



Incerteza na previsão de produção de óleo em plataformas

Vitor Hugo Pinheiro Marques

Orientador: Fernando Luiz Cyrino Oliveira

Coorientador: Antônio Orestes de Salvo Castro



- 1 Introdução
- 2 Contextualização e objetivos
- 3 Revisão Bibliográfica
- 4 Metodologia
- 5 Desenvolvimento do aplicativo
- 6 Análise dos resultados

Análise de incertezas na previsão de produção de petróleo

Empresa de Petróleo envolve complexas decisões devido ao ambiente de incertezas

(ABATH, 2020).



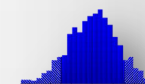
- Arbitrariedade dos reservatórios; equipamentos de Plataformas; e alto problemas complexos em E&P.

Necessidade de quantificar as incertezas em E&P surge naturalmente a partir da modelagem

(GUIMARÃES, 2016).



- As distribuições de probabilidade são usadas para modelar a incerteza de modo estocástico
- Avaliar os atributos de entrada, para caracterizá-las em um modelo, e gerir riscos associados



A maioria das organizações petrolíferas decidem usar procedimentos consistentes de análise de risco, buscando ferramentas para modelos de simulação mais complexos.

SUSLICK e SCHIOZER (2004)



O volume de petróleo que será produzido, denominado meta de produção, influencia diretamente as ações das empresas e tem um crítico impacto em relação ao mercado.

BOMFIM (2017)



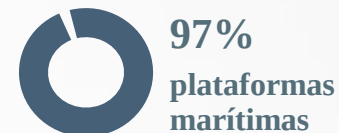
Escopo na produção de óleo de plataformas marítimas:

- Produção de petróleo brasileira no mundo:



Fonte: IEA [2021].

- Produção de Petróleo no Brasil:



Fonte: ANP [2022].

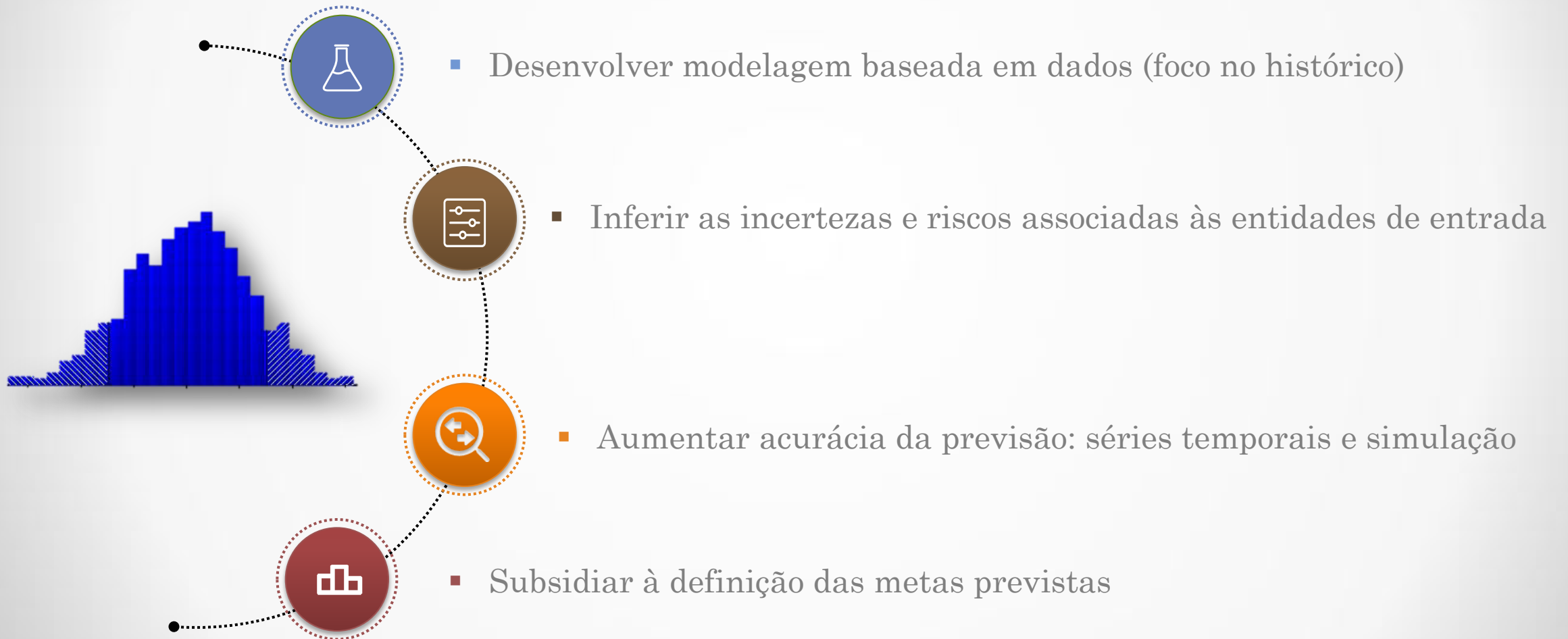


O trabalho analisa a região geográfica da Bacia sedimentar de Campos na empresa Petrobras.

- Área marinha de aproximadamente 100 mil KM²;
- Localização compreende do centro-norte do RJ até o sul ES;
- Campo marítimo mais maduro brasileiro (1977)
 - Grande quantidade dados históricos
 - Pode servir como benchmarking para previsão de novas plataformas marítimas
- Região há 5 anos realiza abaixo da meta (difere dos outros campos da empresa)



Analisar incertezas da previsão de produção de óleo

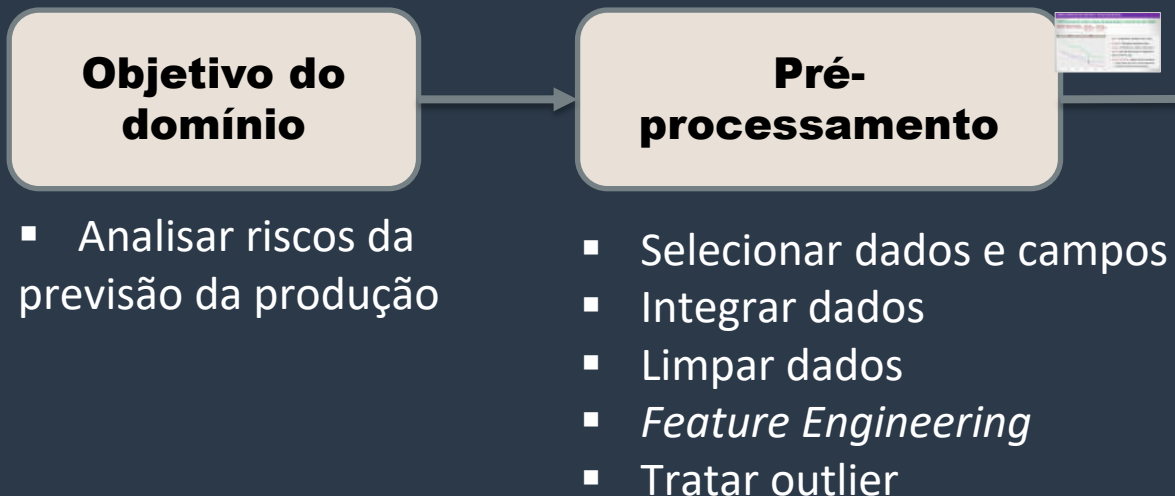


Análise de risco de previsão da meta de produção de óleo

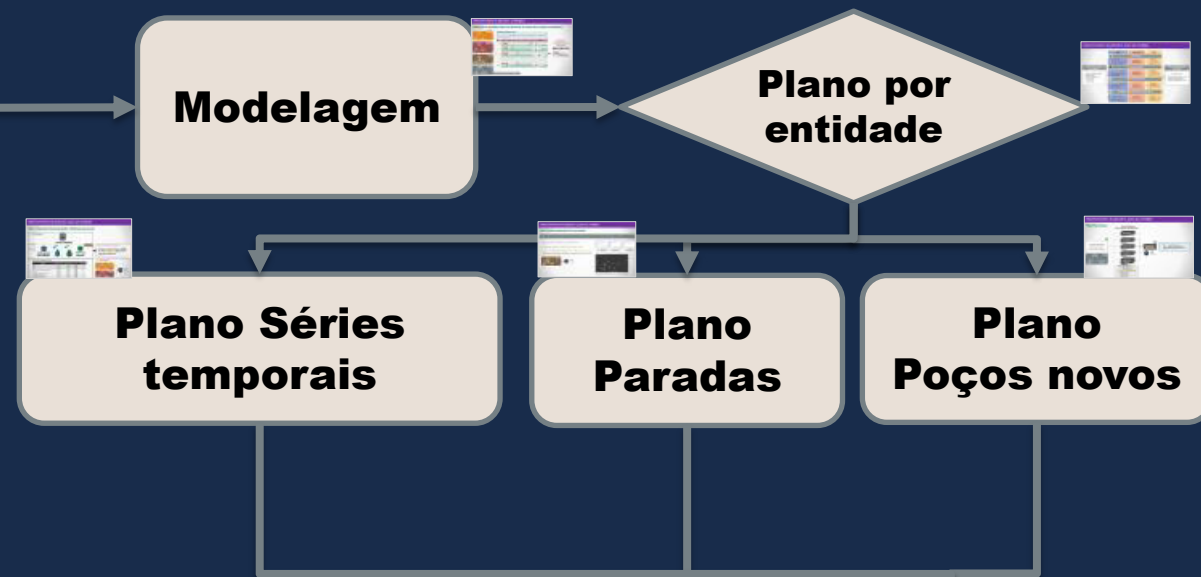


Desenvolvimento do aplicativo

➤ Preparação da base:

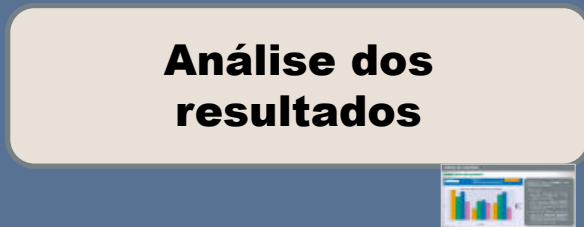


➤ Plano de Experimento:

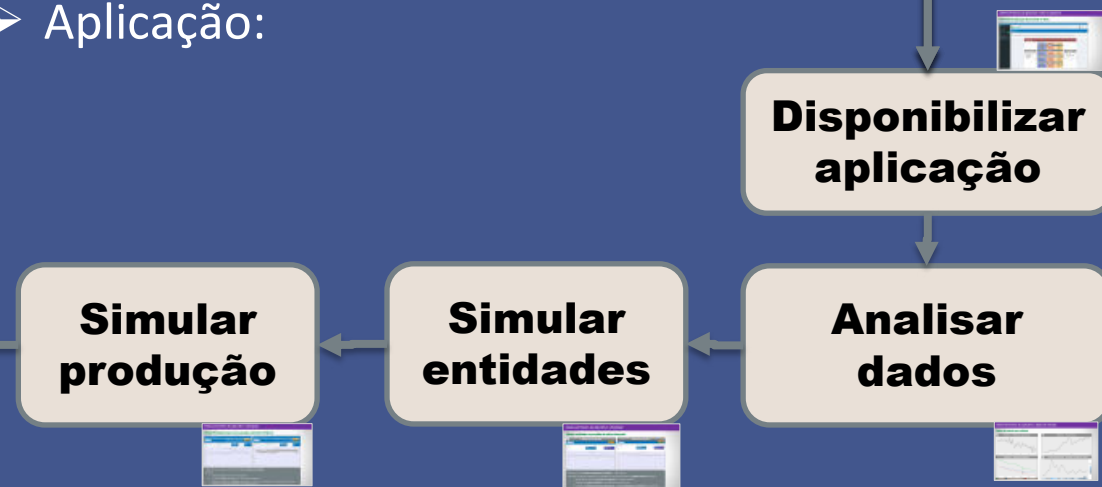


➤ Avaliação dos Modelos:

- Análise de erro das previsões
- Interpretação de riscos
- Suporte a definição da meta



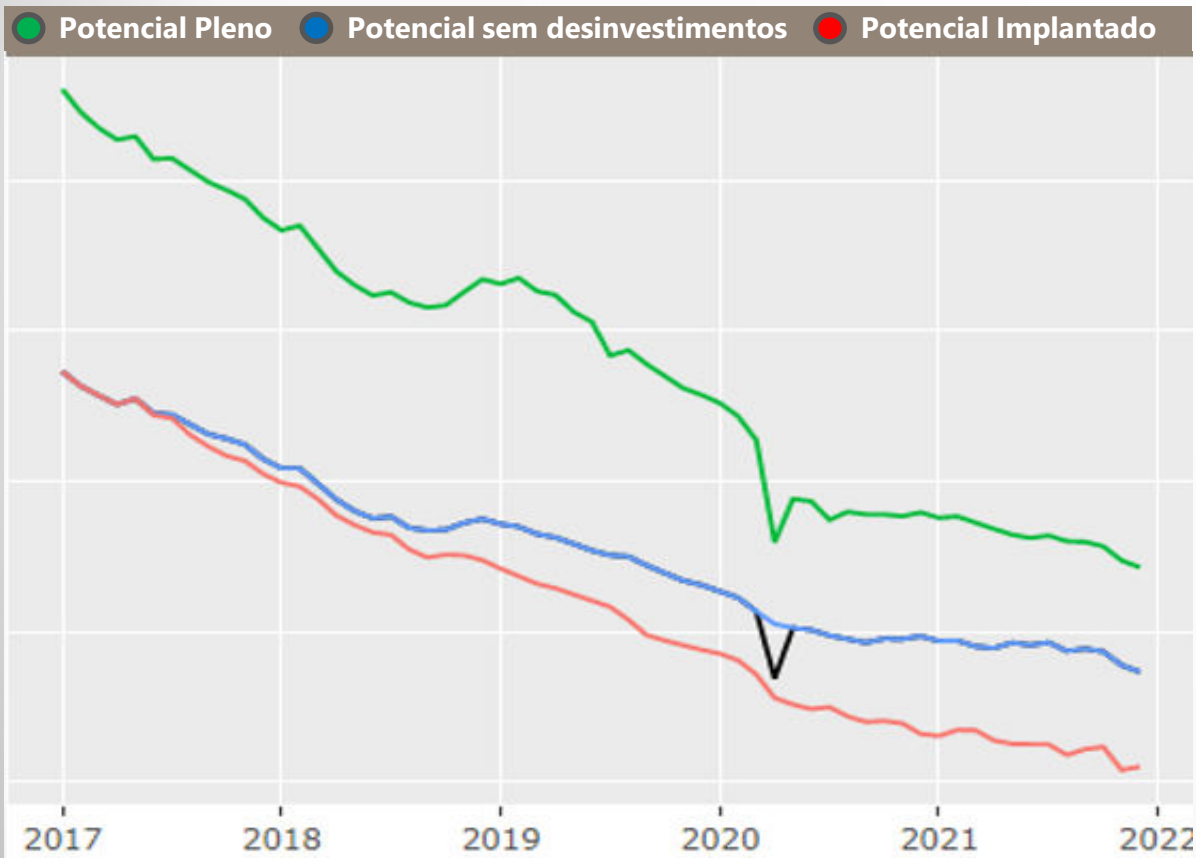
➤ Aplicação:



Desenvolvimento do aplicativo: Pré-processamento

O Pré-Processamento contém a seleção, integração, limpeza e tratamento dos dados iniciais.

Eficiência Histórica **Potencial Histórico** **Perdas Históricas** **Curva de Poços Novos (entraram)** **Curva de Poços Novos (a entrar)**

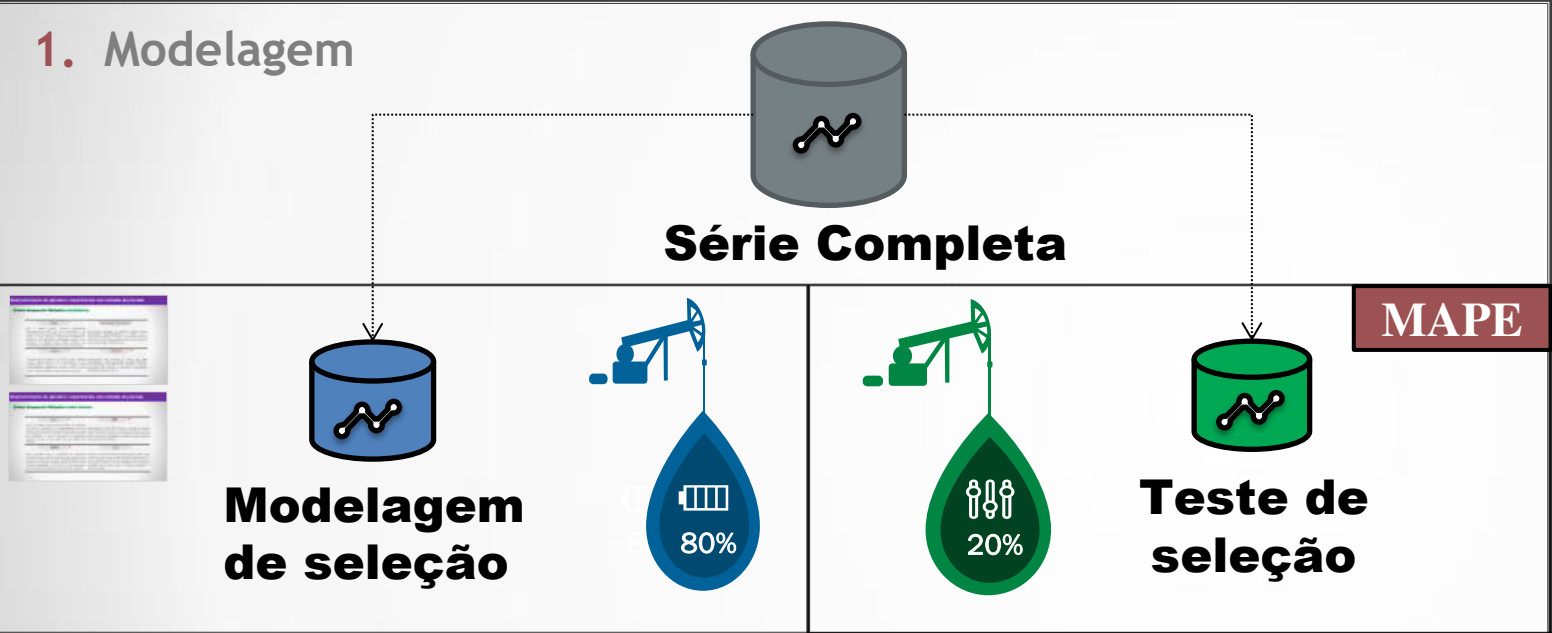


- **Dados:** 22 Plataformas marítimas (2017 a 2021)
- **Integração:** Padronizar nomenclatura Poços
- **Limpeza:** 29 plataformas vendidas ou hibernadas
- **Outlier:** Abr/20 devido COVID-19. Imputação de dados por EWMA (4M)
- **Feature engineering:** *splitting variable* do potencial
 - Seleção métodos de previsão: Potencial Implantado
 - Simulação: Potencial sem desinvestimentos

Desenvolvimento do aplicativo: plano por entidade

Séries Temporais: Potencial produtivo e Eficiência operacional

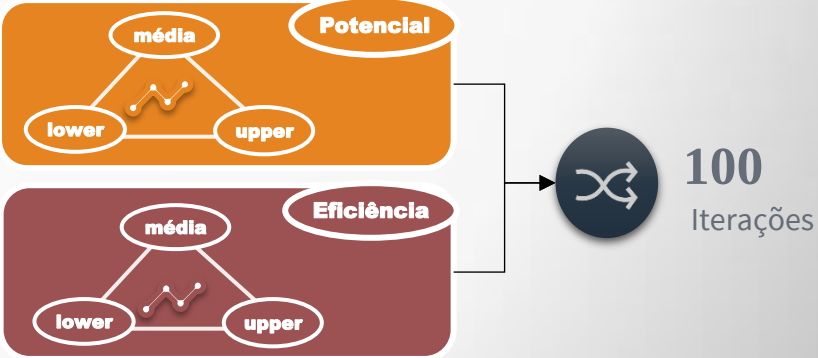
1. Modelagem



2. Seleção dos Métodos

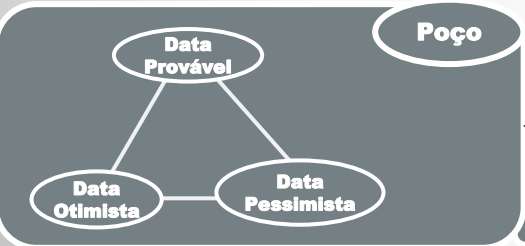
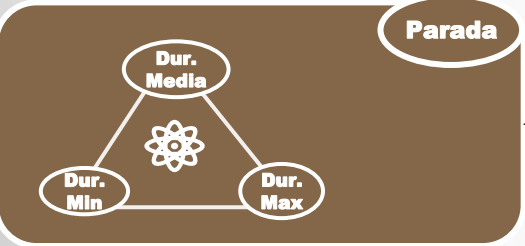
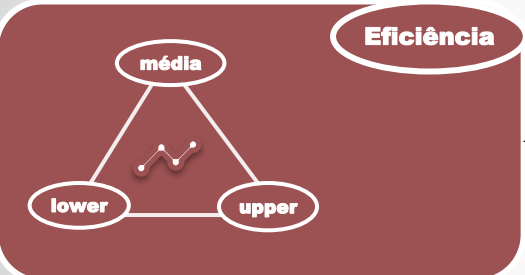
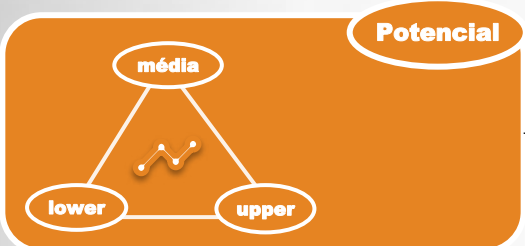
MODELO	Média Eficiência	Média Potencial	Geral	Classificação
TBATS	7,90%	5,18%	6,54%	1
ARIMA	8,09%	6,49%	7,29%	3
ETS	7,91%	7,19%	7,55%	4
Redes Neurais Artificiais Recorrentes	8,91%	6,91%	7,91%	5
Prophet	10,63%	8,78%	9,70%	7
Multilayer Perceptron Network	10,14%	15,29%	12,72%	8
Radial Basis Function	9,64%	9,49%	9,56%	6
Extreme Learning Machine	9,37%	4,94%	7,16%	2

3. Simulação



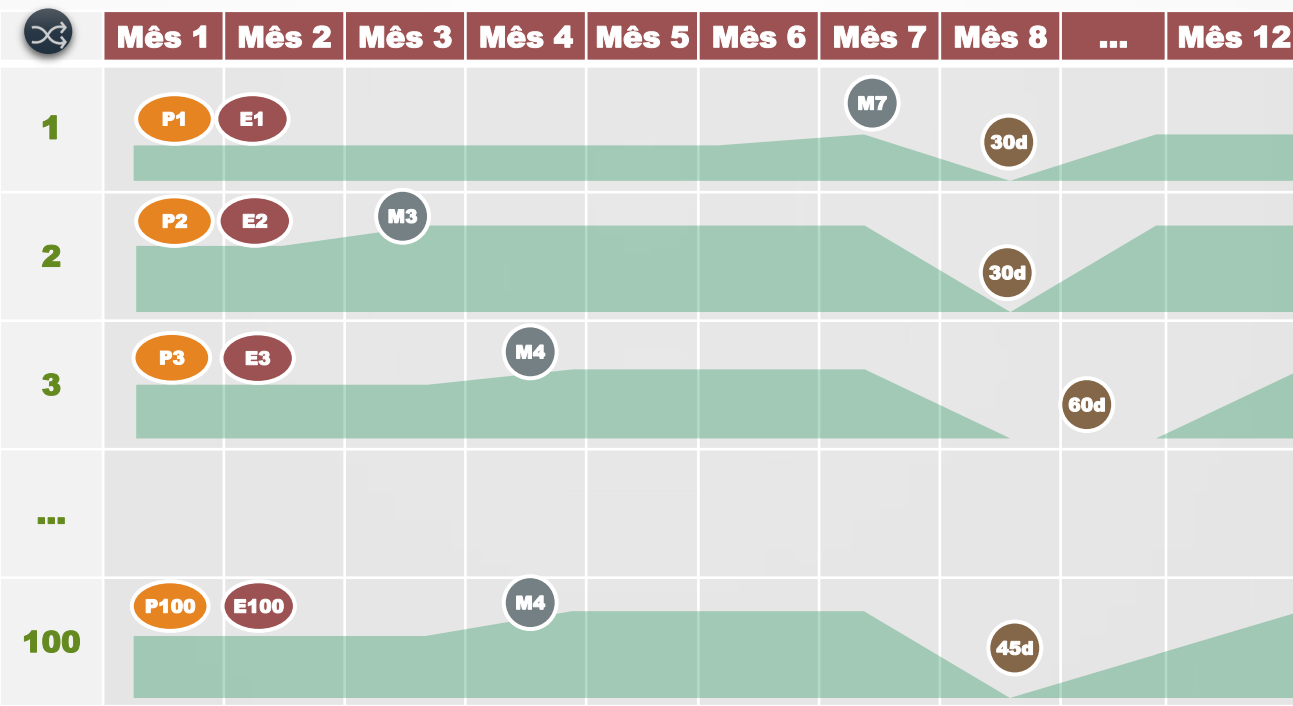
Desenvolvimento do aplicativo: modelagem

Modelagem do simulador: dados das dimensões de entrada até a projeção probabilística



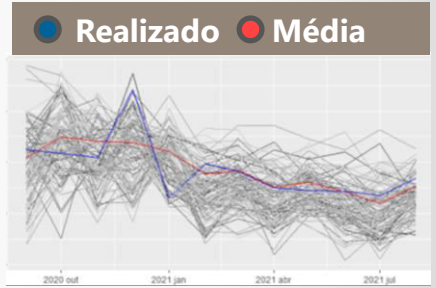
Simulação Monte Carlo

$$\sum_{n=1}^N \sum_{t=1}^T \sum_{p=1}^P \frac{ \left[\left(\text{Potencial}_p[t] + \text{Potencial}_i[t] \right) - \text{Parada}_i[t] \right] * \text{Eficiência}_i[t] }{T}$$



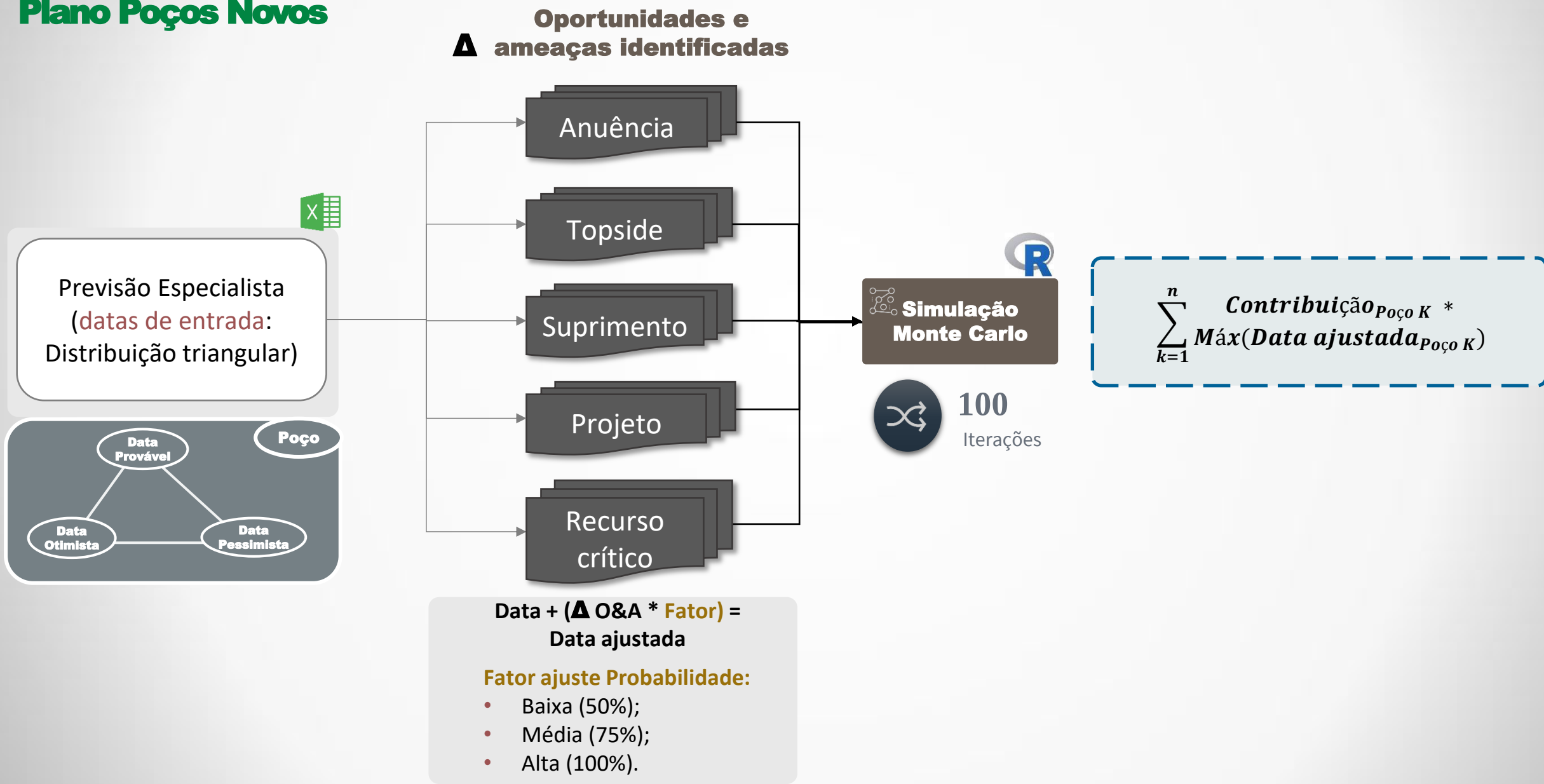
Obs.: único com base em projeção

Projeção Probabilística de Produção



Desenvolvimento do aplicativo: plano por entidade

Plano Poços Novos

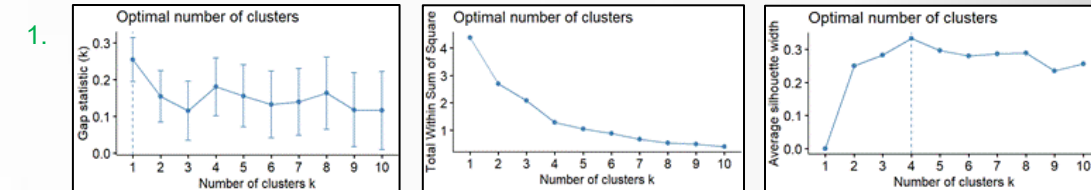
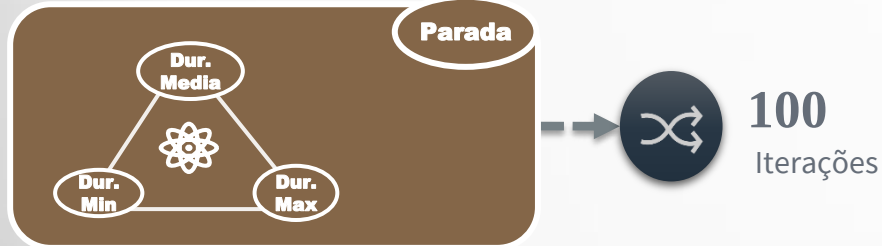


Desenvolvimento do aplicativo: plano por entidade

Plano Paradas programadas para manutenção

Histórico de 5 anos por plataforma é baixo para análises estatísticas (~1/ano): clusters por plataformas similares;

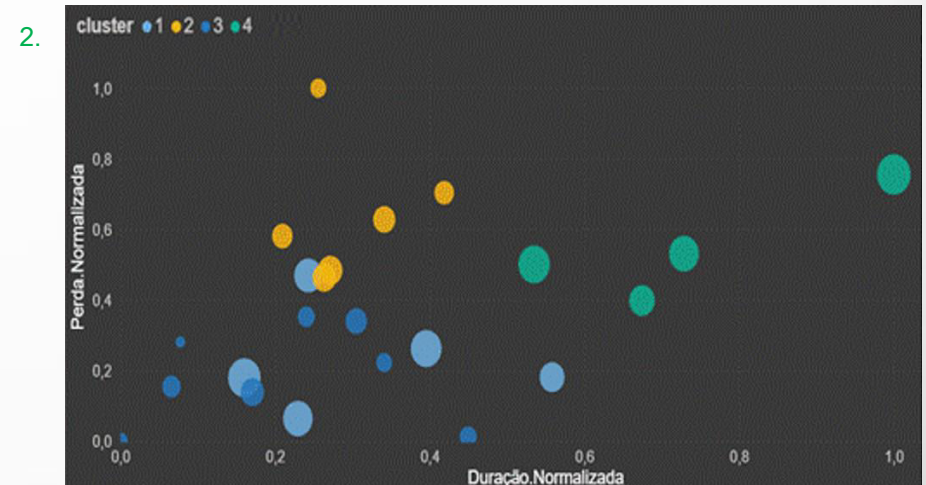
1. Identificar número de clusters: 4 (índices NbClust);
2. Clusterização k-means: Duração, Volume Perda e Idade;
3. Simular com distribuição triangular: mín., média e máx. por cluster
(**delta realizado e planejado da duração dos eventos históricos do cluster**)



(a) Silhouette

(b) Gap stat

(c) WSS



Duração (x), Volume de perda (y) e Idade (circunferência)

Desenvolvimento do aplicativo: sobre o aplicativo

Tela inicial da aplicação desenvolvida em Shiny

Simulador de Producao

Sobre o Aplicativo

Dados de entrada

Simular Entidade

Simular Produção **Resultado**

Analise

» Principais Riscos

» Comparação com Meta

» Comparação do Erro



OBJETIVO

Elaborar modelagem capaz de inferir as incertezas associadas aos dados históricos de potencial de produção e eficiência operacional, e de projeções em poços novos e paradas para manutenção, fazendo uso de séries temporais e simulação para aumentar a acurácia das previsões de produção de óleo em plataformas marítimas da Bacia de Campos, e analisar os principais riscos relacionados.

Metodologia

	Input	Simulação	Saída
Poços Novos	- Estimativas data início: - Oportunidade adiantar - Ameaça postergar	- Monte Carlo - Dist. triangular	Projeções probabilísticas por data início
Potencial	- Série temporal (5 anos) - Modelagem (4 anos) - Teste (1 ano)	- Monte Carlo - Dist. triangular	Projeções probabilísticas por Forecasting
Eficiência (com ajuste das Paradas Programadas)	- Série temporal (5 anos) - Modelagem (4 anos) - Teste (1 ano)	- Monte Carlo - Dist. triangular	Projeções probabilísticas por Forecasting
Paradas	- Eventos (5 anos) - Clusters Plataformas: idade,duração,volume	- Monte Carlo: - Dist. triangular	Projeções por delta de duração previsto e real

Pré-processamento dos dados

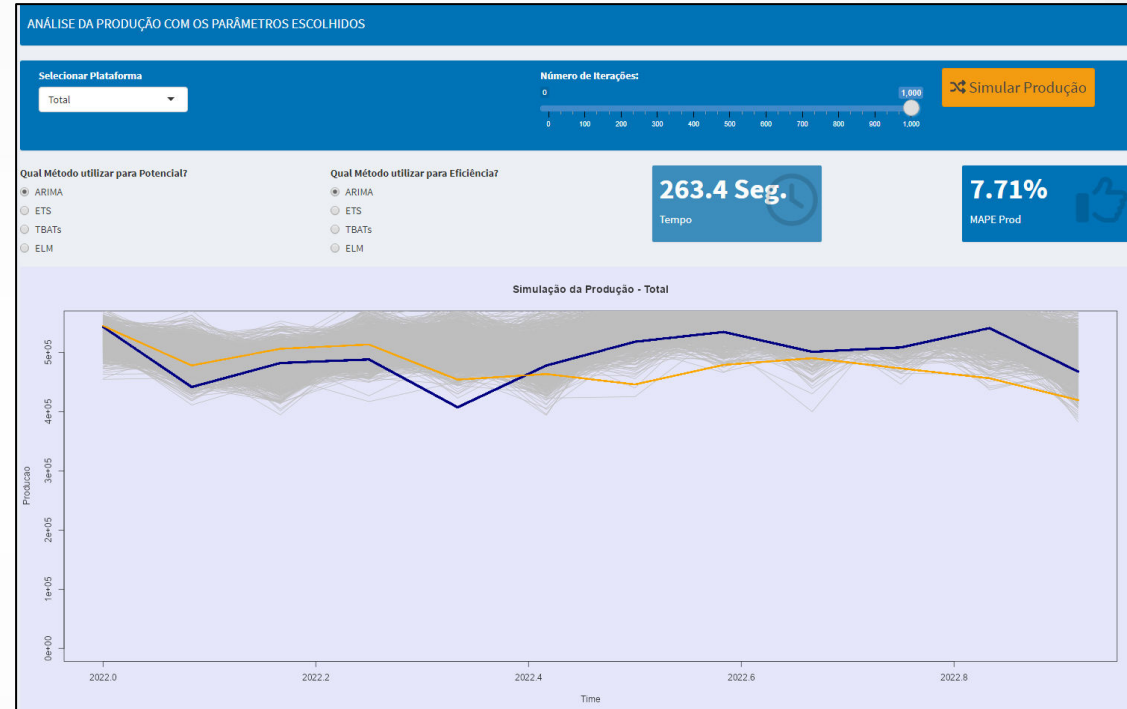
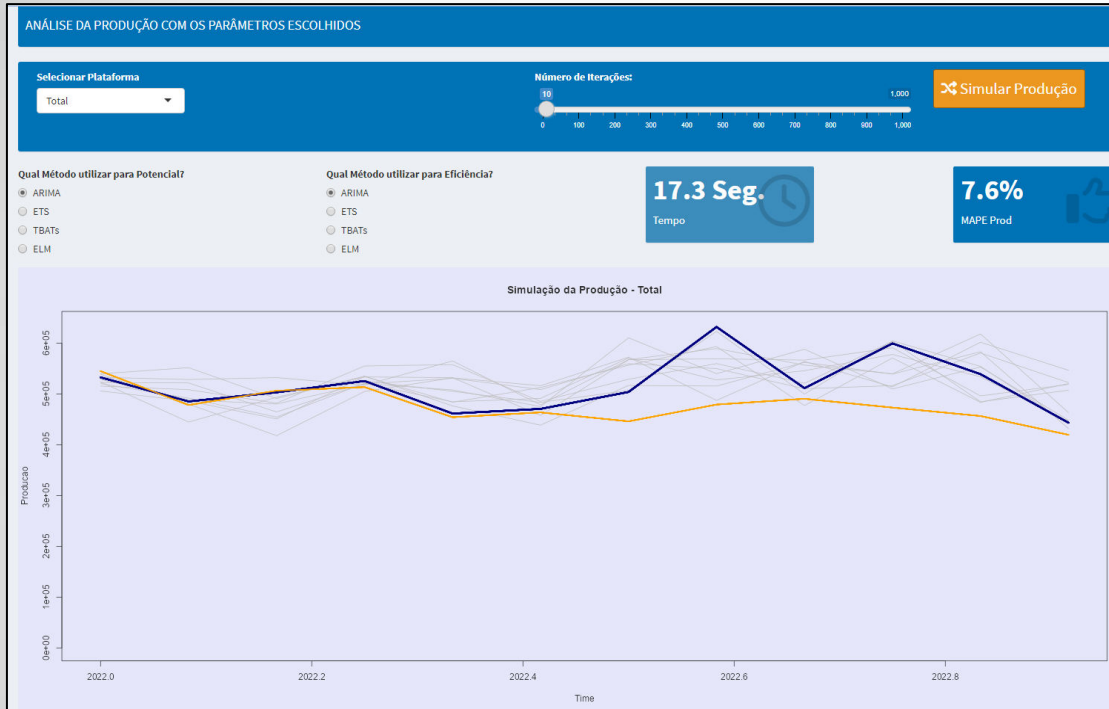
- Selecionar dados/Campos
- Integrar dados
- Feature Engineering
- Limpar dados
- Tratar outliers

Pós-processamento dos dados

- Simulador Shiny
- Interpretação de padrões
- Análise de erro das previsões

Desenvolvimento do aplicativo: simulação

Simular Produção: poços novos, paradas, potencial e eficiência



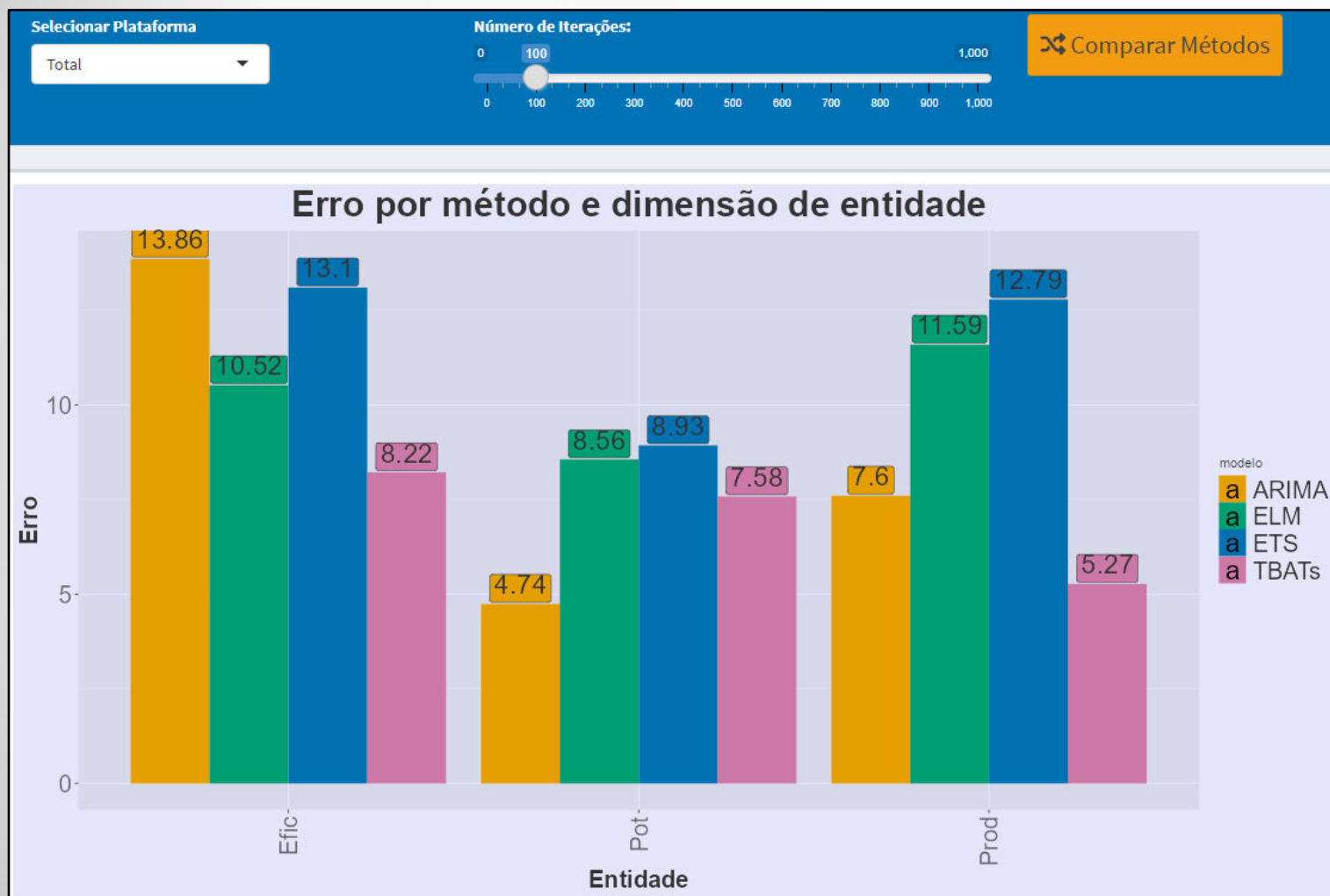
Ambiente para analisar iterações da curva mensal de produção

Permite:

- Visualizar erros (próximo entre exemplos);
- Visualizar tempos (diferença em ~4min. entre exemplos).
- Aumentar número de iterações: recomenda-se, ao final do processo de calibração das curvas, executar a simulação da produção com 1000 iterações

Análise dos resultados

Análise do erro das previsões



Ambiente para comparar os métodos de previsão

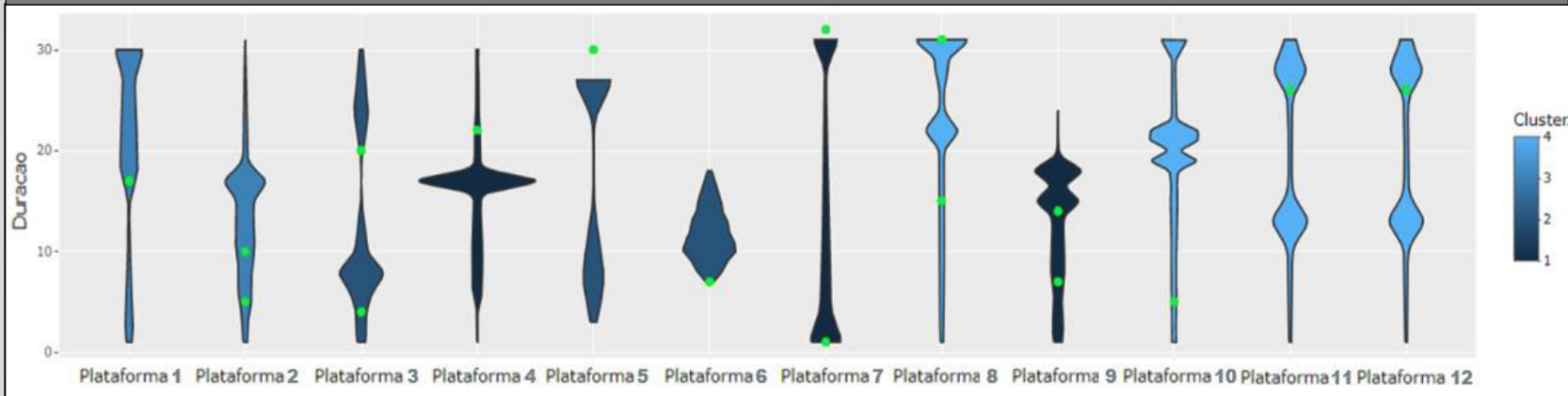
Observa-se:

- Erro das dimensões de entrada eficiência e potencial, e do resultado da produção;
- TBATs apresentou maior acurácia para o cálculo da produção, com um erro de (5,3%) enquanto o ARIMA foi o segundo (7,6%);
- Para todos os métodos utilizados, a variável de potencial apresenta menor erro do que a eficiência, o que denota maior assertividade do seu processo de previsão.

Análise dos resultados

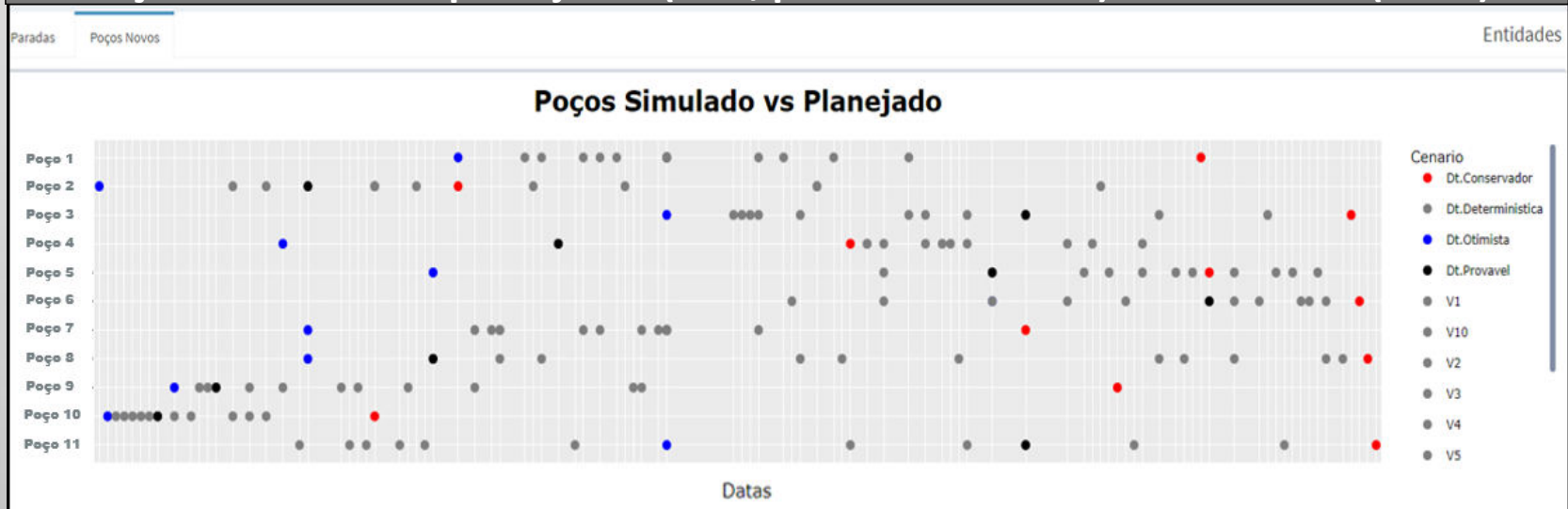
Interpretação de riscos: análise das entidades sem uso de séries temporais

Paradas para manutenção: planejadas (verde) vs simuladas (cinza)



- Plataformas 2, 6, e 10 apresentam planejamentos menores do que a maior parte da densidade das simulações, logo indica maiores durações;
- plataformas 4 e 5 possuem indicação para considerar menores durações.

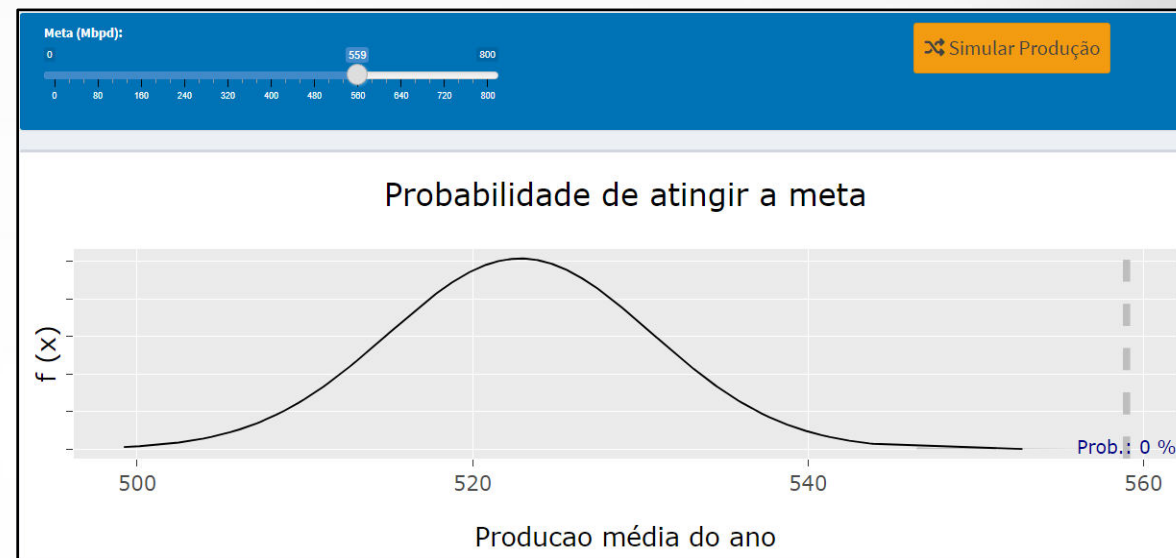
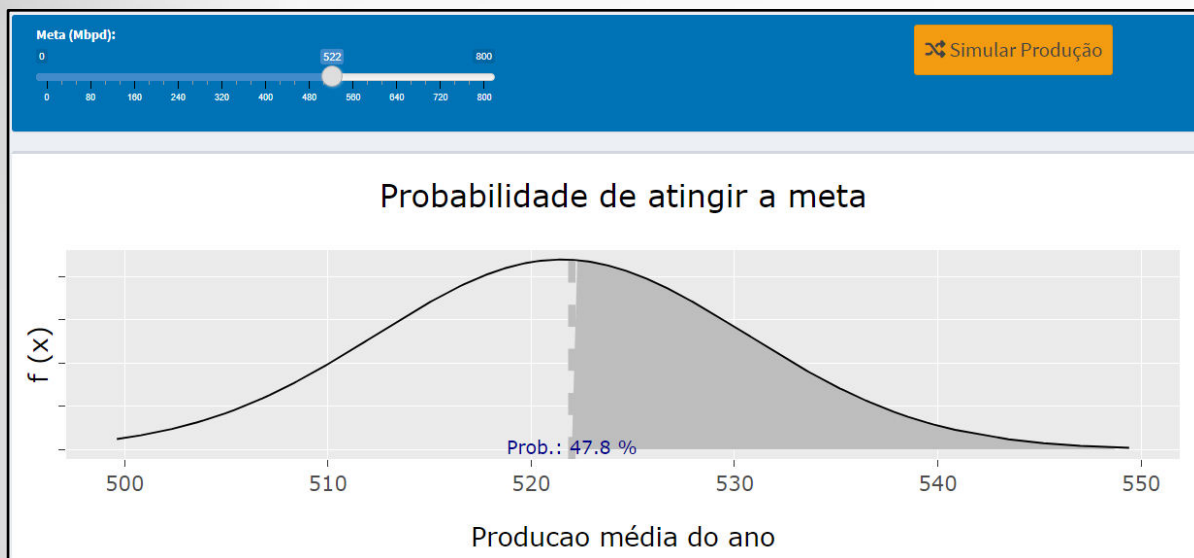
Poços novos: datas planejadas (azul, preto e vermelho) vs simuladas (cinza)



- Poço 5 tem simulação mais próximas da data conservadora, logo tem maior ameaça de postergar a data de entrada;
- Poço 11 o simulador projeta datas mais cedo, logo tem maior oportunidade de antecipar.

Análise dos resultados

Suporte a definição da meta



- **Objetivo do trabalho:** auxiliar na definição da meta com base em dados numéricos e objetivos.
- **Permite** comparar distribuição da simulação em relação à meta almejada
- Necessidade do simulador: de acordo com os riscos identificados, simulador indica 0% de chances de atingir a meta determinística definida para 2022.
- Eficácia do modelo: valor mais provável de produção (~521 Mbpd; 1000 iterações) se aproximou do valor que foi efetivamente realizado no ano de 2022 (526 Mbpd)

Considerações finais



Processo com resultados preliminares mostra possibilidades de aprimorar acurácia

- Histórico do potencial projeta baixo erro;
- Realizado de eficiência ainda tem de expurgar paradas.
- Outros Backlogs: comparar dados com realizado; detalhar inputs no resultado

Modelo atinge os principais objetivos almejados:

- Identifica principais ricos por plataforma e poço;
- Subsidia a determinação da meta com base em dados objetivos, com clara probabilidade de alcance, e auxilia desdobramento por plataforma.

Trabalhos futuros:

- Extrapolar a análise dos dados históricos para poços novos.
- Processo de Eficiência (maior erro) seja aprimorado por meio de estudos de perdas históricas e mapeamento de mudanças nas projeções futuras.
- Manter estudo com o realizado anual, como Ciclo PDCA
- Utilizar outros parâmetros históricos e aplicar análise de sensibilidade.

Referências

- ANP. (2022). Brasil produz cerca de 3 milhões de bbl/d de petróleo em abril. Boletim Mensal de Produção de Petróleo e Gás Natural da Agência Nacional de Petróleo, Agência Nacional do Petróleo, Gás Natural e Biocombustíveis. Publicado em 03/06/2022. Disponível em: https://www.gov.br/anp/pt-br/canais_atendimento/imprensa/noticias-comunicados/brasil-produz-cerca-de-3-milhoes-de-bbl-d-de-petroleo-em-abril
- Bergmeir, C., Hyndman, R. J., Benítez, J.M. (2016). Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation. *International Journal of Forecasting*, 32(2), 303–312.
- Bomfim, G. A. (2017). Modelagem da Data de Entrada em Produção de Poços de Petróleo Utilizando Inferência Fuzzy. PUC-Rio, Rio de Janeiro, RJ.
- Breiman, L. (1996). Bagging predictors. *Machine Learning* 24 (2): 123–140.
- Castro, A. O. de S., Mattos, J. R., Silva, S. B. P. (2010). Job Risk - Análise Probabilística da Previsão de Produção de Óleo. Seminário de Planejamento, Controle e Avaliação do Desempenho do E&P. Anais...Rio de Janeiro.
- Correia, J. et. al. (2021). Comparing ARIMA and LSTM models to predict time series in the oil industry. 129-136. 10.5753/kdmile.2021.17470.
- Dantas, T. M., Oliveira, F. L. C. (2018). Improving time series forecasting: An approach combining bootstrap aggregation, clusters and exponential smoothing. *International Journal of Forecasting*. 34. 748–761.
- EFRON, B. (1979). Bootstrap methods: another look at the Jackknife. *The Annals of Statistics*, 7, 1–26.
- Fan, D. et. al. (2021). Well production forecasting based on arima-lstm model considering manual operations. *Energy* vol. 220, pp. 119708.
- Fayyad, U. M. et al. (1996). From Data Mining to Knowledge Discovery in Databases. *AI Mag.* 17 37-54.
- Fayyad U. M., Stolorz. P. 1997. Data mining and KDD: Promise and challenges. *Future Generation Computer Systems*, Vol. 13, P. 99-115.
- Marins et. al. (2021). Fault detection and classification in oil wells and production/service lines using random forest, *Journal of Petroleum Science and Engineering*, Vol 197.
- Makridakis, S., Spiliotis, E., Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS ONE*. 13. 10.1371/journal.pone.0194889.

Referências

- Mckay, M. D., Beckman, R. J., Conover. W. J. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*.
- Mounir B. A.& Hela L. & C. Kolski & A. Alimi [2010]. A user-centered approach for the design and implementation of KDD-based DSS: A case study in the healthcare domain. *Decision Support Systems*. 50. p. 64-78.
- Pan, Q., Daniel D. "An efficient reliability method combining adaptive support vector machine and Monte Carlo simulation." *Structural Safety* 67 (2017): 85-95
- ROBERT, C. P., CASELLA, G. (1998). *Monte Carlo Statistical Methods*.
- Sinaga, K. P. e Yang, M. "Unsupervised K-Means Clustering Algorithm," in *IEEE Access*, vol. 8, pp. 80716-80727, 2020, doi: 10.1109/ACCESS.2020.2988796.
- Tealab, A. (2018). Time series forecasting using artificial neural networks methodologies: A systematic review, *Future Computing and Informatics Journal*, Vol. 3, Issue 2” (2018): 334-340.
- Tang, L. & Lv, H. & Yang, F. & Yu, L. (2015). Complexity testing techniques for time series data: A comprehensive literature review. *Chaos, Solitons & Fractals*. 81. p.117-135.
- Tariq, Z., Aljawad, M.S., Hasan, A. et al. (2021). A systematic review of data science and machine learning applications to the oil and gas industry. *J Petrol Explor Prod Technol* 11, 4339–4374. <https://doi.org/10.1007/s13202-021-01302-2>.
- Vrugt, Jasper A. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling & Software* 75 (2016): 273-316.
- WAGNER. C. et. Al. 2016. Machine Condition Monitoring and Fault Diagnostics with Imbalanced Data Sets based on the KDD Process. *IFAC journal*. Vol. 49.
- Yanrui N., Hossein K., Pejman T. (2022). A comparative machine learning study for time series oil production forecasting: ARIMA, LSTM, and Prophet, *Computers & Geosciences*, Vol. 164.