



RIO

PUC

Dissertação de Mestrado

Previsão da Produção de Petróleo com Séries Temporais: Estudo de caso no campo Roncador

Daniela Mendonça dos Santos

Pontifícia Universidade Católica do Rio de Janeiro
Centro Técnico Científico
Departamento de Engenharia Industrial

Rio de Janeiro, 22 de setembro de 2025



**Pontifícia
Universidade
Católica do
Rio de Janeiro**

Dissertação de Mestrado

Previsão da Produção de Petróleo com Séries Temporais: Estudo de caso no campo Roncador

Daniela Mendonça dos Santos

Orientador: Prof. Marcelo Xavier Seeling

Co-orientadora: Profa. Nara Angélica Policarpo

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Logística (opção profissional) pelo Programa de Pós-graduação em Engenharia de Produção, no Departamento de Engenharia Industrial da PUC-Rio.

Rio de Janeiro, 22 de setembro de 2025



**Pontifícia
Universidade
Católica do
Rio de Janeiro**

Previsão da Produção de Petróleo com Séries Temporais: Estudo de caso no campo Roncador

Daniela Mendonça dos Santos

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Logística (opção profissional). Aprovada pela Comissão examinadora abaixo:

Professor Marcelo Xavier Seeling

Orientador

Departamento de Engenharia Industrial – PUC-Rio

Professora Nara Angélica Policarpo

Co-Orientadora

USP

Professor Igor Tona Peres

Departamento de Engenharia Industrial – PUC-Rio

Prof. Rodrigo Goyannes Gusmão Caiado

Departamento de Engenharia Industrial – PUC-Rio

Rio de Janeiro, 22 de setembro de 2025



**Pontifícia
Universidade
Católica do
Rio de Janeiro**

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização da universidade, da autora e do orientador.

Graduou-se em Administração com Ênfase em Comércio Exterior pela Universidade Estácio de Sá de Vila Velha - ES, em 2008. Bacharelado em Direito pela Universidade Estácio Campos Maracanã. Em 2008, ingressou na Petrobras, onde permanece até hoje, integrando a equipe de E&P – Exploração e Produção.

Ficha Catalográfica

Santos, Daniela Mendonça dos

Previsão da Produção de Petróleo com Séries

Temporais: Estudo de caso no campo Roncador / Santos, Daniela Mendonça dos; orientador: Marcelo Xavier Seeling; co-orientador: Nara Angélica Policarpo – 2025.

88 f.: color.; 30cm

Dissertação (mestrando) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2025.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Previsão de Produção; Produção de Petróleo; Séries Temporais; Eficiência Operacional. I. Seeling, Marcelo Xavier. II. Policarpo, Nara Angélica. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. IV. Título.

CDD: 658.5

Agradecimentos

Agradeço a Deus pelo dom da vida e pela sua presença forte em mim, por ter cuidado de mim em todos os momentos, por me haver concedido saúde, clareza mental e capacidade de superar os desafios de forma que eu pudesse concretizar mais um objetivo.

Agradeço à minha família, que é a base de tudo, minha fonte de suporte, amor e apoio.

Agradeço à minha amada filha, Maria Alice, por ter compreendido minha ausência, pelo apoio imprescindível nos momentos desafiadores e por depositar em mim olhares de admiração. Eu te amo.

Agradeço à minha mãe, Maria José, que me educou muito bem, ensinando-me as virtudes que um ser humano pode ter, por estar ao meu lado, por se alegrar com os meus avanços e principalmente por me ensinar a ser perseverante e a perseguir os meus objetivos profissionais com foco e determinação. Eu te amo.

Agradeço ao meu pai, Damázio (in memoriam). Mesmo que tenhamos tido pouco tempo de convivência, reconheço o papel importante que você teve em minha vida e na minha existência.

Agradeço ao meu orientador, Marcelo Xavier Seeling, pela orientação objetiva e pragmática, pelas ideias e pelos esclarecimentos que ajudaram no desenvolvimento deste estudo e à minha co-orientadora professora Dra Nara Angélica Policarpo da USP, por me conduzir nesse projeto e colaborar com ensinamentos e orientações para o aperfeiçoamento deste trabalho.

Expresso minha gratidão à minha colega da Petrobras Dra Verônica Gouveia pelos esclarecimentos sobre aspectos da indústria e pelas sugestões.

Expresso minha gratidão aos colegas de trabalho da Petrobras por contribuírem para o meu desenvolvimento profissional, em especial à Flavia Maria Bessa Respeita Sadocco, pela oportunidade, pelo incentivo e por me permitir extrair o melhor de mim.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Código de Financiamento 001.

Resumo

Santos, Daniela Mendonça; Seeling, Marcelo Xavier (Orientador); Policarpo, Nara Angélica (Co-orientadora). **Previsão da Produção de Petróleo com Séries Temporais: Estudo de caso no campo Roncador**. Rio de Janeiro, 2025. 88p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

A previsão da produção de petróleo em campos maduros é essencial para a sustentabilidade técnica e econômica da indústria de óleo e gás. Apesar da relevância do tema, observa-se uma lacuna na literatura quanto à aplicação prática de modelos estatísticos que integrem variáveis exógenas em ambientes *offshore* brasileiros, especialmente em campos maduros como Roncador. Grande parte dos estudos concentra-se em abordagens baseadas apenas em séries endógenas ou em modelos geológicos complexos, que demandam alto custo e tempo de processamento, limitando sua aplicabilidade em cenários operacionais que exigem respostas rápidas. Este estudo busca preencher essa lacuna ao analisar a produção do Campo de Roncador, um dos maiores ativos *offshore* do Brasil, aplicando modelos estatísticos para prever sua produção futura. A pesquisa adota uma abordagem quantitativa, utilizando séries temporais mensais de produção de petróleo e água entre 2015 e 2025. Foram testados seis modelos preditivos: Suavização Exponencial Simples (SES), Holt, Holt-Winters, ETS, ARIMA e ARIMAX, sendo este último o único a incorporar variáveis exógenas. A análise revelou que o modelo ARIMAX apresentou o melhor desempenho, com menor erro quadrático médio (MSE) e maior aderência à série real, destacando-se como o mais robusto para o cenário analisado. A inclusão da variável “produção de água” como fator exógeno foi determinante para a acurácia do modelo, refletindo a maturidade do reservatório e sua influência na produção de petróleo. Os resultados reforçam a importância de modelos que considerem variáveis externas e a necessidade de dados confiáveis para decisões operacionais e econômicas. Este estudo contribui para o aprimoramento das estratégias de previsão em campos maduros, oferecendo uma alternativa prática e de menor custo frente às abordagens tradicionais.

Palavras-chave

Previsão de Produção; Produção de Petróleo; Séries Temporais; Eficiência Operacional.

Abstract

Santos, Daniela Mendonça; Seeling, Marcelo Xavier (Advisor); Policarpo, Nara Angélica (Co-advisor). **Oil Production Forecasting with Time Series: A Case Study in the Roncador Field**. Rio de Janeiro, 2025. 88p. Dissertação de Mestrado – Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Production forecasting in mature oil fields is essential for the technical and economic sustainability of the oil and gas industry. Despite the relevance of the topic, there is a gap in the literature regarding the practical application of statistical models that integrate exogenous variables in Brazilian offshore environments, especially in mature fields such as Roncador. Most studies focus on approaches based solely on endogenous series or complex geological models, which require high cost and processing time, limiting their applicability in operational scenarios that demand quick responses. This study seeks to fill this gap by analyzing the production of the Roncador Field, one of Brazil's largest offshore assets, applying statistical models to forecast its future output. The research adopts a quantitative approach, using monthly time series of oil and water production between 2015 and 2025. Six predictive models were tested: Simple Exponential Smoothing (SES), Holt, Holt-Winters, ETS, ARIMA, and ARIMAX, the latter being the only one to incorporate exogenous variables. The analysis revealed that the ARIMAX model showed the best performance, with the lowest mean squared error (MSE) and the highest adherence to the actual series, standing out as the most robust for the analyzed scenario. The inclusion of the “water production” variable as an exogenous factor was decisive for the model's accuracy, reflecting the reservoir's maturity and its influence on oil production. The results reinforce the importance of models that consider external variables and the need for reliable data to support operational and economic decisions. This study contributes to improving forecasting strategies in mature fields, offering a practical and lower-cost alternative compared to traditional approaches.

Keywords

Production Forecasting; Oil Production; Time Series; Operational Efficiency.

Sumário

1 INTRODUÇÃO	12
1.1 Objetivo	18
1.2 Justificativa	19
1.3 Estrutura da Dissertação	19
2 REFERENCIAL TEÓRICO	21
2.1 Fases da Produção de Petróleo	21
2.1.1 Recuperação primária: produção inicial baseada na energia natural do reservatório	22
2.1.2 Recuperação secundária: técnicas de injeção para sustentação da pressão	23
2.1.3 Recuperação terciária: métodos avançados para mobilização do óleo residual	24
2.1.4 Estimulação de poços: técnicas para aumentar a produtividade local	25
2.2 Previsão da Produção	26
2.2.1 Previsão de Produção no Planejamento Operacional	27
2.2.2 Fatores que Influenciam o Declínio da Produção	28
2.2.3 Previsão de Séries Temporais de Produção de Petróleo	29
2.2.4 Aspectos Econômicos e Decisões Baseadas na Previsão	30
2.3 Séries Temporais	30
2.3.1 Decomposição STL	32
2.3.2 Horizonte x Intervalo de Previsão	33
2.3.3 Função de Autocorrelação	34
2.4 Métodos de Previsão	35
2.4.1 Suavização Exponencial (SES, Holt, Holt-Winters)	36
2.4.2 Modelo Error, Trend, Seasonality (ETS)	38
2.4.3 ARIMA e Regressão Dinâmica (ARIMAX)	41
2.5 Métricas de Avaliação de Modelos	44
2.5.1 Mean Absolute Percentagem Error (MAPE)	45
2.5.2 Mean Squared Error (MSE)	45
2.6 Modelos Preditivos em Campos Maduros	47
2.6.1 Contexto e Relevância dos Campos Maduros	47
2.6.2 Modelos Estatísticos Aplicados à Previsão	48
2.6.3 Limitações, Lacunas e Perspectivas Futuras	48
3 MÉTODOS DE PESQUISA	50
3.1 Etapas	50
3.1.1 Entrevistas	52
3.1.2 Análise de Dados	53
3.1.3 Modelagem dos Dados Coletados	54
3.1.4 Normalização das Séries Temporais	56
3.1.5 Modelos de Previsão Adotados	57
3.1.6 Aplicação e Avaliação de Modelos SES, Holt, Holt-Winters, ETS e ARIMAX	58

4 RESULTADOS E DISCUSSÕES	59
4.1 Análise da Série Histórica	59
4.2 Avaliação dos Modelos	65
4.3 Validação dos Modelos	67
4.4 Desempenho do Modelo ARIMAX	69
5 CONCLUSÃO	73
REFERÊNCIAS BIBLIOGRÁFICAS	77
APÊNDICE 1 - Protocolo de Estudo de Caso	86

Lista de tabelas

Tabela 1 – Estrutura de Modelos Temporais: Tendência, Sazonalidade e Erro	39
Tabela 2 – Relação dos profissionais e pesquisadores entrevistados.	52
Tabela 3 – Descrição e Finalidade das Bibliotecas R Utilizadas na Análise	54
Tabela 4 – Parâmetros do Modelo de Previsão Multivariado de Regressão Dinâmica	67
Tabela 5 – Métricas de Acurácia	68

Lista de figuras e gráficos

Gráfico 1 – Matriz Energética Mundial 2022	13
Gráfico 2 – Matriz Energética Brasileira 2023	14
Figura 1 – Localização da área de estudo no contexto da Bacia de Campos, com ênfase no campo petrolífero de Roncador	15
Gráfico 3 – Série Temporal da Produção de Petróleo (m^3/d)	60
Gráfico 4 – Decomposição da Série Temporal (m^3/d)	60
Gráfico 5 – Previsão de Produção de Petróleo (2015–2025)	61
Gráfico 6 – Séries de Petróleo e Água (Normalizadas)	62
Gráfico 7 – Correlação Cruzada (CCF) entre Petróleo e Gás	63
Gráfico 8 – Função Autocorrelação da Série Alvo – Período Treinamento	63
Gráfico 9 – Produção Período de Treinamento	64
Gráfico 10 – Função Autocorrelação Parcial – Série Treinamento	65
Gráfico 11 – Produção Real vs Previsões (Mar/2024 – Fev/2025)	66
Gráfico 12 – Comparação de MAPE e MSE entre os modelos – Período Validação	68
Gráfico 13 – Previsão de Modelo ARIMAX	69

Lista de siglas

ANP – Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ARIMA – AutoRegressive Integrated Moving Average
ARIMAX – AutoRegressive Integrated Moving Average with Exogenous Variables
BEM – Balanço Energético Nacional
CAPEX – Capital Expenditures (Investimentos de Capital)
CEPETRO – Centro de Estudos em Energia e Petróleo
E&P – Exploração e Produção
EOR – Enhanced Oil Recovery (Recuperação Avançada de Petróleo)
ETS – Error, Trend, Seasonality (Erro, Tendência, Sazonalidade)
FAC – Função de Autocorrelação
FACP – Função de Autocorrelação Parcial
FPSO – Floating Production Storage and Offloading
IBP – Instituto Brasileiro de Petróleo e Gás
IEA – International Energy Agency
LSTM – Long Short-Term Memory
MAPE – Mean Absolute Percentage Error
MAE – Mean Absolute Error
MMboe – Milhões de Barris de Óleo Equivalente
MSE – Mean Squared Error
OPEX – Operational Expenditures
OPEP – Organização dos Países Exportadores de Petróleo
PUC-Rio – Pontifícia Universidade Católica do Rio de Janeiro
RAO – Razão Água-Óleo
RMSE – Root Mean Squared Error
SAR – Sistema de Análise de Reservatórios
SES – Suavização Exponencial Simples
S4D – Sísmica 4D
STL – Seasonal and Trend Decomposition using Loess
TIR – Taxa Interna de Retorno
VPL – Valor Presente Líquido
WAG – Water-Alternating Gas Injection
WC – Water Cut

Introdução

O petróleo consolidou-se como a principal fonte de energia da sociedade moderna, desempenhando papel fundamental no avanço econômico global a partir do século XX. Conforme destaca Yergin (2010), essa matéria-prima não apenas substituiu outras formas de energia, mas também passou a impulsionar a produção de combustíveis e a expansão da indústria petroquímica. Em sua análise, o autor compara o petróleo a um monarca absoluto, cuja generosidade promoveu uma era de crescimento, confiança e prosperidade, transformando profundamente a estrutura econômica e social. Essa influência marcante deu origem à chamada Era do Homem do Hidrocarboneto (Yergin, 2010, p. 610).

O petróleo é considerado um dos recursos naturais mais relevantes e amplamente utilizados globalmente. Sua presença é marcante em diversos setores da economia, como transporte e indústria petroquímica, evidenciando sua importância estratégica. Além de ser indispensável para o funcionamento dos sistemas de transporte, o petróleo serve como matéria-prima para a fabricação de inúmeros compostos químicos e plásticos, essenciais em áreas como agricultura, saúde, eletrônicos e embalagens. A disponibilidade desses derivados é fundamental para impulsionar o desenvolvimento e garantir a competitividade desses segmentos econômicos (MARTINS, 2023).

Considerando os custos associados às emissões de gases de efeito estufa, observa-se uma transição gradual da indústria de petróleo e gás para fontes alternativas, mais eficientes do ponto de vista ambiental. Esse movimento tem impulsionado a diversificação da matriz energética e reduzido a dependência de fontes não renováveis (ENERDATA, 2023). Nesse contexto, além da expansão da geração solar e eólica, destaca-se a importância da implantação de sistemas de armazenamento de energia, essenciais para mitigar os efeitos da intermitência e da sazonalidade dessas fontes, cuja participação na matriz tem crescido de forma significativa (ERBER; MOSSÉ, 2023).

Apesar dos avanços nas tecnologias de energia limpa, os combustíveis fósseis continuam desempenhando papel central na matriz energética global,

especialmente devido à sua ampla disponibilidade e ao custo relativamente baixo. Estimativas indicam que, em 2014, mais de 80% do consumo mundial de energia ainda era proveniente de fontes fósseis, como carvão, petróleo e gás natural (SILVA et al., 2023). Segundo a IEA (2024), petróleo, gás natural e carvão continuam predominantes, especialmente em países em desenvolvimento. A transição energética exige superar essa dependência com investimentos em tecnologias sustentáveis. Nos países em desenvolvimento, o acesso contínuo à energia é essencial para o progresso social e econômico, especialmente em setores como o agronegócio, que depende de energia para irrigação, transporte e processamento (IEA, 2025).

Globalmente, a matriz energética, conforme Gráfico 1, está segmentada entre fontes de energia renováveis e não renováveis, em torno de 20% e 80%, respectivamente.

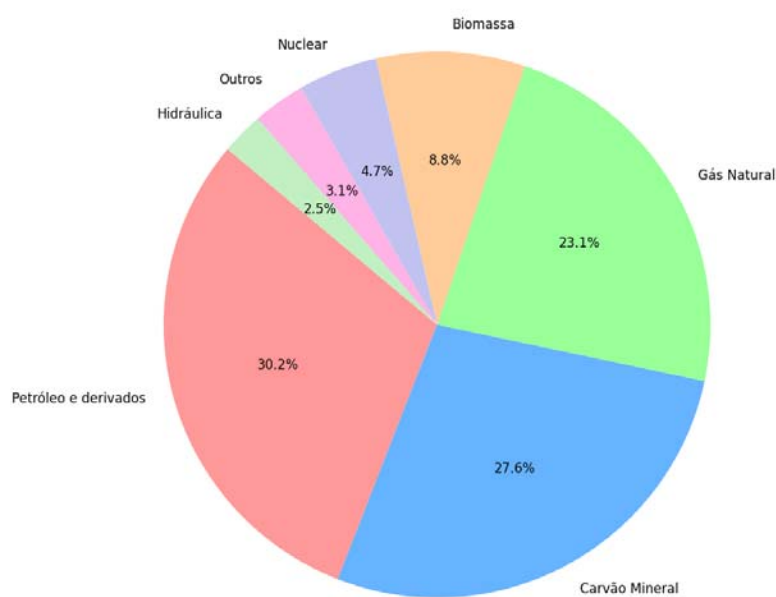


Gráfico 1: Matriz Energética Mundial 2022.
Fonte: Adaptado de *International Energy Agency (IEA)* (2023).

No caso brasileiro, a segmentação da matriz energética apresenta uma configuração significativamente distinta, conforme ilustrado no Gráfico 2, com aproximadamente 49% provenientes de fontes renováveis e 51% de fontes não

renováveis (Matriz Energética Brasileira) (BEM) (EPE, 2025), o que evidencia a diversidade e o potencial do setor energético brasileiro



Gráfico 2: Matriz Energética Brasileira 2023.
Fonte: Adaptado de Bem (2024).

No Brasil, o setor representa 10% do PIB industrial e continuará a contribuir para o crescimento econômico por meio de investimentos previstos em Exploração e Produção, que somam aproximadamente US\$ 180 bilhões entre 2022 e 2031 (INSTITUTO BRASILEIRO DE PETRÓLEO E GÁS - IBP, 2023).

Descoberto em 1996 pelo poço RJS-436 o Campo de Roncador, conforme Figura 1, está localizado a aproximadamente 130 km da costa do estado do Rio de Janeiro, na Bacia de Campos, em lâminas d'água que variam entre 1.500 e 1.900 metros (COSTA et al., 2005). Segundo a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) (2024), a produção total acumulada do campo até o ano de 2024 atingiu 1.766 MMboe. A sigla MMboe refere-se a “Milhões de Barris de Óleo Equivalente” (*Million Barrels of Oil Equivalent*), unidade de medida que contabiliza não apenas volumes de óleo, mas também de gás natural, convertidos em barris de óleo por equivalência energética.

A unidade de MMboe é amplamente utilizada na indústria para integrar volumes de óleo e gás natural com base em equivalência energética, permitindo análises comparativas e consolidadas de produção e reserva. Essa padronização é

essencial para a gestão integrada de ativos, especialmente em ambientes *offshore*, onde a variabilidade operacional e os custos logísticos impõem desafios adicionais à previsibilidade da produção. Ela também permite a consolidação e comparação de dados de produção e reservas de forma integrada, sendo amplamente adotada na indústria de petróleo e gás (MONTEIRO, 2023).

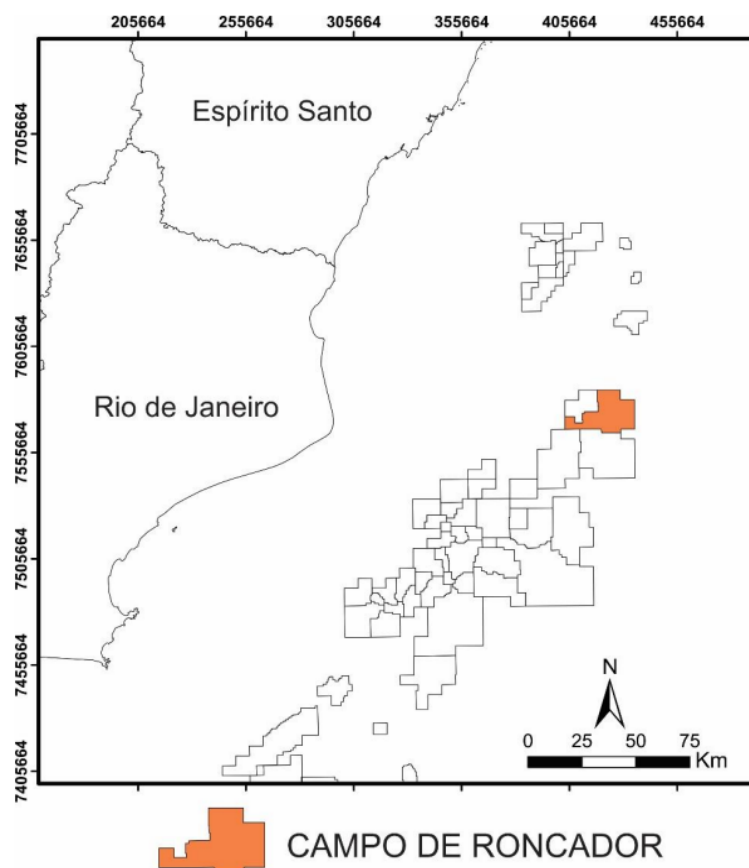


Figura 1: Localização da área de estudo no contexto da Bacia de Campos, com ênfase no campo petrolífero de Roncador.

Fonte: FURLAN, 2017.

Nas últimas três décadas, a Petrobras ampliou significativamente suas atividades de exploração e produção em águas profundas e ultraprofundas, especialmente nas bacias brasileiras. Ainda na década de 1990, foram descobertos e desenvolvidos campos gigantes como Marlim, Albacora e Roncador, este último iniciando produção em 1999. Esses campos representam marcos na indústria petrolífera nacional e continuam sendo revitalizados e integrados a novas tecnologias, mantendo sua importância estratégica para o país. Esse crescimento rápido aumentou as reservas provadas brasileiras. Cerca de 88% dessas reservas e

mais de 80% da produção diária de 2 milhões de boe vêm de arenitos profundos do Cretáceo e Terciário na Bacia de Campos (MUTTI; CARMINATTI, 2012).

Apesar de sua condição de maturidade, a Bacia de Campos permanece como um dos principais polos de desenvolvimento da indústria *offshore* brasileira. De acordo com a ANP (2024a), a Bacia de Campos concentra cerca de 19% das reservas provadas de petróleo do Brasil. Em relação à produção, essa bacia é responsável por aproximadamente 22% do total nacional (ANP, 2024b).

A Petrobras é a maior empresa de petróleo e gás do Brasil e uma das líderes mundiais na exploração em águas profundas e ultraprofundas (AGÊNCIA PETROBRAS, 2024).

Em 2024, a Petrobras e a Equinor iniciaram com sucesso a produção dos dois primeiros poços do projeto de Recuperação Avançada de Óleo (EOR – *Enhanced Oil Recovery*) em Roncador, marcando um avanço significativo na eficiência e sustentabilidade do campo. A produção começou cinco meses antes do previsto e com metade do custo planejado, elevando a produção diária para aproximadamente 150 mil barris de óleo equivalente (EQUINOR, 2022).

Sob a ótica macroeconômica, a produção de petróleo permanece como um vetor estratégico para o desenvolvimento nacional. Projeções da Organização dos Países Exportadores de Petróleo (OPEP) indicam que a produção brasileira poderá atingir 4,4 milhões de barris por dia em 2025, impulsionada por uma combinação de fatores, como a recuperação da atividade industrial, o aumento da demanda energética e a expansão da capacidade instalada (CNN BRASIL, 2024).

Entretanto, em um ambiente competitivo e volátil, caracterizado por rápidas mudanças nos padrões de consumo e pela influência de variáveis macroeconômicas, a tarefa de prever com precisão a demanda futura torna-se cada vez mais desafiadora. À medida que as necessidades dos clientes se tornam mais dinâmicas e complexas, aumentam também as limitações dos modelos preditivos, que frequentemente enfrentam a escassez de séries históricas robustas e a dificuldade de capturar o comportamento errático das variáveis envolvidas (HOFMANN; RUTSCHMANN, 2018).

A previsão de produção, nesse contexto, configura-se como uma etapa crítica no planejamento operacional e financeiro das empresas do setor de Exploração e Produção (E&P) de Óleo e Gás. A utilização de modelos estatísticos

e computacionais permite antecipar o comportamento futuro dos campos, subsidiando decisões estratégicas relacionadas à alocação de recursos, investimentos em infraestrutura e mitigação de riscos operacionais (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998). No caso do Campo de Roncador, objeto de estudo do presente trabalho, a maturidade do reservatório e a complexidade das operações tornam a previsão de produção importante para garantir a sustentabilidade técnica e econômica do ativo.

A previsão da produção desempenha papel central na avaliação da viabilidade econômica de empreendimentos petrolíferos, sendo utilizada como base para a estimativa de indicadores financeiros como o Valor Presente Líquido (VPL), a Taxa Interna de Retorno (TIR) e o payback (FARIA, 2015).

A modelagem de séries temporais em ambientes industriais complexos, como a produção *offshore* de petróleo, demanda abordagens que considerem tanto a estrutura temporal dos dados quanto a influência de variáveis operacionais externas. A identificação de padrões como tendência e sazonalidade é essencial para a construção de modelos preditivos robustos, especialmente em campos maduros, onde a variabilidade e a incerteza são elevadas (HYNDMAN et al., 2008).

Nesse contexto, modelos estatísticos como *AutoRegressive Integrated Moving Average* (ARIMA), *Error, Trend, Seasonal* (ETS) e suas extensões, incluindo o *ARIMA with Exogenous variables* (ARIMAX), têm sido amplamente utilizados na literatura e demonstram desempenho satisfatório na previsão de séries temporais com diferentes níveis de complexidade (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998). O modelo *ETS*, por exemplo, permite a combinação de diferentes formas de erro (aditivo ou multiplicativo), tendência (nula, linear ou amortecida) e sazonalidade (aditiva ou multiplicativa), resultando em uma estrutura altamente flexível e adaptável a diferentes padrões de comportamento da série (HYNDMAN et al., 2008).

Além disso, a incorporação de variáveis exógenas, como o volume de água extraído nos poços durante o processo de produção de petróleo, por meio de modelos como o ARIMAX, contribui significativamente para o aumento da acurácia das previsões. Segundo Yanrui (2022), essa abordagem permite capturar relações determinísticas entre variáveis operacionais, o que é particularmente

relevante em sistemas de produção interdependentes. A produção de água, por exemplo, é um indicador crítico do estágio de maturidade do reservatório e influencia diretamente a eficiência da extração de petróleo (ABRAHAM-ADEJUMO, 2018). A produção de água impacta diretamente a eficiência da elevação e o processamento primário, validando sua inclusão como variável exógena no modelo ARIMAX. Ao integrar essa variável ao modelo, é possível antecipar padrões de declínio e ajustar estratégias operacionais com maior precisão, o que é essencial para a sustentabilidade técnica e econômica de campos maduros (DANTAS; PESSOA, 2020). A variabilidade da produção de água pode ser usada como indicador preditivo da produtividade do poço, por meio da extrapolação da RAO (Razão Água-Óleo) e do declínio da fração de óleo. Além disso, o uso de variáveis exógenas permite que o modelo reaja de forma mais sensível a mudanças nas condições do sistema, como intervenções técnicas ou variações na pressão do reservatório, ampliando sua capacidade de adaptação a cenários reais (BOX et al., 2015). Essa flexibilidade é particularmente valiosa em ambientes *offshore*, onde a tomada de decisão precisa ser rápida, baseada em dados confiáveis e com alto grau de previsibilidade (BIDGOLI, 2018).

1.1

Objetivo

O objetivo deste estudo é analisar a aplicabilidade de modelos de séries temporais para previsão de produção de petróleo no Campo de Roncador como um suporte adicional para tomada de decisão. Para isso, são aplicados modelos estatísticos *Simple Exponential Smoothing* (SES), *Holt's Linear Trend Method* (Holt), *Holt-Winters Exponential Smoothing* (Holt-Winters), *Error, Trend e Seasonality* (ETS), ARIMA ARIMAX, com a finalidade de selecionar aquele que apresenta maior acurácia preditiva, visando apoiar decisões estratégias e operacionais em campos de petróleo *offshore* maduros.

Este trabalho contribui para a literatura do setor de óleo e gás, com aplicação prática de métodos estatísticos de séries históricas para a previsão de produção de petróleo.

1.2

Justificativa

A necessidade de novos modelos preditivos para o Campo de Roncador decorre de sua condição de maturidade e da complexidade operacional associada à produção em águas profundas. Após mais de duas décadas de operação, o campo apresenta declínio acentuado na produção e aumento significativo da razão água-óleo (RAO), fatores que impactam diretamente a eficiência e os custos operacionais (DANTAS; PESSOA, 2020). As abordagens tradicionais, baseadas em simulações geológicas ou modelos univariados, têm se mostrado limitadas para capturar a influência de variáveis operacionais críticas, como a produção de água, que é determinante para a tomada de decisão em estratégias de recuperação avançada (ABRAHAM-ADEJUMO, 2018). Nesse contexto, torna-se essencial aplicar modelos estatísticos que integrem variáveis exógenas, capazes de refletir a dinâmica real do reservatório e oferecer previsões mais precisas para apoiar decisões estratégicas e operacionais (HYNDMAN et al., 2008). Essa necessidade é reforçada pelo cenário de transição energética e pressão por eficiência, no qual previsões confiáveis permitem otimizar recursos, reduzir custos e mitigar riscos (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998).

A Bacia de Campos, apesar de sua maturidade, ainda representa parcela significativa da produção nacional, o que reforça a importância de estratégias preditivas robustas para garantir sua longevidade produtiva (ANP, 2024 a).

Neste contexto, este estudo contribui para a literatura sobre o tema e para os profissionais da área oferecendo uma aplicação real de métodos de previsão estatísticos aplicados à produção de petróleo em ambientes *offshore*, contribuindo para a eficiência operacional e para o desenvolvimento de soluções tecnológicas alinhadas às demandas atuais da indústria.

1.3

Estrutura da dissertação

Este estudo está estruturado em cinco capítulos. No Capítulo 1 de introdução são apresentados o contexto e a relevância do tema pesquisado, o

objetivo do estudo, a justificativa da escolha do tema e a estrutura do trabalho. No Capítulo 2 de Referencial Teórico, são apresentados conceitos fundamentais para a compreensão da pesquisa, como produção de petróleo e técnicas estatísticas de previsão da produção. No Capítulo 3, Método de Pesquisa, é descrito o método de pesquisa adotado neste estudo, incluindo os procedimentos de coleta e tratamento dos dados, as limitações do estudo, bem como os critérios de seleção e avaliação dos modelos utilizados. No Capítulo 4, são apresentados os resultados obtidos e discutido o desempenho dos modelos aplicados. Por fim, no Capítulo 5, Conclusão, são apresentadas as conclusões do estudo e são propostas sugestões para pesquisas futuras.

2

Referencial teórico

Neste capítulo é apresentada a base teórica utilizada para o desenvolvimento do estudo, abordando conceitos fundamentais para a compreensão do tema. O capítulo está estruturado em cinco tópicos principais. O primeiro tópico trata das fases da produção de petróleo, explorando suas classificações, causas e impactos. O segundo aborda a previsão de produção, detalhando as fases que compõem o ciclo de resposta a esses eventos. O terceiro tópico trata dos métodos de previsão explicando a importância e diferenças entre eles. O quarto discorre sobre séries temporais, destacando os modelos abordados neste estudo. Por fim, o quinto tópico apresenta as métricas de avaliação de modelos.

2.1

Fases da produção de petróleo

Nos campos de petróleo, os poços produtores atingem sua máxima produção na forma de picos ou platôs de vazão que podem estar associados a diversas causas. Após esse estágio, é natural que a produção entre em uma fase de declínio, tornando essencial a aplicação de métodos de recuperação para manter a produção economicamente viável. Portanto, é necessário prever a produção futura desses campos (SILVA et al., 2023).

Métodos de recuperação de petróleo são técnicas aplicadas para aumentar a extração de óleo de um reservatório, dividindo-se em três tipos: a recuperação primária, que utiliza a energia natural do reservatório, ou seja, sua pressão interna; a secundária, que injeta água ou gás para manter a pressão que empurra o óleo para a superfície; e a terciária, que emprega agentes químicos, térmicos ou biológicos para mobilizar o óleo residual (FREITAS, 2013). E há os métodos de estimulação que são técnicas aplicadas diretamente nos poços para aumentar a produtividade, principalmente por meio da redução da resistência ao fluxo de fluidos. Destacam-se a acidificação, que remove obstruções minerais nas rochas, e

o fraturamento hidráulico, que cria fraturas artificiais para facilitar o escoamento do óleo. Essas práticas são comuns em poços com baixa produtividade (CASTRO et al., 2011). Os campos petrolíferos, assim como os poços que o compõe, têm um período de pico de produção e depois decaem até o fim de suas vidas produtivas. Isso ocorre, pois a quantidade de óleo da reserva vai esgotando e consequentemente a extração vai diminuindo. A vida útil dos campos e poços pode variar muito, fatores como o volume de reserva, a estrutura instalada e a tecnologia disponível para produzir o hidrocarboneto (DANTAS; PESSOA, 2020).

A fase de produção é a etapa crucial na exploração de petróleo, onde a extração ocorre de maneira contínua. O petróleo é extraído do subsolo utilizando sistemas de elevação artificial ou a pressão natural dos reservatórios, que força o petróleo a subir à superfície. Para aumentar a pressão no reservatório e otimizar a produção, diversas tecnologias podem ser empregadas, como a injeção de água, gás ou outros agentes. Quando um campo petrolífero chega ao fim de sua vida útil, é necessário proceder com a desativação e o abandono. As plataformas de perfuração, os poços e outras estruturas são desativadas, e a área é restaurada ao máximo possível para assegurar que não ocorram vazamentos ou contaminações (NAVSUPPLY, 2024).

2.1.1

Recuperação primária: produção inicial baseada na energia natural do reservatório

Conforme Abraham-Adejumo (2018), a recuperação primária corresponde ao estágio inicial da produção de petróleo, no qual a extração ocorre por meio da energia natural contida no reservatório, como a pressão do gás, a expansão do óleo ou a presença de água associada. Essa energia é responsável por impulsionar os hidrocarbonetos até a superfície, sem a necessidade de intervenções externas. essa etapa apresenta uma eficiência limitada, com recuperação estimada entre 5% e 15% do volume original de óleo, variando conforme as propriedades petrofísicas da rocha e a pressão inicial do sistema. A conectividade entre os poros e a mobilidade dos fluidos são fatores determinantes para o desempenho dessa fase.

Santos (2024) ressalta que a previsão da produção durante essa etapa é essencial para o planejamento da infraestrutura de superfície e para a avaliação da viabilidade econômica do empreendimento. A modelagem de parâmetros como a transmissibilidade e o corte de água (*water cut*) permite antecipar o comportamento do reservatório e orientar decisões operacionais. Bidgoli (2018) destaca que, em ambientes *offshore*, a simulação de plantas de processamento deve considerar os limites da recuperação primária para evitar subdimensionamentos ou investimentos desnecessários. O autor ainda recomenda o uso de ferramentas de simulação numérica e análise de sensibilidade para avaliar o desempenho da produção sob diferentes cenários.

2.1.2

Recuperação secundária: técnicas de injeção para sustentação da pressão

A recuperação secundária é introduzida quando a energia natural do reservatório se torna insuficiente para manter a produção em níveis economicamente sustentáveis. Essa fase é caracterizada pela injeção de fluidos, como água ou gás, com o objetivo de manter a pressão do reservatório e deslocar o óleo remanescente em direção aos poços produtores (ABRAHAM-ADEJUMO, 2018). Ainda segundo o autor a injeção de água é amplamente empregada em reservatórios com boa conectividade e permeabilidade, sendo eficaz na recuperação de volumes adicionais de óleo.

A recuperação secundária é aplicada quando a produção primária se torna ineficaz, utilizando a injeção de água ou gás para aumentar a pressão do reservatório. Essa técnica pode atingir cerca de 33% de eficiência em relação ao volume original de óleo *in situ* (Perera et al., 2016). Além disso, conforme Abraham-Adejumo (2018), a literatura técnica destaca a importância da integração de dados sísmicos, petrofísicos e operacionais para prever o avanço da frente de fluido e evitar a canalização. A recuperação secundária, portanto, representa uma etapa estratégica para prolongar a vida útil dos campos e maximizar o retorno sobre o investimento.

Davolio et al. (2024) propuseram uma metodologia para calibrar o peso entre dados de produção e sísmica 4D, demonstrando melhorias significativas na previsão de desempenho de reservatórios. Essa abordagem é especialmente eficaz em campos maduros com histórico de injeção de água, como Roncador.

O projeto de EOR em Roncador, conduzido pela parceria entre Petrobras e Equinor, exemplifica a aplicação bem-sucedida de técnicas avançadas de recuperação. Os dois primeiros poços adicionaram cerca de 20 mil barris por dia à produção do campo, e os 18 poços planejados devem fornecer recursos recuperáveis adicionais de 160 milhões de barris. A experiência tecnológica combinada dos parceiros foi essencial para reduzir os custos em 50% nos primeiros poços (EQUINOR, 2022).

2.1.3

Recuperação terciária: métodos avançados para mobilização do óleo residual

Segundo Neves (2014), a recuperação terciária, também denominada recuperação avançada de petróleo (EOR – *Enhanced Oil Recovery*), é aplicada quando essa técnica se mostra mais adequada para aumentar a eficiência de extração do óleo presente no reservatório. Essa fase envolve o uso de agentes químicos, térmicos ou biológicos que modificam as propriedades do óleo ou da rocha, facilitando o escoamento dos hidrocarbonetos.

Estudos indicam que a recuperação terciária pode elevar a eficiência total de extração para até 60% do volume original de óleo (SANTOS, 2024). A escolha da técnica mais adequada depende de variáveis como profundidade do reservatório, composição do óleo e condições de temperatura e pressão. A literatura recente enfatiza a importância de testes piloto e simulações numéricas para avaliar a viabilidade técnica e econômica da aplicação de EOR (BIDGOLI, 2018).

Neves (2014) destaca que a injeção de vapor é particularmente eficaz em reservatórios com óleos pesados, pois reduz a viscosidade do fluido e aumenta sua mobilidade. Por outro lado, os métodos químicos, como a injeção de surfactantes

e polímeros, têm como finalidade reduzir a tensão interfacial entre o óleo e a água, melhorando a eficiência da varrida.

As técnicas de recuperação de petróleo, abordadas no contexto da engenharia de reservatórios, são fundamentais para ampliar a eficiência produtiva em campos maduros. Elas possibilitam a extração de volumes adicionais de óleo, contribuindo para a extensão da vida útil dos ativos e para a maximização do aproveitamento das jazidas (ROSA et al., 2006).

2.1.4

Estimulação de poços: técnicas para aumentar a produtividade local

A estimulação de poços é uma técnica complementar às fases de recuperação, aplicada diretamente nos poços produtores com o objetivo de aumentar sua produtividade, especialmente em formações de baixa permeabilidade. Entre os métodos mais utilizados destacam-se a acidificação, que visa dissolver obstruções minerais nos poros da rocha, e o fraturamento hidráulico, que consiste na criação de fraturas artificiais para facilitar o escoamento dos fluidos (TAKINAMI, 2024).

O fraturamento hidráulico tem sido amplamente empregado em formações não convencionais, como o *shale gas*, gás aprisionado em rochas de xisto, devido à sua capacidade de ampliar significativamente a área de drenagem do poço (NEVES, 2014). No entanto, essa técnica levanta preocupações ambientais, como o risco de contaminação de aquíferos e a indução de sismos, o que tem motivado a busca por alternativas menos invasivas (TAKINAMI, 2024). Entre essas alternativas, destaca-se a estimulação por ultrassom, que apresenta menor impacto ambiental e custos operacionais reduzidos. A escolha da técnica de estimulação deve considerar aspectos como a geologia local, a composição dos fluidos e os objetivos de produção (SANTOS, 2024). A integração de dados geomecânicos, petrofísicos e operacionais é essencial para o sucesso dessas intervenções, garantindo maior eficiência e segurança no processo produtivo.

Em estudo recente, Davolio et al. (2025) propuseram um modelo físico para estimar a rigidez da matriz rochosa em carbonatos, contribuindo para a

previsão do comportamento mecânico em reservatórios sujeitos à reinjeção de água.

2.2

Previsão da produção

Segundo Bonfim (2017), a previsão de produção é uma das etapas mais críticas do processo de planejamento de curto prazo das empresas de Exploração e Produção (E&P) de Petróleo. A precisão das previsões de produção pode influenciar diretamente a eficiência operacional, a satisfação do cliente e a rentabilidade de uma empresa (MAKRIDAKIS ET AL., 1998).

Os principais fatores que levam ao início do declínio da produção do campo são a queda de pressão, ou depleção do reservatório, e a produção de concomitante de outros fluidos (gás ou água). A produção entra em uma fase de redução gradual, que pode tornar a exploração inviável economicamente (DANTAS; PESSOA, 2020). De acordo com Souza & Silva (2016), a indústria petrolífera é caracterizada por ser intensiva em capital, apresentado altos riscos nas atividades de exploração, elevados investimentos no desenvolvimento da produção e longos períodos de maturação do investimento. Conforme esses autores, cada decisão, desde a aquisição de um bloco exploratório até o início da produção do campo, deve ser fundamentada em avaliações econômicas robustas. Obidike et. al. (2014) afirmam que a previsão da produção de um campo é um dado essencial para determinar a viabilidade econômica do projeto e, conseqüentemente, para a tomada de decisões na empresa. Segundo Oyekunle (2011), além da curva de produção, existem outros parâmetros importantes para definir a viabilidade econômica de um campo de petróleo, como a previsão dos investimentos necessários (*capital expenditures* – CAPEX) e das despesas operacionais (*operational expenditures* – OPEX), o modelo fiscal e a previsão de preços de venda dos produtos (óleo e gás).

2.2.1

Previsão de Produção no Planejamento Operacional

A previsão constitui uma etapa fundamental no planejamento operacional das empresas do setor de exploração e produção (E&P) de petróleo, uma vez que fornece subsídios técnicos e estratégicos para a tomada de decisão em ambientes marcados por elevada incerteza. Conforme argumentam Makridakis et al. (1998), a capacidade de antecipar o comportamento futuro da produção é determinante para a competitividade organizacional, especialmente em setores com alta volatilidade de mercado. Modelos estatísticos como o SARIMA, conforme demonstrado por Santos (2024), têm se mostrado eficazes na modelagem de séries temporais de produção, permitindo a identificação de padrões sazonais e tendências de longo prazo. Além disso, Hamacher e Ferreira Filho (2015) destacam que a previsibilidade da produção impacta diretamente a gestão da cadeia de suprimentos, a alocação de recursos e o planejamento logístico. No contexto *offshore*, Bidgoli (2018) ressalta que a acurácia das previsões é essencial para a otimização de plantas de processamento, evidenciando a interdependência entre previsão e operação. Dessa forma, a previsão de produção não apenas orienta decisões operacionais, mas também sustenta o planejamento estratégico e financeiro das empresas petrolíferas.

Fernandes et al. (2025) destacam que a previsão confiável da produção é essencial para o planejamento operacional e a gestão de reservatórios, especialmente em campos maduros. O autor propõe uma abordagem baseada em aprendizado de máquina, utilizando apenas dados de produção e injeção, sem depender de modelos geológicos ou propriedades de fluidos, o que torna o método ágil e aplicável em ambientes com restrições de dados.

Dessa maneira, torna-se imprescindível dispor de projeções acuradas que possibilitem às equipes operacionais atuarem de forma estruturada, assegurando a entrega de serviços com qualidade, dentro dos parâmetros de eficiência e com a alocação adequada de recursos. Para tanto, é recomendável a implementação de rotinas sistemáticas de geração de previsões, as quais devem ser incorporadas de forma contínua aos processos decisórios organizacionais. Idealmente, essas projeções devem ser produzidas com base em um volume mínimo de dados, o que

favorece sua geração de maneira ágil, simplificada e com custos reduzidos (WINTERS, 1960).

2.2.2

Fatores que Influenciam o Declínio da Produção

O declínio da produção de petróleo é um fenômeno inerente à dinâmica dos reservatórios, sendo influenciado por uma série de fatores técnicos, geológicos e operacionais. Entre os elementos mais relevantes estão a redução da pressão do reservatório, o aumento da razão gás/óleo (ROG) e a intrusão de água nos poços produtores, os quais comprometem a eficiência da recuperação de hidrocarbonetos (DANTAS; PESSOA, 2020). A razão óleo/gás é um indicador técnico que expressa a proporção entre o volume de óleo e o volume de gás produzido em um reservatório, sendo utilizada para avaliar o tipo de fluido predominante e orientar estratégias de produção (DOU; WEN; WANG, 2024).

A análise de séries temporais, conforme demonstrado por Santos (2024), permite identificar padrões de declínio e antecipar a necessidade de intervenções técnicas. Abraham-Adejumo (2018) enfatiza que parâmetros como o corte de água e a transmissibilidade do reservatório são cruciais para a avaliação da continuidade da produção. Além disso, falhas na gestão integrada do reservatório, corrosão de equipamentos e alterações nas condições operacionais podem acelerar o processo de declínio.

A literatura especializada recomenda o emprego de técnicas de recuperação secundária e terciária, como a injeção de água ou gás, para mitigar os efeitos do esgotamento natural e prolongar a vida útil dos campos (ORDÓÑEZ, 2025). A compreensão aprofundada desses fatores é, portanto, essencial para a sustentabilidade técnica e econômica dos projetos de E&P.

A iniciativa de recuperação avançada em Roncador também visa mitigar os efeitos do declínio natural da produção, com foco em aumentar os recursos recuperáveis em até 1 bilhão de barris de óleo equivalente. Essa estratégia está alinhada com a tendência de prolongar a vida útil dos campos maduros por meio de inovação tecnológica e parcerias estratégicas (EQUINOR, 2022).

2.2.3

Previsão de Séries Temporais de Produção de Petróleo

Modelos estatísticos como o Suavizamento Exponencial Simples (SES), Holt, Holt-Winters, ETS e ARIMAX são amplamente utilizados para identificar padrões de produção, tendências e sazonalidades, permitindo maior precisão na tomada de decisão. Estudos recentes indicam que modelos que incorporam decomposição de séries temporais — como o ETS — apresentam desempenho superior ao extrair componentes de tendência e sazonalidade de forma explícita, o que é particularmente relevante em ambientes com comportamento cíclico ou sazonal (SONG et al., 2025, p. 51).

A produção de petróleo em campos maduros frequentemente apresenta padrões cíclicos e não estacionários. Modelos como Holt-Winters e ETS são eficazes na captura de tendências e sazonalidades, pois conseguem identificar padrões repetitivos e variações ao longo do tempo. A revisão destaca que modelos capazes de lidar com dependências de longo prazo e padrões sazonais são mais adequados para tarefas de previsão em séries temporais complexas (SU et al., 2025, p. 8).

No setor energético, a previsão de séries temporais é utilizada para planejar estratégias de produção e consumo, com destaque para fontes como eletricidade, energia solar e eólica. A produção de petróleo, por exemplo, depende fortemente da capacidade de prever padrões de longo prazo e sazonalidades, especialmente em campos maduros offshore (KONG et al., 2025, p. 5079).

A aplicação desses modelos permite identificar padrões de demanda e oferta, sendo diretamente aplicável à indústria do petróleo (KONG et al., 2025, p. 5100). Essa abordagem é diretamente aplicável à indústria do petróleo, onde prever a produção futura de campos maduros é essencial para decisões operacionais e de investimento (SU et al., 2025, p. 15).

Além disso, modelos como o ARIMAX, que incorporam variáveis exógenas, continuam sendo relevantes, especialmente quando fatores operacionais ou ambientais influenciam diretamente a produção. A análise comparativa entre esses modelos, com base em métricas de erro, permite selecionar aquele com

maior capacidade preditiva, contribuindo para o planejamento de longo prazo e a otimização da produção em campos maduros (SONG et al., 2025, p. 50–51).

A seleção do modelo mais adequado para previsão deve considerar métricas que são amplamente utilizadas para avaliar a acurácia preditiva. Essas métricas permitem comparar o desempenho de modelos como ARIMAX, ETS e Holt-Winters em diferentes horizontes de previsão, sendo fundamentais para aplicações práticas em campos de petróleo (SU et al., 2025, p. 18).

2.2.4

Aspectos Econômicos e Decisões Baseadas na Previsão

Moraes (2013) argumenta que, no contexto do pré-sal brasileiro, a previsão de produção é essencial para estimar os impactos econômicos de longo prazo, inclusive sobre a arrecadação fiscal e a balança comercial. Maluf (2023) reforça que decisões como a perfuração de novos poços e a contratação de serviços logísticos dependem diretamente da confiabilidade dos modelos preditivos.

Além disso, Ordóñez (2025) alerta que a ausência de planejamento baseado em previsões robustas pode levar à dependência externa e à instabilidade fiscal em países exportadores. A previsão também subsidia a formulação de políticas de hedge e a negociação de contratos de fornecimento, contribuindo para a mitigação de riscos financeiros em um mercado caracterizado por oscilações frequentes nos preços do barril.

2.3

Séries Temporais

De acordo com Hyndman et al. (2008), séries temporais podem ser descritas como uma composição de quatro componentes principais: tendência, sazonalidade, ciclo e erro. A tendência representa o comportamento de crescimento ou decrescimento da série ao longo do tempo. A sazonalidade refere-se a flutuações que se repetem em intervalos regulares, como meses ou trimestres. Ainda segundo o autor, de forma semelhante, o ciclo também envolve padrões de repetição, porém sem uma periodicidade definida. Por fim, o erro corresponde aos

fatores imprevisíveis que afetam a série, representando a variabilidade não explicada pelos demais componentes.

Além desses elementos, é importante considerar o nível da série, que representa o valor médio em torno do qual os dados oscilam. Quando esse nível é acompanhado por um padrão de crescimento ou declínio, forma-se a tendência. Portanto, uma série que apresenta apenas o nível, sem variações sistemáticas ao longo do tempo, não possui tendência (HYNDMAN et al., 2008).

A depender da abordagem utilizada para previsão, essas componentes podem ser organizadas em estruturas de espaço de estados. Nessa representação, os parâmetros são armazenados em vetores que descrevem o estado do sistema em diferentes momentos, permitindo a modelagem dinâmica da série ao longo do tempo (HYNDMAN et al., 2008).

Séries temporais consistem em valores numéricos registrados ao longo do tempo, com intervalos regulares. As séries temporais podem ser usadas para inferir valores futuros baseadas nas análises de correlação dos dados coletados. As observações vizinhas são geralmente dependentes, o que torna a ordem dos dados crucial para análise. Análise de séries temporais utiliza dados históricos para criar modelos de previsão, buscando identificar padrões que permitam fazer previsões precisas (FERREIRA & MATIAS, 2011). Como parte das análises da viabilidade econômica, os modelos de previsão usando séries temporais permitem ter estimativas dos valores futuros da produção (SANTOS, 2024).

No contexto do presente trabalho, Marques (2023) empregou técnicas de previsão usando séries temporais das variações históricas da produção de petróleo nas operações de poços, utilizando dados de 65 poços situados na Bacia Denver-Julesburg, nos Estados Unidos.

Nas análises preditivas usando séries temporais existem diversos métodos estatísticos que podem ser usados, entre os quais podem-se citar diversos modelos estatísticos para a previsão de séries temporais, cada um com suas características e aplicações específicas e chegando até os mais sofisticados como *ETS* são utilizados para capturar padrões mais elaborados em séries temporais, incluindo tendências e sazonalidades (JAIN, 2012).

2.3.1

Decomposição STL

Segundo Carmo (2017), a decomposição *Seasonal and Trend Decomposition using Loess* (STL) é uma técnica que permite separar uma série temporal em seus três componentes principais de tendência, sazonalidade e resíduo (CARMO, 2017). Essa abordagem é descrita pela Equação 01:

$$Y_t = T_t + S_t + R_t \quad (01)$$

Em que Y_t é o valor observado da série temporal no tempo t , T_t é componente de tendência, mostra o comportamento de longo prazo, S_t é o componente sazonal, representa padrões que se repetem periodicamente e R_t é o componente residual, inclui variações aleatórias ou não explicadas.

Essa decomposição facilita a análise da série temporal ao isolar o comportamento individual de cada componente. Não é esperado que se observe um padrão sazonal na produção de petróleo, embora pudesse haver na demanda do seu mercado, mas optou-se por manter os modelos que também tratam de sazonalidade junto com os demais, oferecendo um quadro mais completo das ferramentas, deixando que sejam descartados naturalmente, conforme os resultados.

A propriedade de estacionariedade em uma série temporal refere-se à estabilidade de seu comportamento estatístico ao longo do tempo. Em termos práticos, isso significa que, ao se comparar dois intervalos distintos, porém de mesma duração, os padrões observados devem ser semelhantes. Nesses casos, o valor esperado da série permanece constante, independentemente do momento em que é observado (MONTGOMERY et al., 2008). Modelos estacionários assumem que o processo analisado opera sob um equilíbrio estatístico, no qual as características probabilísticas, como média e variância, não se alteram com o tempo. Tais modelos são reconhecidos por apresentarem oscilações em torno de um valor médio fixo e por manterem uma variância constante ao longo de toda a série (BOX et al., 2016).

Séries temporais que exibem tendência ou sazonalidade não atendem ao critério de estacionariedade, pois essas componentes introduzem variações sistemáticas ao longo do tempo. Em contrapartida, séries com padrões cíclicos podem ser consideradas estacionárias, desde que esses ciclos não apresentem periodicidade fixa e ocorram de forma intermitente (HYNDMAN; ATHANASOPOULOS, 2021).

2.3.2

Horizonte x Intervalo de Previsão

De acordo com Montgomery (2008), no processo de construção de previsões estatísticas, dois elementos fundamentais devem ser definidos previamente: o horizonte e o intervalo de previsão. O horizonte, também conhecido como *lead time*, refere-se à quantidade de períodos futuros que se deseja estimar, enquanto o intervalo diz respeito à frequência com que as previsões serão atualizadas.

A escolha adequada desses parâmetros deve considerar os objetivos estratégicos das áreas que utilizarão os dados, garantindo que as projeções geradas sejam úteis e relevantes para a tomada de decisão. Em contextos operacionais, é comum adotar uma abordagem de horizonte móvel, na qual as previsões são atualizadas em intervalos regulares, mesmo que o horizonte de previsão seja mais longo (MONTGOMERY et al., 2008).

Essa estrutura permite que as organizações se adaptem a mudanças no ambiente de negócios, mantendo a acurácia das estimativas ao longo do tempo. Além disso, a definição do horizonte e do intervalo está diretamente relacionada à natureza da série temporal e ao comportamento das variáveis envolvidas. Segundo Box e Jenkins (1976), a eficácia de um modelo preditivo depende não apenas da qualidade dos dados, mas também da adequação do modelo ao contexto temporal e operacional em que será aplicado. Assim, a integração entre planejamento estratégico e modelagem estatística é essencial para garantir previsões consistentes e alinhadas às necessidades organizacionais.

2.3.3

Função de Autocorrelação

A função de autocorrelação (FAC) é uma ferramenta estatística essencial na análise de séries temporais, utilizada para quantificar o grau de dependência linear entre os valores de uma série e suas próprias defasagens (ou *lags*). Em termos práticos, a FAC mede a correlação entre a série original e versões deslocadas dela mesma ao longo do tempo, permitindo identificar padrões estruturais como tendência, ciclos e, especialmente, sazonalidade (VENABLES; RIPLEY, 2002; BOX et al., 2015). Matematicamente, a autocorrelação no *lag* k é definida como:

$$\rho_k = \frac{n - \min(k) \sum_{t=1}^{n-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\max(k) - \min(k) \sum_{t=1}^n (y_t - \bar{y})^2} \quad (02)$$

Onde y_t representa o valor da série no tempo t , $\bar{y}$ é a média da série, e T é o número total de observações. Essa medida varia entre -1 e 1, indicando respectivamente correlação negativa perfeita, ausência de correlação e correlação positiva perfeita.

A análise da FAC é particularmente útil na etapa de diagnóstico e especificação de modelos de séries temporais, como ARIMA e ARIMAX. Por exemplo, a presença de autocorrelações significativas em *lags* específicos pode sugerir a necessidade de incluir termos autorregressivos (AR) ou de média móvel (MA) no modelo (HYNDMAN; ATHANASOPOULOS, 2021).

Além disso, padrões periódicos na função de autocorrelação (FAC) podem indicar a presença de sazonalidade, mesmo quando não explicitamente visível na série original. Já a função de autocorrelação parcial (FACP) é uma ferramenta estatística utilizada na análise de séries temporais para medir a correlação entre os valores da série e suas defasagens, eliminando a influência das defasagens intermediárias. Ao contrário da FAC, que considera todos os efeitos acumulados, a FACP revela apenas o impacto direto de cada *lag* sobre o valor atual da série, sendo especialmente útil na identificação da ordem dos componentes

autorregressivos (AR) em modelos como o ARIMA. Por exemplo, se a FACP mostra correlação significativa até o *lag* 2 e insignificante a partir do *lag* 3, isso sugere que um modelo AR(2) pode ser adequado (BOX; JENKINS, 2008).

Em conjunto com a FAC, a FACP permite construir modelos mais precisos, ajudando a distinguir entre padrões reais e ruído, e contribuindo para previsões mais confiáveis, ou seja, a função de autocorrelação parcial (FACP) é frequentemente utilizada em conjunto com a FAC para distinguir entre efeitos diretos e indiretos de defasagens múltiplas, sendo fundamental na identificação da ordem dos componentes autorregressivos (AR) que representa a influência dos valores passados da própria série sobre o valor atual. e média móvel (MA) que representa a influência dos erros de previsão anteriores sobre o valor atual (MONTGOMERY et al., 2008).

Em contextos industriais, como a previsão de produção de petróleo, a FAC permite avaliar a persistência temporal da série e a presença de choques estruturais ou mudanças de regime (BOX et al., 2015). Segundo Abraham-Adejumo (2018), a análise de autocorrelação é particularmente relevante em campos maduros, onde a variabilidade da produção pode refletir tanto fatores operacionais quanto fenômenos geológicos.

Por fim, a interpretação da FAC deve considerar os limites de significância estatística, geralmente representados por bandas de confiança de 95% em gráficos de correlograma. Autocorrelações que ultrapassam essas bandas indicam dependência estatisticamente significativa, justificando a inclusão de componentes adicionais no modelo preditivo (SHUMWAY; STOFFER, 2017).

2.4

Métodos de previsão

Conforme Hyndman et al. (2008), os modelos estatísticos de previsão permitem a construção de uma distribuição de probabilidade completa para projeções futuras, baseada em um processo estocástico subjacente. Essa distribuição contempla os diversos cenários possíveis para a variável em análise, e a partir dela é possível extrair uma previsão pontual, como a média ou a mediana, que representa o valor esperado para um período futuro.

Já Montgomery, Jennings e Kulahci (2008) destacam que esses modelos também possibilitam a estimativa de intervalos de previsão, os quais indicam, com um nível de confiança previamente definido, a faixa na qual os valores futuros tendem a se situar. Esses intervalos refletem a incerteza inerente ao processo de previsão, sendo fundamentais para a tomada de decisão em ambientes sujeitos a variabilidade.

Ao incorporar tanto a estimativa pontual quanto os intervalos de previsão, os modelos estatísticos oferecem uma abordagem mais completa e informativa, pois consideram não apenas o valor esperado, mas também a dispersão dos possíveis resultados (HYNDMAN et al., 2008). Nesse cenário, os modelos *ETS* e *ARIMA* destacam-se como ferramentas amplamente reconhecidas e eficazes na modelagem e previsão de séries temporais, especialmente em contextos hierárquicos (HYNDMAN et al., 2024).

Diante disso, este capítulo aborda os modelos de Suavização Exponencial (SES, Holt, Holt-Winters), o modelo Error, Trend, Seasonality (ETS) e a Regressão Dinâmica (ARIMAX).

2.4.1

Suavização exponencial (SES, Holt, Holt-Winters)

Segundo Hyndman et al. (2008), a Suavização Exponencial é uma técnica que aplica uma média ponderada aos dados históricos, atribuindo maior peso aos valores mais recentes. Montgomery e Jennings (2008) destacam que a suavização exponencial separa o padrão do ruído, permitindo uma estimativa mais precisa. Essa abordagem é amplamente reconhecida por sua simplicidade e eficácia.

Existem três principais variações do método de suavização exponencial: a Suavização Exponencial Simples (SES), adequada para séries sem tendência ou sazonalidade, que atribui pesos exponencialmente decrescentes às observações passadas (HYNDMAN; ATHANASOPOULOS, 2021); o Modelo de Holt, que estende o SES para séries com tendência linear (HOLT, 2004); e o Modelo Holt-Winters, que incorpora componentes sazonais à estrutura de Holt, permitindo previsões mais precisas em séries com sazonalidade (WINTERS, 1960). A Suavização Exponencial Simples (SES) é uma técnica de previsão que aplica uma

média ponderada aos dados históricos, dando mais peso aos dados mais recentes. É ideal para séries temporais sem tendência ou sazonalidade (LIMAVERDE FILHO et al., 2022). O Modelo Holt estende a SES ao incorporar uma tendência linear. De acordo com Silva (2023), esse modelo é adequado para séries com tendência constante ao longo do tempo, oferecendo previsões mais realistas do que a SES em contextos com crescimento ou declínio.

O Modelo Holt-Winters é uma extensão do modelo de Holt, que incorpora sazonalidade, sendo ideal para séries sazonais bem definidas (SOUZA, 1983). Jain (2012) e Hyndman e Athanasopoulos (2018) destacam sua ampla aplicação em previsões de demanda devido à sua flexibilidade e precisão. Embora o modelo Holt-Winters tenha sido incluído por completude teórica, sua aplicabilidade foi limitada neste caso, dada a ausência de sazonalidade significativa na série de produção. A formulação deste modelo segue o padrão da Equação 03:

$$\hat{Z}_{T+h} = N_t + hT_t + S_{t+h-m(k+1)} \quad (03)$$

$$N_t = \alpha(Z_t - S_{t-m}) + (1 - \alpha)(N_{t-1} + T_{t-1}) \quad (3.1)$$

$$T_t = \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \quad (3.2)$$

$$S_t = \gamma(Z_t - N_{t-1} - T_{t-1}) + (1 - \gamma)S_{t-m} \quad (3.3)$$

$$k = \text{mod}\left(\frac{h-1}{m}\right) \quad (3.4)$$

Onde Z_{T+h} é o valor previsto da série temporal para o tempo futuro $T+h$; N_t é o nível da série no tempo t ; T_t é a tendência da série no tempo t ; S_t é a sazonalidade da série no tempo t ; α é o parâmetro de suavização de nível ($0 \leq \alpha \leq 1$); β é o parâmetro de suavização de tendência ($0 \leq \beta \leq 1$); γ é o parâmetro de suavização de sazonalidade ($0 \leq \gamma \leq 1$); h é o horizonte de previsão; e m é o período de sazonalidade.

Neste método, os parâmetros α , β e γ são ajustados por meio da minimização da soma dos erros quadrados entre os valores previstos e observados, prática comum na modelagem de séries temporais.

2.4.2

Modelo Error, Trend, Seasonality (ETS)

O modelo *ETS* combina suavização exponencial com uma estrutura de espaço de estados, permitindo representar diferentes combinações de erro, tendência e sazonalidade (*Error, Trend, Seasonality*). Ele pode assumir formas mais aditivas ou multiplicativas, sendo ajustado automaticamente aos dados (HYNDMAN et al., 2008).

Diversos métodos eficazes de previsão de séries temporais fundamentam-se na técnica de Suavização Exponencial, reconhecida por sua simplicidade operacional e desempenho robusto em diferentes contextos aplicados (HYNDMAN et al., 2008). Conforme apontam Montgomery e Jennings (2008), a suavização consiste, essencialmente, em um procedimento que isola o padrão sistemático do ruído aleatório presente nos dados, permitindo que o modelo funcione como um filtro, resultando em estimativas mais precisas do comportamento subjacente da série.

A Suavização Exponencial, ou *Exponential Smoothing*, é um algoritmo que calcula médias ponderadas dos dados históricos, atribuindo pesos decrescentes às observações à medida que se distanciam no tempo. Dessa forma, a influência das informações passadas diminui exponencialmente, o que torna o método particularmente eficaz para capturar padrões recentes e projetá-los em horizontes futuros (WINTERS, 1960).

Holt (2004) destaca como principais vantagens dessa técnica sua flexibilidade, baixo custo computacional e reduzida exigência de grandes volumes de dados históricos. Tais características tornam a Suavização Exponencial especialmente adequada para ambientes como o varejo, nos quais há demanda por múltiplas previsões simultâneas e, frequentemente, limitação no histórico de dados disponível.

O modelo *ETS* (*Erro, Tendência e Sazonalidade*), por sua vez, estrutura-se a partir da decomposição da série temporal nesses três componentes fundamentais. Cada um deles pode ser especificado de forma aditiva ou multiplicativa, resultando em diferentes configurações possíveis (HYNDMAN et al., 2008). Na Tabela 1 são apresentadas as diferentes combinações possíveis para

as componentes de tendência e sazonalidade. A componente de erro, por sua vez, pode ser incorporada a cada uma dessas configurações tanto de forma aditiva quanto multiplicativa. A combinação entre essas variações gera até 30 modelos distintos, incluindo os clássicos modelos de Holt e Holt-Winters. Embora modelos com erro aditivo e multiplicativo possam produzir previsões pontuais idênticas, eles diferem na construção dos intervalos de confiança, o que impacta diretamente na avaliação da incerteza associada às projeções (HYNDMAN et al., 2008).

Tabela 1: Estrutura de Modelos Temporais: Tendência, Sazonalidade e Erro.

Tendência (T)	Sazonalidade		
	Nenhuma (N)	Aditivo (A)	Multiplicativo (M)
Nenhuma (N)	(N,N)	(N,A)	(N,M)
Aditiva (A)	(A,N)	(A,A)	(A,M)
Aditivo amortecido (Aa)	(Aa,N)	(Aa,A)	(Aa,M)
Multiplicativa (M)	(M,N)	(M,A)	(M,M)
Multiplicativo (Ma)	(Ma,N)	Ma,A)	(Ma,M)

Fonte: Adaptado de Hyndman e Athanasopoulos (2021).

A análise da Tabela 1 permite identificar que, em determinadas configurações, as componentes de tendência e sazonalidade podem estar ausentes. Em outros casos, essas componentes se manifestam de forma aditiva ou multiplicativa. Além disso, há situações em que a tendência é modelada com amortecimento, uma estratégia desenvolvida para representar séries que apresentam crescimento ou declínio, mas cuja continuidade desse comportamento ao longo do tempo é incerta. Nesses casos, quanto mais distante estiver o horizonte de previsão, menor será a influência da tendência observada nos dados históricos (HYNDMAN et al., 2008).

A configuração mais simples do modelo *ETS* é aquela em que não há tendência nem sazonalidade, conhecida como Suavização Exponencial Simples (SES). Nesse modelo, a previsão para o próximo período baseia-se na média ponderada entre o valor observado e a previsão anterior. Embora seja uma abordagem amplamente utilizada devido à sua simplicidade e adaptabilidade, sua

aplicação pode comprometer a acurácia das previsões em séries que apresentam padrões de tendência ou sazonalidade. Portanto, seu uso é recomendado apenas quando tais componentes não estão presentes de forma significativa. Na prática, a previsão para o tempo $T+1$ corresponde ao nível estimado no período T , calculado por meio da Equação (4), na qual o parâmetro α , com valor entre 0 e 1, determina o peso atribuído às observações mais recentes (WINTERS, 1960).

$$\hat{Y}_{T+1|T} = L_t = \alpha Y_t + (1-\alpha)L_{t-1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \dots + \alpha(1-\alpha)^n Y_{t-n} \quad (04)$$

Segundo Silva (2023), o modelo *ETS* é altamente flexível e eficaz para capturar padrões complexos em séries temporais, sendo uma alternativa robusta aos modelos tradicionais de suavização. O modelo *ETS* pode ser representado pela seguinte Equação:

$$\text{Observação: } Y_t = \ell_{t-1} + b_{t-1} + S_{t-m} + \varepsilon_t \quad (05)$$

$$\text{Nível: } \ell_t = \ell_{t-1} + b_{t-1} + \alpha \varepsilon_t \quad (06)$$

$$\text{Tendência: } b_t = b_{t-1} + \beta \varepsilon_t \quad (07)$$

$$\text{Sazonalidade: } S_t = S_{t-m} + \gamma \varepsilon_t \quad (08)$$

Onde: Y_t = valor observado na série no tempo t ; ℓ_t = nível da série; b_t = tendência; S_t = componente sazonal; ε_t = erro aleatório; α, β, γ = parâmetros de suavização; m = período da sazonalidade.

2.4.3

ARIMA e Regressão dinâmica (ARIMAX)

O ARIMA, conforme descrito por Box & Jenkins (1976), Hyndman & Athanasopoulos (2018) e Morettin & Toloí (2006), é indicado para séries temporais que apresentam autocorrelação e não estacionariedade, permitindo capturar padrões de tendência e sazonalidade por meio de componentes autorregressivos (AR), de integração (I) e de médias móveis (MA). Esse modelo é particularmente útil quando se deseja prever uma série com base exclusivamente em seus valores passados, sem considerar variáveis externas.

A regressão dinâmica é um tipo de modelo que integra a análise de séries temporais com os efeitos de variáveis explicativas exógenas. Esse tipo de abordagem é especialmente indicado quando há uma relação de dependência entre a variável de interesse e as variáveis causais, e quando a estrutura de correlação da série temporal indica que não se pode assumir a independência dos erros (BOX et al., 2015; GOODRICH, 1989;).

O modelo ARIMAX (*AutoRegressive Integrated Moving Average with Exogenous Variables*) representa uma extensão dos modelos ARIMA, incorporando variáveis explicativas externas (exógenas) ao processo de modelagem de séries temporais. Essa abordagem é particularmente eficaz quando se deseja capturar não apenas a estrutura autocorrelacionada da série temporal, mas também os efeitos de variáveis externas que influenciam diretamente a variável de interesse (BOX; JENKINS, 1976). Segundo Box et al. (2015), o ARIMAX é apropriado quando os resíduos do modelo não podem ser considerados independentes e quando há evidência de causalidade entre variáveis. Goodrich (1989) já havia destacado que a regressão dinâmica é útil para modelar relações entre séries temporais e variáveis exógenas, especialmente em contextos econômicos e industriais.

Estudos recentes reforçam a aplicabilidade do ARIMAX em diferentes domínios. Chian (1993) também explora a regressão dinâmica como ferramenta para análise de variância em séries temporais, destacando sua utilidade na modelagem de dependência temporal com múltiplos regressores. Além disso, Nacinben (2023) discute a incorporação de variáveis exógenas em modelos de

volatilidade estocástica, o que reforça a versatilidade do ARIMAX em aplicações financeiras e estatísticas.

O estudo de Fernandes et al. (2025) reforça a importância da escolha adequada das janelas de observação e da reconfiguração periódica dos modelos preditivos, especialmente em séries temporais sujeitas a deriva de conceito. A abordagem proposta inclui regressões lineares com regularização, regressão Ridge, introduzida por Hoerl e Kennard (1970), aplica uma penalização quadrática aos coeficientes, sendo especialmente útil em situações de multicolinearidade. Já a regressão Lasso, desenvolvida por Tibshirani (1996), utiliza uma penalização absoluta que permite a seleção automática de variáveis, tornando o modelo mais parcimonioso. Ambas as técnicas têm sido amplamente aplicadas em contextos de séries temporais e aprendizado de máquina, inclusive em cenários com deriva de conceito. Estudos recentes, como o de Passos et al. (2021), reforçam a importância da escolha adequada das janelas de observação e da reconfiguração periódica dos modelos preditivos, destacando o uso de janelas móveis e reamostragem temporal como estratégias para melhorar a acurácia em ambientes dinâmicos. A estrutura do modelo ARIMAX permite que os coeficientes das variáveis exógenas sejam estimados simultaneamente com os parâmetros autorregressivos e de média móvel, o que o torna uma ferramenta robusta para previsão e inferência. A modelagem adequada requer testes de estacionariedade, identificação da ordem dos componentes ARIMA e análise da significância das variáveis exógenas (YAHYAOUUI et al., 2022).

Davolio et al. (2023) propuseram uma abordagem baseada em importância de atributos e modelos de ensemble para descobrir relações causais em séries temporais. Essa técnica pode ser aplicada na previsão de produção em Roncador, ao identificar variáveis exógenas com maior impacto na série alvo.

Conforme pode ser observado pelos estudos citados, o ARIMAX é uma alternativa robusta para modelagem de séries temporais com múltiplos fatores explicativos, aplicada em muitos estudos clássicos e contemporâneos.

O termo “dinâmico”, conforme Korstanje (2021), refere-se à inclusão de defasagens na série temporal e nas variáveis causais, e não à variação dos parâmetros ao longo do tempo. Esses modelos podem ser representados, de forma resumida, pela seguinte Equação (09):

$$\varphi(B)y_t = \beta x_t + F_t \quad (09)$$

Onde:

- y_t = variável dependente no instante t ;
- β = vetor de coeficientes das variáveis causais, que será estimado por mínimos quadrados;
- x_t = vetor de variáveis causais (exógenas) no instante t independentes e identicamente distribuídos com densidade $N(0, \sigma^2)$;
- $\varphi(B)$ = polinômio autorregressivo de ordem p definido conforme Equação (10), sendo B o operador de atraso (BOX et al., 2015).

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad (10)$$

A equação completa do modelo ARIMAX pode ser expressa como:

$$y_t = \alpha + \sum_{i=1}^p \varphi_i y_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{k=0}^r \beta_k x_{t-k} + \varepsilon_t \quad (11)$$

(HYNDMAN & ATHANASOPOULOS, 2018).

Onde:

- y_t representa o valor do item que se deseja prever no tempo t ;
- x_{t-k} representa a variável explicativa exógena com impacto no resultado com defasagem k ;
- φ_i são os coeficientes autorregressivos (AR);
- θ_j são os coeficientes da média móvel (MA);
- β_k são os coeficientes associados à variável explicativa exógena;
- ε_t é o termo de erro branco.

Esse modelo é particularmente útil quando se deseja prever uma série temporal com base não apenas em seus próprios valores passados, mas também em

variáveis externas que influenciam seu comportamento (HYNDMAN & ATHANASPOULOS, 2018; MORETTIN & TOLOI, 2006).

2.5

Métricas de avaliação de modelos

As métricas de avaliação de modelos de previsão desempenham um papel fundamental na análise de desempenho de algoritmos preditivos, sendo ferramentas estatísticas essenciais para quantificar a discrepância entre os valores observados (ou reais) e aqueles estimados por um modelo. Essas métricas permitem não apenas avaliar a acurácia das previsões realizadas, mas também comparar diferentes modelos entre si, auxiliando na escolha da abordagem mais adequada para determinado problema ou conjunto de dados. Dentre as mais utilizadas estão Erro Quadrático Médio, Erro Percentual Absoluto Médio (*Mean Absolute Percentage Error* - MAPE) e o Erro Quadrático Médio (*Mean Squared Error* - MSE) (MAKRIDAKIS, WHEELWRIGHT & HYNDMAN, 1998).

2.5.1

Mean absolute percentage error (MAPE)

Segundo Tichy (2023), o Erro Percentual Absoluto (APE – *Absolute Percentage Error*) é determinado pela razão entre o desvio absoluto da previsão em relação ao valor real e o próprio valor real. O MAPE, por sua vez, corresponde à média dos APEs calculados ao longo de todo o horizonte de previsão. O MAPE mede a precisão das previsões em termos percentuais. Um MAPE baixo indica previsões mais precisas, sendo um indicador crucial na seleção do modelo mais adequado para uma determinada série temporal (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998). Segue a fórmula do MAPE:

$$\text{MAPE} = \frac{100}{n} \sum \frac{|e_t|}{y_t} \quad (12)$$

Onde:

n : representa o número total de períodos ou observações no horizonte de

previsão;

e_t : erro de previsão no tempo t , ou seja, a diferença entre o valor previsto e o valor real;

y_t : valor real observado da variável no tempo;

Σ : soma dos erros percentuais absolutos ao longo de todos os períodos.

De acordo com Hyndman e Athanasopoulos (2021) e Tichy (2023), embora o MAPE seja amplamente adotado, especialmente em ambientes corporativos, devido à sua linguagem acessível e à facilidade de interpretação dos resultados, sua aplicação apresenta limitações importantes. Em contextos com volumes reduzidos de dados ou em séries temporais com comportamento intermitente, o desempenho dessa métrica pode ser significativamente comprometido. Além disso, o MAPE possui a tendência de amplificar determinados ruídos enquanto oculta outros, o que pode levar os tomadores de decisão a preferirem modelos com viés sistemático, comprometendo a qualidade das previsões.

Tichy (2023) ressalta, portanto, que o MAPE não deve ser utilizado de forma isolada ou sem considerar o contexto específico da análise. Ainda assim, uma de suas principais vantagens é a independência de escala, o que permite a comparação entre erros de previsão de diferentes categorias, mesmo quando os dados analisados estão em escalas distintas. Essa característica torna o MAPE particularmente útil em análises comparativas entre diferentes segmentos ou produtos.

2.5.2

Mean squared error (MSE)

O *Mean Squared Error* (MSE) é uma métrica estatística amplamente utilizada para mensurar a precisão de modelos de regressão. O MSE é obtido calculando-se a média dos quadrados das diferenças entre os valores previstos e os valores reais observados nos dados de treinamento, sendo sensível a grandes

desvios e amplamente utilizado na avaliação de modelos de regressão (LUKE; PORTER; SANTHANAM, 2022).

O cálculo do MSE consiste em obter a média dos quadrados das diferenças entre os valores estimados pelo modelo e os valores efetivamente observados nos dados de treinamento ou teste. Essa abordagem de elevação ao quadrado tem como principal consequência a penalização mais severa de erros maiores, o que torna o MSE particularmente sensível a outliers ou desvios extremos (WILLMOTT; MATSUURA, 2005). Em outras palavras, quanto maior a discrepância entre a previsão e o valor real, maior será a contribuição desse erro para o valor final do MSE.

Essa característica pode ser vantajosa em contextos em que grandes erros são inaceitáveis e devem ser fortemente penalizados, como em aplicações financeiras, médicas ou industriais. No entanto, essa mesma sensibilidade pode ser uma limitação em cenários onde os dados apresentam variabilidade natural ou ruídos que não devem ser supervalorizados (CHAI; DREW, 2014).

Além disso, o MSE é uma métrica dependente da escala dos dados, o que significa que seu valor absoluto não pode ser diretamente comparado entre diferentes conjuntos de dados ou variáveis com unidades distintas. Por esse motivo, é comum que o MSE seja complementado por outras métricas, como o *Root Mean Squared Error* (RMSE) e o *Mean Absolute Error* (MAE), que oferecem diferentes perspectivas sobre o desempenho do modelo. O MAE, por exemplo, é menos sensível a pontos fora da curva (*outliers*), enquanto o RMSE mantém a penalização de grandes erros, mas com a vantagem de estar na mesma unidade da variável de interesse (CHAI; DREW, 2014). O MSE pode ser calculado pela Equação (13).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

Onde:

n : número total de observações

y_i : valor observado (real) na posição i

$\hat{y}_i$:valor previsto (estimado) na posição i

$(y_i - \hat{y}_i)^2$: erro quadrático entre o valor observado e o previsto

Σ : somatório dos erros quadráticos para todas as observações

$\frac{1}{n}$: média dos erros quadráticos

2.6

Modelos Preditivos em Campos Maduros

A previsão de produção em campos maduros é um desafio estratégico para a indústria de óleo e gás, especialmente em ambientes *offshore*, onde o declínio natural dos reservatórios e a variabilidade operacional exigem modelos robustos e adaptáveis. A literatura apresenta diferentes abordagens, que podem ser agrupadas em três categorias: modelos estatísticos clássicos, modelos com variáveis exógenas e modelos baseados em aprendizado de máquina e híbridos. Este estudo concentra-se nos modelos estatísticos, por sua aplicabilidade prática e menor custo computacional, mas reconhece outras abordagens como parte do estado da arte.

2.6.1

Contexto e Relevância dos Campos Maduros

Campos maduros apresentam características específicas que impactam diretamente a previsão de produção, como declínio acentuado, aumento da razão água-óleo (RAO) e maior complexidade operacional. Esses fatores tornam a previsão uma ferramenta essencial para decisões estratégicas e operacionais, permitindo otimizar recursos e mitigar riscos (DANTAS; PESSOA, 2020). A Bacia de Campos, onde se localiza o Campo de Roncador, é um exemplo emblemático, pois concentra ativos com alta maturidade e relevância econômica (ANP, 2024).

2.6.2

Modelos Estatísticos Aplicados à Previsão

Os modelos estatísticos são amplamente utilizados para previsão de séries temporais na indústria de petróleo, devido à sua simplicidade, baixo custo computacional e capacidade de capturar padrões históricos. Entre os mais conhecidos estão os métodos de *Simple Exponential Smoothing* (SES), *Holt's Linear Trend Method* (Holt), *Holt-Winters Exponential Smoothing* (Holt-Winters), *Error, Trend e Seasonality* (ETS) e os modelos da família ARIMA (AutoRegressive Integrated Moving Average), que se destacam pela flexibilidade na modelagem de séries com tendência e sazonalidade (HYNDMAN et al., 2008; BOX et al., 2015).

Em campos maduros, onde há declínio acentuado da produção e aumento da RAO, a previsão baseada apenas em séries endógenas pode ser insuficiente. Por isso, estudos recentes apontam para a necessidade de incorporar variáveis exógenas que reflitam a dinâmica operacional do reservatório. O modelo ARIMAX, uma extensão do ARIMA, permite incluir variáveis externas, como a produção de água, melhorando significativamente a acurácia das previsões (OYINEBIFUN et al., 2024; ABRAHAM-ADEJUMO, 2018). Pesquisas internacionais reforçam essa abordagem, validando sua aplicabilidade em ambientes complexos.

2.6.3

Limitações, Lacunas e Perspectivas Futuras

Apesar da ampla utilização, os modelos estatísticos apresentam limitações, como dependência de séries históricas extensas, sensibilidade a mudanças abruptas e ausência de aprendizado adaptativo (HYNDMAN; ATHANASOPOULOS, 2021). A literatura indica escassez de estudos aplicando ARIMAX com variáveis exógenas em ambientes offshore brasileiros, especialmente em campos maduros, onde grande parte das pesquisas concentra-se em modelos geológicos complexos ou abordagens univariadas (YANRUI et al., 2022; BIDGOLI, 2018). Embora este estudo não utilize aprendizado de máquina

ou modelos híbridos, é importante reconhecer sua presença na literatura como alternativas para cenários com alta complexidade e grandes volumes de dados (FERNANDES et al., 2025). No entanto, essas abordagens demandam maior custo computacional e bases extensas, justificando a escolha por modelos estatísticos neste trabalho.

3

Método de pesquisa

Este trabalho de pesquisa caracteriza-se como um estudo de caso aplicado. Para Yin (2009), o estudo de caso é definido como pesquisa de caráter empírico que investiga fenômenos contemporâneos inseridos em ambientes da vida real. Assim, tem como objetivo investigar, detalhar e explicar o evento, fornecendo uma compreensão ampla e profunda. Para tanto, são analisados os dados concretos de um campo de produção de petróleo, Roncador, e as projeções e os resultados são obtidos e discutidos com base nos instrumentos encontrados na literatura. Sobretudo, o estudo adota uma abordagem quantitativa cujos desdobramentos mensuráveis são fortalecidos por uma estrutura matemática (Gil, 2008).

3.1

Etapas

Segundo Gil (2008), a construção de um estudo de caso envolve o cumprimento de sete etapas, sendo elas:

1. Formulação do problema

A formulação do problema constitui a etapa inicial de uma pesquisa científica e, geralmente, resulta de um processo aprofundado de reflexão e análise crítica, fundamentado em fontes bibliográficas pertinentes ao objeto de estudo.

2. Definição da unidade de análise

A unidade de análise refere-se à delimitação do objeto empírico da investigação, representado por um indivíduo, grupo ou entidade inserido em um contexto específico e relevante para os objetivos da pesquisa.

A unidade-caso deste estudo é o Campo de Roncador, localizado na bacia de Campos, operado pela Petrobras, considerado o quinto maior ativo de produção de petróleo no Brasil, em operação desde 1999.

3. Determinação do número de casos

O estudo de caso pode ser desenvolvido com base em um único caso ou em múltiplos casos. A utilização de múltiplos casos é mais comum, pois permite maior robustez analítica, embora exija maior rigor metodológico e mais tempo para a coleta e análise dos dados. Por sua vez, o estudo de caso único é recomendado quando há restrições no acesso à informação ou quando o fenômeno investigado apresenta características singulares que justificam sua análise isolada.

4. Elaboração do protocolo de pesquisa

O protocolo de pesquisa consiste em um documento estruturado que reúne os instrumentos de coleta de dados e define os procedimentos a serem seguidos durante a investigação. Trata-se de um recurso essencial para assegurar a confiabilidade do estudo de caso. Conforme Yin (2001, p. 89), o protocolo deve conter: (i) uma visão geral do projeto, com a explicitação dos objetivos e do contexto da pesquisa; (ii) os procedimentos de campo, incluindo o acesso às organizações e informantes, bem como os materiais e orientações gerais; (iii) a definição das questões de investigação, acompanhadas das possíveis fontes de evidência; e (iv) um guia para a elaboração do relatório final.

5. Coleta de dados

A coleta de dados deve ser realizada por meio de múltiplas fontes e técnicas, com o objetivo de promover a triangulação das informações e, assim, aumentar a validade e a confiabilidade dos resultados obtidos.

6. Análise dos dados

A análise dos dados, em estudos de caso, é predominantemente qualitativa. Por esse motivo, o pesquisador deve adotar uma postura crítica e cautelosa na interpretação dos achados, evitando conclusões precipitadas ou generalizações indevidas.

7. Redação do relatório

A redação do relatório final deve seguir uma estrutura lógica e coerente, contemplando a apresentação do problema, a descrição da metodologia adotada, a análise dos dados, a discussão dos resultados e as conclusões da pesquisa. O relatório de pesquisa é a presente dissertação de mestrado.

3.1.1

Entrevistas

Foram realizadas entrevistas com especialistas da área de Exploração e Produção (E&P), por meio de questionário técnico estruturado. O objetivo foi coletar percepções práticas sobre o campo Roncador e validar os achados quantitativos do estudo.

As entrevistas foram conduzidas de forma online com profissionais atuantes em diferentes funções estratégicas da Petrobras e da academia, conforme descritos na Tabela 2.

Tabela 2: Relação de Profissionais e Pesquisadores Entrevistados.

Cargo/Função	Empresa	Área de atuação
Especialista sênior em reservatórios	Petrobras	Engenharia de Reservatórios de Poço de Petróleo
Gerente da área de produção do Campo de Roncador	Petrobras	Produção Offshore
Especialista Master em Reservatórios	Petrobras	Engenharia de Reservatórios de Poço de Petróleo
Especialista sênior em reservatórios, elevação e escoamento	Petrobras	Engenharia de Reservatórios de Poço de Petróleo
Diretora associada do centro de estudos em energia e petróleo (CEPETRO)	UNICAMP	Pesquisa Acadêmica em E&P

Fonte: Autora (2025).

As entrevistas foram utilizadas como fonte qualitativa complementar, com o objetivo de validar os achados quantitativos e reforçar a triangulação metodológica, conforme Gil (2008). As percepções dos especialistas foram

integradas ao corpo do texto, conforme categorização temática das respostas.

Os dados obtidos nas entrevistas foram analisados por meio de leitura interpretativa, com categorização temática das respostas. As falas dos especialistas foram incorporadas ao corpo do texto como forma de complementar e validar os achados quantitativos, promovendo a convergência entre diferentes fontes de evidência. Essa abordagem está em conformidade com os princípios de reforçar a consistência e a profundidade da análise por meio da combinação de dados qualitativos e quantitativos conforme descritos por Gil (2008).

3.1.2

Análise de dados

Os dados utilizados para treinamento e validação dos modelos de previsão de séries temporal foram extraídos dos Painéis Dinâmicos de Produção de Petróleo e Gás Natural da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) (2025), organizados em planilha no formato de arquivo .xlsx. Esses painéis estão acessíveis publicamente no site oficial da ANP, por meio do seguinte link:

 <https://www.gov.br/anp/pt-br/centrais-de-conteudo/paineis-dinamicos-da-anp/paineis-dinamicos-sobre-exploracao-e-producao-de-petroleo-e-gas>

Os dados foram organizados em planilhas no formato .xlsx e posteriormente importados e tratados no software R, versão 4.4.3, utilizando o pacote readxl. As variáveis de interesse, totalizando 360 observações mensais, foram convertidas em séries temporais com frequência mensal, abrangendo o período de janeiro de 2015 a fevereiro de 2025.

A variável dependente considerada foi a produção de petróleo em metros cúbicos (m^3), enquanto a produção de água foi utilizada como variável explicativa exógena nos modelos de regressão dinâmica. As séries foram divididas sequencialmente em aproximadamente 80% para treinamento (janeiro/2019 a fevereiro/2024) e 20% para validação (março/2024 a fevereiro/2025).

A análise exploratória foi realizada por meio de gráficos de linha, *seasonal and trend decomposition using Loess* (STL) e função de autocorrelação (FAC),

utilizando os pacotes de *software* *ggplot2*, *zoo*, e *forecast* no R versão 4.4.3. A decomposição STL permitiu identificar componentes de tendência, sazonalidade e ruído.

As informações obtidas nas entrevistas foram trianguladas entre si para garantir confiabilidade e também cruzadas com a pesquisa da literatura, dando embasamento para validação do método de previsão escolhido e interpretação dos resultados.

3.1.3

Modelagem dos dados coletados

O *software* R foi a linguagem escolhida para o desenvolvimento da aplicação prática em cada etapa deste projeto. As bibliotecas utilizadas, conforme descritas na Tabela 3, foram selecionadas com base em sua relevância para a análise de séries temporais, conforme indicado por Hyndman (2022) em seus métodos implementados na linguagem R.

Tabela 3: Descrição e Finalidade das Bibliotecas R Utilizadas na Análise.

Biblioteca R	Finalidade
Ecdat	Fornecer conjuntos de dados econômicos e financeiros amplamente utilizados para ensino e pesquisa em econometria.
Dplyr	Ferramenta poderosa para manipulação de dados. Permite filtrar, selecionar, agrupar e transformar dados de forma eficiente e legível.
Forecast	Usada para modelagem e previsão de séries temporais. Inclui funções como <code>auto.arima()</code> , <code>ets()</code> , e <code>forecast()</code> .
Zoo	Suporte para séries temporais irregulares e regulares. Muito útil para manipulação de dados indexados no tempo.
ggplot2	Sistema de visualização de dados baseado na gramática dos gráficos. Permite criar gráficos sofisticados e personalizados.
MASS	Conjunto de funções e dados do livro <i>Modern Applied Statistics with S</i> . Inclui métodos estatísticos clássicos e modernos.
Dynlm	Permite estimar modelos de regressão linear dinâmica, especialmente úteis em econometria com séries temporais.
Readxl	Leitura de arquivos Excel (.xls e .xlsx) diretamente no R, sem necessidade de Java ou Excel instalado.

Tidyr	Complementa o dplyr para transformar dados entre formatos 'long' e 'wide', além de limpar e organizar dados.
GGally	Extensão do ggplot2 que facilita a criação de gráficos de pares, correlações e visualizações multivariadas.
Ggcorrplot	Especializada em visualização de matrizes de correlação com ggplot2, incluindo mapas de calor e anotações.
Metrics	Fornecer métricas de avaliação de modelos, como MSE, RMSE, MAE, R^2 , entre outras.
gridExtra	Permite combinar múltiplos gráficos ggplot2 em uma única visualização organizada em grades.
Knitr	Ferramenta para geração de relatórios dinâmicos em R Markdown, combinando código, texto e resultados.
Grid	Sistema gráfico de baixo nível usado por ggplot2 e outras bibliotecas para controle detalhado de layouts gráficos.
Scales	Fornecer funções para formatar e escalar eixos em gráficos ggplot2, como transformar valores em porcentagens, log, etc.
Tseries	Conjunto de ferramentas para análise de séries temporais, incluindo testes estatísticos como Dickey-Fuller e KPSS.

Fonte: R CORE TEAM. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2025.

O *software R* foi a linguagem de programação estatística adotada para o desenvolvimento da aplicação prática em todas as etapas deste projeto, devido à sua ampla aceitação na comunidade científica, à sua natureza de código aberto e à sua robustez para análise de dados e modelagem estatística. Desenvolvido inicialmente por Ross Ihaka e Robert Gentleman na Universidade de Auckland, o R evoluiu para se tornar um dos principais ambientes computacionais para análise estatística, sendo mantido atualmente pela *R Foundation for Statistical Computing* (R CORE TEAM, 2024).

A linguagem *R* oferece uma vasta gama de pacotes especializados, que permitem desde tarefas básicas de manipulação de dados até modelagens estatísticas avançadas, aprendizado de máquina e visualização gráfica de alta qualidade. Sua arquitetura extensível e a forte comunidade de desenvolvedores contribuem para a constante atualização e expansão de suas funcionalidades.

No contexto específico da análise de séries temporais, o R destaca-se por oferecer bibliotecas amplamente reconhecidas, como “Forecast” e “tseries”, que

implementam métodos clássicos e modernos de previsão, como ARIMA, *ETS*, modelos sazonais e decomposição de séries. A obra de Hyndman e Athanasopoulos (2021) é uma das principais referências nesse campo, apresentando uma abordagem prática e teórica para previsão estatística com exemplos implementados diretamente em R.

As bibliotecas utilizadas neste trabalho, conforme descritas na Tabela 3, foram selecionadas com base em sua relevância metodológica e aderência às boas práticas recomendadas na literatura especializada. A integração entre os pacotes e a linguagem R permitiu a construção de um fluxo analítico reprodutível, transparente e cientificamente fundamentado.

Para atingir esse objetivo, foram empregados conceitos e técnicas de análise de séries temporais, incluindo modelos de suavização exponencial e regressão dinâmica, que são amplamente aplicados na literatura e na prática de previsão de produção (YANRUI et al., 2022). Os modelos de previsão aplicados foram os modelos de suavização exponencial (SES, Holt, Holt-Winters e *ETS*) e a Regressão Dinâmica (ARIMAX), cada um com suas características e capacidades de ajuste aos dados. Estes métodos são utilizados para capturar padrões mais elaborados em séries temporais, incluindo tendências e sazonalidades (JAIN, 2012). Ainda que não seja esperado observar sazonalidade na produção de petróleo, embora possa haver na sua demanda foram mantidos o modelos que tratam também de sazonalidade junto com os demais, deixando que sejam descartados naturalmente, conforme os resultados, se for o caso.

3.1.4

Normalização das séries temporais

Para possibilitar uma comparação direta entre as séries temporais de produção de petróleo e água, foi aplicada a técnica de normalização min-max, que transforma os valores originais de cada série para uma escala comum entre 0 e 1. Essa transformação é definida pela seguinte equação:

$$x_{\{norm\}} = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (14)$$

A normalização min-max é especialmente útil quando se comparam variáveis com escalas muito diferentes, como a produção de petróleo e de água. Essa técnica transforma os dados para um intervalo comum (geralmente entre 0 e 1), permitindo representá-los no mesmo gráfico e facilitando a análise visual de padrões. Segundo Han, Kamber & Pei (2011), trata-se de uma transformação linear que mantém as proporções originais dos dados, sendo amplamente utilizada em mineração de dados e modelagem estatística.

3.1.5

Modelos de previsão adotados

Foram ajustados cinco modelos de previsão no R: SES, Holt, Holt-Winters, *ETS* e ARIMAX. Os modelos de suavização exponencial foram implementados com o pacote “Forecast”, enquanto o modelo ARIMAX foi ajustado por meio da função `auto.arima()`, utilizando a variável exógena “água” por meio do argumento `xreg`. As previsões foram geradas para o período de validação (12 meses), e os resultados foram avaliados com base nas métricas MAPE e MSE.

A configuração do modelo ARIMAX foi definida automaticamente pela função `auto.arima()` do pacote “Forecast”, com base nos critérios AIC e BIC. Utilizou-se a produção de água como variável exógena (`xreg`), resultando em um modelo ARIMA (0,1,0), sem termos AR ou MA significativos.

Com o objetivo de capturar a influência de variáveis externas sobre a produção de petróleo, foi ajustado automaticamente, por meio da função `auto.arima()` do pacote “Forecast”, um modelo ARIMAX (AutoRegressive Integrated Moving Average with Exogenous Variables), com estrutura ARIMA(0,1,0). A configuração foi selecionada com base nos critérios AIC e BIC, indicando a necessidade de uma diferenciação de primeira ordem, sem termos autoregressivos ou de média móvel significativos. A produção de água foi utilizada como variável exógena.

3.1.6

Aplicação e avaliação de modelos SES, Holt, Holt-Winters, *ETS* e ARIMAX

Após o ajuste dos modelos de previsão descritos no tópico anterior, foram geradas projeções para o período de validação, compreendido entre março de 2024 e fevereiro de 2025. As previsões foram realizadas com base nos modelos SES, Holt, Holt-Winters, *ETS* e ARIMAX, previamente calibrados com os dados de treinamento.

As previsões geradas por cada modelo foram posteriormente avaliadas com base em métricas de desempenho, como o MAPE e o MSE, cujos resultados e análises comparativas são apresentados no Capítulo 4.

4

Resultados e discussões

Neste capítulo, são apresentados e discutidos os resultados dos passos descritos na metodologia. Os resultados obtidos ao aplicar, no problema em estudo, os métodos de previsão detalhados no Referencial Teórico são, portanto, apresentados e avaliados.

Além da análise quantitativa, este capítulo integra percepções qualitativas obtidas em entrevistas com especialistas, visando validar e enriquecer a interpretação dos resultados. Essa triangulação permite compreender não apenas os padrões estatísticos, mas também os fatores operacionais que influenciam a produção.

4.1

Análise da série histórica

No Gráfico 3, pode-se observar o pico de produção de petróleo entre os anos 2018 e 2019, próximo a 60.000 m³/d. Após 2019, observa-se uma tendência de queda contínua. Esse comportamento é típico de campos maduros de petróleo, como o de Roncador, conforme apontado pelo especialista sênior em reservatórios, elevação e escoamento, que destaca a chegada de água como principal gatilho do declínio produtivo indicando possível esgotamento natural dos poços ou redução de investimentos, ou seja, campos maduros frequentemente experimentam declínio na produção devido ao esgotamento do reservatório e desafios operacionais.

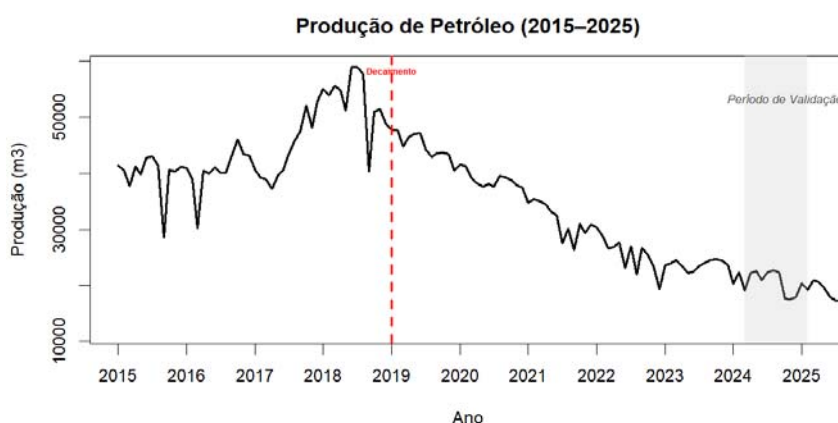


Gráfico 3: Série Temporal da Produção de Petróleo (m³/d).
Fonte: Autora (2025).

Esse comportamento foi confirmado pelo especialista sênior em reservatórios, que apontou a chegada de água como principal gatilho do declínio produtivo, indicando possível esgotamento natural dos poços ou redução de investimentos. Essa percepção reforça a importância de considerar variáveis operacionais no processo de previsão.

O Gráfico 4 apresenta a decomposição da série temporal, o que facilita sua análise ao permitir a observação isolada de cada componente. Foram identificados três componentes principais: tendência, sazonalidade e resíduo. A tendência reflete variações de longo prazo, a sazonalidade evidencia padrões cíclicos anuais e o resíduo representa flutuações aleatórias não explicadas pelos demais componentes.

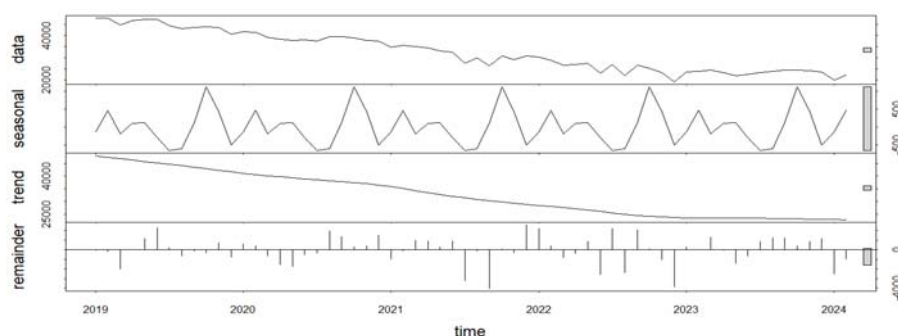


Gráfico 4: Decomposição da Série Temporal (m³/d).
Fonte: Autora (2025).

Para análise dos resultados, foram aplicados modelos de suavização exponencial (SES, Holt, Holt-Winters e *ETS*) e modelos ARIMA, incluindo sua

extensão com variável exógena (ARIMAX), cada um com suas características e capacidades de ajuste aos dados.

No Gráfico 5 são apresentadas a evolução histórica e as projeções de produção de petróleo, expressa em metros cúbicos (m^3), no período compreendido entre os anos de 2015 e 2025. A série histórica, representada pela linha preta contínua, evidencia um pico de produção em 2016, seguido por um declínio acentuado a partir de 2017, que se mantém ao longo dos anos subsequentes. A partir de 2023, são exibidas as projeções geradas pelos modelos SES, Holt, Holt-Winters, ETS, ARIMA e ARIMAX, representadas por linhas coloridas e tracejadas para facilitar a comparação visual.

Podemos notar no Gráfico 5 que os modelos de suavização exponencial (SES, Holt e Holt-Winters) tendem a seguir padrões lineares ou suavizados, enquanto ETS incorpora componentes de tendência e sazonalidade, oferecendo maior flexibilidade. O ARIMA, por sua vez, captura dependências temporais sem variáveis externas, enquanto o ARIMAX se destaca por considerar a influência da variável exógena “água”, resultando em ajustes mais aderentes à dinâmica operacional. Essa diversidade metodológica é relevante para decisões estratégicas, pois previsões mais precisas contribuem para o planejamento de descomissionamento, otimização logística e mitigação de riscos financeiros, reforçando a importância da escolha adequada do modelo conforme o contexto e os objetivos da análise.

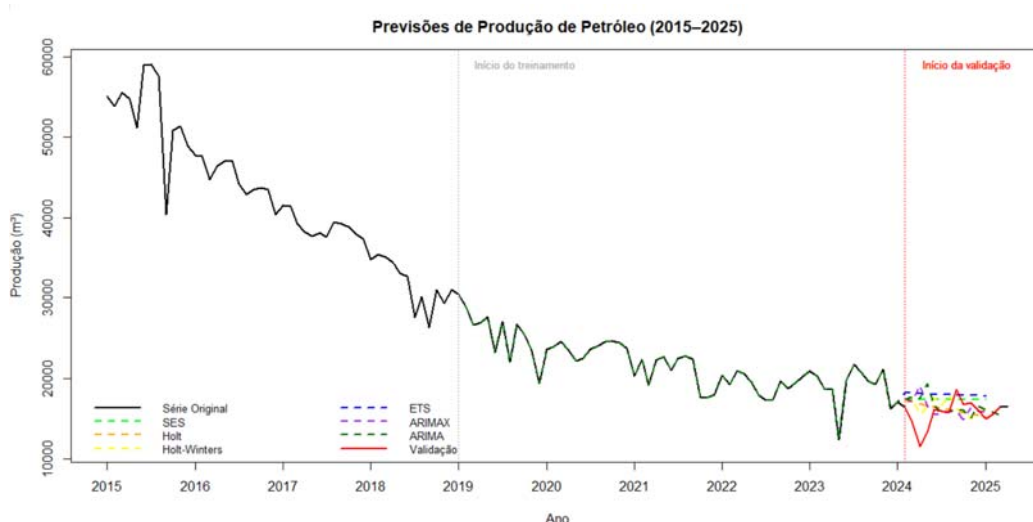


Gráfico 5: Previsão de Produção de Petróleo (2015-2025).
Fonte: Autora (2025).

Após a normalização min-max, de acordo com o Gráfico 6, observaram-se padrões distintos entre as séries de produção de petróleo e água. Enquanto o petróleo apresentou tendência de queda, a produção de água mostrou alta variabilidade. A análise destaca a importância da normalização para comparar variáveis com escalas diferentes e reforça o papel da variável “água” como fator exógeno em modelos como o ARIMAX, especialmente em estratégias voltadas a restaurar a produtividade dos poços ou reduzir a produção de água.

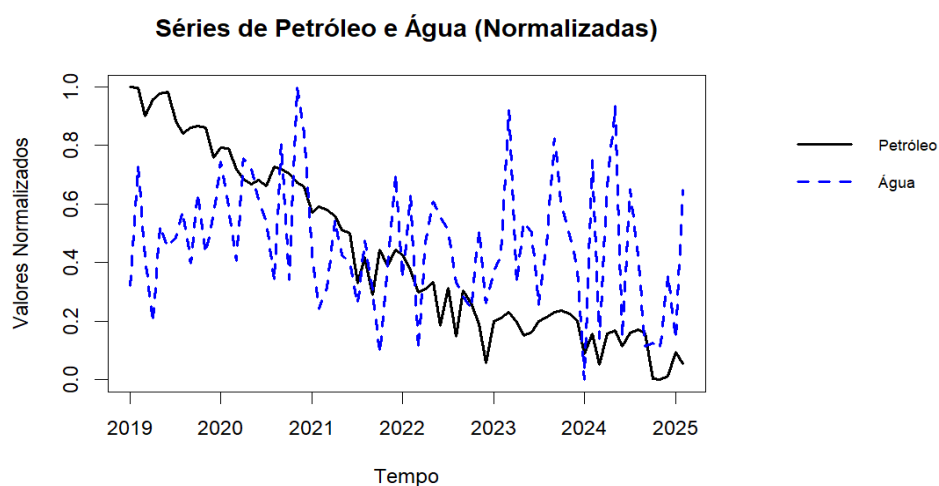


Gráfico 6: Séries de Petróleo e Água (Normalizadas).
Fonte: Autora (2025).

De acordo com o Gráfico 7, a análise da Correlação Cruzada (CCF) entre as séries de produção de petróleo e água revelou padrões relevantes para a compreensão da dinâmica do reservatório. Observa-se correlação positiva significativa em defasagens próximas de zero, indicando que aumentos na produção de água tendem a ocorrer simultaneamente ou com pequena defasagem em relação à redução da produção de petróleo. Esse comportamento é coerente com a literatura, que associa o avanço da frente de água à maturidade do reservatório e ao declínio da produção de óleo.

Em defasagens maiores, a correlação tende a diminuir, sugerindo que os efeitos da produção de água sobre a produção de petróleo são mais pronunciados no curto prazo. Essa evidência reforça a importância da inclusão da variável “água” como fator exógeno em modelos preditivos, como o ARIMAX, pois sua influência direta sobre a série alvo contribui para melhorar a acurácia das previsões. Além disso, a correlação positiva em lags iniciais indica que estratégias

operacionais, como controle da injeção de água, podem impactar rapidamente a produção de petróleo, justificando sua consideração em análises de planejamento e otimização.

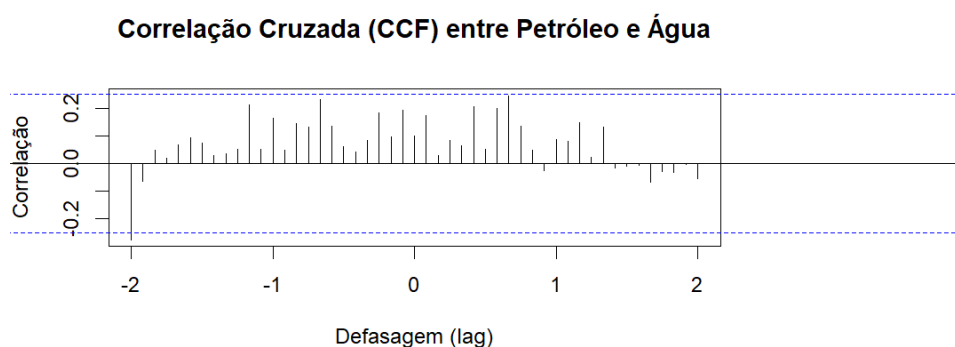


Gráfico 7: Correlação Cruzada (CCF) entre Petróleo e Água.
Fonte: Autora (2025).

No Gráfico 8, é apresentada a Função de Autocorrelação (FAC) aplicada à série de treinamento da produção de petróleo. Observa-se que a maioria dos coeficientes de autocorrelação encontra-se dentro dos limites de confiança de 95%, o que indica ausência de autocorrelação estatisticamente significativa em grande parte dos *lags* analisados. Esse comportamento sugere que os resíduos da série não apresentam estrutura temporal relevante, sendo compatíveis com ruído branco. A presença pontual de autocorrelações significativas em *lags* iniciais, no entanto, justifica a utilização de modelos autorregressivos e de média móvel, como o ARIMA, que são capazes de capturar essas dependências de curto prazo. Assim, a análise da FAC reforça a adequação do modelo ARIMA para a previsão da série, uma vez que este modelo é projetado para lidar com séries estacionárias e com padrões de autocorrelação limitados.

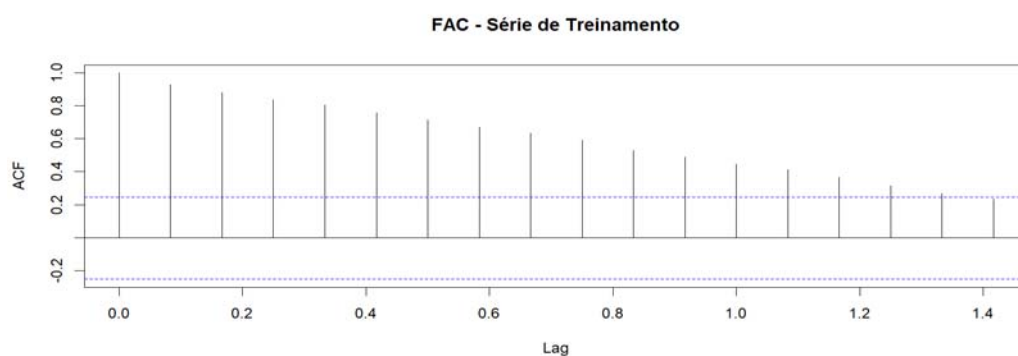


Gráfico 8: Função Autocorrelação da Série Alvo – Período Treinamento.
Fonte: Autora (2025).

Conforme Gráfico 9, a produção de petróleo apresenta uma tendência geral de queda, com oscilações ao longo do tempo. Esse comportamento destaca a importância de modelos preditivos capazes de capturar tanto a tendência negativa quanto variações sazonais. Modelos como ARIMA, ARIMAX, ETS e Hol-Winters são apropriados para esse tipo de série, pois lidam bem com componentes de tendência e suavização. A análise do período de treinamento é essencial para garantir previsões confiáveis no futuro.



Gráfico 9: Produção Período de Treinamento.
Fonte: Autora (2025).

O Gráfico 10 apresenta a Função de Autocorrelação Parcial (PACF) aplicada à série temporal de produção de petróleo equivalente (boed), no período de janeiro de 2019 a fevereiro de 2024, utilizado como conjunto de treinamento para os modelos preditivos.

A PACF mede a correlação direta entre o valor atual da série (X_t) e seus valores defasados (X_{t-k}), eliminando os efeitos das defasagens intermediárias (X_{t-1} , X_{t-2} , ..., $X_{t-(k-1)}$). Essa abordagem permite identificar com precisão os lags que influenciam diretamente o comportamento da série, sendo essencial para definir a ordem autorregressiva (AR) em modelos como o ARIMA.

A análise revelou que os lags 1 e 2 apresentam correlações estatisticamente significativas, indicando que os dois meses anteriores têm influência direta sobre a produção atual. Os lags subsequentes não demonstram correlação relevante, o que justifica a adoção de um modelo AR(2). Essa escolha contribui para otimizar o processo de previsão, reduzindo a complexidade do modelo e aumentando sua eficiência e acurácia.

Do ponto de vista prático, a identificação dos lags mais relevantes orienta a tomada de decisão operacional. Gestores podem priorizar o monitoramento de variáveis críticas nos dois meses anteriores ao planejamento de ações como perfuração de novos poços, ajustes logísticos ou estratégias de descomissionamento. Essa abordagem fortalece a capacidade de resposta às variações produtivas e melhora a alocação de recursos em um setor caracterizado por alta volatilidade, como o de petróleo e gás.

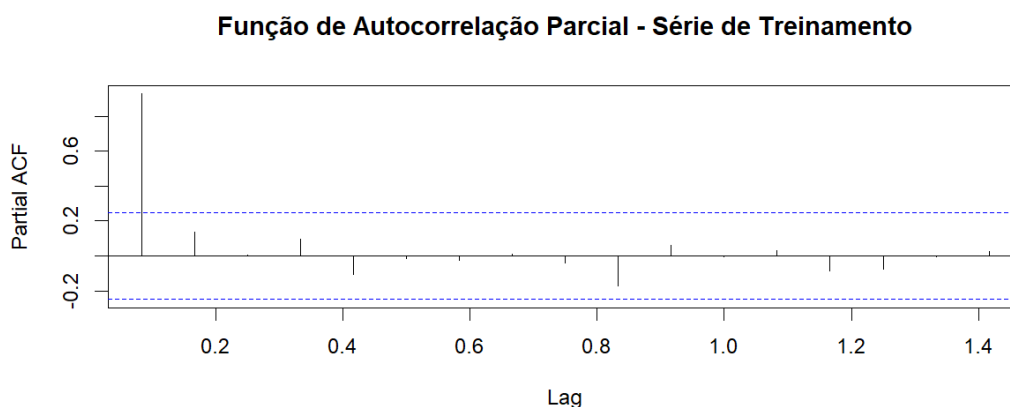


Gráfico 10: Função Autocorrelação da Série Alvo – Período Treinamento.

4.2

Avaliação dos Modelos

O Gráfico 11 apresenta a comparação entre os valores reais de produção de petróleo e as previsões geradas pelos modelos estatísticos SES, Holt, Holt-Winters, *ETS*, *ARIMA* e *ARIMAX* ao longo do período de validação compreendido entre março de 2024 e fevereiro de 2025.

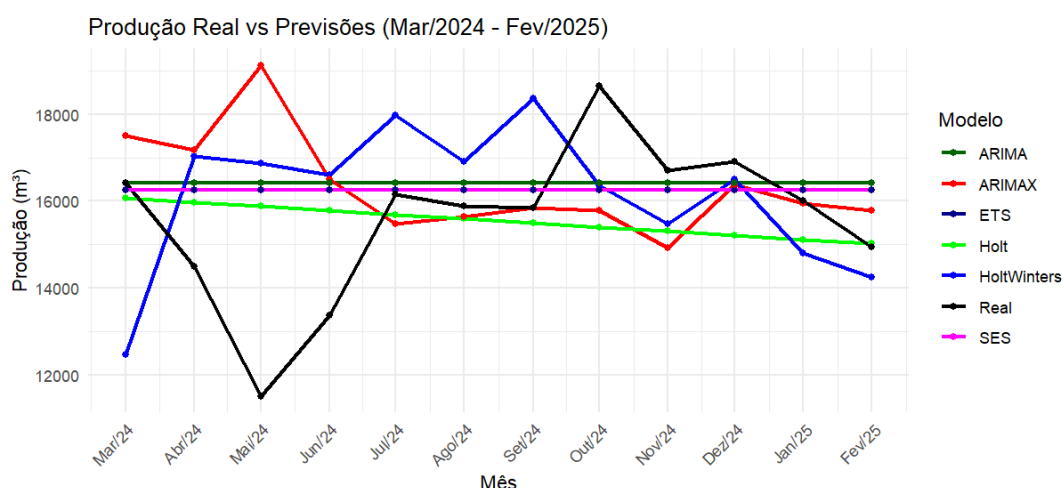


Gráfico 11: Produção Real vs Previsões (Mar/2024 – Fev/2025).

Fonte: Autora (2025).

O modelo SES (Suavização Exponencial Simples) apresentou uma linha praticamente constante ao longo do período de validação, refletindo sua limitação em capturar tendências ou padrões sazonais. Esse comportamento é esperado, dado que o SES não incorpora componentes de tendência nem de sazonalidade. Em contraste, os modelos Holt e ETS demonstraram maior sensibilidade à tendência da série, embora com variações menos acentuadas que a série real. O modelo Holt-Winters, que incorpora componentes sazonais, apresentou oscilações moderadas, mas não conseguiu acompanhar os picos observados nos meses centrais do período.

Observa-se ainda que o ARIMAX apresentou maior variabilidade, aproximando-se parcialmente da dinâmica real, enquanto o ARIMA manteve uma trajetória quase estável, evidenciando sua limitação frente às oscilações abruptas da série.

O modelo ARIMAX ajustado estimou coeficientes estatisticamente significativos tanto para a tendência ($\text{drift} = -473,63$; $p = 0,0053$) quanto para a variável exógena “água” (coeficiente = $0,6817$; $p < 0,001$), ambos ao nível de significância de 5%. Esses resultados indicam que tanto a tendência temporal quanto a produção de água exercem influência relevante sobre a produção de petróleo, justificando sua inclusão no modelo.

A qualidade do ajuste foi avaliada com base na análise dos resíduos, resultando em um Erro Quadrático Médio (MSE) de 1.728.584 e um Erro Percentual Absoluto Médio (MAPE) de 3,37% no conjunto de treinamento. O modelo também apresentou um coeficiente de determinação aproximado (R^2) de 0,9756, sugerindo que cerca de 97,56% da variabilidade da série foi explicada pelos componentes do modelo.

A diretora associada do CEPETRO destacou que a previsão em campos maduros exige considerar interações entre poços e variáveis operacionais, como a produção de água, corroborando a superioridade do ARIMAX observada nos resultados quantitativos.

Tabela 4: Parâmetros do Modelo de Previsão Multivariado de Regressão Dinâmica.

Descrição	Valor	Erro.Padrão	Significância
Drift	-473.6317	169.8136	**
Variável Exógena (xreg)	0.6817	0.0705	***
Erro padrão residual	1314.756		
R^2 aproximado	0.9756		
MSE	1728584		
MAPE	3.3673%		
Teste de Ljung-Box (p-valor)	0.5153		

Fonte: Autora (2025).

4.3

Validação dos Modelos

A avaliação dos modelos preditivos foi baseada nos indicadores MSE e MAPE, permitindo identificar o modelo mais preciso no período de validação (Mar/2024 – Fev/2025). O especialista master em reservatórios reforçou que a reinjeção de água impacta diretamente a injetividade dos poços, justificando a inclusão dessa variável no ARIMAX. Essa percepção prática está alinhada com a redução do erro observada nesse modelo.

Os resultados, apresentados no Gráfico 12 e na Tabela 5, evidenciam diferenças significativas entre as abordagens.

O modelo ARIMAX apresentou o melhor desempenho, com menor MAPE ($\approx 9,5\%$) e menor MSE ($\approx 4,49$ milhões), confirmando sua superioridade frente aos

demais modelos. Essa performance reforça a relevância da inclusão da variável exógena “água”, que contribui para capturar a dinâmica operacional da produção. Em comparação, os modelos ETS e Holt também apresentaram erros relativamente baixos (MAPE $\approx 9,3\%$), enquanto Holt-Winters obteve o pior resultado, com MAPE acima de 12% e MSE próximo a 8,87 milhões, evidenciando limitações na previsão para séries com comportamento irregular.

Segundo o especialista master em reservatórios, a reinjeção de água impacta diretamente a injetividade dos poços, justificando sua utilização como variável explicativa no ARIMAX. Essa análise demonstra que a escolha do modelo deve considerar não apenas a estrutura estatística, mas também fatores operacionais que influenciam a série histórica.

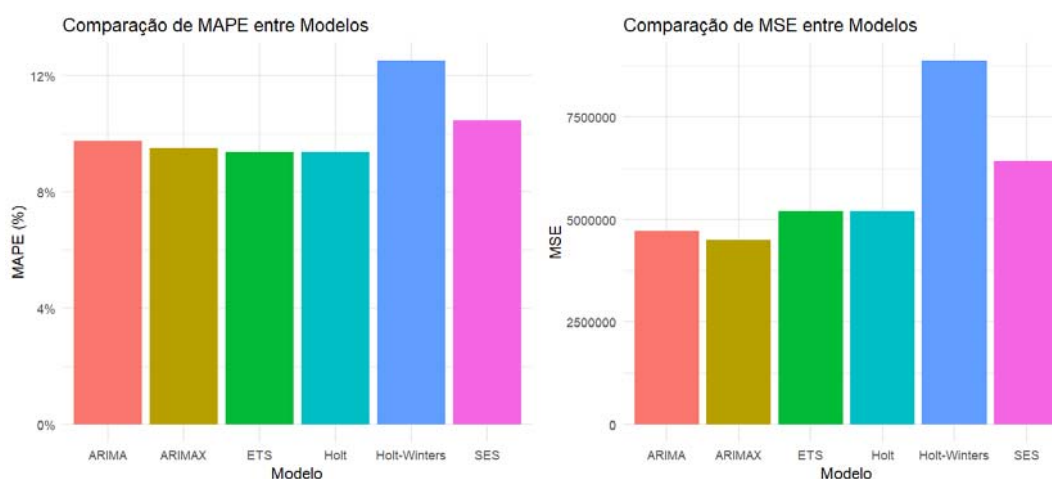


Gráfico 12: Comparação de MAPE e MSE entre os modelos – Período Validação.
Fonte: Autora (2025).

Tabela 5: Métricas de Acurácia.

Modelo	MAPE_Treino	MSE_Treino	MAPE_Valid	MSE_Valid
ARIMA	4.85 %	3025684	12.04 %	8254652
ARIMAX	4.72 %	2920542	9.51 %	4492787
ETS	4.97 %	3252408	9.37 %	5196150
Holt	4.97 %	3252408	9.37 %	5196335
Holt-Winters	4.76 %	2932037	12.52 %	8874093
SES	5.08 %	3838133	10.47 %	6432794

Fonte: Autora (2025).

Acredita-se que pode ser possível gerar resultados ainda mais precisos se o trabalho for realizado em parceria com a empresa operadora, utilizando dados e informações internos, como, por exemplo, registros de paradas de produção, uma vez que esses dados podem fornecer séries temporais explicativas adicionais, com maior correlação com a série temporal alvo da produção.

4.4

Desempenho do Modelo ARIMAX

Conforme evidenciado no Gráfico 13, a linha correspondente ao ARIMAX (em azul escuro) apresenta maior aderência à série real (em laranja), com menor dispersão ao longo do período analisado. Essa proximidade visual é corroborada pelas métricas quantitativas, nas quais o ARIMAX obteve o menor erro quadrático médio (MSE), indicando maior precisão em termos absolutos.

O especialista sênior em reservatórios afirmou que, embora não utilize diretamente modelos estatísticos como o ARIMAX, reconhece a relevância da produção de água como variável preditiva, prática já consolidada em métodos tradicionais de previsão de reservatórios.

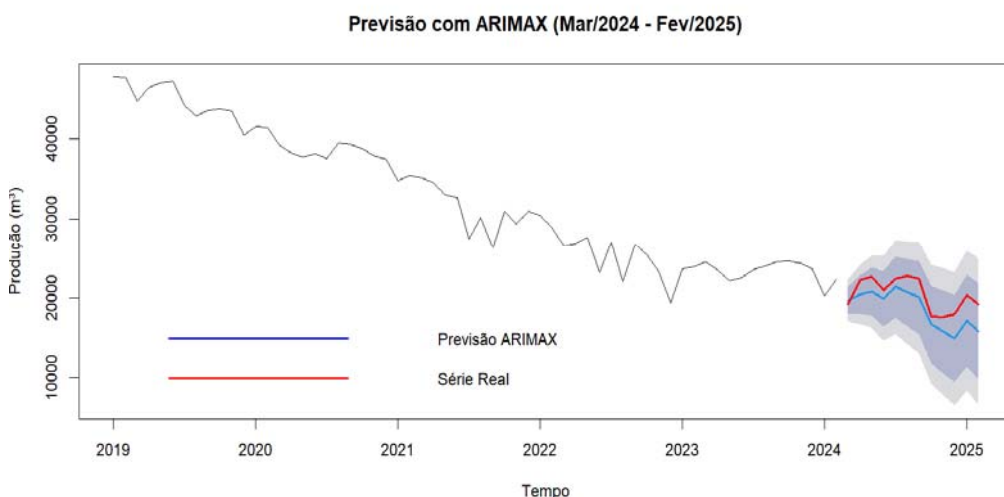


Gráfico 13: Previsão de Modelo ARIMAX.
Fonte: Autora (2025).

Embora não tenha utilizado diretamente o modelo ARIMAX, a diretora associada do Centro de Estudos de Energia e Petróleo (CEPETRO), acredita que a

inclusão da produção de água como variável exógena tem potencial para melhorar a acurácia dos modelos preditivos, especialmente em campos maduros.

Dessa forma, o modelo ARIMAX destaca-se como o mais robusto entre os avaliados para este caso, sendo o mais adequado para aplicações preditivas em séries temporais de produção industrial com influência de variáveis externas.

A metodologia de Fernandes et al. (2025) foi aplicada com sucesso em campos do pré-sal brasileiro, demonstrando que modelos baseados em séries temporais e aprendizado supervisionado são capazes de antecipar comportamentos indesejados, otimizar parâmetros de produção e apoiar decisões operacionais. Os resultados obtidos com regressão Ridge e redes neurais mostraram desempenho superior em cenários com alta variabilidade e complexidade.

A análise dos resultados obtidos ao longo do período de validação evidencia diferenças relevantes entre os modelos avaliados. O modelo ARIMAX, ao incorporar variáveis exógenas, apresentou o menor Erro Quadrático Médio (MSE), o que indica maior precisão em termos absolutos. Esse desempenho é particularmente relevante em contextos em que a minimização de grandes desvios é prioritária, como na previsão de produção industrial com influência de fatores externos.

Por outro lado, os modelos Holt e *ETS* obtiveram os menores valores de Erro Percentual Absoluto Médio (MAPE), ambos com 9,37%, sugerindo desempenho mais consistente em termos relativos. Essa característica pode ser vantajosa em aplicações onde a acurácia percentual é mais relevante do que o erro absoluto.

O modelo Holt-Winters, teoricamente mais adequado para séries com forte componente sazonal, apresentou o pior desempenho entre os avaliados, tanto em MSE quanto em MAPE. Isso pode indicar que a série de produção de petróleo analisada não possui uma sazonalidade significativa, como realmente era o esperado no caso do esgotamento de um poço de petróleo, ou ainda que o modelo tenha superajustado os dados, tornando-se excessivamente sensível a flutuações pontuais.

Visualmente, conforme ilustrado nos Gráficos 9 e 10, observa-se que os modelos Holt e *ETS* seguem uma trajetória de queda suave, alinhada ao comportamento recente da série. O modelo SES, por sua vez, mantém uma linha

praticamente constante, refletindo sua limitação em capturar tendências. Já o modelo Holt-Winters projeta uma queda mais acentuada, o que reforça a hipótese de sobreajuste. O modelo ARIMAX, ao incorporar a variável exógena “água”, apresenta uma trajetória mais oscilante, refletindo maior sensibilidade a variações externas.

A variabilidade da produção de água reforça sua importância como variável exógena em modelos preditivos. A entrevista com o especialista sênior em reservatórios, elevação e escoamento destaca que os principais desafios na previsão de produção em campos maduros incluem o tratamento de problemas operacionais e a melhoria da previsibilidade em cenários com injeção de água. Esse resultado é corroborado pela entrevista com o especialista sênior em reservatórios, que destacou a produção de água como fator crítico para o declínio da produção e para a tomada de decisão operacional. A percepção do especialista master em reservatórios sobre a reinjeção de água e sua influência na injetividade dos poços reforça a relevância da variável “água” como componente exógeno no modelo ARIMAX. Estratégias para restaurar a produtividade dos poços ou reduzir a produção de água são fundamentais para melhorar a eficiência da produção de petróleo, sendo bem representadas por modelos como o ARIMAX, que capturam essas influências externas com maior precisão. Apesar de o modelo Holt-Winters ter apresentado boa aderência visual em determinados trechos, os resultados quantitativos confirmam a superioridade do ARIMAX em termos de precisão absoluta. A linha de previsão do ARIMAX aproxima-se mais da série real ao longo do período de validação, com menor dispersão, o que é corroborado pelo menor MSE (4.492.787).

Os modelos ETS e Holt, embora menos precisos em termos absolutos, ainda oferecem desempenho satisfatório quando a prioridade for a acurácia percentual. Já os modelos SES e Holt-Winters mostraram-se menos adequados para a dinâmica específica da série em questão. Ainda que os modelos Holt e ETS tenham apresentado melhor desempenho percentual, a robustez do ARIMAX o posiciona como a abordagem mais confiável para o cenário analisado.

Portanto, a análise combinada dos resultados estatísticos e visuais sugere que o modelo ARIMAX é o mais robusto entre os avaliados, sendo o mais

indicado para aplicações preditivas em séries temporais de produção de petróleo com influência de variáveis externas.

O especialista sênior em reservatórios afirmou que, embora não utilize diretamente modelos estatísticos como o ARIMAX em seu dia a dia, reconhece o valor da produção de água como variável preditiva. Ele ressaltou que métodos tradicionais de previsão de reservatórios já consideram essa variável como essencial para entender o comportamento dos fluidos e prever o acumulado de produção. A inclusão da água como variável exógena, portanto, é coerente com práticas consolidadas na indústria.

O especialista master em reservatórios mencionou o uso do sistema SAR (Sistema de Análise de Reservatórios), desenvolvido internamente pela Petrobras, como ferramenta complementar para previsões de curto prazo. O SAR incorpora métodos estatísticos e tem sido utilizado em áreas como águas profundas, oferecendo agilidade na tomada de decisão. Embora não substitua a simulação numérica, o SAR representa uma alternativa promissora para cenários com alta variabilidade e necessidade de resposta rápida.

O modelo ARIMAX previu 223.432,84 m³ contra 186.004,08 m³ reais, superestimando em 20,12%, mas ainda foi o mais próximo dos valores observados. Foi considerado o mais confiável para decisões operacionais e econômicas no Campo de Roncador. A análise indica sua robustez para previsões entre mar/2024 e fev/2025 com variáveis externas.

A relevância da variável “produção de água” como fator exógeno foi corroborada pelo especialista sênior em reservatórios, que afirmou que a variabilidade da produção de água pode ser usada como indicador preditivo da produtividade do poço. Métodos como a extrapolação da RAO e o declínio da fração de óleo são baseados na análise conjunta da produção de água e óleo.

Segundo o especialista master em reservatórios, a reinjeção de água impacta diretamente a injetividade dos poços, o que reforça sua inclusão como variável exógena no modelo ARIMAX.

Conclusão

O objetivo deste estudo foi identificar padrões e tendências de produção de petróleo do Campo de Roncador, por meio dos modelos estatísticos SES, Holt, Holt-Winters, *ETS*, ARIMA e ARIMAX onde foi selecionado aquele com maior acurácia preditiva, com foco na aplicação prática em campos de petróleo *offshore* maduros. Para alcançar esse objetivo, foram empregados conceitos e técnicas avançadas de análise de séries temporais, incluindo modelos de suavização exponencial e regressão dinâmica.

A análise comparativa demonstrou que, embora o ARIMA tenha apresentado desempenho satisfatório, sua limitação em não considerar variáveis externas resultou em menor aderência à série real. O ARIMAX, por outro lado, destacou-se como o modelo mais robusto, com menor erro absoluto e maior capacidade de capturar a influência da produção de água, variável crítica em campos maduros. Essa evidência reforça a importância de incorporar variáveis exógenas em modelos de previsão, especialmente em cenários onde fatores operacionais impactam diretamente a produção.

A convergência entre os resultados estatísticos e as percepções dos especialistas e profissionais entrevistados reforça a confiabilidade das conclusões e a aplicabilidade prática do modelo ARIMAX, especialmente para previsões realizadas com rapidez para o curto prazo. O modelo ARIMAX foi o mais adequado para previsão, entre técnicas como SES, Holt, Holt-Winters e *ETS*.

Avaliando os modelos pelos indicadores MSE e MAPE, o ARIMAX se destacou como o mais preciso, com MAPE de 4,72% e MSE de 2.920.542 na fase de treinamento, mantendo bom desempenho na validação (MAPE de 9,51% e MSE de 4.492.787). A inclusão da variável exógena “água” contribuiu significativamente para explicar a variabilidade da série, superando modelos puramente endógenos.

O modelo ARIMAX apresentou o melhor desempenho neste estudo, considerando a série histórica do campo de Roncador e a inclusão da variável exógena ‘produção de água’. No entanto, sua replicação para outros campos deve

ser avaliada caso a caso, considerando as características específicas de cada reservatório e a disponibilidade de variáveis explicativas relevantes.

A previsão neste estudo foi realizada utilizando dados específicos da produção de petróleo do campo de Roncador. Embora os resultados obtidos com o modelo ARIMAX tenham sido promissores, sua aplicação em outros campos deve ser cuidadosamente avaliada, considerando as particularidades operacionais e geológicas de cada ativo, bem como a disponibilidade de variáveis exógenas relevantes.

A implementação do projeto EOR em Roncador representa um exemplo concreto de como a tecnologia e a cooperação internacional podem impulsionar a recuperação de campos maduros. Os resultados obtidos até o momento indicam ganhos significativos em produção, eficiência e sustentabilidade, reforçando a relevância de estratégias preditivas e colaborativas para o futuro da indústria de petróleo (EQUINOR, 2022).

No geral, este caso de estudo, reforça sua robustez, contribui para uma compreensão das dinâmicas de produção no campo Roncador e oferece uma base para futuras pesquisas e aplicações práticas na otimização da produção de petróleo.

Para pesquisas futuras, sugere-se avaliar métodos específicos do setor de petróleo e gás, incorporando dados dos reservatórios, como pressão, temperatura, injetividade e características do fluido, que são determinantes para a dinâmica de produção. Essa abordagem possibilitaria a inclusão de variáveis com maior relevância física e operacional, conferindo maior robustez aos modelos e elevando a confiabilidade das previsões. Além disso, recomenda-se abordagens de *machine learning*.

A entrevista realizada com a diretora associada do Centro de Estudos de Energia e Petróleo (CEPETRO) reforçou a complexidade da previsão de produção em campos maduros, destacando a influência de variáveis operacionais, interações entre poços e o papel da produção de água como fator crítico para decisões estratégicas.

As entrevistas realizadas com especialistas e profissionais da área de E&P reforçaram os achados quantitativos do estudo, especialmente no que diz respeito à influência da produção de água na previsão de petróleo. A percepção técnica dos

profissionais validou o uso de variáveis exógenas e destacou a importância da análise econômica e operacional para a tomada de decisão em campos maduros como Roncador.

A entrevista com o especialista master em reservatórios reforça a importância de considerar variáveis operacionais como a injetividade dos poços e a qualidade da água reinjetada na modelagem preditiva. Além disso, destaca a relevância de ferramentas internas como o SAR, que podem complementar os modelos estatísticos tradicionais, especialmente em previsões de curto prazo. Essas contribuições evidenciam que a integração entre conhecimento técnico e ferramentas analíticas é essencial para a sustentabilidade da produção em campos maduros.

A entrevista com o especialista sênior em reservatórios reforça a importância da integração entre modelos estatísticos e conhecimento técnico-operacional. Sua visão sobre a complexidade do Campo de Roncador, a relevância da produção de água como variável preditiva e os desafios da reinjeção contribuem para validar os achados do estudo. Além disso, sua sugestão de considerar múltiplos cenários e curvas de previsão (P10, P50, P90) aponta para caminhos promissores na gestão da incerteza.

O especialista sênior da área de reservatórios, elevação e escoamento da considera que modelos como o ARIMAX podem ser eficazes em situações mais controladas, desde que se evite o uso de dados futuros na previsão. Martins (2024), por exemplo, compara o desempenho de modelos ARIMA, ARIMAX e redes neurais recorrentes *Long Short-Term Memory* (LSTM) na previsão de séries temporais, demonstrando que a inclusão de variáveis exógenas melhora a acurácia dos modelos em contextos com múltiplos fatores explicativos.

As entrevistas realizadas com especialistas da Petrobras e da UNICAMP revelaram convergências importantes, especialmente quanto à relevância da produção de água como variável preditiva e ao uso de técnicas de recuperação secundária. Todos os profissionais destacaram a complexidade operacional do Campo de Roncador e a necessidade de modelos preditivos robustos para apoiar a tomada de decisão. A triangulação entre os resultados estatísticos e as percepções dos especialistas reforça a robustez das conclusões, evidenciando que a inclusão

da variável ‘produção de água’ não apenas melhora a acurácia preditiva, mas também reflete práticas consolidadas na indústria.

Referências Bibliográficas

ABRAHAM-ADEJUMO, T. Water cut prediction and reservoir performance analysis using machine learning techniques. *Journal of Petroleum Science and Engineering*, v. 171, p. 123–134, 2018.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS. *Boletim da Produção de Petróleo e Gás Natural – 2024: Encarte de Consolidação da Produção – 2024a*. Rio de Janeiro: ANP, 2024.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS. *Boletim de Recursos e Reservas de Petróleo e Gás Natural – 2024b*. Rio de Janeiro: ANP, 2024.

AGÊNCIA NACIONAL DO PETRÓLEO, GÁS NATURAL E BIOCOMBUSTÍVEIS. *Boletim de Recursos e Reservas de Petróleo e Gás Natural – 2024b*. Rio de Janeiro: ANP, 2024.

AGUIAR DE ARAÚJO, D. Gerente da área de Gerenciamento e Maximização do Potencial de Roncador – Petrobras. Entrevista concedida a Daniela Mendonça dos Santos e Marcelo Xavier Seeling. Online, 29 jul. 2025, às 15h30. Entrevista pessoal.

ALLAHYARZADEH BIDGOLI, A. Simulation and optimization of primary oil and gas processing plant of FPSO operating in pre-salt oil field. 2018. 206 f. Tese (Doutorado em Engenharia Mecânica) – Universidade de São Paulo, São Paulo, 2018. Disponível em: <https://teses.usp.br/teses/disponiveis/3/3150/tde-13122018-150547/publico/AliAllahyarzadehBidgoliCorr18.pdf>. Acesso em: 1 set. 2025.

AMARAL FILHO, F.M. do; BANDEIRA, S.G.; ALCALÁ, S.G.S.; BARBOSA, T.M.G. de A. Um estudo comparativo de ensembles híbridos para aplicações de previsão de séries temporais. *Revista Brasileira de Computação Aplicada*, v. 13, n. 2, p. 58-72, 2021. DOI: 10.5335/rbca.v13i2.12067.

ASSIS, J.P.C. de. Especialista em reservatórios da Petrobras. Entrevista concedida a Daniela Mendonça dos Santos e Marcelo Xavier Seeling. Rio de Janeiro, 6 ago. 2025. Entrevista pessoal.

BENALCÁZAR ORDÓÑEZ, J.C. Dependência petrolífera e vulnerabilidade econômica: o caso do Equador. Quito: Universidad Central del Ecuador, 2025. Dissertação (Mestrado em Economia).

BOX, G.E.P.; JENKINS, G.M. Time series analysis: forecasting and control. San Francisco: Holden-Day, 1976.

BOX, G.E.P.; JENKINS, G.M. Time Series Analysis: Forecasting and Control. 4th ed. San Francisco: Holden-Day, 2008.

BOX, G.E.P.; JENKINS, G.M.; REINSEL, G.C.; LJUNG, G.M. Time Series Analysis: Forecasting and Control. 5. ed. Hoboken: Wiley, 2015.

CARMO, E.G.D. Ajuste sazonal de séries temporais: uma abordagem comparativa entre as metodologias X-13 ARIMA SEATS. Monografia (Bacharelado em Economia) – Universidade Federal do Rio de Janeiro, 2017.

CASTRO, A.O. de S.; FERREIRA FILHO, V.J.M. Selecting oil wells for hydraulic fracturing: a comparison between genetic-fuzzy and neuro fuzzy systems. *American Journal of Operations Research*, v. 4, p. 202-216, 2014. Disponível em: <<https://www.scirp.org/journal/ajor>>. DOI: 10.4236/ajor.2014.44020. Acesso em: 1 set. 2025.

CASTRO, M.; MENDES JÚNIOR, P.R.; DAVOLIO, A. Time series causal relationships discovery through feature importance and ensemble models. *Scientific Reports*, v. 13, n. 1, p. 11402, 2023.

CHAI, T.; DREW, D. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, v. 7, p. 1247-1250, 2014. Disponível em: <https://gmd.copernicus.org/articles/7/1247/2014/>. Acesso em: 10 jul. 2025.

CHIAN, C. Análise de variância em séries temporais. Dissertação (Mestrado em Estatística) – Universidade de São Paulo, 1993. Disponível em: <https://www.teses.usp.br/teses/disponiveis/45/45133/tde-20210729-002114/pt-br.php>. Acesso em: 1 set. 2025.

CNN BRASIL. Projeção da OPEP sobre a produção brasileira de petróleo. 2024. Disponível em: <https://www.cnnbrasil.com.br/economia/macroeconomia/opep-preve-aumento-de-demanda-por-petroleo-e-maior-producao-no-brasil-em-2025/>. Acesso em: 1 set. 2025.

CORREIA, J.B.; PIVETTA, M.; NASCIMENTO, G.S.; BECKER, K. Comparing ARIMA and LSTM models to predict time series in the oil industry. Anais do IX Symposium on Knowledge Discovery, Mining and Learning (KDMiLe), 2021.

COSTA, F. de A. et al. Reservoir characterization using high frequency seismic data in Roncador, Campos Basin. Apresentado no 9º Congresso Internacional da Sociedade Brasileira de Geofísica, Salvador, 11–14 set. 2005.

DANTAS, S.W.B.; PESSOA, A.R.P. Estudo sobre os métodos de recuperação de campos maduros onshore. Trabalho de Conclusão de Curso – Universidade Federal Rural do Semiárido, 2020.

DANAEI, S.; MALEKI, M.; SCHIOZER, D.J.; DAVOLIO, A. Using petro-elastic proxy model to integrate 4D seismic in ensemble based data assimilation. *Journal of Petroleum Science and Engineering*, v. 194, p. 107457, 2020.

DAVOLIO, A.; MASCHIO, C.; SCHIOZER, D.J. Pressure and saturation estimation from P and S impedances: a theoretical study. *Journal of Geophysics and Engineering*, v. 9, n. 5, p. 447–460, 2012.

DOU, L.; WEN, Z.; WANG, Z. Global Oil and Gas Resources: Potential and Distribution. Singapore: Springer, 2024. Disponível em: <<https://link.springer.com/book/10.1007/978-981-97-4756-6>>.

EMPRESA DE PESQUISA ENERGÉTICA (EPE). Balanço Energético Nacional 2025: Relatório Síntese. Rio de Janeiro: EPE, 2025. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/balanco-energetico-nacional-2025>. Acesso em: 9 jul. 2025.

ENERDATA. World Energy Statistics – Yearbook 2023. Grenoble: Enerdata, 2023. Disponível em: <https://yearbook.enerdata.net/>. Acesso em: 9 jul. 2025.

EQUINOR. Petrobras e Equinor iniciam produção no projeto de recuperação avançada de petróleo no campo de Roncador. *Equinor Brasil*, 2022. Disponível em: <https://www.equinor.com.br/noticias/petrobras-e-equinor-iniciam-producao-no-projeto-de-recuperacao-avancada-de-petroleo-no-campo-de-roncador>. Acesso em: 1 set. 2025.

ERBER, P.; MOSSÉ, A. Armazenamento de energia. *Valor Econômico*, São Paulo, 28 fev. 2023. Disponível em: <https://valor.globo.com/opiniao/coluna/armazenamento-de-energia-1.ghtml>.

FARIA, L.F.V. de. Análise de projeto de investimento e tomada de decisão com utilização da teoria de opções reais na indústria de petróleo: um modelo de opção de abandono. 2015. Dissertação (Mestrado em Engenharia de Produção) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Campos dos Goytacazes.

FERNANDES, M. de A. Especialista em reservatórios da Petrobras. Entrevista concedida a Daniela Mendonça dos Santos e Marcelo Xavier Seeling. Entrevista online. 1 ago. 2025.

FERNANDES, M.A.; FURLANETTI, M.M.; GILDIN, E.; SAMPAIO, M.A. Data-driven models for production forecasting and decision supporting in petroleum reservoirs. *Preprint. arXiv*, 2025. Disponível em: <https://arxiv.org/abs/2508.18289>. Acesso em: 1 set. 2025.

FERREIRA, F.P.; MATIAS, Í.O. Previsão da produção de campos de petróleo utilizando redes neurais. In: 10th Brazilian Congress on Computational Intelligence (CBIC'2011), Fortaleza. Anais... Fortaleza: SBIC, 2011.

FREITAS, G.R.S. de. Estudo teórico aplicado ao petróleo: estabilidade dos alcenos e recuperação terciária. Dissertação (Mestrado) – Universidade Federal do Rio Grande do Norte, Natal, 2013.

FURLAN, L.M. Caracterização sísmica de aspectos tectonoestratigráficos da região do campo de Roncador (Cretáceo Superior) – Bacia de Campos – Brasil. 63 f. Trabalho de Conclusão de Curso (Graduação em Geologia) – Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio Claro, 2017.

GIL, A.C. Métodos e técnicas de pesquisa social. 6. ed. São Paulo: Atlas, 2008.

GOMES, A.D. Diretora associada do Centro de Estudos de Energia e Petróleo da UNICAMP. Entrevista concedida a Daniela Mendonça dos Santos e Marcelo Xavier Seeling. Entrevista online. 29 jul. 2025.

GOODRICH, R.L. Dynamic Regression Models. Manuscrito não publicado, 1989.

HAMACHER, S.; FERREIRA FILHO, J.B. Cadeia de suprimentos e logística no setor de petróleo e gás. In: ARAGÃO, A.; SCAVARDA, L. *Gestão e Produção*, v. 11, n. 3, p. 299-311, 2004.

HAN, J.; KAMBER, M.; PEI, J. Data mining: concepts and techniques. 3. ed. San Francisco: Morgan Kaufmann, 2011.

HOERL, A.E.; KENNARD, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics*, v. 12, n. 1, p. 55-67, 1970.

HOFMANN, E.; RUTSCHMANN, E. Big data analytics and demand forecasting in supply chains: a conceptual analysis. *The International Journal of Logistics Management*, v. 29, n. 2, p. 739-766, 2018. DOI: 10.1108/IJLM-04-2017-0088.

HOLLYMAN, R. et al. Forecast combination based forecast reconciliation: insights and extensions. *arXiv preprint*, 2021. Disponível em: <https://arxiv.org/abs/2106.05653>. Acesso em: 1 set. 2025.

HOLT, C.C. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, v. 20, n. 1, p. 5-10, 2004.

HYNDMAN, R.J.; ATHANASOPOULOS, G. Forecasting: Principles and Practice. 2. ed. OTexts, 2018. Disponível em: <https://otexts.com/fpp2/>. Acesso em: 1 set. 2025.

HYNDMAN, R.J.; ATHANASOPOULOS, G. Forecasting: principles and practice. 3. ed. Melbourne: OTexts, 2021. Disponível em: <https://otexts.com/fpp3/>. Acesso em: 9 jul. 2025.

HYNDMAN, R.J.; KOEHLE, A.B.; SNYDER, R.D. Forecasting with exponential smoothing: the state space approach. Berlin: Springer, 2008.

HYNDMAN, R.J.; LEE, A.; WANG, E.; WICKRAMASURIYA, S.L. hts: Hierarchical and Grouped Time Series. R package version 6.0.3, 2024. Disponível em: <https://cran.r-project.org/web/packages/hts/hts.pdf>.

IEEE. Application of the Double Smoothing and ARIMAX Methods for the Prediction of Polycrystalline Photovoltaic Generation. In: IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society, 17–20 out. 2022. IEEE Xplore. Disponível em: <https://ieeexplore.ieee.org/abstract/document/9968612>. Acesso em: 1 set. 2025.

INSTITUTO BRASILEIRO DE PETRÓLEO E GÁS (IBP). Panorama geral do setor de petróleo e gás: uma agenda para o futuro. Rio de Janeiro: IBP, 2023. Disponível em: <https://www.ibp.org.br/personalizado/uploads/2023/02/panorama-setor-og-ibp-2023.pdf>. Acesso em: 9 jul. 2025.

INTERNATIONAL ENERGY AGENCY (IEA). World Energy Outlook 2023. Paris: IEA, 2023. Disponível em: <https://www.iea.org/reports/world-energy-outlook-2023>. Acesso em: 9 jul. 2025.

INTERNATIONAL ENERGY AGENCY (IEA). Energy access improving, but international financial support still needed to boost progress and address disparities. Paris: IEA, 2025. Disponível em: <https://www.iea.org/news/energy-access-improving-but-international-financial-support-still-needed-to-boost-progress-and-address-disparities>. Acesso em: 1 set. 2025.

INTERNATIONAL ENERGY AGENCY (IEA). World Energy Outlook 2024. Paris: IEA, 2024. Disponível em: <https://www.iea.org/reports/world-energy-outlook-2024>. Acesso em: 1 set. 2025.

KAMEI, C.A.N. Uma abordagem evolucionária para hibridização de preditores. Tese (Doutorado em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2024.

KARMY, M.; MALDONADO, S. Hierarchical time series forecasting using machine learning models. *Applied Soft Computing*, v. 80, p. 1-12, 2019.

KORSTANJE, J. Advanced Forecasting with Python: Concepts and Applications with Time Series and Sequential Data. Springer, 2021. DOI: <https://doi.org/10.1007/978-1-4842-6994-4>.

LIMAVERDE FILHO, J.O. de Aquino; BESSA, R.; PIMENTA, B.G.; FORTALEZA, E.L.F. Agregação de modelos para prever a produção de reservatórios de petróleo a médio prazo. In: Rio Oil & Gas Expo and Conference, 2022, Rio de Janeiro. Anais [...]. Rio de Janeiro: IBP, 2022.

LUKE, J.; PORTER, D.; SANTHANAM, P. Beyond Algorithms. Taylor & Francis, 2022. Disponível em: <https://taylorandfrancis.com/>. Acesso em: 1 set. 2025.

MAKRIDAKIS, S.; WHEELWRIGHT, S.C.; HYNDMAN, R.J. Forecasting: Methods and Applications. 3. ed. New York: Wiley, 1998.

MALUF, T. Análise comparativa de custos e técnicas de perfuração de poços para a produção de água, energia geotérmica, petróleo e gás natural. Trabalho de Graduação – Universidade Estadual Paulista, Guaratinguetá, 2023. Disponível em: <https://repositorio.unesp.br/items/7f425500-c76a-46af-a91d-04ce2a57b10a>. Acesso em: 1 set. 2025.

MELO FILHO, M.G. de. Especialista em reservatórios da Petrobras. Entrevista concedida a Daniela Mendonça dos Santos e Marcelo Xavier Seeling. Rio de Janeiro, 6 ago. 2025. Entrevista pessoal.

MARQUES, V.H.P. Modelos de simulação para análise de incerteza na previsão de produção de óleo em plataformas da bacia de Campos. Dissertação (Mestrado em Engenharia de Produção) – PUC-Rio, 2023.

MARTINS, D.W. de P. Inteligência artificial aplicada em séries temporais. Dissertação (Mestrado em Estatística) – Universidade de São Paulo,

2024. Disponível em: <https://www.teses.usp.br/teses/disponiveis/45/45133/tde-08042024-094807/>. Acesso em: 1 set. 2025.

MARTINS, T. A importância do petróleo na economia global: impactos e tendências. *Petróleo Energia*, 2023. Disponível em: <https://www.petroleoenergia.com.br/importancia-do-petroleo-na-economia/>.

MATELLI, J.A. Inteligência artificial em engenharia ambiental: aplicações e oportunidades. In: VILANOVA, M.R.N.; SHIINO, M.Y. (org.). *Fronteiras da engenharia e ciências ambientais: perspectivas multidisciplinares*. São Paulo: Editora UNESP, 2020. p. 103–134. Disponível em: <https://editoraunesp.com.br/catalogo/9786557110096,inteligencia-artificial>. Acesso em: 1 set. 2025.

MONTEIRO, T. MMBOE: entenda o que significa essa sigla importante no setor de petróleo e gás. *Dicionário do Petróleo*, 20 mar. 2023. Disponível em: <https://dicionariodopetroleo.com.br/>. Acesso em: 1 set. 2025.

MONTGOMERY, D.; JENNINGS, C.; KULAHCI, M. *Introduction to Time Series Analysis and Forecasting*. Hoboken: John Wiley & Sons, 2008. 1ª ed.

MORAES, M.J.I. de. Impactos do pré-sal na economia brasileira. 210 f. Tese (Doutorado em Economia) – Universidade de São Paulo, Ribeirão Preto, 2013. Disponível em: <https://www.teses.usp.br/teses/disponiveis/96/96131/tde-23082013-110508/?lang=pt-br>. Acesso em: 1 set. 2025.

MORETTIN, P.A.; TOLOI, C.M.C. *Análise de Séries Temporais*. São Paulo: Edgard Blücher, 2006.

MUTTI, E.; CARMINATTI, M. Deep-Water Sands of the Brazilian Offshore Basins. Search and Discovery Article #30219. Apresentado na AAPG International Conference and Exhibition, Milão, 23–26 out. 2011.

NACINBEN, J.P.C.C.S.M. Modelagem de volatilidade estocástica via aproximações de Laplace. Tese (Doutorado em Estatística) – Universidade de São Paulo, 2023. Disponível em: https://www.teses.usp.br/teses/disponiveis/96/96131/tde-05102023-141604/publico/JoaoPedroCCSMNacinben_Corrigida.pdf. Acesso em: 1 set. 2025.

NAVSUPPLY. Fases da exploração do petróleo: etapas cruciais na indústria energética. 2024. Disponível em: <https://navsupply.com.br/>. Acesso em: 1 set. 2025.

NEVES, B.G. Métodos de recuperação em reservatórios de petróleo do pré-sal: análise da injeção alternada de água e gás. Trabalho de Conclusão de Curso (Engenharia de Petróleo) – Universidade Federal Fluminense, Niterói, RJ, 2021. Disponível em: <https://app.uff.br/riuff/bitstream/handle/1/21909/TCC%20-%20FINAL%20-%20BRUNO%20G%20NEVES.pdf>. Acesso em: 1 set. 2025.

OBIDIKE, P.; EMELLE, C.; VLAARDINGERBROEK, R. Integrated Production Forecasting for Business Planning. *Society of Petroleum Engineers*. SPE Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 5-7 ago. 2014.

OYEKUNLE, A.A. Impact of the Petroleum Industry Bill on Deepwater Economics. *Society of Petroleum Engineers*. SPE Nigeria Annual International Conference and Exhibition, Abuja, Nigeria, 30 jul. – 3 ago. 2011.

OYINEBIFUN, E.B.; SYLVA, L.; ONYEBUCHUKWU, E.M. Comparison of ARIMA and ARIMAX Model Forecasts of Monthly Sales of Petroleum Products in Nigeria. *International Journal of Innovative Mathematics, Statistics & Energy Policies*, v. 12, n. 3, p. 1-21, 2024.

PASSOS, L.F.C.; YAGINUMA, K.Y.; PINTO, D.R.C.; MARINS, E.S. Métodos de regularização no aprendizado de máquinas: Ridge e LASSO. Niterói: Universidade Federal Fluminense, 2021. Disponível em: <https://educapes.capes.gov.br/handle/capes/757694>. Acesso em: 3 set. 2025.

PERERA, M.S.A. et al. A review of CO₂-enhanced oil recovery with a simulated sensitivity analysis. *Energies*, Basel, v. 9, n. 7, p. 1-22, 2016. DOI: 10.3390/en9070481. Disponível em: <<https://www.mdpi.com/1996-1073/9/7/481>>. Acesso em: 1 set. 2025.

R CORE TEAM. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2024. Disponível em: <https://www.R-project.org/>. Acesso em: 10 jul. 2025.

RODELLA, F.A. Aplicação de aprendizado de máquina em séries temporais multivariadas para previsão de produção de petróleo. Dissertação (Mestrado em Engenharia de Petróleo) – Universidade Federal do Espírito Santo, Vitória, 2023.

ROSA, A.J.; CARVALHO, R.S.; XAVIER, J.A.D. Engenharia de Reservatórios de Petróleo. Rio de Janeiro: Interciência, 2006.

SANTOS, V.G.A. dos. Previsão da produção de petróleo no Brasil: uma análise comparativa entre modelos de decomposição temporal e Box-Jenkins. 68 f. Trabalho de Conclusão de Curso (Bacharelado em Estatística) – Universidade Estadual Paulista, Presidente Prudente, 2024. Disponível em: <https://repositorio.unesp.br/items/6562e64a-80c1-41ad-8718-23cbb1b83f34>. Acesso em: 1 set. 2025.

SARAPKA, C.E. Controle preditivo baseado em modelo aplicado à otimização de refinarias de petróleo. Dissertação (Mestrado em Engenharia Química) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2020.

SHANG, H.L.; HYNDMAN, R.J. Grouped functional time series forecasting: an application to age-specific mortality rates. *Journal of Computational and Graphical Statistics*, v. 26, n. 2, p. 330-343, 2016.

SHUMWAY, R.H.; STOFFER, D.S. Time series analysis and its applications: with R examples. 4. ed. Cham: Springer, 2017.

SILVA, D.M.G. da et al. Previsão de comportamento da produção de petróleo: estudo para o campo de Furado. Relatório Final (Graduação em Engenharia de Petróleo) – SET/Alagoas, 2023.

SILVA, M. Previsão da produção de petróleo no Brasil: uma análise comparativa entre modelos de decomposição temporal e Box-Jenkins. 2023. Disponível em: <https://repositorio.unesp.br/>. Acesso em: 1 set. 2025.

SILVA, Y.C. de L. e; SILVA FILHO, L.A. da; SILVA, A. de L. e. Crescimento econômico, energia e meio ambiente: efeitos da substituição de combustíveis fósseis na matriz energética mundial. Roraima: Universidade Federal de Roraima, 2023. Disponível em: <https://brsa.org.br/wp-content/uploads/wpcf7-submissions/32251/CRESCIMENTO-ECONOMICO-ENERGIA-E-MEIO-AMBIENTE-EFEITOS-DA-SUBSTITUICAO-DE-COMBUSTIVEIS-FOSSEIS-NA-MATRIZ-ENERGETICA-MUNDIAL.pdf>. Acesso em: 4 set. 2025.

SORIANO-VARGAS, A.; WERNECK, R.; MOURA, R.; DAVOLIO, A. A visual analytics approach to anomaly detection in hydrocarbon reservoir time series data. *Journal of Petroleum Science and Engineering*, v. 206, p. 108988, 2021.

SOUZA, R.C. de. A modelagem univariada por Holt-Winters. PUC-Rio, 1983. Disponível em: https://www.maxwell.vrac.puc-rio.br/8161/8161_3.PDF.

TAKINAMI, N.M. Projeto de bancada para simulação de estimulação de poços de petróleo por ultrassom. Trabalho de Graduação (Engenharia Mecânica) – Faculdade de Engenharia e Ciências, Universidade Estadual Paulista, Guaratinguetá, 2024. Disponível em: <https://repositorio.unesp.br/server/api/core/bitstreams/91a24330-7026-483d-be9e-6d87e967ff9d/content>. Acesso em: 1 set. 2025.

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, 1996.

TICHY, Malter. Time to retire the MAPE. *Foresight: The International Journal of Applied Forecasting*, v. 69, p. 5-12, 2023.

VENABLES, W.N.; RIPLEY, B.D. Modern applied statistics with S. 4. ed. New York: Springer, 2002.

WILLMOTT, C.J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, v. 30, p. 79-82, 2005. Disponível em: <https://www.scirp.org/reference/referencespapers?referenceid=1432886>. Acesso em: 10 jul. 2025.

WINTERS, P.R. Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science*, v. 6, n. 3, p. 324-342, 1960.

YAHYAOU, I.; MARÍÑAS-COLLADO, I.; GARCÍA SIPOLS, A.E.; SIMÓN DE BLAS, C.; RODRÍGUEZ SÁNCHE, C. Application of the Double Smoothing and ARIMAX Methods for the Prediction of Polycrystalline Photovoltaic

Generation. In: IECON 2022 – 48th Annual Conference of the IEEE Industrial Electronics Society, Bruxelas, Bélgica, 17–20 out. 2022. IEEE Xplore, 2022. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/9968612>>.

YANRUI et al. Previsão de produção de petróleo na Bacia DJ. 2022. Disponível em: <https://www.sciencedirect.com/science/article/pii/S014098832300587X/pdf>. Acesso em: 1 set. 2025.

YERGIN, D. O petróleo: uma história mundial de conquistas, poder e dinheiro. São Paulo: Paz e Terra, 2010.

YIN, R.K. Case Study Research: Design and Methods. *SAGE*, 2009.

Apêndice A – Protocolo e Questionário de Entrevista

Primeiramente, os entrevistadores se apresentam, realizam um breve resumo dos principais objetivos do trabalho de conclusão de curso e explicam o tipo de informação que buscam absorver com a entrevista. Em seguida, os entrevistados se apresentam e explicam um pouco sobre as suas funções dentro de suas áreas na indústria de óleo e gás. A entrevista é iniciada pelos entrevistadores e segue os tópicos abaixo por blocos:

Bloco 1 – Produção e Declínio

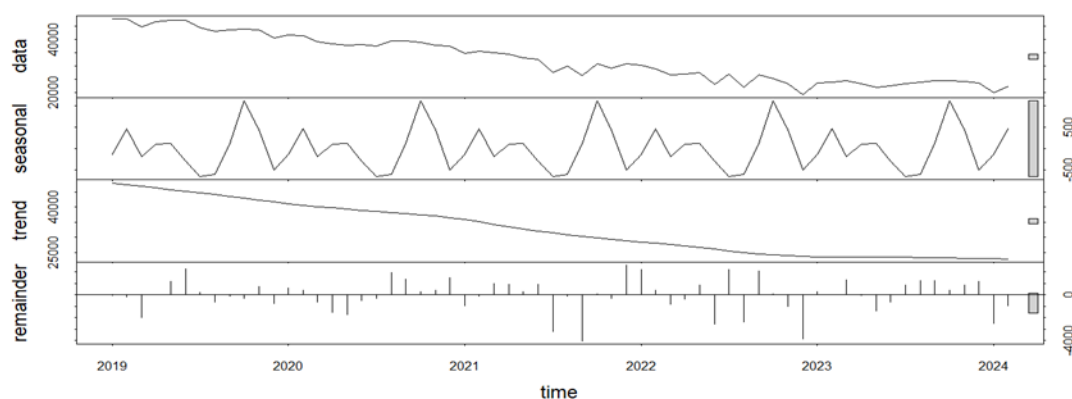
1. Quais são, na sua experiência, os principais fatores que influenciam o início do declínio da produção em campos como Roncador?
2. Quais são as técnicas de recuperação secundária e recuperação terciária utilizadas?
3. Como você avalia a eficácia das técnicas de recuperação secundária e terciárias aplicadas no Roncador?
4. Qual o papel da produção de água na produção do campo?
5. A produção de água tem sido um desafio relevante no campo? Como ela tem sido monitorada e gerenciada?
6. Você acredita que a variabilidade da produção de água pode ser usada como indicador preditivo da produtividade do campo? Por quê?
7. Como podemos caracterizar até quando o poço é economicamente viável? Que indicadores temos?

Bloco 2 – Previsão de Produção

8. Sua equipe utiliza modelos estatísticos ou computacionais para prever a produção de petróleo? Se sim, quais?
9. Já teve contato com modelos que utilizam variáveis exógenas (como ARIMAX)? Como avalia sua aplicabilidade no contexto do Roncador?
10. Como você avalia o final da produção útil do poço?

11. Das análises de tendencias realizadas é possível identificar fatores de sazonalidade tendencias, variações aleatórias na composição da previsão? Por que existem variações aleatórias?
12. Você identifica sazonalidade no gráfico abaixo?

Gráfico 4: Decomposição da Série Temporal (m³/d)



13. Em sua opinião, quais variáveis externas (exógenas) mais impactam a produção de petróleo no Roncador?
14. Você considera que a inclusão da produção de água como variável exógena melhora a acurácia dos modelos de previsão?
15. Como você avalia o uso de modelos como *ETS*, Holt-Winters e ARIMAX em campos maduros? Algum deles já foi testado por sua equipe?

Bloco 3 – Estratégia, Dados e Tomada de Decisão e Considerações Finais

16. Como as previsões de produção influenciam decisões operacionais e de investimento no campo?
17. Quais são os principais desafios enfrentados atualmente na previsão de produção em campos maduros como Roncador?
18. Que sugestões você daria para melhorar a acurácia e aplicabilidade dos modelos preditivos no contexto do campo Roncador?