz de especialistas do poço p . Escrita a equação de regressão, o objetivo é estimar \mathbf{w} por $\boldsymbol{\omega} = (\omega_1, \dots, \omega_N)$ de forma que o erro quadrático (também chamado soma residual dos quadrados) $L(\mathbf{y}_p, \mathbf{M}_p \boldsymbol{\omega}) = (\mathbf{y}_p - \mathbf{M}_p \boldsymbol{\omega})^T (\mathbf{y}_p - \mathbf{M}_p \boldsymbol{\omega})$ seja mínimo. Na época, os algoritmos para análise do erro quadrático ou se tornavam impraticáveis para número de dimensões grande suficiente como é o caso gráficos de contorno, ou se tornavam complexas ao ponto de dificultar interpretação, como é o caso da análise canônica. Isso motivou a criação de um novo tipo de análise, dita análise de Ridge, que eventualmente resultou na regressão de Ridge. Tanto a análise quanto a regressão são apresentadas a seguir.

Análise de Ridge. Um bom ponto inicial para começar a análise consiste em achar $\hat{\boldsymbol{\omega}}$ tal que $\partial L / \partial \hat{\boldsymbol{\omega}} = 0$, dita a estimativa dos quadrados mínimos. Visto que $\mathbf{M}_p^T \mathbf{y}_p = \mathbf{y}_p^T \mathbf{M}_p$, tem-se

$$0 = \frac{\partial L}{\partial \hat{\boldsymbol{\omega}}} \iff \quad (2.26)$$

$$0 = \frac{\partial}{\partial \hat{\boldsymbol{\omega}}} \left(\mathbf{y}_p^T \mathbf{y}_p + \hat{\boldsymbol{\omega}}^T \mathbf{M}_p^T \mathbf{M}_p \hat{\boldsymbol{\omega}} - \mathbf{y}_p^T \mathbf{M}_p \hat{\boldsymbol{\omega}} - \hat{\boldsymbol{\omega}}^T \mathbf{M}_p^T \mathbf{y}_p \right) \Rightarrow \quad (2.27)$$

$$0 = 2 \mathbf{M}_p^T \mathbf{M}_p \hat{\boldsymbol{\omega}} - 2 \mathbf{M}_p^T \mathbf{y}_p \iff \quad (2.28)$$

$$\hat{\boldsymbol{\omega}} = (\mathbf{M}_p^T \mathbf{M}_p)^{-1} \mathbf{M}_p^T \mathbf{y}_p. \quad (2.29)$$

Entretanto, como detalham Arthur E. Hoerl e Kennard (1970), $\hat{\boldsymbol{\omega}}^T \hat{\boldsymbol{\omega}}$ será muito maior que $\mathbf{w}^T \mathbf{w}$ quanto menores forem os autovalores de $\mathbf{M}_p^T \mathbf{M}_p$. Como autovalores pequenos advém da não ortogonalidade entre as colunas de \mathbf{M}_p , pode-se dizer que $\hat{\boldsymbol{\omega}}$ será um estimador tão pior quanto mais correlacionados forem os vetores $\mathbf{m}_{\tau,p}$, $1 \leq \tau \leq t$. No contexto de reservatório, por exemplo, é óbvio que essa hipótese é válida. Cabe então observar o comportamento de

$L(\boldsymbol{\omega}, \hat{\boldsymbol{\omega}})$ conforme $\boldsymbol{\omega}$ se distancia de $\hat{\boldsymbol{\omega}}$. Tem-se

$$L(\mathbf{y}_p, \mathbf{M}_p \boldsymbol{\omega}) = (\mathbf{y}_p - \mathbf{M}_p \boldsymbol{\omega})^T (\mathbf{y}_p, \mathbf{M}_p \boldsymbol{\omega}) \quad (2.30)$$

$$= (\mathbf{y}_p - \mathbf{M}_p \hat{\boldsymbol{\omega}} - \mathbf{M}_p (\boldsymbol{\omega} - \hat{\boldsymbol{\omega}}))^T (\mathbf{y}_p - \mathbf{M}_p \hat{\boldsymbol{\omega}} - \mathbf{M}_p (\boldsymbol{\omega} - \hat{\boldsymbol{\omega}})) \quad (2.31)$$

$$= (\mathbf{y}_p - \mathbf{M}_p \hat{\boldsymbol{\omega}})^T (\mathbf{y}_p - \mathbf{M}_p \hat{\boldsymbol{\omega}}) + (\boldsymbol{\omega} - \hat{\boldsymbol{\omega}})^T \mathbf{M}_p^T \mathbf{M}_p (\boldsymbol{\omega} - \hat{\boldsymbol{\omega}}) \quad (2.32)$$

$$= L(\mathbf{y}_p, \mathbf{M}_p \hat{\boldsymbol{\omega}}) + \Delta(\boldsymbol{\omega}), \quad (2.33)$$

onde $\Delta(\boldsymbol{\omega}) = (\boldsymbol{\omega} - \hat{\boldsymbol{\omega}})^T \mathbf{M}_p^T \mathbf{M}_p (\boldsymbol{\omega} - \hat{\boldsymbol{\omega}})$. Pelo comentado anteriormente, deseja-se que $\boldsymbol{\omega}^T \boldsymbol{\omega}$, a norma 2 de $\boldsymbol{\omega}$, seja tão menor quanto mais correlacionados estão $\mathbf{m}_{1,p}, \dots, \mathbf{m}_{t,p}$. Logo, sendo $L(\mathbf{y}_p, \mathbf{M}_p \hat{\boldsymbol{\omega}})$ fixo, o objetivo se torna minimizar $\boldsymbol{\omega}^T \boldsymbol{\omega}$ assim como $\Delta(\boldsymbol{\omega})$. Fixado $\Delta_0 > 0$, a análise de Ridge observa o caminho traçado pelo conjunto de vetores $\boldsymbol{\omega}$ tais que $\Delta(\boldsymbol{\omega}) = \Delta_0$, procurando $\boldsymbol{\omega}$ cuja norma 2 é mínima. É por causa dessa análise de caminho ou traço que dá o nome do método: do inglês, *ridge* pode ter significado de crista ou relevo, um caminho estreito de destaque numa superfície. Um caminho equivalente consiste em observar os mínimos de $L(\mathbf{y}_p, \mathbf{M}_p \boldsymbol{\omega})$ para os quais $\boldsymbol{\omega}^T \boldsymbol{\omega} = R^2$, $R > 0$ fixo. Cabe agora explorar essa análise. Note que é possível enquadrá-la como um problema de otimização Lagrangiano,

$$\begin{cases} \text{Minimize} & L(\mathbf{y}_p, \mathbf{M}_p \boldsymbol{\omega}) \\ \text{Sujeito a} & \boldsymbol{\omega}^T \boldsymbol{\omega} - R^2 = 0 \end{cases} \quad (2.34)$$

cujo método de solução é conhecido. Denotando por λ o multiplicador de Lagrange, e definindo função de Lagrange

$$\mathcal{L} : \mathbb{R}^{N \times 1} \times \mathbb{R} \rightarrow \mathbb{R} \quad | \quad \mathcal{L}(\boldsymbol{\omega}, \lambda) = L(\mathbf{y}_p, \mathbf{M}_p \boldsymbol{\omega}) + \lambda(\boldsymbol{\omega}^T \boldsymbol{\omega} - R^2) \quad (2.35)$$

Arthur E. Hoerl e Kennard (1970) mostram por derivação parcial de \mathcal{L} em $\boldsymbol{\omega}$ que o mínimo da função se dá em $\boldsymbol{\omega}^* = (\mathbf{M}_p^T \mathbf{M}_p + \lambda \mathbf{I})^{-1} \mathbf{M}_p^T \mathbf{y}_p$, onde λ é tal que $\boldsymbol{\omega}^T \boldsymbol{\omega} = R^2$. Todavia, na prática, pode ser difícil ou impossível definir bons intervalos para R (ainda mais para Δ_0), de forma que a função de Lagrange é ligeiramente modificada, resultando na regressão de Ridge.

Regressão de Ridge. Como comentado, a função de Lagrange em (2.35) pode ser difícil de ser aplicada. A ideia é retirar dela o termo R^2 , de forma que a relação anterior entre o ponto de mínimo e λ deixa de ser válida. A equação abaixo ilustra a função modificada.

$$L(\boldsymbol{\omega}, \lambda) = (\mathbf{y}_p - \mathbf{M}_p \boldsymbol{\omega})^T (\mathbf{y}_p - \mathbf{M}_p \boldsymbol{\omega}) + \lambda \boldsymbol{\omega}^T \boldsymbol{\omega} \quad (2.36)$$

Aqui, como observam Gareth James et al. (2021), um dado valor do multiplicador está associado a um dado ponto de mínimo, sendo portanto crucial determinar um “bom valor” para λ . No contexto de Agregação de Modelos, (2.36) pode ser vista como a função objetivo cujo ponto de mínimo é o vetor de pesos $\mathbf{w}_{t+1,p}$, e o bom valor para o multiplicador é exatamente o hiperparâmetro ótimo $h_{t+1,p}^*$, discutido no início desta seção. Em suma, abandonando a

notação vetorial, (2.36) leva à seguinte fórmula de vetor de pesos:

$$\mathbf{w}_{t+1,p} \in \underset{(\omega_1, \dots, \omega_N)^T \in \mathbb{R}^{N \times 1}}{\operatorname{argmin}} \left\{ h_{t+1,p}^* \sum_{i=1}^N \omega_i^2 + \sum_{\tau=0}^t \left(y_{\tau,p} - \sum_{i=1}^N \omega_i m_{\tau,i,p} \right)^2 \right\} \quad (2.37)$$

Essa foi a fórmula utilizada para gerar os resultados do próximo capítulo. O hiperparâmetro $h_{t+1,p}^*$ é às vezes chamado de parâmetro de ajuste, ou parâmetro de regularização, e pondera a importância de ω pequeno, pelo menos em relação à estimativa dos quadrados mínimos. Por causa disso, Gareth James et al. (2021) chamam $h_{t+1,p}^* \sum_{i=1}^N \omega_i^2$ de penalidade ou penalização de encolhimento. Para $h_{t+1,p}^* = 0$, $\omega = \hat{\omega}$, e a penalidade é nula; se agora $h_{t+1,p}^* \rightarrow \infty$, uma grande ênfase é colocada sobre a norma de ω , diminuindo várias de suas entradas. Do ponto de vista da Agregação de Modelos, essa redução é um ponto positivo na medida em que o vetor de pesos resultante privilegia os especialistas fieis aos dados observados, levando em pouca consideração o restante. Nesse caso, diz-se que o método de agregação é seletivo. A seguir, pontuam-se mais algumas vantagens e desvantagens deste algoritmo.

O leitor é encorajado a ler Roger W. Hoerl (2020), para conhecer mais detalhes sobre a contextualização do Ridge.

Vantagens e desvantagens. Com mais de 50 anos de história, o Ridge é hoje um método consagrado na literatura de regressões lineares, e comum na Agregação de Modelos (vide Tabela 1.1). Arthur E. Hoerl e Kennard (1970) demonstraram que sempre existirá $\lambda > 0$ para o qual o valor esperado da distância entre ω^* e o ótimo \mathbf{w} será menor que aquele da distância entre $\hat{\omega}$ e \mathbf{w} , confirmando a superioridade do método ante o método dos quadrados mínimos. Porém, esse resultado é válido para o problema de otimização na análise de Ridge, e não para a função modificada (2.36). Conforme afirmam Deswarte et al. (2019), a falta de um limitante na norma 2 de ω requer uma análise do valor ótimo do hiperparâmetro, o que acarreta grande carga computacional. Por outro lado, a seletividade do Ridge pode ser aperfeiçoada através de uma pequena modificação em (2.37); essa é a proposta central do segundo método da literatura a ser apresentado, o Lasso.

2.2.2 Lasso

Pensando em como melhorar a seletividade do Ridge, sem abrir mão do encolhimento de coeficiente intrínseco, Tibshirani (1996) modificou a Equação 2.37 trocando a penalidade em norma 2 por penalidade em norma 1. A minimização resultante é mostrada abaixo.

$$\mathbf{w}_{t+1,p} \in \underset{(\omega_1, \dots, \omega_N)^T \in \mathbb{R}^{N \times 1}}{\operatorname{argmin}} \left\{ h_{t+1,p}^* \sum_{i=1}^N |\omega_i| + \sum_{\tau=0}^t \left(y_{\tau,p} - \sum_{i=1}^N \omega_i m_{\tau,i,p} \right)^2 \right\} \quad (2.38)$$

Advém daí o nome do método, do inglês *Least Absolute Shrinkage and Selection Operator*, que poderia ser traduzido como “Operador de Mínimos Encolhimento Absoluto e Seleção”. Lasso é mais simples, e por isso foi mantido no manuscrito.

Exceto para $h_{t+1,p}^* \rightarrow \infty$, a minimização do Ridge sempre irá imputar coeficientes não nulos às entradas de $\mathbf{w}_{t+1,p}$, podendo tornar difícil a interpretação do modelo através de leitura da matriz de pesos.

Ao contrário do Ridge, esta nova maneira de se obter o vetor de pesos tende a ser capaz de, para $h_{t+1,p}^* < \infty$ suficientemente grande, zerar vários dos coeficientes relacionados aos especialistas mais imprecisos. Consequentemente, a matriz de pesos resultante é esparsa, e portanto mais legível e interpretável. Uma justificativa heurística dessa mudança de comportamento foi dada por [Tibshirani \(1996\)](#).

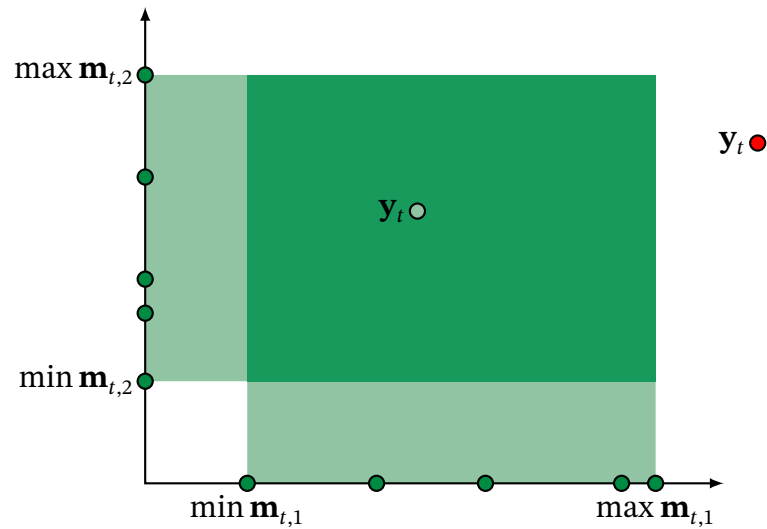
Não obstante a interpretabilidade do resultado, [Gareth James et al. \(2021\)](#) mostram que a seletividade mais radical do Lasso não é interessante no caso em que os especialistas tendem a render boas estimativas do fenômeno observado. Os autores usam um estudo de caso para afirmar que os métodos se complementam: se seletividade maior é necessária, o Lasso é um método mais preciso; caso contrário, o Ridge se sai melhor.

2.2.3 Média Ponderada Exponencial

A MPE, também chamada Hedge em [Erven et al. \(2011\)](#), introduzida por [Littlestone e Warmuth \(1994\)](#) e [Vovk \(1990\)](#), e estudada por, entre outros, [Cesa-Bianchi, Freund et al. \(1997\)](#), [Cesa-Bianchi \(1999\)](#) e [Auer, Cesa-Bianchi e Gentile \(2002\)](#), dá a seguinte fórmula para cada elemento da matriz de pesos,

$$w_{t+1,i,p} = \frac{\exp\left(-h_{t+1,p}^* \sum_{\tau=0}^t (y_{\tau,p} - m_{\tau,i,p})^2\right)}{\sum_{j=1}^N \exp\left(-h_{t+1,p}^* \sum_{\tau=0}^t (y_{\tau,p} - m_{\tau,j,p})^2\right)} \quad (2.39)$$

onde $h_{t+1,p}^* > 0$. Assim, ao contrário do Ridge e do Lasso, os pesos da MPE são sempre convexos, o que garante que a previsão agregada fique entre o menor e o maior valor indicado pelos especialistas. Nesse sentido, pode-se dizer que a MPE gera previsões cautelosas. Isso pode ser uma vantagem considerável no caso em que o alvo da agregação fica dentro do casco convexo gerado pelos especialistas ao longo do tempo, mas se não, é possível que ocorra erro considerável. Para um dado passo t , o casco convexo dos especialistas é o conjunto de pontos (x_1, \dots, x_p) em \mathbb{R}^P tais que $\min \mathbf{m}_{t,p} \leq x_p \leq \max \mathbf{m}_{t,p}$, $1 \leq p \leq P$. A figura abaixo mostra um casco simples, com $p = 2$ e cinco especialistas. A região verde escura corresponde aos possíveis valores de \hat{y}_t . Destacou-se em vermelho o valor observado \mathbf{y}_t que está fora do casco, e portanto não poderia ser atingido pela MPE.

Figura 2.5 – Casco convexo com $p = 2$.

Por outro lado, assim como no Ridge e Lasso, o hiperparâmetro age como um seletor de modelos. De fato, para $h_{t+1,p}^*$ suficientemente pequeno, $w_{t+1,i,p} \sim w_{t+1,j,p}$, $i \neq j$, e portanto $\mathbf{w}_{t+1,p}$ resulta numa ponderação uniforme; se agora $h_{t+1,p}^*$ é suficientemente grande, o peso será tão menor quanto pior for a performance passada do especialista associado, enquanto o conjunto de especialistas de melhor performance terão soma dos pesos próxima a 1. O hiperparâmetro da MPE é considerado uma taxa de aprendizado por causa dessa calibração.

Este capítulo apresentou em detalhe o algoritmo e matemática por trás de cada método de agregação. Para a AMA, (2.5) mostra como o modelo é estimado, e (2.18), como essa matriz é filtrada; (2.19) mostra como o parâmetro da filtragem é calculado. Para o Ridge, Lasso e MPE, as equações (2.37), (2.38) e (2.39) mostram como a matriz de pesos é obtida, nessa ordem. A discussão no início da seção 2.2 explica como o hiperparâmetro desses métodos são encontrados a cada passo de tempo.

Em relação aos objetivos deste manuscrito, a discussão sobre métodos da literatura feita até aqui é suficiente, e pode-se passar para o próximo capítulo, de apresentação dos reservatórios.

3 Reservatórios sintéticos

As simulações de validação serão feitas sobre dois modelos de reservatórios sintéticos, hoje bem conhecidos na literatura: o EGG e o OLYMPUS. Cada modelo contém múltiplas realizações de um reservatório fictício, cujos alguns atributos estruturais ou geológicos diferem entre si para, no caso de serem utilizados para balizar uma estratégia de previsão da produção, serem capazes de capturar as incertezas desse processo. Na Agregação, são essas diferenças que tornam as realizações em especialistas, ou pelo menos candidatos a especialistas.

A seguir, os detalhes de cada reservatório são apresentados. Para cada modelo, dá-se de uma breve contextualização; mostram-se algumas das propriedades geológicas, como geometria, viscosidade dinâmica, etc. e; comentam-se as características das realizações, como número de poços injetores e produtores. Ao final deste capítulo, comparam-se numa tabela os dois modelos nos quesitos relevantes à simulação, e. g., tempo de simulação e diferenças entre realizações dentro de um mesmo modelo.

3.1 OLYMPUS

OLYMPUS surgiu em 2018 para servir de ambiente de teste na competição de otimização de controles e gerência de reservatório conforme proposta pela ASIPP, Abordagem de Sistemas Integrados aplicada à Produção de Petróleo, do inglês *Integrated Systems Approach for Petroleum Production*. A descrição do OLYMPUS dada aqui é a mesma que foi apresentada em [Fonseca et al. \(2020\)](#), com adaptações onde necessário.

3.1.1 Caracterização

Trata-se de um reservatório sintético inspirado numa jazida de petróleo no Mar do Norte, com formato aproximadamente retangular de 9 km por 3 km e 50 m de profundidade dividida entre 16 camadas. Cada célula na malha do modelo dimensões aproximadas de 50 m x 50 m x 3 m. Existem ao todo 341.728 células, sendo 192.750 ativas.

O perímetro do reservatório é definido por uma falha; além desta, seis outras menores estão presentes no reservatório. O OLYMPUS consiste em duas zonas, separadas por uma camada impermeável de xisto (da qual advém a maioria das células inativas). A zona superior do reservatório contém areias de canal fluvial inseridas em folhelhos de planície de inundação, e a zona inferior do reservatório consiste em camadas alternadas de areias grossas, médias e finas com profundidades determinadas

Foram modelados 4 tipos de fácies nas diferentes camadas. Uma visão geral dos diferentes tipos de fácies nas diferentes zonas é fornecida na [Tabela 3.3](#). Propriedades geológicas, como porosidade, permeabilidade e razão Líquido-Bruto (LB), foram geradas utilizando técnicas geo-estatísticas padrão para os diferentes tipos de fácies. Supõe-se que dados insuficientes estão disponíveis nas fases iniciais do desenvolvimento do campo, de forma que nenhuma razão porosidade-permeabilidade foi utilizada. As permeabilidades nas direções x e y são idênticas, enquanto que a em z corresponde a 10% da em x .

Tabela 3.3 – Tipos e intervalos de distribuição de propriedades das fácies nos modelos do OLYMPUS.

Tipo	Zonas de presença	LB	Porosidade	Permeabilidade (mD)
Areia de canal	Superior	0.8-1	0.2-0.35	400-1000
Xisto	Superior e perímetro	0	0.03	1
Areia grossa	Inferior	0.7-0.9	0.2-0.3	150-400
Areia média	Inferior	0.75-0.95	0.1-0.2	75-150
Areia fina	Inferior	0.9-1	0.05-0.1	10-50

Fonte: Adaptado de [Fonseca et al. \(2020, tabela 1\)](#).

Com base nos registros de poços de exploração disponíveis, determinou-se que a profundidade do contato óleo-água é de 2092 m, com uma pressão hidrostática *in situ* de 205 bar. A distribuição inicial de saturação de água é modelada utilizando curvas de pressão capilar. Diferentes curvas de pressão capilar foram atribuídas aos diferentes tipos de fácies. Assim, as alterações no modelo de fácies resultarão em diferentes distribuições iniciais de saturação de água. As propriedades do fluido, similar ao petróleo sem gás dissolvido, usadas no conjunto de modelos são fornecidas na [Tabela 3.4](#). A viscosidade do óleo é modelada como uma função da pressão, logo, uma faixa é fornecida na [Tabela 3.4](#). Da mesma forma, o fator de volume de formação do óleo também é fornecido como função da pressão.

Tabela 3.4 – Propriedades dos fluidos nos modelos do OLYMPUS.

Grandeza	Valor	Unidade
Compressibilidade da rocha	$1,42 \times 10^{-5}$	1/bar
Compressibilidade da água	$3,97 \times 10^{-5}$	1/bar
Viscosidade dinâmica do óleo	2,8 – 3,5	cP
Viscosidade dinâmica da água	0.398	cP
Diâmetro do poço	0,19	m
Tempo de simulação	7200	dia

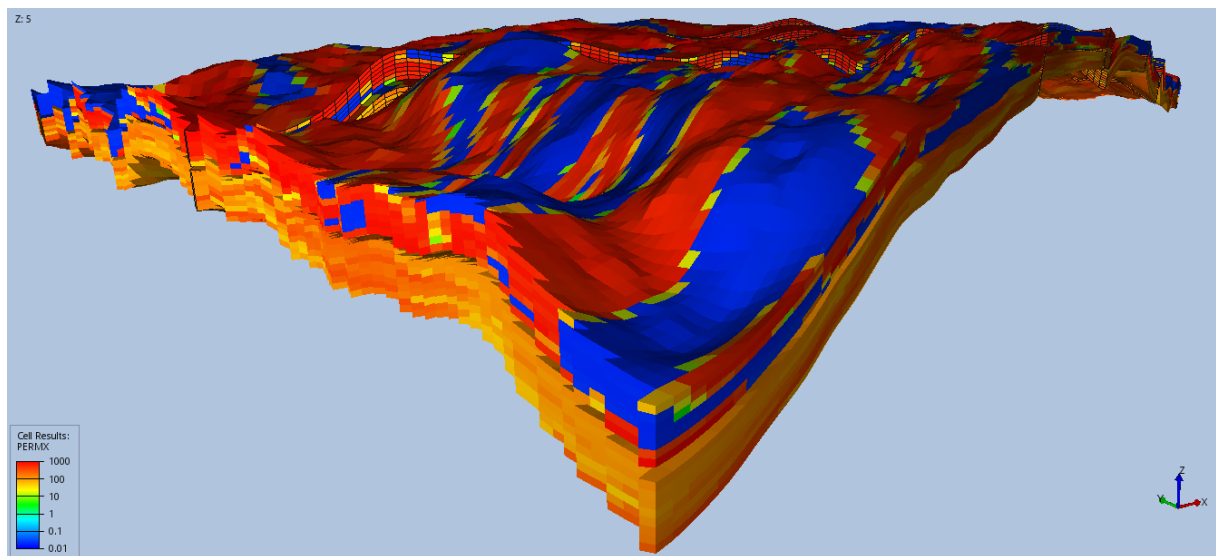
Fonte: Adaptado de [Fonseca et al. \(2020, tabela 2\)](#).

3.1.2 Realizações

Um conjunto de 50 realizações foi gerado em que as fácies são redefinidas alterando a semente aleatória. A geometria da malha, das falhas e o contato óleo-água são os mesmos em todas as realizações. As propriedades incertas são as fácies, porosidade, permeabilidade, razão LB, distribuição inicial da saturação de água, e transmissibilidade no interior das falhas. Eventualmente, algumas realizações apresentam mais passos de tempo, mas a diferença entre a menor quantidade de passos para a maior é de 2. O agregado de modelos foi gerado usando o seguinte procedimento. Um modelo caso-base de alta fidelidade, com aproximadamente 5 milhões de células da grade, foi gerado no primeiro passo. Cinco poços foram perfurados neste modelo, e registros sintéticos foram gerados para cada um desses poços. Esses registros foram então usados para limitar a geração do conjunto de 50 modelos de alta fidelidade, a fim de capturar a incerteza. Cada um desses modelos foi escalonado para fins de simulações de fluxo usando o método de escalonamento baseado em fluxo. Cada realização possui 18 poços ao todo, sendo 11 produtores e 7 injetores.

As figuras abaixo geradas pelo programa de código livre ResInsight mostram a permeabilidade em x de algumas realizações. Nelas, as cores estão em escala logarítmica, e o eixo z foi exagerado em cinco vezes.

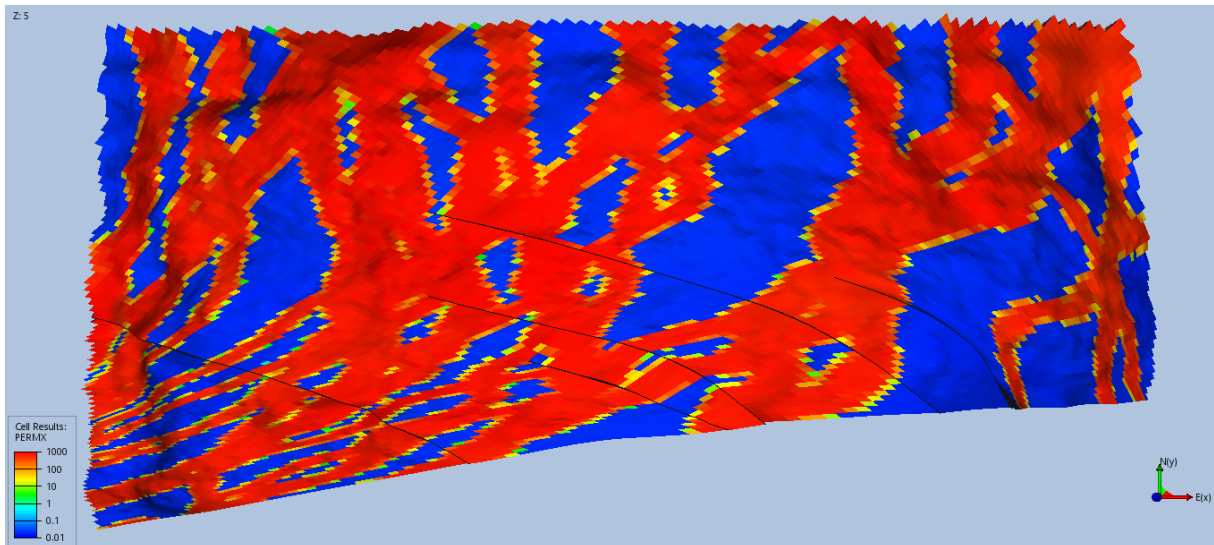
Figura 3.6 – Vista oblíqua da realização 01 do reservatório sintético OLYMPUS.



Fonte: Autores.

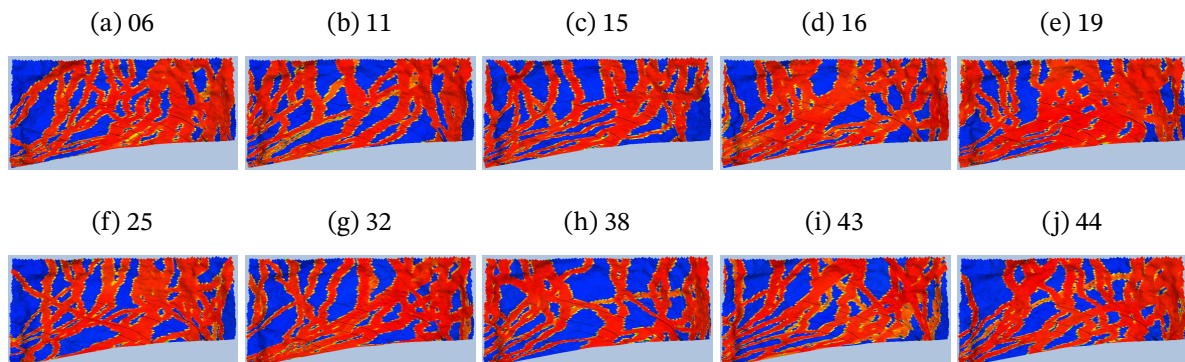
A competição proposta pela ASIPP consiste em três problemas distintos. O primeiro pede otimização da produção através do controle de poços previamente definidos, de iguais posições e direções em cada realização. São os modelos desse problema que foram considerados nas simulações. Os poços produtores do OLYMPUS foram nomeados “PROD- X ”, onde $X \in \{1, 2, \dots, 11\}$.

Figura 3.7 – Vista superior da realização 01 do reservatório sintético OLYMPUS.



Fonte: Autores.

Figura 3.8 – Algumas realizações do reservatório sintético OLYMPUS.



Fonte: Autor.

3.2 EGG

Jansen et al. (2014) buscaram fornecer uma versão unificada do EGG, na forma de características e dados, como será visto a seguir.

3.2.1 Caracterização

Foi desenvolvido como parte da tese de doutorado de Maarten Zandvliet e Gijs van Essen, onde apenas um modelo determinístico foi utilizado. Zandvliet et al. (2007) é a primeira publicação que faz uso do EGG. Desde então, uma versão com agregado de modelos vem sendo usada em vários trabalhos, e. g., Essen et al. (2009), ao passo que o modelo determinístico vem servindo como ambiente de testes para o controle e gestão de reservatórios.

As realizações consistem em campos de permeabilidade discretos numa grade 60 x 60 x 7 com 18.853 células ativas. As não ativas dentro do montante total de 25.200 delineiam o reservatório, atribuindo-o formato oval. Cada célula tem dimensões 8 m x 8 m x 4 m.

Cada um dos campos de permeabilidade em cada uma das sete camadas foi desenhado à mão usando um programa de desenho assistido por computador. Os canais de alta permeabilidade dentro de áreas de baixa permeabilidade representam padrões típicos de rios sinuosos encontrados em ambientes fluviais.

Cada campo é único e, como resultado, a permeabilidade em cada uma das localizações das células pode ser descrita por uma distribuição de probabilidade. No entanto, não existe um modelo matemático subjacente para criar campos adicionais de permeabilidade. Os campos exibem uma orientação clara dos canais com uma distância e sinuosidade de canal típicas. Os valores de permeabilidade não foram estabelecidos nos poços, e a porosidade é assumida como constante.

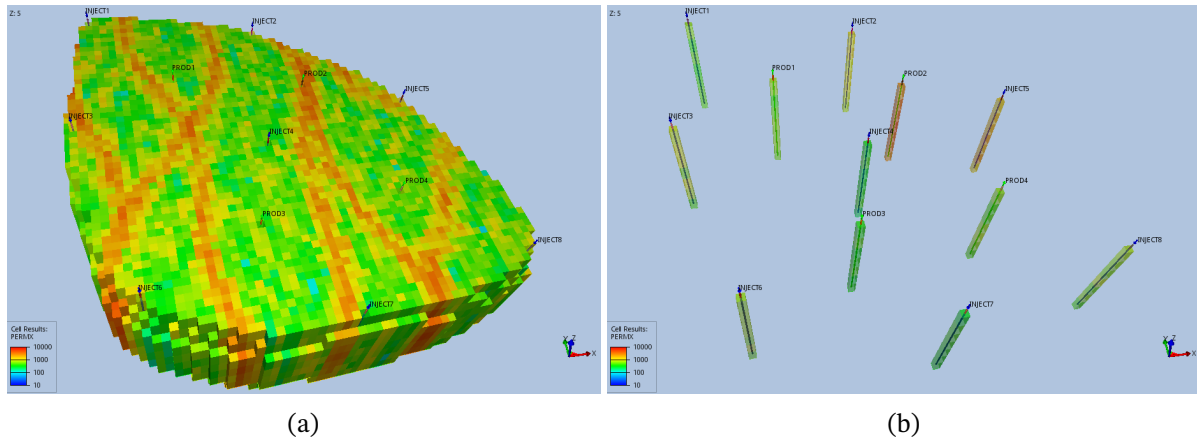
As sete camadas têm forte correlação vertical, de modo que os campos de permeabilidade são praticamente bidimensionais. Na maioria das publicações, o EGG foi usado para simular fluxo bifásico óleo-água. Como o modelo não possui um aquífero nem uma capa de gás, a produção primária é quase insignificante e o mecanismo de produção é a injeção de água com a ajuda de oito poços de injeção e quatro poços de produção.

3.2.2 Realizações

O modelo determinístico junto ao agregado criam um conjunto de 101 realizações de permeabilidade que, ao apresentar as prioridades de fluxo e reservatório na [Tabela 3.5](#), formam o modelo EGG padrão. Cada realização possui 4 poços produtores e 8 injetores, todos verticais.

A [Figura 3.9](#) mostra renderização da permeabilidade no eixo x do modelo EGG base, assim como a configuração espacial dos poços. Cores foram adotadas em escala logarítmica. Lembrando que, por construção, a permeabilidade em y é igual à em x , e que a permeabilidade em z é um décimo daquela em x , a distribuição de cores será a mesma não importando a direção em análise da permeabilidade. Para melhor visualização, o eixo z foi ampliado em cinco vezes. Para gerar as renderizações, usou-se o programa de código livre ResInsight.

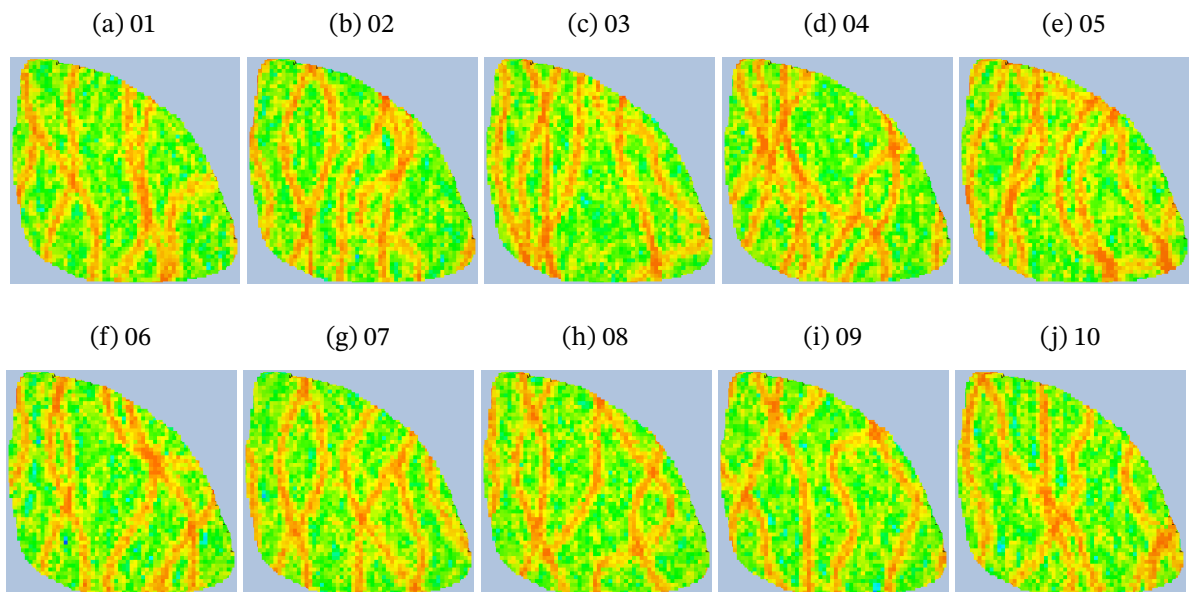
Figura 3.9 – (a) Malha 3D da permeabilidade ao longo do eixo x do modelo EGG base. (b) Distribuição espacial dos poços produtores e injetores do EGG.



Fonte: Autor.

Os poços injetores estão distribuídos ao longo do perímetro da malha, de forma a percolar o óleo presente no reservatório em direção aos poços produtores, que se encontram próximos ao centro, estimulando a produção. A [Figura 3.10](#) mostra, nas mesmas configurações da figura anterior, as renderizações de permeabilidade em x de algumas realizações do conjunto.

Figura 3.10 – Realizações 01 a 10 do reservatório sintético EGG.



Fonte: Autor.

A [Tabela 3.6](#) compara as características mais relevantes entre o EGG e o OLYMPUS no que tange os esforços de simulação importantes para os objetivos gerais.

Tabela 3.5 – Algumas propriedades de geometria e do fluido nos modelos do EGG.

Grandeza	Valor	Unidade
Compressibilidade da rocha	0	1/bar
Compressibilidade da água	$1,0 \times 10^{-5}$	1/bar
Viscosidade dinâmica do óleo	5	cP
Viscosidade dinâmica da água	1	cP
Diâmetro do poço	0,2	m
Tempo de simulação	3600	dia

Fonte: Adaptado de [Jansen et al. \(2014, tabela 1\)](#).

Tabela 3.6 – Comparação entre EGG e o OLYMPUS.

Tópico	EGG	OLYMPUS
Tempo de simulação	3600 dias	7380 dias
Número de produtores	4	11
Número de injetores	8	7
Diferenças entre realizações	Geometria e características das vias permeáveis: distribuição de saturação inicial, distribuição de permeabilidade	Saturação inicial, fácies, porosidade, permeabilidade, razão LB, transmissibilidade nas falhas, quantidade de passos de tempo

Fonte: Adaptado de [Jansen et al. \(2014, tabela 1\)](#) e [Fonseca et al. \(2020, tabela 2\)](#).

Esclarecidos os reservatórios de validação, é hora de mostrar e comparar os resultados das simulações de cada método, função do capítulo subsequente.

4 Resultados e análise

Os capítulos 2 e 3 apresentaram respectivamente os algoritmos de agregação de modelos e os reservatórios sintéticos que servirão para testá-los. Esta seção exibirá os resultados obtidos ao longo desses testes, a começar pela apresentação dos valores das variáveis da AMA que foram fixados durante as simulações.

A efeito de simulação, considerou-se na AMA $c_1 = 1/N^2$, $c_2 = 1$, $v_{\text{OPR}} = 45$ e $v_{\text{WPR}} = 8$. Essas constantes foram determinads de maneira empírica, com o objetivo e destacar o erro individual e a produção de óleo em (2.5). Adotou-se também porcentagem de ruído igual a 5% associada a $\sigma_{\% \text{OPR}}$ e $\sigma_{\% \text{WPR}}$.

A cada passo de tempo da simulação, todos os métodos de agregação citados guardam a matriz de pesos gerada numa estrutura de dados. Terminada a simulação, essa estrutura terá uma matriz de pesos para cada passo de tempo exceto o inicial. Isso ocorre porque ou o método de agregação ou sempre realiza a previsão do passo seguinte, como é o caso da MPE, Ridge e Lasso, ou precisa de pelo menos dois pontos de observação passados num dado instante da simulação, como é o caso da AMA (Equação 2.5).

Dito isso, uma vez preenchida o armazenador de matrizes, a previsão agregada pode ser feita com o modelo de qualquer passo de tempo, conforme a equação abaixo,

$$\hat{\mathbf{y}}_t = \text{DP}(\mathbf{W}_\tau^T \mathbf{M}_t) \quad (4.1)$$

onde τ é o passo de tempo do modelo escolhido. Notadamente, a equação acima não passa de uma adaptação de (2.3) sem a restrição de igualar o passo de tempo da matriz de estimativas ao da matriz de pesos.

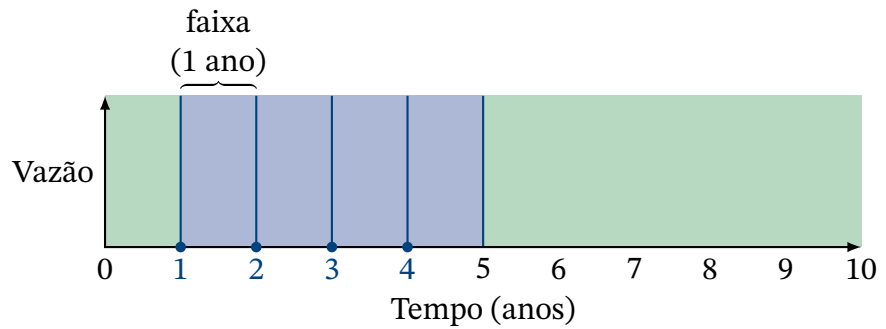
Nos resultados que seguem, para cada reservatório foram fixados 4 valores de passo de tempo igualmente espaçados entre si, e o modelo de cada valor formulou previsões durante cerca de 1 ano.

Para o EGG, cujo tempo foi medido em anos, os modelos escolhidos foram o do ano 1, 2, 3 e 4. Lembra-se ao leitor que esse reservatório possui 10 anos de vida. Foca-se aqui na primeira metade desse período porque, para os últimos 5 anos, a produção tende a ser mínima, e portanto não só o interesse em estimar a produção futura diminuiria, mas também os valores de erro relativo poderiam sofrer instabilidade numérica.

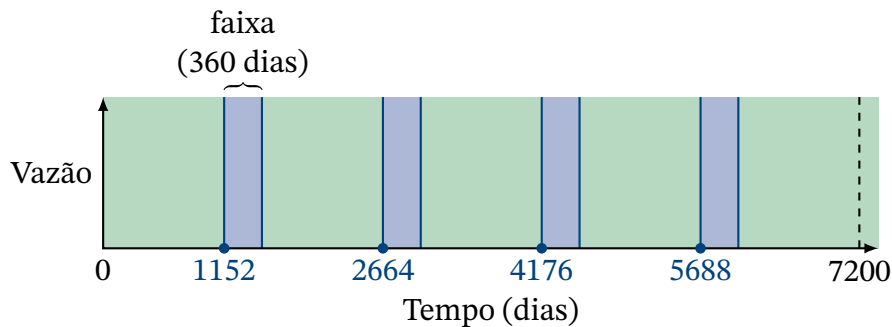
Para o OLYMPUS, cujo tempo foi medido em dias, escolheram-se modelos igualmente espaçados num período de 7200 dias. Esse espaçamento leva em consideração os períodos de previsões de 360 dias. O OLYMPUS possui vida um pouco mais longa que 7200 dias, mas este número é mais compatível com o período de previsão estipulado. A Figura 4.11 ilustra

as faixas de previsão de cada reservatório ao longo do tempo. Áreas em verde e azul são zonas sem e com previsão, nessa ordem. Os instantes de início de uma janela de previsões são números destacados em azul.

Figura 4.11 – Faixas de previsão determinadas para o (a) EGG e (b) OLYMPUS.



(a)



(b)

Cada faixa ou janela de previsão forneceram aproximadamente 13 previsões agregadas por poço e por método. Ponto a ponto, essas previsões foram convertidas ao erro percentual (ou relativo) através da fórmula abaixo,

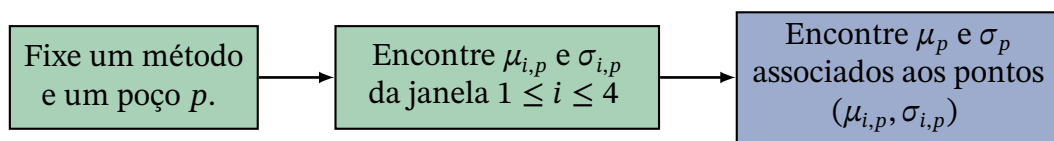
$$e_{\%t,p} = \left| \frac{\tilde{y}_{t,p} - \hat{y}_{t,p}}{\tilde{y}_{t,p}} \right| \times 100 \quad (4.2)$$

onde $\tilde{y}_{t,p}$ representa a real produção, de água ou óleo, do poço p no instante t . A relação entre $\tilde{y}_{t,p}$ e $y_{t,p}$ é que $y_{t,p}$ é o valor observado durante a medida de $\tilde{y}_{t,p}$, e portanto pode sofrer algum ruído de medição. Nas simulações, os valores de referência $\tilde{y}_{t,p}$ correspondem à produção de uma realização pré-determinada que não pertence ao conjunto de especialistas, e tanto água quanto óleo tiveram medidas observadas com ruído branco Gaussiano de desvio padrão igual a 5%.

Esclarecido o papel e origem de $\tilde{y}_{t,p}$, note que a operação em (4.2) deve ser feita para cada método, além de cada poço e passo de tempo na janela. A comparação entre a qualidade dos métodos de agregação se dará pela análise de seus erros percentuais poço a poço. Eis como esses valores podem ser obtido a partir dos erros relativos em cada janela.

Fixado um método de agregação, um poço e uma janela, é tomada a média e desvio padrão de seus erros percentuais, gerando um erro médio e um desvio padrão médio daquela janela para aquele poço. Em seguida, repete-se esse passo para cada janela. Ao fim dessa varredura, obtém-se 4 pontos de média e de desvio padrão, um para cada faixa. Na sequência, calcula-se a média e desvio padrão referentes a esses 4 pontos, o que resulta em um valor para o erro médio e desvio padrão médio do poço fixado. Finalmente, todo esse procedimento é repetido para os outros métodos de agregação e poços do reservatório, até que todos tenham sido considerados. A figura abaixo ilustra as contas descritas aqui de maneira resumida; μ e σ representam médias e desvios padrões.

Figura 4.12 – Sequência de operações para se obter o erro percentual de um poço.



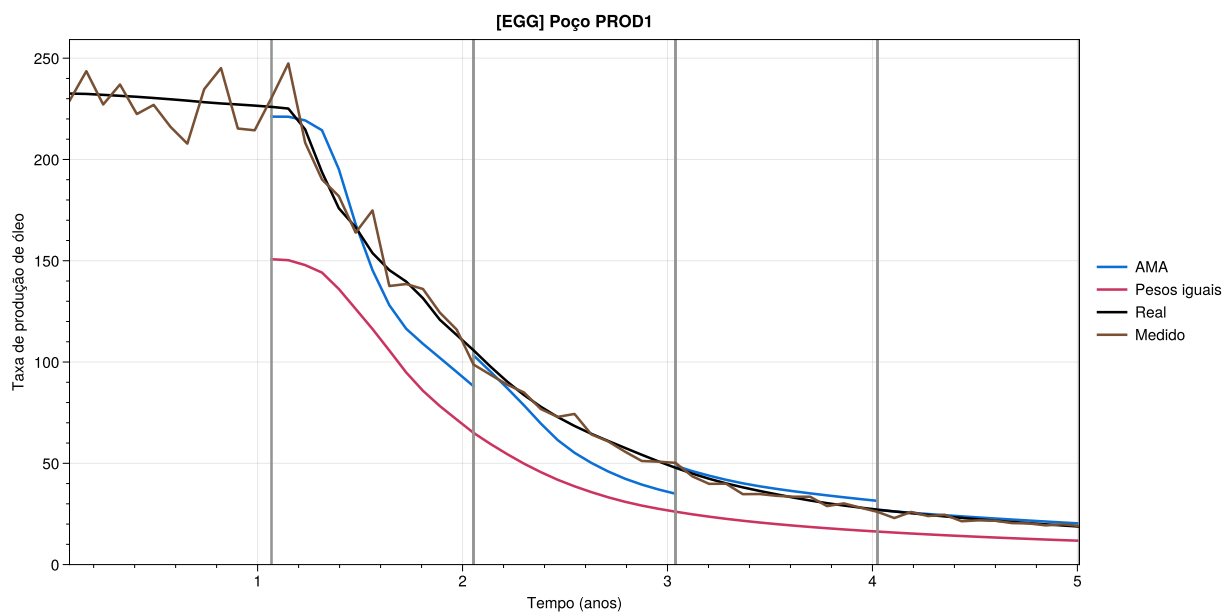
Esses valores foram tabelados para o EGG e OLYMPUS, e são apresentados nas próximas seções. Essas seções também explicitam as realizações utilizadas como referência e como especialistas, assunto omitido nesta introdução.

4.1 Resultados referentes ao EGG

Uma vez que as curvas de permeabilidade do EGG foram geradas de maneira randômica (vide a [subseção 3.2.1](#)), tomou-se as realizações 01 a 10 como especialistas, e a 11 como referência. Perceba que a malha de cada um desses especialistas foi mostrada na [Figura 3.10](#).

A figura [Figura 4.13](#) mostra a previsão de óleo do poço PROD1 quando feita pela AMA, em azul, e feita pela combinação equânime dos especialistas, em vermelho. Por combinação equânime, entende-se matriz de pesos com todos os valores iguais a $1/N$. A curva de produção medida, ou $y_{t,p}$, é marrom, e a curva de produção real, ou $\tilde{y}_{t,p}$, é preta. A previsão é feita no esquema de faixas explicado no início deste capítulo: ao encontrar uma barra vertical cinza, a matriz de pesos desse passo de tempo é fixada para gerar previsões agregadas até que seja satisfeito o período de 1 ano ou uma nova barra seja encontrada (para o EGG, essas condições são equivalentes).

Figura 4.13 – Previsões do óleo feitas pela AMA do 1º ao 5º ano do EGG



Está clara a superioridade da AMA frente à combinação equânime dos especialistas. As figuras 4.14 e 4.15 mostram as previsões agregadas de óleo e água, respectivamente, do poço PROD1 que serão consideradas no erro percentual. A curva de previsão da MPE, Lasso e Ridge são indicadas em violeta, rosa e laranja, respectivamente.

Figura 4.14 – Comparação de previsão de produção de óleo do 1º ao 5º ano do EGG.

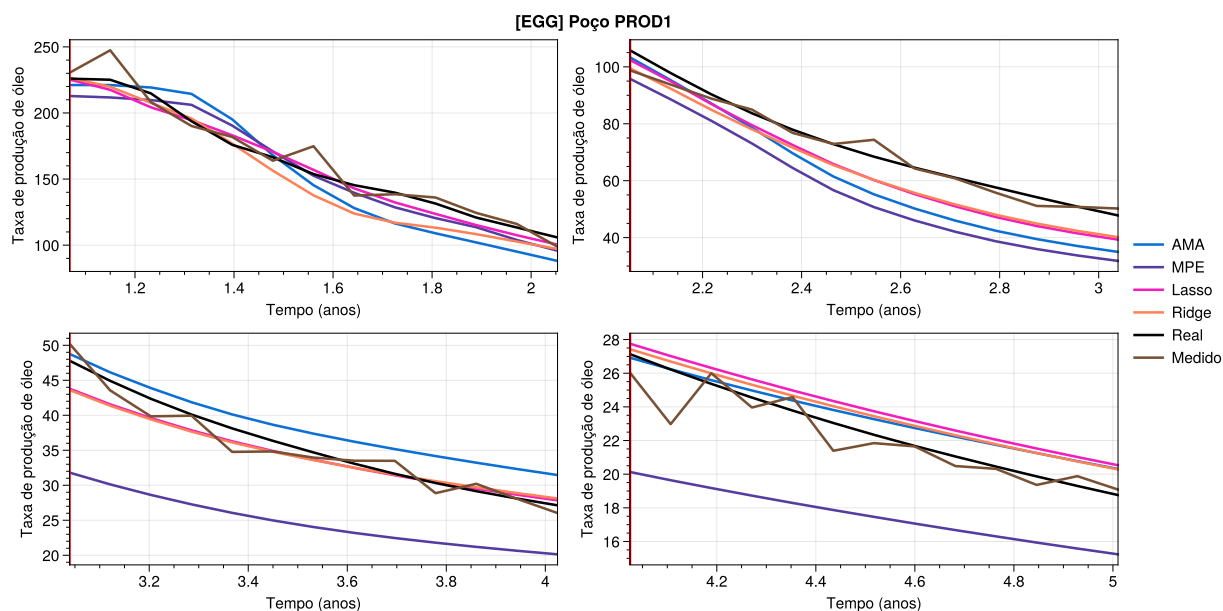
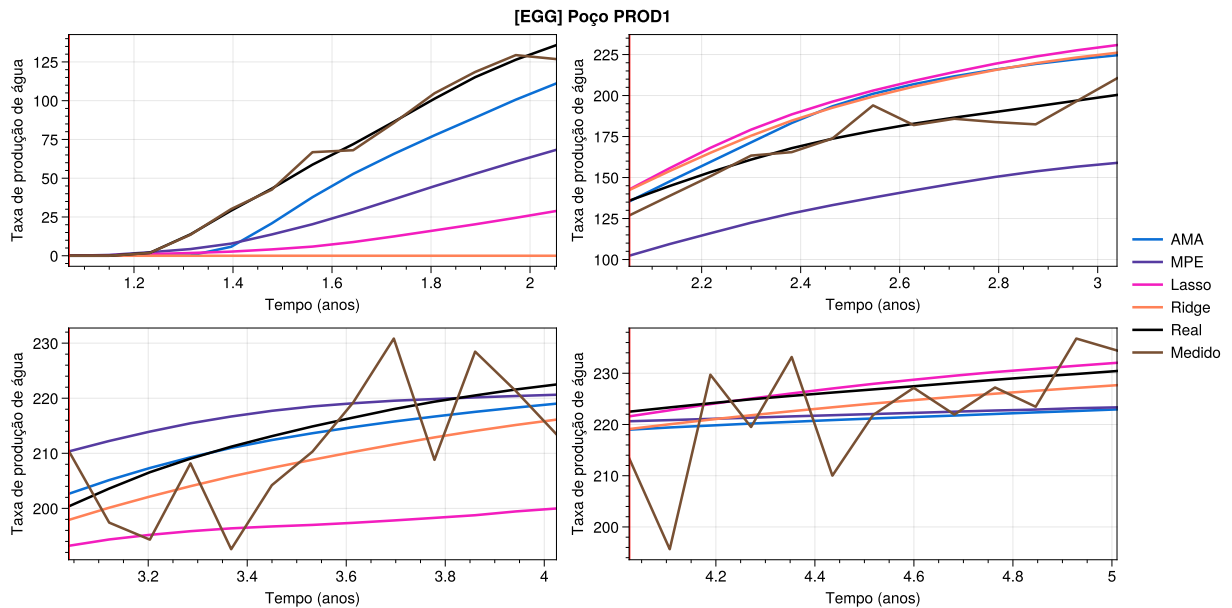


Figura 4.15 – Comparação de previsão de produção de água do 1º ao 5º ano do EGG.



Exemplificados os pontos considerados por meio dos gráficos acima, a tabela a seguir apresenta os erros encontrados para cada poço do EGG. Os melhores resultados de óleo e água, por poço, foram indicados em verde e azul, respectivamente. Para evitar divisão por 0 em (4.2), principalmente na água, só foram considerados para análise de erro pontos em que $\tilde{y}_{t,p} \geq 1$.

A AMA teve a melhor previsão de óleo no PROD4, e as melhores previsões de água no PROD1 e PROD2, colocando-a como uma das melhores estratégias de agregação na tabela. Ela perde apenas para o Lasso, que apresentou as melhores previsões de óleo para o PROD1, PROD2 e PROD3. Essas duas estratégias tiveram resultados próximos no PROD2 e PROD4, mas no PROD1 a AMA teve erro na água bem menor que o Lasso, enquanto que no PROD3, bem maior.

Os bons e maus resultados comparativos da AMA devem advir, majoritariamente, da ponderação feita por v_{OPR} e v_{WPR} em (2.5). Ao considerar na formulação do modelo tanto a produção de óleo quanto água, assume-se que essas produções estão atreladas, ou correlacionadas em algum grau. Se essa correlação for bem modelada por v_{OPR} e v_{WPR} , a AMA tenderia a conseguir um erro percentual razoavelmente baixo para um fluido sem entretanto inflar o erro no outro fluido, pelo menos em comparação com os outros métodos.

Esse fenômeno parece ocorrer por todos os poços, mas principalmente no PROD4, onde o erro no óleo conseguiu se manter baixo sem entretanto deixar crescer demais o erro na água. Em contrapartida, o Ridge, por exemplo, tem o melhor erro de água nesse poço (37,58% contra os 42,63% da AMA), mas seu erro em óleo é maior que o triplo da AMA.

Afim de validar o comentário acima, a Tabela 4.8 compara dados da AMA gerados

sob três diferentes condições de simulação: $(\nu_{\text{OPR}}, \nu_{\text{WPR}}) = (0, 1)$, $(\nu_{\text{OPR}}, \nu_{\text{WPR}}) = (45, 8)$, e $(\nu_{\text{OPR}}, \nu_{\text{WPR}}) = (1, 0)$. Todos os outros valores de ajuste permaneceram os mesmos aos citados anteriormente.

Essa tabela deixa claro que a proporção 45:8 gera melhores previsões de óleo, enquanto que para a TPA tenha obtido apenas uma das 4 melhores médias possíveis. Curiosamente, a ponderação voltada 100% para o óleo nunca obteve a melhor média nesse fluido. Para as combinações que consideram apenas um fluido, os maus resultados se devem, majoritariamente, à incapacidade em se atenuar a variância na medida do fluido em análise. A proporção 1:1 consegue atenuar na previsão final efeito de variância particularmente alta observada em um fluido, e por isso demonstrou melhores resultados que as proporções de fluido único. Por outro lado, a proporção 45:8, que visa modelar a importância econômica das duas proporções, obteve resultados ao todo melhores que os da proporção 1:1.

Tabela 4.7 – Erros na previsão de produção do 1º ao 5º ano do EGG.

Método	Fluido	Erro (%)			
		PROD1	PROD2	PROD3	PROD4
AMA	óleo	9,73 ± 3,20	6,32 ± 2,61	27,30 ± 16,50	10,70 ± 3,27
AMA	água	14,53 ± 7,59	15,33 ± 6,65	38,74 ± 1,50	42,63 ± 31,85
MPE	óleo	20,28 ± 2,63	12,90 ± 4,35	74,01 ± 11,06	24,51 ± 3,87
MPE	água	20,87 ± 3,50	82,03 ± 60,67	18,38 ± 3,66	63,70 ± 39,65
Lasso	óleo	6,16 ± 1,80	5,58 ± 1,68	23,33 ± 3,99	22,91 ± 4,57
Lasso	água	25,65 ± 3,39	15,52 ± 12,68	21,96 ± 2,20	41,08 ± 33,08
Ridge	óleo	6,98 ± 1,94	13,17 ± 3,29	63,68 ± 21,18	34,93 ± 6,31
Ridge	água	28,63 ± 0,72	36,02 ± 2,66	30,80 ± 0,96	37,58 ± 2,27

Fonte: Autor.

Tabela 4.8 – Erros da AMA no EGG para diferentes ajustes em $\nu_{\text{OPR}}/\nu_{\text{WPR}}$.

$\nu_{\text{OPR}} : \nu_{\text{WPR}}$	Fluido	Erro (%)			
		PROD1	PROD2	PROD3	PROD4
45:8	óleo	9,73 ± 3,20	6,32 ± 2,61	27,30 ± 16,50	10,70 ± 3,27
45:8	água	14,53 ± 7,59	15,33 ± 6,65	38,74 ± 1,50	42,63 ± 31,85
1:1	óleo	7,98 ± 2,75	6,46 ± 2,73	40,01 ± 14,19	12,97 ± 3,31
1:1	água	13,41 ± 7,76	18,63 ± 6,78	25,88 ± 1,93	38,57 ± 31,15
1:0	óleo	9,79 ± 3,41	9,78 ± 3,07	28,76 ± 17,51	12,79 ± 3,35
1:0	água	24,48 ± 7,49	36,05 ± 13,51	63,47 ± 1,32	105,92 ± 32,05
0:1	óleo	32,82 ± 3,20	28,76 ± 5,00	155,96 ± 24,13	54,67 ± 6,96
0:1	água	9,61 ± 12,18	44,98 ± 26,54	13,30 ± 8,51	117,84 ± 65,96

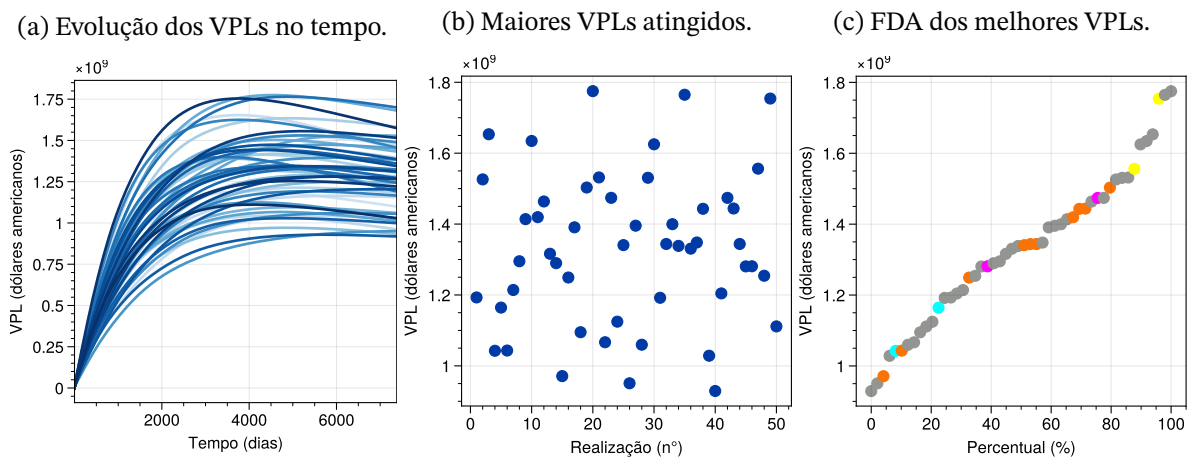
Fonte: Autor.

4.2 Resultados referentes ao OLYMPUS

Nos resultados do OLYMPUS, ao contrário do EGG, 6 referências distintas foram utilizadas para fins de análise. Dessa forma, foi feito um conjunto de previsões para cada referência, sempre com o mesmo grupo de realizações.

Como métrica de escolha dos reservatórios, utilizou-se o lucro acumulado na venda do óleo produzido pelo reservatório menos o custo na produção de água, ao longo da vida do reservatório. Essa grandeza é denominada de Valor Presente Líquido (VPL), do inglês *Net Present Value*. O VPL é calculado a partir da soma dos lucros obtidos a cada passo de tempo no período de produção. No início desse período, a TPO é relativamente grande enquanto a TPA é pequena, de forma que os lucros pontuais são positivos. Conforme o óleo é extraído, a TPO cai e a TPA aumenta, diminuindo os lucros pontuais. Sendo o período de produção longo o suficiente, no instante em que esses pontos deixam de ser positivos e passam a ser negativos (ou seja, começa a existir um prejuízo na operação de extração), o VPL atinge um máximo, e decai a partir daí. A [Figura 4.16a](#) expõe como o VPL de cada realização evolui ao longo do tempo de vida do OLYMPUS. Como comentado no capítulo anterior, todas as 50 realizações do OLYMPUS consideradas no manuscrito não experimentam aplicação de uma estratégia de otimização, ou seja, não há a abertura e fechamento de poços.

Figura 4.16 – Realizações do OLYMPUS e VPLs associados.



Fonte: Autor.

Similarmente ao feito por [Pinto et al. \(2020\)](#), a seleção de especialistas e referências se embasou na ordenação dos VPLs. Dentro do tempo de vida do reservatório, tomou-se o melhor VPL de cada realização, e organizou-se esse conjunto de pontos numa função distribuição acumulada (FDA), vide [Figura 4.16c](#).

Os especialistas foram escolhidos de forma a abranger um grande intervalo de VPLs. Por sua vez, as referências foram escolhidas de forma que, em grupos de 2, um grupo ficasse

com VPLs abaixo dos VPLs da maioria dos especialistas; um grupo ficasse com VPLs acima, e; um grupo ficasse com VPLs razoavelmente próximos.

O grupo de especialistas foi denominado Grupo 1, e cobre intervalo entre 5% e 80%. Destacado em laranja na imagem, corresponde às realizações 16, 15, 25, 06, 44, 32, 11, 43, 19, 38, do menor percentual ao maior. O leitor é convidado a revisitar as malhas dessas realizações que foram apresentadas na [Figura 3.8](#).

Destacado em ciano na [Figura 4.16c](#), o grupo de referências de baixa produção, denominado Grupo 2, está entre os 10% e os 20%, e corresponde às realizações 04 e 05, do menor percentual ao maior. A referência 04 será neste manuscrito chamada também de OLYMPUS A, e a 05, OLYMPUS B.

Destacado em amarelo na [Figura 4.16c](#), o grupo de referências de alta produção, denominado Grupo 3, está próximo dos 95%, e corresponde às realizações 49 e 47, do menor percentual ao maior. A referência 47 será neste manuscrito chamada também de OLYMPUS C, e a 49, OLYMPUS D.

Destacado em magenta na [Figura 4.16c](#), o grupo de referências de produção média, denominado Grupo 2, está entre os 40% e 80%, e corresponde às realizações 42 e 46, do menor percentual ao maior. A referência 42 será neste manuscrito chamada também de OLYMPUS E, e a 46, OLYMPUS F.

As demais realizações aparecem em cinza e não foram utilizadas nas simulações. Como foi feito para o EGG, aqui as figuras [4.17](#) e [4.18](#) ilustram alguns dos pontos de previsão considerados. A referência é o OLYMPUS E, e o poço, PROD-1. O eixo do tempo é agora medido em dias ao invés de anos. O esquema de cores continua o mesmo de antes.

Figura 4.17 – Previsões do óleo feitas por todos os métodos para 5 anos de vida do OLYMPUS.

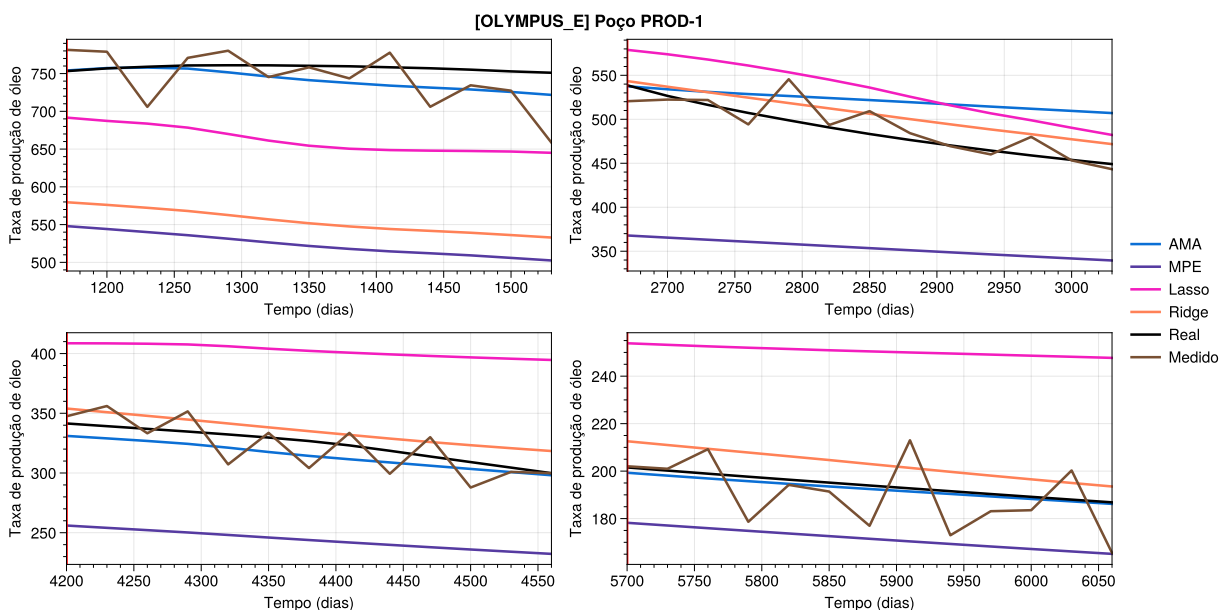
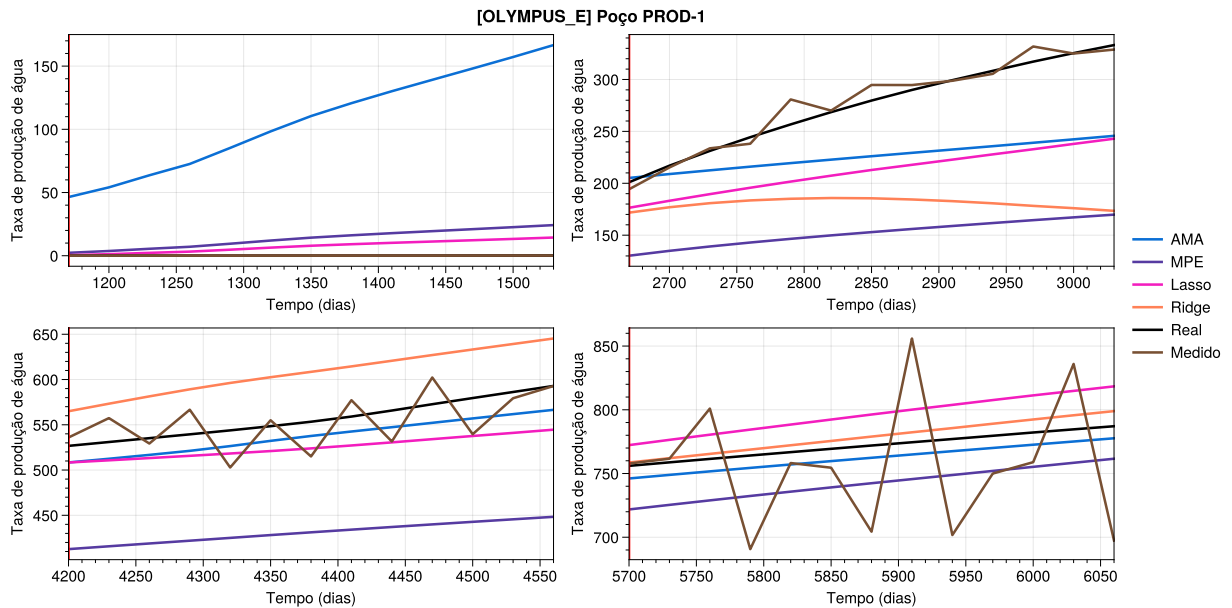


Figura 4.18 – Previsões da água feitas por todos os métodos para 5 anos de vida do OLYMPUS.



Uma vez que o OLYMPUS possui muito mais poços que o EGG, foi necessário quebrar as tabulações de erro por método de agregação, resultando nas tabelas 4.9 a 4.20. Cada subseção a seguir apresenta as tabelas associadas a cada referência.

Similar ao adotado na análise do EGG, destacou-se, para cada poço, o menor erro no óleo em verde, e o menor erro na água em azul. Diferentemente do EGG, o OLYMPUS apresenta poços que não geram vazão significativa de óleo e/ou água durante larga porção de sua vida. Por causa disso, aumentou-se o limiar na Equação 4.2 para 10, ou seja, em cada janela foram considerados apenas os pontos que satisfaziam a desigualdade $\tilde{y}_{t,p} \geq 10$. Por exemplo, a primeira janela na Figura 4.18 não seria considerada para análise de erro. Na situação em que todas as faixas de previsão não superam o limiar, o erro e desvio padrão do poço não são calculados. Nas tabelas, esse fenômeno é indicado por um traço horizontal. A título de transparência e curiosidade, o Apêndice B as consequências de um limiar baixo. Exibem-se tabelas resultados para os quais o limiar usado no EGG não fora ajustado, isto é, permaneceu em 1. Isso leva ao aumento geral dos erros relativos, principalmente na produção de água nos poços PROD-2 e PROD-3 do OLYMPUS F, que com limiar 10 recebem traço horizontal. As figuras A.21 e A.22 mostram os perfis de produção desses poços, assim como as previsões feitas sobre eles.

Em termos de estabilidade numérica do erro, outros poços mais mal comportados foram o PROD-2 do OLYMPUS A (tanto produção de óleo quanto água) e o PROD-6 do OLYMPUS B (produção de água). As figuras A.19 e A.20 no Apêndice A mostram como os métodos de agregação tentam lidar com a ausência de produção.

Com isso dito, a AMA demonstrou dominância na qualidade de previsões de óleo

para todas as referências, tendo no mínimo 6 (OLYMPUS A) e no máximo 11 (OLYMPUS F) dos melhores erros para esse fluido. O mesmo, entretanto, não pode ser dito para a água; no pior caso foi obtido apenas um dos melhores erros (OLYMPUS C), e no melhor caso (OLYMPUS E) foram obtidas apenas 4 dos 11 possíveis.

Dados os baixos erros no óleo, é possível que a AMA tenha feito uma espécie de *overfitting* no óleo em detrimento da água. Em outras palavras, a razão $v_{\text{OPR}}/v_{\text{WPR}}$ é demasiadamente grande, e poderia ser diminuída de forma a melhorar o erro nesse poço (e referência) sem entretanto perder qualidade na previsão de óleo.

Este ajuste em $v_{\text{OPR}}/v_{\text{WPR}}$ é notoriamente pontual, uma vez que, repetindo, a AMA demonstrou resultados competitivos no geral e em particular no óleo. Tendo em mente a análise de erro do EGG, é justo dizer que a AMA é, no mínimo, competitiva frente a outros métodos da agregação. Essa competitividade deve ser tão melhor quanto melhor ajustados forem v_{OPR} e v_{WPR} , mas, a princípio, ênfase no erro do óleo em detrimento da água (proporção 45:8) se demonstrou um bom ponto de partida.

O leitor interessado pode consultar o [Apêndice C](#) para ver que, de fato, essa proporção tende a ser melhor, *ceteris paribus*, que a combinação $(v_{\text{OPR}}, v_{\text{WPR}}) = (1, 1)$ na TPO, e ligeiramente pior na TPA. Não foram consideradas as combinações $(v_{\text{OPR}}, v_{\text{WPR}}) = (1, 0)$ e $(v_{\text{OPR}}, v_{\text{WPR}}) = (0, 1)$ uma vez que, como discutido na análise de resultados do EGG, essas ponderações tendem a ampliar variâncias.

Por fim, vale comentar que a velocidade computacional da AMA é alta se comparado à do Lasso e Ridge, que realizam minimização considerando toda a produção passada com respeito ao instante da minimização. A abordagem adaptativa para encontrar o hiperparâmetro ótimo também introduz carga computacional significativa na simulação. Por outro lado, mesmo com a carga advinda do cálculo do hiperparâmetro, a MPE tem maior velocidade que a AMA, uma vez que, obtido o hiperparâmetro, seus pesos são obtidos de maneira analítica. Entretanto, como mostraram as análises de erro tanto do EGG quanto do OLYMPUS, a MPE não obteve bons resultados de maneira geral, talvez devido à sua convexidade. Logo, é justo afirmar que a AMA possa ser usada em estratégias de otimização de produção que requerem um custo benefício entre erro de previsão e custo computacional, desde que haja disposição para ajustar seus parâmetros $(c_1, c_2, v_{\text{OPR}}, v_{\text{WPR}}, \sigma_{\% \text{OPR}}, \sigma_{\% \text{WPR}}, f_{\sigma})$.

Tabela 4.9 – Erros de previsão para o OLYMPUS A (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	4,24 ± 0,46	3,81 ± 0,82
PROD-2	—	—
PROD-3	3,70 ± 0,83	100,69 ± 8,60
PROD-4	3,22 ± 0,06	2,65 ± 0,19
PROD-5	4,07 ± 0,41	3,83 ± 0,21
PROD-6	10,42 ± 1,26	0,75 ± 0,14
PROD-7	2,27 ± 0,29	2,58 ± 0,31
PROD-8	2,64 ± 0,09	2,80 ± 0,09
PROD-9	4,34 ± 0,40	14,36 ± 0,54
PROD-10	2,14 ± 0,18	20,93 ± 0,13
PROD-11	3,08 ± 0,62	4,11 ± 0,27

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	17,81 ± 0,30	7,81 ± 0,51
PROD-2	—	—
PROD-3	10,16 ± 0,80	26,01 ± 4,45
PROD-4	4,31 ± 0,35	6,75 ± 0,29
PROD-5	4,85 ± 0,41	3,03 ± 0,28
PROD-6	114,55 ± 3,06	10,48 ± 0,23
PROD-7	1,90 ± 0,12	3,23 ± 0,47
PROD-8	8,87 ± 0,24	9,83 ± 0,96
PROD-9	43,76 ± 0,56	22,94 ± 1,09
PROD-10	45,35 ± 0,25	126,11 ± 28,17
PROD-11	18,98 ± 1,22	12,16 ± 0,81

Fonte: Autor.

Tabela 4.10 – Erros de previsão para o OLYMPUS A (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	4,90 ± 0,41	13,33 ± 1,54
PROD-2	—	—
PROD-3	12,94 ± 1,09	18,31 ± 0,97
PROD-4	2,24 ± 0,22	3,75 ± 0,65
PROD-5	2,87 ± 0,81	2,85 ± 0,37
PROD-6	16,75 ± 0,68	2,39 ± 0,50
PROD-7	2,91 ± 0,44	2,40 ± 0,77
PROD-8	11,54 ± 1,44	5,48 ± 1,33
PROD-9	12,06 ± 2,67	11,36 ± 3,44
PROD-10	5,19 ± 0,19	13,21 ± 8,94
PROD-11	13,52 ± 0,80	38,79 ± 21,80

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	11,87 ± 1,07	8,95 ± 0,96
PROD-2	—	—
PROD-3	9,29 ± 0,93	22,48 ± 1,66
PROD-4	2,10 ± 0,06	1,77 ± 0,60
PROD-5	3,38 ± 0,49	2,42 ± 0,66
PROD-6	17,83 ± 3,51	1,91 ± 0,20
PROD-7	1,28 ± 0,34	3,51 ± 0,76
PROD-8	11,60 ± 2,09	4,91 ± 0,73
PROD-9	17,00 ± 4,53	22,08 ± 6,96
PROD-10	3,16 ± 0,93	30,09 ± 9,33
PROD-11	12,25 ± 1,10	44,11 ± 23,09

Fonte: Autor.

Tabela 4.11 – Erros de previsão para o OLYMPUS B (AMA e MPE)

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	3,79 ± 1,34	16,35 ± 0,31
PROD-2	7,36 ± 3,31	20,39 ± 4,46
PROD-3	3,01 ± 0,50	4,22 ± 0,35
PROD-4	2,03 ± 0,27	2,40 ± 0,55
PROD-5	6,22 ± 0,20	11,06 ± 0,31
PROD-6	39,30 ± 1,92	—
PROD-7	6,06 ± 2,40	44,93 ± 3,42
PROD-8	3,14 ± 0,30	9,57 ± 0,24
PROD-9	1,04 ± 0,39	5,13 ± 0,85
PROD-10	7,01 ± 2,43	19,74 ± 7,83
PROD-11	4,18 ± 0,56	9,91 ± 3,01

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	29,94 ± 1,84	20,11 ± 0,88
PROD-2	58,56 ± 3,14	44,43 ± 1,11
PROD-3	16,42 ± 0,41	12,84 ± 0,21
PROD-4	13,74 ± 0,23	7,17 ± 0,35
PROD-5	3,54 ± 0,22	5,84 ± 0,24
PROD-6	348,11 ± 10,31	—
PROD-7	27,42 ± 0,70	4,59 ± 0,39
PROD-8	21,66 ± 0,33	37,91 ± 0,36
PROD-9	18,44 ± 0,73	59,88 ± 0,42
PROD-10	22,87 ± 2,73	19,39 ± 1,40
PROD-11	11,63 ± 0,93	17,33 ± 4,07

Fonte: Autor.

Tabela 4.12 – Erros de previsão para o OLYMPUS B (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	7,51 ± 0,72	3,81 ± 0,75
PROD-2	27,94 ± 1,28	26,47 ± 1,70
PROD-3	4,17 ± 0,24	4,57 ± 1,71
PROD-4	0,76 ± 0,16	0,88 ± 0,20
PROD-5	3,89 ± 0,26	2,28 ± 0,54
PROD-6	121,31 ± 3,39	—
PROD-7	12,24 ± 0,53	2,39 ± 0,70
PROD-8	6,33 ± 0,38	5,15 ± 0,86
PROD-9	12,92 ± 1,52	19,36 ± 6,93
PROD-10	23,55 ± 5,60	14,62 ± 4,39
PROD-11	6,05 ± 2,41	14,20 ± 9,91

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	8,19 ± 1,16	2,29 ± 0,79
PROD-2	22,62 ± 3,65	30,57 ± 2,04
PROD-3	6,26 ± 0,30	1,92 ± 0,26
PROD-4	4,98 ± 0,14	4,19 ± 1,14
PROD-5	10,67 ± 0,24	0,68 ± 0,25
PROD-6	31,76 ± 7,16	—
PROD-7	12,55 ± 0,66	1,70 ± 0,61
PROD-8	7,34 ± 0,44	5,79 ± 1,86
PROD-9	15,58 ± 1,18	12,35 ± 2,21
PROD-10	15,37 ± 6,15	9,26 ± 1,93
PROD-11	6,78 ± 2,02	14,35 ± 6,60

Fonte: Autor.

Tabela 4.13 – Erros de previsão para o OLYMPUS C (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	8,49 ± 3,38	27,58 ± 0,22
PROD-2	4,08 ± 0,83	17,94 ± 0,80
PROD-3	1,62 ± 0,42	2,39 ± 0,20
PROD-4	4,43 ± 0,41	50,40 ± 10,93
PROD-5	2,81 ± 0,21	4,79 ± 0,24
PROD-6	2,82 ± 0,25	2,79 ± 0,06
PROD-7	2,83 ± 0,30	3,82 ± 0,16
PROD-8	4,61 ± 1,47	14,38 ± 2,01
PROD-9	7,14 ± 2,03	24,35 ± 17,92
PROD-10	2,08 ± 0,97	9,24 ± 1,60
PROD-11	2,36 ± 1,11	5,44 ± 0,69

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	38,84 ± 2,92	25,19 ± 1,25
PROD-2	7,57 ± 0,83	16,00 ± 1,58
PROD-3	23,40 ± 0,54	4,22 ± 0,17
PROD-4	15,87 ± 0,53	43,94 ± 0,45
PROD-5	5,35 ± 0,24	11,08 ± 0,24
PROD-6	45,57 ± 0,50	4,82 ± 0,20
PROD-7	15,11 ± 0,43	3,82 ± 0,42
PROD-8	18,58 ± 1,11	10,79 ± 1,99
PROD-9	6,20 ± 1,39	18,81 ± 7,30
PROD-10	28,73 ± 1,95	18,65 ± 2,92
PROD-11	57,80 ± 0,55	11,94 ± 2,14

Fonte: Autor.

Tabela 4.14 – Erros de previsão para o OLYMPUS C (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	13,87 ± 2,24	20,57 ± 1,66
PROD-2	19,00 ± 2,25	19,25 ± 5,83
PROD-3	13,27 ± 0,93	1,26 ± 0,13
PROD-4	7,79 ± 0,44	9,02 ± 1,53
PROD-5	6,60 ± 0,53	3,70 ± 0,75
PROD-6	32,45 ± 1,19	1,83 ± 0,07
PROD-7	11,88 ± 2,13	3,23 ± 2,14
PROD-8	3,82 ± 1,03	8,11 ± 1,65
PROD-9	25,17 ± 2,39	33,97 ± 24,72
PROD-10	6,37 ± 1,12	4,73 ± 2,13
PROD-11	20,21 ± 11,36	8,89 ± 2,24

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	14,68 ± 2,54	26,69 ± 0,99
PROD-2	17,13 ± 2,77	17,03 ± 5,74
PROD-3	14,68 ± 1,51	1,39 ± 0,12
PROD-4	6,05 ± 0,30	23,60 ± 1,31
PROD-5	3,40 ± 0,36	3,60 ± 0,82
PROD-6	31,01 ± 6,53	1,89 ± 0,05
PROD-7	14,46 ± 1,06	2,97 ± 1,43
PROD-8	4,02 ± 0,90	6,42 ± 1,55
PROD-9	20,88 ± 2,35	8,46 ± 3,19
PROD-10	7,63 ± 1,30	20,17 ± 7,95
PROD-11	8,81 ± 2,40	8,04 ± 1,53

Fonte: Autor.

Tabela 4.15 – Erros de previsão para o OLYMPUS D (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	2,60 ± 0,62	4,09 ± 1,86
PROD-2	13,89 ± 0,43	32,03 ± 6,10
PROD-3	4,26 ± 0,86	3,24 ± 0,48
PROD-4	2,19 ± 0,31	2,20 ± 0,11
PROD-5	2,92 ± 0,14	3,51 ± 0,09
PROD-6	3,35 ± 0,45	3,43 ± 0,07
PROD-7	42,70 ± 0,60	18,33 ± 0,58
PROD-8	3,30 ± 0,17	5,31 ± 0,23
PROD-9	2,29 ± 1,23	2,89 ± 0,49
PROD-10	4,25 ± 0,52	6,67 ± 1,92
PROD-11	6,75 ± 0,96	14,72 ± 3,58

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	9,92 ± 0,36	14,29 ± 2,12
PROD-2	29,93 ± 0,98	17,73 ± 3,71
PROD-3	16,90 ± 0,85	12,92 ± 0,48
PROD-4	11,35 ± 0,28	6,19 ± 0,46
PROD-5	3,70 ± 0,15	2,21 ± 0,24
PROD-6	28,14 ± 1,08	4,65 ± 0,16
PROD-7	54,49 ± 0,77	14,19 ± 0,32
PROD-8	13,29 ± 0,39	12,18 ± 0,21
PROD-9	12,17 ± 0,74	25,54 ± 4,39
PROD-10	23,14 ± 1,45	8,77 ± 2,30
PROD-11	39,24 ± 0,61	15,75 ± 2,25

Fonte: Autor.

Tabela 4.16 – Erros de previsão para o OLYMPUS D (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	16,71 ± 2,46	7,15 ± 1,92
PROD-2	9,20 ± 3,14	8,92 ± 2,71
PROD-3	12,43 ± 1,40	2,17 ± 0,57
PROD-4	6,74 ± 0,19	1,47 ± 0,18
PROD-5	3,33 ± 0,45	1,81 ± 0,45
PROD-6	13,30 ± 1,37	2,03 ± 0,65
PROD-7	15,08 ± 1,47	2,77 ± 1,89
PROD-8	6,27 ± 0,29	4,94 ± 1,66
PROD-9	3,65 ± 0,44	5,31 ± 1,78
PROD-10	11,99 ± 1,67	8,29 ± 2,99
PROD-11	9,18 ± 0,77	8,09 ± 3,82

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	12,76 ± 0,66	7,04 ± 3,11
PROD-2	8,99 ± 0,98	6,89 ± 2,54
PROD-3	11,18 ± 1,35	2,19 ± 0,43
PROD-4	4,67 ± 0,19	0,97 ± 0,14
PROD-5	3,54 ± 0,20	1,27 ± 0,27
PROD-6	27,31 ± 0,91	1,06 ± 0,30
PROD-7	12,15 ± 0,75	2,81 ± 0,63
PROD-8	4,69 ± 0,19	1,82 ± 0,43
PROD-9	5,51 ± 1,07	3,91 ± 0,47
PROD-10	10,49 ± 1,04	11,47 ± 4,66
PROD-11	8,65 ± 1,26	4,48 ± 3,19

Fonte: Autor.

Tabela 4.17 – Erros de previsão para o OLYMPUS E (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	3,23 ± 1,10	7,20 ± 2,65
PROD-2	82,53 ± 1,90	56,71 ± 0,30
PROD-3	3,40 ± 1,04	1,54 ± 0,19
PROD-4	2,70 ± 0,20	2,62 ± 0,12
PROD-5	2,67 ± 0,40	3,44 ± 0,34
PROD-6	6,85 ± 1,82	1,94 ± 0,45
PROD-7	6,99 ± 1,57	11,90 ± 1,37
PROD-8	2,54 ± 0,35	5,28 ± 0,88
PROD-9	7,26 ± 1,55	13,88 ± 6,03
PROD-10	5,19 ± 0,35	5,07 ± 0,62
PROD-11	3,05 ± 1,02	5,95 ± 1,00

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	23,59 ± 0,76	23,50 ± 1,45
PROD-2	93,50 ± 3,27	25,15 ± 0,92
PROD-3	13,39 ± 0,30	3,46 ± 0,23
PROD-4	2,77 ± 0,20	5,29 ± 0,41
PROD-5	12,92 ± 0,26	6,75 ± 0,27
PROD-6	29,43 ± 1,97	17,88 ± 0,48
PROD-7	20,19 ± 1,31	18,24 ± 1,04
PROD-8	20,63 ± 0,42	4,26 ± 0,22
PROD-9	14,31 ± 0,60	34,61 ± 2,17
PROD-10	30,26 ± 1,01	16,66 ± 1,18
PROD-11	16,27 ± 1,63	11,89 ± 1,53

Fonte: Autor.

Tabela 4.18 – Erros de previsão para o OLYMPUS E (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	18,90 ± 1,22	10,36 ± 1,60
PROD-2	58,75 ± 6,52	12,47 ± 2,34
PROD-3	18,65 ± 1,52	1,83 ± 0,28
PROD-4	4,44 ± 0,24	2,06 ± 0,44
PROD-5	4,00 ± 0,41	1,27 ± 0,27
PROD-6	7,66 ± 0,47	3,21 ± 0,26
PROD-7	6,67 ± 1,05	3,24 ± 1,44
PROD-8	6,21 ± 1,02	5,14 ± 1,02
PROD-9	13,25 ± 2,23	7,19 ± 2,97
PROD-10	5,87 ± 0,89	2,56 ± 0,43
PROD-11	9,18 ± 1,42	11,48 ± 6,98

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	9,77 ± 0,67	14,29 ± 3,47
PROD-2	82,91 ± 9,46	8,34 ± 3,67
PROD-3	10,90 ± 1,79	1,66 ± 0,17
PROD-4	3,37 ± 0,29	1,86 ± 0,46
PROD-5	5,06 ± 0,35	0,74 ± 0,13
PROD-6	9,89 ± 0,80	2,94 ± 0,28
PROD-7	15,17 ± 1,29	4,39 ± 1,50
PROD-8	3,32 ± 0,41	6,91 ± 2,88
PROD-9	12,99 ± 1,94	21,19 ± 3,00
PROD-10	7,12 ± 0,35	14,54 ± 3,47
PROD-11	5,82 ± 1,36	17,13 ± 8,56

Fonte: Autor.

Tabela 4.19 – Erros de previsão para o OLYMPUS F (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	3,71 ± 1,01	23,83 ± 0,54
PROD-2	4,45 ± 1,08	—
PROD-3	9,16 ± 0,81	201,13 ± 7,90
PROD-4	2,32 ± 0,33	2,05 ± 0,50
PROD-5	3,03 ± 0,26	5,67 ± 0,18
PROD-6	6,54 ± 1,60	3,34 ± 0,35
PROD-7	4,94 ± 0,51	5,56 ± 0,11
PROD-8	4,04 ± 0,58	4,63 ± 0,72
PROD-9	3,25 ± 1,31	2,80 ± 0,54
PROD-10	2,50 ± 0,12	56,87 ± 0,75
PROD-11	3,67 ± 0,32	3,86 ± 0,54

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	21,58 ± 0,44	34,81 ± 3,77
PROD-2	29,98 ± 1,03	—
PROD-3	12,56 ± 0,56	522,88 ± 15,11
PROD-4	11,88 ± 0,24	9,38 ± 0,26
PROD-5	4,53 ± 0,44	8,01 ± 0,17
PROD-6	29,69 ± 1,02	5,46 ± 0,20
PROD-7	9,98 ± 1,40	15,30 ± 1,00
PROD-8	9,29 ± 0,65	18,67 ± 1,78
PROD-9	54,53 ± 1,76	4,67 ± 0,72
PROD-10	14,91 ± 0,15	324,87 ± 5,04
PROD-11	6,45 ± 0,65	5,35 ± 1,91

Fonte: Autor.

Tabela 4.20 – Erros de previsão para o OLYMPUS F (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	15,15 ± 1,29	19,04 ± 3,21
PROD-2	23,59 ± 4,22	—
PROD-3	21,73 ± 1,98	15,66 ± 1,31
PROD-4	7,63 ± 0,96	2,06 ± 0,58
PROD-5	3,60 ± 0,54	0,85 ± 0,25
PROD-6	30,58 ± 1,65	1,14 ± 0,24
PROD-7	24,64 ± 3,29	0,89 ± 0,17
PROD-8	11,28 ± 0,90	10,19 ± 2,97
PROD-9	52,49 ± 6,16	2,65 ± 1,12
PROD-10	10,17 ± 1,15	52,49 ± 2,52
PROD-11	14,27 ± 3,16	14,94 ± 7,97

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	9,83 ± 1,32	30,15 ± 0,99
PROD-2	11,68 ± 0,86	—
PROD-3	13,70 ± 0,94	9,93 ± 1,26
PROD-4	10,29 ± 0,40	4,42 ± 1,80
PROD-5	6,94 ± 0,43	0,64 ± 0,11
PROD-6	40,86 ± 1,65	1,41 ± 0,18
PROD-7	15,78 ± 1,27	2,20 ± 1,90
PROD-8	12,38 ± 1,43	6,40 ± 2,62
PROD-9	41,70 ± 7,44	6,92 ± 1,29
PROD-10	6,87 ± 0,87	17,79 ± 1,79
PROD-11	11,05 ± 2,02	11,73 ± 7,35

Fonte: Autor.

5 Conclusões e trabalhos futuros

5.1 Conclusões

Os resultados e análise do capítulo anterior comprovaram que a AMA é um método de agregação no mínimo competitivo a outros métodos de agregação da literatura, tendo várias vezes obtido o menor erro relativo na TPO e/ou TPA, principalmente no OLYMPUS.

5.2 Trabalhos futuros

O método não vem sem alguns pontos negativos relevantes. Primeiro, como ficou claro na análise de resultados, a AMA carece de uma regra, ou conjunto de regras, para ajustar seus vários parâmetros, em particular os valores de u_{OPR} e u_{WPR} . Cabe ao engenheiro de reservatório que for utilizar o método ajustar esses valores via tentativa e erro. Seria do interesse de trabalhos futuros desenvolver, ou ao mínimo considerar, uma maneira automática de realizar esses ajustes, similarmente ao cálculo do hiperparâmetro ótimo na MPE, Lasso e Ridge.

Segundo, o uso do filtro adaptativo como um todo no contexto de reservatório é fortemente heurístico, uma vez que não é, ou não foi, possível traduzir de maneira direta a versão original do filtro conforme estabelecida por [Fortaleza, Gomes et al. \(2022\)](#). Por exemplo, poderiam ser consideradas diferentes maneiras de se definir o parâmetro da média móvel. No manuscrito, decidiu-se por impor, por passo de tempo, apenas um α por poço. Uma proposição alternativa a ser explorada seria definir, por passo de tempo, um α para cada poço e especialista, ou seja, tornar α_t uma matriz de elementos $\alpha_{t,p,N}$ no lugar de $\alpha_{t,p}$. Nessa abordagem, espera-se que as previsões agregadas fiquem ainda mais precisas, já que conseguem elevar o nível de detalhamento do erro durante a filtragem do modelo. Por outro lado, a própria média móvel poderia ser modificada. Note que ela passou originalmente de uma aplicação direta sobre um sinal de previsão para, no contexto de reservatórios, uma aplicação sobre o vetor de pesos do modelo não filtrado, influenciando a previsão agregada de maneira indireta. Em trabalhos futuros, seria um ponto interessante comparar os resultados da AMA apresentada aqui com uma AMA cuja MME atuasse diretamente sobre a previsão agregada.

Os dois parágrafos acima resumem os pontos mais problemáticos da AMA. Em poucas palavras, trata-se de um método com bom potencial, mas que carece de testes e análises para conquistar uma formulação matemática madura, tal qual o Ridge e a MPE. Por fim, vale dizer que mesmo a forma como a AMA fora validada aqui poderia ser repensada.

Seria interessante, por exemplo, considerar um maior número de métodos da literatura, além dos populares Ridge e MPE, como o GE.

Referências

- AANONSEN, S. I.; NØVDAL, G.; OLIVER, D. S.; REYNOLDS, A. C.; VALLÈS, B. The Ensemble Kalman Filter in Reservoir Engineering—a Review. **SPE Journal**, v. 14, n. 03, p. 393–412, set. 2009. ISSN 1086-055X. DOI: [10.2118/117274-PA](https://doi.org/10.2118/117274-PA). Citado na p. 16.
- ALGULIYEV, R.; ALIGULIYEV, R.; IMAMVERDIYEV, Y.; SUKHOSTAT, L. History matching of petroleum reservoirs using deep neural networks. **Intelligent Systems with Applications**, v. 16, p. 200128, 2022. ISSN 2667-3053. DOI: <https://doi.org/10.1016/j.iswa.2022.200128>. Citado na p. 16.
- AMAT, C.; MICHALSKI, T.; STOLTZ, G. Fundamentals and exchange rate forecastability with simple machine learning methods. **Journal of International Money and Finance**, v. 88, p. 1–24, 2018. ISSN 0261-5606. DOI: <https://doi.org/10.1016/j.jimonfin.2018.06.003>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0261560618303668>>. Citado nas pp. 24, 26, 35.
- AUER, P.; CESA-BIANCHI, N.; GENTILE, C. Adaptive and Self-Confident On-Line Learning Algorithms. **Journal of Computer and System Sciences**, v. 64, n. 1, p. 48–75, 2002. ISSN 0022-0000. DOI: <https://doi.org/10.1006/jcss.2001.1795>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S002200001917957>>. Citado nas pp. 39, 44.
- BAHRAMI, P.; MOGHADDAM, F. S.; JAMES, L. A. A Review of Proxy Modeling Highlighting Applications for Reservoir Engineering. **Energies**, v. 15, n. 14, 2022. ISSN 1996-1073. DOI: [10.3390/en15145247](https://doi.org/10.3390/en15145247). Disponível em: <<https://www.mdpi.com/1996-1073/15/14/5247>>. Citado nas pp. 17–20.
- BROWN, R. G. **Exponential smoothing for predicting demand**. Little, 1956. Citado na p. 35.
- CESA-BIANCHI, N. Analysis of Two Gradient-Based Algorithms for On-Line Regression. **Journal of Computer and System Sciences**, v. 59, n. 3, p. 392–411, 1999. ISSN 0022-0000. DOI: <https://doi.org/10.1006/jcss.1999.1635>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0022000099916355>>. Citado nas pp. 22, 44.
- CESA-BIANCHI, N.; LUGOSI, G. **Prediction, Learning, and Games**. Cambridge University Press, 2006. DOI: [10.1017/CB09780511546921](https://doi.org/10.1017/CB09780511546921). Citado nas pp. 23, 38.

- CESA-BIANCHI, N.; FREUND, Y.; HAUSSLER, D.; HELMBOLD, D. P.; SCHAPIRE, R. E.; WARMUTH, M. K. How to Use Expert Advice. **J. ACM**, Association for Computing Machinery, New York, NY, USA, v. 44, n. 3, p. 427–485, mai. 1997. ISSN 0004-5411. DOI: [10.1145/258128.258179](https://doi.org/10.1145/258128.258179). Disponível em: <https://doi.org/10.1145/258128.258179>. Citado nas pp. 22, 44.
- CESA-BIANCHI, N.; MANSOUR, Y.; STOLTZ, G. Improved second-order bounds for prediction with expert advice. **Machine Learning**, v. 66, n. 2, p. 321–352, mar. 2007. ISSN 1573-0565. DOI: [10.1007/s10994-006-5001-7](https://doi.org/10.1007/s10994-006-5001-7). Disponível em: <https://doi.org/10.1007/s10994-006-5001-7>. Citado na p. 39.
- CHAKRA, N. C. C.; SARAF, D. N. History matching of petroleum reservoirs employing adaptive genetic algorithm. **Journal of Petroleum Exploration and Production Technology**, v. 6, n. 4, p. 653–674, dez. 2016. ISSN 2190-0566. DOI: [10.1007/s13202-015-0216-4](https://doi.org/10.1007/s13202-015-0216-4). Citado na p. 16.
- DESWARTE, R.; GERVAIS, V.; STOLTZ, G.; DA VEIGA, S. Sequential model aggregation for production forecasting. **Computational Geosciences**, v. 23, n. 5, p. 1107–1124, out. 2019. ISSN 1573-1499. DOI: [10.1007/s10596-019-09872-1](https://doi.org/10.1007/s10596-019-09872-1). Disponível em: <https://doi.org/10.1007/s10596-019-09872-1>. Citado nas pp. 24–26, 38–40, 43.
- DEVAINE, M.; GAILLARD, P.; GOUDE, Y.; STOLTZ, G. Forecasting electricity consumption by aggregating specialized experts. **Machine Learning**, v. 90, n. 2, p. 231–260, fev. 2013. ISSN 1573-0565. DOI: [10.1007/s10994-012-5314-7](https://doi.org/10.1007/s10994-012-5314-7). Disponível em: <https://doi.org/10.1007/s10994-012-5314-7>. Citado nas pp. 22, 23, 26, 38–40.
- EMERICK, A. A.; REYNOLDS, A. C. History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations. **Computational Geosciences**, Springer, v. 16, p. 639–659, 2012. Citado na p. 16.
- ERVEN, T. van; GRÜNWARD, P.; KOOLEN, W. M.; ROOIJ, S. de. Adaptive Hedge. In: SHAWE-TAYLOR, J.; ZEMEL, R.; BARTLETT, P.; PEREIRA, F.; WEINBERGER, K. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2011. v. 24. Disponível em: https://proceedings.neurips.cc/paper_files/paper/2011/file/64223ccf70bbb65a3a4aceac37e21016-Paper.pdf. Citado na p. 44.
- ESSEN, G. M. v.; ZANDVLIET, M. J.; HOF, P. M. J. V. d.; BOSGRA, O. H.; JANSEN, J. D. Robust Waterflooding Optimization of Multiple Geological Scenarios. **SPE Journal**, v. 14, n. 01, p. 202–210, mar. 2009. ISSN 1086-055X. DOI: [10.2118/102913-PA](https://doi.org/10.2118/102913-PA). eprint: <https://onepetro.org/SJ/article-pdf/14/01/202/2549702/spe-102913-pa.pdf>. Disponível em: <https://doi.org/10.2118/102913-PA>. Citado na p. 49.

- FONSECA, R. M.; ROSSA, E. D.; EMERICK, A. A.; HANEA, R. G.; JANSEN, J. D. Introduction to the special issue: Overview of OLYMPUS Optimization Benchmark Challenge. **Computational Geosciences**, v. 24, n. 6, p. 1933–1941, dez. 2020. ISSN 1573-1499. DOI: [10.1007/s10596-020-10003-4](https://doi.org/10.1007/s10596-020-10003-4). Disponível em: <https://doi.org/10.1007/s10596-020-10003-4>. Citado nas pp. 46, 47, 52.
- FORTALEZA, E. L. F.; GOMES, L. M.; A. LIMAVERDE FILHO, J. O. de; CAMPOS, M. C. de; LONGHI, L. G. S.; LIMA, M. L. de; TOGNETTI, E. S. Model-free adaptive filter to mitigate actuator wear. **ISA Transactions**, v. 129, p. 493–504, out. 2022. ISSN 0019-0578. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0019057822000829>. Citado nas pp. 7, 8, 34, 36, 69.
- FORTALEZA, E. L. F.; FILHO, J. O. d. A. L.; MUNERATO, F. P. A New Model Aggregation Method for Forecasting Medium-Term Reservoir Production. European Association of Geoscientists & Engineers, v. 2022, p. 1–5, jun. 2022. Paper presented at the 83rd EAGE Annual Conference & Exhibition. DOI: [10.3997/2214-4609.202211007](https://doi.org/10.3997/2214-4609.202211007). Citado nas pp. 25, 26, 32–34.
- GHONIEM, S.; ABDEL ALIEM, S.; SAYYOUH, M.; EL SALALY, M. A simplified method for petroleum reservoir history matching. **Applied Mathematical Modelling**, v. 8, n. 4, p. 282–287, 1984. ISSN 0307-904X. DOI: [https://doi.org/10.1016/0307-904X\(84\)90163-X](https://doi.org/10.1016/0307-904X(84)90163-X). Citado na p. 15.
- GOMES, L. M.; A. LIMAVERDE FILHO, J. O. de; FORTALEZA, E.; GONZALEZ, A. F.; MUNERATO, F. P. Wear Attenuation in Actuators based on a Model-free Adaptive Filter for a Gantry Crane System. In: 2023 IEEE Conference on Control Technology and Applications (CCTA). 2023. P. 1109–1114. DOI: [10.1109/CCTA54093.2023.10252668](https://doi.org/10.1109/CCTA54093.2023.10252668). Citado na p. 37.
- GOMES, L. M.; LONGHI, L.; AQUINO LIMAVERDE FILHO, J. O. de; FORTALEZA, E. L. F. Control Performance Improvement of Light Cut Naphtha Extraction via Model-free Adaptive Filter. In: 27TH INTERNATIONAL CONGRESS OF MECHANICAL ENGINEERING. 2023. Citado na p. 37.
- HEIDARI, L.; GERVAIS, V.; RAVALEC, M. L.; WACKERNAGEL, H. History matching of petroleum reservoir models by the Ensemble Kalman Filter and parameterization methods. **Computers & Geosciences**, v. 55, p. 84–95, 2013. Ensemble Kalman filter for data assimilation. ISSN 0098-3004. DOI: <https://doi.org/10.1016/j.cageo.2012.06.006>. Citado na p. 19.
- HOERL, A. E.; KENNARD, R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. **Technometrics**, Taylor & Francis, v. 12, n. 1, p. 55–67, 1970. DOI: [10.1080/00401706.1970.10488634](https://doi.org/10.1080/00401706.1970.10488634). eprint: <https://www.tandfonline.com/doi/pdf/>

- 10.1080/00401706.1970.10488634. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/00401706.1970.10488634>>. Citado nas pp. 41–43.
- HOERL, R. W. Ridge Regression: A Historical Context. **Technometrics**, Taylor & Francis, v. 62, n. 4, p. 420–425, 2020. DOI: 10.1080/00401706.2020.1742207. eprint: <https://doi.org/10.1080/00401706.2020.1742207>. Disponível em: <<https://doi.org/10.1080/00401706.2020.1742207>>. Citado na p. 43.
- HUNTER, J. S. The Exponentially Weighted Moving Average. **Journal of Quality Technology**, Taylor & Francis, v. 18, n. 4, p. 203–210, 1986. DOI: 10.1080/00224065.1986.11979014. eprint: <https://doi.org/10.1080/00224065.1986.11979014>. Disponível em: <<https://doi.org/10.1080/00224065.1986.11979014>>. Citado na p. 35.
- J. O. A. LIMAVERDE FILHO E. L. F. FORTALEZA, J. G. S.; CAMPOS, M. C. M. M. de. Adaptive Kalman filtering for closed-loop systems based on the observation vector covariance. **International Journal of Control**, Taylor & Francis, v. 95, n. 7, p. 1731–1746, 2022. DOI: 10.1080/00207179.2020.1870158. eprint: <https://doi.org/10.1080/00207179.2020.1870158>. Disponível em: <<https://doi.org/10.1080/00207179.2020.1870158>>. Citado na p. 36.
- JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. **An Introduction to Statistical Learning**: with applications in R. 2. ed.: Cambridge University Press, 2021. (Springer Texts in Statistics). DOI: <https://doi.org/10.1007/978-1-0716-1418-1>. Citado nas pp. 42–44.
- JANSEN, J.-D.; FONSECA, R.-M.; KAHROBAEI, S.; SIRAJ, M.; VAN ESSEN, G.; VAN DEN HOF, P. The egg model—a geological ensemble for reservoir simulation. **Geoscience Data Journal**, Wiley Online Library, v. 1, n. 2, p. 192–195, 2014. Citado nas pp. 49, 52.
- KALMAN, R. E. A New Approach to Linear Filtering and Prediction Problems. **Journal of Basic Engineering**, v. 82, n. 1, p. 35–45, mar. 1960. ISSN 0021-9223. DOI: 10.1115/1.3662552. eprint: https://asmedigitalcollection.asme.org/fluidengineering/article-pdf/82/1/35/5518977/35_1.pdf. Disponível em: <<https://doi.org/10.1115/1.3662552>>. Citado na p. 36.
- KHOR, C. S.; ELKAMEL, A.; SHAH, N. Optimization methods for petroleum fields development and production systems: a review. **Optimization and Engineering**, v. 18, n. 4, p. 907–941, dez. 2017. ISSN 1573-2924. DOI: 10.1007/s11081-017-9365-2. Disponível em: <<https://doi.org/10.1007/s11081-017-9365-2>>. Citado na p. 15.

- KIVINEN, J.; WARMUTH, M. K. Exponentiated Gradient versus Gradient Descent for Linear Predictors. **Information and Computation**, v. 132, n. 1, p. 1–63, 1997. ISSN 0890-5401. DOI: <https://doi.org/10.1006/inco.1996.2612>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0890540196926127>. Citado na p. 22.
- LACERDA, J. M.; EMERICK, A. A.; PIRES, A. P. Using a machine learning proxy for localization in ensemble data assimilation. **Computational Geosciences**, Springer, v. 25, p. 931–944, 2021. Citado na p. 16.
- LALEE, M.; NOCEDAL, J.; PLANTEGA, T. On the implementation of an algorithm for large-scale equality constrained optimization. **SIAM Journal on Optimization**, v. 8, n. 3, p. 682–706, 1998. DOI: [10.1137/S1052623493262993](https://doi.org/10.1137/S1052623493262993). Citado na p. 33.
- LE, D. H.; EMERICK, A. A.; REYNOLDS, A. C. An Adaptive Ensemble Smoother With Multiple Data Assimilation for Assisted History Matching. **SPE Journal**, v. 21, n. 06, p. 2195–2207, dez. 2016. ISSN 1086-055X. DOI: [10.2118/173214-PA](https://doi.org/10.2118/173214-PA). eprint: <https://onepetro.org/SJ/article-pdf/21/06/2195/2355689/spe-173214-pa.pdf>. Disponível em: <https://doi.org/10.2118/173214-PA>. Citado na p. 16.
- LITTLESTONE, N.; WARMUTH, M. K. The Weighted Majority Algorithm. **Information and Computation**, v. 108, n. 2, p. 212–261, 1994. DOI: [10.1006/inco.1994.1009](https://doi.org/10.1006/inco.1994.1009). Citado na p. 44.
- LIU, N.; OLIVER, D. S. Ensemble Kalman filter for automatic history matching of geologic facies. **Journal of Petroleum Science and Engineering**, v. 47, n. 3, p. 147–161, 2005. ISSN 0920-4105. DOI: <https://doi.org/10.1016/j.petrol.2005.03.006>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0920410505000550>. Citado na p. 18.
- LIU, P.; ZHANG, K.; YAO, J. Reservoir automatic history matching: Methods, challenges, and future directions. **Advances in Geo-Energy Research**, v. 7, n. 2, p. 136–140, 2023. DOI: [10.46690/ager.2023.02.07](https://doi.org/10.46690/ager.2023.02.07). Disponível em: <https://www.sciopen.com/article/10.46690/ager.2023.02.07>. Citado na p. 17.
- LORENTZEN, R. J.; FJELDE, K. K.; FRØYEN, J.; LAGE, A. C.; NÆVDAL, G.; VEFRING, E. H. Underbalanced and low-head drilling operations: Real time interpretation of measured data and operational support. In: SPE. SPE Annual Technical Conference and Exhibition? 2001. spe-71384. Citado na p. 16.
- MA, Y. Z. **Quantitative Geosciences: Data Analytics, Geostatistics, Reservoir Characterization and Modeling**. 1. ed.: Springer Cham, 2019. DOI: <https://doi.org/10.1007/978-3-030-17860-4>. Citado na p. 16.

- MALLET, V.; STOLTZ, G.; MAURICETTE, B. Ozone ensemble forecast with machine learning algorithms. **Journal of Geophysical Research: Atmospheres**, Wiley Online Library, v. 114, n. D5, 2009. Citado nas pp. 23, 24, 26, 35, 38.
- MIRZAEI-PAIAMAN, A.; SANTOS, S. M.; SCHIOZER, D. J. A review on closed-loop field development and management. **Journal of Petroleum Science and Engineering**, v. 201, p. 108457, 2021. ISSN 0920-4105. DOI: <https://doi.org/10.1016/j.petrol.2021.108457>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0920410521001169>>. Citado na p. 15.
- OLIVEIRA, R. G. B. de; AQUINO LIMAVERDE FILHO, J. O. de; PIMENTA, B. G.; FORTALEZA, E. L. F.; MUNERATO, F. P.; MELO FILHO, L. S. de. Agregação de modelos para prever a produção de reservatórios de petróleo a médio prazo, 2022. Paper presented at the 36th Rio Oil & Gas Expo and Conference, Sep. 26, 2022, Rio de Janeiro. ISSN 2525-7579. DOI: [10.48072/2525-7579.rog.2022.036](https://doi.org/10.48072/2525-7579.rog.2022.036). Citado nas pp. 25, 26, 32–34.
- OLIVER, D. S.; CHEN, Y. Recent progress on reservoir history matching: a review. **Computational Geosciences**, v. 15, n. 1, p. 185–221, jan. 2011. ISSN 1573-1499. DOI: [10.1007/s10596-010-9194-2](https://doi.org/10.1007/s10596-010-9194-2). Disponível em: <<https://doi.org/10.1007/s10596-010-9194-2>>. Citado nas pp. 15–18.
- PINTO, J. W. O.; TUEROS, J. A. R.; HOROWITZ, B.; SILVA, S. M. B. A. da; WILLMERSDORF, R. B.; OLIVEIRA, D. F. B. de. Gradient-free strategies to robust well control optimization. **Computational Geosciences**, v. 24, n. 6, p. 1959–1978, dez. 2020. ISSN 1573-1499. DOI: [10.1007/s10596-019-09888-7](https://doi.org/10.1007/s10596-019-09888-7). Disponível em: <<https://doi.org/10.1007/s10596-019-09888-7>>. Citado na p. 59.
- RAZAK, S. M.; JAFARPOUR, B. Convolutional neural networks (CNN) for feature-based model calibration under uncertain geologic scenarios. **Computational Geosciences**, Springer, v. 24, n. 4, p. 1625–1649, 2020. Citado na p. 16.
- RIAZI, S. H.; ZARGAR, G.; BAHARIMOGHADAM, M.; MOSLEMI, B.; SHARIFI DARANI, E. Fractured reservoir history matching improved based on artificial intelligent. **Petroleum**, v. 2, n. 4, p. 344–360, 2016. ISSN 2405-6561. DOI: <https://doi.org/10.1016/j.petlm.2016.09.001>. Citado na p. 16.
- SATIJA, A.; CAERS, J. Direct forecasting of subsurface flow response from non-linear dynamic data by linear least-squares in canonical functional principal component space. **Advances in Water Resources**, v. 77, p. 69–81, 2015. ISSN 0309-1708. DOI: <https://doi.org/10.1016/j.advwatres.2015.01.002>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0309170815000044>>. Citado na p. 21.

- SATIJA, A.; SCHEIDT, C.; LI, L.; CAERS, J. Direct forecasting of reservoir performance using production data without history matching. **Computational Geosciences**, v. 21, n. 2, p. 315–333, abr. 2017. ISSN 1573-1499. DOI: [10.1007/s10596-017-9614-7](https://doi.org/10.1007/s10596-017-9614-7). Disponível em: <https://doi.org/10.1007/s10596-017-9614-7>. Citado na p. 21.
- SAYYOUH, M. Goal programming: A new tool for optimization in petroleum reservoir history matching. **Applied Mathematical Modelling**, v. 5, n. 4, p. 223–226, 1981. ISSN 0307-904X. DOI: [https://doi.org/10.1016/S0307-904X\(81\)80069-8](https://doi.org/10.1016/S0307-904X(81)80069-8). Citado na p. 15.
- SCHEIDT, C.; RENARD, P.; CAERS, J. Prediction-focused subsurface modeling: investigating the need for accuracy in flow-based inverse modeling. **Mathematical Geosciences**, Springer, v. 47, p. 173–191, 2015. Citado na p. 21.
- SUN, W.; DURLOFSKY, L. J. A New Data-Space Inversion Procedure for Efficient Uncertainty Quantification in Subsurface Flow Problems. **Mathematical Geosciences**, v. 49, n. 6, p. 679–715, ago. 2017. ISSN 1874-8953. DOI: [10.1007/s11004-016-9672-8](https://doi.org/10.1007/s11004-016-9672-8). Disponível em: <https://doi.org/10.1007/s11004-016-9672-8>. Citado na p. 21.
- TIBSHIRANI, R. Regression Shrinkage and Selection Via the Lasso. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 58, n. 1, p. 267–288, jan. 1996. ISSN 0035-9246. DOI: [10.1111/j.2517-6161.1996.tb02080.x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x). eprint: https://academic.oup.com/jrsssb/article-pdf/58/1/267/49098631/jrsssb_58_1_267.pdf. Disponível em: <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>. Citado nas pp. 43, 44.
- VOVK, V. Aggregating strategies. In: PROCEEDINGS of the 3rd annual workshop on computational learning theory (COLT). 1990. P. 371–383. Disponível em: <https://cir.nii.ac.jp/crid/1573105973883251712>. Citado na p. 44.
- VOVK, V. Competitive On-line Statistics. **International Statistical Review**, v. 69, n. 2, p. 213–248, 2001. DOI: <https://doi.org/10.1111/j.1751-5823.2001.tb00457.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1751-5823.2001.tb00457.x>. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1751-5823.2001.tb00457.x>. Citado na p. 38.
- WILSON, K. C.; DURLOFSKY, L. J. Computational Optimization of Shale Resource Development using Reduced-Physics Surrogate Models. In: All Days. (SPE Western Regional Meeting), spe-152946–ms. DOI: [10.2118/152946-MS](https://doi.org/10.2118/152946-MS). Citado na p. 20.
- ZANDVLIET, M.; BOSGRA, O.; JANSEN, J.; VAN DEN HOF, P.; KRAAIJEVANGER, J. Bang-bang control and singular arcs in reservoir flooding. **Journal of Petroleum Science and Engineering**, v. 58, n. 1, p. 186–200, 2007. ISSN 0920-4105. DOI:

<https://doi.org/10.1016/j.petrol.2006.12.008>. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0920410507000101>>. Citado na p. 49.

ZHANG, Z.; HE, X.; ALSINAN, M.; KWAK, H.; HOTEIT, H. Robust Method for Reservoir Simulation History Matching Using Bayesian Inversion and Long-Short-Term Memory Network-Based Proxy. **SPE Journal**, v. 28, n. 03, p. 983–1007, jun. 2023. ISSN 1086-055X. DOI: 10.2118/203976-PA. eprint: <https://onepetro.org/SJ/article-pdf/28/03/983/3120348/spe-203976-pa.pdf>. Disponível em: <<https://doi.org/10.2118/203976-PA>>. Citado nas pp. 20, 24.

Anexos

Anexo A – Comportamento anômalo no OLYMPUS

Figura A.19 – Previsões para o óleo do PROD-2 no OLYMPUS A

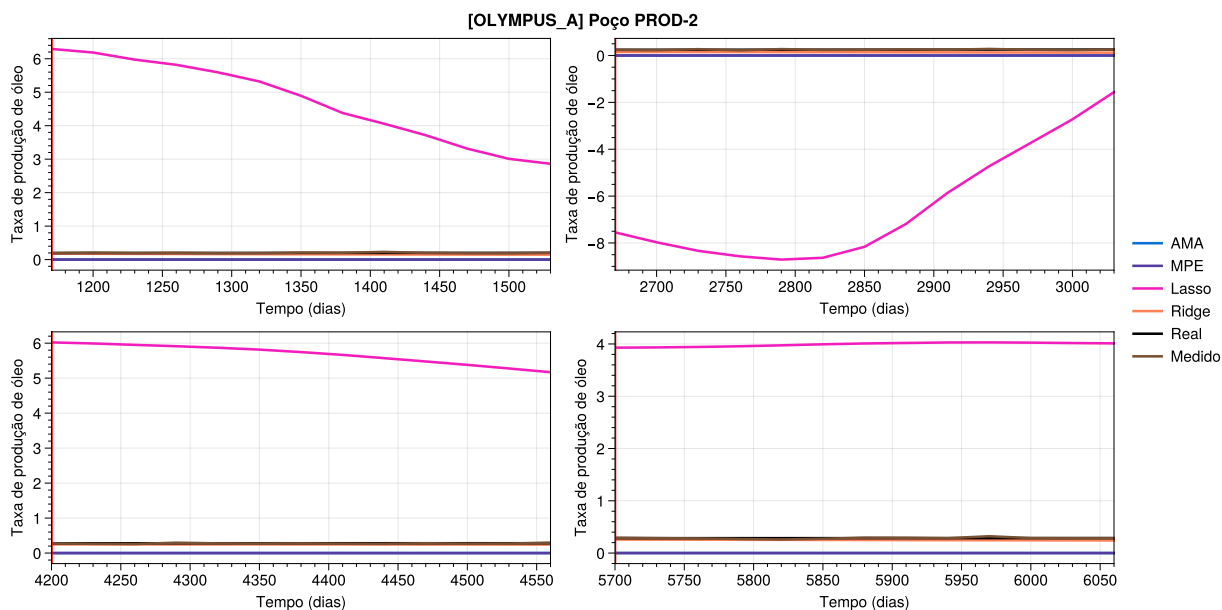


Figura A.20 – Previsões para a água do PROD-6 no OLYMPUS B

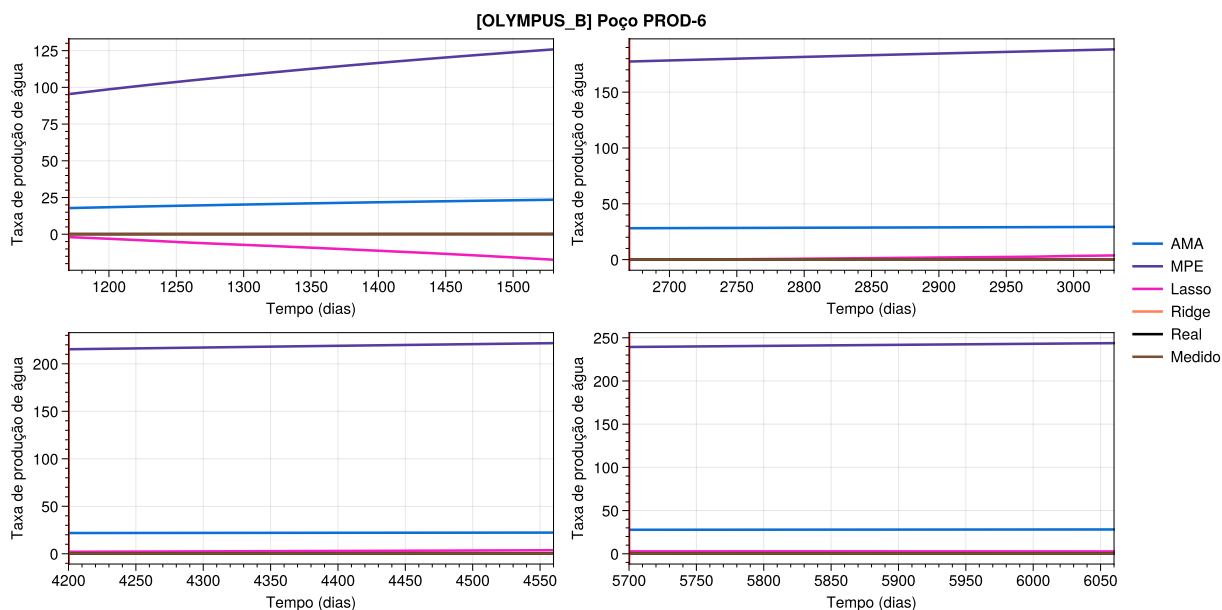


Figura A.21 – Previsões para a água do PROD-2 no OLYMPUS F

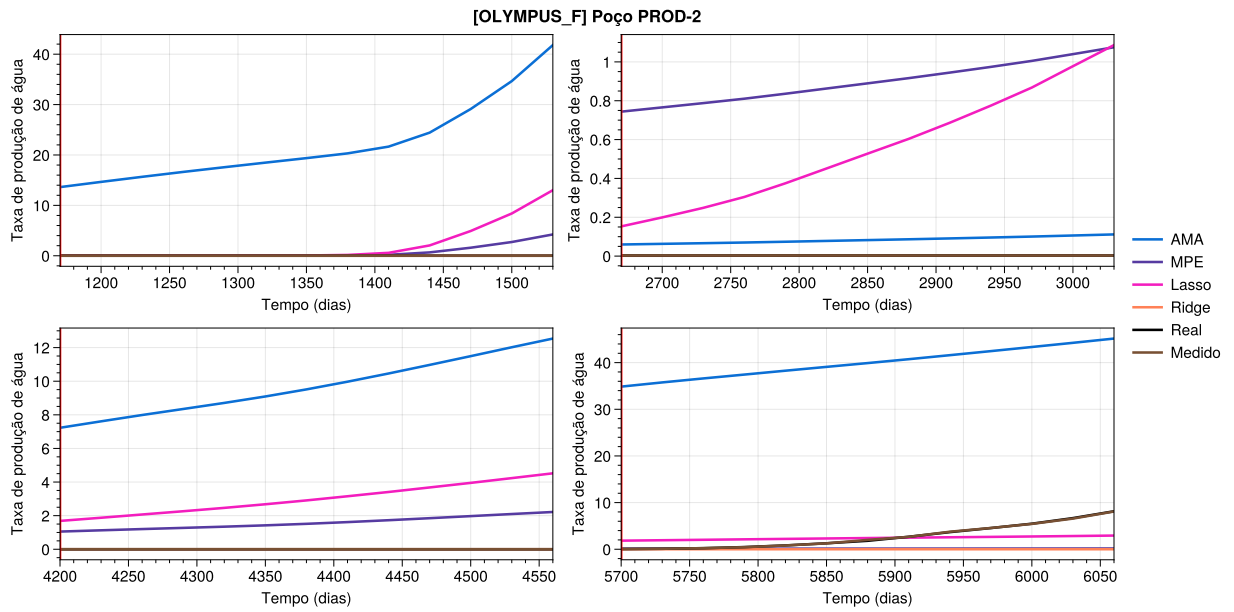
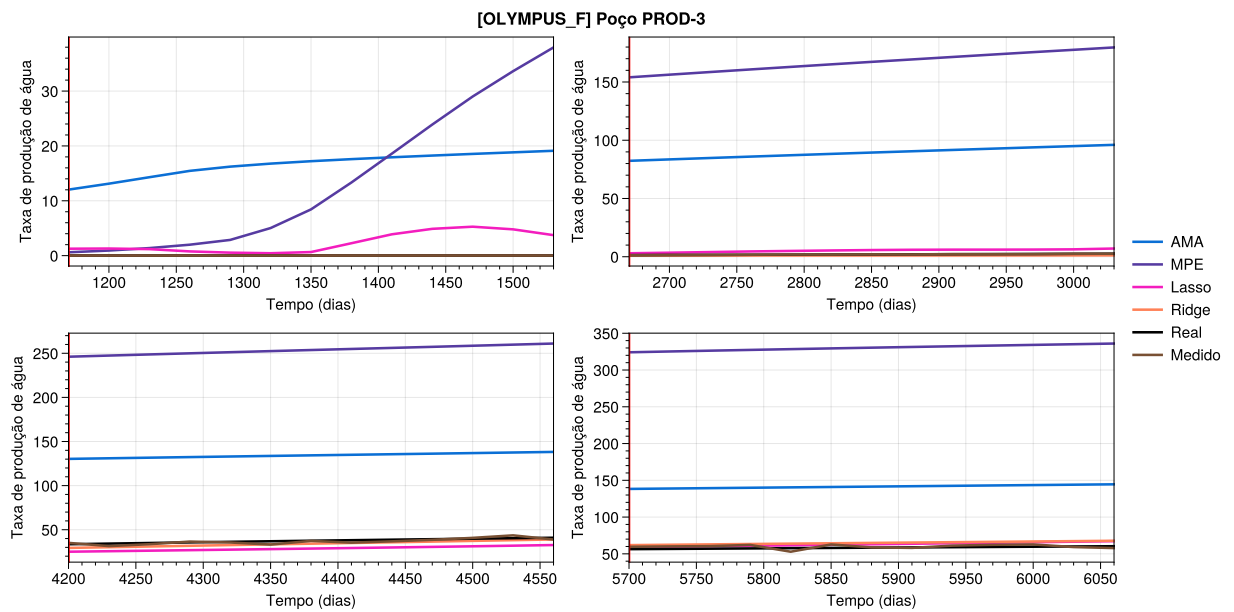


Figura A.22 – Previsões para a água do PROD-3 no OLYMPUS F



Anexo B – Erros relativos no OLYMPUS com limiar 1

OLYMPUS demonstrou sensibilidade ao limiar do erro relativo. No capítulo de resultados o limiar foi tomado como 10. Este apêndice mostra os erros quando o limiar é abaixado para 1.

Tabela B.21 – Erros de previsão para o OLYMPUS A (AMA e MPE).

Poço	Erro da AMA (%)		Poço	Erro da MPE (%)	
	óleo	água		óleo	água
PROD-1	4,24 ± 0,46	428,23 ± 219,24	PROD-1	17,81 ± 0,30	206,41 ± 92,14
PROD-2	—	—	PROD-2	—	—
PROD-3	3,70 ± 0,83	100,69 ± 8,60	PROD-3	10,16 ± 0,80	26,01 ± 4,45
PROD-4	3,22 ± 0,06	2,65 ± 0,19	PROD-4	4,31 ± 0,35	6,75 ± 0,29
PROD-5	4,07 ± 0,41	3,83 ± 0,21	PROD-5	4,85 ± 0,41	3,03 ± 0,28
PROD-6	18,54 ± 0,95	0,75 ± 0,14	PROD-6	145,92 ± 2,84	10,48 ± 0,23
PROD-7	2,27 ± 0,29	2,58 ± 0,31	PROD-7	1,90 ± 0,12	3,23 ± 0,47
PROD-8	2,64 ± 0,09	2,80 ± 0,09	PROD-8	8,87 ± 0,24	9,83 ± 0,96
PROD-9	4,34 ± 0,40	29,22 ± 0,45	PROD-9	43,76 ± 0,56	38,55 ± 22,15
PROD-10	2,14 ± 0,18	20,93 ± 0,13	PROD-10	45,35 ± 0,25	126,11 ± 28,17
PROD-11	3,08 ± 0,62	4,11 ± 0,27	PROD-11	18,98 ± 1,22	12,16 ± 0,81

Fonte: Autor.

Fonte: Autor.

Tabela B.22 – Erros de previsão para o OLYMPUS A (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	4,90 ± 0,41	127,29 ± 58,68
PROD-2	—	—
PROD-3	12,94 ± 1,09	18,31 ± 0,97
PROD-4	2,24 ± 0,22	3,75 ± 0,65
PROD-5	2,87 ± 0,81	2,85 ± 0,37
PROD-6	21,71 ± 0,67	2,39 ± 0,50
PROD-7	2,91 ± 0,44	2,40 ± 0,77
PROD-8	11,54 ± 1,44	5,48 ± 1,33
PROD-9	12,06 ± 2,67	10,21 ± 3,31
PROD-10	5,19 ± 0,19	13,21 ± 8,94
PROD-11	13,52 ± 0,80	38,79 ± 21,80

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	11,87 ± 1,07	36,83 ± 1,40
PROD-2	—	—
PROD-3	9,29 ± 0,93	22,48 ± 1,66
PROD-4	2,10 ± 0,06	1,77 ± 0,60
PROD-5	3,38 ± 0,49	2,42 ± 0,66
PROD-6	16,92 ± 2,65	1,91 ± 0,20
PROD-7	1,28 ± 0,34	3,51 ± 0,76
PROD-8	11,60 ± 2,09	4,91 ± 0,73
PROD-9	17,00 ± 4,53	28,56 ± 11,43
PROD-10	3,16 ± 0,93	30,09 ± 9,33
PROD-11	12,25 ± 1,10	44,11 ± 23,09

Fonte: Autor.

Tabela B.23 – Erros de previsão para o OLYMPUS B (AMA e MPE)

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	3,79 ± 1,34	16,35 ± 0,31
PROD-2	7,36 ± 3,31	19,24 ± 5,86
PROD-3	3,01 ± 0,50	4,22 ± 0,35
PROD-4	2,03 ± 0,27	2,40 ± 0,55
PROD-5	6,22 ± 0,20	11,06 ± 0,31
PROD-6	41,94 ± 1,44	—
PROD-7	6,06 ± 2,40	44,93 ± 3,42
PROD-8	3,14 ± 0,30	9,57 ± 0,24
PROD-9	1,04 ± 0,39	77,19 ± 5,52
PROD-10	7,01 ± 2,43	19,74 ± 7,83
PROD-11	4,18 ± 0,56	9,91 ± 3,01

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	29,94 ± 1,84	20,11 ± 0,88
PROD-2	58,56 ± 3,14	44,23 ± 1,29
PROD-3	16,42 ± 0,41	12,84 ± 0,21
PROD-4	13,74 ± 0,23	7,17 ± 0,35
PROD-5	3,54 ± 0,22	5,84 ± 0,24
PROD-6	369,45 ± 7,80	—
PROD-7	27,42 ± 0,70	4,59 ± 0,39
PROD-8	21,66 ± 0,33	37,91 ± 0,36
PROD-9	18,44 ± 0,73	123,64 ± 58,80
PROD-10	22,87 ± 2,73	19,39 ± 1,40
PROD-11	11,63 ± 0,93	17,33 ± 4,07

Fonte: Autor.

Tabela B.24 – Erros de previsão para o OLYMPUS B (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	7,51 ± 0,72	3,81 ± 0,75
PROD-2	27,94 ± 1,28	26,04 ± 2,22
PROD-3	4,17 ± 0,24	4,57 ± 1,71
PROD-4	0,76 ± 0,16	0,88 ± 0,20
PROD-5	3,89 ± 0,26	2,28 ± 0,54
PROD-6	151,93 ± 2,59	—
PROD-7	12,24 ± 0,53	2,39 ± 0,70
PROD-8	6,33 ± 0,38	5,15 ± 0,86
PROD-9	12,92 ± 1,52	44,31 ± 28,64
PROD-10	23,55 ± 5,60	14,62 ± 4,39
PROD-11	6,05 ± 2,41	14,20 ± 9,91

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	8,19 ± 1,16	2,29 ± 0,79
PROD-2	22,62 ± 3,65	30,17 ± 2,42
PROD-3	6,26 ± 0,30	1,92 ± 0,26
PROD-4	4,98 ± 0,14	4,19 ± 1,14
PROD-5	10,67 ± 0,24	0,68 ± 0,25
PROD-6	31,64 ± 5,38	—
PROD-7	12,55 ± 0,66	1,70 ± 0,61
PROD-8	7,34 ± 0,44	5,79 ± 1,86
PROD-9	15,58 ± 1,18	13,38 ± 1,90
PROD-10	15,37 ± 6,15	9,26 ± 1,93
PROD-11	6,78 ± 2,02	14,35 ± 6,60

Fonte: Autor.

Tabela B.25 – Erros de previsão para o OLYMPUS C (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	8,49 ± 3,38	27,52 ± 0,25
PROD-2	4,08 ± 0,83	17,94 ± 0,80
PROD-3	1,62 ± 0,42	2,39 ± 0,20
PROD-4	4,43 ± 0,41	53,38 ± 14,73
PROD-5	2,81 ± 0,21	4,79 ± 0,24
PROD-6	2,82 ± 0,25	2,79 ± 0,06
PROD-7	2,83 ± 0,30	3,82 ± 0,16
PROD-8	4,61 ± 1,47	14,38 ± 2,01
PROD-9	7,14 ± 2,03	20,17 ± 16,18
PROD-10	2,08 ± 0,97	9,24 ± 1,60
PROD-11	2,63 ± 0,73	26,07 ± 1,00

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	38,84 ± 2,92	24,18 ± 2,67
PROD-2	7,57 ± 0,83	16,00 ± 1,58
PROD-3	23,40 ± 0,54	4,22 ± 0,17
PROD-4	15,87 ± 0,53	43,93 ± 0,45
PROD-5	5,35 ± 0,24	11,08 ± 0,24
PROD-6	45,57 ± 0,50	4,82 ± 0,20
PROD-7	15,11 ± 0,43	3,82 ± 0,42
PROD-8	18,58 ± 1,11	10,79 ± 1,99
PROD-9	6,20 ± 1,39	17,38 ± 6,45
PROD-10	28,73 ± 1,95	18,65 ± 2,92
PROD-11	44,01 ± 0,31	18,50 ± 1,99

Fonte: Autor.

Tabela B.26 – Erros de previsão para o OLYMPUS C (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	13,87 ± 2,24	19,06 ± 3,95
PROD-2	19,00 ± 2,25	19,25 ± 5,83
PROD-3	13,27 ± 0,93	1,26 ± 0,13
PROD-4	7,79 ± 0,44	8,77 ± 1,71
PROD-5	6,60 ± 0,53	3,70 ± 0,75
PROD-6	32,45 ± 1,19	1,83 ± 0,07
PROD-7	11,88 ± 2,13	3,23 ± 2,14
PROD-8	3,82 ± 1,03	8,11 ± 1,65
PROD-9	25,17 ± 2,39	25,34 ± 24,31
PROD-10	6,37 ± 1,12	4,73 ± 2,13
PROD-11	16,78 ± 5,73	11,50 ± 2,99

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	14,68 ± 2,54	25,95 ± 2,06
PROD-2	17,13 ± 2,77	17,03 ± 5,74
PROD-3	14,68 ± 1,51	1,39 ± 0,12
PROD-4	6,05 ± 0,30	23,25 ± 1,76
PROD-5	3,40 ± 0,36	3,60 ± 0,82
PROD-6	31,01 ± 6,53	1,89 ± 0,05
PROD-7	14,46 ± 1,06	2,97 ± 1,43
PROD-8	4,02 ± 0,90	6,42 ± 1,55
PROD-9	20,88 ± 2,35	7,37 ± 3,13
PROD-10	7,63 ± 1,30	20,17 ± 7,95
PROD-11	11,40 ± 1,44	21,70 ± 6,88

Fonte: Autor.

Tabela B.27 – Erros de previsão para o OLYMPUS D (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	2,60 ± 0,62	4,09 ± 1,86
PROD-2	13,89 ± 0,43	32,03 ± 6,10
PROD-3	4,26 ± 0,86	3,24 ± 0,48
PROD-4	2,19 ± 0,31	2,20 ± 0,11
PROD-5	2,92 ± 0,14	3,51 ± 0,09
PROD-6	3,35 ± 0,45	3,43 ± 0,07
PROD-7	42,70 ± 0,60	18,33 ± 0,58
PROD-8	3,30 ± 0,17	5,31 ± 0,23
PROD-9	2,29 ± 1,23	8,63 ± 1,92
PROD-10	4,25 ± 0,52	6,67 ± 1,92
PROD-11	6,75 ± 0,96	14,72 ± 3,58

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	9,92 ± 0,36	14,29 ± 2,12
PROD-2	29,93 ± 0,98	17,73 ± 3,71
PROD-3	16,90 ± 0,85	12,92 ± 0,48
PROD-4	11,35 ± 0,28	6,19 ± 0,46
PROD-5	3,70 ± 0,15	2,21 ± 0,24
PROD-6	28,14 ± 1,08	4,65 ± 0,16
PROD-7	54,49 ± 0,77	14,19 ± 0,32
PROD-8	13,29 ± 0,39	12,18 ± 0,21
PROD-9	12,17 ± 0,74	56,05 ± 20,76
PROD-10	23,14 ± 1,45	8,77 ± 2,30
PROD-11	39,24 ± 0,61	15,75 ± 2,25

Fonte: Autor.

Tabela B.28 – Erros de previsão para o OLYMPUS D (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	16,71 ± 2,46	7,15 ± 1,92
PROD-2	9,20 ± 3,14	8,92 ± 2,71
PROD-3	12,43 ± 1,40	2,17 ± 0,57
PROD-4	6,74 ± 0,19	1,47 ± 0,18
PROD-5	3,33 ± 0,45	1,81 ± 0,45
PROD-6	13,30 ± 1,37	2,03 ± 0,65
PROD-7	15,08 ± 1,47	2,77 ± 1,89
PROD-8	6,27 ± 0,29	4,94 ± 1,66
PROD-9	3,65 ± 0,44	39,54 ± 37,22
PROD-10	11,99 ± 1,67	8,29 ± 2,99
PROD-11	9,18 ± 0,77	8,09 ± 3,82

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	12,76 ± 0,66	7,04 ± 3,11
PROD-2	8,99 ± 0,98	6,89 ± 2,54
PROD-3	11,18 ± 1,35	2,19 ± 0,43
PROD-4	4,67 ± 0,19	0,97 ± 0,14
PROD-5	3,54 ± 0,20	1,27 ± 0,27
PROD-6	27,31 ± 0,91	1,06 ± 0,30
PROD-7	12,15 ± 0,75	2,81 ± 0,63
PROD-8	4,69 ± 0,19	1,82 ± 0,43
PROD-9	5,51 ± 1,07	20,06 ± 10,10
PROD-10	10,49 ± 1,04	11,47 ± 4,66
PROD-11	8,65 ± 1,26	4,48 ± 3,19

Fonte: Autor.

Tabela B.29 – Erros de previsão para o OLYMPUS E (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	3,23 ± 1,10	7,20 ± 2,65
PROD-2	82,53 ± 1,90	56,71 ± 0,30
PROD-3	3,40 ± 1,04	1,54 ± 0,19
PROD-4	2,70 ± 0,20	2,62 ± 0,12
PROD-5	2,67 ± 0,40	3,44 ± 0,34
PROD-6	6,85 ± 1,82	1,94 ± 0,45
PROD-7	6,99 ± 1,57	11,90 ± 1,37
PROD-8	2,54 ± 0,35	5,28 ± 0,88
PROD-9	7,26 ± 1,55	104,18 ± 14,40
PROD-10	5,19 ± 0,35	5,07 ± 0,62
PROD-11	3,05 ± 1,02	5,95 ± 1,00

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	23,59 ± 0,76	23,50 ± 1,45
PROD-2	93,50 ± 3,27	25,15 ± 0,92
PROD-3	13,39 ± 0,30	3,46 ± 0,23
PROD-4	2,77 ± 0,20	5,29 ± 0,41
PROD-5	12,92 ± 0,26	6,75 ± 0,27
PROD-6	29,43 ± 1,97	17,88 ± 0,48
PROD-7	20,19 ± 1,31	18,24 ± 1,04
PROD-8	20,63 ± 0,42	4,26 ± 0,22
PROD-9	14,31 ± 0,60	121,70 ± 52,64
PROD-10	30,26 ± 1,01	16,66 ± 1,18
PROD-11	16,27 ± 1,63	11,89 ± 1,53

Fonte: Autor.

Tabela B.30 – Erros de previsão para o OLYMPUS E (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	18,90 ± 1,22	10,36 ± 1,60
PROD-2	58,75 ± 6,52	12,47 ± 2,34
PROD-3	18,65 ± 1,52	1,83 ± 0,28
PROD-4	4,44 ± 0,24	2,06 ± 0,44
PROD-5	4,00 ± 0,41	1,27 ± 0,27
PROD-6	7,66 ± 0,47	3,21 ± 0,26
PROD-7	6,67 ± 1,05	3,24 ± 1,44
PROD-8	6,21 ± 1,02	5,14 ± 1,02
PROD-9	13,25 ± 2,23	22,07 ± 13,39
PROD-10	5,87 ± 0,89	2,56 ± 0,43
PROD-11	9,18 ± 1,42	11,48 ± 6,98

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	9,77 ± 0,67	14,29 ± 3,47
PROD-2	82,91 ± 9,46	8,34 ± 3,67
PROD-3	10,90 ± 1,79	1,66 ± 0,17
PROD-4	3,37 ± 0,29	1,86 ± 0,46
PROD-5	5,06 ± 0,35	0,74 ± 0,13
PROD-6	9,89 ± 0,80	2,94 ± 0,28
PROD-7	15,17 ± 1,29	4,39 ± 1,50
PROD-8	3,32 ± 0,41	6,91 ± 2,88
PROD-9	12,99 ± 1,94	40,89 ± 2,25
PROD-10	7,12 ± 0,35	14,54 ± 3,47
PROD-11	5,82 ± 1,36	17,13 ± 8,56

Fonte: Autor.

Tabela B.31 – Erros de previsão para o OLYMPUS F (AMA e MPE).

Poço	Erro da AMA (%)	
	óleo	água
PROD-1	3,71 ± 1,01	19,63 ± 4,56
PROD-2	4,45 ± 1,08	1257,96 ± 814,97
PROD-3	9,16 ± 0,81	1639,80 ± 230,97
PROD-4	2,32 ± 0,33	2,05 ± 0,50
PROD-5	3,03 ± 0,26	5,67 ± 0,18
PROD-6	6,54 ± 1,60	3,34 ± 0,35
PROD-7	4,94 ± 0,51	5,56 ± 0,11
PROD-8	4,04 ± 0,58	4,63 ± 0,72
PROD-9	3,25 ± 1,31	53,79 ± 9,47
PROD-10	2,50 ± 0,12	56,61 ± 0,76
PROD-11	3,67 ± 0,32	3,86 ± 0,54

Fonte: Autor.

Poço	Erro da MPE (%)	
	óleo	água
PROD-1	21,58 ± 0,44	86,22 ± 50,52
PROD-2	29,98 ± 1,03	95,13 ± 2,87
PROD-3	12,56 ± 0,56	3194,24 ± 432,06
PROD-4	11,88 ± 0,24	9,38 ± 0,26
PROD-5	4,53 ± 0,44	8,01 ± 0,17
PROD-6	29,69 ± 1,02	5,46 ± 0,20
PROD-7	9,98 ± 1,40	15,30 ± 1,00
PROD-8	9,29 ± 0,65	18,67 ± 1,78
PROD-9	54,53 ± 1,76	43,11 ± 29,75
PROD-10	14,91 ± 0,15	248,70 ± 59,62
PROD-11	6,45 ± 0,65	5,35 ± 1,91

Fonte: Autor.

Tabela B.32 – Erros de previsão para o OLYMPUS F (Lasso e Ridge).

Poço	Erro do Lasso (%)	
	óleo	água
PROD-1	15,15 ± 1,29	61,19 ± 40,13
PROD-2	23,59 ± 4,22	45,51 ± 21,32
PROD-3	21,73 ± 1,98	65,10 ± 6,96
PROD-4	7,63 ± 0,96	2,06 ± 0,58
PROD-5	3,60 ± 0,54	0,85 ± 0,25
PROD-6	30,58 ± 1,65	1,14 ± 0,24
PROD-7	24,64 ± 3,29	0,89 ± 0,17
PROD-8	11,28 ± 0,90	10,19 ± 2,97
PROD-9	52,49 ± 6,16	25,20 ± 24,13
PROD-10	10,17 ± 1,15	32,24 ± 14,22
PROD-11	14,27 ± 3,16	14,94 ± 7,97

Fonte: Autor.

Poço	Erro do Ridge (%)	
	óleo	água
PROD-1	9,83 ± 1,32	30,14 ± 0,99
PROD-2	11,68 ± 0,86	99,85 ± 0,10
PROD-3	13,70 ± 0,94	21,52 ± 2,67
PROD-4	10,29 ± 0,40	4,42 ± 1,80
PROD-5	6,94 ± 0,43	0,64 ± 0,11
PROD-6	40,86 ± 1,65	1,41 ± 0,18
PROD-7	15,78 ± 1,27	2,20 ± 1,90
PROD-8	12,38 ± 1,43	6,40 ± 2,62
PROD-9	41,70 ± 7,44	6,90 ± 1,43
PROD-10	6,87 ± 0,87	14,36 ± 2,47
PROD-11	11,05 ± 2,02	11,73 ± 7,35

Fonte: Autor.

Anexo C – Resultados da AMA para diferentes razões v_{OPR}/v_{WPR}

Tabela C.33 – Erros de previsão da AMA para o OLYMPUS A: 45:8 e 1:1.

Poço	Erro da AMA 45:8 (%)	
	óleo	água
PROD-1	4,24 ± 0,46	3,81 ± 0,82
PROD-2	—	—
PROD-3	3,70 ± 0,83	100,69 ± 8,60
PROD-4	3,22 ± 0,06	2,65 ± 0,19
PROD-5	4,07 ± 0,41	3,83 ± 0,21
PROD-6	10,42 ± 1,26	0,75 ± 0,14
PROD-7	2,27 ± 0,29	2,58 ± 0,31
PROD-8	2,64 ± 0,09	2,80 ± 0,09
PROD-9	4,34 ± 0,40	14,36 ± 0,54
PROD-10	2,14 ± 0,18	20,93 ± 0,13
PROD-11	3,08 ± 0,62	4,11 ± 0,27

Fonte: Autor.

Poço	Erro da AMA 1:1 (%)	
	óleo	água
PROD-1	4,46 ± 0,74	4,13 ± 0,85
PROD-2	—	—
PROD-3	23,06 ± 0,62	30,21 ± 4,71
PROD-4	3,64 ± 0,08	2,36 ± 0,14
PROD-5	3,95 ± 0,44	3,65 ± 0,27
PROD-6	99,78 ± 3,70	1,40 ± 0,10
PROD-7	2,02 ± 0,16	2,68 ± 0,30
PROD-8	6,06 ± 0,29	2,35 ± 0,19
PROD-9	4,29 ± 0,23	5,10 ± 0,48
PROD-10	1,29 ± 0,16	2,52 ± 0,23
PROD-11	3,50 ± 0,66	4,45 ± 0,32

Fonte: Autor.

Tabela C.34 – Erros de previsão da AMA para o OLYMPUS B: 45:8 e 1:1

Poço	Erro da AMA 45:8 (%)	
	óleo	água
PROD-1	3,79 ± 1,34	16,35 ± 0,31
PROD-2	7,36 ± 3,31	20,39 ± 4,46
PROD-3	3,01 ± 0,50	4,22 ± 0,35
PROD-4	2,03 ± 0,27	2,40 ± 0,55
PROD-5	6,22 ± 0,20	11,06 ± 0,31
PROD-6	39,30 ± 1,92	—
PROD-7	6,06 ± 2,40	44,93 ± 3,42
PROD-8	3,14 ± 0,30	9,57 ± 0,24
PROD-9	1,04 ± 0,39	5,13 ± 0,85
PROD-10	7,01 ± 2,43	19,74 ± 7,83
PROD-11	4,18 ± 0,56	9,91 ± 3,01

Fonte: Autor.

Poço	Erro da AMA 1:1 (%)	
	óleo	água
PROD-1	7,08 ± 0,57	4,93 ± 0,88
PROD-2	6,46 ± 3,22	19,30 ± 3,82
PROD-3	2,22 ± 0,44	4,19 ± 0,32
PROD-4	5,09 ± 0,23	2,67 ± 0,29
PROD-5	16,56 ± 0,30	5,45 ± 0,32
PROD-6	92,81 ± 0,47	—
PROD-7	13,81 ± 0,96	15,73 ± 1,65
PROD-8	5,12 ± 0,28	4,15 ± 0,22
PROD-9	1,52 ± 0,42	2,34 ± 0,34
PROD-10	3,62 ± 0,83	8,74 ± 3,01
PROD-11	5,46 ± 1,17	12,57 ± 3,98

Fonte: Autor.

Tabela C.35 – Erros de previsão da AMA para o OLYMPUS C: 45:8 e 1:1.

Poço	Erro da AMA 45:8 (%)	
	óleo	água
PROD-1	8,49 ± 3,38	27,58 ± 0,22
PROD-2	4,08 ± 0,83	17,94 ± 0,80
PROD-3	1,62 ± 0,42	2,39 ± 0,20
PROD-4	4,43 ± 0,41	50,40 ± 10,93
PROD-5	2,81 ± 0,21	4,79 ± 0,24
PROD-6	2,82 ± 0,25	2,79 ± 0,06
PROD-7	2,83 ± 0,30	3,82 ± 0,16
PROD-8	4,61 ± 1,47	14,38 ± 2,01
PROD-9	7,14 ± 2,03	24,35 ± 17,92
PROD-10	2,08 ± 0,97	9,24 ± 1,60
PROD-11	2,36 ± 1,11	5,44 ± 0,69

Fonte: Autor.

Poço	Erro da AMA 1:1 (%)	
	óleo	água
PROD-1	9,56 ± 3,33	25,28 ± 1,04
PROD-2	5,02 ± 0,79	16,69 ± 0,83
PROD-3	2,58 ± 0,39	2,35 ± 0,20
PROD-4	5,82 ± 0,26	13,45 ± 5,63
PROD-5	6,61 ± 0,32	3,51 ± 0,33
PROD-6	13,60 ± 0,43	2,73 ± 0,06
PROD-7	2,93 ± 0,23	2,82 ± 0,18
PROD-8	5,51 ± 0,91	7,19 ± 2,04
PROD-9	4,96 ± 0,60	4,72 ± 1,10
PROD-10	2,82 ± 1,18	4,30 ± 1,37
PROD-11	9,34 ± 2,69	5,46 ± 0,68

Fonte: Autor.

Tabela C.36 – Erros de previsão da AMA para o OLYMPUS D: 45:8 e 1:1.

Poço	Erro da AMA 45:8 (%)	
	óleo	água
PROD-1	2,60 ± 0,62	4,09 ± 1,86
PROD-2	13,89 ± 0,43	32,03 ± 6,10
PROD-3	4,26 ± 0,86	3,24 ± 0,48
PROD-4	2,19 ± 0,31	2,20 ± 0,11
PROD-5	2,92 ± 0,14	3,51 ± 0,09
PROD-6	3,35 ± 0,45	3,43 ± 0,07
PROD-7	42,70 ± 0,60	18,33 ± 0,58
PROD-8	3,30 ± 0,17	5,31 ± 0,23
PROD-9	2,29 ± 1,23	2,89 ± 0,49
PROD-10	4,25 ± 0,52	6,67 ± 1,92
PROD-11	6,75 ± 0,96	14,72 ± 3,58

Fonte: Autor.

Poço	Erro da AMA 1:1 (%)	
	óleo	água
PROD-1	3,28 ± 0,75	5,43 ± 1,98
PROD-2	23,27 ± 0,87	20,38 ± 9,13
PROD-3	4,80 ± 0,82	3,16 ± 0,41
PROD-4	3,43 ± 0,28	2,98 ± 0,33
PROD-5	3,09 ± 0,16	2,55 ± 0,08
PROD-6	6,29 ± 0,45	3,28 ± 0,07
PROD-7	82,35 ± 0,72	2,91 ± 0,60
PROD-8	3,65 ± 0,16	3,82 ± 0,24
PROD-9	3,49 ± 1,34	3,41 ± 0,85
PROD-10	6,91 ± 0,52	8,13 ± 1,61
PROD-11	9,79 ± 0,88	6,42 ± 2,09

Fonte: Autor.

Tabela C.37 – Erros de previsão da AMA para o OLYMPUS E: 45:8 e 1:1.

Poço	Erro da AMA 45:8 (%)	
	óleo	água
PROD-1	3,23 ± 1,10	7,20 ± 2,65
PROD-2	82,53 ± 1,90	56,71 ± 0,30
PROD-3	3,40 ± 1,04	1,54 ± 0,19
PROD-4	2,70 ± 0,20	2,62 ± 0,12
PROD-5	2,67 ± 0,40	3,44 ± 0,34
PROD-6	6,85 ± 1,82	1,94 ± 0,45
PROD-7	6,99 ± 1,57	11,90 ± 1,37
PROD-8	2,54 ± 0,35	5,28 ± 0,88
PROD-9	7,26 ± 1,55	13,88 ± 6,03
PROD-10	5,19 ± 0,35	5,07 ± 0,62
PROD-11	3,05 ± 1,02	5,95 ± 1,00

Fonte: Autor.

Poço	Erro da AMA 1:1 (%)	
	óleo	água
PROD-1	2,73 ± 0,78	8,15 ± 2,07
PROD-2	207,57 ± 2,99	27,86 ± 0,61
PROD-3	4,25 ± 0,96	1,49 ± 0,16
PROD-4	2,76 ± 0,20	2,64 ± 0,17
PROD-5	1,86 ± 0,38	2,97 ± 0,23
PROD-6	6,02 ± 1,63	1,79 ± 0,33
PROD-7	12,80 ± 2,27	5,84 ± 1,91
PROD-8	1,96 ± 0,35	4,64 ± 0,90
PROD-9	5,98 ± 0,91	4,12 ± 1,31
PROD-10	10,28 ± 0,62	1,98 ± 0,55
PROD-11	5,41 ± 1,16	6,21 ± 1,36

Fonte: Autor.

Tabela C.38 – Erros de previsão da AMA para o OLYMPUS F: 45:8 e 1:1.

Poço	Erro da AMA 45:8 (%)	
	óleo	água
PROD-1	3,71 ± 1,01	23,83 ± 0,54
PROD-2	4,45 ± 1,08	—
PROD-3	9,16 ± 0,81	201,13 ± 7,90
PROD-4	2,32 ± 0,33	2,05 ± 0,50
PROD-5	3,03 ± 0,26	5,67 ± 0,18
PROD-6	6,54 ± 1,60	3,34 ± 0,35
PROD-7	4,94 ± 0,51	5,56 ± 0,11
PROD-8	4,04 ± 0,58	4,63 ± 0,72
PROD-9	3,25 ± 1,31	2,80 ± 0,54
PROD-10	2,50 ± 0,12	56,87 ± 0,75
PROD-11	3,67 ± 0,32	3,86 ± 0,54

Fonte: Autor.

Poço	Erro da AMA 1:1 (%)	
	óleo	água
PROD-1	4,36 ± 1,21	27,19 ± 1,06
PROD-2	4,22 ± 0,99	—
PROD-3	43,06 ± 0,56	33,45 ± 3,39
PROD-4	2,10 ± 0,31	3,35 ± 0,70
PROD-5	5,75 ± 0,57	4,52 ± 0,23
PROD-6	19,55 ± 1,57	3,47 ± 0,35
PROD-7	6,94 ± 0,54	3,80 ± 0,23
PROD-8	4,33 ± 0,66	4,63 ± 0,57
PROD-9	3,15 ± 0,60	2,99 ± 0,72
PROD-10	5,93 ± 0,14	33,99 ± 0,69
PROD-11	4,35 ± 0,36	4,65 ± 0,57

Fonte: Autor.