

**UNIVERSIDADE DO ESTADO DO AMAZONAS
ESCOLA SUPERIOR DE TECNOLOGIA
CURSO DE ENGENHARIA DE PRODUÇÃO**

MARCOS VINÍCIUS DUTRA SOARES

**MODELO PREDITIVO DE CONSUMO INDUSTRIAL: aplicação de aprendizado de
máquina na previsão de consumo de um insumo em uma linha de produção
localizada no Polo Industrial de Manaus**

**MANAUS
2026**

MARCOS VINÍCIUS DUTRA SOARES

MODELO PREDITIVO DE CONSUMO INDUSTRIAL: aplicação de aprendizado de máquina na previsão de consumo de um insumo em uma linha de produção localizada no Polo Industrial de Manaus

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia de Produção da Escola Superior de Tecnologia da Universidade do Estado do Amazonas, como parte dos requisitos para a obtenção do grau de Bacharel em Engenharia de Produção.

Orientador: Prof. Dr. Alessandro de Souza Bezerra, PhD em Ciências.

MANAUS
2026

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).
Sistema Integrado de Bibliotecas da Universidade do Estado do Amazonas.

S676m	Soares, Marcos MODELO PREDITIVO DE CONSUMO INDUSTRIAL : aplicação de aprendizado de máquina na previsão de consumo de um insumo em uma linha de produção localizada no Polo Industrial de Manaus / Marcos Soares. Manaus : [s.n], 2026. 102 f.: color.; 21.0 cm. TCC - Graduação em Engenharia de Produção- Universidade do Estado do Amazonas, Manaus, 2026. Inclui Bibliografia. Orientador: Alessandro de Souza Bezerra. 1. estoque. 2. ressuprimento. 3. previsão. 4. aprendizado de máquina. 5. XGBoost. I. Alessandro de Souza Bezerra (Orient.) II. Universidade do Estado do Amazonas. III. Título CDU(1997)658.5
-------	--

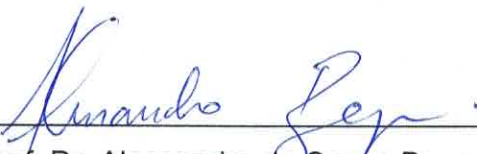
MARCOS VINICIUS DUTRA SOARES

**MODELO PREDITIVO DE CONSUMO INDUSTRIAL: aplicação de
aprendizado de máquina na previsão de consumo de um insumo em uma
linha de produção localizada no Polo Industrial de Manaus**

Trabalho apresentado ao curso de Engenharia de Produção da
Universidade do Estado do Amazonas, como requisito parcial para a obtenção
do grau de Bacharel em Engenharia de Produção.

Data de aprovação: Manaus (AM), 22 de junho de 2026.

Banca examinadora:



Prof. Dr. Alessandro de Souza Bezerra
Universidade do Estado do Amazonas



Profa. Dra. Kátia Cilene Neles da Silva
Universidade do Estado do Amazonas



Prof. Dr. Marcio Palheta Piedade
Universidade do Estado do Amazonas

DEDICATÓRIA

À minha mãe, Vera Lucia, pelo amor incondicional, pelo cuidado constante e por tantas vezes ter feito o possível para que eu pudesse seguir em busca dos meus sonhos. Sua força, sua fé e suas palavras de incentivo sempre me guiaram nos momentos mais difíceis. Com ela, aprendi que, por mais dura que a vida seja, somos capazes de continuar, pois Deus está conosco e nos fortalece diante dos desafios.

Ao meu pai, Eduardo Laercio, que já não está presente fisicamente, dedico esta conquista com profunda gratidão e saudade. Agradeço por nunca ter deixado faltar o pão de cada dia, por ter se desdobrado em diversos trabalhos e por ter feito tantos sacrifícios para que nada nos faltasse. Sou grato pelo homem que foi e pelo exemplo que deixou.

Dedico também este trabalho à minha família, especialmente ao meu irmão Lucas Dutra, que também me motiva a buscar os melhores caminhos em vida. Que esta conquista sirva como exemplo de que os estudos podem nos levar a lugares extraordinários. Aos meus primos, tios e demais familiares, agradeço por sempre estarem presentes e por serem uma base tão importante em minha trajetória.

Aos colegas de faculdade, que compartilharam comigo desafios, aprendizados e amadurecimentos, deixo minha dedicação e gratidão. O companheirismo de vocês tornou esta caminhada mais leve e significativa.

Aos professores que fizeram parte da minha formação, não apenas na universidade, mas em todas as etapas da minha vida, deixo meu reconhecimento. Cada ensinamento contribuiu, de alguma forma, para a construção do profissional e da pessoa que sou hoje.

Por fim, dedico este trabalho, com profundo amor e saudade, à memória da minha avó. Ela foi a base da minha vida e de toda a nossa família. Muito do que somos nasceu do cuidado, da força e do amor que ela nos ofereceu. Se hoje chegamos até aqui, é porque antes dela vieram o exemplo, a entrega e o amor que nos sustentaram.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me permitir concluir esta etapa tão importante da minha vida, e à minha mãe do céu, Maria, que me cobriu com seu manto e me protegeu em cada segundo desta caminhada.

Aos meus pais, Vera Lucia e Eduardo Laercio, agradeço por sempre investirem em mim, mesmo quando isso significou abrir mão de algo por amor. Obrigado por todos os sacrifícios, por cada incentivo e por me mostrarem o caminho que eu deveria seguir.

Ao meu irmão, Lucas Dutra, agradeço por acreditar em mim e por despertar em mim o desejo de ser o melhor referencial possível. Sua presença também foi motivação para que eu buscasse seguir em frente e alcançar esta conquista.

Aos meus familiares, agradeço por sempre estarem presentes na vida da minha família, sendo exemplos de união, força e orientação. Vocês foram parte essencial da base que me ensinou os caminhos que eu deveria trilhar.

À minha namorada, Grazyela Azevedo, agradeço pelo amor, carinho, apoio, motivações e por me ouvir nos momentos de desabafo. Obrigado por sempre me lembrar dos motivos pelos quais eu luto e por tornar os dias mais leves. “Perto de você eu sou muito mais eu nada não é tão vulgar, quanto parece sem você”.

Aos amigos que estiveram comigo durante a graduação, agradeço pelas risadas, lágrimas, caronas, lanches, incentivos e por todo companheirismo ao longo dessa jornada. Obrigado, Thiago Mesquita, Katarina Jeovana, Rodrigo Moraes, Hadassa Feitosa, João Pessoa e Pedro Luka. A caminhada foi muito mais leve porque vocês fizeram parte dela.

Aos meus amigos de vida, agradeço pelas conversas, pelas brincadeiras e pelas gameplays nos dias mais exaustivos e frustrantes. Sem esses momentos, certamente a vida teria sido muito mais pesada. Obrigado, Caio Guilherme, Gabriel Henrique, José Moreira, Natanael e Thiago Lira. Espero que a vida sorria para vocês sempre.

Ao meu orientador, professor Alessandro Bezerra, e à professora Rejane Gomes, agradeço por toda paciência, orientação, ensinamentos e contribuições ao longo desta trajetória acadêmica. Cada conselho e cada correção foram importantes para o meu crescimento.

Os homens esqueceram essa verdade”, disse a raposa. “Mas tu não a deves esquecer. Tu te tornas eternamente responsável por aquilo que cativas. Foi o tempo que dedicaste à tua rosa que a fez tão importante. — O Pequeno Príncipe, Antoine de Saint-Exupéry

RESUMO

Este trabalho aborda o desenvolvimento de uma solução preditiva aplicada à previsão de consumo e ressuprimento de fita de aço em kg em uma linha de fabricação de lâminas de barbear localizada no Polo Industrial de Manaus. A pesquisa teve como objetivo propor uma abordagem baseada em aprendizado de máquina para apoiar a análise do comportamento de abastecimento, a estimativa do intervalo até o próximo ressuprimento e o fortalecimento da tomada de decisão no ambiente fabril. Metodologicamente, a investigação foi caracterizada como pesquisa aplicada, de abordagem quantitativa, desenvolvida por meio de estudo de caso. Foram utilizados dados históricos reais de abastecimento e consumo da fita de aço, posteriormente tratados, organizados e complementados por meio da simulação de *Monte Carlo*, empregada como estratégia metodológica para ampliação da base e representação de cenários compatíveis com o comportamento operacional observado. A modelagem preditiva foi estruturada com o algoritmo *Extreme Gradient Boosting (XGBoost)*, considerando variáveis temporais e históricas, como registros anteriores, *lags* e médias móveis, com a variável-alvo definida como o intervalo em dias até o próximo abastecimento. Os resultados indicaram que a simulação de *Monte Carlo* contribuiu para complementar a base analítica, mantendo coerência estatística e plausibilidade operacional em relação aos dados reais. Já a modelagem com *XGBoost* apresentou desempenho moderado, com erro médio inferior a um dia, mas limitada capacidade explicativa global da variável-alvo, evidenciando a necessidade de bases históricas mais amplas e maior diversidade de variáveis operacionais para o aprimoramento das previsões. Conclui-se que a proposta apresenta potencial metodológico para apoiar o planejamento de ressuprimento da fita de aço, oferecendo uma estrutura inicial para uso de aprendizado de máquina na tomada de decisão industrial.

Palavras-chave: estoque; ressuprimento; previsão; aprendizado de máquina; *XGBoost*; *Monte Carlo*.

ABSTRACT

This study addresses the development of a predictive solution applied to the forecasting of steel strip consumption and replenishment in a razor blade manufacturing line located in the Manaus Industrial Pole, Brazil. The research aimed to propose a machine learning-based approach to support the analysis of replenishment behavior, the estimation of the time interval until the next replenishment event, and decision-making in the manufacturing environment. Methodologically, the investigation was characterized as applied research with a quantitative approach, developed through a case study. Real historical data related to steel strip supply and consumption were collected, processed, organized, and complemented through Monte Carlo simulation, employed as a methodological strategy to expand the database and generate scenarios compatible with the observed operational behavior. Predictive modeling was developed using the Extreme Gradient Boosting (XGBoost) algorithm, considering temporal and historical variables, such as previous records, lag features, and moving averages. The target variable was defined as the time interval, in days, until the next replenishment event. The results indicated that Monte Carlo simulation contributed to expanding the analytical database while maintaining statistical coherence and operational plausibility in relation to the real observations. The XGBoost model presented moderate predictive performance, with an average prediction error below one day, although its overall explanatory capacity remained limited. These findings suggest that broader historical databases and additional operational variables may contribute to improving predictive accuracy in future applications. It is concluded that the proposed approach presents methodological potential to support steel strip replenishment planning, providing an initial framework for the application of machine learning techniques in industrial decision-making processes.

Keywords: inventory; replenishment; forecasting; machine learning; XGBoost; Monte Carlo.

LISTA DE ILUSTRAÇÕES

Figura 1 – Funcionamento da base de dados de uma cadeia de suprimentos	19
Figura 2 – Fluxo do desenvolvimento do estudo de caso	43
Figura 3 – Estrutura inicial dos dados operacionais para modelagem preditiva	45
Figura 4 – Fluxo de coleta e consolidação da base real	49
Figura 5 – Base de dados utilizada	50
Figura 6 – Quantidade de fita de aço requisitada por mês/ano	51
Figura 7 – Trecho do código responsável pela parametrização da simulação a partir da base real	53
Figura 8 – Trecho do código responsável pela geração dos eventos simulados utilizando <i>Monte Carlo</i>	54
Figura 9 – Estruturas de saída	55
Figura 10 – Amostra da base simulada gerada por <i>Monte Carlo</i>	55
Figura 11 – Base de dados gerada	56
Figura 12 – Carregamento da base de dados no Python	57
Figura 13 – Conversão dos dados	58
Figura 14 – Trecho do código responsável pelo cálculo da variável-alvo	59
Figura 15 – Criação das <i>features</i> temporais	60
Figura 16 – Definição das variáveis numéricas e categóricas	62
Figura 17 – Divisão temporal da base em treino e teste	65
Figura 18 – Tratamento das variáveis numéricas e categóricas	66
Figura 19 – Configuração do modelo <i>XGBoost</i>	67
Figura 20 – Amostra de previsões geradas pelo modelo <i>XGBoost</i> padrão	68
Figura 21 – <i>Input</i> manual de quantidade e parâmetros operacionais	68
Figura 22 – Saída da simulação com <i>input</i> manual	69
Figura 23 – Fluxo de integração entre base de dados	72
Figura 24 – Fluxo geral da arquitetura proposta para aplicação da solução preditiva	74
Figura 25 – Simulação da estrutura das listas de dados no <i>SharePoint</i>	75
Figura 26 – Tela do Microsoft Azure	77
Figura 27 – Protótipo da tela inicial e tela de registro de abastecimento no <i>Power Apps</i>	78
Figura 28 – Protótipo tela de registro de abastecimento no <i>Power Apps</i>	79
Figura 29 – Estrutura da base real e da base expandida por simulação de <i>Monte Carlo</i>	81
Figura 30 – Comportamento temporal do consumo real e simulado de fitas de aço	83
Figura 31 – Gráfico de aderência entre valores observados e previstos na configuração preliminar do <i>XGBoost</i>	86
Figura 32 – <i>Output</i> dos indicadores da configuração sem vazamentos de dados do <i>XGBoost</i>	88
Figura 33 – Gráfico de aderência entre valores observados e previstos sem vazamentos de dados do <i>XGBoost</i>	89

LISTA DE QUADROS

Quadro 1 – Estrutura dos dados reais coletados	46
Quadro 2 – Classificação das variáveis utilizadas na modelagem	48
Quadro 3 – Justificativas das variáveis	60
Quadro 4 – Variáveis usadas no modelo	63
Quadro 5 – Estrutura das listas do <i>SharePoint</i> utilizadas na solução proposta	70
Quadro 6 – Campos utilizados para visualização das previsões no <i>Power Apps</i>	71
Quadro 7 – Estrutura proposta das listas no <i>SharePoint</i>	75
Tabela 1 – Principais indicadores e seus resultados	82
Tabela 2 – Resultados dos indicadores da configuração preliminar do <i>XGBoost</i> ...	85
Tabela 3 – Resultados dos indicadores da configuração sem vazamentos de dados do <i>XGBoost</i>	87
Tabela 4 – Síntese dos principais resultados obtidos	90

LISTA DE TABELAS

Tabela 1 – Principais indicadores e seus resultados	82
Tabela 2 – Resultados dos indicadores da configuração preliminar do <i>XGBoost</i> ...	85
Tabela 3 – Resultados dos indicadores da configuração sem vazamentos de dados do <i>XGBoost</i>	87
Tabela 4 – Síntese dos principais resultados obtidos	90

SUMÁRIO

1	INTRODUÇÃO	13
1.1	CONTEXTUALIZAÇÃO, PROBLEMA E JUSTIFICATIVA DA PESQUISA ..	13
1.2	HIPÓTESES	14
1.3	OBJETIVOS	14
1.3.1	Objetivo geral	14
1.3.2	Objetivos específicos	15
1.4	ESTRUTURAÇÃO DO TRABALHO	15
2	REFERENCIAL TEÓRICO	17
2.1	CADEIA DE SUPRIMENTOS	17
2.1.1	Previsão de insumos dentro da cadeia de suprimentos	19
2.1.2	Previsão de insumos dentro da indústria	20
2.2	GERAÇÃO DE DADOS SINTÉTICOS	21
2.2.1	A simulação de <i>Monte Carlo</i>	23
2.2.2	Aplicações industriais de dados sintéticos	24
2.3	APRENDIZADO DE MÁQUINA	26
2.3.1	Aplicação de <i>machine learning</i> dentro do ambiente fabril	27
2.3.2	Modelos de previsão	28
2.3.2.1	Regressão linear simples	28
2.3.2.2	Regressão linear múltipla	30
2.3.2.3	<i>Extreme Gradient Boosting</i> — <i>XGBoost</i>	31
2.3.2.4	Comparativo entre modelos tradicionais e <i>machine learning</i>	33
2.4	VALIDAÇÃO ESTATÍSTICA DE MODELOS PREDITIVOS	34
2.4.1	Erro Absoluto Médio — MAE	34
2.4.2	Raiz do Erro Quadrático Médio — RMSE	36
2.4.3	Vazamento de dados	37
2.4.4	Separação de treino-teste	38
2.5	<i>MICROSOFT POWER APPS</i> NA TRANSFORMAÇÃO DIGITAL.....	38
2.5.1	O ambiente de desenvolvimento do <i>software Power Apps</i>	39
2.5.2	<i>Power Automate</i> e Azure	40
2.5.3	Integração digital industrial e sistemas de apoio à decisão	41
3	PROCEDIMENTOS METODOLÓGICOS	42
3.1	NATUREZA DA PESQUISA	42

4	DESENVOLVIMENTO DO ESTUDO DE CASO	43
4.1	DESCRIÇÃO DO PROCESSO PRODUTIVO ESTUDADO	44
4.1.1	Insumos e recursos considerados	45
4.1.2	Coleta e consolidação dos dados reais	46
4.1.3	Estruturação das variáveis operacionais e preditivas	47
4.2	COMPILAÇÃO E ESTRUTURAÇÃO DOS DADOS	49
4.2.1	Tratamento e preparação dos dados	49
4.2.2	Base real para expansão	50
4.2.3	Simulação da base de dados da fita de aço	51
4.2.4	Preparação da base de dados para o modelo <i>XGBoost</i>	56
4.2.5	Treinamento da configuração preliminar do <i>XGBoost</i>	59
4.2.6	Treinamento do modelo <i>XGBoost</i> sem vazamento de dados	63
4.2.7	Integração com <i>Power Apps</i>	69
4.3	ESTRUTURAÇÃO DO SISTEMA E DO FLUXO DE PROCESSOS	71
4.3.1	Arquitetura do ecossistema para aplicação	72
4.3.2	Estrutura das bases no <i>SharePoint</i>	73
4.3.3	Fluxo de alimentação e atualização dos dados	76
4.3.4	Execução do modelo <i>XGBoost</i> no <i>Azure</i>	77
4.3.5	Visualização dos resultados no <i>Power Apps</i>	78
5	ANÁLISE E DISCUSSÃO DOS RESULTADOS	80
5.1	RESULTADOS DA SIMULAÇÃO DE MONTE CARLO	81
5.2	RESULTADOS DA CONFIGURAÇÃO PRELIMINAR DO <i>XGBOOST</i>	84
5.3	RESULTADOS DO MODELO <i>XGBOOST</i> SEM VAZAMENTO DE DADOS	86
5.4	SÍNTESE DOS RESULTADOS DOS INDICADORES	90
6	CONSIDERAÇÕES FINAIS	92
	REFERÊNCIAS	94

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO, PROBLEMA E JUSTIFICATIVA DA PESQUISA

A crescente competitividade no setor industrial, especialmente no Polo Industrial de Manaus, evidencia limitações nos modelos tradicionais de gestão de recursos produtivos, particularmente no controle de materiais de alto impacto operacional e financeiro. Nesse contexto, a ausência de ferramentas analíticas robustas pode comprometer a previsibilidade do consumo, a eficiência operacional e o planejamento de abastecimento, reforçando a necessidade de adoção de abordagens baseadas em ciência de dados e aprendizado de máquina, alinhadas aos princípios da Indústria 4.0.

Diante desse cenário, esta pesquisa tem como foco o desenvolvimento e a aplicação de modelos preditivos voltados à previsão de consumo e ressuprimento da fita de aço, em kg, em uma linha de produção de lâminas de barbear. A escolha desse recurso justifica-se por sua relevância no processo produtivo e por representar um dos materiais de maior impacto econômico na operação analisada. Assim, a delimitação do estudo permite concentrar a análise em um recurso específico, favorecendo maior controle metodológico, melhor interpretação dos dados e maior aderência à realidade operacional observada.

Para apoiar essa proposta, foram utilizados dados históricos reais de consumo e abastecimento, posteriormente complementados por meio da simulação de *Monte Carlo*, com o objetivo de representar cenários compatíveis com o comportamento operacional observado. A partir dessa base estruturada, a modelagem preditiva foi direcionada ao uso do algoritmo *XGBoost*, buscando estimar o intervalo em dias até o próximo abastecimento da fita de aço e apoiar decisões relacionadas ao planejamento de ressuprimento. A delimitação do trabalho concentrou-se em uma unidade fabril do Polo Industrial de Manaus, considerando registros reais e simulados compatíveis com a operação analisada.

Nesse sentido, a problemática da pesquisa foi consolidada na seguinte questão: em que medida o uso de técnicas de aprendizado de máquina pode contribuir para apoiar a previsão de consumo e o planejamento de reposição da fita de aço em kg em uma linha de produção de lâminas de barbear no Polo Industrial de Manaus,

considerando a redução de incertezas operacionais, a melhoria da previsibilidade e o controle de custos associados ao abastecimento desse recurso?

1.2 HIPÓTESES

A partir do problema de pesquisa proposto, formulam-se as seguintes hipóteses, que buscam estabelecer relações entre o uso de técnicas de aprendizado de máquina e a otimização da gestão da fita de aço em kg no ambiente industrial:

I) A implementação de um sistema de previsão baseado em técnicas de aprendizado de máquina contribui para antecipar necessidades de ressuprimento da fita de aço, reduzindo incertezas operacionais e apoiando a continuidade da linha de produção.

II) O uso de modelos estatísticos e algoritmos preditivos contribui para melhorar o planejamento de compras e reposição da fita de aço, fornecendo informações que podem auxiliar na redução de decisões emergenciais, excesso ou escassez desse recurso.

III) A integração de um sistema de previsão inteligente à rotina produtiva melhora a disponibilidade de informações para a tomada de decisão gerencial, ao proporcionar dados mais estruturados sobre o comportamento de consumo e ressuprimento da fita de aço.

1.3 OBJETIVOS

1.3.1 Objetivo geral

O objetivo geral deste estudo consiste em desenvolver uma solução preditiva, com base nos dados históricos fornecidos pela empresa no período de janeiro de 2024 a março de 2026, para apoiar a previsão do comportamento de consumo e o planejamento de ressuprimento da fita de aço em kg em uma linha de produção de lâminas de barbear localizada no Polo Industrial de Manaus, por meio da aplicação de técnicas de aprendizado de máquina integradas a uma solução computacional. Além disso, busca-se estruturar uma abordagem capaz de transformar registros operacionais em informações analíticas úteis ao processo decisório, contribuindo para a redução de incertezas no abastecimento e para o fortalecimento do controle sobre o recurso analisado.

1.3.2 Objetivos específicos

- a) Analisar métodos de previsão aplicados a processos industriais, com foco na simulação de *Monte Carlo* como estratégia complementar de expansão da base de dados e no algoritmo *XGBoost* para apoio à previsão do ressuprimento da fita de aço;
- b) Coletar, tratar e organizar dados históricos de consumo e abastecimento da fita de aço em kg, referentes à linha de produção analisada;
- c) Desenvolver um protótipo de sistema/aplicativo para suporte à tomada de decisão, integrando a solução preditiva à plataforma *Microsoft Power Apps*;
- d) Avaliar o desempenho preditivo do modelo reestruturado por meio de métricas estatísticas de erro e da análise de aderência entre valores observados e previstos.

1.4 ESTRUTURAÇÃO DO TRABALHO

Este trabalho está estruturado em cinco seções, organizadas de forma sequencial para apresentar o desenvolvimento da pesquisa, desde a contextualização do problema até a análise dos resultados e as considerações finais. A primeira seção compreende os elementos introdutórios, apresentando a contextualização, o problema e a justificativa da pesquisa, além das hipóteses, do objetivo geral e dos objetivos específicos.

A segunda seção apresenta o referencial teórico que fundamenta o estudo, abordando conceitos relacionados à cadeia de suprimentos, previsão de consumo, gestão preditiva de estoques industriais, aprendizado de máquina, modelos de previsão, geração de dados sintéticos, simulação de *Monte Carlo*, validação estatística de modelos preditivos e integração digital por meio de ferramentas como *Power Apps*, *Power Automate* e *Azure*.

A terceira seção descreve os procedimentos metodológicos adotados, contemplando a natureza da pesquisa, a caracterização do processo produtivo estudado, a definição da fita de aço em kg como objeto de análise, a coleta e consolidação dos dados reais, a estruturação das variáveis, o tratamento dos dados, a expansão da base por meio da simulação de *Monte Carlo*, a modelagem preditiva com *XGBoost* e a estruturação do sistema integrado de apoio ao ressuprimento.

A quarta seção apresenta o desenvolvimento da pesquisa aplicada, demonstrando a organização da base real, a construção da base expandida por simulação de *Monte Carlo* e a aplicação do modelo *XGBoost* para previsão do intervalo em dias até o próximo abastecimento da fita de aço. Nessa etapa, a simulação é tratada como estratégia complementar para ampliação e representação de cenários, preservando a coerência com o comportamento observado na operação.

A quinta seção apresenta a análise e discussão dos resultados, contemplando a coerência da base simulada em relação à realidade operacional observada, os indicadores de desempenho do modelo preditivo reestruturado e as limitações identificadas na aplicação da solução proposta. Ao final, são retomados os objetivos e hipóteses da pesquisa, sintetizando os principais resultados alcançados, as limitações observadas e as possibilidades de continuidade em estudos futuros.

2 REFERENCIAL TEÓRICO

O referencial teórico deste trabalho está estruturado a partir dos principais fundamentos conceituais e metodológicos que sustentam a proposta desenvolvida. Inicialmente, são discutidos os conceitos relacionados à gestão da cadeia de suprimentos, com ênfase na previsão de consumo como elemento central para o alinhamento entre oferta e demanda em ambientes produtivos. Em seguida, são apresentados os modelos de previsão, abrangendo tanto abordagens estatísticas tradicionais quanto técnicas de aprendizado de máquina, incluindo algoritmos baseados em *boosting*.

Também são abordadas a geração de dados sintéticos e a simulação de *Monte Carlo* como estratégias complementares para ampliação da base de dados, representação de cenários e tratamento de incertezas. Além disso, o referencial contempla a validação estatística de modelos preditivos, destacando métricas de erro, separação entre dados de treinamento e teste e cuidados relacionados ao vazamento de dados, conhecido como *data leakage*. Ao final, são tratados os aspectos relacionados à transformação digital no contexto industrial, com destaque para o uso do *Microsoft Power Apps*, *Power Automate* e *Azure* no desenvolvimento da solução integrada.

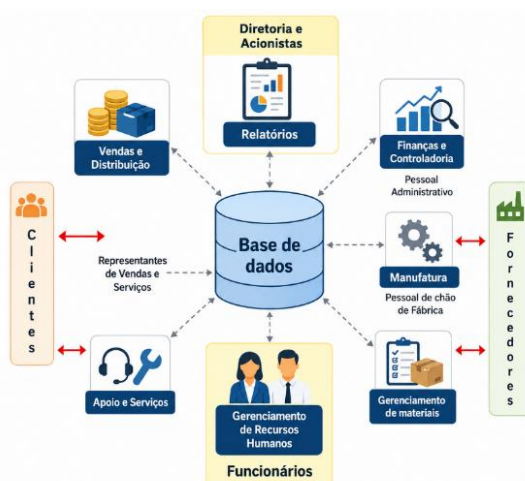
2.1 CADEIA DE SUPRIMENTOS

A gestão da cadeia de suprimentos (*Supply Chain Management – SCM*) constitui um campo multidisciplinar voltado à otimização do fluxo de materiais e informações ao longo dos processos produtivos. Sua eficácia depende do alinhamento entre demanda e capacidade produtiva, o que exige coordenação entre áreas como compras, produção, logística e *marketing*. Conforme destacam Melo e Alcântara (2011), essa integração enfrenta desafios contínuos, sobretudo pela dificuldade de sincronizar previsões de consumo com as variações de mercado e as limitações operacionais. Em uma perspectiva mais recente, Moraes e Moori (2024) apontam que o desempenho da cadeia de suprimentos passou a depender cada vez mais da aplicação de dados, aprendizado de máquina e integração entre os elos operacionais. Assim, a falta de alinhamento estratégico tende a gerar desperdícios, estoques excessivos e perda de competitividade no ambiente industrial.

Entre os fatores que comprometem a eficiência da cadeia está a volatilidade da demanda, fenômeno amplamente discutido por Oliva e Albano (2006) ao tratarem do chamado “efeito chicote”. Pequenas oscilações no consumo final podem provocar grandes distorções nos pedidos realizados pelos diferentes elos da cadeia, comprometendo a previsibilidade e elevando custos. Para reduzir tais distorções, torna-se essencial promover maior colaboração entre fornecedores e fabricantes, além de adotar modelos de previsão mais precisos e integrados. Nesse sentido, Schmidt e Schmidt (2023) destacam que a Logística 4.0 incorpora tecnologias digitais, como inteligência artificial, internet das coisas e *big data*, com o objetivo de ampliar a visibilidade, a rastreabilidade e a eficiência dos processos logísticos.

Outro desafio relevante está no equilíbrio entre objetivos conflitantes, como reduzir custos operacionais e, ao mesmo tempo, garantir altos níveis de serviço ao cliente. Conforme aponta Chopra (2003), lidar com esses *trade-offs* requer o emprego de modelagens quantitativas e analíticas que apoiem decisões em ambientes complexos e dinâmicos. De acordo com Pires (2009) e Fleury, Wanke e Figueiredo (2000), a integração tecnológica e colaborativa constitui um dos pilares para melhorar a eficiência das operações e fortalecer a competitividade das cadeias produtivas. De forma complementar, Cavalcante (2024) ressalta que a inteligência artificial vem sendo aplicada à cadeia de suprimentos como instrumento de apoio à eficiência, à sustentabilidade e à redução de incertezas nos processos decisórios. Essa integração entre dados, processos e tomada de decisão pode ser observada conceitualmente na figura abaixo, que apresenta o funcionamento da base de dados de uma cadeia de suprimentos.

Figura 1 - Funcionamento da base de dados de uma cadeia de suprimentos.



Fonte: Daválos e Lopez (2005)

2.1.1 Previsão de insumos dentro da cadeia de suprimentos

A previsão de consumo na cadeia de suprimentos é uma etapa essencial para garantir o equilíbrio entre oferta e demanda, evitando rupturas e otimizando o uso dos materiais disponíveis. Conforme aponta Ballou (2006), o planejamento eficiente deve considerar o volume necessário, o tempo de reposição e os fatores externos que influenciam a disponibilidade do item analisado, como sazonalidade e transporte. Quando a previsibilidade é insuficiente, surgem reações emergenciais que elevam custos e comprometem o fluxo produtivo. Dessa forma, o emprego de sistemas de previsão baseados em dados históricos e técnicas analíticas torna-se indispensável para reduzir a variabilidade e apoiar decisões estratégicas de longo prazo.

O estudo desenvolvido por Borges (2025) demonstra que grande parte das ineficiências observadas nas cadeias produtivas resulta da falta de integração entre o planejamento físico e o cronograma de suprimentos. Segundo o autor, compras emergenciais e estoques excessivos derivam, muitas vezes, de previsões imprecisas sobre as reais necessidades de produção. Essa desconexão entre planejamento e execução operacional é recorrente em diferentes setores, como a construção civil e a manufatura, nos quais o ritmo produtivo depende diretamente da disponibilidade adequada dos materiais críticos. Assim, aprimorar a previsão de demandas com apoio de ferramentas tecnológicas e analíticas é fundamental para fortalecer a coordenação e a resiliência da cadeia.

Além dos aspectos técnicos, a previsão de consumo enfrenta também desafios estruturais e logísticos. De acordo com Formoso *et al.* (2002), a ausência de um planejamento sistemático e de processos colaborativos entre as áreas de produção, compras e engenharia gera desperdícios significativos e desalinhamento entre materiais e prazos. Christopher (2010) observa que a previsibilidade depende não apenas de tecnologia, mas de integração organizacional, uma vez que a eficiência da cadeia está vinculada à comunicação e ao compartilhamento de informações. Nesse contexto, o uso de modelos preditivos deve estar associado a uma cultura colaborativa, em que diferentes setores contribuam com dados e percepções complementares.

Pesquisas recentes indicam que a incorporação de tecnologias digitais e algoritmos de aprendizado de máquina eleva a precisão das previsões e reduz incertezas operacionais. Borges (2025) e Lu *et al.* (2024) destacam que sistemas inteligentes de gestão de suprimentos, como sistemas *ERP* (Enterprise Resource

Planning), ou sistemas de Planejamento dos Recursos Empresariais, integrados a modelos preditivos, permitem identificar padrões de consumo e antecipar desvios logísticos. Essa integração fortalece o planejamento estratégico, otimiza estoques e incentiva o uso racional do item analisado, impactando positivamente a produtividade e a sustentabilidade. Dessa forma, o aprimoramento da previsão de consumo representa um avanço não apenas operacional, mas também estratégico para o aumento da competitividade industrial.

2.1.2 Previsão de insumos dentro da indústria

No ambiente industrial contemporâneo, especialmente em sistemas produtivos de alta complexidade operacional, como o Polo Industrial de Manaus, a gestão de estoques deixou de se limitar a práticas reativas baseadas apenas em históricos de consumo ou controles estáticos de reposição. A crescente variabilidade da demanda, os riscos associados à ruptura de abastecimento e os custos operacionais decorrentes de estoques excessivos ou insuficientes exigem abordagens mais estruturadas, capazes de integrar previsibilidade, análise estatística e suporte tecnológico à tomada de decisão.

Diante desse cenário, a gestão preditiva de estoques surge como uma evolução estratégica dos sistemas tradicionais, incorporando modelos matemáticos, mineração de dados e algoritmos de aprendizado de máquina para antecipar demandas futuras, otimizar níveis de re suprimento e fortalecer a eficiência logística. Nesta pesquisa, aplicada à previsão de consumo da fita de aço em uma linha industrial de lâminas de barbear, essa abordagem torna-se relevante por transformar dados históricos de consumo, frequência operacional e *lead time* em informações analíticas capazes de reduzir incertezas, evitar falhas de abastecimento e aprimorar o planejamento produtivo.

Sob essa perspectiva, Lima et al. (2022), ao desenvolverem um panorama sistemático sobre a aplicação de mineração de dados e aprendizagem de máquina na gestão de estoques, demonstram que a incorporação dessas tecnologias representa um avanço substancial em relação aos modelos convencionais de planejamento. O estudo evidencia que técnicas analíticas modernas permitem identificar padrões ocultos de consumo, prever oscilações de demanda e apoiar decisões mais precisas sobre políticas de reposição, níveis de estoque e alocação de materiais. Além disso, os autores ressaltam que a integração entre sistemas de dados e algoritmos preditivos

contribui para maior resiliência operacional, redução de desperdícios e fortalecimento estratégico das cadeias produtivas, sobretudo em contextos industriais marcados por elevada complexidade logística.

Ao consolidar evidências sobre aplicações práticas dessas ferramentas, Lima et al. (2022) reforçam que a gestão preditiva não se limita à automação de controles, mas constitui uma transformação estrutural orientada por inteligência analítica. Essa perspectiva aproxima-se da proposta deste trabalho, no qual a utilização combinada de simulação de *Monte Carlo*, modelagem preditiva por *XGBoost* e integração digital busca estruturar uma abordagem de apoio ao planejamento de ressuprimento da fita de aço, sustentando decisões operacionais com base em dados.

Assim, a gestão preditiva de estoques industriais consolida-se como elemento relevante dentro da transformação digital da cadeia de suprimentos, conectando previsão analítica, eficiência operacional e sustentabilidade produtiva. Sua aplicação no ambiente fabril não se limita à automação de controles, mas representa uma mudança estrutural na forma como organizações planejam o abastecimento, antecipam riscos e constroem vantagem competitiva em mercados industriais cada vez mais dinâmicos.

2.2 GERAÇÃO DE DADOS SINTÉTICOS

A geração de dados sintéticos tem se tornado uma alternativa relevante em contextos nos quais a coleta de dados reais é limitada, custosa ou insuficiente para representar adequadamente todos os cenários possíveis de operação. Em ambientes industriais, essa necessidade é ainda mais evidente, pois muitas vezes os registros disponíveis não contemplam situações extremas, variações sazonais ou eventos de baixa frequência, mas que podem impactar diretamente o desempenho produtivo. Segundo Nogueira et al. (2024), a principal motivação para o uso de dados sintéticos está justamente na dificuldade de reunir bases reais amplas, atualizadas e estatisticamente equilibradas. Assim, ao reproduzir características dos dados originais, essas amostras artificiais permitem ampliar o conjunto analisado e fortalecer a construção de modelos preditivos.

Essa lógica é observada no estudo de Nogueira et al. (2024), que apresenta a ferramenta *MalSynGen*, baseada em redes neurais generativas adversariais condicionais, conhecidas como *cGANs*, para geração de dados tabulares no contexto

de detecção de *malware Android*. Embora o estudo esteja voltado à área computacional, sua contribuição metodológica é importante por demonstrar como dados artificiais podem ser avaliados a partir de critérios de fidelidade e utilidade. Em outras palavras, não basta gerar novos registros; é necessário verificar se eles preservam propriedades estatísticas relevantes e se continuam úteis para o treinamento de modelos preditivos. Essa preocupação também se aplica ao ambiente industrial, onde dados sintéticos precisam representar comportamentos plausíveis do processo real.

De forma complementar, Schena *et al.* (2022) demonstram a aplicação de dados sintéticos em manutenção preditiva de sistemas de propulsão naval, aproximando os dados simulados das condições reais de operação. Os autores destacam que dados gerados por simuladores podem ser úteis, mas tendem a não representar completamente ruídos de medição, incertezas operacionais e desbalanceamentos comuns em ambientes industriais. Por esse motivo, foram incorporadas perturbações gaussianas e variações controladas para simular falhas e erros de sensores, resultando em bases mais aderentes ao comportamento observado em campo. Esse tipo de ajuste evidencia que a qualidade dos dados sintéticos depende não apenas do volume gerado, mas da capacidade de representar variações realistas do sistema analisado.

Nesse sentido, a validade dos dados sintéticos está diretamente relacionada ao equilíbrio entre fidelidade e variabilidade. Caso os registros gerados sejam muito próximos dos dados originais, pouco acrescentam ao modelo; por outro lado, se forem excessivamente aleatórios, podem distorcer a realidade operacional. Conforme Nogueira *et al.* (2024), a utilidade desses dados depende de sua capacidade de preservar relações estatísticas importantes sem se limitar a simples cópias da base real. Por isso, técnicas de aprendizado de máquina, simulação e métodos estatísticos devem ser acompanhadas de validações que confirmem a coerência dos dados gerados em relação ao fenômeno estudado.

No contexto desta pesquisa, a geração de dados sintéticos se conecta diretamente à simulação de *Monte Carlo*, utilizada para complementar a base histórica de consumo e abastecimento da fita de aço e representar cenários compatíveis com o comportamento operacional observado. A partir dos dados reais coletados no processo produtivo, a simulação permitiu gerar novas ocorrências respeitando limites operacionais, padrões de consumo e variabilidade identificada na fábrica. Dessa

forma, os dados sintéticos não substituem os registros reais, mas atuam como complemento analítico para ampliar a representação dos cenários estudados, fornecendo suporte à etapa de modelagem preditiva e à análise do planejamento de ressuprimento da fita de aço.

2.2.1 A simulação de *Monte Carlo*

A simulação de Monte Carlo é uma técnica estatística utilizada para representar cenários de incerteza por meio da geração repetida de valores aleatórios (Rubinstein; Kroese, 2017). Seu princípio central consiste em simular múltiplas possibilidades para uma variável de interesse e, a partir dos resultados obtidos, estimar seu comportamento esperado. Diferentemente de modelos determinísticos, que produzem uma única resposta fixa, esse método permite analisar uma faixa de resultados possíveis, sendo especialmente útil em problemas industriais sujeitos a variabilidade, como consumo de insumos, atrasos de abastecimento, falhas operacionais e oscilações de demanda. Segundo Paula (2014) e Morettin e Singer (2022), a repetição de cenários probabilísticos permite avaliar riscos e apoiar decisões em situações nas quais o comportamento futuro não pode ser previsto com absoluta certeza.

Do ponto de vista matemático, a estimativa obtida por *Monte Carlo* pode ser representada pela média dos resultados simulados, conforme a equação abaixo:

$$E[X] \approx (1/N) \times \sum X_i$$

Em que $E[X]$ representa o valor esperado da variável analisada, N corresponde ao número total de simulações realizadas e X_i representa cada resultado gerado na i -ésima simulação. Quanto maior o número de repetições, maior tende a ser a estabilidade da estimativa, pois os efeitos aleatórios se compensam ao longo do processo. Além disso, a dispersão dos resultados pode ser avaliada por medidas como variância e desvio padrão, permitindo compreender não apenas o valor médio esperado, mas também o grau de incerteza associado ao fenômeno simulado.

Em aplicações industriais, a simulação de *Monte Carlo* pode ser associada à simulação de eventos discretos, na qual o sistema evolui por meio de acontecimentos específicos ao longo do tempo. Eventos como consumo de materiais, chegada de novos lotes, falhas de máquinas ou reposições de estoque alteram o estado do sistema em momentos definidos. Esse processo é acompanhado por um “relógio” de

simulação, que avança de evento em evento e registra variáveis como nível de estoque, tempo de operação, frequência de abastecimento e risco de ruptura, aproximando a representação computacional da dinâmica real do processo produtivo.

A definição das distribuições probabilísticas é uma etapa essencial nesse tipo de simulação. Em bases industriais com número limitado de observações, a distribuição triangular costuma ser útil por exigir apenas três parâmetros principais: valor mínimo, valor mais provável e valor máximo (Paula, 2014). Sua estrutura pode ser representada conceitualmente por:

$$X \sim \text{Triangular}(a, c, b)$$

Em que a representa o menor valor observado, c corresponde ao valor mais provável e b representa o maior valor possível ou observado. Essa distribuição é adequada quando se deseja representar cenários otimistas, normais e críticos, preservando limites operacionais conhecidos. Assim, ela permite gerar dados sintéticos coerentes com a realidade do processo, evitando simulações totalmente aleatórias ou desconectadas dos padrões produtivos.

Para este estudo, a simulação de *Monte Carlo* foi utilizada como estratégia complementar para ampliar a representação da base histórica de consumo e abastecimento da fita de aço, a partir dos dados reais coletados no processo produtivo. Como a pesquisa passou a considerar registros reais no período de janeiro de 2024 a março de 2026, a técnica permitiu gerar cenários adicionais compatíveis com o comportamento observado na fábrica, especialmente em relação aos intervalos entre abastecimentos, ao *lead time* e aos padrões de consumo.

Dessa forma, os dados simulados não substituem a base real, mas funcionam como complemento analítico para ampliar a representação dos cenários estudados. Essa estratégia contribuiu para estruturar a etapa de modelagem preditiva e apoiar a análise do planejamento de ressuprimento da fita de aço, respeitando as limitações inerentes ao uso de dados simulados.

2.2.2 Aplicações industriais de dados sintéticos

As aplicações industriais de dados sintéticos têm se expandido em função da necessidade de testar modelos preditivos, validar sistemas inteligentes e representar cenários operacionais que nem sempre aparecem com frequência nos dados reais. Em ambientes produtivos, situações como falhas críticas, atrasos de abastecimento,

variações extremas de consumo ou rupturas de estoque podem ser raras, mas possuem impacto elevado sobre a operação. Sob essa perspectiva, os dados sintéticos permitem criar cenários controlados e estatisticamente coerentes, possibilitando o treinamento, teste e validação de modelos antes de sua aplicação definitiva no processo real.

Um estudo brasileiro recente que exemplifica essa aplicação foi desenvolvido por Braz (2025), em um projeto voltado à integração entre inteligência artificial e Internet das Coisas, do inglês *Internet of Things (IoT)*, para manutenção preditiva no contexto da Indústria 4.0. O trabalho teve como objetivo analisar como essas tecnologias poderiam contribuir para o monitoramento de ativos, identificação antecipada de falhas e automação de processos de manutenção. Para isso, foram utilizadas bases de dados sintéticos, simulando diferentes condições de sensores industriais, como temperatura, pressão e vibração, permitindo validar a lógica do sistema em cenários normais, intermediários e críticos.

No desenvolvimento do estudo de Braz (2025), os dados sintéticos foram organizados em três cenários de teste: funcionamento normal, mistura de estados e falha múltipla. Esses registros foram carregados em um sistema ERPNext e processados por uma lógica de inteligência artificial baseada em regras, responsável por classificar as leituras como “Ok” ou “Alerta” e gerar registros automáticos de manutenção preditiva. Essa estrutura demonstra que os dados artificiais podem ser úteis não apenas para ampliar bases de análise, mas também para testar o comportamento de sistemas inteligentes em diferentes condições operacionais antes da implantação em ambiente real.

Para este estudo, essa aplicação se aproxima da proposta metodológica ao evidenciar que dados sintéticos podem funcionar como apoio à construção de soluções preditivas em ambientes industriais com limitação de dados históricos. Assim como no estudo de Braz (2025), em que cenários simulados foram utilizados para validar um sistema de manutenção, neste trabalho os dados sintéticos gerados por simulação de *Monte Carlo* são empregados para complementar a base histórica de consumo e abastecimento da fita de aço, representando variações futuras compatíveis com o comportamento operacional observado. Com isso, a aplicação industrial dos dados sintéticos contribui para ampliar a base analítica, apoiar a modelagem preditiva e reduzir incertezas no planejamento de ressuprimento.

2.3 APRENDIZADO DE MÁQUINA

O aprendizado de máquina tem ganhado destaque por sua capacidade de transformar dados em informações aplicáveis à tomada de decisão. Essa área da Inteligência Artificial permite desenvolver algoritmos capazes de reconhecer padrões, aprender com experiências anteriores e realizar previsões em cenários complexos, nos quais os dados apresentam múltiplas variáveis e relações dinâmicas (Jeon; Van Roy, 2025). Nesse processo, a estatística exerce papel fundamental ao fornecer bases para compreender incertezas, probabilidades e inferências (Cover; Thomas, 2012).

Um dos princípios centrais do aprendizado de máquina é a generalização, isto é, a capacidade de aplicar o conhecimento obtido a partir de dados históricos em situações ainda não observadas. Conforme Valiant (1984), essa propriedade é essencial para que algoritmos sejam considerados confiáveis, principalmente quando suas previsões são utilizadas para apoiar decisões futuras.

Nos processos industriais, essa lógica é relevante porque consumo, produção e demanda podem apresentar tendências, sazonalidades e variações operacionais. Assim, modelos preditivos e técnicas de aprendizado sequencial permitem identificar padrões temporais e apoiar decisões relacionadas ao planejamento de reposição da fita de aço, contribuindo para reduzir incertezas no ambiente fabril.

2.3.1 Aplicação de *machine learning* dentro do ambiente fabril

No ambiente industrial, o uso de aprendizado de máquina possibilita transformar as informações geradas diariamente no processo produtivo em conhecimento operacional útil. Dados sobre produção, histórico de abastecimento, consumo da fita de aço e variações de desempenho deixam de ser apenas registros históricos e passam a compor uma base analítica para previsão e apoio à decisão. De acordo com Jeon e Van Roy (2025), essa capacidade de identificar padrões e antecipar necessidades torna os modelos de aprendizado de máquina relevantes para apoiar o planejamento e reduzir incertezas operacionais.

Os modelos preditivos aplicados ao ambiente fabril podem integrar múltiplas variáveis do processo produtivo, como volume de produção, turnos, histórico de abastecimento, consumo registrado e frequência operacional. Conforme Cover e Thomas (2012), a combinação estruturada dessas informações amplia a capacidade

de análise, sobretudo quando os dados apresentam relações dinâmicas ao longo do tempo. Dessa forma, decisões relacionadas à reposição, ao gerenciamento de estoques e ao planejamento produtivo passam a ser fundamentadas em evidências quantitativas, e não apenas em percepções operacionais ou médias históricas.

Além da previsão direta de demanda, o aprendizado de máquina também permite avaliar cenários alternativos e antecipar impactos antes que mudanças sejam aplicadas ao processo real. Essa possibilidade é especialmente relevante em indústrias de grande escala, nas quais atrasos na reposição, variações no consumo ou mudanças no ritmo produtivo podem gerar perdas expressivas. Jeon e Van Roy (2025) observam que o uso de modelos analíticos nesse contexto auxilia na identificação de padrões, no ajuste de políticas de estoque e na construção de ações preventivas.

Assim, a aplicação de aprendizado de máquina no ambiente fabril ultrapassa a simples automação de cálculos ou geração de previsões isoladas. Segundo Jeon e Van Roy (2025), essa prática representa uma etapa relevante da transformação digital industrial, ao integrar inteligência analítica ao cotidiano das operações. No contexto de sistemas produtivos complexos, essa abordagem pode contribuir para melhorar o planejamento de reposição, apoiar decisões fundamentadas em dados e reduzir incertezas na gestão da fita de aço.

2.3.2 Modelos de previsão

Os modelos de previsão constituem instrumentos matemáticos e computacionais voltados à estimativa de comportamentos futuros a partir de padrões históricos. No contexto industrial, sua aplicação é relevante para antecipar demandas, apoiar o planejamento de reposição e reduzir incertezas operacionais. Segundo James et al. (2013), tais modelos permitem identificar relações entre variáveis e traduzir grandes volumes de dados em previsões úteis para a tomada de decisão. Na indústria de manufatura, especialmente em linhas de produção contínuas, a previsão de consumo contribui para reduzir riscos de ruptura, desperdícios e desequilíbrios nos níveis de estoque.

A escolha do modelo preditivo mais adequado depende da natureza dos dados, da complexidade das relações envolvidas e do objetivo analítico pretendido. Moraes e Souza (2024) destacam que técnicas estatísticas tradicionais, como regressão linear, permanecem eficazes para relações mais simples e de fácil interpretação, enquanto

algoritmos de aprendizado de máquina, como *Random Forest*, Redes Neurais e *XGBoost*, podem ser aplicados em cenários com interações não lineares e múltiplas variáveis correlacionadas. Essas abordagens, quando adequadamente estruturadas, permitem construir sistemas preditivos capazes de apoiar decisões estratégicas e operacionais.

No caso específico desta pesquisa, o uso de modelos de previsão tem como propósito apoiar a análise do consumo e do ressuprimento da fita de aço em uma fábrica de lâminas de barbear no Polo Industrial de Manaus. A partir de dados históricos de abastecimento, registros temporais e variáveis históricas, como *lags* e médias móveis, o modelo busca estimar o intervalo em dias até o próximo abastecimento. Essa capacidade de antecipação oferece suporte ao planejamento de reposição, alinhando a gestão da fita de aço às necessidades do processo produtivo.

2.3.2.1 Regressão linear simples

A regressão linear simples representa um dos modelos estatísticos fundamentais para análise preditiva, sendo amplamente empregada para investigar relações funcionais entre uma variável dependente e uma variável independente. Seu objetivo principal consiste em modelar matematicamente a influência de uma variável explicativa sobre uma variável resposta, permitindo estimar tendências futuras a partir de padrões históricos observados. Dentro da evolução dos modelos de predição, essa abordagem ocupa posição central por fornecer uma estrutura analítica interpretável, capaz de sintetizar relações quantitativas entre fenômenos observáveis e apoiar processos decisórios em diferentes áreas industriais e científicas (JAMES *et al.*, 2013).

A formulação matemática da regressão linear simples é tradicionalmente expressa por:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Em que Y representa a variável dependente, X corresponde à variável independente, β_0 representa o intercepto da reta, β_1 corresponde ao coeficiente angular associado à influência de X sobre Y , e ε representa o erro residual do modelo. Essa estrutura busca minimizar a soma dos desvios quadráticos entre valores observados e previstos, utilizando o método dos mínimos quadrados como base de ajuste estatístico. Conforme James *et al.* (2013), esse modelo constitui uma das bases

matemáticas mais relevantes para sistemas preditivos, enquanto Montgomery e Runger (2014) destacam sua robustez conceitual para análises iniciais de previsão e inferência estatística.

Para Morais (2024) utilização da regressão linear simples depende de pressupostos específicos, como linearidade entre variáveis, independência dos resíduos, homoscedasticidade e distribuição aproximadamente normal dos erros. Sob essas condições, o modelo fornece previsões consistentes e interpretações quantitativas relevantes, permitindo identificar tendências e relações causais com relativa precisão. Apesar de sua simplicidade, essa abordagem apresenta limitações quando aplicada a cenários altamente complexos ou não lineares, exigindo modelos mais sofisticados para representar interações múltiplas ou padrões operacionais mais dinâmicos.

No contexto desta pesquisa, a regressão linear simples assume papel introdutório como base conceitual para a compreensão matemática das relações entre variáveis operacionais e previsão de consumo. Embora a complexidade do sistema produtivo estudado exija métodos mais avançados, como a simulação de *Monte Carlo* e o *XGBoost*, a regressão linear simples oferece suporte relevante para análises exploratórias iniciais, compreensão da lógica preditiva e interpretação das relações entre dados históricos e comportamento futuro. Dessa forma, sua abordagem contribui para a construção progressiva do entendimento estatístico necessário à aplicação de modelos preditivos mais robustos no apoio à gestão de estoques e pontos de ressuprimento.

2.3.2.2 Regressão linear múltipla

No contexto da previsão industrial, o consumo de insumos dificilmente é explicado por um único fator isolado, pois sofre influência simultânea de variáveis como volume produzido, turno de operação, frequência de abastecimento, número de operadores, sazonalidade e condições produtivas. Nesse sentido, a regressão linear múltipla amplia a lógica da regressão simples ao permitir a análise conjunta de diferentes variáveis explicativas, tornando-se mais adequada para situações em que diversos fatores atuam sobre uma mesma variável de interesse. Conforme destaca Chein (2019), esse modelo permite estimar os efeitos parciais de cada regressor, mantendo constantes os demais fatores relevantes, princípio conhecido como *ceteris paribus*.

A formulação matemática da regressão linear múltipla pode ser representada por:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Em que Y representa a variável dependente, $X_1, X_2, \dots, X_n$ correspondem às variáveis independentes, β_0 representa o intercepto, $\beta_1, \beta_2, \dots, \beta_n$ indicam os coeficientes associados a cada variável explicativa, e ε representa o erro residual do modelo. A principal vantagem dessa estrutura está na possibilidade de mensurar o efeito individual de cada variável sobre o resultado previsto, considerando a presença simultânea das demais. Segundo Chein (2019), essa característica fortalece a capacidade analítica do modelo, pois reduz o risco de interpretações distorcidas que poderiam ocorrer quando apenas uma variável explicativa é considerada.

No campo estatístico, a estimação da regressão linear múltipla é frequentemente realizada pelo método dos Mínimos Quadrados Ordinários (MQO), cujo objetivo é minimizar a soma dos quadrados dos resíduos entre os valores observados e os valores previstos pelo modelo. Entretanto, para que os estimadores sejam consistentes e interpretáveis, algumas hipóteses devem ser observadas, como linearidade, exogeneidade, ausência de colinearidade perfeita entre os regressores, homoscedasticidade e independência dos erros. Conforme Chein (2019), a violação dessas condições pode comprometer a qualidade da inferência estatística, especialmente em casos de multicolinearidade, quando variáveis explicativas apresentam forte correlação entre si e dificultam a identificação do efeito individual de cada fator.

Em aplicações industriais, a regressão linear múltipla permite avaliar simultaneamente diferentes elementos que influenciam o consumo de materiais, a produtividade e o desempenho operacional. Por exemplo, na previsão de demanda de insumos, é possível considerar não apenas o volume de produção, mas também o turno, o número de operadores, a frequência de paradas, o histórico de abastecimento e a sazonalidade. Conforme Chein (2019), a interpretação dos coeficientes, associada a medidas como R^2 e R^2 ajustado, possibilita compreender o grau de explicação do modelo e verificar quais variáveis possuem maior influência sobre o comportamento analisado. Dessa forma, o modelo contribui para decisões mais fundamentadas sobre reposição, planejamento de compras e controle de estoques.

No contexto desta pesquisa, a regressão linear múltipla contribui como base estatística para compreender a relação entre múltiplas variáveis operacionais e a previsão de consumo em uma linha de produção de lâminas de barbear no Polo Industrial de Manaus. Embora o trabalho avance para abordagens mais robustas, como a simulação de *Monte Carlo* e o *XGBoost*, a regressão múltipla permite estabelecer uma etapa intermediária importante entre os modelos lineares tradicionais e os algoritmos de aprendizado de máquina.

Assim, ela auxilia na interpretação inicial dos fatores que influenciam o consumo, no entendimento da lógica preditiva e na fundamentação matemática do sistema proposto para apoio à gestão de estoques industriais.

2.3.2.3 Extreme Gradient Boosting — XGBoost

O *Extreme Gradient boosting (XGBoost)* é reconhecido como uma das técnicas mais eficientes no campo do aprendizado de máquina supervisionado, destacando-se por sua escalabilidade, velocidade de processamento e capacidade de lidar com grandes conjuntos de dados. De acordo com Chen (2023) e Feitosa *et al.* (2024), essa abordagem implementa o método de *gradient boosting* de forma otimizada, utilizando árvores de decisão como modelos fracos que são combinados sequencialmente para formar uma estrutura preditiva mais robusta. Cada nova árvore é treinada com o objetivo de corrigir os erros cometidos pelas árvores anteriores, permitindo que o modelo melhore progressivamente sua capacidade de previsão.

Do ponto de vista matemático, o algoritmo constrói suas estimativas de forma aditiva, ou seja, a previsão final é obtida pela soma das contribuições de várias árvores de decisão. Essa lógica pode ser representada pela seguinte expressão:

$$\hat{y}_i = \sum f_k(x_i), \text{ com } k = 1, 2, \dots, K$$

Em que $\hat{y}_i$ representa o valor previsto para a observação i , x_i corresponde ao conjunto de variáveis de entrada, f_k representa cada árvore de decisão e K indica o número total de árvores utilizadas. Assim, em vez de depender de uma única estrutura decisória, o método combina sucessivas árvores menores, de modo que cada uma contribua para reduzir parte do erro residual deixado pela anterior.

A função objetivo desse modelo combina dois elementos principais: a função de perda, responsável por medir o erro entre o valor real e o valor previsto, e o termo

de regularização, utilizado para controlar sua complexidade. Essa estrutura pode ser expressa como:

$$Obj = \sum l(y_i, \hat{y}_i) + \sum \Omega(f_k)$$

Nessa equação, $l(y_i, \hat{y}_i)$ representa a função de perda, isto é, a diferença entre o valor observado y_i e o valor previsto $\hat{y}_i$. Já $\Omega(f_k)$ corresponde ao termo de regularização aplicado às árvores, evitando que a estrutura se ajuste excessivamente aos dados de treinamento. Esse controle reduz o risco de *overfitting* e favorece melhor generalização para novos dados. Conforme Feitosa et al. (2024), a presença de mecanismos de regularização constitui um diferencial importante dessa técnica, pois penaliza configurações excessivamente complexas e favorece previsões mais estáveis. No caso do *XGBoost*, essas penalizações podem ser associadas aos termos L1 e L2: a regularização L1 contribui para reduzir a influência de variáveis ou pesos pouco relevantes, enquanto a L2 suaviza os pesos do modelo, evitando variações muito elevadas nas estimativas.

No contexto atual a adoção desse algoritmo justifica-se pela necessidade de estruturar uma abordagem preditiva voltada ao ressuprimento da fita de aço, considerando variáveis temporais e históricas associadas ao comportamento dos abastecimentos. Diferentemente de modelos lineares tradicionais, essa abordagem permite capturar relações não lineares e interações entre variáveis, tornando-se compatível com a complexidade do ambiente fabril analisado. Assim, o modelo atua como núcleo analítico da solução proposta, contribuindo para avaliar a possibilidade de estimar o intervalo até o próximo abastecimento e apoiar o planejamento de ressuprimento da fita de aço.

2.3.2.4 Comparativo entre modelos tradicionais e *machine learning*

O comparativo entre modelos tradicionais de previsão e técnicas de aprendizado de máquina é relevante porque cada abordagem apresenta vantagens e limitações conforme a natureza dos dados analisados. Conforme Moraes e Souza (2024), métodos tradicionais, como médias móveis, suavização exponencial, modelo ARIMA (*AutoRegressive Integrated Moving Average*) e regressões lineares, são amplamente utilizados por sua simplicidade, facilidade de interpretação e menor exigência computacional. No entanto, em processos com maior variabilidade,

múltiplas variáveis explicativas e relações não lineares, esses modelos podem apresentar limitações para capturar padrões mais complexos.

As técnicas de aprendizado de máquina, por sua vez, permitem que o modelo aprenda padrões diretamente dos dados, considerando interações entre diferentes variáveis. Segundo Gomes et al. (2023), algoritmos como *Random Forest*, redes neurais e *XGBoost* podem ser aplicados em cenários produtivos mais dinâmicos, nos quais o consumo é influenciado por fatores como volume de produção, turno, *lead time*, sazonalidade e frequência de abastecimento.

Como exemplo aplicado ao contexto brasileiro, Gomes et al. (2023) desenvolveram um estudo de caso em uma empresa de bens de consumo, comparando modelos de séries temporais e técnicas de aprendizado de máquina para previsão de demanda. A pesquisa utilizou cinco anos de histórico de vendas, realizou tratamento dos dados, remoção de *outliers* e treinamento de diferentes modelos preditivos. Os resultados indicaram redução do erro médio de previsão, diminuição dos estoques e aumento do nível de serviço, evidenciando que modelos mais avançados podem superar métodos tradicionais em cenários com maior variabilidade de demanda.

Neste trabalho, esse comparativo justifica a adoção do *XGBoost* como alternativa para estruturar uma modelagem preditiva aplicada ao ressuprimento da fita de aço. Embora os métodos tradicionais sejam importantes como base conceitual e estatística, o ambiente fabril estudado apresenta fatores operacionais e variações que exigem maior capacidade analítica. Assim, o uso de aprendizado de máquina mostra-se coerente com a proposta de reduzir incertezas e apoiar o planejamento de ressuprimento com base em dados.

2.4 VALIDAÇÃO ESTATÍSTICA DE MODELOS PREDITIVOS

A validação estatística de modelos preditivos constitui uma etapa essencial em pesquisas que utilizam dados históricos, bases simuladas e técnicas computacionais para estimar comportamentos futuros. Essa avaliação permite verificar o quanto os valores previstos se aproximam dos valores reais observados. Como destacam Bussab e Morettin (2010), a estatística fornece instrumentos para organizar, descrever e interpretar dados. Nesse sentido, o Erro Absoluto Médio (*Mean Absolute Error* — MAE) e a Raiz do Erro Quadrático Médio (*Root Mean Squared Error* — RMSE) são

utilizados para avaliar a qualidade das previsões, comparar valores reais e estimados e identificar a magnitude dos desvios produzidos pelo modelo.

2.4.1 Erro Absoluto Médio — MAE

O Erro Absoluto Médio, conhecido pela sigla *MAE*, mede a magnitude média dos erros de previsão (James *et al.*, 2013). Seu cálculo considera a diferença absoluta entre cada valor real e seu respectivo valor previsto. Assim, erros positivos e negativos não se anulam durante a avaliação. Dessa forma, o MAE indica o quanto, em média, o modelo se afasta dos valores observados.

A fórmula do MAE é representada por:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Em que y_i representa o valor real observado, $\hat{y}_i$ representa o valor previsto pelo modelo e n corresponde ao número total de observações avaliadas.

A utilização de médias como forma de resumir o comportamento de um conjunto de dados é amplamente discutida na estatística descritiva. Segundo Magalhães e Lima (2015), medidas de resumo são importantes porque permitem representar numericamente um conjunto de observações, facilitando sua interpretação e comparação. No caso do MAE, essa lógica é aplicada aos erros absolutos, resultando em um valor médio que expressa a distância geral entre previsão e realidade. Por permanecer na mesma unidade da variável analisada, esse indicador facilita a interpretação dos resultados em aplicações práticas.

Em estudos de previsão, o MAE é especialmente útil quando o objetivo é compreender o erro médio do modelo sem atribuir peso excessivo a desvios extremos. Diferentemente de métricas baseadas em quadrados dos erros, o MAE trata todos os desvios de forma proporcional, oferecendo uma visão equilibrada do desempenho geral da previsão. Essa característica é relevante em contextos industriais, nos quais o erro médio pode ser interpretado diretamente em termos operacionais, como quilogramas de insumo, litros consumidos ou tempo até necessidade de reposição.

Essa interpretação prática também se relaciona à estatística aplicada, que busca transformar dados numéricos em informações úteis para a tomada de decisão.

Martins e Domingues (2017) destacam que os métodos estatísticos auxiliam na análise de fenômenos reais ao permitir a organização, comparação e interpretação quantitativa dos dados. Nesse sentido, o MAE contribui para avaliar a precisão média do modelo preditivo, servindo como referência para verificar se as previsões apresentam proximidade aceitável em relação aos valores observados.

O *MAE* foi considerado adequado por permitir avaliar, de forma direta, o afastamento médio entre os valores observados e os valores previstos pelo modelo de consumo. Como a proposta envolve apoiar decisões de ressuprimento em ambiente industrial, a interpretação simples desse indicador facilita a análise do desempenho preditivo. Quanto menor o valor do *MAE*, maior tende a ser a proximidade média entre as previsões geradas e os registros utilizados na base analisada.

2.4.2 Raiz do Erro Quadrático Médio — RMSE

A Raiz do Erro Quadrático Médio, conhecida pela sigla RMSE, é uma métrica utilizada para avaliar a magnitude dos erros de previsão com maior sensibilidade a desvios elevados. Sua lógica consiste em calcular a diferença entre os valores reais e previstos, elevar esses erros ao quadrado, obter a média e, posteriormente, aplicar a raiz quadrada. Com isso, erros maiores passam a exercer maior influência sobre o resultado final, tornando o RMSE uma métrica relevante quando desvios elevados representam risco para a interpretação do modelo.

A fórmula do RMSE é representada por:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Em que y_i representa o valor real observado, $\hat{y}_i$ representa o valor previsto pelo modelo e n corresponde ao número total de observações avaliadas.

A lógica do RMSE está associada ao conceito de dispersão dos dados, uma vez que trabalha com os quadrados dos desvios entre valores observados e estimados. Bussab e Morettin (2010) destacam que medidas de variabilidade são fundamentais para compreender o grau de afastamento dos dados em relação a uma referência central. No caso do RMSE, essa lógica é aplicada à avaliação de previsões,

permitindo identificar se o modelo apresenta erros mais intensos em determinados pontos da base analisada.

Em aplicações industriais, essa sensibilidade a erros maiores é particularmente importante. Um pequeno desvio na previsão de consumo pode ser absorvido pela operação, mas um erro elevado pode gerar ruptura de estoque, atraso no abastecimento ou reposição acima da necessidade real. Dessa forma, o RMSE complementa o MAE ao evidenciar a presença de erros mais expressivos, oferecendo uma visão mais rigorosa sobre a estabilidade do modelo preditivo. Essa combinação entre indicadores permite avaliar tanto o erro médio geral quanto a ocorrência de desvios pontuais relevantes.

A utilização de métricas complementares também se aproxima da lógica da ciência de dados, na qual diferentes indicadores são empregados para avaliar modelos de forma mais completa. Conforme Morettin e Singer (2022), a análise de dados moderna envolve não apenas a construção de modelos, mas também sua validação por meio de critérios quantitativos capazes de verificar a qualidade dos resultados. Nesse sentido, o RMSE contribui para identificar situações em que o modelo apresenta maior instabilidade, mesmo quando outros indicadores apontam desempenho satisfatório.

O *RMSE* foi adotado como indicador complementar ao *MAE* para avaliar a qualidade das previsões de consumo e dos pontos de ressuprimento. Sua aplicação permite verificar se o modelo apresenta erros elevados que possam comprometer decisões operacionais no ambiente fabril. Enquanto o *MAE* fornece uma medida direta do erro médio, o *RMSE* amplia a análise ao destacar desvios mais críticos, contribuindo para uma avaliação mais robusta do desempenho preditivo.

2.4.3 Vazamento de dados

O vazamento de dados (*data leakage*) ocorre quando informações que não estariam disponíveis no momento da previsão são incorporadas ao processo de treinamento do modelo, resultando em estimativas artificialmente otimistas. Segundo James *et al.* (2013), esse problema pode comprometer a avaliação do desempenho preditivo, uma vez que o algoritmo passa a utilizar informações diretamente relacionadas ao resultado que se deseja prever. Como consequência, métricas como erro médio e coeficiente de determinação podem apresentar valores excessivamente favoráveis, sem refletir a real capacidade de generalização do modelo.

Por esse motivo, a identificação e prevenção do vazamento de dados constituem etapas fundamentais na construção de modelos preditivos. Conforme Chen (2023), a utilização de variáveis que contenham informações futuras ou diretamente associadas à variável-alvo pode induzir o algoritmo a reproduzir relações já conhecidas, em vez de aprender padrões genuínos dos dados. Dessa forma, a validação adequada dos atributos utilizados no treinamento contribui para garantir maior confiabilidade aos resultados obtidos.

2.4.4 Separação de treino-teste

A separação entre dados de treinamento e teste é uma prática amplamente utilizada na validação de modelos preditivos. Seu objetivo é avaliar a capacidade de generalização do algoritmo, verificando seu desempenho em registros que não participaram do processo de aprendizagem. Segundo James *et al.* (2013), essa estratégia reduz o risco de avaliações excessivamente otimistas e permite estimar de forma mais realista o comportamento do modelo em situações futuras.

No contexto de séries temporais e dados industriais, essa separação deve respeitar a ordem cronológica dos registros, evitando que informações futuras sejam utilizadas durante o treinamento. Conforme Cover e Thomas (2012), a preservação da sequência temporal é importante para garantir que as previsões sejam realizadas em condições semelhantes às encontradas na prática operacional. Assim, a divisão entre conjuntos de treinamento e teste contribui para uma avaliação mais consistente da capacidade preditiva do modelo.

2.5 MICROSOFT POWER APPS NA TRANSFORMAÇÃO DIGITAL

O *Microsoft Power Apps* é uma plataforma de desenvolvimento de aplicações de baixo código (*low-code*) pertencente ao ecossistema da *Microsoft Power Platform*, criada para permitir que organizações desenvolvam soluções personalizadas sem necessidade de programação avançada. Segundo a *Microsoft* (2024a), o *Power Apps* possibilita a criação de aplicativos empresariais capazes de conectar-se a diversas fontes de dados, como *Microsoft Dataverse*, *SharePoint*, *SQL Server* e *Dynamics 365*, com foco em agilidade, integração e governança corporativa. Dessa forma, a ferramenta tem sido amplamente utilizada em ambientes industriais e corporativos

para automatizar fluxos de trabalho, digitalizar operações e centralizar a coleta de informações operacionais.

2.5.1 O ambiente de desenvolvimento do *software Power Apps*

O *back-end* do *Power Apps* é estruturado a partir dos serviços da *Power Platform*, com destaque para o *Dataaverse*, que permite organizar tabelas, colunas, relacionamentos e regras de negócio. Segundo a Microsoft (2024b), essa camada garante a persistência e a integridade das informações utilizadas pelos aplicativos, além de oferecer recursos de autenticação, segurança e controle de acesso no ambiente corporativo.

Além do armazenamento, a plataforma permite a criação de regras e validações por meio do *Power Fx*, linguagem declarativa inspirada nas fórmulas do Excel. De acordo com a Microsoft (2024c), essa linguagem possibilita definir comportamentos automáticos, cálculos e validações em tempo real. Complementarmente, o *Power Automate* pode ser integrado para executar fluxos automatizados, notificações e processos baseados em eventos.

O *Power Apps* também dispõe de conectores que ampliam sua integração com diferentes fontes de dados, como SharePoint, bancos SQL, sistemas corporativos e serviços em nuvem. Essa conectividade é relevante em ambientes industriais, pois permite integrar informações operacionais a plataformas de gestão e visualização.

O *front-end* do *Power Apps* corresponde à camada de interação com o usuário, responsável pela construção das telas, formulários e painéis de acompanhamento. Segundo a Microsoft (2024a), a plataforma oferece compatibilidade com navegadores, tablets e dispositivos móveis, facilitando o acesso de operadores e gestores às informações do processo.

2.5.2 *Power Automate e Azure.*

O *Power Automate* é uma ferramenta da Microsoft voltada à automação de fluxos de trabalho entre diferentes aplicativos, sistemas e serviços digitais. Sua utilização permite criar rotinas automatizadas para atualização de dados, envio de notificações, integração entre plataformas e execução de tarefas repetitivas, reduzindo a necessidade de intervenção manual em processos operacionais e administrativos. Por utilizar uma abordagem de desenvolvimento de baixo código (*low-*

code), a ferramenta facilita a criação de fluxos automatizados mesmo por usuários com diferentes níveis de conhecimento técnico (MICROSOFT, 2026).

O Microsoft Azure, por sua vez, é uma plataforma de computação em nuvem que disponibiliza recursos para armazenamento, processamento, segurança, integração e análise de dados. Em ambientes industriais, sua aplicação pode apoiar a centralização de informações, a conexão entre sistemas e a estruturação de soluções baseadas em dados. Quando associado ao *Power Automate*, o Azure amplia as possibilidades de integração digital, permitindo que dados coletados em diferentes etapas do processo sejam armazenados, tratados e utilizados em sistemas de apoio à decisão (MICROSOFT, 2026).

No presente trabalho, essas ferramentas foram consideradas recursos de apoio à integração digital da solução proposta. O *Power Automate* pode estruturar fluxos automáticos de atualização e integração das informações, enquanto o Azure pode atuar como ambiente de suporte para armazenamento, processamento e disponibilização dos dados utilizados pelo modelo preditivo. Assim, essas tecnologias fundamentam a proposta de automação, integração e apoio à tomada de decisão no contexto do planejamento de ressuprimento da fita de aço.

2.5.3 Integração digital industrial e sistemas de apoio à decisão

A integração digital industrial consiste na conexão entre dados operacionais, sistemas de registro, ferramentas analíticas e interfaces de visualização, permitindo que informações antes dispersas sejam transformadas em suporte à tomada de decisão. No ambiente fabril, essa integração é importante porque aproxima o chão de fábrica da gestão, reduzindo a dependência de controles manuais, planilhas isoladas e análises tardias. Nesse cenário, sistemas de apoio à decisão contribuem para transformar dados produtivos em indicadores, alertas e informações acionáveis, favorecendo respostas mais rápidas diante de desvios, gargalos, falhas de abastecimento ou variações no consumo.

A proposta de integração digital será desenvolvida com apoio do *Microsoft Power Apps*, permitindo que os dados relacionados ao consumo, aos abastecimentos e aos intervalos de ressuprimento da fita de aço possam ser visualizados e operacionalizados em uma aplicação de uso prático. A ferramenta funcionará como interface entre a base de dados, o modelo preditivo e o usuário final, facilitando o acesso às informações geradas e apoiando decisões de reposição no ambiente

industrial. Com isso, o *Power Apps* não atua apenas como aplicativo de registro, mas como parte de uma solução digital integrada, conectando dados, previsão e tomada de decisão operacional.

Como trabalho relacionado, Luna (2024) desenvolveu um sistema de gestão da informação em uma mineradora de São Luís — MA, utilizando *Business Intelligence* por meio do *Microsoft Power BI* para gerenciamento de indicadores ligados ao processo de embarque. A proposta teve como objetivo desenvolver *dashboards* para acompanhamento de indicadores como Horas Paradas Operacionais (HPO) e OEE, permitindo identificar problemas operacionais e oportunidades de melhoria com maior rapidez. A autora relata que os *dashboards* foram construídos com dados organizados em Excel, tratados no *Power Query* e compartilhados por meio de ambiente digital, contribuindo para uma gestão mais eficiente, assertiva e orientada por dados.

Embora a ferramenta principal adotada nesta pesquisa seja o *Power Apps*, e não o *Power BI*, a contribuição de Luna (2024) reforça a importância da integração digital como mecanismo de apoio à decisão industrial. Enquanto o caso da mineradora concentrou-se na visualização e acompanhamento de indicadores operacionais, a aplicação proposta neste trabalho direciona essa integração para o acompanhamento do consumo e do ressuprimento da fita de aço em uma linha produtiva.

Assim, a lógica permanece semelhante: transformar dados produtivos em informação acessível para gestores e operadores, reduzir atrasos na análise, melhorar a visibilidade do processo e fortalecer a tomada de decisão baseada em dados.

3 PROCEDIMENTOS METODOLÓGICOS

3.1 NATUREZA DA PESQUISA

Segundo Gil (2002), o estudo de caso consiste em um método voltado à análise aprofundada de uma ou poucas unidades, como uma organização, processo ou setor específico. Essa abordagem é adequada para compreender fenômenos inseridos em seu ambiente real, considerando suas particularidades e evidências disponíveis. Na Engenharia de Produção, Miguel-Cauchick (2007) destaca que esse método é relevante para investigar fenômenos contemporâneos em sistemas produtivos, desde que sua condução apresente clareza quanto à unidade de análise, coleta de dados e critérios de avaliação.

Ainda conforme Gil (2002), o estudo de caso pode assumir caráter exploratório, descritivo ou explicativo, conforme os objetivos da pesquisa. Em investigações aplicadas à Engenharia de Produção, essa estratégia permite observar processos, tecnologias ou métodos de gestão em condições próximas da realidade operacional. Sob essa perspectiva, Miguel-Cauchick (2007) reforça que essa abordagem aproxima a investigação científica dos problemas reais das organizações, permitindo analisar limitações práticas e possibilidades de melhoria em ambientes industriais.

Quanto à abordagem, a pesquisa caracteriza-se como quantitativa, pois utiliza dados mensuráveis e técnicas estatísticas para analisar os fenômenos observados. Em aplicações recentes, Sampaio e Santos (2025) indicam que o estudo de caso pode apoiar o desenvolvimento de protótipos industriais, integrando dados, tecnologias digitais e avaliação operacional.

Assim, este trabalho adota uma abordagem quantitativa, desenvolvida por meio de estudo de caso em uma linha de produção de lâminas de barbear localizada no Polo Industrial de Manaus. A investigação concentra-se no desenvolvimento de uma solução preditiva voltada ao consumo e ao ressuprimento da fita de aço em kg, recurso produtivo de alto custo e elevada relevância para a continuidade do fluxo operacional. Para isso, são utilizados dados históricos reais, simulação de *Monte Carlo* e modelagem preditiva com *XGBoost*, buscando apoiar o planejamento de ressuprimento, reduzir incertezas operacionais e oferecer suporte à tomada de decisão no contexto produtivo.

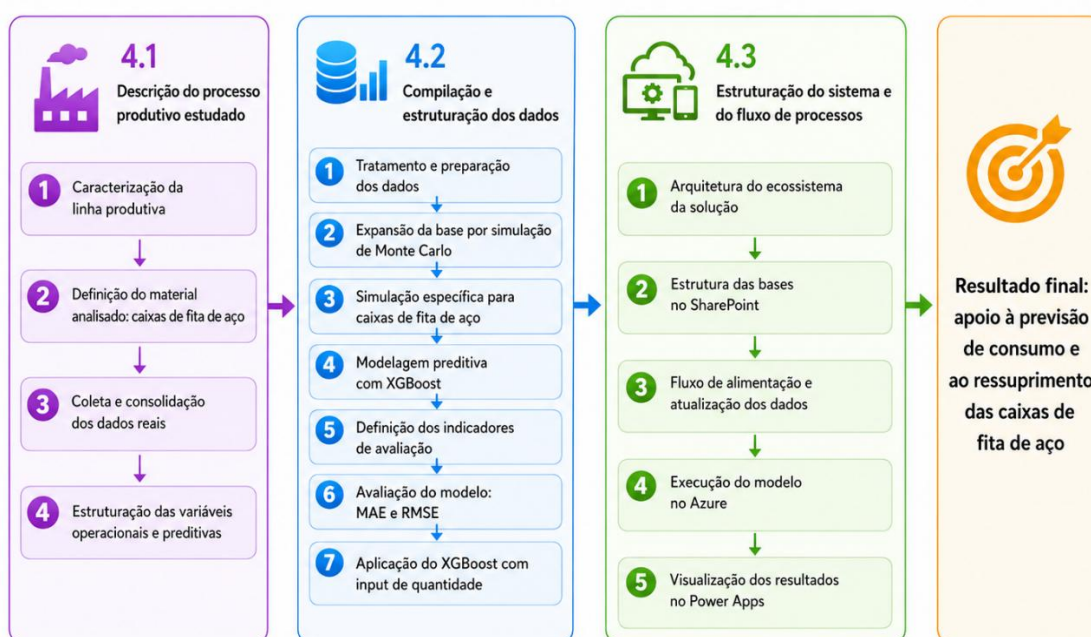
4 DESENVOLVIMENTO DO ESTUDO DE CASO

Esta seção apresenta a caracterização do processo produtivo analisado, com foco na linha industrial de fabricação de lâminas de barbear localizada no Polo Industrial de Manaus. Inicialmente, descreve-se o contexto operacional em que a fita de aço em kg está inserida, destacando sua relevância para a continuidade da produção e para o planejamento de abastecimento. Em seguida, são abordados os principais aspectos relacionados à coleta, organização e utilização dos dados desse recurso, os quais serviram de base para a estruturação da análise preditiva desenvolvida neste estudo.

4.1 DESCRIÇÃO DO PROCESSO PRODUTIVO ESTUDADO

A descrição do processo produtivo tem como finalidade apresentar o ambiente em que os dados foram coletados, destacando a importância do material analisado, a lógica de abastecimento observada e sua relação com as variáveis utilizadas na modelagem preditiva. Essa etapa é fundamental para garantir que a análise de consumo e ressuprimento esteja alinhada às condições reais do processo avaliado.

Figura 2 - Fluxo do desenvolvimento do estudo de caso



A linha produtiva analisada pertence ao segmento de bens de consumo e está inserida em um ambiente fabril localizado no Polo Industrial de Manaus. Por se tratar de um processo industrial contínuo e com elevada exigência de controle operacional, a disponibilidade da fita de aço em kg exerce influência direta sobre a manutenção do fluxo produtivo. Por esse motivo, a caracterização da linha foi realizada de forma geral, sem detalhamento técnico sensível do processo, concentrando-se nos aspectos relacionados ao consumo, ao abastecimento e ao suporte à tomada de decisão.

De maneira geral, o processo produtivo depende do abastecimento periódico da fita de aço para manter a continuidade das atividades. Esse recurso possui grande importância operacional, pois está diretamente associado ao fluxo de fabricação das lâminas de barbear. À medida que a linha opera, o material disponível é consumido gradualmente, exigindo reposição em momentos adequados para reduzir riscos de interrupções, atrasos ou perdas de eficiência produtiva.

A análise do processo produtivo não tem como objetivo descrever etapas específicas de fabricação, mas compreender o comportamento de consumo da fita de aço e sua relação com a dinâmica operacional da fábrica. Assim, o foco recai sobre a frequência de abastecimento, o intervalo entre os eventos de reposição e os padrões históricos que podem apoiar decisões mais estruturadas de ressuprimento.

Essa abordagem permite preservar informações internas do processo, ao mesmo tempo em que fornece base suficiente para a construção do modelo preditivo proposto. Ao direcionar a análise para a fita de aço em kg, recurso crítico, de alto custo e grande relevância produtiva, torna-se possível desenvolver uma análise mais objetiva, alinhada à realidade operacional e voltada à redução das incertezas relacionadas ao abastecimento.

4.1.1 Insumos e recursos considerados

Para a construção da base de análise, foi considerado como foco principal o comportamento de abastecimento da fita de aço em kg, material essencial ao funcionamento da linha produtiva de lâminas de barbear. A escolha ocorreu em razão de sua importância operacional, de seu alto custo e de sua influência direta na continuidade do processo produtivo.

Dentro do recorte definido para a pesquisa, a aplicação foi direcionada a esse item crítico, buscando tornar a análise mais objetiva e alinhada às necessidades do processo estudado. Essa delimitação permitiu concentrar a modelagem preditiva em

um material cuja previsão do intervalo até o próximo abastecimento pode contribuir diretamente para evitar interrupções e melhorar o planejamento de ressuprimento.

A relação entre o material analisado e o processo produtivo associado foi organizada na base de dados a partir dos registros históricos de abastecimento. Esses registros permitiram observar a frequência de reposição, os intervalos entre abastecimentos e os padrões temporais associados ao consumo da fita de aço em kg. A partir dessa estrutura, foram desenvolvidas as etapas posteriores de simulação por *Monte Carlo*, construção das variáveis temporais, aplicação do modelo *XGBoost* e posterior integração à solução digital proposta em *Power Apps*.

4.1.2 Coleta e consolidação dos dados reais

A coleta e consolidação dos dados reais constituiu a etapa inicial de formação da base analítica utilizada no desenvolvimento do modelo preditivo. Nessa fase, foram reunidas informações provenientes do sistema de Planejamento das Necessidades de Materiais, do inglês *Material Requirements Planning (MRP)*, e do acompanhamento operacional da linha produtiva.

Os registros reais compreenderam o período de janeiro de 2024 a março de 2026, totalizando 27 meses de acompanhamento. Esse recorte corresponde ao intervalo de dados que pôde ser coletado junto à empresa e representa a base histórica observada utilizada nesta pesquisa. A ampliação do período foi necessária devido à baixa frequência mensal de abastecimento da fita de aço em kg, que ocorre, em média, poucas vezes ao mês. Assim, mesmo com mais de dois anos de registros, a quantidade de observações disponíveis ainda exigiu tratamento criterioso para aplicação em um modelo preditivo.

Esse procedimento permitiu integrar registros sistêmicos com dados observados no ambiente fabril, aumentando a confiabilidade da base utilizada na análise. A estrutura inicial dos dados operacionais foi organizada em etapas, partindo dos registros reais, passando pelo tratamento e transformação, até sua formatação para leitura em ambiente *Python*, conforme apresentado na Figura 3. Essa organização possibilitou padronizar os dados coletados, reduzir inconsistências entre os registros e preparar as informações para as etapas subseqüentes de simulação e modelagem preditiva.

Figura 3 – Estrutura inicial dos dados operacionais para modelagem preditiva



Fonte: o autor, 2026.

A escolha desse material ocorreu em razão de sua relevância operacional, alto custo e impacto direto sobre a continuidade do fluxo produtivo. Cada registro foi estruturado de acordo com os principais parâmetros necessários à caracterização da base histórica e à posterior construção das variáveis utilizadas na modelagem preditiva.

Essa estrutura pode ser visualizada no quadro abaixo.

Quadro 1 – Estrutura dos dados reais coletados

Campo	Descrição	Finalidade na pesquisa
Insumo	Material acompanhado no processo	Identificar o item consumido
Recurso	Equipamento ou ponto produtivo associado	Relacionar consumo ao local de uso
Quantidade padrão	Volume padrão de abastecimento	Definir referência de reposição
Unidade	Unidade de medida do material	Padronizar os registros
Recipiente	Forma de acondicionamento	Caracterizar a operação de abastecimento
Quantidade mínima para funcionamento	Volume mínimo necessário para operação	Identificar limite crítico
Limite de armazenamento no processo	Capacidade máxima disponível no recurso	Definir restrição operacional
Lead time	Tempo relacionado ao consumo ou reposição	Apoiar previsão de ressurgimento
Frequência de abastecimento	Intervalo entre reposições	Caracterizar padrão de consumo

Fonte: o autor, 2026.

Após a coleta, os dados foram organizados considerando a relação entre a fita de aço em kg e o recurso produtivo analisado. Essa estrutura permitiu representar a dinâmica histórica de abastecimento do material no processo produtivo, contemplando principalmente a sequência temporal dos registros, a quantidade recebida em kg e os intervalos observados entre os eventos de reposição.

A base real consolidada serviu como referência para as etapas seguintes da metodologia, incluindo o tratamento dos dados, a simulação por *Monte Carlo*, a

construção das variáveis temporais, a criação de *lags* e médias móveis, o treinamento do modelo *XGBoost* e a posterior integração dos resultados ao *Power Apps*. Com isso, essa etapa foi fundamental para transformar registros operacionais em uma base organizada e compatível com técnicas de análise estatística e *machine learning*.

4.1.3 Estruturação das variáveis operacionais e preditivas

A estruturação das variáveis foi realizada com o objetivo de transformar os registros coletados em uma base adequada à modelagem preditiva. Conforme Hair et al. (2009), a correta definição das variáveis contribui para a consistência da análise, redução de ruídos e maior confiabilidade na interpretação dos resultados.

Na presente pesquisa, as variáveis foram organizadas com foco no comportamento de abastecimento da fita de aço em kg, considerando os registros históricos disponíveis entre janeiro de 2024 e março de 2026. A estruturação da base considerou informações associadas à sequência temporal dos abastecimentos, permitindo representar a dinâmica de reposição do material ao longo do período analisado.

Diferentemente da versão inicial da modelagem, as variáveis quantidade em kg, consumo médio por hora e tempo de *setup* não foram utilizadas como entradas diretas do modelo preditivo, a fim de evitar vazamento de alvo e tornar a previsão mais aderente à realidade operacional. Assim, o modelo passou a prever um evento efetivo do processo: o intervalo, em dias, até o próximo abastecimento da fita de aço em kg.

A partir dos registros históricos, foram extraídas variáveis temporais, como dia, mês, ano e dia da semana, além de variáveis derivadas baseadas no comportamento passado da série, como *lags*, intervalo anterior e médias móveis. Essas variáveis foram construídas considerando apenas informações disponíveis antes de cada evento previsto, preservando a ordem temporal dos dados e evitando que informações futuras influenciassem o treinamento do modelo.

Do ponto de vista da modelagem, as variáveis foram classificadas em variáveis de entrada, denominadas *features*, e variável-alvo, denominada *target*. As *features* correspondem aos atributos utilizados pelo modelo para realizar o aprendizado, enquanto o *target* representa o resultado a ser previsto. Neste estudo, o *target* foi definido como o intervalo em dias até o próximo abastecimento, permitindo que o modelo *XGBoost* estimasse quando um novo evento de reposição tenderia a ocorrer.

Com essa organização, os registros operacionais foram convertidos em uma estrutura analítica compatível com a simulação por *Monte Carlo*, o treinamento do modelo *XGBoost* e a posterior integração dos resultados ao *Power Apps*. A classificação das variáveis utilizadas na modelagem é apresentada no Quadro 2, permitindo visualizar de forma sistematizada a função de cada informação dentro da base analítica.

Quadro 2 – Classificação das variáveis utilizadas na modelagem

Variável	Tipo estatístico	Classificação	Descrição
Ano	Numérica	<i>Feature</i>	Representa o ano associado ao registro de abastecimento
Mês	Numérica	<i>Feature</i>	Representa o mês em que ocorreu o abastecimento
Dia do mês	Numérica	<i>Feature</i>	Representa o dia do mês associado ao evento de abastecimento
Dia da semana	Numérica/Categórica	<i>Feature</i>	Indica o dia da semana associado ao evento de abastecimento
Trimestre	Numérica	<i>Feature</i>	Representa o trimestre em que ocorreu o abastecimento
Lag	Numérica	<i>Feature</i>	Representa intervalos anteriores de abastecimento utilizados como histórico
Média móvel	Numérica	<i>Feature</i>	Representa a média dos intervalos anteriores dentro de uma janela temporal
Desvio móvel	Numérica	<i>Feature</i>	Representa a variação dos intervalos anteriores em uma janela temporal
Mínimo móvel	Numérica	<i>Feature</i>	Representa o menor intervalo observado em uma janela temporal anterior
Máximo móvel	Numérica	<i>Feature</i>	Representa o maior intervalo observado em uma janela temporal anterior
Tendência histórica	Numérica	<i>Feature</i>	Representa variações recentes entre intervalos anteriores
Origem simulada	Categórica/Binária	<i>Feature</i>	Indica se o registro pertence à base simulada
Intervalo até o próximo abastecimento	Numérica	<i>Target</i>	Representa o intervalo, em dias, até o próximo abastecimento da fita de aço

Fonte: o autor, 2026.

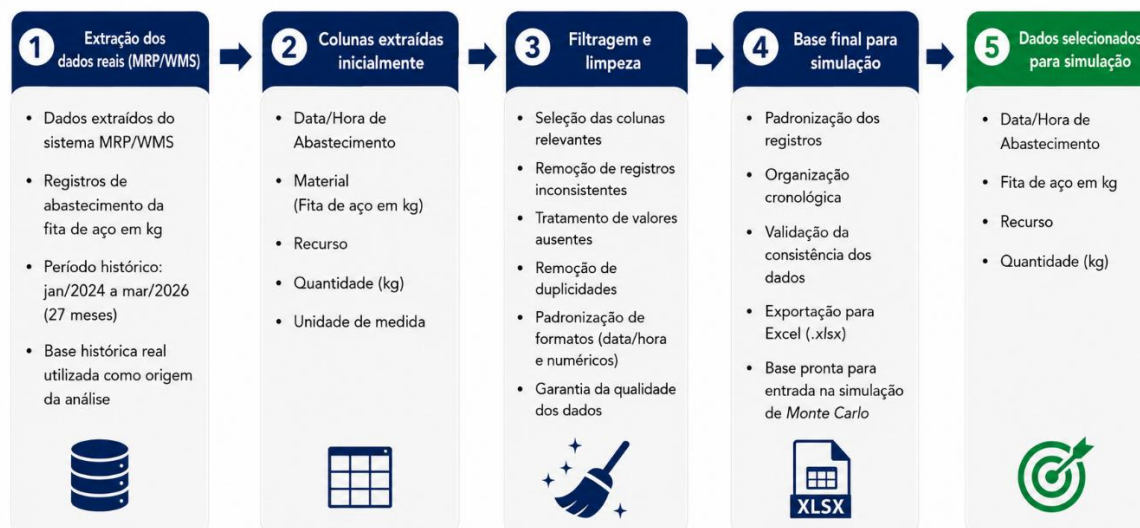
4.2 COMPILAÇÃO E ESTRUTURAÇÃO DOS DADOS

4.2.1 Tratamento e preparação dos dados

O tratamento e a preparação dos dados foram organizados em um fluxo sequencial, desde a extração inicial dos registros no sistema utilizado pela produção até a definição das variáveis finais consideradas na modelagem. Essa organização envolveu a seleção das colunas relevantes, a limpeza dos registros e a estruturação das informações em planilhas, garantindo que a base estivesse adequada para as

etapas posteriores de simulação e treinamento do modelo preditivo. O encadeamento dessas atividades é apresentado na Figura 4.

Figura 4 – Fluxo de coleta e consolidação da base real



Fonte: o autor, 2026.

Após a extração dos registros, a primeira etapa de filtragem foi realizada no próprio Excel, a partir da planilha gerada pelo sistema. Nessa fase, foram mantidas apenas as colunas necessárias para a construção da base histórica utilizada na pesquisa, considerando informações relacionadas à data e hora do abastecimento, ao material analisado, ao recurso produtivo associado e à quantidade registrada em quilogramas. Embora a base original apresentasse outras informações complementares, como código do material, número da requisição, unidade de medida, turno, usuário responsável, setor solicitante e observações, essas variáveis não foram consideradas relevantes para os objetivos desta investigação.

Em seguida, foram removidos registros com valores vazios ou nulos em campos essenciais, principalmente nas colunas de data e quantidade, por serem informações indispensáveis à organização cronológica dos eventos de abastecimento. Também foram tratados os casos de duplicidade, nos quais lançamentos registrados para um mesmo evento foram analisados individualmente, mantendo-se apenas os registros válidos e consistentes para a composição da base histórica.

Após esses ajustes, os dados foram organizados em planilhas eletrônicas, formando a base real consolidada utilizada nas etapas seguintes da pesquisa. Essa estrutura serviu como ponto de partida para a aplicação da simulação por *Monte Carlo*, responsável pela ampliação da base de análise a partir dos padrões observados nos dados reais.

Dessa forma, o tratamento dos dados teve como finalidade transformar a extração original do sistema em uma base organizada, consistente e adequada para as etapas de simulação, análise estatística e modelagem preditiva, conforme ilustrado na Figura 5.

Figura 5 – Base de dados utilizada.

Recebimento no Sistema	Dia	Insumo	Recurso	Quantidade kg	Consumo kg/h	Setup min	Origem	Custo	Mês	Ano
2024-01-02 10:00:00	2024-01-02 10:00:00	Fita de Aço	Prensa Lâmina	1466,2	48,67	20,26	Real	7624,24	2024-01-02 10:00:00	2024-01-02 10:00:00
2024-01-08 01:00:00	2024-01-08 01:00:00	Fita de Aço	Prensa Lâmina	1447,84	49,6	14,19	Real	7528,768	2024-01-08 01:00:00	2024-01-08 01:00:00
2024-01-14 07:00:00	2024-01-14 07:00:00	Fita de Aço	Prensa Lâmina	1420,1	51,58	22,5	Real	7384,52	2024-01-14 07:00:00	2024-01-14 07:00:00
2024-01-18 02:00:00	2024-01-18 02:00:00	Fita de Aço	Prensa Lâmina	1444,99	47,35	23,36	Real	7513,948	2024-01-18 02:00:00	2024-01-18 02:00:00
2024-01-23 07:00:00	2024-01-23 07:00:00	Fita de Aço	Prensa Lâmina	1424,1	50,99	16,86	Real	7405,32	2024-01-23 07:00:00	2024-01-23 07:00:00
2024-02-01 02:00:00	2024-02-01 02:00:00	Fita de Aço	Prensa Lâmina	1458,83	50,19	11,28	Real	7585,916	2024-02-01 02:00:00	2024-02-01 02:00:00
2024-02-08 12:00:00	2024-02-08 12:00:00	Fita de Aço	Prensa Lâmina	1452,69	49,87	11,4	Real	7553,988	2024-02-08 12:00:00	2024-02-08 12:00:00
2024-02-14 07:00:00	2024-02-14 07:00:00	Fita de Aço	Prensa Lâmina	1455,03	52,35	14,62	Real	7566,156	2024-02-14 07:00:00	2024-02-14 07:00:00
2024-02-19 03:00:00	2024-02-19 03:00:00	Fita de Aço	Prensa Lâmina	1432,04	51,79	17,23	Real	7446,608	2024-02-19 03:00:00	2024-02-19 03:00:00
2024-02-23 22:00:00	2024-02-23 22:00:00	Fita de Aço	Prensa Lâmina	1469,67	50,18	16,97	Real	7642,284	2024-02-23 22:00:00	2024-02-23 22:00:00
2024-03-02 10:00:00	2024-03-02 10:00:00	Fita de Aço	Prensa Lâmina	1461,62	49,56	18,38	Real	7600,424	2024-03-02 10:00:00	2024-03-02 10:00:00
2024-03-08 16:00:00	2024-03-08 16:00:00	Fita de Aço	Prensa Lâmina	1453,84	53,79	18,67	Real	7559,968	2024-03-08 16:00:00	2024-03-08 16:00:00
2024-03-14 21:00:00	2024-03-14 21:00:00	Fita de Aço	Prensa Lâmina	1439,44	48,54	11,39	Real	7485,088	2024-03-14 21:00:00	2024-03-14 21:00:00
2024-03-19 15:00:00	2024-03-19 15:00:00	Fita de Aço	Prensa Lâmina	1412,61	48,78	19,58	Real	7345,572	2024-03-19 15:00:00	2024-03-19 15:00:00
2024-03-25 18:00:00	2024-03-25 18:00:00	Fita de Aço	Prensa Lâmina	1454,82	53,24	19,04	Real	7565,064	2024-03-25 18:00:00	2024-03-25 18:00:00
2024-04-01 18:00:00	2024-04-01 18:00:00	Fita de Aço	Prensa Lâmina	1424,6	49,8	19,19	Real	7407,92	2024-04-01 18:00:00	2024-04-01 18:00:00
2024-04-09 00:00:00	2024-04-09 00:00:00	Fita de Aço	Prensa Lâmina	1454,07	49,27	18,31	Real	7561,164	2024-04-09 00:00:00	2024-04-09 00:00:00
2024-04-14 03:00:00	2024-04-14 03:00:00	Fita de Aço	Prensa Lâmina	1437,03	47,25	19,57	Real	7472,556	2024-04-14 03:00:00	2024-04-14 03:00:00
2024-04-19 15:00:00	2024-04-19 15:00:00	Fita de Aço	Prensa Lâmina	1446,91	48,8	21,74	Real	7523,932	2024-04-19 15:00:00	2024-04-19 15:00:00
2024-04-26 22:00:00	2024-04-26 22:00:00	Fita de Aço	Prensa Lâmina	1428,66	50,92	18,15	Real	7429,032	2024-04-26 22:00:00	2024-04-26 22:00:00
2024-05-02 09:00:00	2024-05-02 09:00:00	Fita de Aço	Prensa Lâmina	1479,86	50,89	19,42	Real	7695,272	2024-05-02 09:00:00	2024-05-02 09:00:00
2024-05-08 07:00:00	2024-05-08 07:00:00	Fita de Aço	Prensa Lâmina	1425,6	50,21	15,19	Real	7413,12	2024-05-08 07:00:00	2024-05-08 07:00:00
2024-05-14 14:00:00	2024-05-14 14:00:00	Fita de Aço	Prensa Lâmina	1473,25	49,55	13,88	Real	7660,9	2024-05-14 14:00:00	2024-05-14 14:00:00
2024-05-20 03:00:00	2024-05-20 03:00:00	Fita de Aço	Prensa Lâmina	1457,02	46,14	16,59	Real	7576,504	2024-05-20 03:00:00	2024-05-20 03:00:00
2024-05-26 06:00:00	2024-05-26 06:00:00	Fita de Aço	Prensa Lâmina	1430,71	50,57	17,73	Real	7439,692	2024-05-26 06:00:00	2024-05-26 06:00:00
2024-05-31 22:00:00	2024-05-31 22:00:00	Fita de Aço	Prensa Lâmina	1430,99	48,92	18,29	Real	7441,148	2024-05-31 22:00:00	2024-05-31 22:00:00
2024-06-06 03:00:00	2024-06-06 03:00:00	Fita de Aço	Prensa Lâmina	1421,19	49,14	15,51	Real	7390,188	2024-06-06 03:00:00	2024-06-06 03:00:00
2024-06-12 00:00:00	2024-06-12 00:00:00	Fita de Aço	Prensa Lâmina	1492,2	49,22	23,46	Real	7759,44	2024-06-12 00:00:00	2024-06-12 00:00:00
2024-06-18 06:00:00	2024-06-18 06:00:00	Fita de Aço	Prensa Lâmina	1416,59	49,66	18,9	Real	7366,268	2024-06-18 06:00:00	2024-06-18 06:00:00
2024-06-23 01:00:00	2024-06-23 01:00:00	Fita de Aço	Prensa Lâmina	1452,18	50,86	13,15	Real	7551,336	2024-06-23 01:00:00	2024-06-23 01:00:00
2024-06-29 15:00:00	2024-06-29 15:00:00	Fita de Aço	Prensa Lâmina	1436,15	52,12	17,61	Real	7467,98	2024-06-29 15:00:00	2024-06-29 15:00:00

Fonte: o autor, 2026.

4.2.2 Base real para expansão

A etapa de simulação foi iniciada a partir da organização da base real de abastecimento da fita de aço em kg utilizada no processo produtivo estudado. Os dados históricos foram coletados diretamente da operação industrial e compreenderam registros reais realizados entre janeiro de 2024 e março de 2026, totalizando 27 meses de observação.

Durante essa etapa, buscou-se preservar a estrutura operacional dos dados, mantendo as informações necessárias à caracterização dos eventos de abastecimento, como data e hora do registro, material analisado, recurso produtivo

associado e quantidade registrada em quilogramas. Essa organização permitiu construir uma base histórica real compatível com a análise temporal dos abastecimentos.

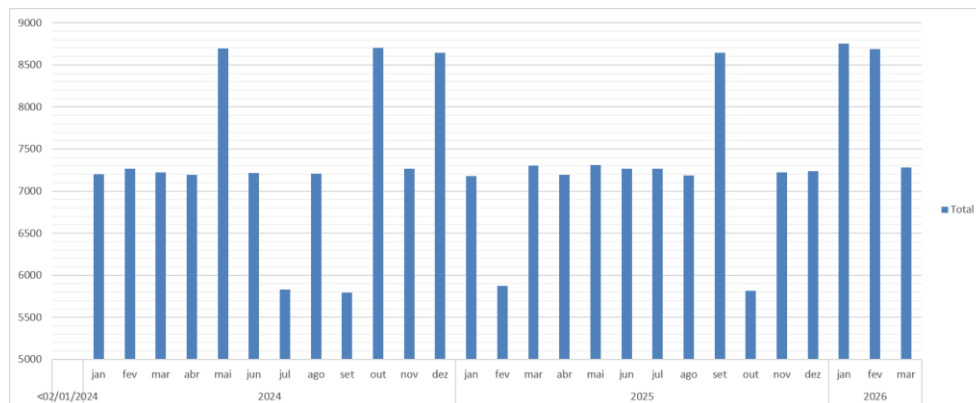
Embora o período coletado represente uma ampliação significativa em relação à base inicialmente utilizada, observou-se que a frequência mensal de abastecimento da fita de aço em kg é relativamente baixa, ocorrendo, em média, entre cinco e sete vezes por mês. Dessa forma, mesmo com 27 meses de registros reais, a quantidade de observações permaneceu limitada para o treinamento de um modelo preditivo baseado em *machine learning*.

A partir dessa limitação, a simulação por *Monte Carlo* foi aplicada como estratégia complementar de ampliação da base de análise. O objetivo não foi substituir os registros reais, mas gerar cenários adicionais compatíveis com o comportamento operacional observado, ampliando a variabilidade da amostra e fornecendo maior volume de registros para as etapas posteriores de modelagem. Foram simulados mais 27 meses de comportamento operacional, mantendo coerência com os padrões identificados na base real. Com isso, a base passou a contemplar o horizonte histórico real de janeiro de 2024 a março de 2026 e um horizonte simulado complementar, utilizado para ampliar a representatividade dos eventos de abastecimento da fita de aço em kg.

A partir dessa visualização inicial, foi possível observar a distribuição dos registros ao longo dos meses coletados e compreender como os abastecimentos ocorreram dentro do intervalo analisado. Essa etapa não teve como finalidade estabelecer uma comparação direta entre dados reais e simulados, mas reconhecer o comportamento operacional da base real e utilizá-lo como referência para a geração de cenários complementares. Com isso, a simulação por *Monte Carlo* passou a ser aplicada como uma estratégia de expansão controlada da base, preservando a coerência estatística e a compatibilidade com a dinâmica observada na operação.

Na Figura 6, é apresentado o comportamento temporal da base real utilizada na pesquisa. O gráfico permite visualizar a distribuição das quantidades de fita de aço requisitadas ao longo do período analisado, evidenciando as variações observadas entre os meses e contribuindo para a compreensão do padrão de abastecimento registrado na operação.

Figura 6 – Quantidade de fita de aço requisitada por mês/ano



Fonte: próprio autor, 2026

A Figura 6 apresenta a quantidade requisitada de fita de aço em kg ao longo do período de janeiro de 2024 a março de 2026. Observa-se que os volumes mensais apresentaram variações ao longo do horizonte analisado, refletindo as oscilações naturais da demanda e da dinâmica de abastecimento do processo produtivo.

Durante o período avaliado, os volumes mensais oscilaram entre aproximadamente 5.796,51 kg, registrado em setembro de 2024, e 8.754,23 kg, observado em janeiro de 2026. De forma geral, a maior parte dos registros concentrou-se na faixa entre 7.000 kg e 8.700 kg por mês, evidenciando relativa estabilidade operacional, porém com flutuações compatíveis com a rotina produtiva da fábrica.

O comportamento observado demonstra que os abastecimentos da fita de aço em kg não ocorreram de forma completamente uniforme ao longo do tempo, apresentando variações entre os meses analisados. Essa característica reforça a necessidade de utilização de técnicas capazes de identificar padrões temporais e apoiar a previsão dos eventos futuros de abastecimento.

4.2.3 Simulação da base de dados da fita de aço em kg

Após a organização da base real de abastecimento da fita de aço em kg, iniciou-se a etapa de ampliação dos dados por meio da simulação por *Monte Carlo*. Os registros históricos utilizados na pesquisa compreenderam o período entre janeiro de

2024 e março de 2026, totalizando 27 meses de observação real, coletados diretamente da operação industrial estudada.

Durante essa etapa, buscou-se preservar a estrutura operacional da base original, mantendo as colunas essenciais à análise, como data e hora do abastecimento, material analisado, recurso produtivo associado e quantidade registrada em quilogramas. Essa preservação foi necessária para que os cenários gerados mantivessem compatibilidade com a lógica observada no ambiente fabril real.

Embora a base histórica tenha sido ampliada em relação à versão inicial da pesquisa, a frequência mensal de abastecimento da fita de aço em kg permaneceu relativamente baixa, ocorrendo, em média, entre cinco e sete vezes ao mês. Dessa forma, mesmo com 27 meses de registros reais, a quantidade de observações ainda se mostrou limitada para o treinamento de um modelo preditivo baseado em *machine learning*. Por esse motivo, identificou-se a necessidade de ampliar a base por meio de cenários complementares, preservando a coerência estatística dos dados observados.

A ampliação da base foi realizada utilizando a técnica de simulação por *Monte Carlo* em ambiente *Python*, com apoio das bibliotecas *Pandas* e *NumPy* para leitura, tratamento e geração dos dados. A lógica da simulação foi construída a partir dos próprios parâmetros observados na base real, preservando aderência operacional e coerência com o comportamento identificado no processo produtivo estudado.

Inicialmente, foram extraídos da base real os parâmetros estatísticos utilizados na geração dos novos registros, considerando os valores mínimo, médio e máximo das quantidades requisitadas, além da frequência mensal de eventos observada na operação. Esses parâmetros serviram como referência para a definição dos limites operacionais da simulação de *Monte Carlo*, permitindo que os cenários gerados mantivessem coerência com o comportamento histórico da fita de aço em kg. A partir dessa estrutura, foram simulados mais 27 meses de registros complementares, ampliando a base analítica utilizada nas etapas posteriores de tratamento e modelagem com *XGBoost*.

Figura 7 – Trecho do código responsável pela parametrização da simulação a partir da base real.

```
# =====  
# PARÂMETROS DA BASE REAL  
# =====  
  
quant_min = float(df_real["Quantidade kg"].min())  
quant_media = float(df_real["Quantidade kg"].mean())  
quant_max = float(df_real["Quantidade kg"].max())  
  
consumo_min = float(df_real["Consumo kg/h"].min())  
consumo_media = float(df_real["Consumo kg/h"].mean())  
consumo_max = float(df_real["Consumo kg/h"].max())  
  
setup_min = float(df_real["Setup min"].min())  
setup_media = float(df_real["Setup min"].mean())  
setup_max = float(df_real["Setup min"].max())  
  
insumo_fixo = df_real["Insumo"].mode()[0] if "Insumo" in df_real.columns else "Fita de Aço"  
recurso_fixo = df_real["Recurso"].mode()[0] if "Recurso" in df_real.columns else "Prensa Lâmina"
```

Fonte: o autor, 2026.

Posteriormente, foi utilizada uma distribuição triangular para a geração dos novos registros simulados. A escolha dessa distribuição ocorreu por permitir a geração de valores dentro de limites operacionais conhecidos, utilizando como referência os valores mínimos, médios e máximos previamente identificados na base real. Essa abordagem foi considerada adequada devido à baixa frequência de ressuprimentos da fita de aço em kg, que ocorre poucas vezes ao mês, limitando a quantidade de amostras disponíveis para a modelagem preditiva. Dessa forma, a simulação foi parametrizada a partir do próprio comportamento da operação, e não definida de maneira arbitrária.

Na lógica adotada, a quantidade requisitada da fita de aço em kg foi simulada com base nos limites observados na base real. Além disso, foram considerados parâmetros operacionais associados ao consumo em kg/h, ao tempo de *setup* e aos intervalos reais entre abastecimentos, permitindo que os registros simulados preservassem coerência com a dinâmica histórica da operação.

A frequência dos eventos simulados também foi orientada pelo comportamento observado na base real, considerando os intervalos entre abastecimentos e a sequência temporal dos registros. Com isso, os dados gerados por *Monte Carlo* não tiveram como finalidade substituir os dados reais, mas ampliar a base de análise por meio de cenários estatisticamente coerentes e compatíveis com o comportamento operacional observado, isso pode ser observado na figura 8.

Figura 8 – Trecho do código responsável pela geração dos eventos simulados utilizando *Monte Carlo*.

```
# =====  
# MONTE CARLO SEQUENCIAL  
# =====  
  
registros_simulados = []  
log_simulacao = []  
  
ultimo_real = df_real.sort_values("Recebimento no Sistema").iloc[-1]  
  
quant_ref = float(ultimo_real["Quantidade kg"])  
consumo_ref = float(ultimo_real["Consumo kg/h"])  
setup_ref = float(ultimo_real["Setup min"])  
  
# Gera primeiro intervalo a partir do último real  
intervalo_inicial = gerar_intervalo_sequencial(  
    quant_ref,  
    consumo_ref,  
    setup_ref  
)  
  
data_atual = ultima_data_real + pd.to_timedelta(intervalo_inicial, unit="D")  
  
# Avança até entrar no período de simulação  
while data_atual < data_inicio_simulacao:  
    quantidade_temp = gerar_triangular(quant_min, quant_media, quant_max)  
    consumo_temp = gerar_triangular(consumo_min, consumo_media, consumo_max)  
    setup_temp = gerar_triangular(setup_min, setup_media, setup_max)  
  
    intervalo_temp = gerar_intervalo_sequencial(  
        quantidade_temp,  
        consumo_temp,  
        setup_temp  
    )  
  
    data_atual = data_atual + pd.to_timedelta(intervalo_temp, unit="D")  
    data_atual = ajustar_para_dia_valido(data_atual, feriados)  
  
contador = 1  
  
while data_atual <= data_fim_simulacao:  
  
    data_atual = ajustar_para_dia_valido(data_atual, feriados)  
  
    if data_atual > data_fim_simulacao:  
        break
```

Fonte: o autor, 2026.

Assim, os registros simulados mantiveram a mesma estrutura da base real, preservando as colunas essenciais, a organização cronológica e as relações operacionais observadas nos dados coletados na empresa. Essa decisão teve como objetivo garantir aderência entre os cenários gerados por *Monte Carlo* e a lógica existente no ambiente fabril, evitando que a base ampliada se distanciasse do comportamento histórico da fita de aço em kg.

Figura 9 – Estruturas de saída

```
# BASE FINAL
# =====

colunas_saida = [
    "Recebimento no Sistema",
    "Dia",
    "Insumo",
    "Recurso",
    "Quantidade kg",
    "Consumo kg/h",
    "Setup min",
    "Origem",
    "Custo",
    "Mês",
    "Ano"
]
```

Fonte: o autor, 2026.

A figura abaixo apresenta uma amostra da base gerada após a simulação por *Monte Carlo*, evidenciando a manutenção da estrutura operacional observada na base real e a continuidade dos registros ao longo dos meses simulados. Essa visualização permite verificar que os campos principais da base foram preservados, como data de recebimento no sistema, material analisado, recurso produtivo associado, quantidade em kg, origem do registro, custo, mês e ano, como pode ser visto na figura 10.

A base resultante foi exportada para Excel (figura 11), com o objetivo de facilitar a conferência, a organização e a visualização dos dados utilizados nas etapas posteriores da pesquisa. Essa exportação também possibilitou a separação entre registros reais e simulados, favorecendo a validação da estrutura dos dados antes da aplicação das etapas seguintes de análise e modelagem preditiva.

Figura 10 – Amostra da base simulada gerada por *Monte Carlo*.

```
Arquivo Excel: BaseAco_Real_Simulada_54Meses.xlsx
Arquivo gráfico: Grafico_Real_Simulado_54Meses.png

Registros reais: 137
Registros simulados: 135
Total de registros: 272

Volume real total: 198,478.16 kg
Volume simulado total: 195,625.82 kg
Volume total: 394,103.98 kg

Registros por origem:
Origem
Real      137
Simulado  135
Name: count, dtype: int64

Resumo mensal:
  Mês  Origem  Registros  Quantidade_total_kg  Consumo_medio_kgh  Setup_medio_min  Custo_total
2024-01  Real      5      7203.23      49.638000      19.434000      37456.796
2024-02  Real      5      7268.26      50.876000      14.380000      37794.952
2024-03  Real      5      7222.33      50.782000      17.412000      37556.116
2024-04  Real      5      7191.27      49.208000      19.392000      37394.604
2024-05  Real      6      8697.43      49.380000      16.850000      45226.636
2024-06  Real      5      7218.31      50.280000      17.726000      37535.212
2024-07  Real      4      5831.33      50.677500      17.482500      30322.916
2024-08  Real      5      7209.76      51.460000      22.290000      37490.752
2024-09  Real      4      5796.51      50.047500      19.360000      30141.852
2024-10  Real      6      8705.86      50.848333      17.555000      45270.472
2024-11  Real      5      7264.64      49.402000      17.074000      37776.128
2024-12  Real      6      8647.87      50.651667      16.401667      44968.924
```

Fonte: o autor, 2026.

Figura 11 – Base de dados gerada

Recebimento no Sistema	Dia	Insumo	Recurso	Quantidade kg	Consumo kg/h	Setup min	Origem	Custo	Mês	Ano
2025-12-12 06:00:00	2025-12-12 06:00:00	Fita de Aço	Prensa Lâmina	1442,16	49,68	11,98	Real	7499,232	2025-12-12 06:00:00	2025-12-12 06:00:00
2025-12-16 19:00:00	2025-12-16 19:00:00	Fita de Aço	Prensa Lâmina	1474,31	48,24	16,37	Real	7666,412	2025-12-16 19:00:00	2025-12-16 19:00:00
2025-12-22 21:00:00	2025-12-22 21:00:00	Fita de Aço	Prensa Lâmina	1424,28	50,33	19,36	Real	7406,256	2025-12-22 21:00:00	2025-12-22 21:00:00
2025-12-28 12:00:00	2025-12-28 12:00:00	Fita de Aço	Prensa Lâmina	1408,34	52,65	13,23	Real	7323,368	2025-12-28 12:00:00	2025-12-28 12:00:00
2026-01-02 16:00:00	2026-01-02 16:00:00	Fita de Aço	Prensa Lâmina	1487,71	50,58	22,22	Real	7736,092	2026-01-02 16:00:00	2026-01-02 16:00:00
2026-01-08 04:00:00	2026-01-08 04:00:00	Fita de Aço	Prensa Lâmina	1461,2	47,56	14,26	Real	7598,24	2026-01-08 04:00:00	2026-01-08 04:00:00
2026-01-14 07:00:00	2026-01-14 07:00:00	Fita de Aço	Prensa Lâmina	1459,23	52,44	19,95	Real	7587,996	2026-01-14 07:00:00	2026-01-14 07:00:00
2026-01-18 09:00:00	2026-01-18 09:00:00	Fita de Aço	Prensa Lâmina	1444,85	48,02	19,83	Real	7513,22	2026-01-18 09:00:00	2026-01-18 09:00:00
2026-01-24 05:00:00	2026-01-24 05:00:00	Fita de Aço	Prensa Lâmina	1452,72	52,41	13,31	Real	7554,144	2026-01-24 05:00:00	2026-01-24 05:00:00
2026-01-29 14:00:00	2026-01-29 14:00:00	Fita de Aço	Prensa Lâmina	1448,52	49,72	19,02	Real	7532,304	2026-01-29 14:00:00	2026-01-29 14:00:00
2026-02-03 16:00:00	2026-02-03 16:00:00	Fita de Aço	Prensa Lâmina	1453,72	51,67	18,29	Real	7559,344	2026-02-03 16:00:00	2026-02-03 16:00:00
2026-02-08 07:00:00	2026-02-08 07:00:00	Fita de Aço	Prensa Lâmina	1441,74	52,94	18,68	Real	7497,048	2026-02-08 07:00:00	2026-02-08 07:00:00
2026-02-14 02:00:00	2026-02-14 02:00:00	Fita de Aço	Prensa Lâmina	1449,62	49,59	22,4	Real	7538,024	2026-02-14 02:00:00	2026-02-14 02:00:00
2026-02-17 23:00:00	2026-02-17 23:00:00	Fita de Aço	Prensa Lâmina	1449,15	52,49	13,29	Real	7535,58	2026-02-17 23:00:00	2026-02-17 23:00:00
2026-02-23 15:00:00	2026-02-23 15:00:00	Fita de Aço	Prensa Lâmina	1420	52,42	12,52	Real	7384	2026-02-23 15:00:00	2026-02-23 15:00:00
2026-02-27 18:00:00	2026-02-27 18:00:00	Fita de Aço	Prensa Lâmina	1476,74	46,45	23,41	Real	7679,048	2026-02-27 18:00:00	2026-02-27 18:00:00
2026-03-06 17:00:00	2026-03-06 17:00:00	Fita de Aço	Prensa Lâmina	1461,35	52,25	19,64	Real	7599,02	2026-03-06 17:00:00	2026-03-06 17:00:00
2026-03-13 13:00:00	2026-03-13 13:00:00	Fita de Aço	Prensa Lâmina	1433,67	53,84	19,57	Real	7455,084	2026-03-13 13:00:00	2026-03-13 13:00:00
2026-03-18 10:00:00	2026-03-18 10:00:00	Fita de Aço	Prensa Lâmina	1462,17	50,55	19,54	Real	7603,284	2026-03-18 10:00:00	2026-03-18 10:00:00
2026-03-25 03:00:00	2026-03-25 03:00:00	Fita de Aço	Prensa Lâmina	1478,46	49,53	18,57	Real	7687,992	2026-03-25 03:00:00	2026-03-25 03:00:00
2026-03-30 07:00:00	2026-03-30 07:00:00	Fita de Aço	Prensa Lâmina	1447,61	50,64	18,22	Real	7527,572	2026-03-30 07:00:00	2026-03-30 07:00:00
2026-04-04 08:00:00	2026-04-04 08:00:00	Fita de Aço	Prensa Lâmina	1431,33	47,59	20,96	Simulado	7442,916	2026-04-04 08:00:00	2026-04-04 08:00:00
2026-04-10 21:00:00	2026-04-10 21:00:00	Fita de Aço	Prensa Lâmina	1459,58	47	22,75	Simulado	7589,816	2026-04-10 21:00:00	2026-04-10 21:00:00
2026-04-17 06:00:00	2026-04-17 06:00:00	Fita de Aço	Prensa Lâmina	1433,27	49,46	17,96	Simulado	7453,004	2026-04-17 06:00:00	2026-04-17 06:00:00
2026-04-22 20:00:00	2026-04-22 20:00:00	Fita de Aço	Prensa Lâmina	1439,75	50,9	14,72	Simulado	7486,7	2026-04-22 20:00:00	2026-04-22 20:00:00
2026-04-28 01:00:00	2026-04-28 01:00:00	Fita de Aço	Prensa Lâmina	1464,22	48,83	17,89	Simulado	7613,944	2026-04-28 01:00:00	2026-04-28 01:00:00
2026-05-02 18:00:00	2026-05-02 18:00:00	Fita de Aço	Prensa Lâmina	1420,89	50,88	15,08	Simulado	7388,628	2026-05-02 18:00:00	2026-05-02 18:00:00
2026-05-07 14:00:00	2026-05-07 14:00:00	Fita de Aço	Prensa Lâmina	1426,53	51,26	17,39	Simulado	7417,956	2026-05-07 14:00:00	2026-05-07 14:00:00
2026-05-13 13:00:00	2026-05-13 13:00:00	Fita de Aço	Prensa Lâmina	1449,31	47,26	21,56	Simulado	7536,412	2026-05-13 13:00:00	2026-05-13 13:00:00
2026-05-18 14:00:00	2026-05-18 14:00:00	Fita de Aço	Prensa Lâmina	1450,38	50,6	15,24	Simulado	7541,976	2026-05-18 14:00:00	2026-05-18 14:00:00
2026-05-26 11:00:00	2026-05-26 11:00:00	Fita de Aço	Prensa Lâmina	1463,57	53,14	21,35	Simulado	7610,564	2026-05-26 11:00:00	2026-05-26 11:00:00
2026-05-31 17:00:00	2026-05-31 17:00:00	Fita de Aço	Prensa Lâmina	1438,66	52,14	16,78	Simulado	7481,032	2026-05-31 17:00:00	2026-05-31 17:00:00
2026-06-05 06:00:00	2026-06-05 06:00:00	Fita de Aço	Prensa Lâmina	1451,38	48,4	20,23	Simulado	7547,176	2026-06-05 06:00:00	2026-06-05 06:00:00
2026-06-11 03:00:00	2026-06-11 03:00:00	Fita de Aço	Prensa Lâmina	1460,78	51,76	13,79	Simulado	7596,056	2026-06-11 03:00:00	2026-06-11 03:00:00
2026-06-17 12:00:00	2026-06-17 12:00:00	Fita de Aço	Prensa Lâmina	1428,15	52,4	18,66	Simulado	7426,38	2026-06-17 12:00:00	2026-06-17 12:00:00
2026-06-23 00:00:00	2026-06-23 00:00:00	Fita de Aço	Prensa Lâmina	1441,53	51,51	18,77	Simulado	7495,956	2026-06-23 00:00:00	2026-06-23 00:00:00
2026-06-29 06:00:00	2026-06-29 06:00:00	Fita de Aço	Prensa Lâmina	1440,24	49,73	19,28	Simulado	7521,368	2026-06-29 06:00:00	2026-06-29 06:00:00

Fonte: o autor, 2026.

Em síntese, a aplicação da simulação de *Monte Carlo* permitiu ampliar a base histórica de abastecimento da fita de aço em kg por meio da geração de cenários compatíveis com o comportamento operacional observado. A utilização de parâmetros extraídos dos próprios registros reais contribuiu para preservar a coerência temporal e estatística dos dados gerados, mantendo a aderência da base expandida à realidade do processo produtivo. Dessa forma, a base simulada passou a atuar como complemento analítico aos dados reais, fornecendo suporte às etapas posteriores de validação e modelagem preditiva.

4.2.4 Preparação da base de dados para o modelo *XGBoost*

Para a aplicação do modelo *XGBoost*, inicialmente foi realizada a preparação da base de dados, com o objetivo de organizar as informações em uma estrutura adequada para o treinamento do algoritmo. Essa etapa foi necessária para garantir que os registros fossem corretamente reconhecidos pelo ambiente *Python* e pudessem ser utilizados no processo de modelagem preditiva sem inconsistências de leitura, formatação ou interpretação dos dados.

A preparação da base também teve a finalidade de padronizar os registros antes da aplicação do modelo, assegurando que as informações temporais e operacionais fossem organizadas de forma compatível com as bibliotecas utilizadas no código. Nessa etapa, os dados foram ordenados cronologicamente, permitindo a construção da variável-alvo e das variáveis derivadas a partir do histórico de abastecimento da fita de aço em kg.

Dessa forma, a base foi estruturada para permitir que o modelo previsse o intervalo, em dias, até o próximo abastecimento. Para isso, foram consideradas variáveis temporais e informações derivadas de eventos anteriores, como intervalos passados, *lags* e médias móveis. Essa organização foi fundamental para evitar que o modelo utilizasse informações futuras durante o treinamento, reduzindo o risco de vazamento de alvo e tornando a previsão mais aderente à dinâmica real da operação.

A base utilizada foi importada a partir de uma planilha em *Excel* denominada BaseAco1.xlsx, considerando sua primeira aba. Após a leitura da planilha, os nomes das colunas foram padronizados por meio da remoção de espaços desnecessários, evitando inconsistências durante a execução do código e facilitando a identificação das variáveis utilizadas na modelagem, como é visto na figura 12.

Figura 12 - Carregamento da base de dados no *Python*

```
# =====  
# 4) LEITURA DA BASE  
# =====  
  
try:  
    df = pd.read_excel(arquivo_entrada, sheet_name=sheet_name)  
    print(f"Aba utilizada: {sheet_name}")  
except:  
    df = pd.read_excel(arquivo_entrada, sheet_name=0)  
    print("Aba Base_Final não encontrada. Primeira aba utilizada.")  
  
df.columns = [str(c).strip() for c in df.columns]  
  
print("\nColunas encontradas:")  
print(df.columns.tolist())
```

Fonte: o autor, 2026.

Em seguida, foram realizados tratamentos nos tipos de dados. A coluna referente ao recebimento no sistema foi convertida para formato temporal, permitindo a ordenação cronológica dos registros e o cálculo do intervalo até o próximo

abastecimento. Já as variáveis numéricas presentes na base, como quantidade em kg, consumo em kg/h, tempo de *setup* e custo, foram convertidas para formato numérico, garantindo a correta leitura dos dados pelo ambiente *Python*.

Embora quantidade, consumo e *setup* não tenham sido utilizadas como *features* diretas do modelo *XGBoost*, sua conversão foi necessária para a preparação da base, construção de variáveis históricas e conferência dos registros. Esse procedimento contribuiu para evitar inconsistências de formatação e assegurar que a base estivesse adequada às etapas posteriores de engenharia de atributos, treinamento e avaliação do modelo preditivo.

A Figura 13 apresenta o trecho do código responsável pela conversão dos campos da base de dados, evidenciando o tratamento aplicado às colunas numéricas e temporais antes da construção das variáveis utilizadas na modelagem.

Figura 13 - Conversão dos dados

```
# =====  
# 5) TRATAMENTO INICIAL  
# =====  
  
df["Recebimento no Sistema"] = pd.to_datetime(  
    df["Recebimento no Sistema"],  
    errors="coerce"  
)  
  
for col in ["Quantidade kg", "Consumo kg/h", "Setup min", "Custo"]:  
    if col in df.columns:  
        df[col] = df[col].apply(parse_numero_br)  
  
df = df.dropna(  
    subset=[  
        "Recebimento no Sistema",  
        "Quantidade kg",  
        "Consumo kg/h",  
        "Setup min"  
    ]  
)  
.reset_index(drop=True)  
  
if "Origem" not in df.columns:  
    df["Origem"] = "Real"  
  
df["Origem"] = df["Origem"].astype(str)  
df["Origem_norm"] = df["Origem"].apply(normalizar_texto)  
  
if "Custo" not in df.columns:  
    df["Custo"] = df["Quantidade kg"] * valor_kg  
  
df["Custo"] = df["Custo"].fillna(df["Quantidade kg"] * valor_kg)
```

Fonte: o autor, 2026.

Também foram mantidos apenas os registros válidos nas variáveis essenciais à preparação da base e à construção da previsão. Assim, linhas sem data de recebimento, quantidade em kg, consumo em kg/h ou tempo de *setup* foram removidas, evitando que valores ausentes comprometessem a organização dos dados, a criação das variáveis analíticas e o treinamento do modelo preditivo. A variável-alvo foi definida como o intervalo, em dias, entre um abastecimento e o próximo registro de abastecimento da fita de aço em kg. Para isso, os registros foram inicialmente ordenados em sequência temporal e, em seguida, foi identificada a data do próximo recebimento para cada evento observado.

$$Y = \text{Intervalo até o próximo abastecimento (dias)}$$

A equação abaixo foi utilizada para calcular o intervalo entre abastecimentos consecutivos na base histórica. Assim, o modelo passou a utilizar o histórico de reposições da fita de aço em kg para identificar padrões temporais associados a novos abastecimentos.

$$Y = (\text{Data do próximo abastecimento} - \text{Data do abastecimento atual})$$

Como as datas são manipuladas em formato temporal, o resultado da diferença é obtido diretamente em dias. Dessa forma, o modelo *XGBoost* passou a prever um evento operacional real, isto é, o intervalo esperado até o próximo abastecimento, em vez de estimar a duração teórica do consumo do material.

Figura 14 – Trecho do código responsável pelo cálculo da variável-alvo

```
if len(group_cols) > 0:
    df["proximo_recebimento"] = (
        df.groupby(group_cols)["Recebimento no Sistema"].shift(-1)
    )
else:
    df["proximo_recebimento"] = df["Recebimento no Sistema"].shift(-1)

df["intervalo_proximo_abastecimento_dias"] = (
    df["proximo_recebimento"] - df["Recebimento no Sistema"]
).dt.total_seconds() / 86400
```

Para a construção do modelo *XGBoost*, as variáveis de entrada foram definidas com base em informações temporais e históricas dos abastecimentos da fita de aço em kg. A proposta foi permitir que o modelo identificasse padrões na sequência dos eventos de reposição, sem utilizar diretamente variáveis que pudessem antecipar o resultado a ser previsto.

A partir da coluna *Recebimento no Sistema*, foram extraídos atributos como mês, dia do mês, dia da semana, semana do ano e ordem do evento. Essas variáveis permitiram representar a posição temporal de cada abastecimento na série analisada, fornecendo ao modelo informações sobre possíveis padrões de reposição ao longo do tempo.

Além das variáveis temporais, foram criadas variáveis históricas calculadas a partir de registros anteriores, como quantidade anterior, consumo anterior, *setup* anterior, intervalo anterior e médias móveis dos três últimos eventos. Essas variáveis foram estruturadas com deslocamento temporal, de modo que cada registro utilizasse apenas informações disponíveis antes do evento previsto.

Dessa forma, a base de modelagem foi organizada para preservar a coerência temporal da previsão, evitando que informações futuras fossem utilizadas no treinamento. A criação dessas variáveis é apresentada na Figura X, demonstrando a transformação dos registros de data, hora e histórico operacional em informações estruturadas para uso no modelo preditivo.

Figura 15 – Criação das *features* temporais

```
ts = df_model["Recebimento no Sistema"]

df_model["mes"] = ts.dt.month
df_model["dia_mes"] = ts.dt.day
df_model["dia_semana"] = ts.dt.weekday
df_model["semana_ano"] = ts.dt.isocalendar().week.astype(int)

if len(group_cols) > 0:
    df_model["ordem_evento"] = df_model.groupby(group_cols).cumcount() + 1
else:
    df_model["ordem_evento"] = np.arange(1, len(df_model) + 1)
```

Fonte: o autor, 2026.

A variável *Origem* foi utilizada para identificar a procedência dos registros da base, diferenciando dados reais e dados gerados por *Monte Carlo*. Sua função foi

permitir a rastreabilidade das informações e apoiar a separação dos conjuntos utilizados no treinamento e no teste do modelo.

No entanto, essa variável não foi empregada como entrada direta do modelo preditivo, evitando que o algoritmo aprendesse apenas a distinguir registros reais e simulados.

O quadro 3 apresenta as principais variáveis utilizadas na preparação da base de modelagem e suas respectivas justificativas.

Quadro 3 – Justificativas das variáveis

Variável	Tipo	Justificativa
Mês	Entrada temporal	Permite identificar padrões de abastecimento associados aos meses do ano.
Dia do mês	Entrada temporal	Auxilia na identificação de variações relacionadas à posição do evento dentro do mês.
Dia da semana	Entrada temporal	Permite observar possíveis diferenças no comportamento dos abastecimentos ao longo da semana.
Semana do ano	Entrada temporal	Representa a posição do evento no calendário anual, contribuindo para a leitura temporal da série.
Ordem do evento	Entrada sequencial	Indica a posição do abastecimento dentro da sequência histórica analisada.
Quantidade anterior kg	Entrada histórica	Representa a quantidade registrada no abastecimento anterior, permitindo ao modelo considerar o comportamento passado da operação.
Consumo anterior kg/h	Entrada histórica	Indica o consumo associado ao evento anterior, sem utilizar o consumo atual como entrada direta.
Setup anterior min	Entrada histórica	Representa o tempo de preparação observado no evento anterior, preservando a lógica temporal da base.
Intervalo anterior dias	Entrada histórica	Indica o intervalo observado entre abastecimentos anteriores, servindo como referência para a previsão do próximo intervalo.
Média móvel da quantidade 3	Entrada histórica	Representa a média das quantidades registradas nos três eventos anteriores.
Média móvel do consumo 3	Entrada histórica	Representa a média dos consumos observados nos três eventos anteriores.
Média móvel do setup 3	Entrada histórica	Representa a média dos tempos de preparação observados nos três eventos anteriores.
Média móvel do intervalo 3	Entrada histórica	Representa a média dos intervalos entre abastecimentos nos três eventos anteriores.
Origem	Variável de controle	Identifica se o registro é real ou gerado por <i>Monte Carlo</i> , sendo utilizada para rastreabilidade e separação dos conjuntos, mas não como entrada direta do modelo.
Intervalo até o próximo abastecimento	Variável-alvo Y	Representa o intervalo, em dias, entre o abastecimento atual e o próximo registro de abastecimento da fita de aço em kg.

Fonte: o autor, 2026.

Após a definição da variável-alvo e das variáveis explicativas, a base foi organizada entre variáveis numéricas e categóricas. Essa separação foi necessária para garantir que cada tipo de informação recebesse o tratamento adequado durante

o processamento do modelo, evitando inconsistências na leitura dos dados e permitindo a aplicação correta das etapas de preparação antes do treinamento.

As variáveis numéricas foram compostas pelos atributos temporais e históricos utilizados na previsão, como mês, dia do mês, dia da semana, semana do ano, ordem do evento, intervalo anterior e médias móveis. Esses atributos permitiram representar a sequência dos abastecimentos da fita de aço em kg e fornecer ao modelo informações sobre o comportamento passado da operação.

Entre as variáveis categóricas, foram considerados apenas os campos Insumo e Recurso, por apresentarem variação relevante na base. A variável Origem foi mantida apenas para controle e rastreabilidade dos registros, diferenciando dados reais e simulados, mas não foi utilizada como entrada do modelo para evitar influência da procedência dos dados no aprendizado.

A Figura 16 apresenta o trecho do código responsável pela definição das variáveis utilizadas na preparação da base para treinamento e avaliação do modelo *XGBoost*.

Figura 16 – Definição das variáveis numéricas e categóricas

```
# =====  
# 10) FEATURES PERMITIDAS  
# =====  
  
num_cols = [  
    "mes",  
    "dia_mes",  
    "dia_semana",  
    "semana_ano",  
    "ordem_evento",  
    "quantidade_anterior_kg",  
    "consumo_anterior_kgh",  
    "setup_anterior_min",  
    "intervalo_anterior_dias",  
    "media_movel_qtd_3",  
    "media_movel_consumo_3",  
    "media_movel_setup_3",  
    "media_movel_intervalo_3"  
]  
  
num_cols = [c for c in num_cols if c in df_model.columns]  
  
cat_cols = []  
  
for col in ["Insumo", "Recurso"]:  
    if col in df_model.columns and df_model[col].nunique() > 1:  
        cat_cols.append(col)  
  
if "Origem" in cat_cols:  
    cat_cols.remove("Origem")
```

Fonte: o autor, 2026.

No modelo *XGBoost*, a previsão da variável-alvo Y , correspondente ao intervalo até o próximo abastecimento da fita de aço em kg, é obtida pela soma dos resultados de múltiplas árvores de decisão construídas de forma sequencial. Cada árvore busca corrigir os erros produzidos pelas árvores anteriores, melhorando gradualmente a capacidade preditiva do modelo.

$$\hat{Y}_i = \sum_{k=1}^K f_k(X_i)$$

Onde:

- $\hat{Y}_i$ = valor previsto pelo modelo;
- K = número de árvores de decisão;
- f_k = k -ésima árvore construída pelo algoritmo;
- X_i = vetor de variáveis de entrada do registro analisado.

Considerando as variáveis utilizadas nesta pesquisa, o vetor de entrada do modelo pode ser representado da seguinte forma:

$$X_i = (M_i, DM_i, DS_i, SA_i, OE_i, IA_i, MMQ_i, MMC_i, MMS_i, MMI_i)$$

Em que:

Quadro 4 – Variáveis usadas no modelo

Símbolo	Variável	Descrição
M_i	Mês	Mês em que o abastecimento foi registrado.
DM_i	Dia do mês	Dia do mês associado ao evento de abastecimento.
DS_i	Dia da semana	Dia da semana correspondente ao registro analisado.
SA_i	Semana do ano	Semana do calendário anual em que ocorreu o abastecimento.
OE_i	Ordem do evento	Posição sequencial do abastecimento na série histórica.
IA_i	Intervalo anterior (dias)	Intervalo observado entre os dois abastecimentos anteriores.
MMQ_i	Média móvel da quantidade	Média das quantidades registradas nos três abastecimentos anteriores.
MMC_i	Média móvel do consumo	Média dos consumos observados nos três abastecimentos anteriores.

MMS_i	Média móvel do <i>setup</i>	Média dos tempos de <i>setup</i> observados nos três abastecimentos anteriores.
MMI_i	Média móvel do intervalo	Média dos intervalos entre abastecimentos observados nos três eventos anteriores.
Símbolo	Variável	Descrição
M_i	Mês	Mês em que o abastecimento foi registrado.
DM_i	Dia do mês	Dia do mês associado ao evento de abastecimento.
DS_i	Dia da semana	Dia da semana correspondente ao registro analisado.
SA_i	Semana do ano	Semana do calendário anual em que ocorreu o abastecimento.
OE_i	Ordem do evento	Posição sequencial do abastecimento na série histórica.
IA_i	Intervalo anterior (dias)	Intervalo observado entre os dois abastecimentos anteriores.
MMQ_i	Média móvel da quantidade	Média das quantidades registradas nos três abastecimentos anteriores.
MMC_i	Média móvel do consumo	Média dos consumos observados nos três abastecimentos anteriores.
MMS_i	Média móvel do <i>setup</i>	Média dos tempos de <i>setup</i> observados nos três abastecimentos anteriores.
MMI_i	Média móvel do intervalo	Média dos intervalos entre abastecimentos observados nos três eventos anteriores.

Fonte: o autor, 2026.

Ou seja, o *XGBoost* estima o intervalo em dias até o próximo abastecimento da fita de aço em kg a partir da combinação entre variáveis temporais e informações históricas dos abastecimentos anteriores.

No código, essa estrutura representa, de forma sintetizada, os principais grupos de variáveis utilizados na modelagem preditiva, enquanto a variável-alvo é representada por *target_col = "intervalo_proximo_abastecimento_dias"*.

4.2.5 Treinamento do *XGBoost* Padrão

Após a preparação da base e a definição da variável-alvo, foi realizada a aplicação do modelo *XGBoost*. Essa etapa teve como objetivo verificar se o algoritmo seria capaz de estimar o intervalo em dias até o próximo abastecimento da fita de aço em kg, a partir de variáveis temporais e históricas estruturadas anteriormente.

Para isso, os valores reais da variável-alvo foram comparados com os valores previstos pelo modelo, permitindo avaliar o grau de aderência entre o comportamento observado e o resultado estimado. Dessa forma, essa aplicação teve a função de validar tecnicamente o desempenho do modelo antes de sua utilização em uma etapa

prática com entrada manual, garantindo maior segurança na interpretação dos resultados gerados.

A divisão da base foi realizada respeitando a ordem cronológica dos registros reais. Inicialmente, 70% dos registros reais mais antigos foram destinados ao treinamento, enquanto os 30% reais mais recentes foram reservados para teste. Essa proporção foi adotada por representar uma prática recorrente em aplicações computacionais de modelagem preditiva, pois oferece ao algoritmo uma quantidade maior de dados para aprender os padrões históricos, ao mesmo tempo em que preserva uma parcela independente para avaliar sua capacidade de generalização.

No contexto desta pesquisa, a adoção da divisão 70/30 também se justifica pela quantidade limitada de eventos reais de abastecimento, já que a fita de aço em kg apresenta baixa frequência de ressuprimento mensal. Assim, utilizar uma parcela maior para treinamento foi necessário para que o modelo tivesse volume suficiente de registros para aprender padrões temporais e históricos, enquanto a parcela de teste permaneceu reservada para avaliar o desempenho em dados não utilizados no treinamento.

Os registros simulados por *Monte Carlo* foram incorporados apenas ao conjunto de treinamento, com o objetivo de ampliar a variabilidade disponível ao modelo sem contaminar a avaliação final. Essa decisão foi importante porque os dados simulados atuaram como complemento analítico para o aprendizado do algoritmo, mas não foram utilizados para medir seu desempenho final.

Dessa forma, o teste foi realizado exclusivamente com registros reais posteriores, permitindo verificar o desempenho do modelo em uma condição mais próxima da aplicação operacional. Essa estratégia preservou a lógica temporal da previsão, pois o modelo foi treinado com dados anteriores e avaliado em eventos reais futuros, evitando que informações geradas artificialmente influenciassem diretamente a validação dos resultados.

A estrutura dessa divisão é apresentada na figura, evidenciando a separação entre registros reais utilizados no treinamento, registros simulados incorporados ao treino e registros reais reservados para teste.

Figura 17 – Divisão temporal da base em treino e teste

```
# =====  
# 13) SPLIT 70% REAL + 30% SIMULADO  
# =====  
  
if len(real_df) < 5:  
    raise ValueError("Poucos registros reais para separar treino e teste.")  
  
cut_real = int(np.floor(len(real_df) * 0.70))  
  
real_train = real_df.iloc[:cut_real].copy()  
real_test = real_df.iloc[cut_real:].copy()  
  
n_real_train = len(real_train)  
n_sim_necessario = int(round(n_real_train * (30 / 70)))  
  
if len(sim_df) > 0:  
    n_sim_usado = min(len(sim_df), n_sim_necessario)  
  
    sim_train = sim_df.sample(  
        n=n_sim_usado,  
        random_state=random_state  
    ).copy()  
  
else:  
    n_sim_usado = 0  
    sim_train = pd.DataFrame(columns=df_model.columns)  
  
train_df = pd.concat([real_train, sim_train], ignore_index=True)  
train_df = train_df.sample(frac=1, random_state=random_state).reset_index(drop=True)  
  
test_df = real_test.copy().reset_index(drop=True)
```

Fonte: próprio autor, 2026

Após a divisão da base em conjuntos de treinamento e teste, foi estruturado um *pipeline* de pré-processamento com o objetivo de padronizar o tratamento das variáveis antes da aplicação do modelo *XGBoost*. Nessa etapa, as variáveis numéricas passaram por imputação pela mediana, estratégia utilizada para preencher eventuais valores ausentes sem provocar distorções relevantes na distribuição dos dados.

Já as variáveis categóricas, quando presentes, passaram por imputação do valor mais frequente e, em seguida, foram transformadas por meio da técnica *One-Hot Encoding*. Esse procedimento permitiu converter informações textuais em representações numéricas adequadas para leitura pelo algoritmo.

A adoção do *pipeline* foi necessária para garantir que todos os dados fossem processados de maneira padronizada antes do treinamento do modelo, reduzindo inconsistências e assegurando maior estabilidade na etapa de aprendizagem computacional. A organização do tratamento aplicado às variáveis numéricas e categóricas é apresentada na Figura 18, evidenciando a estrutura do *pipeline* de pré-processamento utilizado antes da modelagem preditiva. Essa etapa demonstra como cada tipo de variável recebeu um tratamento específico, garantindo que a base fosse transformada em uma estrutura adequada para o treinamento do modelo *XGBoost*.

Figura 18 – Tratamento das variáveis numéricas e categóricas

```
# =====  
# 15) PIPELINE DE PRÉ-PROCESSAMENTO  
# =====  
  
num_transformer = Pipeline(steps=[  
    ("imputer", SimpleImputer(strategy="median"))  
])  
  
cat_transformer = Pipeline(steps=[  
    ("imputer", SimpleImputer(strategy="most_frequent")),  
    ("onehot", OneHotEncoder(handle_unknown="ignore"))  
])  
  
transformers = []  
  
if len(num_cols) > 0:  
    transformers.append(("num", num_transformer, num_cols))  
  
if len(cat_cols) > 0:  
    transformers.append(("cat", cat_transformer, cat_cols))  
  
preprocess = ColumnTransformer(  
    transformers=transformers,  
    remainder="drop"  
)
```

Fonte: próprio autor, 2026

Esse tratamento foi necessário porque a base utilizada no modelo é composta por variáveis de diferentes naturezas, exigindo procedimentos específicos para cada tipo de dado. As variáveis numéricas, como atributos temporais, intervalos anteriores e médias móveis, foram tratadas para evitar problemas relacionados a valores ausentes, garantindo maior consistência no treinamento do modelo. Já as variáveis categóricas, como Insumo e Recurso, quando presentes e relevantes, precisaram ser convertidas para uma representação numérica, uma vez que o algoritmo não interpreta diretamente informações em formato textual.

Após essa etapa de pré-processamento, foi configurado o modelo *XGBoost* por meio da classe *XGBRegressor*, adequada para problemas de regressão. Essa escolha está alinhada ao objetivo da pesquisa, pois a variável-alvo corresponde a um valor numérico contínuo, representado pelo intervalo em dias até o próximo abastecimento da fita de aço em kg.

A configuração dos principais parâmetros utilizados no modelo é apresentada abaixo, demonstrando a etapa de definição do algoritmo empregado na modelagem preditiva.

Figura 19 – Configuração do modelo *XGBoost*

```
# =====  
# 16) MODELO 1 - XGBOOST DIRETO  
# =====  
  
model_direto = XGBRegressor(  
    n_estimators=120,  
    learning_rate=0.04,  
    max_depth=2,  
    subsample=0.90,  
    colsample_bytree=0.90,  
    reg_lambda=8.0,  
    reg_alpha=0.3,  
    min_child_weight=3,  
    objective="reg:squarederror",  
    random_state=random_state,  
    n_jobs=-1  
)  
  
pipe_direto = Pipeline(steps=[  
    ("prep", preprocess),  
    ("model", model_direto)  
)  
  
pipe_direto.fit(X_train, y_train)  
  
pred_direto = pipe_direto.predict(X_test)  
pred_direto = np.maximum(pred_direto, 0)  
  
mae_direto, rmse_direto, r2_direto = calcular_metricas(y_test, pred_direto)
```

Fonte: próprio autor, 2026

Os principais parâmetros definidos no modelo foram o número de árvores, a taxa de aprendizado, a profundidade máxima das árvores e os critérios de regularização. Esses parâmetros controlam a forma como o *XGBoost* aprende os padrões presentes nos dados, buscando equilibrar a capacidade preditiva do modelo com a redução de erros e o controle de possíveis ajustes excessivos aos dados de treinamento. Após essa configuração, o pré-processamento e o modelo foram integrados em um único pipeline, permitindo que todas as etapas, desde o tratamento das variáveis até o treinamento do algoritmo, fossem executadas de forma organizada e padronizada.

Em seguida, o modelo foi treinado com os dados de treinamento, possibilitando que o *XGBoost* aprendesse as relações entre as variáveis explicativas e a duração total até o esgotamento da fita de aço. Com o modelo treinado, foram geradas previsões para o conjunto de teste, as quais foram comparadas com os valores reais observados.

Ao final, foi gerada uma amostra das previsões realizadas no conjunto de teste, apresentando os valores reais, os valores previstos e os horários estimados de término do material.

Figura 20 – Amostra de previsões geradas pelo modelo *XGBoost* padrão

=====							
MOSTRA DAS PREVISÕES							
=====							
recebimento no Sistema	proximo_recebimento	Quantidade kg	Consumo kg/h	Setup min	Origem	intervalo_proximo_abastecimento_dias	p
2025-08-03 16:00:00	2025-08-08 09:00:00	1431.09	51.14	16.91	Real	4.708333	
2025-08-08 09:00:00	2025-08-15 10:00:00	1437.63	49.74	15.66	Real	7.041667	
2025-08-15 10:00:00	2025-08-21 15:00:00	1437.94	48.28	13.06	Real	6.208333	
2025-08-21 15:00:00	2025-08-26 09:00:00	1445.81	50.80	22.13	Real	4.750000	
2025-08-26 09:00:00	2025-09-02 00:00:00	1430.72	50.00	20.24	Real	6.625000	
2025-09-02 00:00:00	2025-09-08 00:00:00	1487.45	49.83	19.95	Real	6.000000	
2025-09-08 00:00:00	2025-09-13 00:00:00	1421.01	47.20	14.06	Real	5.000000	
2025-09-13 00:00:00	2025-09-19 15:00:00	1422.43	51.40	18.11	Real	6.625000	
2025-09-19 15:00:00	2025-09-25 09:00:00	1461.54	51.80	16.82	Real	5.750000	
2025-09-25 09:00:00	2025-09-30 20:00:00	1433.67	49.13	15.82	Real	5.458333	

Fonte: próprio autor, 2026

A tabela permite observar, registro por registro, a proximidade entre o intervalo real e o intervalo previsto até o próximo abastecimento. Além disso, a estimativa do próximo abastecimento transforma o resultado do modelo em uma informação operacional útil, pois permite visualizar uma provável data de reposição da fita de aço em kg com base no comportamento histórico identificado pelo *XGBoost*.

4.2.6 Aplicação do *XGBoost* com *input* de quantidade

Após o treinamento, avaliação e escolha do melhor resultado do modelo, foi incluída no mesmo código uma etapa complementar com entrada manual de quantidade. Essa funcionalidade permite que o usuário informe a quantidade disponível de fita de aço em kg, possibilitando a geração de uma estimativa prática associada ao próximo abastecimento.

O objetivo dessa aplicação não é utilizar a quantidade informada como variável direta de treinamento do modelo, mas transformar o resultado preditivo em uma informação de apoio à decisão operacional. Para isso, o código utiliza o modelo previamente treinado e as variáveis históricas mais recentes da base, estimando o intervalo provável até o próximo abastecimento e a respectiva data prevista de reposição.

Além disso, o código calcula uma autonomia operacional de referência com base na quantidade informada, no consumo mediano e no tempo de *setup* mediano da base. Essa informação funciona apenas como apoio interpretativo, enquanto a previsão principal permanece vinculada ao modelo treinado com variáveis temporais e históricas, isso é pode ser observado na figura 21.

Figura 21 – *Input* manual de quantidade e parâmetros operacionais

```
rodar_input = input("\nDeseja fazer uma simulação com input manual de quantidade? Digite S ou N: ")
sim_out = pd.DataFrame()
if rodar_input.strip().upper() == "S":
    entrada = input("Digite a quantidade disponível de fita de aço em kg. Ex: 1500 ou 1.500,00: ")
    novo_input_kg = parse_numero_br(entrada)
    ultimo = df_model.sort_values("Recebimento no Sistema").iloc[-1]
```

Fonte: próprio autor, 2026

A principal função dessa etapa está na criação de um registro de entrada para simulação operacional, construído a partir da quantidade informada pelo usuário. Esse registro recebe a data e hora atual e utiliza informações históricas recentes da base, como último abastecimento, intervalo anterior e médias móveis, para compor as variáveis necessárias à previsão.

A quantidade digitada é utilizada para calcular uma autonomia operacional de referência, com apoio do consumo mediano e do tempo de setup mediano observados na base. No entanto, a previsão do próximo abastecimento permanece vinculada ao modelo treinado, que utiliza variáveis temporais e históricas para estimar o intervalo provável até o próximo evento de reposição da fita de aço em kg.

Após a criação do registro de entrada, o modelo previamente treinado realiza a previsão do intervalo provável até o próximo abastecimento da fita de aço em kg. Essa etapa permite transformar as informações inseridas e os dados históricos recentes em uma estimativa operacional de reposição, indicando quando um novo abastecimento tende a ocorrer com base no comportamento aprendido pelo modelo.

Em seguida, o código calcula a data prevista para o próximo abastecimento a partir da data e hora de referência utilizada na simulação. Dessa forma, o resultado deixa de ser apenas um valor numérico em dias e passa a ser apresentado como uma informação operacional mais prática, capaz de apoiar o planejamento de ressuprimento e a tomada de decisão no processo produtivo.

A saída do modelo apresenta a quantidade informada, o horário de recebimento, a duração prevista até o esgotamento, o término previsto e o momento recomendado para acionamento do pedido, o seu output pode ser visto na figura 22:

Figura 22 – Saída da simulação com *input* manual

```
=====
SIMULAÇÃO COM INPUT MANUAL
=====
Modelo usado: Baseline - média móvel 3
Quantidade informada: 1.400,00 kg
Consumo usado apenas como referência: 50.48 kg/h
Setup usado apenas como referência: 17.89 min
Autonomia operacional de referência: 1.17 dias
Data/hora de referência: 03/06/2026 21:15:30

Intervalo previsto até o próximo abastecimento: 6.18 dias
Intervalo previsto em horas: 148.33 horas
Próximo abastecimento previsto: 10/06/2026 01:35:30
```

Fonte: próprio autor, 2026

4.2.7 Integração com *Power Apps*

A integração com o *Power Apps* foi estruturada como a etapa tecnológica responsável por transformar os resultados do modelo preditivo em uma ferramenta operacional de apoio à decisão. Após a consolidação da base de dados e a aplicação do modelo *XGBoost*, as informações foram organizadas para permitir que o aplicativo atuasse como interface de entrada, consulta e visualização dos dados relacionados à fita de aço em kg.

Dessa forma, o aplicativo não foi tratado apenas como uma tela de exibição, mas como parte de um fluxo digital integrado entre registro de dados, armazenamento, previsão e apoio à tomada de decisão. A proposta considera que o usuário possa consultar informações do abastecimento, visualizar resultados estimados e utilizar a previsão do próximo abastecimento como suporte ao planejamento operacional.

A lógica proposta considera a utilização de listas no *SharePoint* para armazenar as informações estruturadas. Uma lista pode ser destinada ao registro histórico de abastecimentos da fita de aço em kg, contendo informações como data do abastecimento, recurso produtivo associado, quantidade registrada e origem do dado. Outra lista pode ser utilizada para armazenar os resultados preditivos gerados pelo modelo, como intervalo previsto até o próximo abastecimento, data estimada de reposição e data de atualização da previsão.

Essa separação evita misturar os registros operacionais brutos com os resultados gerados pelo modelo preditivo, tornando o fluxo mais organizado e facilitando a consulta pelo *Power Apps*. Além disso, permite maior rastreabilidade das informações, uma vez que os dados históricos permanecem armazenados de forma independente dos resultados de previsão. Com isso, a solução digital proposta

contribui para aproximar a modelagem preditiva da rotina operacional, oferecendo uma forma mais prática de acompanhar o comportamento da fita de aço em kg e apoiar decisões de ressuprimento.

Quadro 5 – Estrutura das listas do *SharePoint* utilizadas na solução proposta

Lista no <i>SharePoint</i>	Finalidade
Base de Abastecimento	Armazenar dados reais, simulados e novos registros do <i>Power Apps</i>
Base de Previsão <i>XGBoost</i>	Armazenar a última previsão por insumo e recurso

Fonte: o autor, 2026.

Na lista de abastecimento, os principais campos considerados são aqueles necessários para registrar e acompanhar os eventos operacionais da fita de aço em kg, como data e hora do abastecimento, insumo, recurso produtivo associado, quantidade em kg, origem do dado e custo. Esses campos compõem a base operacional de acompanhamento e permitem manter o histórico dos registros utilizado como referência para atualização e análise dos abastecimentos.

Além desses campos, podem ser armazenadas informações auxiliares, como consumo médio e tempo de *setup*, quando aplicável, com a finalidade de apoiar análises operacionais e conferências internas. No entanto, a previsão do modelo permanece baseada nas variáveis temporais e históricas construídas a partir da sequência dos abastecimentos, como intervalos anteriores e médias móveis.

Já na lista de previsão, os campos são voltados à tomada de decisão, contendo informações como intervalo previsto até o próximo abastecimento, data estimada de reposição, status operacional, data da última atualização e observações relacionadas ao resultado gerado. Essa estrutura permite separar os registros operacionais da saída preditiva, facilitando a consulta das informações pelo *Power Apps* e apoiando o planejamento de ressuprimento da fita de aço em kg, conforme apresentado no quadro abaixo:

Quadro 6 - Campos utilizados para visualização das previsões no *Power Apps*

Campo	Aplicação
Insumo	Identificação do material analisado
Recurso	Máquina ou processo associado
Quantidade atual	Situação atual do insumo

Tempo previsto até ruptura	Saída principal do modelo
Lead time médio	Parâmetro de comparação para reposição
Margem de segurança	Critério adicional de decisão
Ponto de ressuprimento	Indicação de necessidade de abastecimento
Status	Classificação operacional: Normal, Atenção ou Abastecer
Data prevista de ruptura	Momento estimado de esgotamento
Data recomendada de abastecimento	Momento ideal para nova reposição

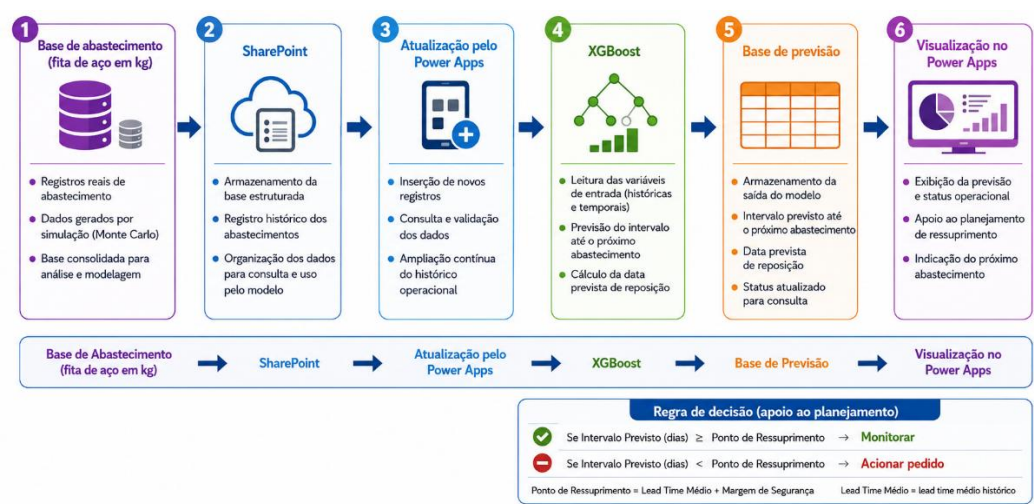
Fonte: o autor, 2026.

O fluxo de funcionamento foi representado pela sequência das principais etapas executadas pelo sistema, desde o registro dos dados operacionais até a geração da previsão do próximo abastecimento da fita de aço em kg.

Essa estrutura permite visualizar, de forma simplificada, como as informações percorrem todo o processo proposto, partindo do registro inicial dos abastecimentos, passando pelo armazenamento da base, organização dos dados e aplicação do modelo preditivo. A partir desse fluxo, os registros operacionais são transformados em variáveis temporais e históricas, que servem de entrada para o modelo *XGBoost* estimar o intervalo provável até o próximo evento de reposição.

Com isso, a solução busca transformar registros operacionais em informações úteis para o planejamento de ressuprimento, disponibilizando ao usuário final uma previsão mais objetiva sobre o próximo abastecimento da fita de aço em kg, conforme expresso pela figura 23.

Figura 23 - Fluxo de integração entre base de dados



Fonte: o autor, 2026.

Assim, a integração proposta permitiu conectar a etapa analítica do modelo preditivo ao ambiente operacional da produção. A base estruturada no *SharePoint* funcionou como repositório central dos dados, enquanto o *XGBoost* atuou na geração das previsões e o *Power Apps* na visualização dos resultados.

Com essa estrutura, as informações relacionadas ao intervalo previsto até o próximo abastecimento, à data estimada de reposição e ao status operacional puderam ser apresentadas de forma simples ao usuário. Dessa maneira, a solução proposta contribui para transformar registros históricos de abastecimento da fita de aço em kg em informações práticas para o planejamento de ressuprimento, apoiando decisões mais rápidas e orientadas por dados.

4.3 ESTRUTURAÇÃO DO SISTEMA E DO FLUXO DE PROCESSOS

4.3.1 Arquitetura do ecossistema para aplicação

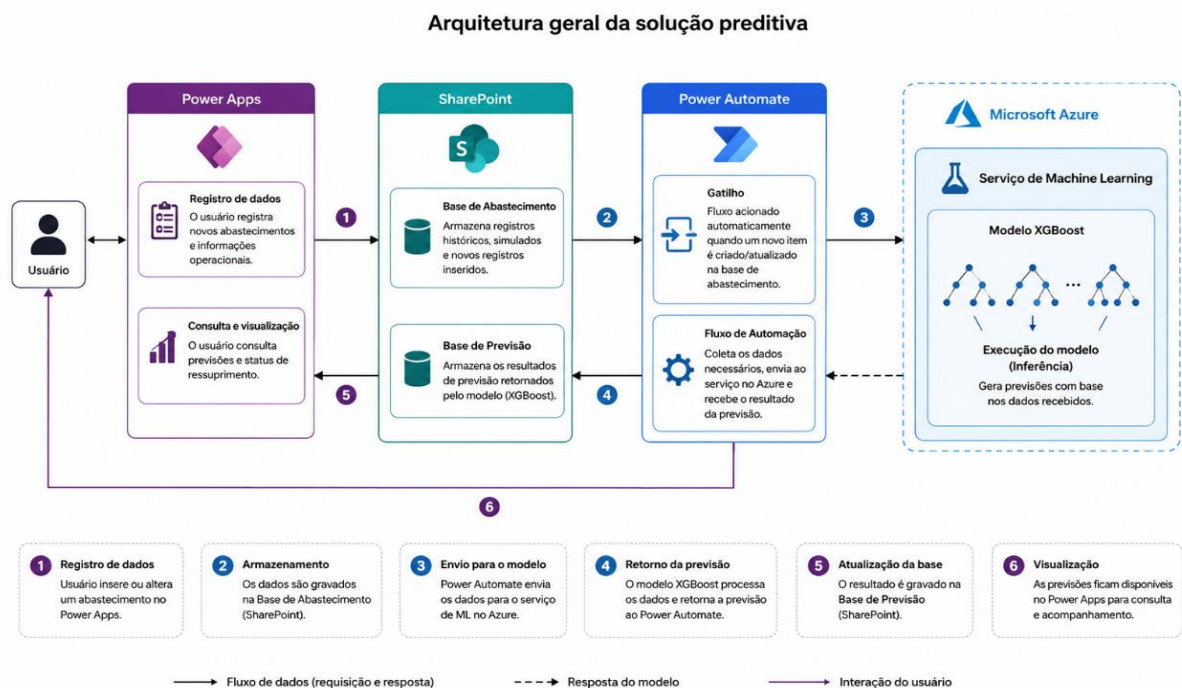
A arquitetura do ecossistema para aplicação da solução preditiva foi estruturada com o objetivo de integrar o registro operacional dos abastecimentos de fita de aço, o armazenamento dos dados, a execução do modelo preditivo e a visualização dos resultados em um aplicativo. Para isso, foi proposta uma integração entre *Power Apps*, *SharePoint*, *Power Automate* e *Azure*, permitindo que os dados inseridos pelo usuário sejam utilizados como entrada para a atualização das previsões de ressuprimento.

Nessa estrutura, o *Power Apps* funciona como a interface de interação com o usuário, possibilitando o registro de novas entradas de fita de aço e a consulta dos resultados gerados pela previsão. O *SharePoint* atua como repositório central dos dados, armazenando tanto os registros operacionais de abastecimento quanto os resultados retornados pelo modelo. Com isso, as informações permanecem organizadas em listas, facilitando o acesso, a atualização e a visualização dos dados utilizados no acompanhamento da disponibilidade do material.

O *Power Automate* é utilizado como mecanismo de automação do processo, acionando o fluxo sempre que um novo item é criado ou modificado na lista do *SharePoint*. Após essa alteração, os dados necessários são enviados ao serviço responsável pela execução do modelo preditivo. Esse fluxo reduz a necessidade de intervenção manual e permite que a previsão seja atualizada conforme os registros operacionais são inseridos pelo usuário.

A execução do modelo *XGBoost* foi prevista em ambiente externo ao *Power Apps*, por meio de um *endpoint* no Azure, responsável por receber os dados, realizar a inferência e retornar o resultado ao SharePoint. Em seguida, o *Power Apps* apresenta ao usuário final as informações de previsão de autonomia e o status de ressuprimento da fita de aço. Esse funcionamento integrado é sintetizado na Figura 24, que representa a comunicação entre *Power Apps*, SharePoint, *Power Automate*, Azure e o modelo preditivo.

Figura 24 - Fluxo geral da arquitetura proposta para aplicação da solução preditiva.



Fonte: o autor, 2026.

4.3.2 Estrutura das bases no *SharePoint*

O *SharePoint* foi definido como repositório central das informações utilizadas no protótipo, por permitir a criação de listas estruturadas e a integração direta com *Power Apps* e *Power Automate*. A base foi planejada em listas separadas, de modo a evitar a mistura entre os dados operacionais registrados pelo usuário, os parâmetros de apoio e os resultados gerados pelo modelo preditivo.

A organização proposta contempla quatro listas principais: Base de Abastecimento, Base de Previsão, Parâmetros Operacionais e Histórico de

Atualizações. A primeira representa a entrada principal do sistema, armazenando registros históricos, simulados e novos apontamentos realizados pelo usuário. A segunda concentra os resultados retornados pelo modelo *XGBoost*, incluindo informações de previsão e status de ressuprimento. As demais listas atuam como apoio ao processo, reunindo parâmetros utilizados na interpretação dos dados e registros de rastreabilidade das atualizações realizadas.

Quadro 7 – Estrutura proposta das listas no *SharePoint*

Lista	Finalidade	Principais campos
Base de Abastecimento	Armazenar registros históricos, simulados e novos registros do <i>Power Apps</i>	Data do abastecimento, insumo, recurso, quantidade, origem, usuário
Base de Previsão	Armazenar os resultados retornados pelo modelo <i>XGBoost</i>	Insumo, recurso, quantidade prevista, tempo até ressuprimento, data prevista, status
Parâmetros Operacionais	Armazenar informações de apoio ao cálculo e à interpretação	Insumo, recurso, consumo médio, <i>lead time</i> , margem de segurança
Histórico de Atualizações	Registrar alterações e execuções do fluxo automatizado	Data da atualização, tipo de alteração, usuário, status do processamento

Fonte: o autor, 2026.

O quadro 7 sintetiza a estrutura lógica das listas utilizadas no *SharePoint*, evidenciando a separação entre dados de entrada, resultados preditivos, parâmetros operacionais e histórico de processamento. Essa separação facilita a manutenção da base, melhora a organização das informações e permite que o *Power Apps* consulte apenas os dados necessários em cada tela do aplicativo. Como pode ser visto na figura abaixo:

Figura 25 – Simulação da estrutura das listas de dados no *SharePoint*

Base de Abastecimento							
ID	Data do Abastecimento	Insuno	Recurso Produtivo	Quantidade Abastecida	Origem do Registro	Usuário	Observações
1012	13/05/2025	Aço SAE 1020	Torno CNC 01	250,00	Real	joao.pereira	Abastecimento programado
1011	10/05/2025	Aço SAE 1020	Torno CNC 01	300,00	Simulado	sistema	Carga inicial
1010	07/05/2025	Óleo de Corte	Fresadora 02	20,00	Real	maria.souza	Reposição emergencial
1009	05/05/2025	Aço SAE 1045	Centro Usinagem 01	400,00	Real	joao.pereira	Abastecimento programado
1008	02/05/2025	Ferramenta Pastilha	Torno CNC 02	50,00	Simulado	sistema	Carga inicial

Base de Previsão							
ID	Insuno	Recurso Produtivo	Quantidade Prevista	Tempo até Ressuprimento (dias)	Data Prevista	Status de Ressuprimento	Última Atualização
2012	Aço SAE 1020	Torno CNC 01	278,45	7	20/05/2025	OK	13/05/2025 10:35
2011	Óleo de Corte	Fresadora 02	18,30	5	18/05/2025	Atenção	13/05/2025 10:35
2010	Aço SAE 1045	Centro Usinagem 01	120,50	3	16/05/2025	Crítico	13/05/2025 10:35
2009	Ferramenta Pastilha	Torno CNC 02	15,00	2	15/05/2025	Crítico	13/05/2025 10:35
2008	Aço Inox 304	Fresadora 01	60,20	6	19/05/2025	OK	13/05/2025 10:35

Parâmetros Operacionais						
ID	Insuno	Recurso Produtivo	Consumo Médio (unid/dia)	Lead Time (dias)	Margem de Segurança (%)	Observações
3001	Aço SAE 1020	Torno CNC 01	36,50	2	10%	Consumo baseado nos últimos 30 dias
3002	Óleo de Corte	Fresadora 02	3,60	1	15%	Reposição semanal
3003	Aço SAE 1045	Centro Usinagem 01	40,20	2	10%	Consumo baseado nos últimos 30 dias
3004	Ferramenta Pastilha	Torno CNC 02	7,50	2	20%	Alta variabilidade de consumo
3008	Aço Inox 304	Fresadora 01	10,10	3	10%	Consumo baseado nos últimos 30 dias

Histórico de Atualizações						
ID	Data da Atualização	Tipo de Alteração	Item Afetado	Usuário	Status do Processamento	Detalhes
4005	13/05/2025 10:35	Previsão Atualizada	Aço SAE 1020 - Torno CNC 01	sistema	Concluído	Previsão recalculada com sucesso
4004	13/05/2025 10:33	Novo Registro	Aço SAE 1020 - Torno CNC 01	joao.pereira	Concluído	Abastecimento registrado via Power Apps
4003	10/05/2025 14:20	Previsão Atualizada	Óleo de Corte - Fresadora 02	sistema	Concluído	Previsão recalculada com sucesso
4002	07/05/2025 09:15	Novo Registro	Óleo de Corte - Fresadora 02	maria.souza	Concluído	Reposição emergencial registrada
4001	05/05/2025 16:40	Previsão Atualizada	Aço SAE 1045 - Centro Usinagem 01	sistema	Parcial	Lead time atualizado manualmente

Fonte: o autor, 2026.

A Figura 25 ilustra como essa estrutura pode ser visualizada no ambiente do *SharePoint*, demonstrando a organização das listas e dos principais campos utilizados para armazenamento, previsão e acompanhamento das atualizações do sistema.

4.3.3 Fluxo de alimentação e atualização dos dados

A alimentação dos dados ocorre inicialmente a partir da carga da base histórica e da base simulada no *SharePoint*. Esses dados formam o conjunto inicial de registros utilizados pelo protótipo. Após essa etapa, novos abastecimentos podem ser registrados diretamente no *Power Apps*, ampliando a base de dados operacional e atualizando o histórico disponível para futuras previsões.

Quando um novo registro é criado ou modificado na lista de abastecimento, o *Power Automate* é acionado automaticamente. Em seguida, o fluxo coleta os dados do item registrado e envia as informações necessárias para o *endpoint* do modelo no *Azure*. Após o processamento, o resultado da previsão é retornado ao *Power Automate*, que atualiza a lista de previsão no *SharePoint*. Essa estrutura permite que o aplicativo trabalhe com dados atualizados sem exigir que o usuário execute manualmente o modelo preditivo. O processo de atualização fica automatizado,

enquanto o *Power Apps* permanece responsável pela interface de entrada e visualização.

4.3.4 Execução do modelo *XGBoost* no *Azure*

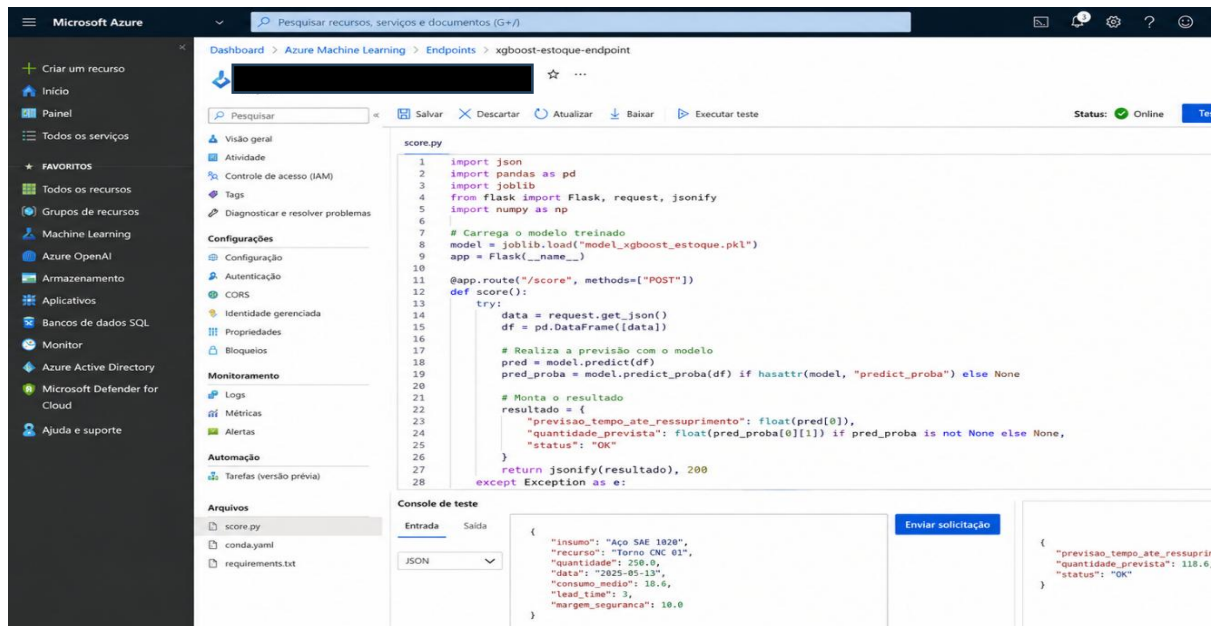
A execução do modelo *XGBoost* foi posicionada em ambiente *Azure* por exigir processamento externo ao *Power Apps*, que atua principalmente como interface de entrada, consulta e visualização dos dados. Nessa proposta, o modelo treinado é disponibilizado como serviço de inferência, recebendo do *Power Automate* informações como fita de aço, recurso produtivo, quantidade, data de abastecimento e parâmetros operacionais associados.

A partir desses dados, o *XGBoost* gera previsões relacionadas ao ressuprimento, como o tempo estimado até nova necessidade de material, a autonomia prevista da fita de aço e a classificação da condição operacional. Em seguida, o resultado é gravado no *SharePoint*, permitindo que o *Power Apps* consulte os campos atualizados e apresente as informações ao usuário final.

O Quadro 7 sintetiza a estrutura lógica das listas utilizadas no *SharePoint*, evidenciando a separação entre dados de entrada, resultados preditivos, parâmetros operacionais e histórico de processamento. Essa separação facilita a manutenção da base, melhora a organização das informações e permite que o *Power Apps* consulte apenas os dados necessários em cada tela do aplicativo. Além disso, essa organização contribui para reduzir inconsistências no registro das informações e torna o fluxo de dados mais claro para o usuário. Também permite maior rastreabilidade das previsões geradas, pois cada informação fica associada à sua respectiva etapa do processo.

Outro benefício dessa estrutura é a possibilidade de realizar atualizações e ajustes em listas específicas sem comprometer o funcionamento das demais partes do sistema. Dessa forma, o ambiente torna-se mais flexível para futuras adaptações e expansões da aplicação. A segmentação das informações também favorece a integração entre os diferentes componentes da solução, garantindo maior eficiência na comunicação entre a base de dados, o modelo preditivo e a interface utilizada pelos usuários. Como pode ser visto na figura abaixo:

Figura 26 – Tela do Microsoft Azure



Fonte: o autor, 2026.

4.3.5 Visualização dos resultados no *Power Apps*

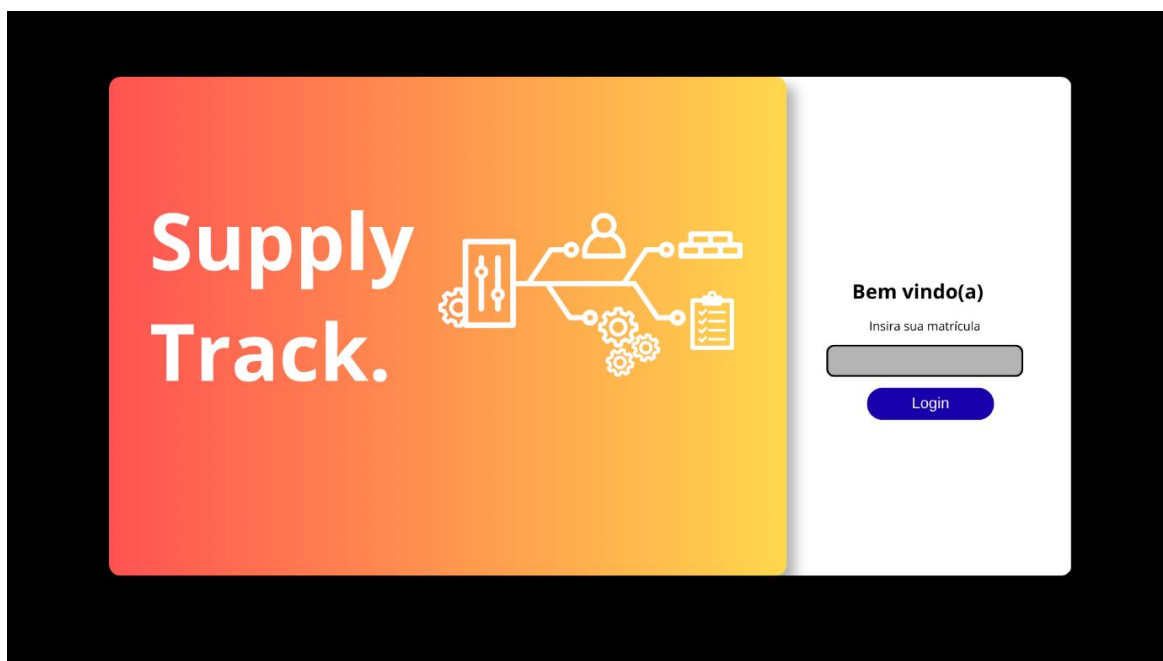
O *Power Apps* foi definido como a interface de interação do usuário com a solução proposta, concentrando as funções de acesso, registro de abastecimentos e acompanhamento da fita de aço. A estrutura do aplicativo foi desenvolvida com foco na simplicidade operacional, priorizando uma navegação intuitiva e a rápida visualização das informações mais relevantes para o usuário. Assim, o aplicativo atua como uma camada operacional de entrada e consulta dos dados, evitando que o usuário precise acessar diretamente as listas do SharePoint ou o ambiente Azure.

A primeira etapa da aplicação corresponde à tela inicial de acesso, responsável pela identificação do usuário por meio da matrícula. Após o login, o sistema direciona o usuário para a tela de registro de abastecimento, utilizada para inserção das informações operacionais. Nessa tela, podem ser registrados dados como fita de aço, recurso ou máquina, quantidade abastecida, data e hora do abastecimento e observações adicionais.

Além do registro dos dados, a interface apresenta indicadores resumidos relacionados ao status da fita de aço, à previsão de necessidade de ressuprimento, ao estoque atual, ao consumo previsto e ao último abastecimento realizado. Com isso,

o *Power Apps* permite que o usuário acompanhe, de forma simples e centralizada, tanto os dados operacionais inseridos quanto as informações geradas pela solução preditiva.

Figura 27 – Protótipo da tela inicial e tela de registro de abastecimento no *Power Apps*



Fonte: o autor, 2026.

A Figura 28 apresenta as principais telas desenvolvidas no *Power Apps* para interação com a solução proposta. À esquerda, observa-se a tela inicial de acesso ao sistema, responsável pela identificação do usuário.

À direita, é apresentada a tela de registro de abastecimento, utilizada para o lançamento das informações operacionais relacionadas à fita de aço e para a consulta dos indicadores disponibilizados pela aplicação, o que irá direcionar para a próxima tela, como pode ser visto na figura 28.

Figura 28 – Protótipo tela de registro de abastecimento no *Power Apps*

O protótipo da tela de registro de abastecimento no Power Apps, intitulada "Tela 2 Input de insumos", apresenta uma interface com uma barra lateral esquerda em azul escuro contendo o logotipo "Supply Track .", os botões "Abastecimento" (em laranja) e "Tracking" (em cinza), e um botão "Sair" (em vermelho) na base. O conteúdo principal da tela é dividido em duas seções principais. A seção central, intitulada "Registro de Abastecimento", contém campos de entrada para "Insumo" e "Recurso / Máquina", ambos com ícones de menu hambúrguer à direita. Abaixo, há campos para "Quantidade Abastecida" (com ícone de menu hambúrguer) e "Data e Hora do Abastecimento" (com ícone de calendário). Um campo de texto para "Comentários" está posicionado abaixo desses. Na base desta seção, há três botões: "Registrar" (em azul escuro), "Limpar" (em azul claro) e "Atualizar" (em azul claro). À direita, a seção "Status" exibe a data e hora da última atualização ("Última atualização às 00:00 DD/MM/AAAA") e o "Status do Insumo" como "Condição normal dentro do esperado" com um círculo verde. Abaixo, há dois cartões de informação: "Próxima Necessidade" mostrando "00h" com uma data "DD/MM/AAAA" e um ícone de calendário, e "Consumo WIP" mostrando "00%" com uma barra de progresso verde. Na base, há um cartão "Estoque Atual" mostrando "00Kg" e um cartão "Último abastecimento" com a data "00:00 DD/MM/AAAA" e a quantidade "0000,00 KG". Na base da tela, há o texto "Bem vindo (a) USUÁRIO" e um indicador "Sistema Online" com um ponto verde.

Fonte: o autor, 2026.

A terceira tela do aplicativo foi estruturada para o acompanhamento visual dos recursos produtivos, permitindo monitorar o comportamento da fita de aço em diferentes máquinas ou processos. Essa tela tem como objetivo apresentar, de forma resumida, a condição operacional do material, facilitando a identificação de situações que demandam atenção do usuário.

O status de ressuprimento foi definido como a principal forma de traduzir os resultados do modelo preditivo em uma orientação prática para o ambiente operacional. Essa classificação considera as previsões retornadas pelo *XGBoost*, o *lead time* da fita de aço e uma margem de segurança operacional. Com isso, o sistema pode indicar situações de condição normal, atenção ou possível ruptura, auxiliando o usuário na tomada de decisão relacionada ao acompanhamento e ao abastecimento do material.

5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Os resultados são apresentados considerando a sequência lógica desenvolvida ao longo da metodologia. Inicialmente, analisa-se a coerência da base expandida por simulação de *Monte Carlo*, observando se os cenários gerados preservam compatibilidade estatística e operacional com o comportamento identificado nos registros reais da fita de aço. Essa análise não tem como finalidade estabelecer uma comparação direta entre base real e base simulada, mas verificar a plausibilidade dos dados gerados como complemento metodológico para ampliar a representação dos cenários de consumo.

Em seguida, avalia-se o desempenho do modelo preditivo *XGBoost*, utilizando os indicadores de erro definidos no referencial teórico e aplicados na etapa de validação do modelo. Nessa etapa, a análise considera a variável-alvo corrigida, definida como o intervalo em dias até o próximo abastecimento, de modo que o modelo passe a estimar um evento futuro da operação, e não uma relação matemática previamente inserida na base.

Posteriormente, os resultados contemplam a aplicação prática da solução desenvolvida, com destaque para a forma como as informações geradas pelo modelo podem apoiar a decisão de ressuprimento da fita de aço. Assim, a apresentação dos resultados foi estruturada em três blocos principais: análise da coerência da base expandida por simulação, avaliação crítica do desempenho do modelo preditivo e demonstração da aplicação prática da solução como ferramenta de apoio à gestão industrial da fita de aço.

5.1 RESULTADOS DA SIMULAÇÃO DE MONTE CARLO

A etapa de simulação teve como finalidade complementar a base histórica de consumo da fita de aço a partir dos registros reais organizados no processo produtivo. Na versão revisada da base, os dados reais passaram a contemplar o período de janeiro de 2024 a março de 2026, ampliando a representação inicial da operação. A simulação de *Monte Carlo* foi utilizada como estratégia complementar para gerar cenários adicionais de abastecimento, permitindo ampliar a variabilidade observada e representar possíveis comportamentos futuros do consumo da fita de aço.

Inicialmente, foi organizada uma visualização contendo a base real e a base simulada, com o objetivo de apresentar a estrutura dos dados utilizados na análise.

Essa figura permite observar a continuidade entre os registros observados e os cenários simulados, mantendo campos como data de recebimento, quantidade em quilogramas, consumo médio, tempo de *setup* e origem do registro. Essa organização demonstra que a simulação não alterou a natureza da base, mas ampliou a quantidade de registros disponíveis para a análise e para a etapa posterior de modelagem.

Nesse sentido, a simulação não deve ser interpretada como substituição dos dados reais, mas como uma estratégia de complementação metodológica. Seu papel foi ampliar a representação de cenários operacionais possíveis, preservando a coerência com os parâmetros extraídos da base real e contribuindo para uma análise mais estruturada do comportamento da fita de aço ao longo do tempo.

Figura 29 – Estrutura da base real e da base expandida por simulação de *Monte Carlo*

Carlo

Recebimento no Sistema	Dia	Insumo	Recurso	Quantidade kg	Consumo kg/h	Setup min	Origem	Custo	Mês	Ano
01/10/2025 08:30	01/10/2025	Fita de Aço	Prensa Lâmina	1.477,33	50,00	25,00	Real	R\$ 7.682,12	out	2025
02/10/2025 09:00	02/10/2025	Fita de Aço	Prensa Lâmina	1.466,51	50,00	25,00	Real	R\$ 7.625,85	out	2025
07/10/2025 08:00	07/10/2025	Fita de Aço	Prensa Lâmina	1.433,56	50,00	25,00	Real	R\$ 7.454,51	out	2025
14/10/2025 09:00	14/10/2025	Fita de Aço	Prensa Lâmina	1.438,74	50,00	25,00	Real	R\$ 7.481,45	out	2025
20/10/2025 07:00	20/10/2025	Fita de Aço	Prensa Lâmina	1.474,86	50,00	25,00	Real	R\$ 7.669,27	out	2025
25/10/2025 07:00	25/10/2025	Fita de Aço	Prensa Lâmina	1.405,13	50,00	25,00	Real	R\$ 7.306,68	out	2025
02/11/2025 08:00	02/11/2025	Fita de Aço	Prensa Lâmina	1.470,10	50,00	25,00	Real	R\$ 7.644,52	nov	2025
09/11/2025 08:30	09/11/2025	Fita de Aço	Prensa Lâmina	1.468,15	50,00	25,00	Real	R\$ 7.634,38	nov	2025
15/11/2025 09:45	15/11/2025	Fita de Aço	Prensa Lâmina	1.448,37	50,00	25,00	Real	R\$ 7.531,52	nov	2025
23/11/2025 07:00	23/11/2025	Fita de Aço	Prensa Lâmina	1.438,93	50,00	25,00	Real	R\$ 7.482,44	nov	2025
28/11/2025 09:00	28/11/2025	Fita de Aço	Prensa Lâmina	1.437,31	50,00	25,00	Real	R\$ 7.474,01	nov	2025
05/12/2025 07:00	05/12/2025	Fita de Aço	Prensa Lâmina	1.435,70	50,00	25,00	Real	R\$ 7.465,64	dez	2025
10/12/2025 08:30	10/12/2025	Fita de Aço	Prensa Lâmina	1.447,17	50,00	25,00	Real	R\$ 7.525,28	dez	2025
15/12/2025 07:45	15/12/2025	Fita de Aço	Prensa Lâmina	1.450,23	50,00	25,00	Real	R\$ 7.541,20	dez	2025
27/12/2025 08:30	27/12/2025	Fita de Aço	Prensa Lâmina	1.452,75	50,00	25,00	Real	R\$ 7.554,30	dez	2025
02/01/2026 08:30	02/01/2026	Fita de Aço	Prensa Lâmina	1.495,26	50,00	25,00	Real	R\$ 7.775,35	jan	2026
03/01/2026 09:00	03/01/2026	Fita de Aço	Prensa Lâmina	1.467,80	50,00	25,00	Real	R\$ 7.632,56	jan	2026
05/01/2026 07:00	05/01/2026	Fita de Aço	Prensa Lâmina	1.456,54	50,00	25,00	Real	R\$ 7.574,01	jan	2026
12/01/2026 07:00	12/01/2026	Fita de Aço	Prensa Lâmina	1.492,57	50,00	25,00	Real	R\$ 7.761,36	jan	2026
21/01/2026 07:00	21/01/2026	Fita de Aço	Prensa Lâmina	1.432,81	50,00	25,00	Real	R\$ 7.450,61	jan	2026
24/01/2026 09:00	24/01/2026	Fita de Aço	Prensa Lâmina	1.428,30	50,00	25,00	Real	R\$ 7.427,16	jan	2026
30/01/2026 07:00	30/01/2026	Fita de Aço	Prensa Lâmina	1.455,99	50,00	25,00	Real	R\$ 7.571,15	jan	2026
02/02/2026 09:00	02/02/2026	Fita de Aço	Prensa Lâmina	1.414,82	50,00	25,00	Real	R\$ 7.357,06	fev	2026
09/02/2026 07:00	09/02/2026	Fita de Aço	Prensa Lâmina	1.413,36	50,00	25,00	Real	R\$ 7.349,47	fev	2026
15/02/2026 09:00	15/02/2026	Fita de Aço	Prensa Lâmina	1.450,75	50,00	25,00	Real	R\$ 7.543,90	fev	2026
23/02/2026 07:00	23/02/2026	Fita de Aço	Prensa Lâmina	1.448,28	50,00	25,00	Real	R\$ 7.531,06	fev	2026
27/02/2026 09:00	27/02/2026	Fita de Aço	Prensa Lâmina	1.479,65	50,00	25,00	Real	R\$ 7.694,18	fev	2026
03/03/2026 09:00	03/03/2026	Fita de Aço	Prensa Lâmina	1.456,94	50,00	25,00	Real	R\$ 7.576,09	mar	2026
05/03/2026 10:00	05/03/2026	Fita de Aço	Prensa Lâmina	1.450,71	50,00	25,00	Real	R\$ 7.543,69	mar	2026
08/03/2026 07:00	08/03/2026	Fita de Aço	Prensa Lâmina	1.449,84	50,00	25,00	Real	R\$ 7.539,17	mar	2026
14/03/2026 08:00	14/03/2026	Fita de Aço	Prensa Lâmina	1.435,18	50,00	25,00	Real	R\$ 7.462,94	mar	2026
19/03/2026 07:00	19/03/2026	Fita de Aço	Prensa Lâmina	1.407,68	50,00	25,00	Real	R\$ 7.319,94	mar	2026
25/03/2026 09:00	25/03/2026	Fita de Aço	Prensa Lâmina	1.431,02	50,00	25,00	Real	R\$ 7.441,30	mar	2026
30/03/2026 07:00	30/03/2026	Fita de Aço	Prensa Lâmina	1.460,76	50,00	25,00	Real	R\$ 7.595,95	mar	2026
11/04/2026 08:00	11/04/2026	Fita de Aço	Prensa Lâmina	1.425,06	50,00	25,00	Simulado	R\$ 7.410,31	abr	2026
15/04/2026 08:00	15/04/2026	Fita de Aço	Prensa Lâmina	1.471,69	50,00	25,00	Simulado	R\$ 7.652,79	abr	2026
29/04/2026 07:00	29/04/2026	Fita de Aço	Prensa Lâmina	1.428,95	50,00	25,00	Simulado	R\$ 7.430,54	abr	2026
29/04/2026 08:00	29/04/2026	Fita de Aço	Prensa Lâmina	1.414,17	50,00	25,00	Simulado	R\$ 7.353,68	abr	2026
01/05/2026 07:00	01/05/2026	Fita de Aço	Prensa Lâmina	1.455,11	50,00	25,00	Simulado	R\$ 7.586,57	mai	2026
02/05/2026 09:45	02/05/2026	Fita de Aço	Prensa Lâmina	1.446,56	50,00	25,00	Simulado	R\$ 7.522,11	mai	2026
06/05/2026 08:30	06/05/2026	Fita de Aço	Prensa Lâmina	1.450,85	50,00	25,00	Simulado	R\$ 7.544,42	mai	2026
12/05/2026 10:00	12/05/2026	Fita de Aço	Prensa Lâmina	1.428,67	50,00	25,00	Simulado	R\$ 7.429,08	mai	2026
21/05/2026 07:00	21/05/2026	Fita de Aço	Prensa Lâmina	1.484,82	50,00	25,00	Simulado	R\$ 7.721,06	mai	2026
29/05/2026 10:00	29/05/2026	Fita de Aço	Prensa Lâmina	1.447,70	50,00	25,00	Simulado	R\$ 7.578,04	mai	2026

Fonte: o autor, 2026.

A partir dessa estrutura, foram selecionados indicadores estatísticos capazes de verificar a coerência da base expandida. Para essa análise, consideraram-se a média, o desvio padrão, o valor mínimo, o valor máximo e a quantidade de registros. A média foi utilizada para observar a tendência central das quantidades abastecidas, enquanto o desvio padrão permitiu avaliar o grau de dispersão dos dados em relação a essa tendência. Já os valores mínimo e máximo contribuíram para verificar

se os cenários gerados permaneceram dentro de uma faixa compatível com os limites observados na operação.

O quadro abaixo apresenta os principais indicadores obtidos para a base real e para a base expandida por simulação.

Tabela 1 – Principais indicadores e seus resultados

Indicador	Base Real	Base Simulada
Média (kg)	1448,75	1449,08
Desvio padrão	18,40	16,07
Valor mínimo (kg)	1408,34	1413,18
Valor máximo (kg)	1492,20	1486,18
Quantidade de registros	137	135

Fonte: o autor, 2026.

Os resultados apresentados indicam que a base simulada manteve comportamento compatível com os registros observados. A média da base simulada foi de 1449,08 kg, valor muito próximo da média da base real, que foi de 1448,75 kg. Essa proximidade demonstra que a simulação preservou a tendência central das quantidades abastecidas, evitando deslocamentos artificiais que poderiam comprometer a coerência da base.

Em relação à variabilidade, o desvio padrão da base simulada apresentou valor de 16,07 kg, enquanto a base real registrou 18,40 kg. Essa diferença indica que os cenários simulados apresentaram dispersão ligeiramente menor, porém ainda compatível com o comportamento operacional analisado. Esse resultado é aceitável, pois a finalidade da simulação não foi reproduzir exatamente cada oscilação observada, mas gerar cenários plausíveis dentro de uma faixa coerente com o histórico disponível.

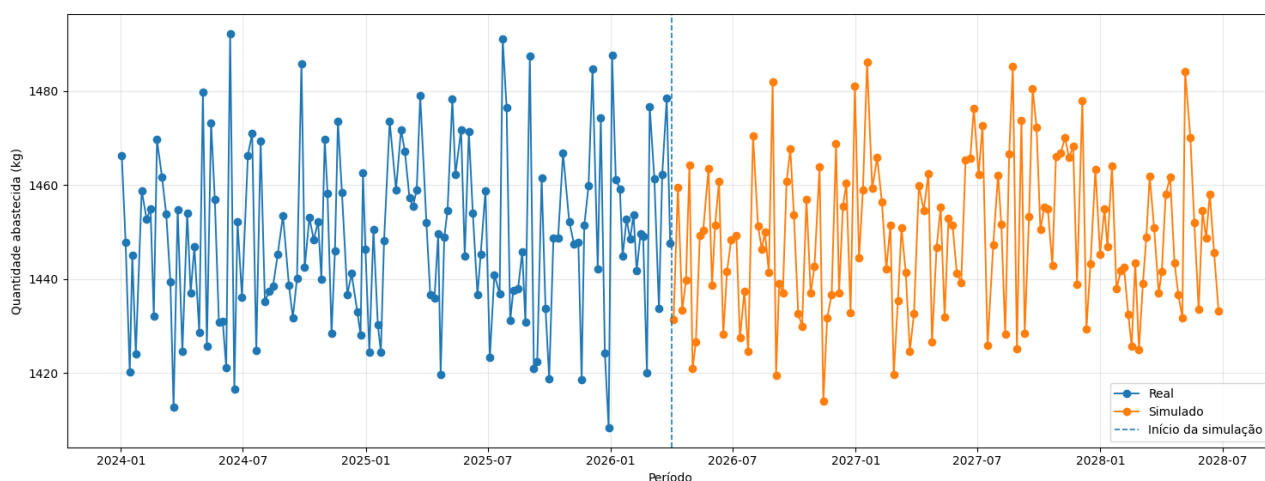
Os valores mínimo e máximo também reforçam essa compatibilidade. Na base real, a menor quantidade registrada foi de 1408,34 kg e a maior foi de 1492,20 kg. Na base simulada, os valores permaneceram entre 1413,92 kg e 1486,18 kg. Essa faixa demonstra que os registros gerados não extrapolaram os limites operacionais

identificados, mantendo-se próximos ao comportamento observado no período analisado.

Outro ponto relevante está na quantidade de registros. A base real contou com 137 ocorrências, enquanto a base simulada apresentou 135 registros, totalizando 272 registros. Esse volume demonstra a função complementar da simulação de *Monte Carlo*, que permitiu ampliar a representação da base sem descaracterizar sua estrutura. Com isso, a expansão contribuiu para oferecer maior volume de informações à etapa posterior de treinamento do modelo preditivo.

O comportamento temporal dos registros também foi analisado por meio da representação gráfica das quantidades reais e simuladas ao longo do período considerado. Essa visualização permitiu observar a distribuição dos dados no tempo e verificar se os cenários gerados mantiveram uma sequência compatível com a dinâmica de abastecimento da fita de aço no processo produtivo.

Figura 30 – Comportamento temporal do consumo real e simulado de fitas de aço



Fonte: o autor, 2026.

A figura acima evidencia que os dados simulados acompanharam uma trajetória compatível com a faixa de variação observada na base real, sem apresentar oscilações excessivas ou valores incompatíveis com o processo analisado. Essa representação reforça que a simulação foi aplicada como estratégia de complementação da base, e não como substituição dos registros coletados.

Em síntese, os resultados da base expandida por simulação de *Monte Carlo* indicam coerência estatística e plausibilidade operacional. A proximidade entre as médias, a manutenção de limites compatíveis e a dispersão controlada dos dados demonstram que os cenários gerados preservaram características relevantes do comportamento observado. Portanto, a base expandida mostrou-se adequada para apoiar a etapa seguinte da pesquisa, voltada ao treinamento e à avaliação crítica do modelo preditivo.

5.2 RESULTADOS DA CONFIGURAÇÃO PRELIMINAR DO MODELO XGBOOST

Após a preparação da base analítica e a definição inicial da variável-alvo, foi realizado o treinamento de uma configuração preliminar do modelo *XGBoost*. Essa versão foi utilizada como primeira etapa de avaliação da modelagem, antes da reestruturação posterior da variável-alvo e das variáveis explicativas. O objetivo inicial foi verificar se o algoritmo seria capaz de estimar a duração total até o esgotamento da fita de aço a partir das variáveis operacionais e temporais estruturadas nas etapas anteriores.

Para essa aplicação, a base foi dividida em dados de treinamento e dados de teste. Os registros de treinamento foram utilizados para que o modelo aprendesse as relações entre as variáveis explicativas e a duração total estimada do material, enquanto os registros de teste permitiram avaliar o comportamento do algoritmo em dados não utilizados durante o treinamento. Essa separação foi importante para observar a aderência inicial entre os valores previstos e os valores calculados para a variável-alvo.

Com o modelo treinado, foram geradas previsões para o conjunto de teste. Os valores previstos foram analisados em relação aos valores observados da variável-alvo, permitindo verificar a proximidade entre a duração calculada e a duração estimada pelo algoritmo. Os principais indicadores obtidos nessa configuração preliminar são apresentados na tabela 2.

Tabela 2 – Resultados dos indicadores da configuração preliminar do XGBoost

Indicador	Resultado
-----------	-----------

Total de registros	272
Registros de treinamento	190
Registros de teste	82
RMSE (h)	0,134
MAE (h)	0,100
R²	0,984

Fonte: próprio autor, 2026

Os resultados apresentados indicaram desempenho aparentemente elevado da configuração preliminar. O RMSE de 0,134 h demonstra que os desvios mais relevantes permaneceram baixos quando analisados em horas, enquanto o MAE de 0,100 h indica reduzido afastamento médio entre as previsões geradas e os valores observados. Em um primeiro momento, esses resultados sugeriam que o modelo havia conseguido estimar a duração total até o esgotamento da fita de aço com baixo erro médio no conjunto de teste.

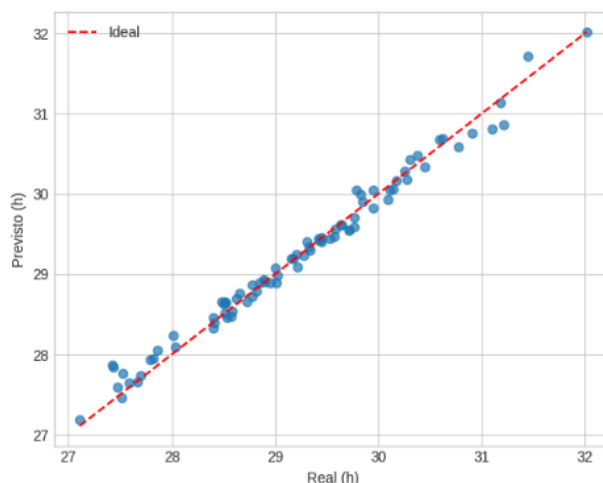
O valor de R² igual a 0,984 também reforçava essa percepção inicial. Esse indicador demonstrava que grande parte da variação da variável-alvo estava sendo representada pelo *XGBoost*, sugerindo elevada aderência entre os valores previstos e observados. Em termos práticos, os resultados indicavam potencial para apoiar estimativas relacionadas ao tempo de autonomia do material.

Entretanto, a magnitude desses indicadores motivou uma análise mais detalhada da estrutura da modelagem. Embora resultados elevados sejam desejáveis em aplicações preditivas, valores excessivamente próximos da perfeição podem indicar a necessidade de uma avaliação metodológica mais aprofundada, especialmente em problemas operacionais sujeitos a variações de abastecimento, consumo e comportamento temporal.

Durante essa revisão, verificou-se que a variável-alvo utilizada na configuração preliminar era calculada diretamente a partir da quantidade disponível, do consumo médio horário e do tempo de *setup*. Ao mesmo tempo, essas mesmas variáveis também compunham o conjunto de entradas fornecidas ao algoritmo. Essa condição caracterizava um vazamento de dados, uma vez que o modelo recebia como variáveis explicativas os próprios elementos utilizados para construir o valor que deveria ser previsto.

Na sequência, foi analisado o gráfico de aderência entre os valores observados e os valores previstos pelo modelo nessa configuração preliminar. Essa visualização permite verificar, de forma gráfica, a proximidade entre a duração calculada e a duração estimada no conjunto de teste.

Figura 31 – Gráfico de aderência entre valores observados e previstos na configuração preliminar do XGBoost



Fonte: próprio autor, 2026

No gráfico, a linha tracejada representa a condição ideal, em que os valores previstos seriam iguais aos valores observados. A concentração dos pontos próxima a essa linha indica elevada aderência entre as previsões realizadas e os registros avaliados. Esse comportamento visual confirma os resultados numéricos apresentados no Quadro 9, especialmente os baixos valores de erro e o elevado coeficiente de determinação.

Embora essa aderência demonstre a capacidade do modelo em reproduzir os padrões presentes na base utilizada, a análise metodológica indicou que parte desse desempenho estava associada à própria estrutura das variáveis empregadas na modelagem.

Dessa forma, os resultados dessa configuração preliminar foram tratados como etapa exploratória da modelagem, e não como evidência final de desempenho preditivo. A partir dessa constatação, tornou-se necessária a reestruturação do modelo, com a redefinição da variável-alvo para o intervalo em dias até o próximo abastecimento da fita de aço e a exclusão das variáveis diretamente associadas ao cálculo anterior.

5.3 RESULTADOS DA MODELAGEM REESTRUTURADA DO XGBOOST

Após a revisão metodológica da configuração preliminar, foi desenvolvida uma nova estrutura de modelagem para o algoritmo *XGBoost*. Nessa etapa, a variável-alvo passou a representar o intervalo em dias até o próximo abastecimento da fita de aço, permitindo que o modelo estimasse um evento futuro da operação a partir de informações históricas disponíveis antes da ocorrência do abastecimento seguinte.

Para essa nova configuração, foram removidas as variáveis diretamente associadas ao cálculo da duração do material, evitando que o modelo utilizasse informações que determinavam previamente o valor a ser previsto. Em seu lugar, passaram a ser consideradas variáveis temporais e históricas, como mês, ano, dia da semana, registros anteriores de abastecimento (*lags*) e médias móveis calculadas a partir dos eventos passados.

A base utilizada na modelagem foi composta por 272 registros, sendo formada por dados reais e simulados obtidos na etapa de expansão da base. Para a avaliação do modelo, os registros foram organizados de forma temporal, preservando a sequência cronológica dos eventos e permitindo que o treinamento fosse realizado com informações anteriores ao período utilizado para teste.

Com o treinamento concluído, foram geradas previsões para o conjunto de teste e calculados os indicadores de desempenho do modelo, como pode ser visto na tabela 3.

Tabela 3 – Resultados dos indicadores da configuração sem vazamentos de dados do *XGBoost*

Indicador	Resultado
Registros após criação das <i>features</i>	261
Total de <i>features</i> utilizadas	34
Registros de treinamento	180
Registros de teste	78
RMSE (dias)	0,835
MAE (dias)	0,646
R ²	0,001

Fonte: próprio autor, 2026

Na modelagem reestruturada, a avaliação deixou de considerar uma entrada manual de quantidade como principal referência de previsão. Em vez disso, o modelo passou a estimar o intervalo em dias até o próximo abastecimento da fita de aço,

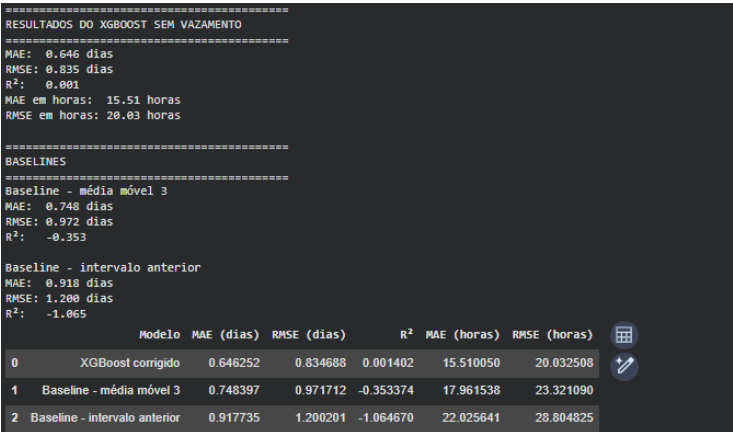
utilizando informações temporais e históricas disponíveis antes da ocorrência do evento previsto. Essa alteração tornou a aplicação mais adequada ao objetivo da pesquisa, pois passou a tratar o ressuprimento como um evento futuro da operação.

Os resultados obtidos indicaram MAE de 0,646 dias e RMSE de 0,835 dias, o que corresponde, respectivamente, a aproximadamente 15,51 horas e 20,03 horas. Esses valores demonstram que, em média, as previsões apresentaram diferença inferior a um dia em relação aos intervalos observados no conjunto de teste. Embora o R^2 tenha apresentado valor próximo de zero, esse resultado deve ser interpretado com cautela, pois a variável-alvo envolve oscilações operacionais difíceis de explicar apenas com variáveis temporais e históricas.

Diferentemente da configuração preliminar, cuja elevada aderência estava associada à presença de variáveis diretamente relacionadas ao cálculo da duração do material, a modelagem reestruturada foi conduzida sem vazamento de dados. Dessa forma, os resultados obtidos representam uma avaliação mais realista da capacidade preditiva do algoritmo, ainda que com indicadores mais modestos.

Além dos indicadores numéricos, foi gerado o gráfico de aderência entre os valores observados e os valores previstos no conjunto de teste. Essa visualização permite analisar o comportamento do modelo em relação aos intervalos reais de abastecimento, demonstrando de forma gráfica a proximidade e as limitações das previsões geradas.

Figura 32 – Output dos indicadores da configuração sem vazamentos de dados



```
=====
RESULTADOS DO XGBOOST SEM VAZAMENTO
=====
MAE: 0.646 dias
RMSE: 0.835 dias
R²: 0.001
MAE em horas: 15.51 horas
RMSE em horas: 20.03 horas

=====
BASELINES
=====
Baseline - média móvel 3
MAE: 0.748 dias
RMSE: 0.972 dias
R²: -0.353

Baseline - intervalo anterior
MAE: 0.918 dias
RMSE: 1.200 dias
R²: -1.065
```

	Modelo	MAE (dias)	RMSE (dias)	R²	MAE (horas)	RMSE (horas)
0	XGBoost corrigido	0.646252	0.834688	0.001402	15.510050	20.032508
1	Baseline - média móvel 3	0.748397	0.971712	-0.353374	17.961538	23.321090
2	Baseline - intervalo anterior	0.917735	1.200201	-1.064670	22.025641	28.804825

Fonte: próprio autor, 2026

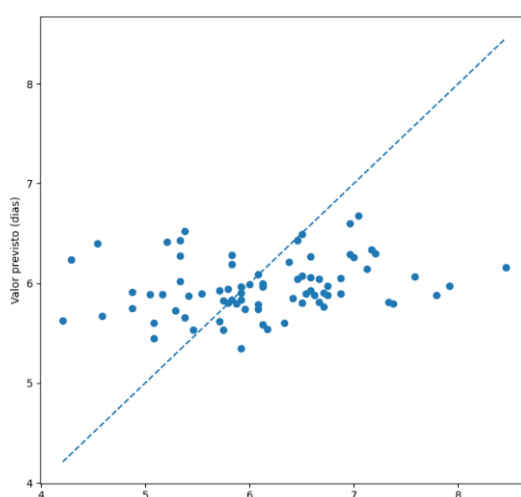
A visualização dos resultados permite observar o comportamento do modelo reestruturado diante dos dados de teste, evidenciando tanto sua capacidade de acompanhar parte da dinâmica dos intervalos de abastecimento quanto suas

limitações preditivas. Diferentemente da configuração preliminar, os pontos não apresentam aderência perfeita à linha ideal, o que indica um comportamento mais compatível com uma previsão operacional real, sujeita a variações e incertezas.

Esse resultado demonstra que a modelagem corrigida passou a representar um problema preditivo mais adequado, pois o modelo deixou de reproduzir uma relação matemática previamente definida e passou a estimar o intervalo até o próximo abastecimento com base em informações históricas. No entanto, os indicadores obtidos também mostram que a base disponível ainda impõe restrições ao desempenho do algoritmo, principalmente em razão do volume limitado de registros e da variabilidade dos eventos de abastecimento.

Assim, os resultados não devem ser interpretados como uma previsão plenamente consolidada para aplicação operacional imediata, mas como uma etapa metodológica válida para estruturar o uso de aprendizado de máquina no apoio ao planejamento de ressuprimento da fita de aço. Com uma base histórica mais ampla, maior diversidade de condições operacionais e continuidade na alimentação dos registros, espera-se que o modelo possa capturar padrões com maior consistência e apresentar estimativas mais precisas. Na figura abaixo é a evidência de aderência desses valores.

Figura 33 – Gráfico de aderência entre valores observados e previstos sem vazamentos de dados do *XGBoost*



Fonte: próprio autor, 2026

5.3 SÍNTESE DOS RESULTADOS DOS INDICADORES

Após a apresentação dos resultados da base expandida por simulação de *Monte Carlo* e da modelagem reestruturada do *XGBoost* sem vazamento de dados, foi elaborada uma síntese dos principais indicadores obtidos. A tabela 4 reúne os resultados mais relevantes da pesquisa, permitindo visualizar, de forma integrada, a coerência da base simulada e o desempenho do modelo adotado na análise final.

Tabela 4 – Síntese dos principais resultados obtidos

Etapa analisada	Indicador	Resultado	Interpretação
Base expandida por <i>Monte Carlo</i>	Média da base real	1448,75 kg	Representa a tendência central dos registros observados
Base expandida por <i>Monte Carlo</i>	Média da base simulada	1449,08 kg	Indica proximidade com o comportamento observado
Base expandida por <i>Monte Carlo</i>	Desvio padrão real	18,40 kg	Demonstra a variabilidade dos registros reais
Base expandida por <i>Monte Carlo</i>	Desvio padrão simulado	16,07 kg	Aponta dispersão controlada nos cenários simulados
Base expandida por <i>Monte Carlo</i>	Registros reais e simulados	137 e 135	Evidencia a complementação da base analítica
Modelagem reestruturada do <i>XGBoost</i>	Registros após criação das <i>features</i>	261	Base efetivamente utilizada após criação de <i>lags</i> e médias móveis
Modelagem reestruturada do <i>XGBoost</i>	Total de <i>features</i> utilizadas	34	Representa o conjunto de variáveis temporais e históricas usadas no modelo
Modelagem reestruturada do <i>XGBoost</i>	Treinamento	180 registros	Base composta por 70% de registros reais e 30% de registros simulados
Modelagem reestruturada do <i>XGBoost</i>	Teste	78 registros	Base utilizada para avaliar o desempenho em período posterior
Modelagem reestruturada do <i>XGBoost</i>	MAE	0,646 dias	Indica erro médio inferior a um dia
Modelagem reestruturada do <i>XGBoost</i>	RMSE	0,835 dias	Mostra maior penalização dos desvios mais elevados
Modelagem reestruturada do <i>XGBoost</i>	R ²	0,001	Demonstra capacidade explicativa global limitada da variável-alvo
Modelagem reestruturada do <i>XGBoost</i>	MAE em horas	15,51 h	Conversão do erro médio para interpretação operacional
Modelagem reestruturada do <i>XGBoost</i>	RMSE em horas	20,03 h	Conversão do erro quadrático médio para interpretação operacional

Fonte: próprio autor, 2026

A síntese apresentada demonstra que a simulação de *Monte Carlo* contribuiu para complementar a base de análise sem descaracterizar o comportamento operacional observado. A proximidade entre as médias da base real e da base simulada, associada à manutenção de valores mínimos, máximos e dispersão compatíveis, reforça a coerência estatística dos cenários gerados.

Em relação à modelagem reestruturada do *XGBoost*, os indicadores apontaram um desempenho mais moderado, porém metodologicamente mais adequado. O MAE de 0,646 dias e o RMSE de 0,835 dias indicam erro médio inferior a um dia no conjunto de teste, enquanto o R^2 de 0,001 demonstra limitada capacidade explicativa global da variável-alvo. Esse resultado evidencia que o modelo conseguiu gerar estimativas próximas em termos médios, mas ainda apresenta limitações para explicar toda a variabilidade dos intervalos até o próximo abastecimento.

Esses resultados devem ser interpretados considerando as limitações da base disponível e a complexidade do evento previsto. A previsão do intervalo até o próximo abastecimento depende de fatores operacionais que nem sempre estão plenamente representados na base, como ritmo de produção, decisões de antecipação de ressuprimento, variações de demanda e condições específicas da rotina fabril.

Em resumo, os indicadores analisados mostram que a proposta apresentou coerência metodológica e evolução na estruturação da modelagem preditiva. A simulação fortaleceu a base analítica, e a reestruturação do *XGBoost* permitiu avaliar o modelo sem vazamento de dados, oferecendo uma leitura mais realista da capacidade preditiva. Embora os resultados ainda não indiquem elevada precisão, eles demonstram um caminho tecnicamente válido para o uso de aprendizado de máquina no apoio ao planejamento de ressuprimento da fita de aço.

6 CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo desenvolver uma solução preditiva voltada ao apoio da gestão de consumo e ressuprimento da fita de aço em uma linha de produção de lâminas de barbear localizada no Polo Industrial de Manaus. Para isso, foram abordados fundamentos sobre cadeia de suprimentos, previsão de consumo, gestão preditiva de estoques, geração de dados sintéticos, simulação de *Monte Carlo*, aprendizado de máquina, modelo *XGBoost* e integração digital. A partir dessa estrutura, foi possível transformar registros operacionais em uma base analítica voltada à previsão e ao apoio à decisão.

Em relação à primeira hipótese, que afirmava que o uso de técnicas de aprendizado de máquina contribui para antecipar necessidades de ressuprimento da fita de aço, verificou-se confirmação parcial. A modelagem reestruturada passou a estimar o intervalo em dias até o próximo abastecimento, utilizando variáveis temporais e históricas, como *lags* e médias móveis, sem empregar atributos que provocassem vazamento de dados. Conforme apresentado na Tabela 3 e na Figura 33, o modelo obteve MAE de 0,646 dias, RMSE de 0,835 dias e R^2 de 0,001. Esses resultados indicam erro médio inferior a um dia, mas também demonstram baixa capacidade explicativa global da variável-alvo, evidenciando que a antecipação do ressuprimento ainda depende de maior volume de dados e de variáveis operacionais adicionais.

Quanto à segunda hipótese, relacionada à contribuição dos modelos preditivos para melhorar o planejamento de compras e reposição da fita de aço, observou-se também confirmação parcial e indireta. A simulação de *Monte Carlo* contribuiu para complementar a base histórica e representar cenários compatíveis com o comportamento operacional observado. Os resultados apresentados na Tabela 1 e na Figura 30 indicaram proximidade entre os dados reais e simulados, com médias e dispersões compatíveis. Entretanto, como não foi realizada uma validação financeira direta nem mensuração de redução efetiva de decisões emergenciais, excesso ou escassez do recurso, essa hipótese não pode ser confirmada de forma plena.

A terceira hipótese, referente à melhoria da disponibilidade de informações para a tomada de decisão gerencial por meio da integração de uma solução preditiva à rotina produtiva, apresentou maior aderência à proposta desenvolvida. A arquitetura baseada em *Power Apps*, *SharePoint*, *Power Automate*, *Azure* e *XGBoost* demonstrou

como os dados operacionais podem ser organizados, processados e disponibilizados ao usuário de forma estruturada. Assim, mesmo com limitações no desempenho preditivo, a solução proposta contribuiu para ampliar a visibilidade sobre o comportamento de consumo e ressuprimento da fita de aço.

De modo geral, os resultados sintetizados na Tabela 4 mostram que a proposta apresentou coerência metodológica e evolução na estruturação da modelagem preditiva. A simulação de *Monte Carlo* contribuiu para complementar a base analítica, enquanto a reestruturação do *XGBoost* permitiu avaliar o modelo sem vazamento de dados, oferecendo uma leitura mais realista da capacidade preditiva. Embora os resultados não indiquem elevada capacidade explicativa global, o erro médio inferior a um dia demonstra potencial de uso como apoio à decisão operacional.

Assim, conclui-se que o trabalho atingiu seu objetivo ao propor uma abordagem metodológica capaz de integrar dados históricos, simulação, modelagem preditiva e visualização digital em uma solução voltada ao planejamento de ressuprimento da fita de aço. A pesquisa também evidenciou que, com as variáveis atualmente disponíveis, o intervalo até o próximo abastecimento apresenta baixa explicabilidade por padrões temporais e históricos isolados. Esse resultado, longe de invalidar a proposta, reforça a importância de ampliar a base real, incorporar novas variáveis operacionais e aprimorar a integração entre dados produtivos e modelos preditivos.

Como recomendações para trabalhos futuros, sugere-se a ampliação do volume de registros reais, a inclusão de variáveis relacionadas ao ritmo de produção, turnos, paradas, consumo por máquina, lead time de fornecedores e decisões de antecipação de compra. Também se recomenda a validação da solução em ambiente operacional, com acompanhamento dos resultados gerados pelo sistema e comparação com decisões reais de ressuprimento. Dessa forma, a proposta desenvolvida neste trabalho pode servir como base para estudos futuros voltados ao uso de aprendizado de máquina na gestão preditiva de estoques industriais.

REFERÊNCIAS

ALBANO, João Francisco; OLIVA, Gláucia Michel de. **Modelos de previsão no Supply Chain Management**. In: SEMANA DE ENGENHARIA DE PRODUÇÃO SUL-AMERICANA, 6., 2006, Florianópolis. Anais [...]. Porto Alegre: Fundação Empresa-Escola de Engenharia da Universidade Federal do Rio Grande do Sul, 2006.

Disponível em:

https://www.producao.ufrgs.br/arquivos/publicacoes/58_051_Seprosul.pdf. Acesso em: 16 set. 2025.

BALLOU, Ronald H. **Gerenciamento da cadeia de suprimentos: logística empresarial**. 5. ed. Porto Alegre: Bookman, 2006. Disponível em:

<https://books.google.com/books?id=Y5JRAAAACAAJ>. Acesso em: 18 nov. 2025.

BISHOP, Christopher M. **Pattern recognition and machine learning**. New York: Springer, 2006. Disponível em: <https://link.springer.com/book/10.1007/978-0-387-45528-0>. Acesso em: 22 nov. 2025.

BORGES, Luan Barbosa. **Planejamento e controle de suprimentos e logística para redução do custo de construção**. 2025. Trabalho de Conclusão de Curso (Graduação em Engenharia Civil) – Universidade Federal de Uberlândia, Uberlândia, 2025. Disponível em: <https://repositorio.ufu.br/handle/123456789/45828>. Acesso em: 05 dez. 2025.

BRAZ, Caique Martins. **IA e IoT aplicado em manutenção preditiva e segurança de dados na Indústria 4.0**. 2025. Trabalho de Conclusão de Curso (Curso Superior de Tecnologia em Segurança da Informação) – Faculdade de Tecnologia de Americana “Ministro Ralph BIASI”, Americana, 2025. Disponível em:

<https://ric.cps.sp.gov.br/handle/123456789/35848>. Acesso em: 12 fev. 2026.

BUSSAB, Wilton de Oliveira; MORETTIN, Pedro Alberto. **Estatística básica**. 6. ed. São Paulo: Saraiva, 2010. Disponível em:

<https://books.google.com/books?id=6vYwPwAACAAJ>. Acesso em: 25 nov. 2025.

CARNELOSSI, Ana Carolina Neves; GALDAMEZ, Edwin Vladimir Cardoza; LIMA, Rafael Henrique Palma. **Gestão de estoques apoiada pela mineração de dados e aprendizagem de máquina: um panorama geral**. In: CONGRESSO DE ADMINISTRAÇÃO, SOCIEDADE E INOVAÇÃO, 14., 2022, evento on-line. Anais [...]. Volta Redonda: Universidade Federal Fluminense, 2022. Disponível em: <https://www.even3.com.br/anais/14casi/474627-gestao-de-estoques-apoiada-pela-mineracao-de-dados-e-aprendizagem-de-maquina-um-panorama-geral/>. Acesso em: 18 nov. 2025.

CARVALHO, Admilson Rodrigues de. **Método Monte Carlo e suas aplicações**.

2017. Dissertação (Mestrado Profissional em Matemática em Rede Nacional – PROFMAT) – Universidade Federal de Roraima, Boa Vista, 2017. Disponível em: https://sca.profmatt-sbm.org.br/busca_tcc_det.php?id=150070513. Acesso em: 22 set. 2025.

CAVALCANTE, Anne de Souza. **Mapeamento da inteligência artificial na cadeia de suprimentos sustentáveis**. Revista Eletrônica Científica Inovação e Tecnologia,

Medianeira, v. 15, n. 37, 2024. Disponível em:
<https://periodicos.utfpr.edu.br/recit/article/view/19142>. Acesso em: 18 nov. 2025.

CHEIN, Flávia. **Introdução aos modelos de regressão linear: um passo inicial para compreensão da econometria como uma ferramenta de avaliação de políticas públicas**. Brasília: Escola Nacional de Administração Pública, 2019. Disponível em:

https://bibliotecadigital.enap.gov.br/bitstream/1/4788/1/Livro_Regress%C3%A3o%20Linear.pdf. Acesso em: 14 out. 2025.

CHEN, Tianqi; GUESTRIN, Carlos. **XGBoost: a scalable tree boosting system**. In: ASSOCIATION FOR COMPUTING MACHINERY SPECIAL INTEREST GROUP ON KNOWLEDGE DISCOVERY AND DATA MINING INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING, 22., 2016, San Francisco. Proceedings [...]. New York: Association for Computing Machinery, 2016. p. 785-794. DOI: 10.1145/2939672.2939785. Disponível em:
<https://doi.org/10.1145/2939672.2939785>. Acesso em: 20 set. 2025.

CHOPRA, Sunil; MEINDL, Peter. **Gerenciamento da cadeia de suprimentos: estratégia, planejamento e operação**. São Paulo: Prentice Hall, 2003. Disponível em: <https://books.google.com/books?id=8x4FAQAAIAAJ>. Acesso em: 03 dez. 2025.

CHRISTOPHER, Martin. **Logística e gerenciamento da cadeia de suprimentos: criando redes que agregam valor**. 1. ed. São Paulo: Cengage Learning, 2010. Disponível em: <https://books.google.com/books?id=8xvYAAAACAAJ>. Acesso em: 10 dez. 2025.

COVER, Thomas M.; THOMAS, Joy A. **Elements of information theory**. 2. ed. New Jersey: Wiley, 2012. Disponível em:
<https://onlinelibrary.wiley.com/doi/book/10.1002/047174882X>. Acesso em: 15 jan. 2026.

DÁVALOS, Ricardo Villarroel; LÓPEZ, Oscar Ciro. **Uma abordagem da implantação de um ERP visando apoio às atividades administrativas e de ensino**. In: CONFERÊNCIA DA ASSOCIAÇÃO PORTUGUESA DE SISTEMAS DE INFORMAÇÃO, 3., 2002, Coimbra. Anais eletrônicos [...]. Coimbra: Associação Portuguesa de Sistemas de Informação, 2002. Disponível em:
<http://revista.apsi.pt/index.php/capsi/article/viewFile/331/315>. Acesso em: 15 set. 2025.

DORNIER, Philippe-Pierre et al. **Logística e operações globais**. São Paulo: Atlas, 2000. Disponível em: <https://books.google.com/books?id=Q4YwAAAACAAJ>. Acesso em: 20 dez. 2025.

FEITOSA, Rafael Divino Ferreira; SANTOS, Danilo Turkievicz dos; SOUSA, Rafael Teixeira. **Introdução a machine learning e redes neurais**. Goiânia: Centro Editorial e Gráfico da Universidade Federal de Goiás, 2024. E-book. Disponível em:
<https://portaldelivros.ufg.br/index.php/cegrafufg/catalog/book/644>. Acesso em: 06 nov. 2025.

FLEURY, Paulo Fernando; WANKE, Peter; FIGUEIREDO, Kleber Fossati. **Logística empresarial: a perspectiva brasileira**. São Paulo: Atlas, 2000. Disponível em:
<https://books.google.com/books?id=7x4FAQAAIAAJ>. Acesso em: 08 jan. 2026.

FORMOSO, Carlos Torres et al. **Material waste in building industry: main causes and prevention**. Journal of Construction Engineering and Management, v. 128, n. 4, p. 316-325, 2002. DOI: 10.1061/(ASCE)0733-9364(2002)128:4(316). Disponível em: [https://doi.org/10.1061/\(ASCE\)0733-9364\(2002\)128:4\(316\)](https://doi.org/10.1061/(ASCE)0733-9364(2002)128:4(316)). Acesso em: 22 set. 2025.

GÉRON, Aurélien. **Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow**. 2. ed. Sebastopol: O'Reilly Media, 2019. Disponível em: <https://www.oreilly.com/library/view/hands-on-machine-learning/9781492032632/>. Acesso em: 12 jan. 2026.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002. Disponível em: <https://books.google.com/books?id=Y4YwAAAACAAJ>. Acesso em: 28 nov. 2025.

GOMES, Guilherme et al. **Aplicação de machine learning e séries temporais para previsão de demanda: um estudo de caso em uma empresa de bens de consumo**. Brazilian Journal of Development, Curitiba, v. 9, n. 8, p. 23110-23132, ago. 2023. DOI: 10.34117/bjdv9n8-001. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BRJD/article/view/61838>. Acesso em: 10 out. 2025.

HAIR, Joseph F. et al. **Análise multivariada de dados**. 6. ed. Porto Alegre: Bookman, 2009. Disponível em: <https://books.google.com/books?id=Q5JRAAAACAAJ>. Acesso em: 18 dez. 2025.

INTERNATIONAL BUSINESS MACHINES CORPORATION. **What is XGBoost?** IBM Think, 2025. Disponível em: <https://www.ibm.com/think/topics/XGBoost>. Acesso em: 12 set. 2025.

INTERNATIONAL BUSINESS MACHINES CORPORATION. **XGBoost Tree node (watsonx.ai)**. IBM Documentation, 2025. Disponível em: <https://www.ibm.com/docs/en/watsonx/saas?topic=modeling-XGBoost-tree-node>. Acesso em: 12 set. 2025.

JAMES, Gareth et al. **An introduction to statistical learning: with applications in R**. New York: Springer, 2013. Disponível em: <https://www.statlearning.com>. Acesso em: 05 dez. 2025.

JEON, Hong Jun; VAN ROY, Benjamin. **Information-theoretic foundations for machine learning**. arXiv, 2024. Disponível em: <https://arxiv.org/abs/2407.12288>. Acesso em: 14 jan. 2026.

LUNA, Raiane de Almeida. **Sistema de gestão da informação para a coordenação de monitoramento e controle da produção através do Business Intelligence em uma mineradora em São Luís - MA**. 2024. 49 f. Monografia (Bacharelado em Engenharia de Produção) – Universidade Estadual do Maranhão, São Luís, 2024. Disponível em: <https://repositorio.uema.br/jspui/handle/123456789/3138>. Acesso em: 08 dez. 2025.

MAGALHÃES, Marcos Nascimento; LIMA, Antônio Carlos Pedroso de. **Noções de probabilidade e estatística**. 7. ed. São Paulo: EDUSP, 2015. Disponível em: <https://books.google.com/books?id=R5JRAAAACAAJ>. Acesso em: 20 jan. 2026.

MARTINS, Gilberto de Andrade; DOMINGUES, Osmar. **Estatística geral e aplicada**. 5. ed. São Paulo: Atlas, 2017. Disponível em: <https://books.google.com/books?id=V5JRAAAACAAJ>. Acesso em: 25 jan. 2026.

MELO, Daniela de Castro; ALCÂNTARA, Rosane Lúcia Chicarelli. **A gestão da demanda em cadeias de suprimentos: uma abordagem além da previsão de vendas**. Gestão & Produção, São Carlos, v. 18, n. 4, p. 809-824, 2011. DOI: 10.1590/S0104-530X2011000400009. Disponível em: <https://doi.org/10.1590/S0104-530X2011000400009>. Acesso em: 18 set. 2025.

MORETTIN, Pedro Alberto; SINGER, Julio da Motta. **Estatística e ciência de dados**. Rio de Janeiro: LTC, 2022. Disponível em: <https://www.grupogen.com.br/estatistica-e-ciencia-de-dados>. Acesso em: 18 fev. 2026.

PAULA, Rafael. **Simulação e modelagem de sistemas industriais**. São Paulo: Atlas, 2014. Disponível em: <https://books.google.com/books?id=W5JRAAAACAAJ>. Acesso em: 22 fev. 2026.

PIRES, Sílvio Roberto Ignácio. **Gestão da cadeia de suprimentos: conceitos, estratégias, práticas e casos**. São Paulo: Atlas, 2009. Disponível em: <https://books.google.com/books?id=X5JRAAAACAAJ>. Acesso em: 28 fev. 2026.

PIVA, Gustavo Dibbern; OLIVEIRA, Wilson José de. **Análise e gerenciamento de dados**. São Paulo: Fundação Padre Anchieta; Centro Paula Souza, 2010. v. 3. (Manual de Informática). Disponível em: <https://www.cpscetec.com.br>. Acesso em: 05 mar. 2026.

RIBEIRO, Eduardo Augusto Werneck et al. **Modelo preditivo para a proliferação do Aedes aegypti em Itajaí (Santa Catarina): uma abordagem integrando fatores climáticos locais e globais**. Estrabão, v. 5, n. 1, p. 81-91, 2024. Disponível em: <https://revista.estrabao.press/index.php/estrabao/article/view/207>. Acesso em: 05 out. 2025.

RUBINSTEIN, Reuven Y.; KROESE, Dirk P. **Simulation and the Monte Carlo Method**. 3. ed. Hoboken: John Wiley & Sons, 2017. Disponível em: <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118631980>. Acesso em: 10 mar. 2026.

SAMPAIO, Pedro Antunes; SANTOS, Renan Vitor Miranda. **Desenvolvimento de um dispositivo de baixo custo para apontamento de produção**. Revista Produção Online, v. 25, n. 2, 2025. DOI: 10.14488/1676-1901.v25i2.5446. Disponível em: <https://www.producaoonline.org.br/rpo/article/view/5446>. Acesso em: 23 maio 2026.

SCHENA, Francesco et al. **Synthetic data generation for predictive maintenance of naval propulsion systems**. 2022. Disponível em: https://www.researchgate.net/publication/398573003_Investigate_the_degradation_of_marine_propulsion_systems_under_predictive_maintenance_Comparison_of_Machine_Learning_Techniques. Acesso em: 15 jan. 2026.

SCHMIDT, Paulo; SCHMIDT, Ana Paula. **Logística 4.0**. Revista Gestão, Sustentabilidade e Negócios, Porto Alegre, número especial 1, p. 49-66, nov. 2023. Disponível em:

<https://www.saofranciscodeassis.edu.br/ojs/index.php/RGSN/article/view/44>. Acesso em: 20 nov. 2025.

VALIANT, Leslie Gabriel. ***A theory of the learnable***. Communications of the Association for Computing Machinery, v. 27, n. 11, p. 1134-1142, 1984. DOI: 10.1145/1968.1972. Disponível em: <https://doi.org/10.1145/1968.1972>. Acesso em: 25 set. 2025.

ZHOU, Zhi-Hua. ***Ensemble methods: foundations and algorithms***. Boca Raton: CRC Press, 2012. Disponível em: <https://www.crcpress.com/Ensemble-Methods-Foundations-and-Algorithms/Zhou/p/book/9781439830031>. Acesso em: 15 mar. 2026.