

Análise de Dados do Sistema SCADA

Relatório Final do Projeto Seguindo o CRISP-DM

Thiago Macedo, Luciano Farias, Pedro Pagnussat e João Vitor Abitante
05/07/2024

1. Entendimento do Negócio

1.1. Elipse Software

Elipse Software é uma empresa especializada no desenvolvimento de ferramentas de software para gestão de informações em tempo real em processos industriais, energia, infraestrutura, água e saneamento. Seus produtos, como o Elipse E3 e Elipse Power, integram diferentes sistemas em uma arquitetura única, proporcionando valor e produtividade. A Elipse Software tem presença global e seus produtos são utilizados por importantes empresas ao redor do mundo, oferecendo suporte técnico, plataformas flexíveis e comunicação com qualquer dispositivo.

1.2. PowerSysLab

PowerSysLab (PSL) é uma empresa brasileira especializada em soluções tecnológicas para o setor elétrico. Fundada em 2006, a PSL utiliza pesquisa acadêmica de ponta para desenvolver softwares que melhoram a eficiência e competitividade de seus clientes. Com uma equipe multidisciplinar experiente, a empresa oferece produtos que otimizam o planejamento, operação e proteção de redes elétricas. A PSL também colabora com centros de pesquisa renomados e, desde 2008, mantém uma parceria com a Elipse Software para desenvolver a plataforma Elipse Power.

1.3. O Sistema SCADA e o Elipse Water

Sistema SCADA

O Sistema de Controle e Aquisição de Dados (SCADA) é uma plataforma essencial para o controle e monitoramento de processos industriais complexos. Composto por sensores, controladores e uma estação de controle, o SCADA permite a coleta de dados em tempo real, análise e controle de processos. Os sensores captam dados como temperatura e pressão, os controladores processam esses dados e executam ações, enquanto a estação de controle permite a interface com os operadores, exibindo dados e alarmes, e possibilitando controle manual. Essa estrutura centraliza o monitoramento contínuo e a automação de processos, garantindo eficiência operacional e segurança.

Elipse Water

O Elipse Water é uma solução SCADA desenvolvida especificamente para o setor de saneamento básico. Ele gerencia e controla sistemas de abastecimento de água e tratamento de esgoto, oferecendo ferramentas avançadas para monitoramento em tempo real, controle de processos, otimização de recursos, análise de dados e geração de relatórios. Este sistema melhora

significativamente a eficiência operacional e a gestão dos recursos hídricos. Um exemplo prático é o monitoramento de uma planta de distribuição de água, onde o Elipse Water controla bombas, monitorando vazão e nível do reservatório, otimizando o uso de energia e garantindo um abastecimento eficiente.

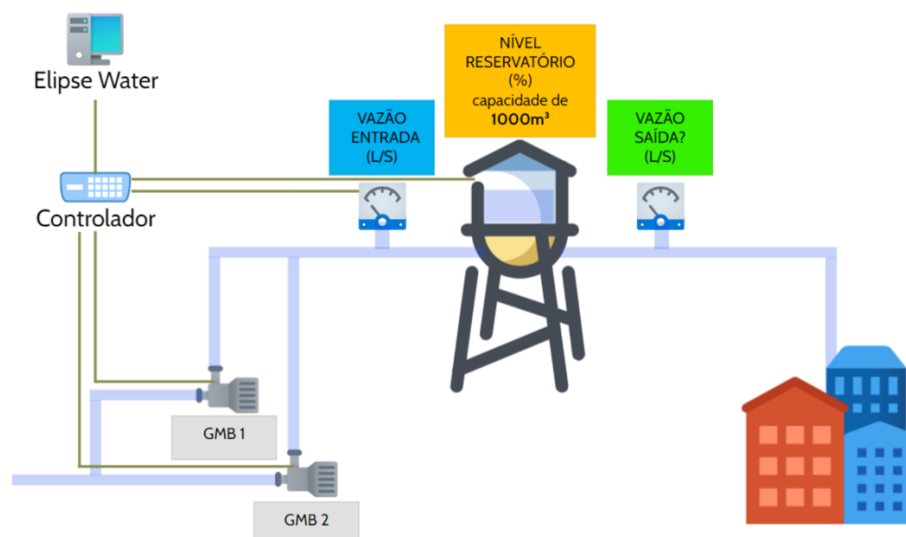
1.4. Objetivo do Projeto

O objetivo deste projeto é utilizar os dados coletados pelo sistema SCADA, especificamente pelo Elipse Water, para responder a diversas perguntas críticas relacionadas à operação e otimização de sistemas de saneamento básico. As perguntas propostas incluem:

1. Qual é a vazão de saída do reservatório a cada momento?
2. Qual a curva típica da saída do reservatório ao longo de 24h durante dias úteis e finais de semana?
3. Qual o tempo de uso das bombas em horários de ponta e fora de ponta?
4. Qual a previsão da vazão de saída para as próximas 24h?
5. Para um determinado dia, horário e nível do reservatório, em quanto tempo ele ser esvaziará caso falte energia?
6. Há correlação entre a temperatura e o consumo de água?
7. Qual a previsão da vazão de saída para as próximas 24h em um determinado dia e horário, considerando uma previsão de temperatura?
8. Para um determinado dia, horário, nível do reservatório e previsão da vazão de saída para as próximas 24h, qual deveria ser a programação horária de cada bomba, de modo a minimizar o uso delas durante o horário de ponta, enquanto mantém o nível do reservatório entre 20% e 95%?

Sistema Exemplo

O sistema exemplo monitorado pelo Elipse Water consiste em uma planta de distribuição de água tratada, incluindo duas bombas elevatórias, um reservatório e um conjunto de consumidores. O sistema SCADA coleta dados em tempo real sobre o status das bombas, a vazão de entrada no reservatório e seu nível, permitindo uma análise detalhada e a tomada de decisões estratégicas.



Este projeto visa maximizar a eficiência e a eficácia das operações de saneamento básico por meio de uma análise detalhada e estratégica dos dados coletados pelo sistema SCADA, oferecendo insights valiosos para a gestão dos recursos hídricos.

2. Entendimento dos Dados

2.1. Os Dados Originais

Dois conjuntos de dados foram disponibilizados para o projeto. Um deles contendo os dados obtidos pelo sistema SCADA, onde cada linha apresenta o estado do sistema em um determinado momento do dia. O segundo conjunto apresenta diversos dados meteorológicos, também em um determinado momento.

Amostra do conjunto de dados do sistema SCADA:

| DATA/HORA | VAZÃO ENTRADA (L/S) | NÍVEL RESERVATÓRIO (%) | PRESSÃO (mca) | GMB 1 | GMB 2 |
|---------------------|---------------------|------------------------|---------------|-------|-------|
| 2023-03-17 11:27:06 | 68.59 | 29.86 | 38.20 | 10 | 90 |
| 2023-03-17 12:28:56 | 66.05 | 35.86 | 38.20 | 10 | 90 |

Informações importantes sobre o conjunto da SCADA:

- Número de linhas: 125.073
- Colunas:
 - DATA/HORA: Horário do registro no tipo timestamp
 - VAZÃO ENTRADA (L/S): Litros de água que entram no reservatório por segundo
 - NÍVEL RESERVATÓRIO (%): Porcentagem do nível do reservatório
 - PRESSÃO (mca): Pressão dentro do reservatório
 - GMB 1: Bomba de água 1 (10 para desligada, 90 para ligada)
 - GMB 2: Bomba de água 2 (10 para desligada, 90 para ligada)

O segundo conjunto de dados disponibilizado contém dados sobre o clima. Esse conjunto foi dividido em dois, o primeiro contendo os dados de 2023 e o segundo contendo os dados de 2024. Ambos com as mesmas informações. As colunas desse conjunto são as seguintes:

- Data
- Hora UTC
- PRECIPITAÇÃO TOTAL, HORÁRIO (mm)
- PRESSAO ATMOSFERICA AO NIVEL DA ESTACAO, HORARIA (mB)
- PRESSÃO ATMOSFERICA MAX.NA HORA ANT. (AUT) (mB)
- PRESSÃO ATMOSFERICA MIN. NA HORA ANT. (AUT) (mB)
- RADIACAO GLOBAL (Kj/m²)
- TEMPERATURA DO AR - BULBO SECO, HORARIA (°C)
- TEMPERATURA DO PONTO DE ORVALHO (°C)
- TEMPERATURA MÁXIMA NA HORA ANT. (AUT) (°C)
- TEMPERATURA MÍNIMA NA HORA ANT. (AUT) (°C)
- TEMPERATURA ORVALHO MAX. NA HORA ANT. (AUT) (°C)

- TEMPERATURA ORVALHO MIN. NA HORA ANT. (AUT) (°C)
- UMIDADE REL. MAX. NA HORA ANT. (AUT) (%)
- UMIDADE REL. MIN. NA HORA ANT. (AUT) (%)
- UMIDADE RELATIVA DO AR, HORARIA (%)
- VENTO, DIREÇÃO HORARIA (gr) (° (gr))
- VENTO, RAJADA MAXIMA (m/s)
- VENTO, VELOCIDADE HORARIO (m/s)

2.2. Exploração Inicial e Checagem de Qualidade

Após uma avaliação simples, foi percebido que o conjunto de dados da SCADA apresentava alguns problemas marcantes e que, se deixados sem tratamento, causariam dificuldades significativas no processo de modelagem, além de gerar resultados não confiáveis nas respostas das perguntas. Os principais problemas encontrados foram:

- Bombas da água ligadas, porém, vazão de entrada zerada.
- Bombas da água desligadas, porém, vazão de entrada significativa
- Nível do reservatório zerado de forma inconsistente:
 - Pequenas janelas: de uma a cinco linhas onde o nível do reservatório estava zerado, mas registros adjacentes mostravam continuidade normal.
 - Grandes janelas: centenas ou milhares de linhas com nível zerado, apesar de vazão de entrada significativa e bombas ligadas.
- Registros sem pressão no reservatório, mesmo com vazão de entrada e bombas da água ligadas.

3. Preparação dos Dados

3.1. Limpeza e Remoção dos Problemas

O conjunto de dados da SCADA teve um processo de limpeza e processamento extenso. As etapas foram as seguintes:

- Renomeação das colunas para nomes mais simples e troca dos status da bomba para 0: desligado, 1: ligado
- Criação de colunas de datas: ano, mês, dia, dia da semana, semana do mês, hora e minuto
- Correção das linhas onde as bombas estavam desligadas e havia vazão de entrada: linhas foram alteradas e as bombas foram ligadas
- Correção das linhas onde as bombas estavam ligadas e não havia vazão de entrada: foi adicionada uma vazão de entrada com base na bomba ligada e na vazão de entrada média nessa data.
- Para corrigir pequenas janelas de inconsistência no nível do reservatório, foi necessário ajustar gradualmente os valores entre as linhas adjacentes. Por exemplo, se na linha 10 o nível era xx e na linha 14 era $x-10x-10$, os níveis entre as linhas 11 e 13 foram ajustados dividindo a diferença total (10) por 4 (número de linhas problemáticas mais a próxima linha normal). Assim, a linha 11 recebeu $x-2.5x-2.5$, a linha 12 $x-5x-5$, a linha 13 $x-7.5x-7.5$ e a linha 14 manteve $x-10x-10$. Isso garantiu uma transição suave e lógica nos dados.
- Correção das grandes janelas com o nível do reservatório inconsistente: A correção das grandes janelas foi extremamente complexa. A solução utilizada na correção das pequenas janelas não poderia ser aplicada pois, nesses grandes intervalos, as bombas eram ligadas e desligadas periodicamente, logo, o nível do reservatório aumentava e diminuía

periodicamente, acompanhando o comportamento das bombas. Para deixar os dados consistentes, as grandes janelas foram separadas em janelas menores, entre sub-janelas com as bombas ligadas, onde o nível do reservatório imputado de forma gradual, e sub-janelas onde as bombas estavam desligadas, onde o nível do reservatório foi diminuído gradualmente.

- Criação das colunas necessárias: Novas colunas foram criadas, contendo:
 - Nível do reservatório em litros
 - Tempo decorrido em segundos
 - Total de litros que entraram no reservatório: vazão de entrada multiplicada pelo tempo decorrido
 - Litros que efetivamente entraram no reservatório: total de litros que entraram menos os litros que saíram
 - Litros que saíram do reservatório: diferença do nível do reservatório anterior e nível atual (zero caso negativo)
 - Vazão de saída da água em litros por segundo: litros que saíram dividido pelo tempo decorrido

Dessa forma, o conjunto de dados resultante manteve consistência e padronização, impedindo erros matemáticos que futuramente causariam problemas no processo de modelagem e resposta das perguntas.

Para o conjunto de dados meteorológicos, a limpeza apenas consistiu na remoção das colunas julgadas menos úteis. Foram mantidas apenas as colunas da umidade relativa, temperatura e precipitação total naquele momento, além das colunas referentes às datas: ano, mês, dia e hora.

3.2. Agrupamento

Com ambos os conjuntos limpos e sem problemas, foi realizado o processo de agrupamento. O conjunto de dados meteorológicos foi agregado ao conjunto da SCADA. Depois, o conjunto resultante foi agrupado por hora. Anteriormente, havia um registro a cada 2 minutos e 30 segundos (em média).

Com o conjunto de dados agregado por hora, novas colunas foram criadas para maximizar a informação. As colunas do conjunto final são as seguintes:

- timestamp
- year
- month
- day
- day_of_week
- week_of_year
- hour
- input_flow_rate_first
- input_flow_rate_last
- input_flow_rate_mean
- reservoir_level_percentage_first
- reservoir_level_percentage_last
- reservoir_level_percentage_mean
- output_flow_rate_first
- output_flow_rate_last
- output_flow_rate_mean
- pressure_first
- pressure_last
- pressure_mean
- pump_1_duration_sum

- pump_2_duration_sum
- air_temp_c
- total_precip_mm
- relative_humidity_percentage
- input_flow_rate_diff
- output_flow_rate_diff
- reservoir_level_change
- pressure_change
- pump_1_active
- pump_2_active
- change_from_last_hour_output_flow_rate_mean

3.3. Imputação de Valores Faltantes

Com o conjunto de dados agregado por hora, a etapa final do tratamento foi realizada: a imputação dos valores faltantes. O conjunto inteiro ficou com 6158 linhas. No intervalo de tempo contido nos dados, havia 2478 linhas faltantes, aproximadamente 103 dias.

Durante o processo de imputação de dados, cada linha do conjunto de dados foi avaliada para identificar valores ausentes. Se uma linha apresentasse valores ausentes, os valores das colunas foram preenchidos utilizando informações das linhas adjacentes.

Para os índices das linhas anteriores e posteriores, as taxas de vazão de entrada e saída, bem como o nível do reservatório e a pressão, foram ajustadas. Especificamente, as taxas de vazão e os níveis de reservatório foram imputados com base nos valores das linhas anteriores e seguintes. Se a linha seguinte também estivesse faltando dados, os valores foram imputados com base na linha atual. A média entre os valores inicial e final foi calculada para representar a média da coluna correspondente.

Além disso, os dados climáticos, como temperatura do ar, precipitação total e umidade relativa, foram imputados usando a média das três linhas anteriores. As mudanças nas taxas de vazão de entrada e saída, bem como nas pressões, foram calculadas e imputadas de acordo com a diferença entre os valores iniciais e finais. A duração das bombas também foi ajustada com base nas linhas anteriores.

Por fim, uma nova coluna foi criada para identificar se a linha teve dados imputados, e outra para calcular a mudança na taxa de vazão de saída em comparação com a hora anterior. Este processo garantiu a consistência dos dados, permitindo análises mais precisas e confiáveis.

4. Modelagem

4.1. Modelos e Abordagem

Para responder às perguntas propostas, criamos 48 modelos, sendo 24 para prever a vazão de saída de água sem dados meteorológicos e 24 com dados meteorológicos, cada um prevendo uma das 24 horas seguintes. Utilizamos o algoritmo XGBoost para todos os modelos, que foram treinados nos dados de 2023 e testados nos dados de 2024. Cada modelo previa a diferença da vazão de saída em relação à linha anterior, mantendo a série temporal centrada em 0. Por exemplo, o modelo 1 previa a diferença entre a vazão de saída original e a vazão de saída 1 hora depois. Da mesma forma, o modelo 2 previa a diferença entre a vazão de saída da hora 1 e a vazão de saída da hora 2 e assim sucessivamente até 24 horas, utilizando todos os 24 modelos.

4.2. Conjunto de Treino

O conjunto de treino foi criado a partir do dataset gerado pela fase do tratamento dos dados. Utilizamos um script em Python que gera características de defasagem (lag features) e janelas móveis (window features). Essas características incluíam taxas de vazão de entrada e saída, níveis de reservatório, pressão, temperatura do ar, precipitação e umidade relativa. Além disso, criamos alvos (targets) para cada uma das próximas 24 horas.

As colunas utilizadas no treinamento são:

Com os seguintes lags: 1, 2, 3, 6, 12, 24, 26, 48, 72:

- {lag}_hours_ago_input_flow_rate_first
- {lag}_hours_ago_input_flow_rate_last
- {lag}_hours_ago_input_flow_rate_mean
- {lag}_hours_ago_reservoir_level_percentage_first
- {lag}_hours_ago_reservoir_level_percentage_last
- {lag}_hours_ago_reservoir_level_percentage_mean
- {lag}_hours_ago_output_flow_rate_first
- {lag}_hours_ago_output_flow_rate_last
- {lag}_hours_ago_output_flow_rate_mean
- {lag}_hours_ago_pressure_first
- {lag}_hours_ago_pressure_last
- {lag}_hours_ago_pressure_mean
- {lag}_hours_ago_pump_1_duration_sum
- {lag}_hours_ago_pump_2_duration_sum
- {lag}_hours_ago_temperature (somente nos modelos com dados meteorológicos)
- {lag}_hours_ago_precipitation (somente nos modelos com dados meteorológicos)
- {lag}_hours_ago_humidity (somente nos modelos com dados meteorológicos)

Além das features das linhas passadas (lag), também utilizamos as médias móveis de algumas colunas, utilizando as seguintes janelas: 2, 3, 6, 12, 24, 36, 48, 72. As colunas geradas foram as seguintes:

- {janela}_hours_rolling_input_flow_rate_diff
- {janela}_hours_rolling_output_flow_rate_diff
- {janela}_hours_rolling_reservoir_level_change
- {janela}_hours_rolling_pressure_change
- {janela}_hours_rolling_change_from_last_hour_output_flow_rate_mean
- {janela}_hours_rolling_temperature
- {janela}_hours_rolling_precipitation
- {janela}_hours_rolling_humidity

Por fim, cada nova linha também tinha 24 variáveis alvo, cada uma contendo a diferença entre vazão de entrada média daquela hora e vazão de entrada média da hora seguinte correspondente.

4.3. Treinamento

O treinamento dos modelos foi realizado utilizando o XGBoost, com a configuração dos parâmetros básica e validação cruzada em série temporal (TimeSeriesSplit). Os modelos foram treinados nos dados de 2023 e testados nos dados de 2024. Cada modelo foi ajustado para minimizar o erro absoluto médio (MAE), com avaliações adicionais de erro quadrático médio (MSE), raiz do erro quadrático médio (RMSE) e R^2 .

Para realizá-lo, foram elaborados 24 scripts de treinamento, onde cada script treinava dois modelos: um deles com os dados meteorológicos e outro sem. Os 48 modelos treinados foram salvos para usos posterior.

4.4. Avaliação

Seguem abaixo as métricas de cada um dos 48 modelos:

| Modelo | MAE: Com dados meteorológicos | MAE: Sem dados meteorológicos | RMSE: Com dados meteorológicos | RMSE: Sem dados meteorológicos |
|-------------|-------------------------------|-------------------------------|--------------------------------|--------------------------------|
| Xgboost 1h | 8.09 | 6.67 | 10.95 | 9.19 |
| Xgboost 2h | 8.96 | 7.21 | 12.1 | 10.01 |
| Xgboost 3h | 7.96 | 7.41 | 11.03 | 10.42 |
| Xgboost 4h | 7.61 | 7.07 | 10.66 | 9.97 |
| Xgboost 5h | 12.79 | 6.77 | 19.27 | 9.7 |
| Xgboost 6h | 9.86 | 7.44 | 13.82 | 10.43 |
| Xgboost 7h | 7.07 | 6.91 | 10.2 | 9.87 |
| Xgboost 8h | 7.72 | 8.09 | 10.76 | 10.96 |
| Xgboost 9h | 7.43 | 7.05 | 10.29 | 9.85 |
| Xgboost 10h | 8.17 | 7.43 | 11.2 | 10.34 |
| Xgboost 11h | 7.69 | 6.93 | 10.61 | 9.88 |
| Xgboost 12h | 7.71 | 7.15 | 10.66 | 10.07 |
| Xgboost 13h | 8.53 | 7.04 | 11.73 | 10.06 |
| Xgboost 14h | 7.84 | 6.75 | 10.94 | 9.76 |
| Xgboost 15h | 7.56 | 7.42 | 10.67 | 10.55 |
| Xgboost 16h | 7.77 | 7.7 | 10.81 | 10.79 |
| Xgboost 17h | 8.22 | 7.17 | 11.15 | 10.24 |
| Xgboost 18h | 7.64 | 6.74 | 10.73 | 9.74 |
| Xgboost 19h | 7.13 | 7.01 | 10.06 | 10.0 |
| Xgboost 20h | 8.97 | 7.47 | 12.41 | 10.66 |
| Xgboost 21h | 9.94 | 6.88 | 14.07 | 9.8 |
| Xgboost 22h | 11.09 | 6.72 | 16.61 | 9.66 |
| Xgboost 23h | 7.32 | 8.12 | 10.3 | 11.07 |
| Xgboost 24h | 7.43 | 7.36 | 10.32 | 10.27 |

5. Resolução das Perguntas Propostas

5.1. Pergunta 1

A primeira pergunta: “Qual a vazão de saída do reservatório a cada momento?” é a mais simples de se responder. Respondida no próprio tratamento dos dados, o seu valor é a quantidade total de água que saiu do reservatório menos a quantidade total de água que entrou no reservatório dividida pelos segundos que se passaram entre a linha atual e a linha anterior.

$$Volume = \left(\frac{PctReserva}{100} \right) * 10^6$$

$$\Delta Volume = \Delta Vi = Vi - Vi-1$$

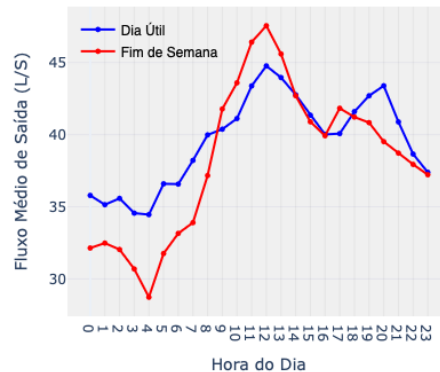
$$\Delta Tempo = \Delta T_{(s)} = t_i - t_{i-1}$$

$$VazaoSaida = -\frac{\Delta Volume}{\Delta t} + VazaoEntrada$$

5.2. Pergunta 2

A segunda pergunta: "Qual a curva típica da saída do reservatório ao longo de 24h durante dias úteis e finais de semana?" foi respondida através de uma simple manipulação nos conjunto de dados processado. Primeiro, o conjunto de dados foi dividido entre dias úteis e finais de semana. Depois, cada um desses dois sub-conjuntos de dados foi agrupado por hora do dia, gerando a média da vazão de saída para cada hora do dia, tanto nos dias úteis quanto nos finais de semana. Seguem os resultados:

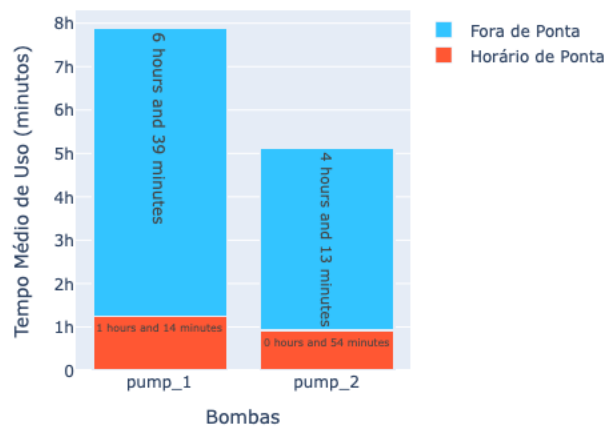
Comparação do Fluxo Médio de Saída por Hora



5.3. Pergunta 3

A terceira pergunta: "Qual o tempo de uso das bombas em horários de ponta e fora de ponta?" foi respondida através da análise do conjunto de dados de consumo de água. Primeiro, os dados foram filtrados para separar os horários de ponta (18h às 21h) e fora de ponta. Em seguida, os tempos de operação das bombas 1 e 2 foram somados diariamente para ambos os períodos. A média diária de uso foi calculada para cada bomba e período, e os resultados foram convertidos para horas e minutos para facilitar a interpretação. Seguem os resultados:

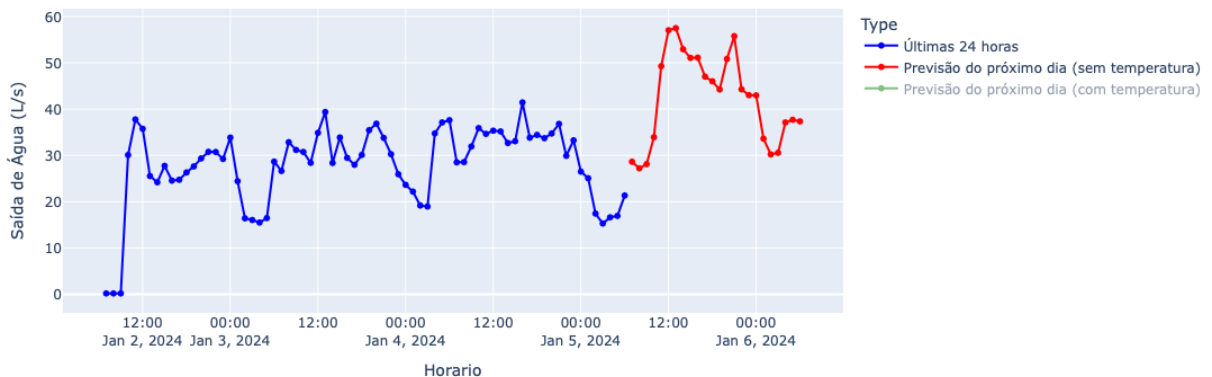
Tempo Médio de Uso das Bombas



5.4. Pergunta 4

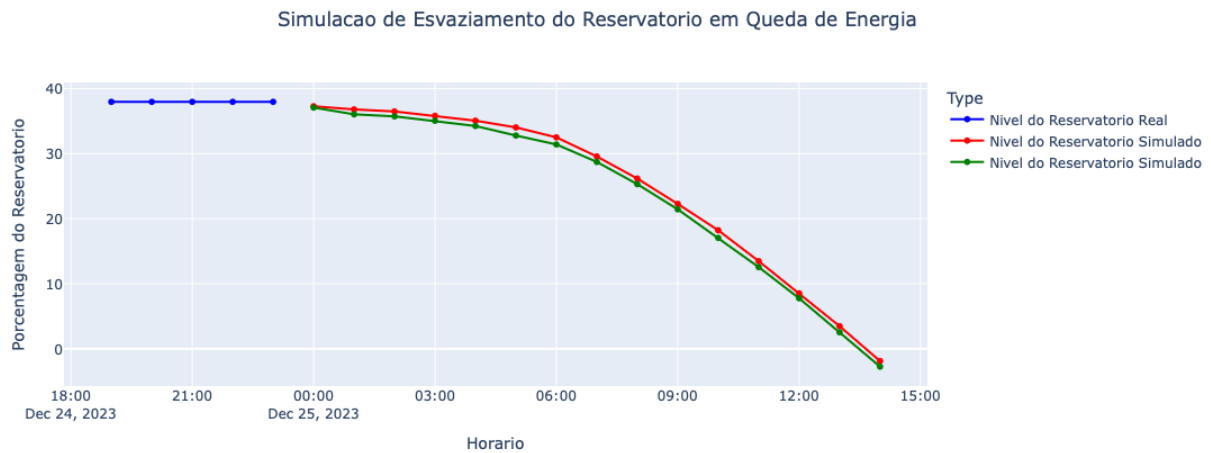
A quarta pergunta: "Qual a previsão da vazão de saída para as próximas 24h?" foi respondida através da criação de 24 modelos, cada um prevendo a vazão de saída de água para cada uma das próximas 24 horas, sem considerar dados meteorológicos. Esses modelos foram treinados utilizando dados históricos e a diferença da vazão de saída em relação à hora anterior. Seguem os resultados:

Previsão da Saída de Água



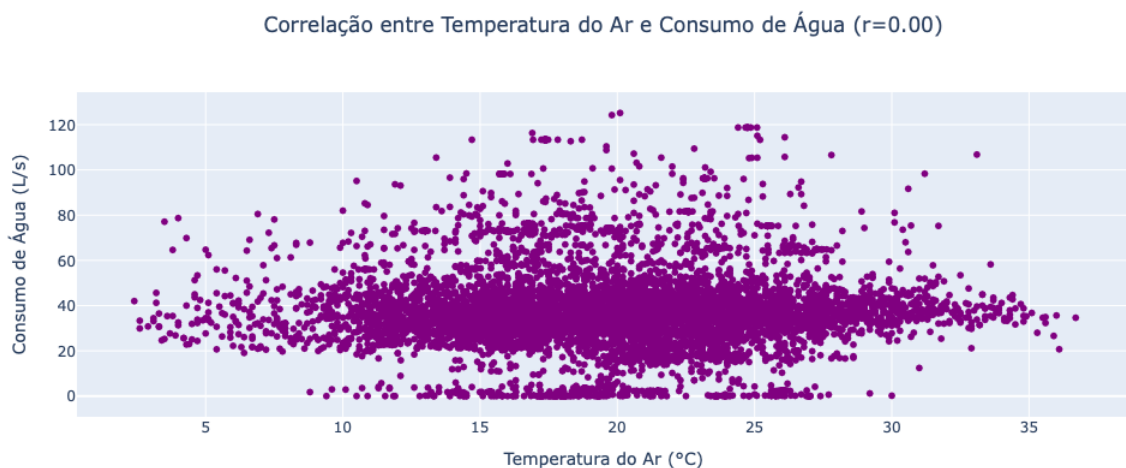
5.5. Pergunta 5

A quinta pergunta: "Para um determinado dia, horário e nível do reservatório, em quanto tempo ele se esvaziará caso falte energia?" foi respondida simulando o esvaziamento do reservatório. Utilizamos previsões da vazão de saída para as próximas 24 horas, tanto com quanto sem dados meteorológicos, e calculamos o tempo necessário até que o reservatório se esvaziasse, considerando a taxa de saída de água e o nível inicial do reservatório. Seguem os resultados:



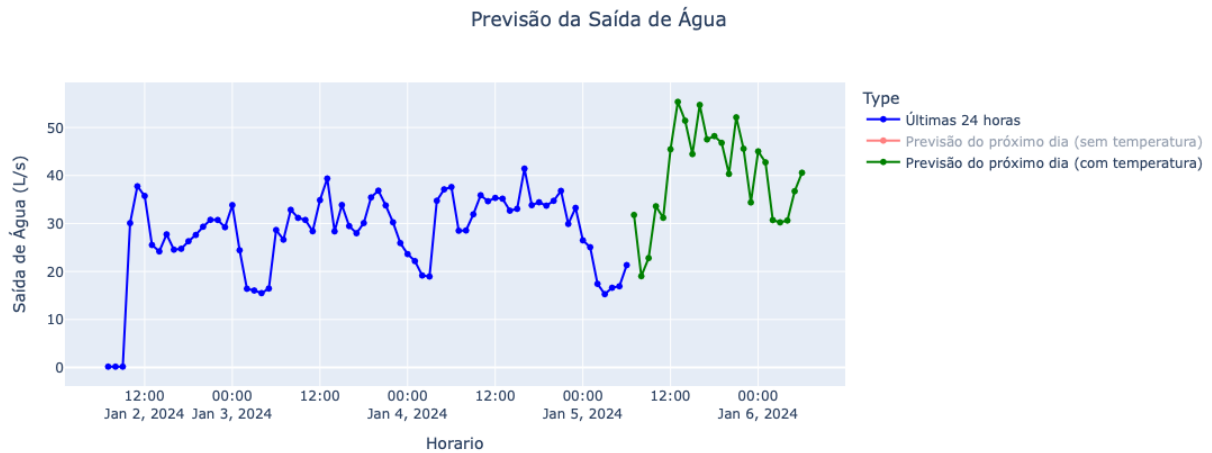
5.6. Pergunta 6

A sexta pergunta: "Há correlação entre a temperatura e o consumo de água?" foi respondida através da análise gráfica. Foram plotados gráficos de correlação entre a vazão de saída e variáveis climáticas, como temperatura, umidade e precipitação. Esses gráficos permitiram visualizar e analisar a relação entre as condições climáticas e o consumo de água, facilitando a identificação de possíveis correlações. Através da análise gráfica, não foi encontrada nenhuma correlação. Seguem os resultados:



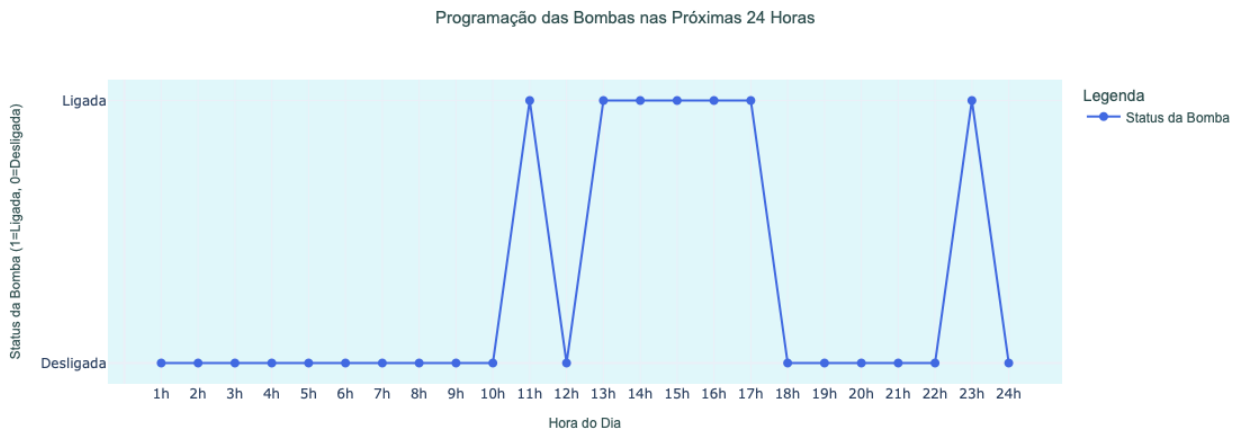
5.7. Pergunta 7

A sétima pergunta: "Qual a previsão da vazão de saída para as próximas 24h, considerando dados meteorológicos?" foi respondida com a criação de outros 24 modelos, incorporando dados meteorológicos. Esses modelos também previam a diferença da vazão de saída em relação à hora anterior, permitindo uma análise mais precisa ao considerar variáveis climáticas. Seguem os resultados:



5.8. Pergunta 8

A oitava pergunta: "Para um determinado dia, horário, nível do reservatório e previsão da vazão de saída para as próximas 24h, qual deveria ser a programação horária de cada bomba, de modo a minimizar o uso delas durante o horário de ponta, enquanto mantém o nível do reservatório entre 20 e 95%?" foi respondida utilizando um algoritmo genético para encontrar a melhor programação das bombas. Esse algoritmo considerou a previsão da vazão de saída, o nível do reservatório e penalidades por uso durante horários de ponta para otimizar a programação das bombas de forma eficiente. Seguem os resultados:



O algoritmo genético desenvolvido segue as seguintes configurações:

- **Função de Recompensa:** Avalia a programação das bombas, como a manutenção do nível do reservatório entre 20% e 95% de sua capacidade, a penalização pelo uso das bombas durante o horário de ponta (18h às 21h), e recompensas pelo uso durante horas não de pico (0h às 5h). A função também considera a taxa de esgotamento do reservatório para adicionar ou subtrair pontos à recompensa total.

- **Inicialização da População:** Cria uma população inicial de possíveis programações das bombas (cromossomos), onde cada programação é uma sequência de 24 horas, com cada hora podendo ter a bomba ligada (1) ou desligada (0).
- **Seleção de Pais:** Seleciona duas programações da população com base em suas recompensas, utilizando um método que quanto maior a recompensa, maior a probabilidade de ser selecionado.

6. Deploy da Solução Final

Nossa solução final é um dashboard interativo desenvolvido com Plotly e Flask, que permite visualizar e analisar dados de consumo de água e desempenho dos modelos. O dashboard serve para responder às perguntas do projeto, fornecendo gráficos e previsões interativas.

Funcionalidades do Dashboard

- **Visualização de Gráficos Estáticos:** Inclui gráficos para perguntas como a curva típica da vazão de saída (pergunta 2), horário de uso das bombas (pergunta 3) e a correlação entre variáveis climáticas e consumo de água (pergunta 6).
- **Previsões e Simulações:** Permite ao usuário inserir uma data e hora específicas para gerar previsões da vazão de saída nas próximas 24 horas (perguntas 4 e 7), simular o esvaziamento do reservatório (pergunta 5) e otimizar a programação das bombas (pergunta 8).
- **Avaliação de Modelos:** Exibe gráficos de comparação de desempenho dos modelos com e sem dados meteorológicos, ajudando a entender a eficácia de cada abordagem.

Utilidade do Dashboard

O dashboard é uma ferramenta essencial para gerentes de operações e engenheiros, permitindo monitorar e otimizar o uso de recursos hídricos em tempo real. Com ele, é possível prever o comportamento do sistema de água, identificar correlações importantes e tomar decisões informadas para garantir a eficiência e sustentabilidade das operações.

Prints do Dashboard:

Bem-vindo ao Dashboard de Vazão de Água

Gráficos Estáticos

Explore várias visualizações estáticas que ilustram os padrões típicos de vazão de água em diferentes períodos, incluindo tendências diárias, semanais e sazonais.

Gráficos de Previsão

Veja previsões de vazão de água com base em diferentes modelos. Esses gráficos ajudam a prever tendências futuras e tomar decisões informadas.

Desempenho do Modelo

Analise o desempenho de diferentes modelos usados para prever a vazão de água. Compare precisão, eficiência e outras métricas.

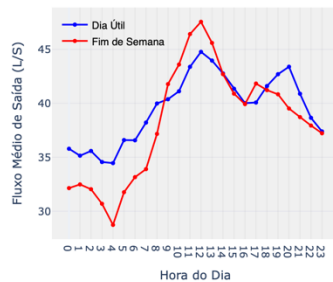


© 2024 Elipse Software & PowerSysLab. Todos os direitos reservados.

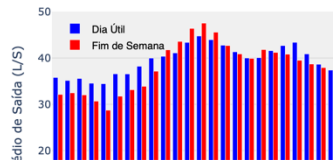
Gráficos Estáticos

Qual a curva típica da saída do reservatório ao longo de 24h durante os dias úteis? E nos finais de semana?

Comparação do Fluxo Médio de Saída por Hora

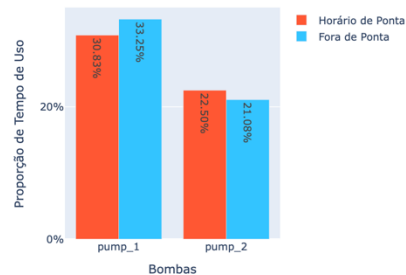


Comparação do Fluxo Médio de Saída por Hora

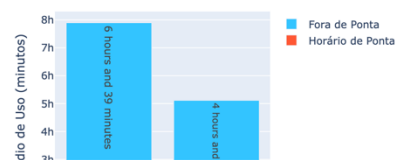


Qual o tempo de uso das bombas em horário de ponta e fora de ponta?

Proporção de Tempo de Uso das Bombas



Tempo Médio de Uso das Bombas



Gráficos de Previsão

Year

2024

Month

1

Day

12

Hour

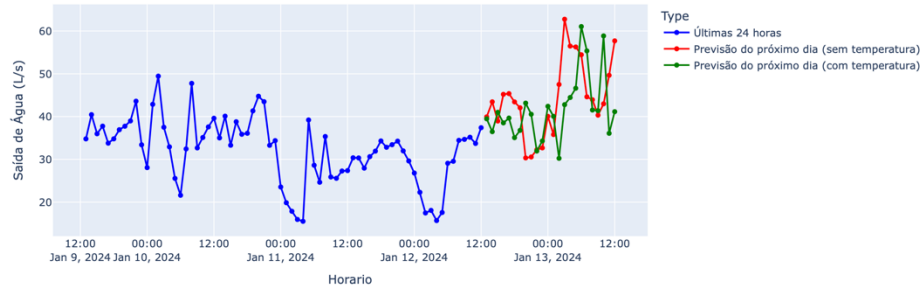
12

Minute

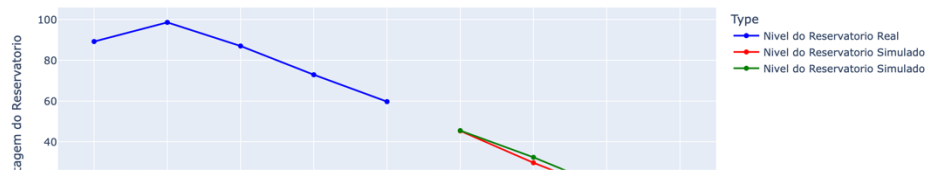
30

Forecast

Previsão da Saída de Água

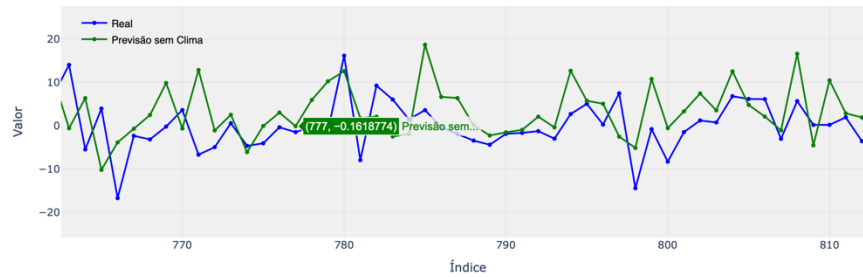


Simulação de Esvaziamento do Reservatório em Queda de Energia

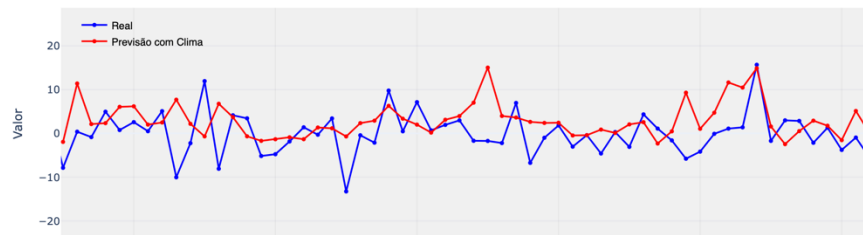


Desempenho dos Modelos

Comparação da mudança na saída de água com Previsão sem Clima



Comparação da mudança na saída de água com Previsão com Clima



7. Conclusão

O projeto utilizou dados coletados pelo sistema SCADA para responder a diversas perguntas críticas sobre a gestão e operação de sistemas de saneamento básico. Através da criação e treinamento de modelos de previsão, análises de correlação e simulações, foi possível obter insights valiosos sobre o comportamento do sistema de água. A solução final, um dashboard interativo, facilita a visualização desses dados e previsões, permitindo uma gestão mais eficiente e informada dos recursos hídricos. Essa abordagem integrativa e baseada em dados melhora a tomada de decisões e a sustentabilidade operacional.