

DESENVOLVIMENTO DE MODELO PREDITIVO DE EMISSÕES DIFUSAS EM PÁTIOS DE PELOTAS DE USINAS DE PELOTIZAÇÃO *

Alexandre Gonçalves Andrade ¹
Tatiane Ferreira Alvarenga ²
Crisley Dellaqua Pacheco ³
Aline Maria Santoro Nasaré⁴
Gabriel Miranda Teixeira⁵

Resumo

Emissões de fontes difusas são aquelas diretamente relacionadas à suspensão de material particulado de estruturas civis, metálicas, materiais carreados pelo vento em correias transportadoras, operações de estocagem e recuperação e processos onde os sistemas de captação não são suficientes para realizar a limpeza e tratamento. A Vale SA, Diretoria de Pelotização, possui um sistema exclusivo para quantificação das emissões de fontes difusas, a Rede Automática de Monitoramento de Poeira (RAMPs), formando uma malha de análise ao longo de toda unidade operacional de Tubarão, em Vitória, ES. Para atuar de forma a prevenir as emissões difusas nos pátios de pelotas das usinas de pelotização da unidade de Tubarão, foi desenvolvido um sistema preditivo com base nas informações de variáveis de processo e informações climáticas disponíveis, utilizando ferramentas de Big Data e analítica avançada. Os modelos se mostraram aderentes às taxas reais de emissões medidos pelas RAMPs e se mostraram uma ferramenta eficiente para prever emissões acima da meta com até 4 horas de antecedência.

Palavras-chave: Emissões Difusas; Pelotização; Modelo Preditivo; Big Data

DEVELOPMENT OF PREDICTIVE MODEL FOR DIFFUSE EMISSIONS IN PELLETS STOCKYARD

Abstract

Diffuse emissions are that directly related to particulate matter suspended from civil structures, materials disposed in conveyor belts carried out by wind, stack e reclaimers operations and that one where dust collectors are not able to act correctly. Vale implemented in the last years a unique system capable to measure this kind of dust, called RAMPs (Dust Monitoring Automatic Network), which creates an analytical net in Tubarao industrial site, located in Vitória, ES. In order to prevent diffuse emissions in Tubarao's pellets stockyard, it was proposed and developed a predictive model based on operational and climate variables available on the site, using big data and advanced analytics tools. The models developed shows to be adherent to the real emissions rates measured by the RAMPs and proved to be an efficient tool to predict emissions over the internal target with 4 hours of precedence.

Keywords: Diffuse Emissions; Pelletizing; Predictive Model; Big Data.

¹ Engenharia Metalúrgica, bacharel, Engenheiro, Engenharia de Processo e Sustentabilidade, Vale S/A, Vitória, ES, Brasil.

² Engenharia de Produção, bacharel, Analista Operacional, Engenharia de Processo e Sustentabilidade, Vale S/A, Vitória, ES, Brasil.

³ Tecnologia em Processamento de Dados, bacharel, Líder de projetos de inteligência artificial, AI Center, Vale S/A, Vitória, ES, Brasil.

⁴ Engenharia de Produção, bacharel, Engenheira, Unidade Técnica Brasil, Vale S/A, Vitória, ES, Brasil

⁵ Engenharia Metalúrgica, graduando, Instituto Federal do Espírito Santo, Vitória, ES, Brasil.

1 INTRODUÇÃO

O processo de pelletização de minério de ferro, assim como outras atividades metalúrgicas, gera impactos ambientais no meio no qual está inserido. Um dos problemas mais recorrentes está ligado com a dispersão de material particulado, gerado principalmente em processos de abrasão e degradação durante o manuseio e transporte de materiais granulados [1]. Diversas ações são tomadas com o objetivo de minimizar a dispersão desse material, como aplicação de reagentes químicos durante a estocagem e nas pilhas de minério e pelotas, que aglomeram a partículas finas e impedem sua dispersão [2], além de ações durante a etapa de queima, direcionando os fluxos de ar dos fornos para os precipitadores eletrostáticos, reduzindo a quantidade de finos em suspensão [3]. Entretanto, ações climáticas ainda impactam fortemente na geração de material particulado [4], com medições de emissões excessivas em períodos do dia de fortes ventos e altas temperaturas, o que faz necessário o monitoramento contínuo das taxas de emissões e condições meteorológicas das usinas e pátios de estocagem da pelletização.

As usinas de pelletização da unidade de Tubarão, localizadas em Vitória - ES, possuem um sistema único de monitoramento de particulados de fontes difusas, denominado RAMP (Rede Automática de Monitoramento de Poeira). Esse sistema é composto por torres distribuídas ao longo da unidade, formando uma rede de monitoramento para toda a unidade industrial. As torres possuem 16 metros de altura e contêm 3 medidores de particulado, distribuídos na base, meio e pico da torre, conforme Figura 1.

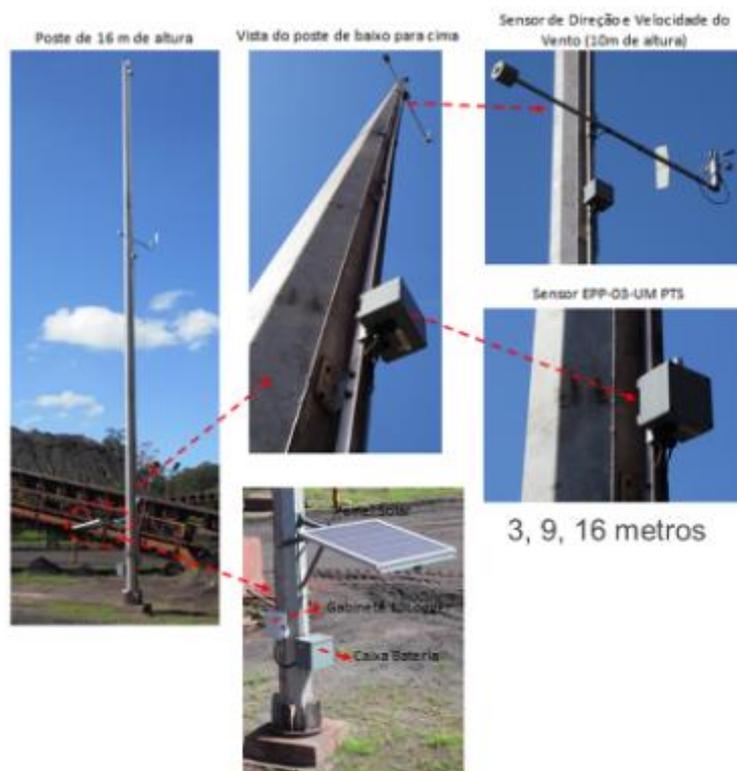


Figura 1: Torre da RAMP.

As torres de medição são compostas por sensores que medem a concentração de partículas totais em suspensão de forma on line e interpolam os valores a cada 15 minutos, além de sensores de direção e velocidade do vento permitindo a construção dos campos de ventos que atuam sobre o setor. A associação das informações de concentrações de partículas aos dados de direção e velocidade do vento permite calcular o fluxo de material particulado em cada um dos setores de monitoramento de emissões difusas de poeira.

No total, A Vale (SA) Unidade de Tubarão, possui 37 torres de RAMP, distribuídas conforme Figura 2.



Figura 2: Localização das torres RAMP.

Para calcular a taxa de emissão (kg/h) dos setores, são utilizados dois métodos de medição de forma combinada: Método de Medição do Perfil de Exposição “Exposure Profiling” e Perfiladores Móveis, como esquematizado na Figura 3. A técnica do perfil de exposição (“Exposure Profiling”) para testes em fontes de material particulado abertas é baseada no conceito do perfil isocinético que é utilizado nas medições convencionais de chaminés. A passagem do poluente emitido e carregado pelo vento imediatamente à jusante da fonte é medida diretamente através da amostragem simultânea em múltiplos pontos distribuídos na seção normal à pluma de poeira proveniente da fonte aberta. A metodologia “exposure-profiling method” (método do perfil de exposição) envolve o monitoramento simultâneo das concentrações de partículas e da velocidade do vento na seção transversal à pluma de emissão, em

diferentes alturas acima do nível do solo, seguido da integração matemática dos perfis obtidos, para obtenção da taxa de emissão final [5].

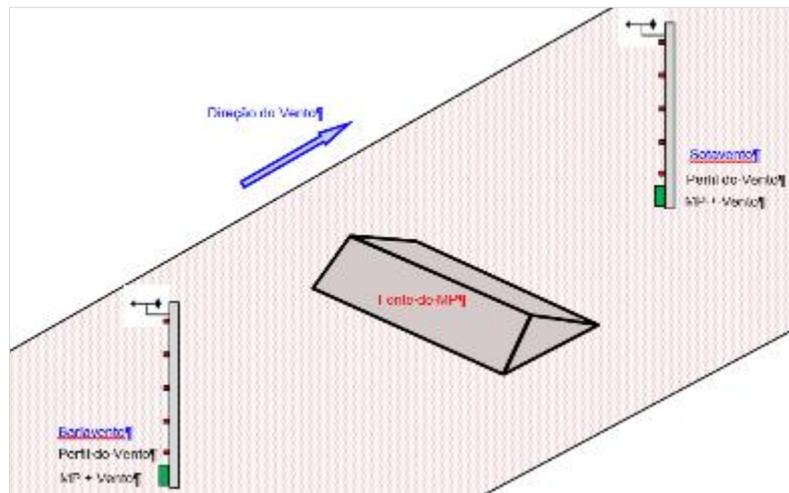


Figura 3: Metodologia de cálculo de taxa de emissão.

Os sensores são capazes de monitorar, simultaneamente, as concentrações das variáveis medidas por cada modelo em intervalos de 2 segundos. O princípio de medição baseia-se na contagem das partículas suspensas no fluxo de ar que atravessam um feixe de laser, sendo o tamanho das partículas determinado pelo sinal percebido pelo detector posicionado a 90° do feixe luminoso. O microprocessador totalmente dedicado ao sensor é capaz de detectar milhões de partículas por litro de ar. Um algoritmo de integração das partículas amostradas determina as concentrações das frações medidas. O monitor utiliza uma bomba interna para aspirar a amostra de ar a uma vazão controlada, e uma sonda de amostragem aquecida, controlada de acordo com a umidade relativa do ar e temperatura ambiente, garantindo um condicionamento adequado do fluxo de ar amostrado e reduzindo a interferência da umidade do ar.

As torres são dotadas de sistema de comunicação remota (modem GPRS celular) sendo os dados enviados para o servidor, onde são realizados os cálculos do campo de ventos e conseqüentemente da taxa de emissão de cada setor. Os dados calculados de cada sensor são enviados diretamente para o sistema de monitoramento onde é realizado o acompanhamento das áreas operacionais e de meio ambiente.

Uma vez que o sistema de monitoramento RAMP fornece uma análise contínua das taxas de emissão em toda a unidade industrial, é possível utilizar esses dados para a construção de modelos preditivos das emissões baseado em dados climáticos e operacionais, utilizando o conceito de Análise avançada de dados (*Advanced Analytics*). A análise avançada de dados está sendo usada em todos os setores da indústria para prever eventos futuros, para descobrir insights mais profundos, fazer previsões ou gerar recomendações, provando ser uma ferramenta de grande utilidade para as operações [6, 7].

Modelos preditivos para estudo da emissão de poluentes já foram desenvolvidos anteriormente e se provaram confiáveis nas previsões [8, 9]. Devido ao grande volume de sistemas de controle e sensores nas usinas, o volume de dados operacionais é grande o suficiente para que técnicas de *Big Data* sejam aplicadas para correlacionar tais dados com os dados meteorológicos.

Visando atuar antecipadamente nos eventos de emissões acima das metas dos setores dos pátios de pelotas das usinas de pelotização, o objetivo do projeto foi desenvolver um sistema de monitoramento preditivo das emissões de fontes difusas baseado em análise de dados operacionais e climáticos, permitindo desta forma, atuar com ações preventivas antes que as emissões ocorram de fato.

2 DESENVOLVIMENTO

2.1 Metodologia CRISP-DM (*Cross Industry Standard Process for Data Mining*)

A metodologia utilizada para o desenvolvimento do projeto seguiu os moldes definidos pelo CRISP-DM (*Cross Industry Standard Process for Data Mining*). O CRISP-DM é um modelo processo que define as fases que o projeto de análise de dados deve seguir. A figura 4 apresenta as etapas da metodologia.

A metodologia apresenta um formato dinâmico de forma que em diversos momentos torna-se necessário retornar nas etapas anteriores conforme o amadurecimento do projeto [10]. Além disso, o círculo externo apresentado na Figura 4 mostra que o processo é contínuo, visto que os resultados obtidos em um projeto podem desencadear novas necessidades para o negócio.

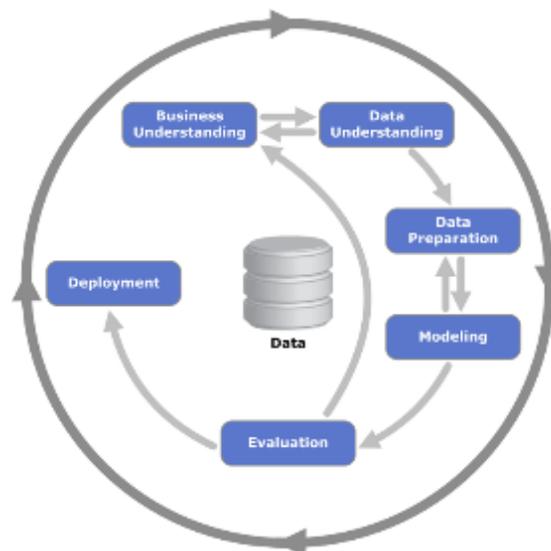


Figura 4: Etapas da metodologia CRISP-DM.

As etapas envolvidas no CRISP-DM são:

- Entendimento de negócio (*Business Understanding*): definir os objetivos, requisitos e ações do projeto, além de entender sobre o contexto de negócio do projeto;

- Entendimento dos dados (*Data Understanding*): etapa de coleta dos dados e análise exploratória, detectando padrões e valores discrepantes;
- Preparação dos dados (*Data Preparation*): etapa que visa entregar o conjunto de final de dados através dos dados brutos, realizando diversas transformações nas bases de dados;
- Modelagem (*Modeling*): aplicação de técnicas de modelagem e de calibração de parâmetros para (a) obter os melhores resultados dos modelos;
- Avaliação (*Evaluation*): verificar se os resultados dos modelos correspondem às expectativas da área de negócio e às métricas definidas;
- Implantação (*Deployment*): etapa final de entrega do resultado, podendo ser um relatório, dashboard, plano de ação, etc.

2.2 Coleta de Dados

O desenvolvimento do projeto iniciou com o mapeamento das principais variáveis do processo que impactam nas emissões difusas, listadas na Tabela 1, identificadas por meio de consultas com os especialistas do processo de pelotização e de verificação das telas do supervisor de operação.

Tabela 1. Lista de variáveis selecionadas para o modelo

Nome da Variável	Unidade
Abrasão Pelotas Queimadas	%
Temperatura Pelota Queimada nas Correias	°C
Produção de Pelotas na Usina	t/h
Vazão de Glicerina	t/h
Vazão Água de Refrigeração	m ³ /h
Vazão Água de Diluição	m ³ /h
Nível Silo da Camada de Forramento	%
Peso na Balança de Finos	t/h
Temperatura	°C
Radiação	W/m ²
Direção do vento	
Velocidade do vento	m/s

Inclui apenas parte das variáveis selecionadas.

Após avaliação inicial das variáveis operacionais preditoras através do modelo de árvore de decisão, foram definidas aproximadamente 30 variáveis operacionais importantes para os modelos preditivos. O *software* utilizado para avaliação inicial dos dados foi o SPSS *Modeler* da IBM. A Figura 5 apresenta os principais softwares utilizados e suas interações. Nota-se que os dados operacionais foram retirados da plataforma PIMS, os dados das RAMPS foram retirados da plataforma ATMOS e dados climáticos foram retirados do site *The Weather Company*. Os dados tiveram uma interface com o Oracle e foram avaliados através dos *softwares* Python e SPSS *Modeler*. A saída do modelo foi publicada utilizando o *software* Power BI.

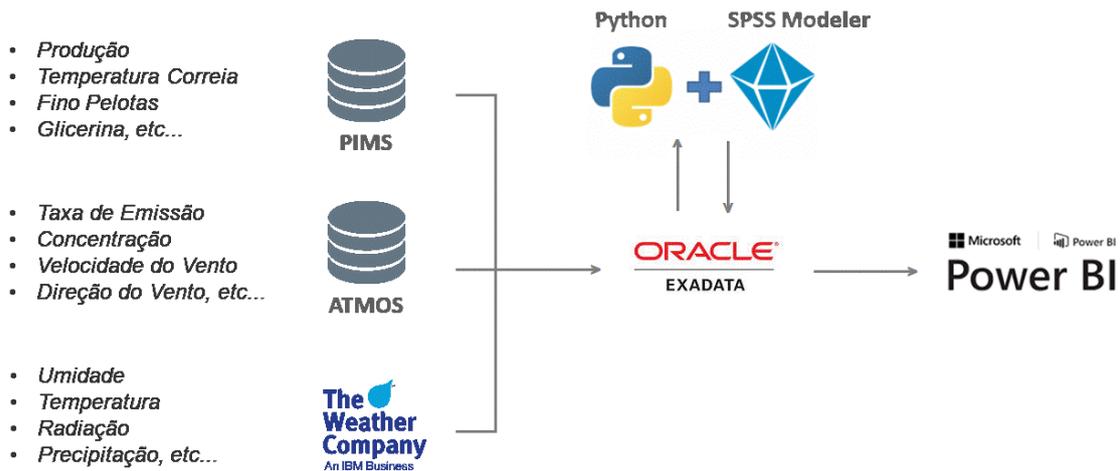


Figura 5: Arquitetura para desenvolvimento do projeto.

2.3 Construção do Modelo

A metodologia adotada para a construção do modelo é ilustrada na Figura 6, onde o principal produto do modelo diagnóstico foi uma análise mais refinada criando modelos para avaliar a correlação das variáveis ambientais e operacionais, para descoberta de *insights* de efeito causa/consequência. As principais perguntas a serem respondidas foram: Quais variáveis mais impactam a emissão de particulado? Quais cenários (conjunto de ações) resultam em efeito crítico no aumento de emissão? Quais são os valores limites ótimos das variáveis operacionais para minimizar os efeitos na emissão? Quais são os valores de alerta para monitoramento das variáveis ambientais?

Na etapa do modelo preditivo, que é o objetivo principal do trabalho, os modelos desenvolvidos apresentaram resultados de emissões de particulado de fontes difusas com 4 horas de antecedência, levando em consideração os dados de entrada fornecidos, ou seja, dados operacionais das usinas de pelotização, dados climáticos da região e dados das RAMP's.

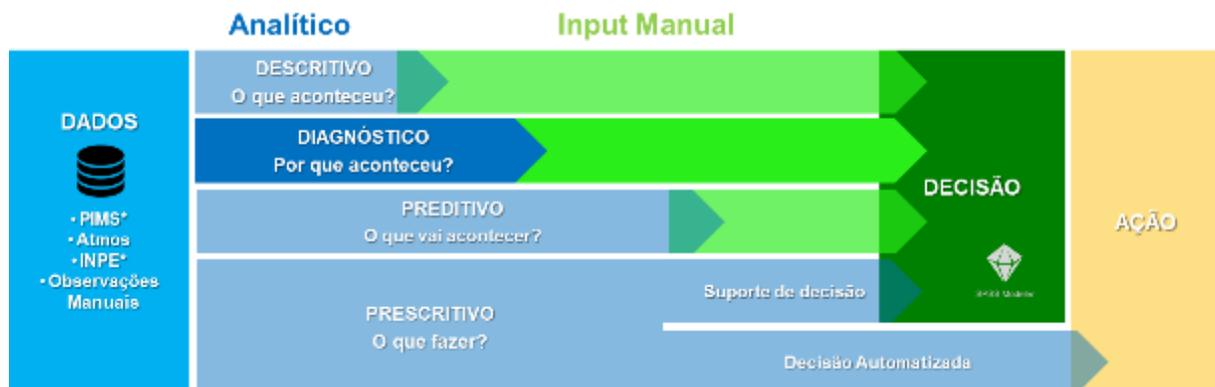


Figura 6: Modelo de desenvolvimento de análise de dados.

2.3.1 Descritivo

O projeto de *Advanced Analytics* iniciou – se por uma etapa de estudo descritivo. Nessa etapa foram realizadas análises com os dados passados para entender “o que” aconteceu através de análises gráficas e estatísticas. A Figura 7 apresenta um exemplo de análise através do nó de Auditoria dos Dados (SPSS).

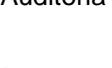
Campo	Gráfico de A...	Medição	Min.	Máx.	Média	Erro Dev	Assimetria	Exclusivo	Válido
Data_Hora		Contínuo	12-12-...	10-06-2019 13:0...	--	--	--	--	4320
Particulas_PP-U5-7		Contínuo	0.000	23.882	2.171	2.676	2.168	--	4319
Abrasao_U5_%		Contínuo	3.437	12.300	5.158	0.628	1.944	--	4318
Abrasao_U6_%		Contínuo	3.233	7.671	4.907	0.597	0.415	--	4318
Abrasao_U7_%		Contínuo	0.714	9.249	5.314	0.744	0.485	--	4202
Concentracao_Torre19-16		Contínuo	0.000	0.530	0.052	0.053	2.442	--	4319
Concentracao_Torre19-3		Contínuo	0.000	0.823	0.074	0.073	2.280	--	4319
Concentracao_Torre19-9		Contínuo	0.021	0.717	0.135	0.068	1.240	--	4319
Concentracao_Torre21-16		Contínuo	0.000	1.283	0.129	0.105	1.780	--	4319
Concentracao_Torre21-3		Contínuo	0.002	1.076	0.120	0.101	1.619	--	4319

Figura 7: Auditoria dos dados de entrada no IBM SPSS Modeler.

2.3.2 Diagnóstico

A modelagem diagnóstica foi o segundo passo do projeto de *Advanced Analytics*. Através de análises estatísticas e modelagens por árvores de decisão (algoritmo CHAID) buscou-se entender “os porquês” das emissões utilizando a base de dados existente. Os resultados das análises foram apresentados e discutidos com os especialistas da área de negócio, visando aprimorar a modelagem, além reforçar conceitos e trazer novos conhecimentos para auxiliar na tomada de decisão. A Figura 8 apresenta um exemplo de análise realizada através de uma árvore de decisão CHAID.

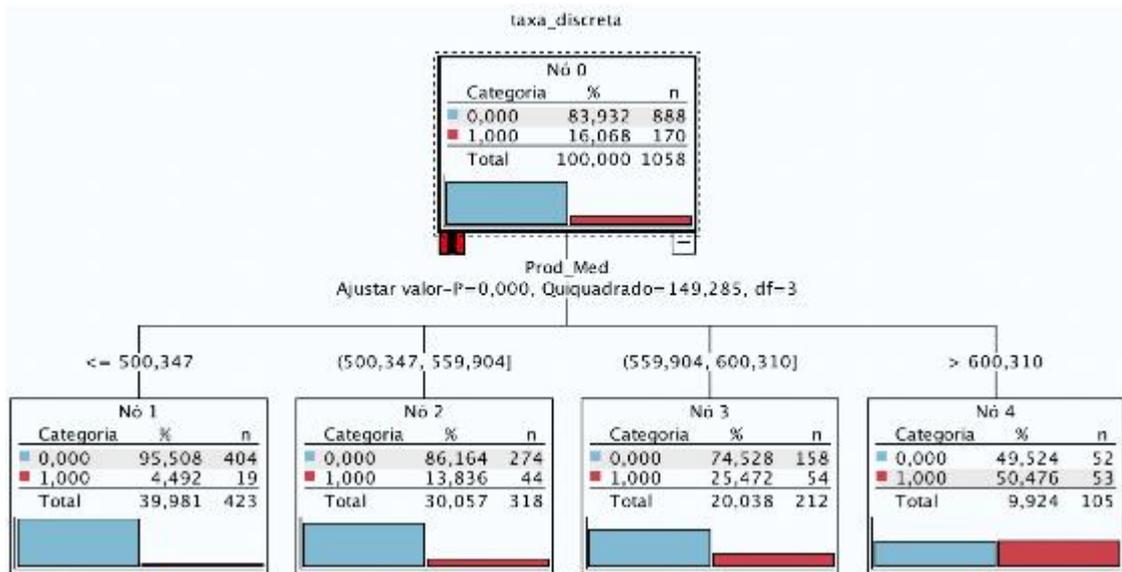


Figura 8: Análise Diagnóstica com árvore de decisão CHAID

2.3.3 Preditivo

Na etapa de modelagem preditiva o objetivo foi encontrar um modelo robusto que seja capaz de realizar previsões para atender as necessidades do negócio. O acúmulo de conhecimento das etapas anteriores foi utilizado nessa etapa, uma vez que com as informações obtidas podemos modelar as variáveis para explicar o máximo de efeitos observados no dia a dia das operações.

No processo de modelagem preditiva os dados passam por uma sequência de tratamentos para serem ajustados. Entre os ajustes realizados, pode-se citar: retirada de *outliers*, tratamento de valores nulos, tratamento de valores negativos (quando não fazem sentido), transformação de variáveis (Ex: Direção Escalar do Vento para Quadrante do Vento), cálculo de médias (Ex: Média das concentrações nas 3 alturas de uma torre RAMP), ajuste de dados no horário de verão, entre outras.

A escolha do algoritmo para a realização das modelagens é um passo muito importante. Como existem diversos algoritmos disponíveis, é importante entender qual melhor se adequa com o seu processo. Para o projeto em questão os algoritmos que melhor atenderam as expectativas foram os de Árvores de Decisão, principalmente por lidar bem com fenômenos não lineares e com variáveis de diversos tipos (contínuas, categóricas e discretas). A Figura 9 apresenta alguns nós de modelagem utilizados.

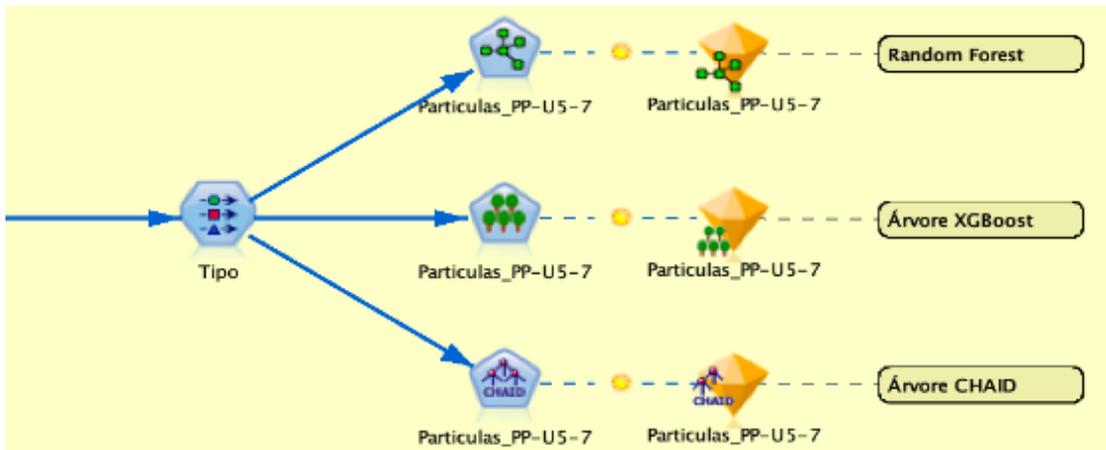


Figura 9: Nós de modelagem para o modelo preditivo.

2.3.4 Prescritivo

Uma vez realizada a modelagem preditiva, a última etapa do projeto foi a etapa prescritiva. Essa etapa teve como objetivo utilizar a previsão realizada na etapa anterior para fomentar uma tomada de decisão, seja ela com o informativo em um *dashboard*, no disparo de um alerta ou até mesmo no acionamento de um equipamento. Junto com os especialistas de área, foram desenvolvidos indicadores que mostram o nível das principais variáveis climáticas e de processo e o seu impacto nas emissões, além de elencar as variáveis mais importantes para cada uma das previsões. A Figura 10 e Figura 11 apresentam exemplo do *dashboard* construído no *software Power BI* contendo essas informações. A figura 10 representa o modelo desenvolvido para visualização dos operadores das usinas e pátios, enquanto que a figura 11 apresenta a tela desenvolvida para a equipe de equipamentos móveis, visando atuação emergencial com equipamentos auxiliares para redução dos níveis de emissão.



Figura 10: Exemplo de *dashboard* para equipe de operação das usinas.



Figura 11: Exemplo de *dashboard* para equipe de equipamentos móveis.

2.4 Resultados

Os modelos apresentaram boa precisão entre os dados preditos e os valores reais, todos com valores acima de 80%. Isso significa que o modelo acerta 80% das vezes de quando diz que a taxa de emissão do pátio vai ultrapassar o limite de emissões. A Figura 12 apresenta os valores de precisão e recall do modelo de predição (base de março/19). Recall representa a porcentagem de valores que ultrapassaram a meta de emissão e o modelo conseguiu prever.

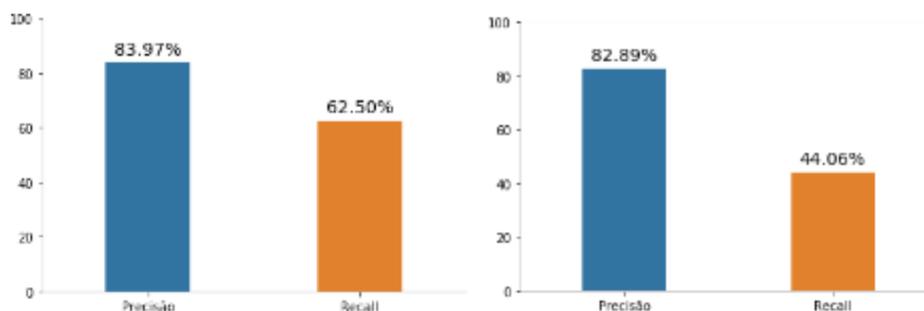


Figura 12: Precisão e recall do modelo do pátio de pelotas das usinas 1 a 4 e usinas 5 a 7.

Com os modelos preditivos criados, o próximo passo foi acompanhar o modelo por 3 semanas e tomar ações sugeridas pelo modelo para reduzir os níveis de emissões dos pátios de pelotas das usinas de pelotização. O teste foi conduzido por equipe dedicada de Processo e Operação, onde uma pessoa acompanhou os resultados previstos pelo modelo durante o turno administrativo e indicou atuação para a sala de controle baseado nas principais preditoras impactantes no resultado. As atuações

foram realizadas somente quando os valores de emissões do modelo ultrapassaram a meta de emissões de fontes difusas de cada pátio.

Apesar dos problemas verificados no período, relacionados a falhas de comunicação entre os sistemas de dados, foi possível identificar que o modelo acertou os momentos de ultrapassagem do valor limite do setor do pátio de pelotas das usinas (valor limite estimado em 3,15 kg/h) e orientou corretamente quais ações corretivas adotar para reduzir os valores para o limite.

Através do gráfico da Figura 13, é possível verificar uma redução média em torno de 6,4 kg/h para os momentos de ultrapassagem da taxa do setor da RAMP, adotando-se medidas corretivas sugeridas pelo modelo preditivo desenvolvido. Porém, o mais importante é manter as emissões sempre abaixo das metas de emissões das RAMPs, fazendo dessa forma que a média diária de emissões seja reduzida.

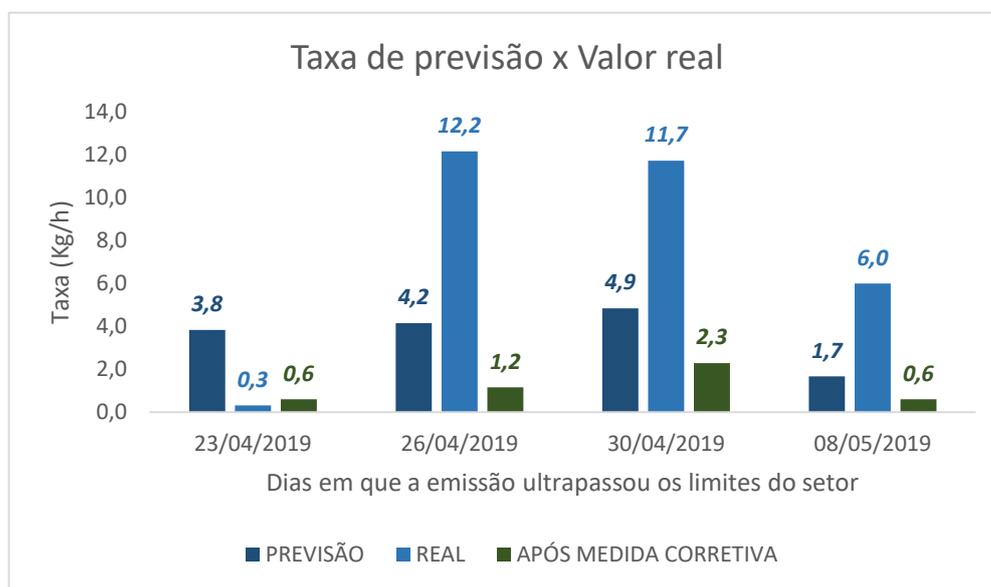


Figura 13: Resultados do teste piloto com o modelo preditivo do pátio de pelotas da usina.

3 CONCLUSÃO

Os modelos preditivos de emissões difusas dos pátios de pelotas das usinas de pelletização apresentaram boa precisão aos valores reais medidos pelas RAMPs, acima de 80%. Além disso, durante os testes industriais, os modelos indicaram corretamente as variáveis preditoras impactantes nos resultados acima das metas dos setores em teste. Portanto, pode ser considerado uma importante ferramenta, a ser utilizada para redução dos picos de emissões de particulado de fontes difusas e consequentemente a média diária de emissões dos pátios de pelotas das usinas de pelletização da unidade de Tubarão - ES.

REFERÊNCIAS

- 1 Halt JA, Nitz MC, Kawatra SK, Dubé M. Iron Ore Pellet Dustiness Part I: Factors Affecting Dust Generation. *Mineral Processing and Extractive Metallurgy Review*. 2015; 36:258-266.
- 2 Copeland CR, Kawatra SK. Dust Suppression in Iron Ore Processing Plants. *Minerals & Metallurgy Processing*. 2005;22(4):177-191.
- 3 Bolen J. Modern Air Pollution Control for Iron Ore Induration. *Minerals & Metallurgical Processing*. 2014; 31(2):103-114.
- 4 Elminir HK. Dependence of Urban Air Pollutants on Meteorology. *Science of the Total Environment*. 2005; 350:225-237.
- 5 Jia Q. Estimation of Industrial Dust Using Exposure-profiling Method [Tese]. Lulea: Lulea University of Technology; 2014. Disponível em: <https://www.diva-portal.org/smash/get/diva2:991661/FULLTEXT01.pdf>
- 6 Ge Z, Song Z, Ding SX, Huang B. Data Mining and Analytics in the Process Industry: The Role of Machine Learning. *IEEE Access*. 2017; (5): 20590-20616
- 7 Krumeich J, Werth D, Loos P, Schimmelpfennig J, Jacobi S. Advanced Planning and Control of Manufacturing Processes in Steel Industry Through Big Data Analytics: Case Study and Architecture proposal. *IEEE International Conference on Big Data (Big Data)*; 27-30 Outubro 2014, Washington DC, EUA.
- 8 Zhu D, Cai C, Yang T, Zhou X. A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization. *Big Data and Cognitive Computing*. 2018;2(5).
- 9 Suleiman A. Multivariate Study of Vehicle Exhaust Particles Using Machine Learning and Statistical Techniques [Tese]. Birmingham: The University of Birmingham; 2016. Disponível em: etheses.bham.ac.uk/6945/1/Suleiman16PhD.pdf
- 10 Azevedo A, Santos MF. KDD, SEMMA and CRIP-DM: A Parallel Overview. *IADIS European Conference Data Mining*; 24-26 Julho 2008, Amsterdam..