

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOCOMBUSTÍVEIS**

FERNANDO CESAR CALSONI

TRABALHO DE DISSERTAÇÃO

**Etanol de biomassa de milho - Utilização de aprendizagem de máquina no estudo de
casos de caldeiras**

Uberlândia/MG

2020

FERNANDO CESAR CALSONI

TRABALHO DE DISSERTAÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Biocombustíveis da Universidade Federal de Uberlândia, como requisito parcial para a obtenção do título de Mestre em Ciência e Tecnologia em Biocombustíveis.

Orientador: Prof. Dr. Ubirajara Coutinho Filho

Uberlândia/MG

2020

FERNANDO CESAR CALSONI

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

C165 Calsoni, Fernando César, 1962-
2020 Etanol de biomassa de milho - Utilização de aprendizagem de máquina no estudo de casos de caldeiras [recurso eletrônico] / Fernando César Calsoni. - 2020.

Orientador: DR UBIRAJARA COUTINHO FILHO.
Dissertação (Mestrado) - Universidade Federal de Uberlândia,
Pós-graduação em Biocombustíveis.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.di.2020.464>
Inclui bibliografia.
Inclui ilustrações.

1. Biocombustível. I. COUTINHO FILHO, DR UBIRAJARA ,1970-, (Orient.). II. Universidade Federal de Uberlândia. Pós-graduação em Biocombustíveis. III. Título.

CDU: 662.756

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Coordenação do Programa de Pós-Graduação em
Biocombustíveis

Av. João Naves de Ávila, 2121 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902
Telefone: (34) 3239-4385 / 4208 - www.iq.ufu.br - ppbic@iqufu.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Biocombustíveis			
Defesa de:	Mestrado Acadêmico, 40, PPBIC			
Data:	03 de abril de 2020	Hora de início:	09:00	Hora de encerramento:
Matrícula do Discente:	11812PGB001			
Nome do Discente:	Fernando César Calsoni			
Título do Trabalho:	Etanol de biomassa de milho - utilização de aprendizagem de máquina no estudo de casos de caldeira			
Área de concentração:	Biocombustíveis			
Linha de pesquisa:	Processos químicos e bioquímicos			
Projeto de Pesquisa de vinculação:				

Reuniu-se remotamente através da sala de conferencia meet.google.com/dnn-xfqp-zpr os Doutores: Prof. Moíton Ribeiro Franco, da Universidade Federal de Uberlândia; Prof. Carlos André Gonçalves, do Instituto Luterano de Ensino Superior; e Prof. Ubirajara Coutinho Filho, orientador do candidato.

Iniciando os trabalhos o presidente da mesa, Dr. Ubirajara Coutinho Filho, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(as) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente

ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Ubirajara Coutinho Filho, Professor(a) do Magistério Superior**, em 03/04/2020, às 11:07, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Moilton Ribeiro Franco Junior, Professor(a) do Magistério Superior**, em 03/04/2020, às 11:07, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Carlos André Gonçalves, Usuário Externo**, em 03/04/2020, às 21:18, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1926905** e o código CRC **D42F7859**.

Referência: Processo nº 23117.018425/2020-15

SEI nº 1926905

**Etanol de biomassa de milho- Utilização de modelos de aprendizagem de máquina no
estudo de casos de caldeiras**

Dissertação apresentada ao Programa de Pós-Graduação em Biocombustíveis da Universidade Federal de Uberlândia, como requisito parcial para a obtenção do título de Mestre em Ciência e Tecnologia em Biocombustíveis.

Orientador: Prof. Dr. Ubirajara Coutinho Filho.

Data de aprovação: ____/____/____

Banca examinadora:

Prof. Dr.

Instituição:

Prof. Dr.

Instituição:

Prof. Dr.

Instituição:

Prof. Dr.

Instituição:

DEDICATÓRIA

A meus pais João Calsoni e Maria de Lourdes Dessimoni Calsoni, pessoas simples, que me ensinaram ser uma pessoa melhor a cada dia.

AGRADECIMENTOS

A Deus, por me guiar , me orientar nos momentos difíceis e por me proporcionar uma vida de muito trabalho, vitórias e aprendizado na derrotas.

Aos meus familiares e amigos que sempre me apoiam, em especial à minha esposa e filhos. *In memoriam* ao meus pais ,João Calsoni e Maria de Lourdes Dessimoni Calsoni.

À Universidade Federal de Uberlândia (UFU) e à Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM) por oferecerem o curso de Mestrado e Doutorado na área de Biocombustíveis e oportunizar qualificação aos estudantes de diversos cursos.

Ao meu orientador, Prof. Dr. Ubirajara Coutinho Filho, pelo incentivo, orientação, dedicação e por acreditar neste trabalho, tendo a paciência em ajudar a desenvolvê-lo.

Ao Programa de Pós-Graduação em Biocombustíveis (PPGB) e todos os docentes por colaborarem com todo aprendizado que obtive durante estes dois anos.

RESUMO

Nesta dissertação, trabalhamos com um gerador de vapor (caldeira) operando dentro de uma planta industrial produtora de biocombustíveis (no caso, etanol de milho) utilizando como ferramenta a aprendizagem de máquina supervisionada . O etanol de cana de açúcar é exaustivamente estudado , porém no caso do milho, no Brasil, ainda é embrionário – com históricos escassos, de um equipamento fundamental que é a caldeira, operando com queima específica de chip de eucalipto, buscando obter eficiências térmicas otimizadas e entendendo o seu comportamento, principalmente de unidade fabril” full” de milho. Foram obtidos, em campo através do DCS- Digital Control System (60 pontos de duas caldeiras por 30 dias) , dados de variáveis de processo, trabalhados com o código Python e ML (*Machine Learning*) que utilizamos como metodologia. Os resultados foram positivos, demonstrando que esse importante equipamento industrial pode ser estudado e seu comportamento previsto com a metodologia usada, auxiliando e fazendo com que possamos alcançar melhores eficiências e entender melhor o seu comportamento técnico.

Palavras-chave: Geradores de vapor. Biocombustíveis. Aprendizagem de máquina.

ABSTRACT

In this dissertation, we work with a steam generator (boiler) operating within an industrial plant that produces biofuels (in this case, corn ethanol) using supervised machine learning as a tool. Sugarcane ethanol is extensively studied, but in the case of corn, in Brazil, it is still embryonic - with scarce records, of a fundamental equipment that is the boiler, operating with specific burning of eucalyptus chip, seeking to obtain optimized thermal efficiencies and understanding its behavior, mainly of a “full” corn plant. Data from process variables were obtained in the field through the DCS- Digital Control System (60 points from two boilers for 30 days), worked with Python and ML (Machine Learning) code that we use as a methodology. The results were positive, demonstrating that this important industrial equipment can be studied and its behavior predicted with the methodology used, helping and enabling us to achieve better efficiencies and better understand its technical behavior.

Keywords: Steam boilers. Biofuels. Machine Learning.

LISTA DE FIGURAS

Figura 1 – Biomassa cavaco de eucalipto.....	14
Figura 2 – Um dos modelos de centrais térmicas de uma usina.....	15
Figura 3 – Mudança de estado.....	19
Figura 4 – Exemplificação de um processo de combustão.....	20
Figura 5 – Fluxo do ar, combustível e gases queimados.....	21
Figura 6 – Fluxo de vapor e água.....	22
Figura 7 – Fluxo dos produtos químicos.....	22
Figura 8 – Ilustração de uma turbina a vapor.....	23
Figura 9 – Caldeira Flamotubular.....	25
Figura 10 – Caldeira Aquotubular.....	25
Figura 11 – Blocos da AI (Artificial Intelligence).....	27
Figura 12 – Modelo de aprendizado supervisionado.....	28
Figura 13 – Estados que formam o <i>Corn Belt</i> (Cinturão do Milho) nos Estados Unidos.....	31
Figura 14 – O processo de refino do milho.....	33
Figura 15 – Conversão do milho em produtos e coprodutos com suas respectivas relações de produção.....	35
Figura 16 – Aproveitamento de produtos e coprodutos da cana-de-açúcar.....	36
Figura 17 – Fluxograma de Fabricação: Etanol, Açúcar e Energia.....	37
Figura 18 – Principais estados brasileiros produtores de milho.....	41
Figura 19 – Diagrama do sistema de produção de uma usina de milho no Brasil.....	42
Figura 20 – Controle geral da caldeira.....	44

LISTA DE GRÁFICOS

Gráfico 1 – Produção de etanol, no Brasil, em mil metros cúbicos, entre 1980 e 2018.....	17
Gráfico 2 – Produção de vapor	47
Gráfico 3 – Umidade da biomassa	48
Gráfico 4 – PCI da biomassa	49
Gráfico 5 – Percentual de cinzas	50
Gráfico 6 – Consumo de biomassa	51
Gráfico 7 – Temperatura dos Gases- Distribuição % dos pontos analisados.....	52
Gráfico 8 - Temperatura dos Gases- Importância das variáveis.....	53
Gráfico 9 – Teor de Oxigênio (%)– Distribuição % dos pontos analisados.....	54
Gráfico 10 - Teor de oxigênio (%) - Importância das variáveis.....	55
Gráfico 11 – Temperatura da Fornalha- Distribuição % dos pontos analisados.....	56
Gráfico 12 – Temperatura da Fornalha – Importância das variáveis.....	57
Gráfico 13 – Vazão de Vapor- Distribuição % dos pontos analisados.....	58
Gráfico 14 – Vazão de Vapor- Importância das variáveis.....	59

LISTA DE ABREVIATURAS E SIGLAS

ABNT – Associação Brasileira de Normas Técnicas

ANN – Artificial Neural Network

CAPEX – Custo de capital para implantação

COFINS – Contribuição para o Financiamento da Seguridade Social

Corn Belt – Cinturão de Milho

DDGS – Distilled Dried Grains with Solubles

DCS – Digital Control System

DOU - Diário Oficial da União

Embrapa – Empresa Brasileira de Pesquisa Agropecuária

EN – European standard

Eq – Equivalente

EUA – Estados Unidos da América

Flex – Flexível

GEE – Gases Efeito Estufa

IA – Inteligência Artificial

ICMS – Imposto sobre Circulação de Mercadorias e Serviços

IMEA – Instituto Mato Grossense de Economia Agropecuária

ML - Machine Learning

MP – Material Particulado

MTE – Ministério do Trabalho e Emprego

Nox – Óxido Nitroso

NR – Norma Regulamentadora
OPEX – Custo de capital para operação
PCI – Poder Calorífico Inferior
PPM – Parte por milhão
PIS – Programa de Integração Social
RNA – Redes Neurais Artificiais
SC – Saco de 60 kg
TC – Temperatura em grau Celsius
1G – Etanol de Primeira Geração
2G – Etanol de Segunda Geração
3G – Etanol de Terceira Geração
4G – Etanol de Quarta Geração

SUMÁRIO

1 INTRODUÇÃO	12
1.1 BIOCOMBUSTÍVEIS.....	16
2 REVISÃO BIBLIOGRÁFICA	18
2.1 ASPECTOS GERAIS RELACIONADOS ÀS CALDEIRAS	18
2.2 PRINCIPAIS CONCEITOS DA TERMODINÂMICA RELACIONADOS À OPERAÇÃO DE CALDEIRAS	19
2.3 <i>Tipos, classificação das caldeiras e seus acessórios principais.</i>	24
2.4 USO DE APRENDIZAGEM DE MÁQUINA E MÉTODOS ESTATÍSTICOS.....	27
2.4.1 INTELIGÊNCIA ARTIFICIAL	27
2.4.2 <i>XGBoost.</i>	28
2.4.3 <i>Quimiometria</i>	29
2.5 MODELAGEM E ESTUDO DE CALDEIRAS UTILIZANDO APRENDIZADO DE MÁQUINA.	30
2.6.1 ASPECTOS RELEVANTES DAS PRODUÇÕES DO ETANOL DE MILHO NOS EUA E DA CANA-DE-AÇÚCAR NO BRASIL	31
2.6.2 PRODUÇÃO DO ETANOL DE MILHO	32
2.6.3 VIABILIDADE DA PRODUÇÃO DE ETANOL DE MILHO NOS EUA	33
2.6.4 VANTAGENS E DESVANTAGENS – MILHO X CANA-DE-AÇÚCAR NO BRASIL.....	35
2.6.5 EQUILÍBRIO ENERGÉTICO DO ETANOL DE MILHO X CANA-DE-AÇÚCAR NO BRASIL.....	39
2.6.6 ANÁLISE SOCIOECONÔMICA E AMBIENTAL DA PRODUÇÃO DE ETANOL DE MILHO NO CENTRO OESTE BRASILEIRO.....	39
3 OBJETIVOS	44

4 MATERIAIS E METODOLOGIA	45
4.1 COLETA DE DADOS HISTÓRICOS	45
5 RESULTADOS E DISCUSSÕES.....	49
5.1 ANÁLISE DO DESEMPENHO DA CALDEIRA SEGUNDO O XGBOOST	54
CONCLUSÕES	63
REFERÊNCIAS	64
APÊNDICE –A	69

1 INTRODUÇÃO

Pode ser dito que estamos vivendo o momento da “biomania”, com grandes incertezas no volume de fornecimento e política de preços do petróleo, praticadas pelo mundo e o potencial exaustão desse recurso natural não renovável; a biorrefinaria tornar-se-á então uma realidade, a menos que encontremos uma forma mais efetiva de capturar a energia solar, que não seja por meio da fotossíntese (CREUTZIG,2015).

O uso de combustíveis fósseis traz consigo um aumento considerável nos gases de efeito estufa (GEE) na atmosfera, além de estar em processo de redução de volume, em que os custos de produção somente tenderão a aumentar; razão principal e tema atual do interesse crescente na captura da bioenergia na produção de biocombustíveis. Os biocombustíveis, atualmente num pequeno percentual e num futuro breve com alto percentual, serão o complemento e a resposta da sociedade ao mundo competitivo e que busca equilíbrio ambiental e sustentabilidade, além de aspectos sócio- econômicos (DEBONI,2019). No contexto da sustentabilidade da produção de biocombustíveis, um dos aspectos importantes é o aproveitamento da energia da biomassa da melhor e mais racional forma possível, sendo que uma das formas de tal aproveitamento está relacionada à geração e à utilização de vapor, que na prática está associado ao uso de caldeiras (CUNHA,*et al*, 1989).

Caldeiras ou geradores de vapor d’água são grandes, complexos e dispendiosos equipamentos e de suma importância no contexto geral da unidade produtora. O perfeito entendimento desse equipamento traz resultados econômico-financeiros que afetam a viabilidade técnica e econômica da atividade. Então, diante do fato de as caldeiras serem fundamentais nos processos industriais diversos e em particular nas unidades produtoras de biocombustíveis, tem-se que esta dissertação se justifica pois vai estudar, trabalhar e modelar as caldeiras no contexto das biorrefinarias.

O estudo de caldeiras tem uma complexidade associada ao seu rendimento térmico, emissão gasosa e segurança do processo. Uma das formas de estudos de sistemas complexos, como caldeiras, envolve o uso de modelos matemáticos e computacionais, sendo estes os modelos de inteligência artificial e aprendizado de máquina, cada vez mais utilizados pela capacidade de modelar sistemas complexos. O emprego de caldeiras implica na presença de riscos como explosões, incêndios, choques elétricos, intoxicações, quedas, ferimentos diversos, etc (FANTAZZINE,*et al* 2002). Os riscos de explosões são, entretanto, os mais importantes

pelas seguintes razões:encontra-se presente durante todo o tempo de funcionamento, sendo imprescindível seu controle de forma contínua, ou seja, sem interrupções,em razão da violência com que as explosões acontecem. Na maioria dos casos, suas consequências são catastróficas em virtude da enorme quantidade de energia liberada instantaneamente (BOTELHO & BIFANO, 2016). Envolve não só os operadores como também as pessoas que trabalham nas redondezas e instalações a sua volta. A prevenção deve ser considerada em todas as fases: projeto, fabricação, operação, manutenção, inspeção e outras. Aumento de pressão decorrente de falhas diversas, que podem ser operacionais ou não.

O principal objetivo desta dissertação é avaliar o desempenho de uma caldeira aquatubular- a primeira do Brasil com grelha vibratória com tecnologia Dan Power- de 67 kgf/cm² de pressão, temperatura do vapor produzido em 520 °C queimando biomassa (chip de eucalipto), operando dentro da uma planta industrial autônoma, produtora de biocombustíveis (no caso, etanol de milho), utilizando como ferramentas o Aprendizado de Máquina e a Quimiometria para avaliar o efeito das variáveis: temperatura dos gases efluentes, temperatura da fornalha, vazão de vapor e teor de oxigênio dos gases efluentes, no desempenho da caldeira.

Como objetivos específicos, será feito:

- a) Coleta,organização e descrição dos dados de entrada que representam o comportamento do equipamento;
- b) Seleção de metodologia de aprendizado de máquina e métodos estatísticos ;
- c) Avaliação da caldeira segundo as modelagens selecionadas;

As caldeiras à biomassa, como outras quaisquer, têm sempre associadas a elas uma capacidade de produção em toneladas/hora de vapor a uma determinada temperatura e pressão de trabalho. A biomassa pode ser uma mistura de resíduos de madeira/eucalipto, cascas, material vegetal ou bagaço de cana, que foram gerados internamente no processo ou adquirido de fontes externas (CORTEZ,*et al*,2008).

A Figura 1 ilustra a biomassa utilizada nesta dissertação.

Figura 1 – Biomassa cavaco de eucalipto.



Fonte: O autor (2017).

A Figura 2 ilustra um dos modelos de centrais térmicas de uma usina.

Figura 2 – Um dos modelos de centrais térmicas de uma usina.



Fonte: O autor (2017).

1.1 Biocombustíveis

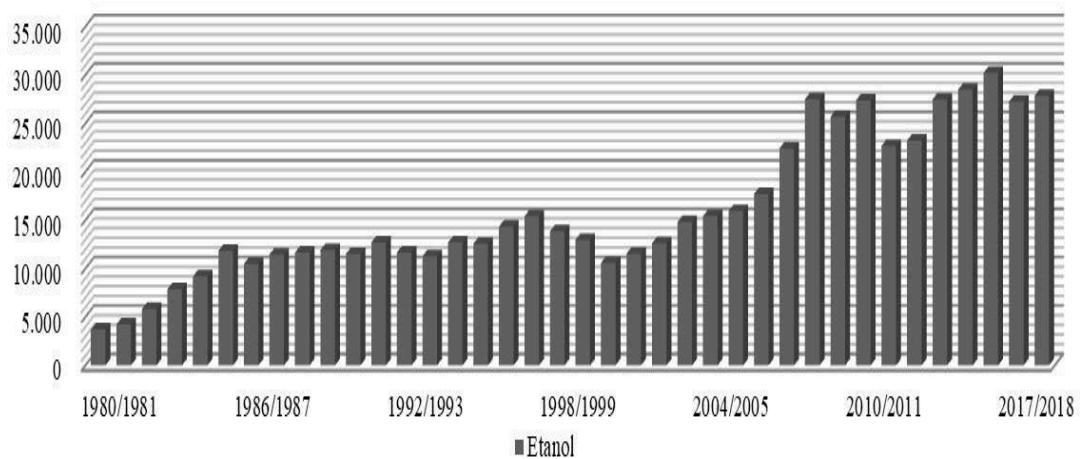
Biocombustíveis são combustíveis de origem biológica, fabricados a partir de vegetais e/ou resíduos de animais. Os biocombustíveis podem ser usados integralmente em veículos ou misturados com combustíveis fósseis (etanol + gasolina, biodiesel + diesel) (SCIENCEBLOGS,2019). Principal componente dos biocombustíveis, a biomassa é toda fonte de energia renovável que vem de materiais orgânicos. Servem como biomassa, por exemplo, os vegetais, restos de madeira, galhos e folhas de árvore, óleo vegetal e lixo urbano.Os principais biocombustíveis são: etanol (produzido da cana-de-açúcar e do milho), biogás (produzido da biomassa), biodiesel (produzido de vegetais ou sebo animal), entre outros.(HOTTA,2019).

Os Estados Unidos produzem etanol de milho e é o maior produtor do mundo com a produção ao redor de 61 bilhões de litros por ano. Já o Brasil está em segundo lugar no ranking com a produção em torno de 29 bilhões de litros anuais, a partir da cana-de-açúcar (ÚNICA, 2018). Ambos os países são criticados por desenvolverem “certa” monocultura para a produção do biocombustível, e, ainda, no caso do milho, por usar alimento para a geração de combustível. Mas as críticas são rebatidas, apenas 20% do milho americano é destinado à indústria sucroenergética, alguns estados americanos – no *Corn Belt* – produzem mais milho que todo o Brasil. Não muito diferente, o Brasil usa apenas parte da plantação de cana para o processamento de etanol, a outra fração é direcionada para a produção de açúcar (CANAL RURAL, 2019). A maior dificuldade em ambas as plantações está nas questões climáticas. Por mais que a tecnologia e a ciência avancem, ainda é difícil prever muito antecipadamente secas, geadas e chuvas e como são produtos agrícolas, ambos dependem das condições climáticas para um bom desempenho. O petróleo continua sendo a fonte de energia primária mundial.(IEA, 2020) Por ser uma fonte não renovável, a medida em que a demanda por combustíveis aumenta, intensificam-se as discussões sobre temas ambientais, como o aumento dos gases de efeito estufa (GEE) emitidos pela queima dos combustíveis fósseis e a busca por fontes renováveis de energia. Esse cenário traz a perspectiva de aumento na demanda por etanol, tanto no mercado nacional quanto no internacional, impulsionado pelo uso de veículos automovidos leves “*flexfuel*” e híbridos (INGLEDEW,2009).A produção de biocombustíveis, especialmente o etanol, pode ser dividida atualmente em quatro gerações. São considerados os de 1^a geração (1G) os que utilizam açúcares ou amidos, de 2^a e 3^a geração (2G e 3G) os que utilizam biomassa lignocelulósica, inclusive resíduos agrícolas e industriais, árvores e determinadas espécies de gramíneas, algas. A 4^a geração (4G) , é a integração dos processos de produção das demais

gerações, que utiliza de alterações genéticas da cultura a ser empregada e de microrganismos industriais, contribuindo para aumentar a eficiência dos processos de 2^a e 3^a gerações para produção de etanol (DEBONI, 2019).

A produção mundial desse combustível é da ordem de 100 bilhões de litros – o Brasil é responsável pela fabricação de cerca de 29 bilhões de litros (UNICA, 2018). O Gráfico 1 ilustra séries históricas desde 1980/1981 até 2018 da produção de etanol no Brasil.

Figura/ Gráfico 1 – Produção de etanol, no Brasil, em mil metros cúbicos, entre 1980 e 2018.



Fonte: (<http://www.unica.com.br/documentos> acesso em 23 de abril de 2019).

2 REVISÃO BIBLIOGRÁFICA

2.1 Aspectos gerais relacionados às caldeiras

A primeira tentativa do homem em produzir vapor na história da evolução da humanidade foi no século II a.C., quando Heron de Alexandria ,projetou um aparelho que vaporizava água e movimentava uma esfera em torno de seu eixo. Esse foi o aparelho precursor das caldeiras e das turbinas a vapor. Entretanto, foi na época da Revolução Industrial que teve impulso o uso do vapor sob pressão para movimentar as máquinas. Muitos – cientistas, artífices e operários – ocuparam-se por longos anos na evolução dos geradores de vapor. Os mais notáveis trabalhos nesse campo devem-se a Denis Papin, na França, a James Watt, na Escócia, e a Wilcox, nos Estados Unidos (ALTAFINI, 2002).

Caldeira é o nome popular dado aos equipamentos geradores de vapor cuja aplicação tem sido ampla no meio industrial e também na geração de energia elétrica nas chamadas centrais termoelétricas. As caldeiras podem utilizar-se de diferentes tipos de combustível, que têm como função principal gerar vapor para a produção e servir ao processo de cogeração de energia produzindo, além de vapor, energia elétrica consumida internamente e exportada, parcialmente, para a rede consumidora. No indústria sucroenergética, as caldeiras de força utilizam como seu principal combustível a biomassa. As caldeiras representam um grande gasto de capital, tanto no CAPEX quanto no OPEX . Sua operação e manutenção correta, segura e eficaz é frequentemente crítica para garantir lucratividade. Portanto, é essencial o treinamento,

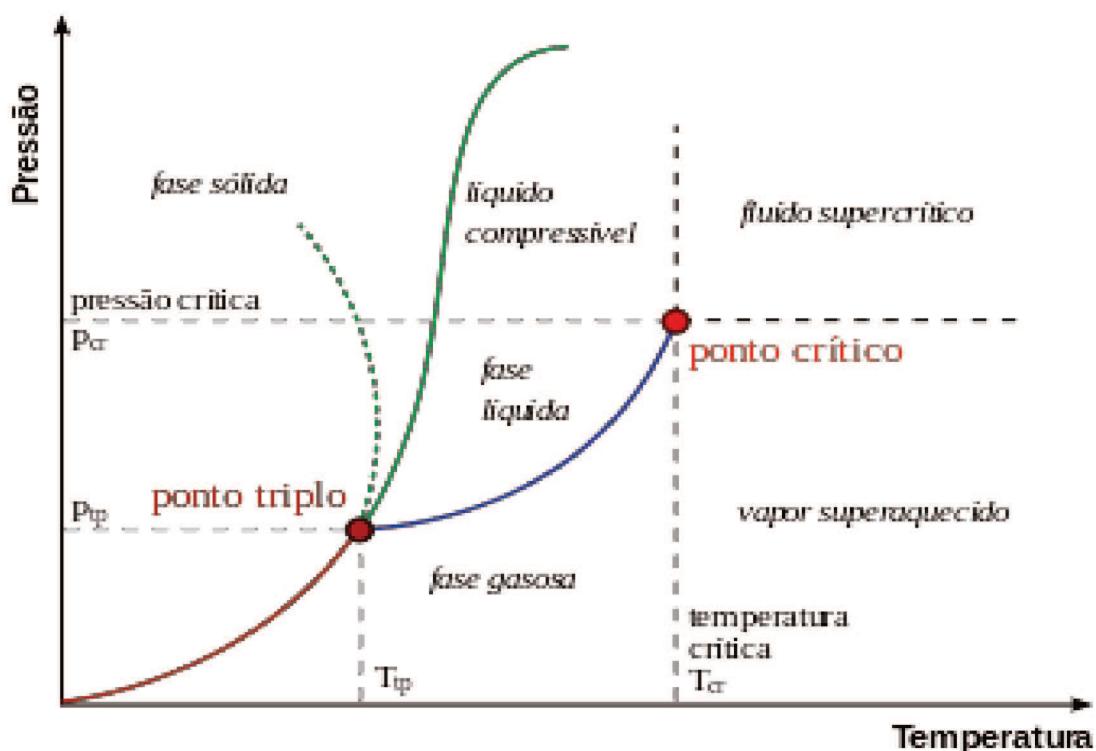
a capacitação e o desenvolvimento do pessoal responsável por esses equipamentos. Vários são os aspectos relacionados ao funcionamento das caldeiras (DUTRA,*et al.*,2006).

Será tratado de maneira geral os principais conceitos da termodinâmica envolvidos na operação de caldeiras; tipos e classificação das caldeiras e seus acessórios principais; riscos de explosões e da Norma Regulamentadora NR 13 - Caldeiras, Vasos de Pressão e Tubulação, publicação D.O.U. Portaria MTE n.º 3.214, de 08 de junho de 1978. Neste trabalho, o foco é a Aprendizagem de Máquina, aplicado para um determinado tipo e característica de caldeira.

2.2 Principais conceitos da termodinâmica relacionados à operação de caldeiras

De modo geral, as substâncias podem existir em diferentes fases, que são elas: sólida, líquida e gasosa. Assim, fase é definida como uma porção homogênea de matéria.

Figura 3 – Mudança de estado.



Fonte: (ÇENGEL,2013).

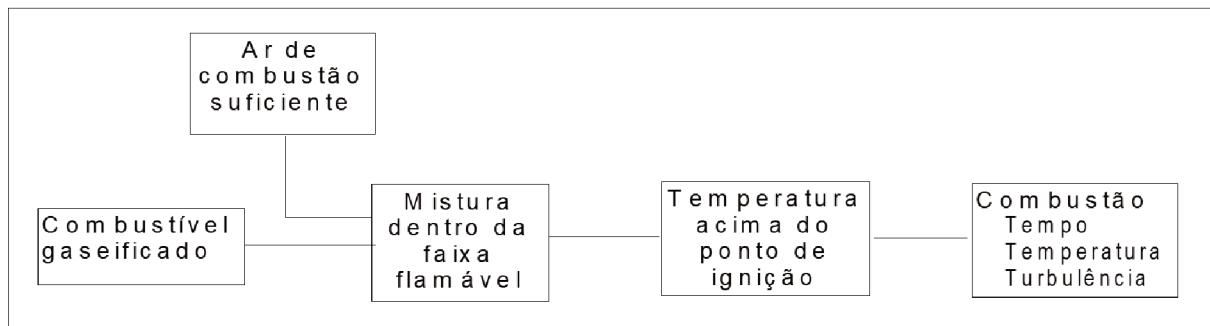
A Figura 3 exemplifica a fase gasosa da substância. Utiliza-se com frequência o nome vapor para essa fase, quando a substância está próxima de um estado em que parte desta pode se condensar. O comportamento pressão, volume e temperatura é expresso pela equação $PV = nRT$ para os chamados gases perfeitos, já para o vapor, que é considerado um gás real, essa equação não representa muito bem o comportamento mencionado(ÇENGEL & BOLES,2013). As equações de estado utilizadas para expressar o comportamento dos gases reais são, em geral, muito complexas, inviabilizando de forma rápida os seus usos. Para tanto, utiliza-se na maioria das aplicações em engenharia os softwares, diagramas e as tabelas termodinâmicas para as diferentes fases das substâncias. Nesses recursos, especialmente para as fases líquida e gasosa (vapor), são apresentados diversos valores das propriedades termodinâmicas como a entalpia e a entropia.

É importante destacar que o vapor d'água é utilizado como agente transportador de energia em diversos processos industriais e nas centrais termelétricas.

A água é a substância mais abundante na Terra, possuindo grande conteúdo energético (entálpico) e sendo pouco corrosiva, além de não tóxica, não inflamável e não explosiva (ALTAFINI,2002). Em geral, o vapor é empregado para aquecimento e para a produção de trabalho mecânico. Para aquecimento, o vapor pode ser usado direta ou indiretamente. No processo de aquecimento direto, o vapor entra em contato direto com o material a ser aquecido, exemplo disso é o aquecimento de água ou outros líquidos com injeção direta de vapor. Outros exemplos: lavagem de garrafas, curtimento de couro, esterilização, engomagem de tecidos, etc. No uso indireto, o vapor não entra em contato com o material a ser aquecido e fica, portanto, separado por uma superfície. Esse método é empregado quando for necessária uma grande quantidade de calor e/ou em processos que devem ser livres de contaminação. Em nosso estudo, trabalhamos com as variáveis obtidas do processo de combustão, exemplificado na Figura 4, que é responsável por mais de 85% da energia utilizada pelo homem, presente em quase todos os setores da sociedade (CHAVES,2013). É resultado de uma série de reações

químicas muito complexas e rápidas e os produtos formados dependem de vários fatores, com a utilização do aprendizado de máquina torna-se mais fácil compreendê-lo.

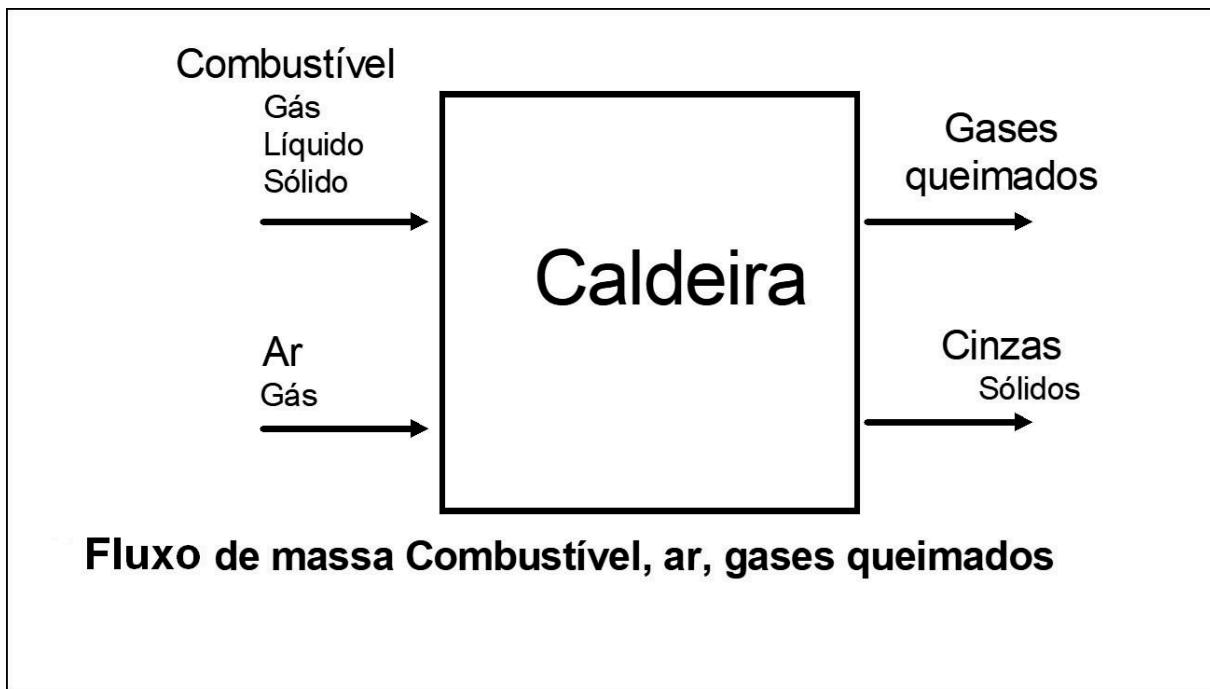
Figura 4 – Exemplificação de um processo de combustão.



Fonte:(BEGA,2003)

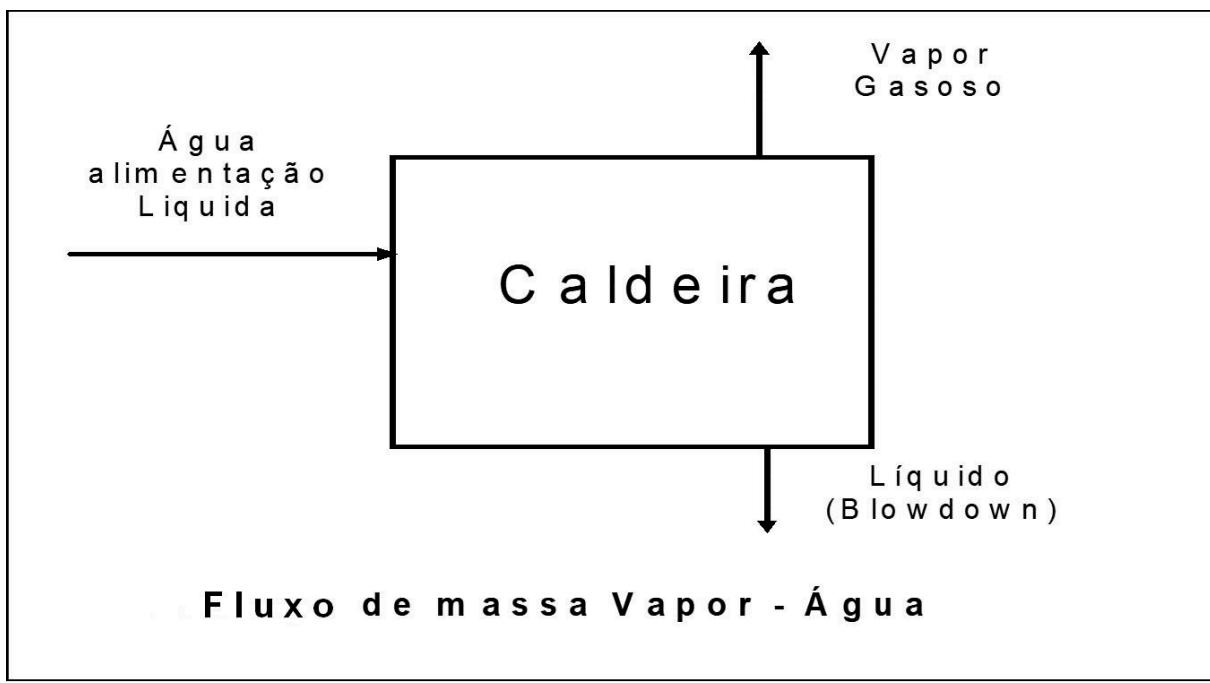
Nas caldeiras, de modo geral, os fluxos de entradas e saídas são vitais ao seu funcionamento e performance. As Figuras 5, 6, 7 exemplificam de maneira sucinta esses fluxos considerados em uma caldeira.

Figura 5 – Fluxo do ar, combustível e gases queimados.



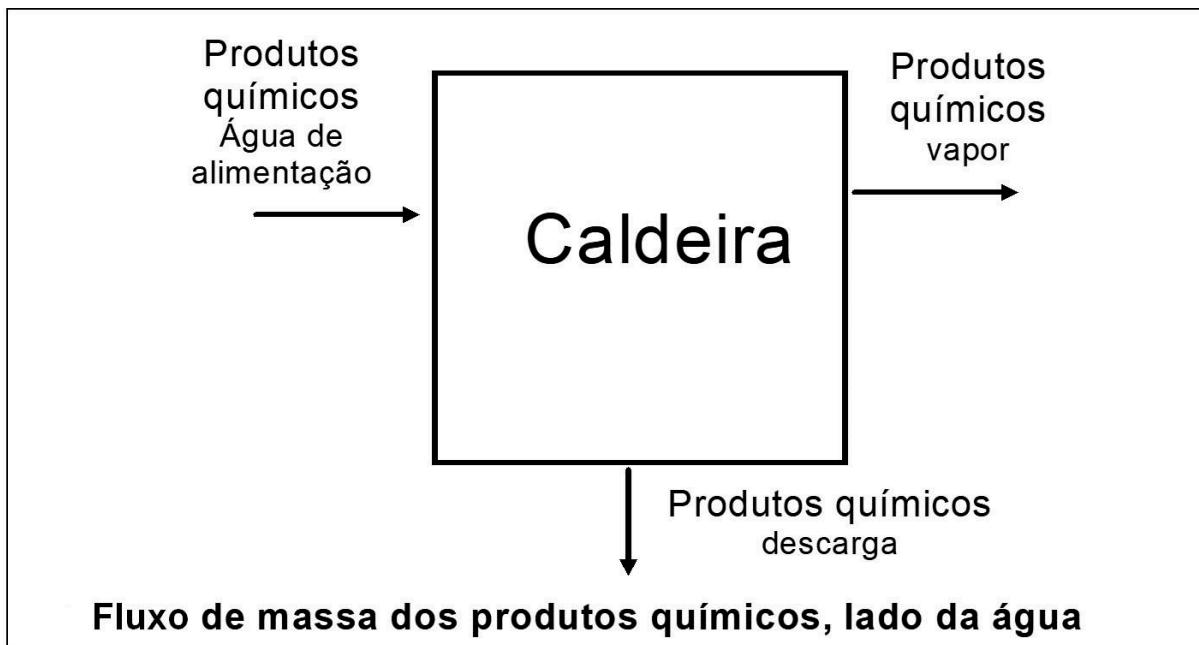
Fonte: (BEGA,2003)

Figura 6 – Fluxo de vapor e água.



Fonte: (BEGA,2003)

Figura 7 – Fluxo dos produtos químicos.



Fonte: (BEGA,2003)

Para a produção de potência, o vapor é utilizado em máquinas alternativas e em turbinas (Figura 8), sendo que nestas o vapor é em geral superaquecido. Exemplos de máquinas alternativas: prensas, martelo para forjaria, locomóveis, locomotivas, etc.

Figura 8 – Ilustração de uma turbina a vapor.



Fonte: Catálogo Ilustrativo GE

O vapor pode ser empregado também para extração de gases não condensáveis, tais como o ar, dos espaços evacuados, nas chamadas bombas de jato. Outra aplicação muito comum do vapor é no bombeio da água de alimentação de caldeiras nos chamados injetores e turbo bombas.

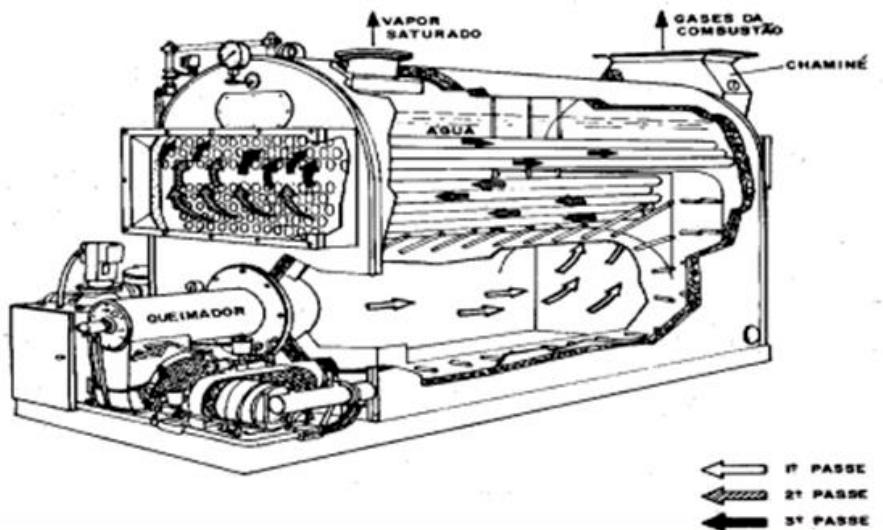
2.3 Tipos, classificação das caldeiras e seus acessórios principais

Nos geradores de vapor, a energia térmica é liberada por meio da queima de um combustível sólido (carvão mineral ou vegetal , lenha, biomassa, xisto , turfa), líquido (óleos, derivados do petróleo) ou gasoso (gás natural, hidrogênio, biogás);resistências elétricas (eletro termia) além de fontes não convencionais, como a fissão nuclear, energia solar, energia geotérmica (TURNS,2000).

A energia térmica liberada na queima de qualquer uma das fontes acima deve ser adequadamente transferida para as superfícies de absorção de calor.Existem diversos tipos de caldeiras, as quais podem ser classificadas segundo diversos critérios. Segundo uma classificação mais genérica, as caldeiras classificam-se em: flamotubulares ou fogotubulares e aquotubulares. As caldeiras fogotubulares ou flamotubulares caracterizam-se pela circulação

interna dos gases de combustão, ou seja, os tubos conduzem os gases por todo o interior da caldeira, como mostra a Figura 9. Estas são construídas para operar com pressões limitadas, pois as partes internas submetidas à pressão são relativamente grandes, inviabilizando o emprego de chapas de espessuras maiores.

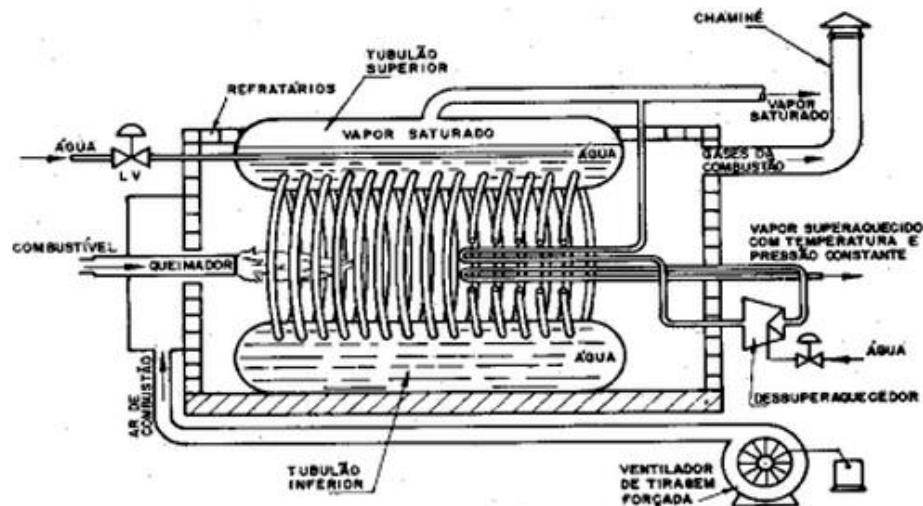
Figura 9 – Caldeira Flamotubular.



Fonte: (BEGA,2003)

As caldeiras aquotubulares caracterizam-se pela circulação externa dos gases de combustão, os tubos conduzem massa de água e vapor. A Figura 10 ilustra esse tipo de caldeira, cuja produção de vapor é maior que a das fogotubulares. As caldeiras aquotubulares são de utilização mais ampla, pois possuem internamente vasos pressurizados e de menores dimensões relativas. Isso viabiliza econômica e tecnicamente o emprego de maiores espessuras e, portanto, a operação em pressões mais elevadas.

Figura 10 – Caldeira Aquotubular.



Fonte: (BEGA,2003)

Existem, embora sejam raras, caldeiras que possuem partes fogotubulares e partes aquotubulares, constituindo, dessa forma, o que se pode denominar de caldeiras mistas. Finalmente, para os propósitos da NR 13, as caldeiras são classificadas em três categorias , conforme segue:

- Caldeiras da categoria A: são aquelas cuja pressão de operação é igual ou superior a 1960 kPa (19,6 bar = 19,98 kgf/cm²).
- Caldeiras da categoria C: são aquelas cuja pressão de operação é igual ou inferior a 588 kPa (5,88 bar = 5,99 kgf/cm²) e o volume interno é igual ou inferior a 100 litros.
- Caldeiras da categoria B: são todas aquelas que não se enquadram nas categorias anteriores. (NR 13,2018). As unidades geradoras de vapor necessitam de equipamentos para controle e operação, tais como:
 - Sistema de controle da água de alimentação.
 - Sistema de controle da combustão.
 - Indicadores de nível.
 - Válvulas de segurança.
 - Sopradores de fuligem.
 - Injetores.
 - Outros.

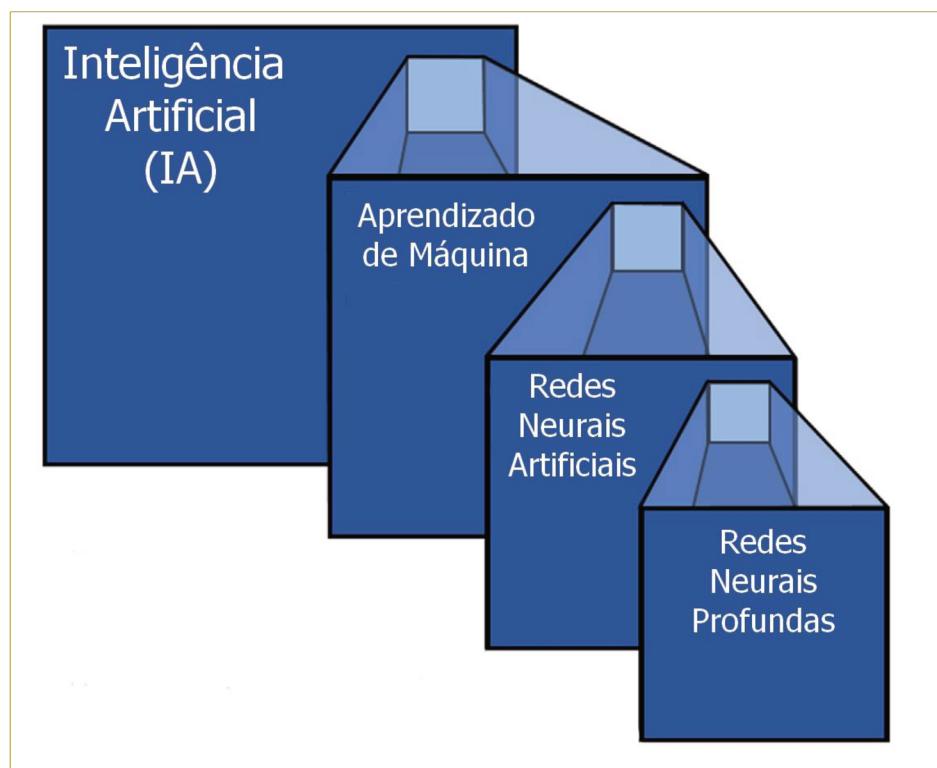
No caso da produção de etanol de milho, é importante ter conhecimento e domínio das tecnologias empregadas no processo de produção, que apresentam diferenças significativas em relação à produção de etanol de cana-de-açúcar. O setor de Utilidades (água, vapor, ar comprimido, energia elétrica, etc., ou seja, formas de energia utilizadas na produção final), a energia primária e secundárias são vitais e as caldeiras destacam-se dentro do sistema de cogeração (PROCKNOR,2020). No estudo das caldeiras, podem existir diferentes resultados em função do tipo de combustível e das entradas e saídas, por se tratar de um processo muito dinâmico e com muitas variáveis com efeitos cruciais no custo de produção e eficiência do sistema. Os equipamentos são de alto custo de implantação e exigem continuidade operacional, segurança do equipamento, treinamento e capacitação pessoal, excelência operacional, além de um sistema automático de monitoramento da manutenção, operação e processo.

2.4 Uso de aprendizagem de máquina e métodos estatísticos

2.4.1 Inteligência Artificial

A Figura 11 mostra os blocos originários da IA (Inteligência Artificial), que é uma disciplina acadêmica fundada em 1956, destinada à simulação da inteligência humana com máquinas. O aprendizado de máquina é uma das metas da IA e um campo da ciência computacional, que evoluíram dos campos de reconhecimento de padrões e teoria de aprendizado computacional. Aplica-se a técnicas estatísticas para permitir modelos que progressivamente melhoraram a performance. Tem-se as redes neurais artificiais e as redes neurais profundas (RUSSELL, *et al.*, 2004).

Figura 11 – Blocos da AI (Artificial Intelligence).



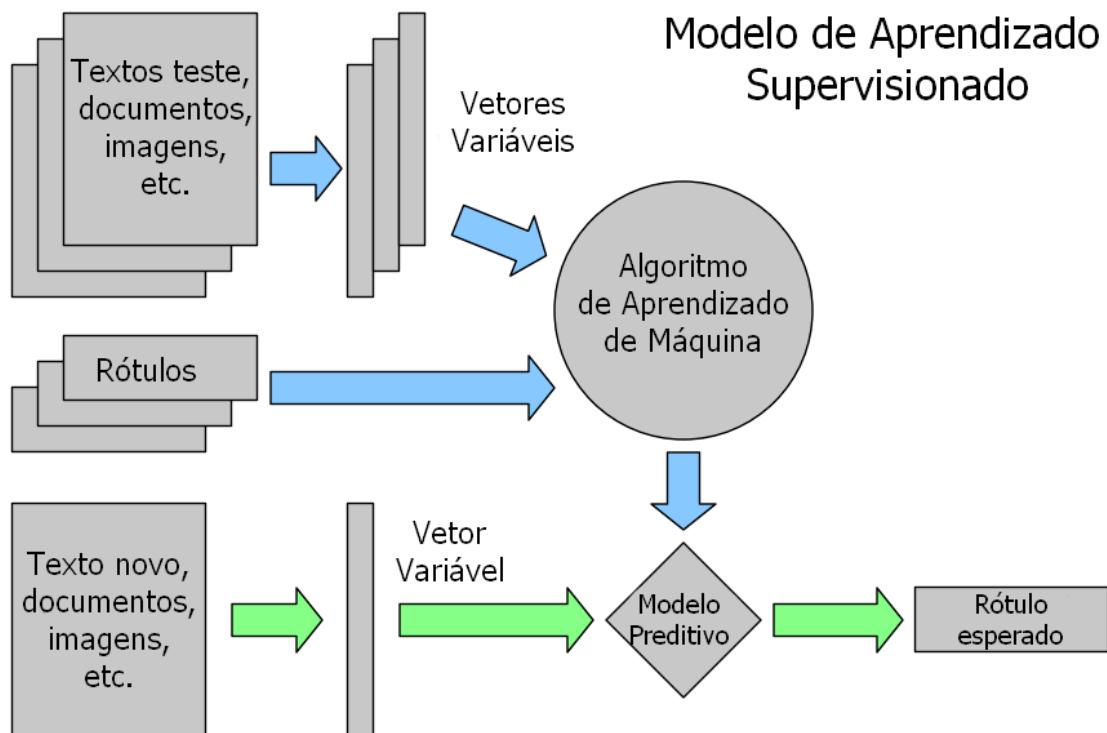
Fonte: Bioorganic and Medicinal Chemistry Letters 28, 2018.

Inteligência Artificial: Disciplina Acadêmica, fundada em 1956, com a finalidade de simular a inteligência humana com máquinas.

Aprendizado de Máquina: Um dos objetivos da IA e um campo da ciência da computação que evoluiu, dos campos de reconhecimento de padrões e da teoria de aprendizado computacional. Aplica técnicas estatísticas que permitem modelos que melhora a performance progressivamente.

Redes Neurais Artificiais: Um algoritmo de aprendizagem levemente inspirado na redes neurais biológicas: são ferramentas de modelos de dados estatísticos não lineares, aplicados em tarefas de Aprendizado de Máquina. Tem-se várias formas de tratar os dados no aprendizado de máquina. Abaixo na Figura 12 temos o modelo de aprendizado supervisionado que hora foi utilizado nesta dissertação.

Figura 12 – Modelo de aprendizado supervisionado.



Fonte: RUSSELL,2004

2.4.2 XGBoost

O XGBoost- eXtreme Gradient Boosting- é um sistema de aprendizado de máquina escalável para o aumento da previsibilidade. XGBoost, como outros modelos de produtividade, combina as previsões de vários modelos para resultar em um modelo com melhor desempenho preditivo (RIDGEWAY,2019). XGBoost (como resultado do trabalho como modelo não

paramétrico sem premissas formais de distribuição) pode manipular dados assimétricos e multimodais, bem como dados categóricos ordinais ou não ordinais (BCCVL,2019). As vantagens que podemos destacar deste sistema. Pode ser usado com vários tipos de resposta (binomial, gaussiano, poisson).

Estocástico, melhora o desempenho preditivo, o melhor ajuste é detectado automaticamente pelo algoritmo. O modelo representa o efeito de cada preditor após contabilizar os efeitos de outros preditores. Robusto para valores, com alto desvio e fora da curva (BCCVL, 2019).

2.4.3 Quimiometria

A quimiometria é uma área que se refere à aplicação de métodos estatísticos e matemáticos, assim como aqueles baseados em lógica matemática, a problemas de origem química. Com a sofisticação crescente das técnicas instrumentais, impulsionada pela invasão de microprocessadores e microcomputadores no laboratório químico, tornaram-se necessários tratamentos de dados mais complexos do ponto de vista matemático e estatístico, a fim de relacionar os sinais obtidos (intensidades por exemplo) com os resultados desejados (concentrações). (UNICAMP, 2005 apud NASCIMENTO, 2005)

“De todos os ramos da química clássica, talvez a química analítica tenha sido a mais afetada pelo desenvolvimento recente da instrumentação química associada a computadores”. (BRUNS; FAIGLE, 1985 apud VALDERRAMA, 2015)

De fato, a “Chemometrics Society”, organização internacional dedicada ao uso e desenvolvimento de métodos em Quimiometria, é composta principalmente por químicos interessados em problemas analíticos. Atualmente é muito raro de encontrar qualquer periódico respeitável sobre pesquisas em química analítica, que não traga artigos reportando dados obtidos com o auxílio de microcomputadores, ou tratados por matemática multivariada ou métodos estatísticos, sempre com o objetivo de melhorar a qualidade dos resultados ou facilitar a sua interpretação. (FERREIRA, et al., 1999). Quimiometria é o campo da química que utiliza ferramentas estatísticas e matemáticas para o planejamento e otimização das condições experimentais, e para a extração de informação química relevante de dados químicos multivariados. A diferença entre dados univariados e multivariados é que nos primeiros, a análise dos resultados é feita pela observação do comportamento de uma única variável de cada vez, por exemplo, a concentração de uma espécie de interesse ou uma propriedade físico-química (densidade, viscosidade, ponto de fusão, ponto de ebulição). Nos dados multivariados, é possível analisar mais de uma variável simultaneamente, e assim identificar a correlação entre elas. Este tipo de análise permite um entendimento mais completo e sistematizado dos resultados analíticos. (RIBEIRO, et al., 2007, pg. 14-15).

Vale ressaltar que o uso da Quimiometria está centrado no nível acadêmico (superior) e industrial.

2.5 Modelagem e estudo de caldeiras utilizando aprendizado de máquina.

Martinelli; Neitzel; Vieira (2010) utilizaram duas redes neurais que foram programadas em plataforma Excel para possibilitar o monitoramento contínuo do processo produtivo. Para a alimentação contínua dos dados de processo no sistema de monitoramento em Excel, foi utilizado o programa comercial *Plant Information* (PI). (MARTINELLI; NEITZEL; VIEIRA, 2010)

Tan *et al.* (2019) analisaram que a emissão de gases e particulados na saída da caldeira representa um desafio substancial para a operação desse equipamento. Um modelo preciso de previsão de emissões de particulados e gases sob estados estáveis e transitórios é crítico para resolver essa questão. (TAN *et al.*, 2019)

Garami *et al.* (2018) demonstraram a utilização de um sistema de monitoramento de combustão baseado em uma visão de máquina em caldeiras de biomassa, monitorou-se uma caldeira operada em cargas variadas e com qualidade de combustível flutuante, evitando emissões excessivas e para melhorar a economia. (GARAMI *et al.*, 2018)

Para otimizar a combustão em uma caldeira de carvão utilizando Redes Neurais Artificiais (RNA), modelos são estabelecidos para prever as propriedades de operação e emissão da caldeira. As previsões da eficiência térmica e das emissões de NOx mostram uma boa concordância com os dados da planta, com erros médios relativamente baixos respectivamente. (SHI *et al.*, 2019)

Deboni (2019), em um estudo, tiveram como objetivo analisar a evolução da qualidade de diferentes tipos de resíduos florestais e madeiras utilizados para geração de energia em uma planta de cogeração, especificamente a umidade da biomassa e a geração de cinzas. (DEBONI *et al.*, 2019).

Embora existam estudos para caldeiras utilizadas na produção de etanol de cana-de-açúcar, a realidade para o etanol de milho no Brasil ainda é muito recente. Normalmente, usamos o chip de eucalipto para queima nas fornalhas se a planta é autônoma, ou seja, somente produz etanol de milho; se temos uma planta “flex” utiliza-se o bagaço de cana , que sobrou da safra normal da cana de açúcar, para queima na fornalha. Nos EUA e Argentina , queima-se nas fornalhas desses equipamentos o gás natural. Para viabilizar a produção com o milho no Brasil, é de extrema importância um sistema térmico equilibrado e a cogeração, que é a obtenção de duas formas de energia (térmica e elétrica) a partir de uma fonte de energia primária

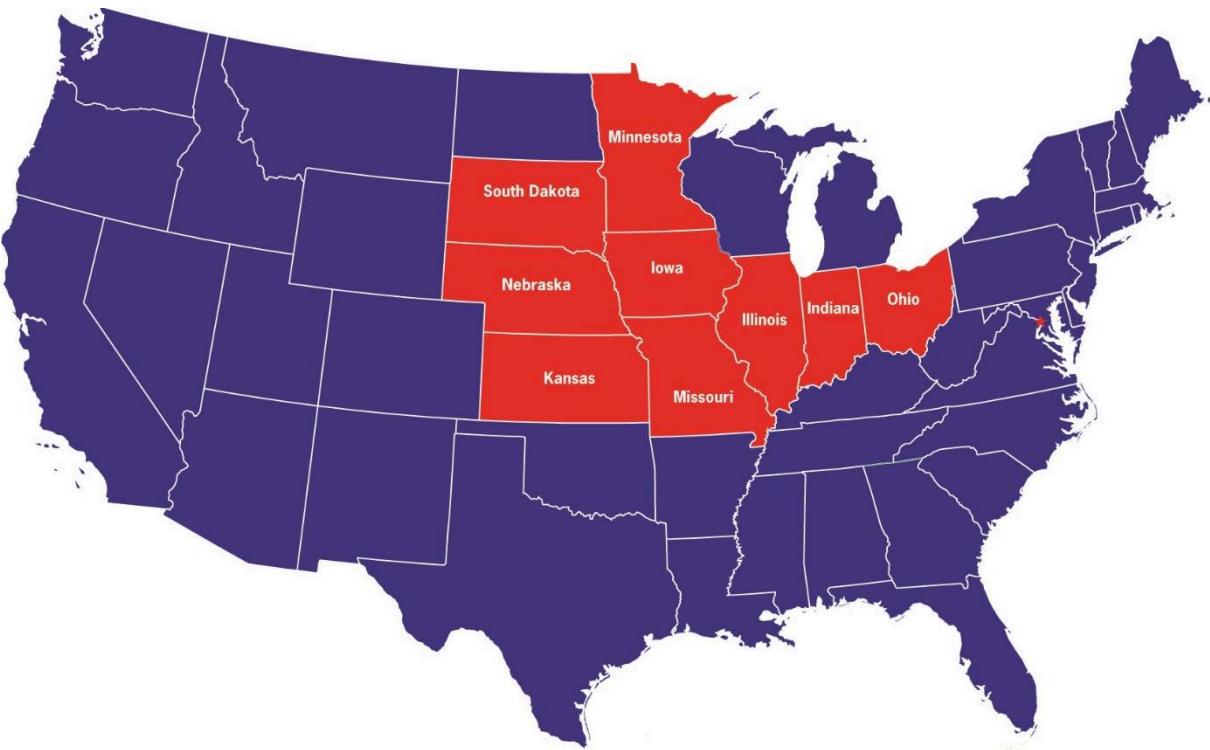
(combustível), e, nesse contexto, as caldeiras e turbinas têm que ser as mais eficientes possíveis, pois compra-se energia primária de fontes renováveis ou não e esta tem que ser muito bem aproveitada. A utilização de biomassa em caldeiras oferece muitos benefícios econômicos, sociais e ambientais, conservação de recursos de combustíveis fósseis, criação de oportunidades de emprego e redução de emissões de CO₂ e NOx.

2.6 ETANOL E BIOMASSA

2.6.1 Aspectos relevantes das produções do etanol de milho nos EUA e da cana-de-açúcar no Brasil

As duas fontes mais comuns na produção do etanol são a cana-de-açúcar e o milho. O Brasil lidera a produção com o etanol da cana e os Estados Unidos do milho (UNEM, 2020) . O etanol de milho é o tipo mais comum de etanol nos Estados Unidos, mas é considerado menos eficiente do que outros tipos de etanol (de cana-de-açúcar aqui no Brasil), porque apenas o grão é usado e muitos produtos e insumos à base de petróleo, fertilizantes, pesticidas, etc. são utilizados na sua produção.O milho utilizado para a produção de etanol é retirado de uma área conhecida como Cinturão do Milho (*Corn Belt*), uma região localizada nos EUA e especializada no cultivo desse grão. A região foi determinada pelo governo pois possui o clima mais adequado para essa plantação. (WORDPRESS ,2019)

Figura 13 – Estados que formam o *Corn Belt* (Cinturão do Milho) nos Estados Unidos.



Fonte: <https://intercambiadasoja.wordpress.com/2011/08/25/os-cinturoes-agricolas-dos-estados-unidos/>. Acesso em 22 de agosto de 2019.

O Cinturão do Milho (*Corn Belt*) inclui partes da Dakota do Sul, Nebraska, Kansas, Minnesota, Missouri, Iowa, Indiana, Illinois e Ohio, conforme mostra a Figura 13 acima. Aproximadamente 50% de todo o milho plantado nos EUA é proveniente dos quatro últimos estados mencionados. Quanto às colheitas no Brasil, o milho tem seu cultivo variado em função da época de chuva, produzindo, portanto, por um tempo menor. A safra principal ocorre durante o verão e a chamada safrinha, após o verão (UNEM ,2020). No Brasil, os estados de Goiás, Mato Grosso e Paraná estão a frente, tanto nas usinas “flex”, que moem cana-de-açúcar e milho (após a safra normal da cana-de-açúcar) produzindo etanol e coprodutos, quanto nas usinas denominadas “autônomas ou “full milho”” que só moem milho, produzindo etanol ,coprodutos do milho e energia elétrica o ano todo.

2.6.2 Produção do etanol de milho

O etanol é produzido do milho como uma biomassa industrial por meio de fermentação,tratamento químico e destilação. A produção de etanol pode ocorrer por dois métodos de transformação do milho: moagem seca ou úmida. No processo à seco, são obtidos produtos tais como: canjica especial, canjicas para cereais matinais e para produção de pipocas

expandidas, canjicão, *gritz* de milho, óleo de milho bruto e refinado e a sêmola de milho. Esse processo não requer muita tecnologia e também não gera muitos produtos sofisticados. Os derivados tecnologicamente superiores são oriundos do processamento via moagem úmida (ADEN,2008).

No processo via moagem úmida, o etanol é obtido após a hidrólise do amido, liberando as moléculas de açúcares que são transformados em álcool pelo processo de fermentação. Nesse processo, o milho passa por uma moagem e sai em forma de pó. Uma mistura feita com o pó desse grão, água e enzimas entram em um tanque de aquecimento alto, em que é liquefeita. Duas enzimas (gluco e alfa amilase) ajudam a quebrar o composto do grão no processo de liquefação e convertem o amido em açúcares, que podem ser fermentados para fabricar álcool. A levedura é acrescentada à mistura de açúcar para iniciar o processo de fermentação. Os açúcares são quebrados em etanol e em dióxido de carbono. A mistura fermentada é destilada. O etanol é separado dos sólidos e em um processo de desidratação para gerar o etanol anidro (ADEN,2008).

A Figura 14 abaixo ilustra o processo do refino do milho.

Figura 14 – O processo de refino do milho.



Fonte: <http://sinueloagropecuaria.com.br/wp-content/uploads/2016/09/o-processamento-do-milho-agro-ind.>

2.6.3 Viabilidade da produção de etanol de milho nos EUA

Estudos concluem que a indústria do etanol de milho apenas prospera e se torna viável nos EUA, cuja produção média por hectare é de 10 toneladas, por causa dos subsídios (o fornecimento de fundos monetários a certas indústrias e que possuem o intuito de abaixar o preço final dos produtos vendidos por tais companhias, para que esses produtos possam competir com os produzidos em outros países a preços menores) à prática, que atingiram U\$ 12 bilhões em 2018 (CANALRURAL,2019). A maioria das usinas de etanol depende da queima de gás natural ou, cada vez mais, de carvão mineral para gerar o vapor utilizado na destilação, adicionando emissões de combustíveis fósseis ao dióxido de carbono produzido pela levedura.(CANALRURAL,2019).

O cultivo de milho requer, ainda, fertilizantes com nitrogênio, fabricados com gás natural e o emprego intensivo de equipamentos agrícolas movidos a óleo diesel. Alguns estudos sobre o equilíbrio energético do etanol de milho apontam a irracionalidade do processo, que requer mais combustível fóssil emissor de carbono do que aquele que viria a substituir. Outros estudos, porém, indicam uma ligeira vantagem positiva. (SF AGRO, 2017).

A Agência de Proteção Ambiental, a Secretaria de Agricultura e a Secretaria de Energia do governo norte-americano anunciaram as novas referências, segundo as quais o etanol de milho reduz em 12% as emissões de poluentes (em comparação com a gasolina), enquanto o de cana reduz em 76% a depender do processo produtivo utilizado e o tipo da matéria-prima(GREGO,2018). Esses índices levam em conta a emissão de poluentes durante o transporte e a distribuição, pela queima de combustível nas usinas. Mas consideram, também, o controverso cálculo do uso indireto da terra: com o aumento da demanda por milho para produzir etanol, aumenta o preço dessas commodities e cresce a área cultivada em outros lugares, o que causa desmatamento e, consequentemente, emissão de poluentes.(GREGO,2011) . Para ser "avançado", o combustível precisa reduzir em pelo menos 50% a emissão de poluentes, com tolerância de 10 pontos porcentuais, que estão na meta da lei de combustíveis renováveis. O etanol de milho ficou de fora, com apenas 12%. Para ser qualificado apenas como “renovável”, o combustível precisa reduzir em pelo menos 20% a emissão de poluentes, como o etanol de milho só reduz emissões em 12%, está comprometida toda a justificativa ambiental do “lobby” do milho para receber subsídios e manter tarifas sobre o etanol importado.(SF Agro, 2017). Segundo pesquisadores, os produtores ,sob pressão econômica para produzir biocombustíveis irão usar cada vez mais áreas de florestas e pastos, liberando muito do carbono anteriormente armazenado nas plantas e solos por meio de decomposição ou queimadas. Em um nível global, mais áreas de pasto e floresta serão

convertidas ao cultivo de lavouras para substituir a perda de grãos quando os produtores dos Estados Unidos converterem mais terras para os biocombustíveis. (CANAL RURAL, 2014)

2.6.4 Vantagens e Desvantagens – milho x cana-de-açúcar no Brasil

Ambos os processos, são feitos por meio da fermentação. No caso do milho, é necessário quebrar as grandes moléculas de amido antes de fermentá-lo até se transformar em açúcar, o que requer mais tecnologia para a produção. Já com a cana, o processo é direto, assim, agilizando o procedimento (BNDES & CGEE, 2008) . A fertilização do solo da lavoura de cana-de-açúcar é complementada com os resíduos do processo de industrialização, como a torta de filtro e a vinhaça. Os resíduos do etanol de milho são aproveitados na fabricação de ração animal (*Distilled Dried Grains with Solubles - DDGS*) e óleo de milho. Quanto às colheitas e plantio da cana, a cana-de-açúcar é uma cultura semi-perene, por isso, só precisa ser replantada a cada seis anos em média, o que ajuda na conservação do solo, diferentemente do grão, que é replantado anualmente. O período de produção do etanol da cana-de-açúcar varia de local para local, mas ocorre normalmente no Brasil, geralmente de abril a novembro. A cana colhida não pode ser estocada, devendo ser imediatamente processada. Já o milho tem seu cultivo variado em função da época de chuva, produzindo, portanto, por um tempo menor. A safra principal ocorre durante o verão e a chamada safrinha, após o verão. A vantagem é que ele pode ser estocado em silos.(IMEA,2019).

Apesar de a tonelada de milho render em média 430 litros de etanol e a da cana apenas 85 litros – a depender do tipo de processo/tecnologia, mix de produção e equipamentos utilizados-, a área explorada pela cana é menor. Como as áreas são pequenas e a produtividade é alta (planta-se mais vezes no ano), o impacto na agricultura é menor. No Brasil, cada hectare plantado com cana de açúcar, produz de 60 a 120 toneladas de cana, dependendo da fertilidade do solo, perfil varietal, manejo no cultivo, condições edafoclimáticas, dentre outros fatores; já o milho como média geral Brasil, um hectare produz cerca 6 toneladas (EMBRAPA,2020) . Na Figura 15, demonstramos a conversão do milho em produtos e coprodutos com suas respectivas relações de produção.

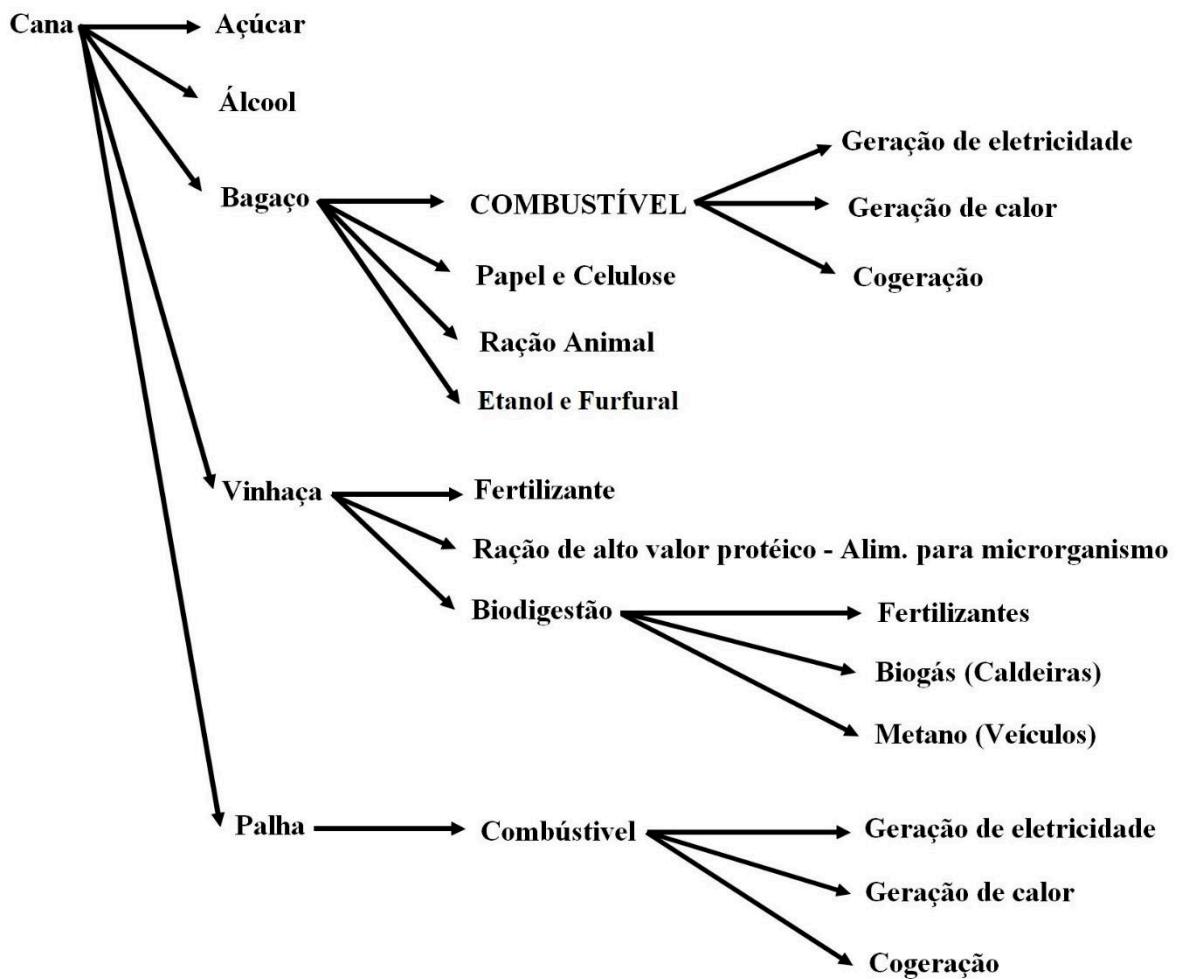
Figura 15 – Conversão do milho em produtos e coprodutos com suas respectivas relações de produção.



Fonte: Reprodução de figura do Catálogo Comercial da FS Bioenergia. Imagens de demonstração retiradas do site <http://www.fsbioenergia.com.br/pt-br/produto/2/59>.

Na Figura 16, demonstramos o aproveitamento de produtos e coprodutos da cana-de-açúcar.

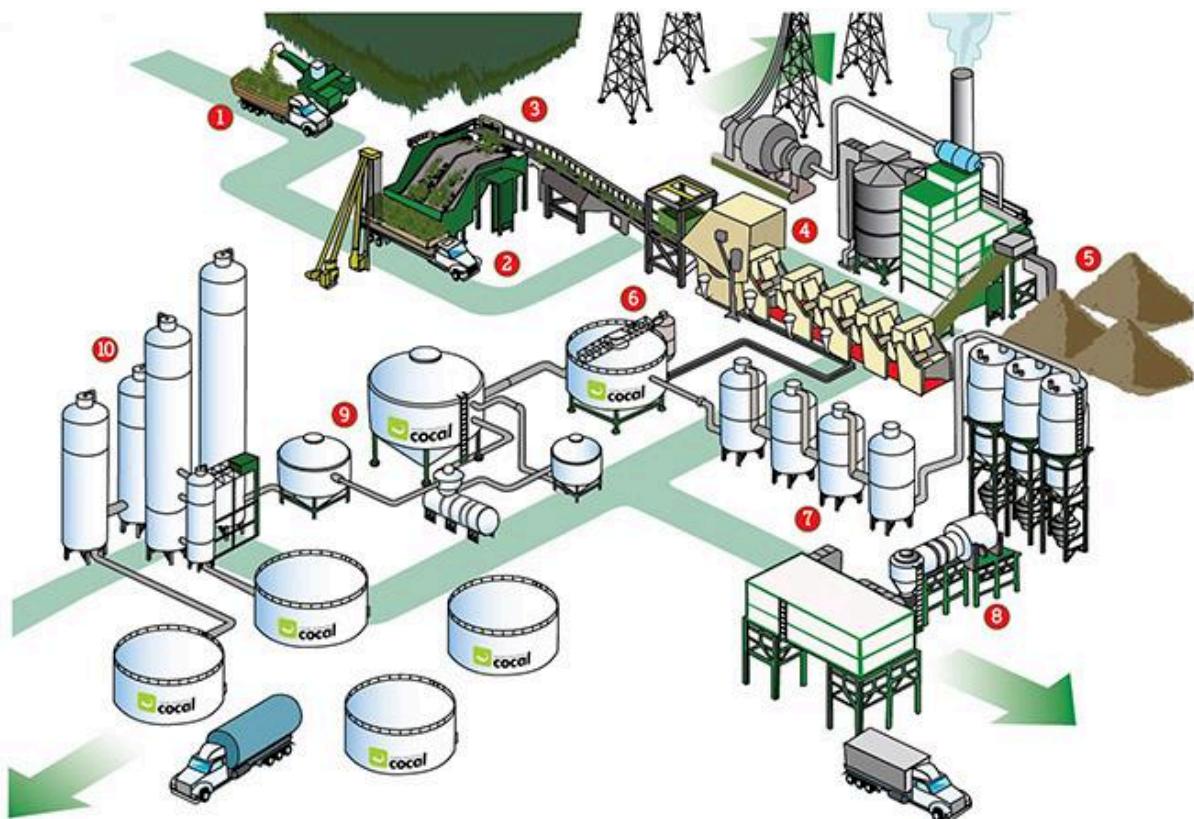
Figura 16 – Aproveitamento de produtos e coprodutos da cana-de-açúcar.



Fonte: (BNDES&CGEE,2008)

Na Figura 17, demonstramos as etapas de um fluxograma básico, de uma indústria típica ,de fabricação do etanol, açúcar e energia provenientes da cana-de-açúcar.

Figura 17 – Fluxograma de Fabricação: Etanol, Açúcar e Energia.



Fonte: <http://www.cocal.net.br/conteudo/21/3/fluxograma-de-fabricacao.html>

- ① Colheita da cana-de-açúcar.
- ② Transporte da cana-de-açúcar até a área industrial.
- ③ A cana-de-açúcar é descarregada, passando pelo picador e desfibrador, e transportada para moagem, onde se inicia o processo de fabricação de etanol, de energia elétrica e do açúcar.
- ④ A cana-de-açúcar é preparada para a extração do caldo, por meio da moenda ou difusor. No final dessa fase o caldo é separado do bagaço.
- ⑤ O bagaço é queimado nas caldeiras para gerar o vapor. Este vapor de alta pressão é enviado às turbinas dos geradores de eletricidade, transformando grande parte da energia térmica em energia elétrica. Cerca de 35% de energia elétrica gerada é utilizada na unidade para a fabricação de açúcar e etanol e 65% é destinada para a iluminação de residências, cidades ou indústrias de diversos segmentos.
- ⑥ O caldo que sai do processo de extração passa por um tratamento para remover as impurezas minerais. São utilizados aquecedores, decantadores, turbo filtros e filtros para clarificar o caldo, deixando-o preparado para o processo produtivo de etanol e açúcar.

- 7 O caldo clarificado entra em um conjunto de evaporadores de múltiplo efeito para concentração do caldo onde se evapora grande parte da água, tornando-o na consistência de um xarope. Este xarope é bombeado aos tachos de cozimento para a cristalização do açúcar (sacarose).
- 8 O uso das centrífugas proporciona a separação dos cristais de açúcar do líquido açucarado, denominado mel ou melaço. Posteriormente, o açúcar passa no secador para a retirada da umidade contida nos cristais. Na saída do secador, o açúcar é enviado por esteiras de uso alimentício até o silo ou armazéns, de onde é feito o ensacamento ou a expedição a granel.
- 9 A fermentação é em sistema de batelada e utiliza o caldo tratado das leveduras (fermento). A reação de transformação dos açúcares em álcool e gás carbônico ocorre nos fermentadores, este é um processo bioquímico realizado pelas leveduras. Após a fermentação, a levedura é separada do vinho e recuperada por meio de um processo de centrifugação. As leveduras recuperadas em pré-fermentadores são tratadas e retornam ao processo fermentativo.
- 10 O vinho em torno de 9%/11% de álcool, sem leveduras, vai para os aparelhos de destilação onde o álcool é separado, concentrado e purificado até 96% para o etanol hidratado ou 99,6% para o etanol anidro quando se utiliza processos de desidratação do etanol. Em seguida, o combustível é armazenado em grandes tanques distintos, cuja capacidade alcança 20 mil/m³.

2.6.5 Equilíbrio energético do etanol de milho x cana-de-açúcar no Brasil

O balanço energético é o rendimento da fonte de energia primária e resultante, ou seja, estima-se a quantidade de energia fóssil gasta para a produção ,além do necessário para gerar o etanol, o que dá um índice alto, como é o caso da cana, com 9,0. Já o milho corresponde a 1,5, ou seja, quase seis vezes inferior. (GREGO, 2011; PIMENTEL, *et al.*, 2005)

2.6.6 Análise socioeconômica e ambiental da produção de etanol de milho no centro oeste brasileiro

A produção de etanol de milho no Brasil chama a atenção pela velocidade de expansão e pelos volumes de investimentos. A tecnologia ainda é uma novidade em solo brasileiro, cujos impactos são pouco conhecidos. Como um dos objetivos ,será analisar os impactos socioeconômicos, as emissões de Gases de Efeito Estufa (GEE) e uso da terra da implantação da indústria de etanol e de milho no Centro Oeste brasileiro, tendo como referência o caso concreto de uma indústria autônoma instalada em Mato Grosso (MT). A análise socioeconômica tem foco no médio prazo. Ela faz uso de uma matriz de insumo-produto inter-regional e avalia os impactos da instalação de uma usina de etanol de milho que produz 500

milhões de litros de etanol ao ano, além dos coprodutos DDGs (*Dried Distillers Grains*), óleo de milho bruto e eletricidade (HOFFMAN ,2011). Assumiu-se o cenário no qual tais produtos deslocam a demanda de etanol importado, ração animal de origem doméstica, óleo de soja bruto e eletricidade a partir do gás natural.

Outras possibilidades (como maior suplementação de animais) são possíveis, mas não foram apresentadas no estudo. Estima-se que a fase de investimentos pode gerar um total de aproximadamente 8,5 mil empregos diretos e indiretos ao longo de sua duração, sendo grande parte vinda de outros estados brasileiros. São movimentados aproximadamente R\$ 1,5 bilhão em toda a economia doméstica, com um valor da produção de R\$ 660 milhões em nível nacional e R\$ 80 milhões em impostos indiretos líquidos e impostos diretos. (MOREIRA; ARANTES, 2018). A operação da planta gera anualmente, base 2017, um valor de produção total de R\$ 2,5 bilhões e um PIB de R\$ 910 milhões. Quase 80% desses valores ficam dentro do estado do Mato Grosso, já a arrecadação aumenta em R\$ 73 milhões. Cabe relembrar que se trata de um acréscimo de tributação causado pelos efeitos indiretos na economia, adicionais aos R\$ 130 milhões anuais em ICMS e PIS-COFINS gerados diretamente pela atividade da usina de etanol de milho. O total de empregos aumenta em aproximadamente 4,5 mil postos de trabalho. Em uma abordagem atribucional, a pegada de carbono do etanol de milho no Brasil é calculada em 18 e 25,5 g CO₂eq/MJ (dois cenários distintos para a avaliação das emissões do milho). Esses valores representam uma redução de mais de 70% em comparação com a gasolina e são significativamente menores do que o etanol de milho produzido nos Estados Unidos. As principais razões são o uso de biomassa de eucalipto como fonte de energia e do milho de segunda safra, que otimiza recursos na sua rotação com a soja. Já numa abordagem consequencial, a pegada de carbono é de 3,1 g CO₂eq/MJ. A diferença é marcante devido ao tratamento dos coprodutos (bioeletricidade) e pelos efeitos diretos e indiretos de uso da terra. (MOREIRA; ARANTES, 2018). A expansão da área de eucalipto aumenta os estoques de carbono em comparação com outros usos antrópicos da terra. Por sua vez, a coprodução de DDGs desloca o uso de insumos utilizados para a produção de ração animal então existente, dentre eles, a soja. Uma menor expansão da área de soja leva à manutenção de áreas de pastagens, que tem maiores estoques de carbono quando comparados com lavouras anuais. Muito embora o modelo identifique conversão indireta de vegetação nativa, esse efeito é muito menor (em termos de carbono) que os efeitos apresentados anteriormente. (DONKE *et al* ,2017). Os resultados permitem a conclusão de que a indústria de etanol de milho tem grande potencial para gerar renda e emprego no estado de MT e no Brasil. Ao mesmo tempo, a análise de ciclo

de vida evidencia um potencial importante para a redução de emissões de GEE quando comparado com gasolina e até mesmo com etanol de milho dos EUA (RAUCCI *et al*,2015). Em particular, quando considerado o aspecto de mudança de uso da terra, são ressaltadas especificidades de agricultura no centro-oeste brasileiro, pouco compreendidas na literatura internacional. Há, portanto, evidências suficientes para recomendar o fomento da indústria de etanol de milho no Brasil nas condições avaliadas neste estudo. (UNEM,2018) . A tecnologia de conversão de milho em etanol e outros coprodutos é a tecnologia dominante nos EUA. Lá o processo de produção é caracterizado pelo uso de milho primeira safra, por usinas grandes e uso de fontes de energia tradicionais, em grande parte fósseis (i.e., carvão mineral e gás natural). Tais características são diferentes do setor de biocombustíveis nacional, esse último com claras vantagens em uso da terra e uso de fontes renováveis. Avanços recentes em tecnologia agrícola em ambiente tropical possibilitaram a disseminação de sistemas de produção em múltiplas safras, dos quais se destaca o sistema que combina soja de ciclo curto e milho de segunda safra. Além de uma maior produção por área, lavouras de segunda safra permitem uma melhor proteção do solo e a otimização dos recursos das fazendas. A adoção e aprimoramento desse pacote tecnológico levou à rápida expansão dos volumes produzidos de grãos no Centro Oeste do Brasil, que não foi acompanhada por sistemas de escoamento logístico. Como resultado, observou-se o acúmulo de estoques e a necessidade de acionar políticas para manutenção de preços. (MOREIRA; ARANTES, 2018). O etanol de milho foi inicialmente adotado no Brasil em usinas “flex”, aproveitando instalações, mão de obra e energia de usinas de cana-de-açúcar. Com os volumes atuais de produção de milho e preços relativamente baixos “na porteira das fazendas”, investidores passaram a dar passos mais ousados transferindo ao Centro Oeste brasileiro e, em particular ao estado de MT, novos pacotes tecnológicos estruturados em usinas grandes (produção de 250 a 500 milhões de litros por ano), que deverão se adaptar às condições locais. Pairam, porém, dúvidas sobre os efetivos impactos da implementação dessa indústria, que só o tempo nos mostrará.

Figura 18 – Principais estados brasileiros produtores de milho.



Fonte: CONAB, 2019.

A Figura 18 mostra os principais estados brasileiros produtores de milho.

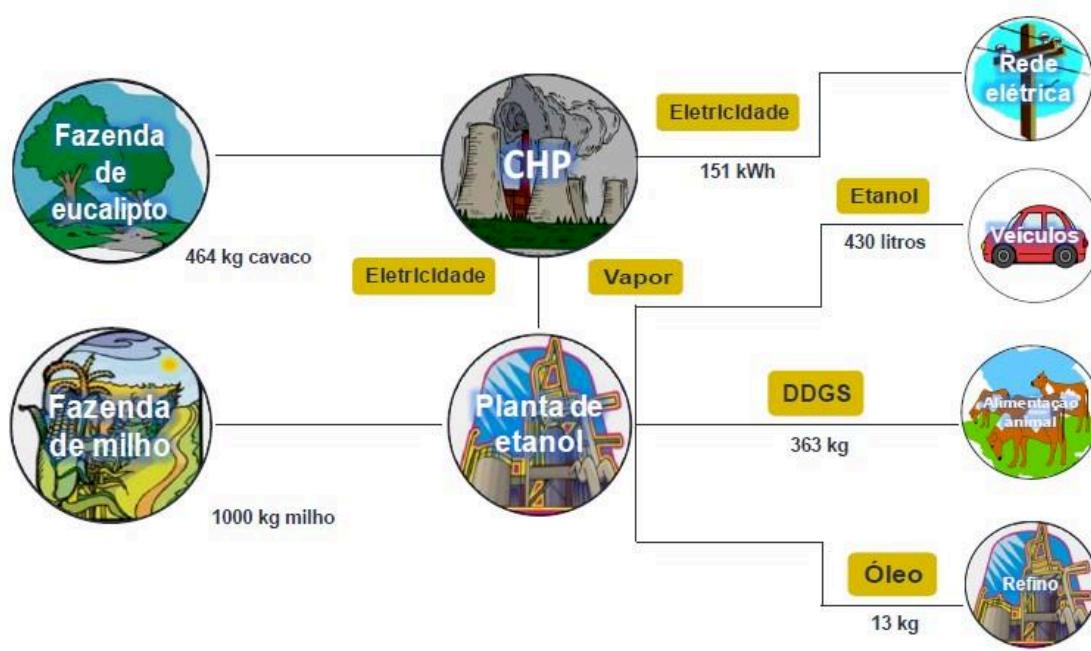
Os investimentos em etanol de milho estão concentrados em MT por diferentes razões. O estado ocupa a primeira posição em termos de produção de milho no país, registrando, na safra 2016/2017, um recorde de 29 milhões de toneladas. No estado, quase todo o milho é de segunda safra e encontra dificuldades logísticas para escoamento. Nesse contexto, há uma oportunidade para um modelo de negócios que tem como estratégia o beneficiamento local do milho em etanol e coprodutos, agregando valor ao insumo agrícola. Os principais mercados consumidores são o próprio estado de MT, assim como as regiões do Arco Norte brasileiro (IMEA,2017).

O processo de fabricação no Brasil guarda semelhanças com a tecnologia nos EUA, porém com algumas diferenças importantes. O milho é moído e liquefeito em água, em seguida, segue para a etapa de fermentação que, com adição de enzimas e leveduras, transforma o amido em açúcares e, posteriormente, em etanol. A Tecnologia de Separação de Fibras permite recuperar fibras e proteínas, resultando na coprodução de produtos ricos em nutrientes comumente conhecidos como DDGs em seu acrônimo em inglês (FS Bioenergia,2017). A energia utilizada pelo processo, que é o vapor e eletricidade e gerada por uma termoelétrica (planta de cogeração) que utiliza biomassa de eucalipto e que fica anexa à fábrica de etanol.

O sistema energético deverá ser otimizado de maneira que a eletricidade gerada pela termoelétrica deverá ser superior à demanda dos processos de produção. A eletricidade excedente é vendida à rede distribuidora.

De acordo com dados da FS-Bioenergia, cada 1.000 quilos de milho produzem 430 litros de etanol anidro, 363 kg de DDGs e 13 quilos de óleo de milho bruto. Para a mesma quantidade de milho, são necessários 464 kg de cavaco de eucalipto para atender à demanda de energia da usina, as quais ainda permitem exportar 151 kWh de eletricidade excedente. (FS Bioenergia,2017) O diagrama do fluxo de materiais e energia é apresentado na Figura 19 a seguir.

Figura 19 – Diagrama do sistema de produção de uma usina de milho no Brasil.



Fonte: FS Bioenergia

3 OBJETIVOS

O principal objetivo desta dissertação é avaliar o desempenho de uma caldeira aquatubular- a primeira do Brasil com grelha vibratória com tecnologia Dan Power- de 67 kgf/cm² de pressão, temperatura do vapor produzido em 520 °C queimando biomassa (chip de eucalipto), operando dentro da uma planta industrial autônoma, produtora de biocombustíveis (no caso, etanol de milho), utilizando como ferramentas o Aprendizado de Máquina e a Quimiometria para avaliar o efeito das variáveis: temperatura dos gases efluentes, temperatura da fornalha, vazão de vapor e teor de oxigênio dos gases efluentes, no desempenho da caldeira.

Como objetivos específicos, será feito:

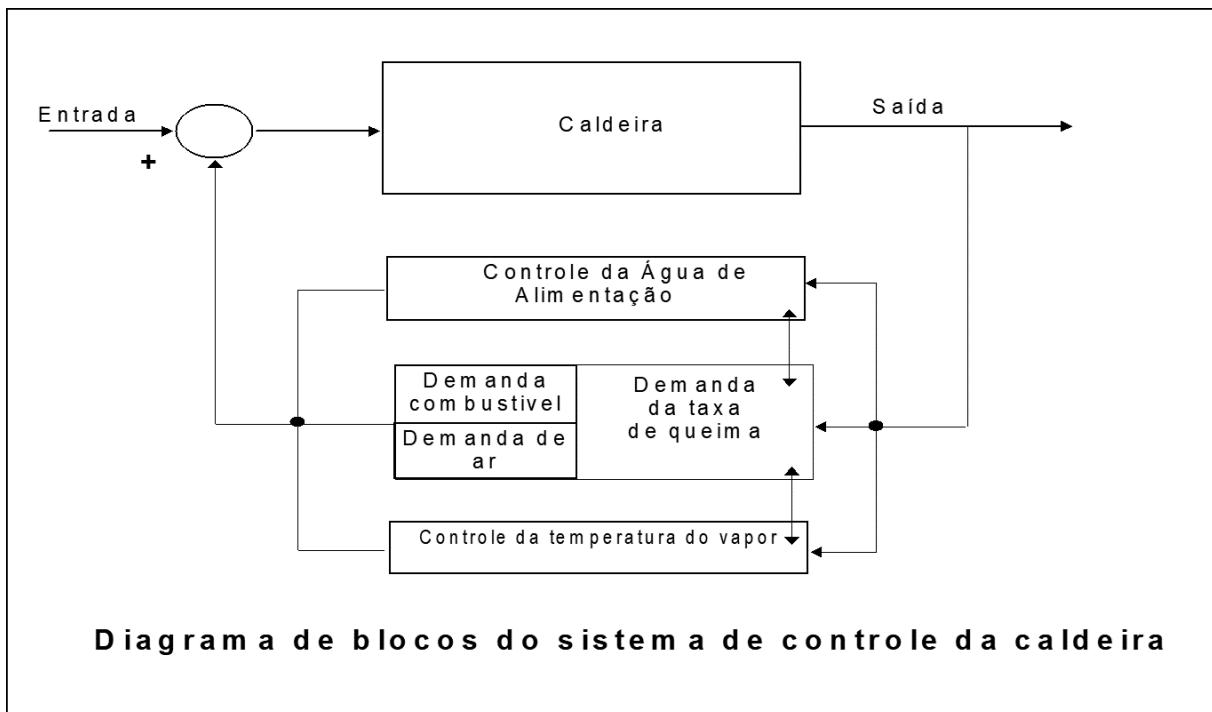
- d) Coleta, organização e descrição dos dados de entrada que representam o comportamento do equipamento;
- e) Seleção de metodologia de aprendizado de máquina e métodos estatísticos ;
- f) Avaliação da caldeira segundo as modelagens selecionadas;

4 MATERIAIS E METODOLOGIA

4.1 Coleta de dados históricos

A Figura 20 abaixo mostra um diagrama de bloco do sistema de controle da caldeira, em que grandes quantidades de dados de medição do processo de produção e conversão de energia de uma planta industrial moderna, no caso a FS Bioenergia- primeira planta industrial “full” etanol de milho localizada em Lucas do Rio Verde/MT; os dados estão disponíveis por meio do *Digital Control System* (DCS)/Supervisório e que podem ser usadas para determinar um modelo de *Machine Learning*(ML). Quando as ordens de carga da rede elétrica/processo alternam-se, a caldeira vai seguir essa ordem para ajustar a capacidade/demandas necessárias. Neste trabalho, os dados das caldeiras para um mês de operação foram obtidos a um intervalo de 20 segundos cada, que “guardam” as variáveis trabalhadas obtidas provenientes do DCS das caldeiras. Os parâmetros considerados incluem a carga da planta.. As coletas dos 60 pontos (30 dias) foram realizadas na safra /2019.

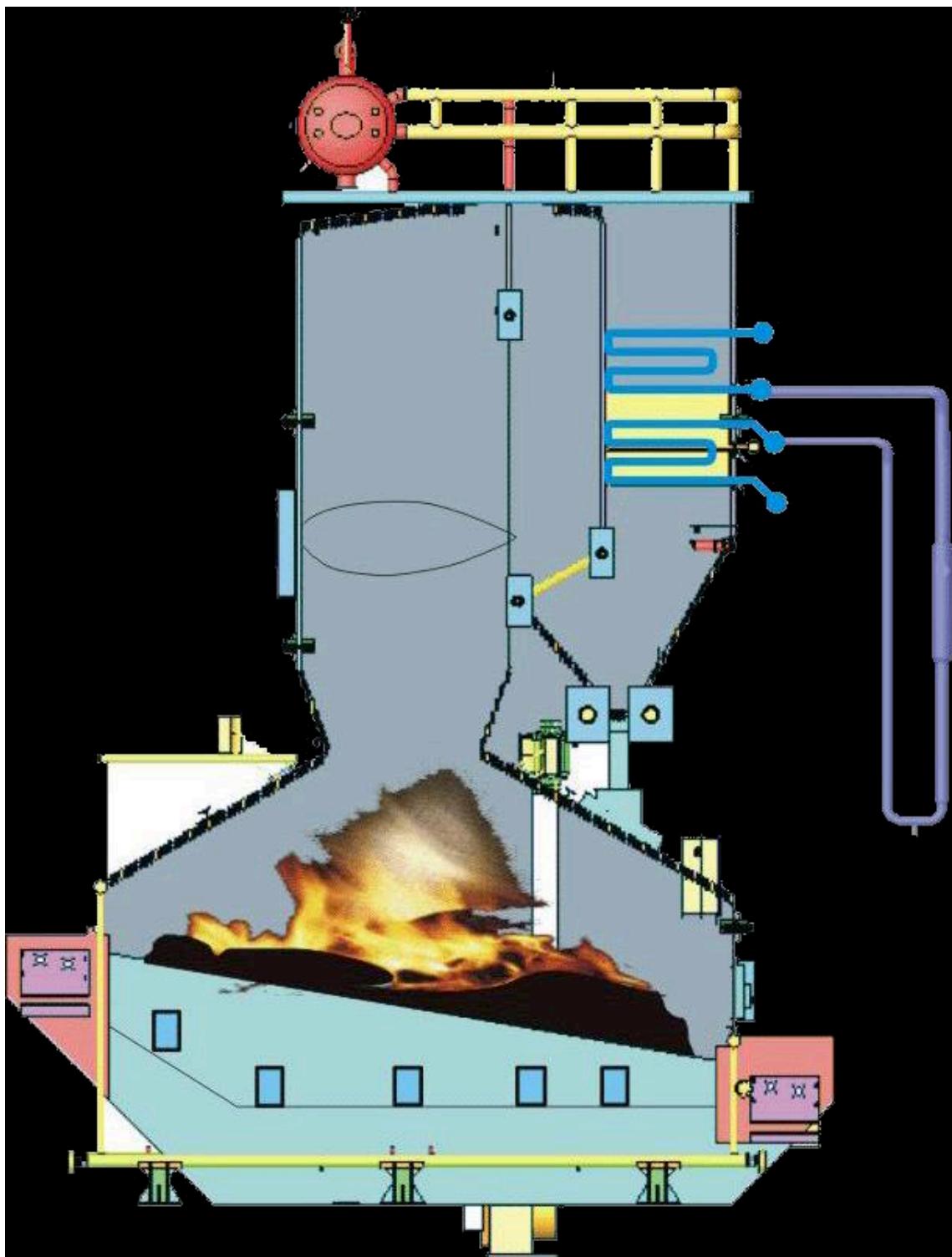
Figura 20 – Controle geral da caldeira.



Fonte: BEGA,2003

A Figura 21 a seguir, mostra o modelo da caldeira em estudo com o aprendizado de máquinas.

Figura 21 – Modelo da caldeira de estudo.



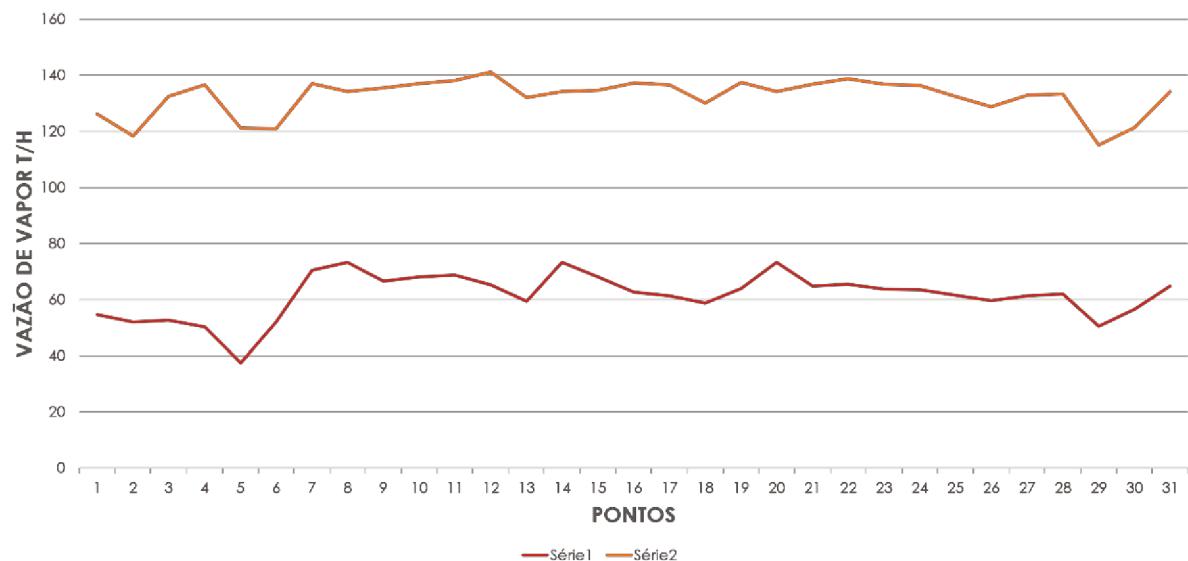
Fonte: O autor (2017).

Dados Técnicos da Caldeira para cavaco com 40% de umidade:

- 1- Tipo: Caldeira Alta Pressão
- 2- Capacidade nominal de vapor: 100.000 kg/h
- 3- Demanda atual de vapor: 96.083 kg/h
- 4- Pressão do vapor: Saída da caldeira: 67 Bar(g)
- 5- Temperatura do vapor: Saída da caldeira: 520 °C
- 6- Entalpia do vapor: 826,5 Kcal/kg
- 7- Rendimento: 85 %
- 8- Perda Descarga Fundo/Sais: 3%
- 9- PCI do combustível: 2.400 kcal/kg
- 10- Carga Térmica Líquida do Vapor: 79.203 Mcal/h
- 11- Consumo de combustível para Vapor Saturado : 27.203 kg/h
- 12- Consumo de combustível para Vapor Superaquecido: 33.837 kg/h
- 13- Consumo de combustível para Cogeração: 6.634 kg/h

5 RESULTADOS E DISCUSSÕES

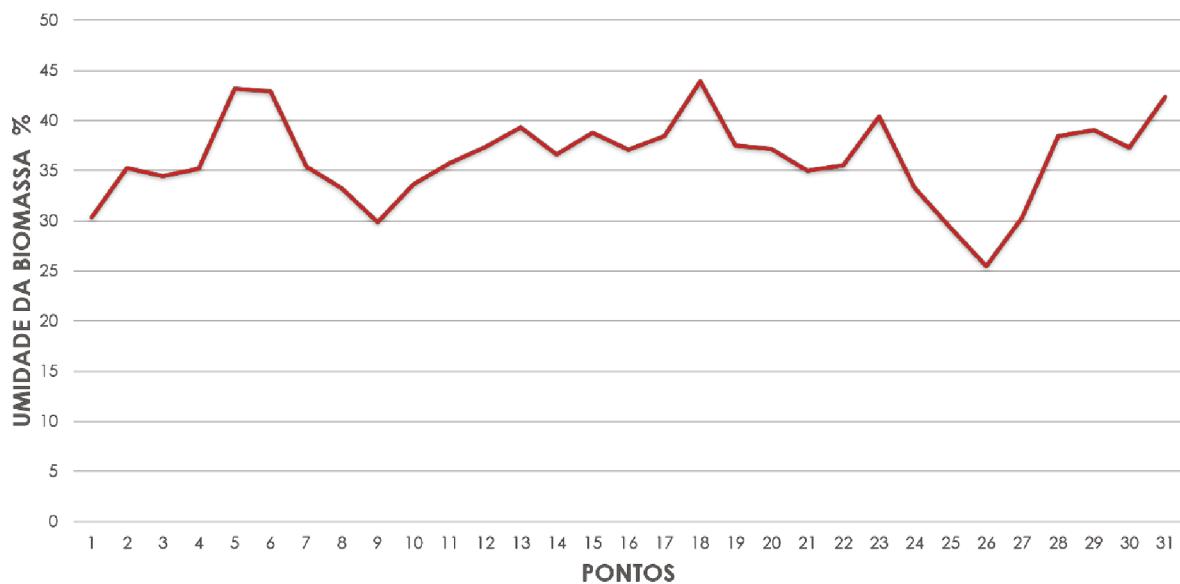
Gráfico 2 – Produção de vapor



Fonte: O autor (2019).

O Gráfico 2 apresenta a produção de vapor. A linha série 1 representa a produção de vapor de uma caldeira individualmente e a linha série 2 representa a soma das duas caldeiras. Cada ponto representa a média do dia. Verifica-se variações entre 115,20 t/h a 141,15 t/h, que é função das demandas do processo e requerimento da planta. Produção horária (t/h) de vapor de alta pressão e temperatura para o processo de cogeração de energia elétrica e uso geral na planta industrial. A caldeira pode modular essas quantidades em função de sua especificação técnica – limite de circulação de água – bem como em função dos controles e automatismo utilizado no equipamento e processo. Na indústria do etanol de milho, plantas autônomas, o consumo médio de t vapor /t de milho é de 1,47- considerando a planta em sua totalidade. Na indústria do etanol de cana de açúcar, este consumo, t vapor/t cana de açúcar moída, fica na ordem de 0,4 .

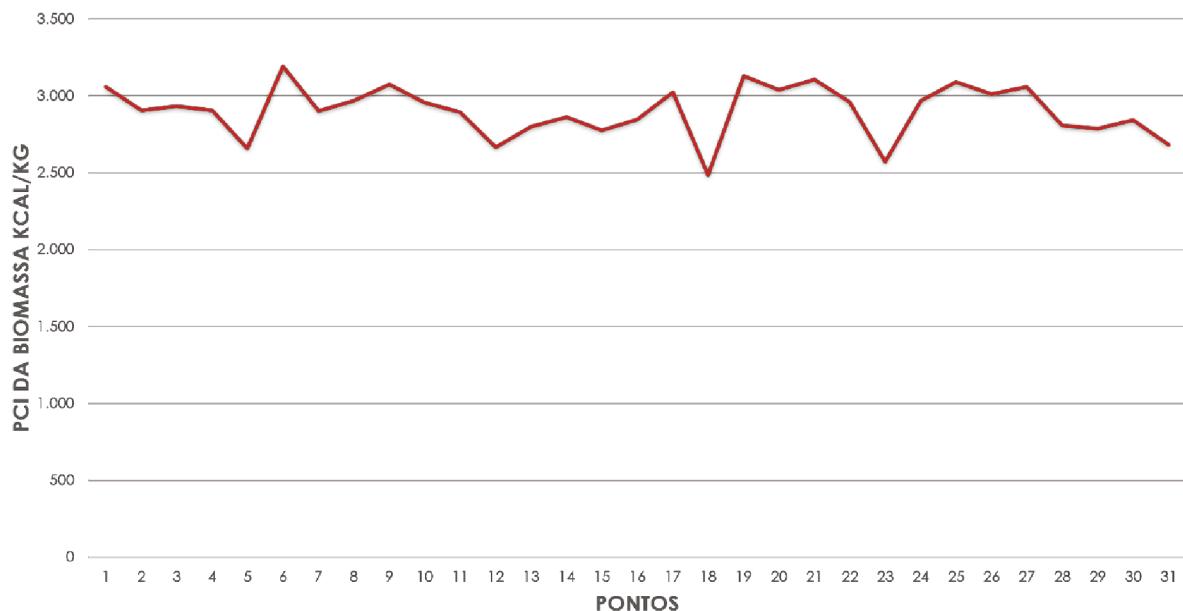
Gráfico 3 – Umidade da biomassa.



Fonte: O autor (2019).

No Gráfico 3, observa-se variações entre um mínimo de 25,52% a 43,96%. Cada ponto é a média diária. A umidade da biomassa é um fator muito importante em todo o processo, pois interfere diretamente na quantidade de energia contida no combustível utilizado, ou seja, o PCI (Poder Calorífico Inferior). O comportamento da caldeira e a quantidade a ser consumida desse combustível, ou seja, a combustão propriamente dita é resultado da quantidade de água (umidade) desse material. O tamanho da partícula – chip – também tem grande influência na combustão dentro da fornalha da caldeira. O preço no mercado desse combustível, também está ligado à umidade deste e será mais valorado se a quantidade de água diminuir. A umidade média da biomassa – bagaço de cana – na indústria do etanol de cana de açúcar, fica ao redor de 50%; enquanto que a biomassa – chip de eucalipto – utilizado na indústria do etanol de milho para plantas autônomas, fica ao redor de 35% a 40% (KLITZKE,2007).

Gráfico 4 – PCI da biomassa.



Fonte: O autor (2019).

No Gráfico 4, as variações do PCI estão entre um mínimo de 2484 kcal/kg e um máximo de 3192 kcal/kg- determinados por bomba calorimétrica IKA (C/2000 Basic) metologia ASTM D 2015-00-(2000).O poder calorífico é definido como a quantidade de calor produzido pela combustão completa de uma unidade de massa de combustível. Nesse caso, o combustível (biomassa de chip de eucalipto) é constituído, sobretudo, de hidrogênio e carbono, tendo o hidrogênio o poder calorífico de 28.700 kcal/kg, enquanto o do carbono é de 8.140 kcal/kg, e, por isso, quanto mais rico em hidrogênio for o combustível maior será o seu poder calorífico. O poder calorífico é distinguido em poder calorífico superior e poder calorífico inferior (BRIANE & DOAT, 1985).

O cavaco é um recurso renovável composto por lascas obtidas de toras de madeira e na sua maioria destina-se à produção de energia em fornos e caldeiras. Apresenta um valor de poder calorífico inferior de 2.500 a 3.500 kcal/kg, massa específica de 300 a 450 kg/m³ e teor de enxofre em valor mínimo (BRIANE & DOAT,1985).

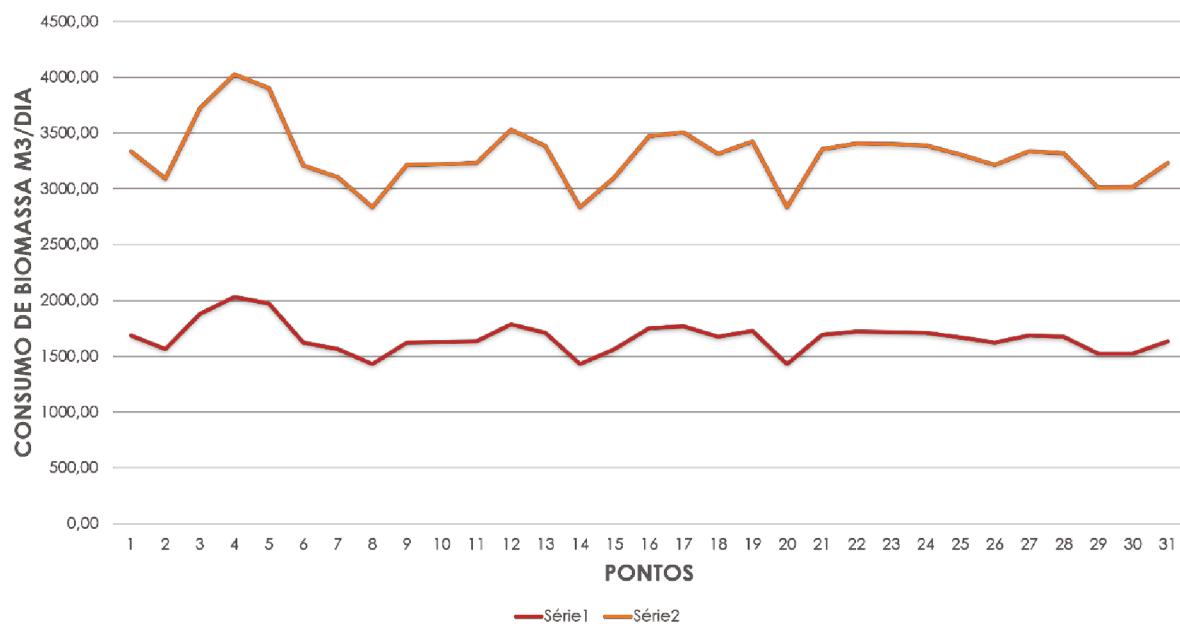
Gráfico 5 – % de cinzas.



Fonte: O autor (2019).

O Gráfico 5 mostra que as variações dos não queimados, basicamente material mineral (argila/terra/sílica e carbono não queimado), ficaram entre 78,11% e 99,11%, ou seja, as cinzas basculhadas para fora do equipamento variaram entre 21,89% a 0,89%. Cada ponto é a média do dia. A quantidade de cinzas é função da qualidade da combustão (ar estequiométrico x real e qualidade do combustível). Existe a influência do tipo de fornalha utilizada – tipo de grelha – e local onde o combustível é queimado. O teor de carbono “não queimado” deverá ser o mínimo possível. Quanto menor a quantidade de cinzas (inclui carbono mal queimado) melhor é o aproveitamento do combustível (CUNHA, *et al*, 1989).

Gráfico 6 – Consumo de biomassa.

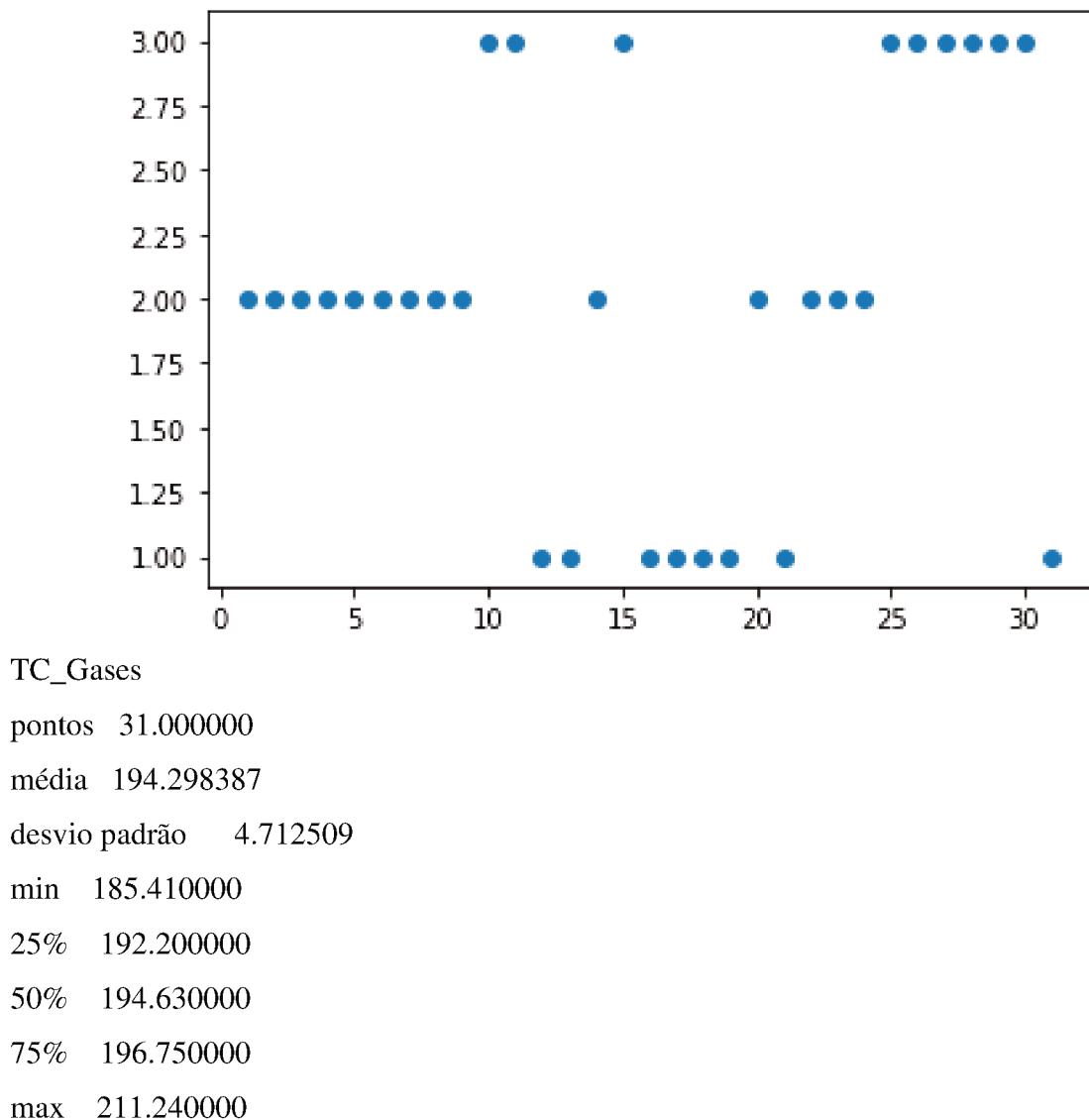


Fonte: O autor (2019).

No Gráfico 6, a série 1 indica o consumo de biomassa na primeira caldeira e a série 2 é a soma das duas caldeiras, o consumo total final fica entre um mínimo de 2834,52 m³/dia e 3532,52 m³/dia. O consumo da biomassa está diretamente relacionado à carga das caldeiras, à umidade do combustível, à qualidade da combustão (ar real x estequiométrico), à composição química do combustível , em última instância o PCI e ao teor de cinzas resultantes na grelha (SOARES, *et al*,2006).

5.1 Análise do desempenho da caldeira segundo o XGboost

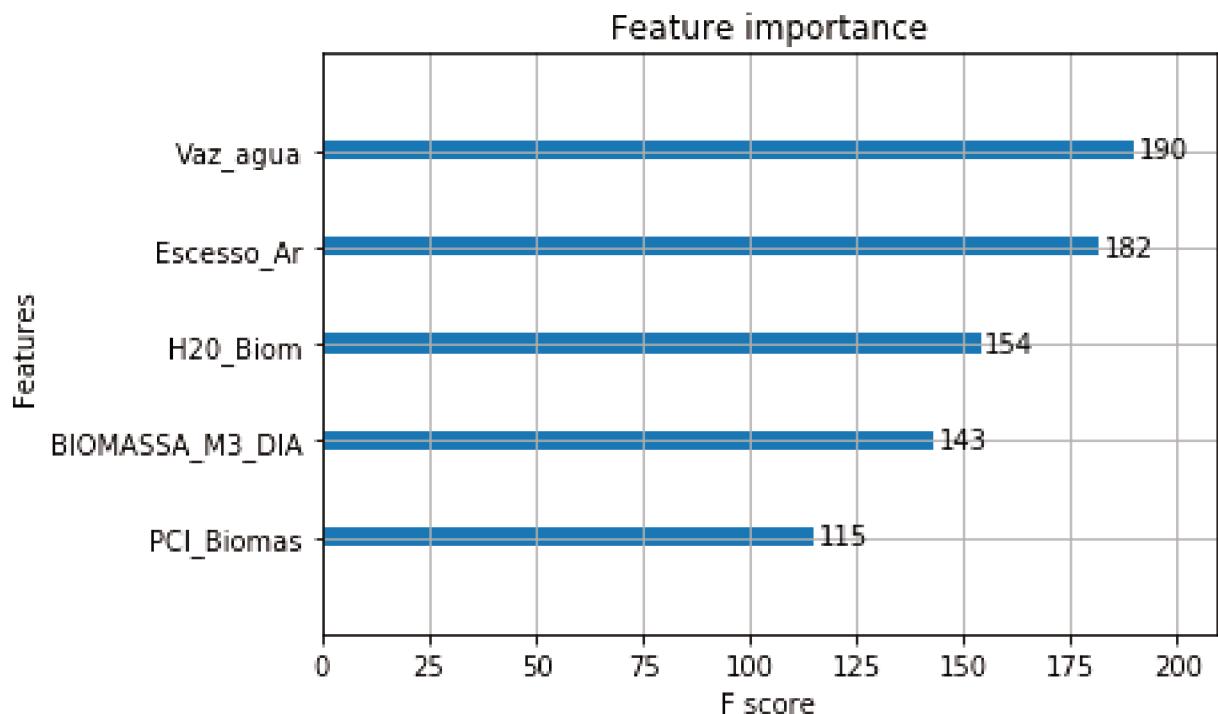
Gráfico 7 – Temperatura dos Gases – Distribuição % dos pontos analisados



Como demonstrado acima ,temos como critério, baixo (1) 25% dos pontos até a temperatura de 192.20 C. Como critério médio (2) temos o intervalo de 192.21 até 196.74 C com 50% dos

pontos analisados. Acima de 196.75 C temos o critério alto (3) com os outros 25% dos pontos. A precisão do modelo foi de 50% para esta variável- conforme apêndice -A.

Gráfico 8 – Temperatura dos Gases- Importância das variáveis

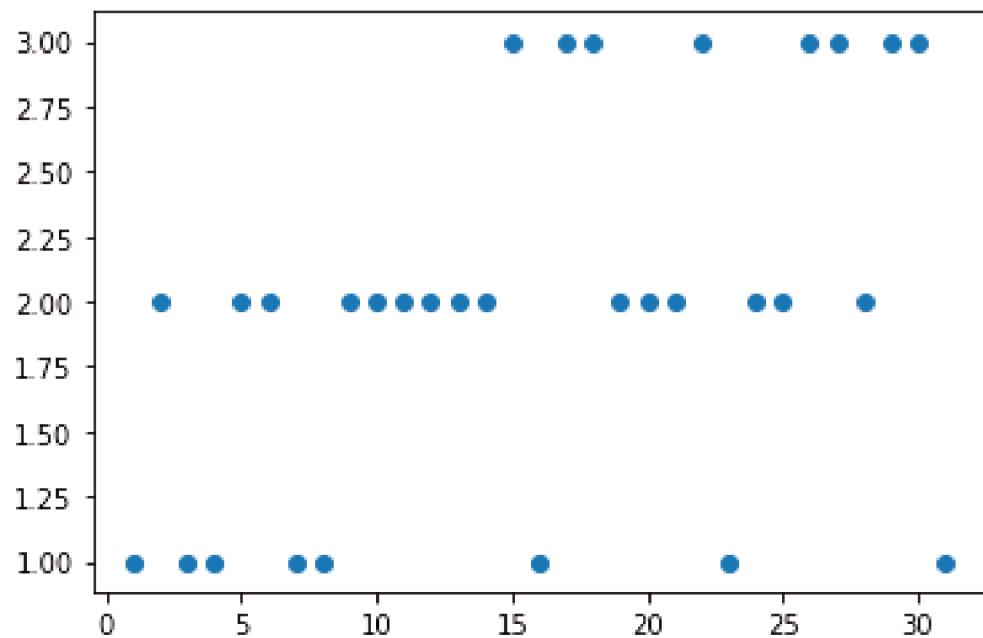


Fonte: O autor.

Como demonstrado no Gráfico 8, temos como influência maior, da importância das variáveis (*Feature Importance*), a vazão de água relacionada à Temperatura dos Gases, conforme pontuação (*F score*), o que é observado e verdadeiro, visto que a carga da caldeira está diretamente relacionada a essa variável (*Features*). Não tanto com menor importância o excesso de ar para a fornalha e queima, fica em segundo lugar; a água contida na biomassa- que vaporiza dentro da fornalha antes dessa biomassa entrar em combustão completa- e o consumo

da biomassa em m³/dia , estão respectivamente em terceiros e quartos lugares. O Poder Calorífico Inferior (PCI) da biomassa, alcançou a última classificação.

Gráfico 9 – Teor Oxigênio (%) – Distribuição % dos pontos analisados



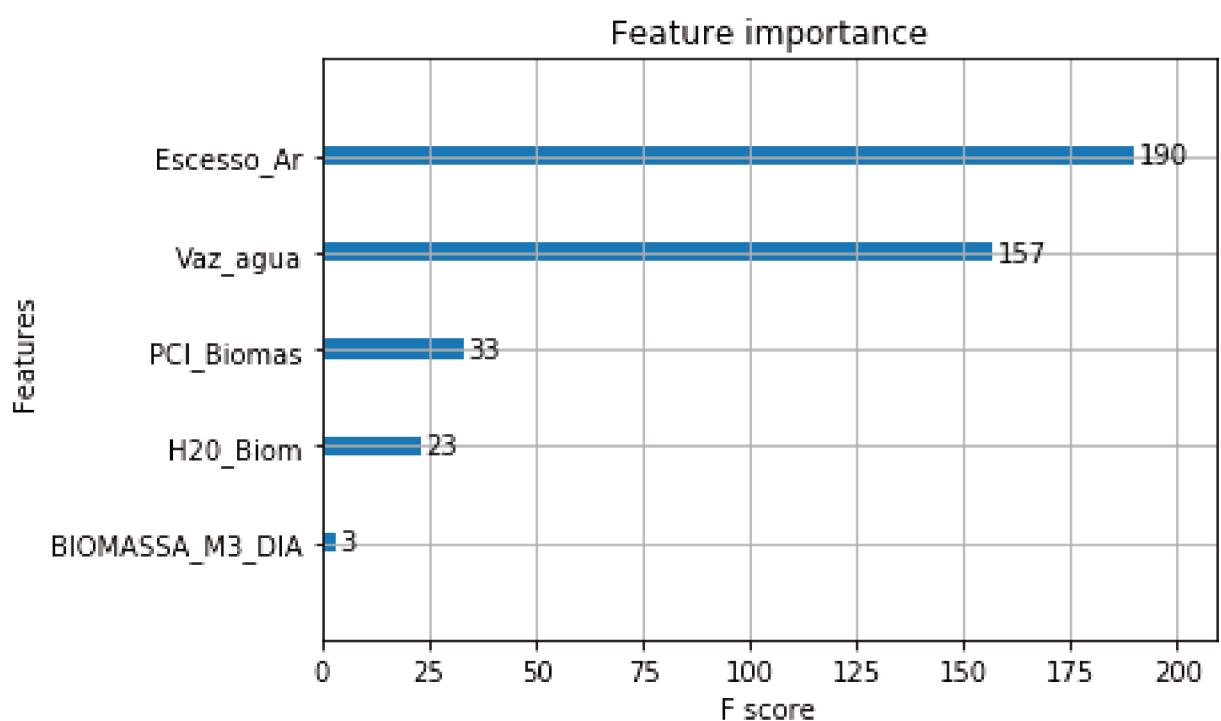
Fonte: O autor

Oxigenio
pontos 31.000000
média 4.960968
desvio padrão 0.360595
min 4.150000
25% 4.785000
50% 5.000000
75% 5.155000

max 6.000000

Como demonstrado acima ,temos como critério, baixo (1) 25% dos pontos até o percentual de oxigênio em 4.77%. Como critério médio (2) temos o intervalo de 4.785 até 5.15% com 50% dos pontos analisados. Acima de 5.16 temos o critério alto (3) com os outros 25% dos pontos. A precisão do modelo foi de 90% para esta variável; conforme apêndice –A.

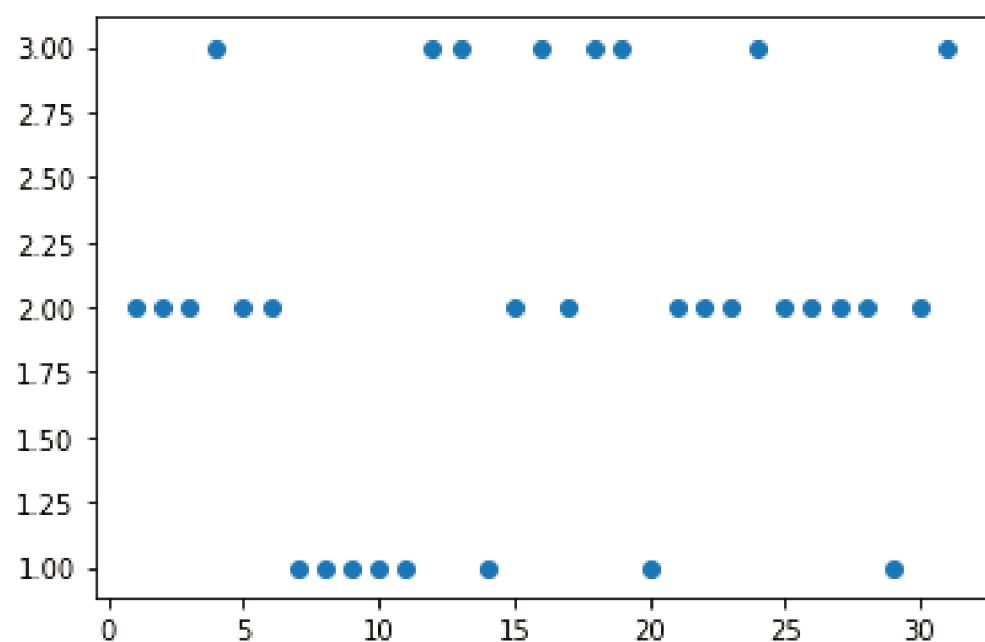
Gráfico 10 – Teor de oxigênio (%) – Importância das variáveis.



Fonte: O autor.

O Gráfico 10 mostra como influência maior o excesso de ar relacionado ao teor de oxigênio, o que é observado nas práticas de campo, visto que a quantidade de ar estequiométrico e real está diretamente relacionada a essa variável. Em segundo lugar na classificação geral , fica a vazão de água para a caldeira , variável que tem influência direta no oxigênio, visto que está relacionado à carga da caldeira. Menos relevantes estão :o PCI da biomassa,a água contida no combustível e o consumo desta biomassa em m³/dia (em última posição)

Gráfico 11 – Temperatura da Fornalha- Distribuição % dos pontos analisados



Fonte: O autor

TC_Fornalha

Pontos 31

Media 738.951935

Desvio padrão 37.789074

min 670.120000

25% 720.185000

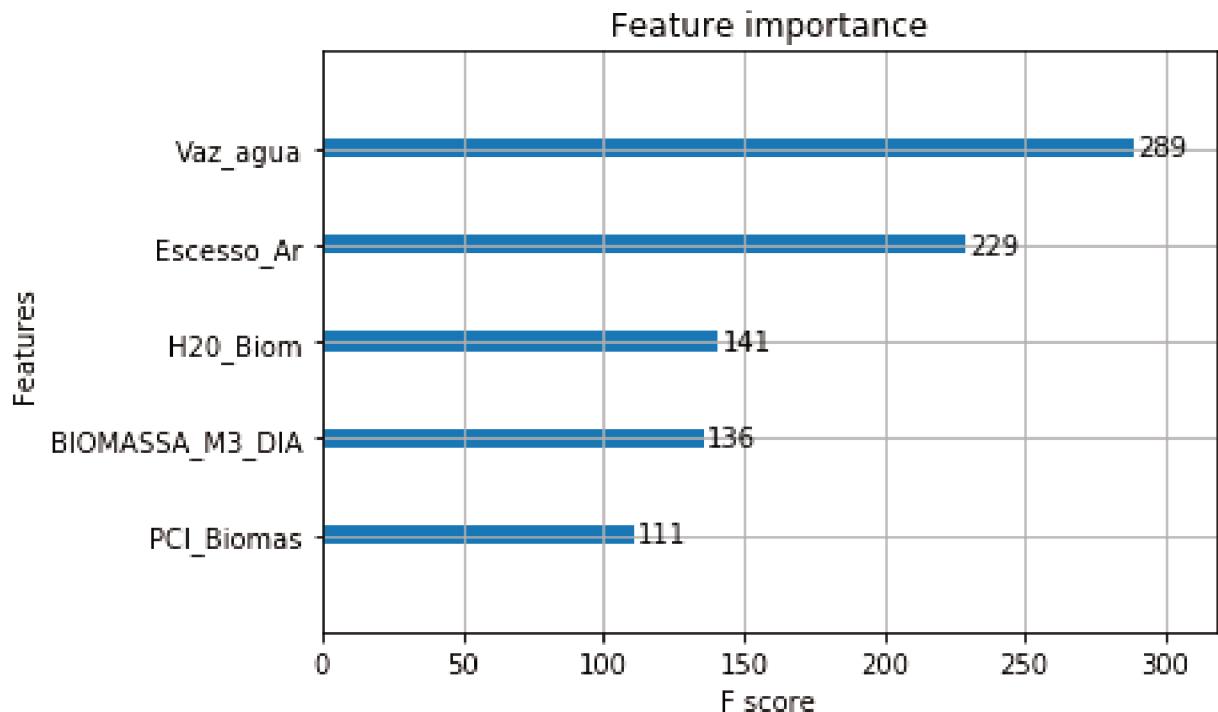
50% 741.290000

75% 768.625000

max 812.730000

Como demonstrado acima ,temos como critério, baixo (1) 25% dos pontos até o percentual de oxigênio em 720.18 C. Como critério médio (2) temos o intervalo de 720.19 até 768.62 C com 50% dos pontos analisados. Acima de 768.625 C temos o critério alto (3) com os outros 25% dos pontos. A precisão do modelo foi de 50% para esta variável;conforme apêndice –A.

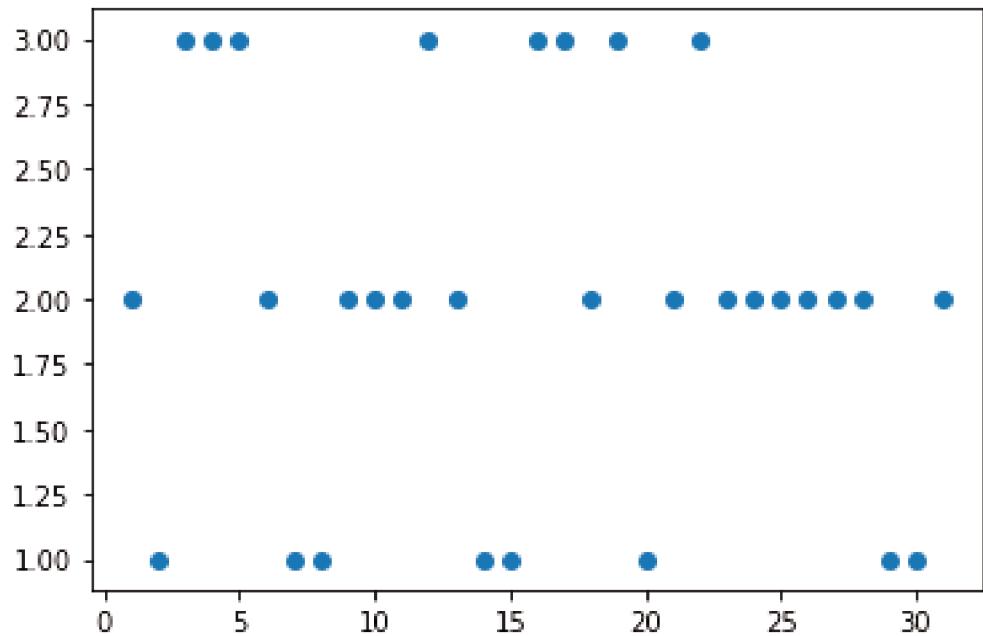
Gráfico 12 – Temperatura da Fornalha – Importância das variáveis.



Fonte: O autor.

Como demonstrado no Gráfico 12, temos como influência maior a vazão de água para o equipamento, o que é observado na prática, visto que a temperatura da fornalha , devido à carga térmica na fornalha, está diretamente relacionada à essa variável. O excesso de ar para a combustão da caldeira fica em segundo lugar indicando uma influência também forte na temperatura da fornalha. Em terceito lugar e muito próximo do quarto lugar temos a água da biomassa e o consumo em m³/dia desta biomassa respectivamente . O PCI da biomassa, em quinto lugar, indica que também tem importância relativa, visto que aqui se representa a qualidade/riqueza do combustível, aumentando a temperatura da fornalha de acordo com sua carga energética.

Gráfico 13 – Vazão de Vapor- Distribuição % dos pontos analisados



Fonte: O autor.

Vaz_Vap

Pontos 31.000000

Média 70.799032

Desvio Padrão 5.839929

min 60.830000

25% 67.745000

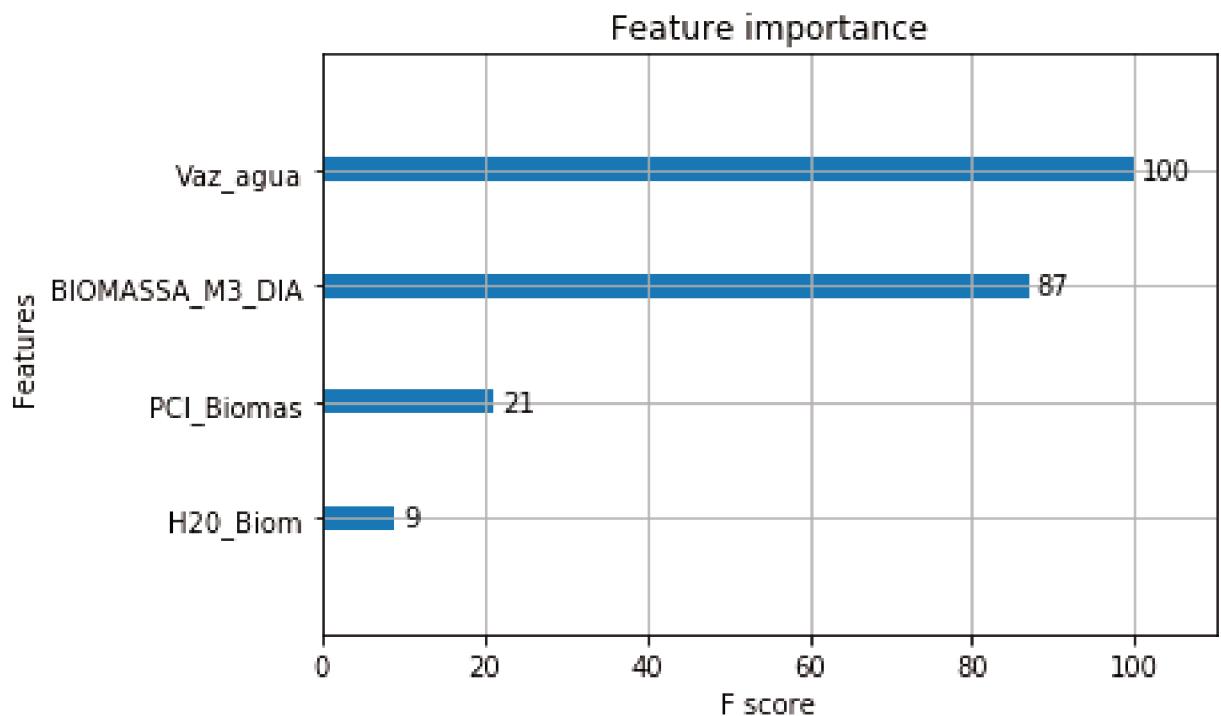
50% 71.160000

75% 73.035000

max 86.330000

Como demonstrado acima ,temos como critério, baixo (1) 25% dos pontos até a vazão de vapor em 67.745 t/h. Como critério médio (2) temos o intervalo de 67.746 até 73.03 com 50% dos pontos analisados. Acima de 73.035 C temos o critério alto (3) com os outros 25% dos pontos. A precisão do modelo foi de 100% para esta variável; conforme apêndice -A.

Gráfico 14 – Vazão de Vapor- Importância das variáveis



Fonte: O autor.

No Gráfico 14 acima, temos como influência maior na vazão de vapor a vazão de água, o que é observado na prática, visto que a carga da caldeira está diretamente relacionada a essa variável. Não tanto com menor importância, o consumo da biomassa em m³/dia ,fica em segundo lugar; já o PCI da biomassa e a água contida nesta biomassa aparecem em terceiros e quartos lugares respectivamente- água esta que vaporiza dentro da fornalha , antes dessa biomassa entrar em combustão completa.

CONCLUSÕES

O Algoritmo escolhido- XGboost dentro do Aprendizado de Máquinas Supervisionado-prediz, com entradas e saídas adequadas e com os fatores bem interpretados, o comportamento dessas máquinas. Portanto, verifica-se que o modelo é satisfatório e coerente, indicando que as variáveis que mais influenciam no resultado estão convergentes com a prática industrial desse equipamento. A vazão de água aparece como a variável mais impactante em três gráficos e em segundo em somente um. Pode-se e deve-se melhorar o resultado e o estudo utilizando mais dados e/ ou trabalhando com outras variáveis de entrada e saída, portanto a Aprendizagem de Máquina, utilizando o XGboost como mecanismo de reenforço é uma ferramenta que pode ser utilizada para predizer de modo satisfatório o comportamento desse importante equipamento industrial no contexto dos biocombustíveis.

REFERÊNCIAS

ADEN, A. **Biochemical Production of Ethanol from Corn Stover:** 2007 State of Technology Model. Colorado: U.s. Department Of Energy, 2008. 15 p.
<https://doi.org/10.2172/928256>

ALTAFINI, C. R. Apostila Sobre Caldeiras. Curso de Engenharia Mecânica, Disciplina de Máquinas Térmicas, Universidade de Caxias do Sul, 2002. Disponível em <https://fluidcontrols.com.br/wp-content/uploads/2019/05/2-Apostilas-sobre-Caldeiras-Prof%C2%BA-Carlos-Alberto-Altafini.pdf>. Acesso em 5 de agosto de 2019.

BCCVL. Biodiversity and Climate Change Virtual Laboratory. **Generalized Boosting Model.** Disponível em: <https://support.bccvl.org.au/support/solutions/articles/6000083212-generalized-boosting-model>. Acesso em 5 de agosto de 2019.

BEGA, E. A., **Instrumentação Aplicada ao Controle de Caldeiras,** 3º edição, Editora Interciênciac, Rio de Janeiro, 2003.

BOTELHO, M. H. C. e BIFANO, H. M. **Operação de Caldeiras – gerenciamento, controle e manutenção.** São Paulo:Editora Blucher, 2016- 2 Edição.

BNDES E CGEE. (Org.). **Bioetanol de cana-de-açúcar: energia para o desenvolvimento sustentável.** Rio de Janeiro: Bndes, 2008. 316 p. DC, 2007. Disponível em: .Acesso em: 06 dez. 2019.

BRASIL. Ministério do Trabalho. **NR 13 - Caldeiras e Vasos de Pressão.** Portaria GM n.º 3.214, de 08 de junho de 1978.

BRIANE.D.;DOAT,J. **Guide technique de la carbonisation: la fabricacion de charbon de bois.** Aix-en Provence, ÉDISUD, 1985\180p.

CANALRURAL.UOL. 2014. Disponível em:
<https://canalrural.uol.com.br/noticias/estudo-mostra-que-subsidios-dos-estados-unidos-sao-mais-prejudiciais-agronegocio-brasileiro-23907/>. Acesso em: em 23 de julho 2019.

ÇENGEL Y. A. ;BOLES M.A. **Termodinâmica 7 edição** – Cap. 14 . Ed. MacGraw-Hill. 2013- Nova York./USA.

CHAVES,A.M.B. **Características energéticas da madeira e do carvão vegetal de clones de Eucalyptus spp.** Encyclopédia Bioesfera. Centro Científico Conhecer- Goiânia v .9 n 17 p 533-542,2013.

CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: **Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining**. ACM, 2016. Disponível em: <https://arxiv.org/pdf/1603.02754.pdf>. Acesso em 5 de agosto de 2019.
<https://doi.org/10.1145/2939672.2939785>

CONAB. **Companhia Nacional de Abastecimento**. Disponível em: www.conab.gov.br. Acesso em: 29 de novembro de 2019.

CORTEZ,Luis Augusto Barbosa;LORA, Electo Silva; GÓMEZ, Edgardo Olivares, **Biomassa para Energia. Capítulo 1- Biomassa no Brasil e no mundo**, Campinas,SP; UNICAMP,p 11-29,2008.

CREUTZIG, Felix *et al.* **Bioenergy and climate change mitigation: an assessment**. Gcb Bioenergy, v. 7, n. 5, p. 916-944, 2015.
<https://doi.org/10.1111/gcbb.12205>

CUNHA,M.P.S.C.;PONTES, C.L.F.; CRUZ ,I.A;CABRAL,M.T.F.D; CUNHA NETO,Z.B.; BARBOSA,A.P.R. **Estudo Químico de 55 espécies lenhosas para geração de energia em caldeiras.In 3 encontro Brasileiro em madeiras e em estruturas de madeira**. Anais V2. P 93-121, São Carlos 1989.

DEBONI, T. L. *et al.* **Evolution of the quality of forest biomass for energy generation in a cogeneration plant**. Renewable energy, v. 135, p. 1291-1302, 2019.
<https://doi.org/10.1016/j.renene.2018.09.039>

DONKE A, Nogueira A, Matai P, Kulay L. **Environmental and Energy Performance of Ethanol Production from the Integration of Sugarcane, Corn, and Grain Sorghum in a Multipurpose Plant**. Resources 2017, 6, 1.
<https://doi.org/10.3390/resources6010001>

DUTRA, Aldo Cordeiro et AL. **Manual Técnico de Caldeiras e Vasos de Pressão**. Brasília: MTE, SIT, DSST, 2006.124 p.

EMBRAPA- [https://www.embrapa.br/milho e sorgo /-](https://www.embrapa.br/milho-e-sorgo/) Acesso em 17 de fevereiro 2020.

FANTAZZINE, M. L. & SERPA, R. R. **Aspectos gerais de segurança e Elementos de Gerenciamento de Riscos**. Rio de Janeiro – ITSEMAP do Brasil, Serviços Tecnológicos MAPFRE, 2002.

FERREIRA, M. C. *et al.* **Quimiometria I: calibração multivariada, um tutorial**. Química Nova, v. 22, n. 5, p. 724-731, 1999.
<https://doi.org/10.1590/S0100-40421999000500016>

FSBIOENERGIA. Disponível em: <http://www.fsbioenergia.com.br/pt-br/>. Acesso em 14 de novembro de 2019.

GARAMI, A. *et al.* **Reaction zone monitoring in biomass combustion.** Control Engineering Practice, v. 74, p. 95-106, 2018.

<https://doi.org/10.1016/j.conengprac.2018.02.009>

GREGO,M,B,L, **Eficiencia Energética do Etanol e sua produção: Análise comparativa entre o Brasil e EUA.**, ESALQ/USP Ano 2011, acesso internet em 03 de abril de 2018.

HOFFMAN, Linwood A., and Allen Baker. **Estimating the Substitution of Distillers' Grains for Corn and Soybean Meal in the U.S. Feed Complex.** Department of Agriculture, Economic Research Service, October 2011.

HOTTA, C. **Biocombustíveis: cana-de-açúcar vs. milho.** Laboratório de Fisiologia Molecular de Plantas. 2008. Disponível em:
<http://www.carloshotta.com.br/brontossauros/2008/8/1/biocombustiveis-cana-de-acucar-vs-milho.html>. Acesso em 17 ago 2019.

HU, J. *et al.* **Co-gasification of coal and biomass: synergy, characterization and reactivity of the residual char.** Bioresource technology, v. 244, p. 1-7, 2017.

<https://doi.org/10.1016/j.biortech.2017.07.111>

IEA- International Energy Agency acesso em <https://www.iea.org/topics/world-energy-outlook>, em 02 de abril de 2020.

IMEA, 2017. **Clusters de Etanol de Milho.** Instituto Mato-grossense de Economia Agropecuária (IMEA) – Cuiabá/MT: 2017

INGLEDEW, W. M. *et al.* **The alcohol industry: How has it changed and matured. The Alcohol Textbook–A Reference for the Beverage, Fuel and Industrial Alcohol Industries,** v. 5, p. 1-6, 2009.

JONES, A. M.; INGLEDEW, W. M. **Fuel alcohol production: optimization of temperature for efficient very-high-gravity fermentation.** Appl. Environ. Microbiol., v. 60, n. 3, p. 1048-1051, 1994.

<https://doi.org/10.1128/AEM.60.3.1048-1051.1994>

KLITZKE, Ricardo Jorge. **Curso de secagem da madeira** – Módulo 01 p 1 – 119. 2007.

LEE, K. M.; ZANIL, M. F. **Interval type-2 fuzzy logic with Karmen-Mendel algorithm for sequential ionic liquid dissolution–solid acid saccharification.** Journal of Chemical Technology & Biotechnology, v. 94, n. 4, p. 1073-1081, 2019.

<https://doi.org/10.1002/jctb.5854>

MARTINELLI, S. H. S.; NEITZEL, I.; VIEIRA, O. **Detecção de defeitos via redes neurais: aplicação-sistema de ar secundário de uma caldeia de recuperação química.** Revista Ciência e Tecnologia, v. 11, n. 18, 2010.

MOREIRA M. M. R; ARANTES S. M. **Análise socioeconômica e ambiental da produção de etanol de milho no Centro Oeste Brasileiro.** Policy Brief. São Paulo, 2018.

NASCIMENTO, A. P. **Desenvolvimento e validação de metodologia para medicamentos contendo dipirona sódica e cloridrato de papaverina isolados e em associação.** 2005. 141 f. Tese (Doutorado) - Curso de Faculdade de Ciências Farmacêuticas, Universidade de São Paulo, São Paulo, 2005. Disponível em: https://teses.usp.br/teses/disponiveis/9/9139/tde-01082012-162353/publico/TESEAndreia_Peraro_Nascimento.pdf. Acesso em: 10 nov. 2019.

PIMENTEL, D.; PATZEK, T. W. **Ethanol production using corn, switchgrass, and wood; biodiesel production using soybean and sunflower.** Natural resources research, v. 14, n. 1, p. 65-76, 2005.

<https://doi.org/10.1007/s11053-005-4679-8>

PROCKNOR C, acesso em 02 de abril de 2020 no endereço

<https://www.procknor.com.br/br/artigos/utilidades>

RAUCCI GS, Moreira CS, Alves PA, Mello FFC, Frazão LA, Cerri CEP e Cerri CC. **Greenhouse gas assessment of Brazilian soybean production: a case study of Mato Grosso State.** Journal of Cleaner Production 2015, 96: 418-425
<https://doi.org/10.1016/j.jclepro.2014.02.064>

RIBEIRO, F.A.L. *et al.* **Quimiometria: inovação no desenvolvimento e validação de métodos analíticos para controle de qualidade na indústria química.** Informativo CRQ 4ª Região, n. 85, p. 14-15, 2007.

RIDGEWAY, G. The state of boosting. **Computing Science and Statistics**, 1999. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.22.276&rep=rep1&type=pdf>. Acesso em 5 de agosto de 2019.

RUSSELL,STUART J, NORVIG ,P. **Inteligencia Artificial**, Segunda Edição, Elsevier, Rio de Janeiro,2004.

SCIENCEBLOGS.com.br/brontossauros/2008/08/biocombustiveis-cana-de-acucar-vs-milho.php –acesso em 13 de outubro de 2019.

SF AGRO. **SF Agro - Farming Brasil.** Artigo original do site SF Agro - Farming Brasil. <http://sfagro.uol.com.br/etanol-de-segunda-geracao-deve-custar-80-do-preco-do-combustivel-de-cana/>. Acesso em: agosto de 2017.

SHI, Y. et al. **Combustion optimization of ultra supercritical boiler based on artificial intelligence.** Energy, v. 170, p. 804-817, 2019.
<https://doi.org/10.1016/j.energy.2018.12.172>

SOARES, Vássia Carvalho et at. **Correlações entre as propriedades da madeira e do carvão vegetal de híbridos de eucalipto.** Revista Árvore. Viçosa, v 38 n.3, p 543-549, 2014.
<https://doi.org/10.1590/S0100-67622014000300017>

TAN, P. et al. **Dynamic modeling of NOx emission in 660 MW coal-fired boiler with long short-term memory.** Energy, v. 176, p. 429-439, 2019.
<https://doi.org/10.1016/j.energy.2019.04.020>

TURNS, S. R., **An Introduction to combustion: Concepts and Applications**, McGraw-Hill, 2000.

UNEM . **União Nacional do Etanol de Milho.** Disponível em:
<http://www.etanoldemilho.com.br/>. Acesso em 2 fevereiro de 2020.

UNICA. **União da Indústria de Cana-de-Açúcar.** 2018. Disponível:
<https://www.unica.com.br/>. Acesso em abril de 2019.

VALDERRAMA, L. **Análise discriminante em Química Forense: Aplicações em documentoscopia.** 2015. 43 f. TCC (Graduação) - Curso de Química, Universidade Técnica Federal do Paraná, Campo Mourão, 2015. Disponível em:
<http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/5259/3/CM_COLIQ_2015_1_01.pdf>. Acesso em: 10 nov. 2019.

WANG, X. et al. **Numerical Investigation of Solid-Fueled Chemical Looping Combustion Process Utilizing Char for Carbon Capture.** Processes, v. 7, n. 9, p. 603, 2019.
<https://doi.org/10.3390/pr7090603>

WORDPRESS,<https://intercambiobadasoja.wordpress.com/2011/08/25/os-cinturoes-agricolas-dos-estados-unidos/>. Acesso em 22 de agosto de 2019.

APÊNDICE A

6/3/2020 Boiler_Part_09_open_Fer_June_03_2020.ipynb - Colaboratory

```
# carrega as bibliotecas do python
```

```
%matplotlib inline
from matplotlib import pyplot as plt
import os
from pandas import read_excel
import numpy as np
from sklearn.decomposition import PCA
import statistics as st
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.model_selection import KFold
from sklearn.model_selection import cross_val_score
```

```
# carrega o arquivo de nome Fernando_Caldeira_1_maio_10_2019.xlsx
```

```
from google.colab import files
def getLocalFiles():
    _files = files.upload()
    if len(_files) >0:
        for k,v in _files.items():
            open(k,'wb').write(v)
getLocalFiles()
```

→ Choose Files Fernando_C..._2019.xlsx

- **Fernando_Caldeira_1_maio_10_2019.xlsx**(application/vnd.openxmlformats-officedocument.spreadsheetml.sheet) - 21 done

Saving Fernando Caldeira 1 maio 10 2019.xlsx to Fernando Caldeira 1 maio 10 2019 (7).xlsx

```
# leitura dos dados
```

```
# find your sheet name at the bottom left of your excel file and assign
# it to sheet_name
my_sheet = 'Planilha1'
file_name = 'Fernando_Caldeira_1_maio_10_2019.xlsx' # name of your excel file
df = read_excel(file_name, sheet_name = my_sheet)
print(df.head(3)) # shows headers with top 5 rows
```

→

	Data	TC	Patm\n	Vaz_Vap	...	Cinzas	PCI_Cinzas	Temp_Desc	TC_Gases
0	2017-12-01	505.19	67.88	71.66	...	94.04	791.0	110.0	194.97
1	2017-12-02	503.29	68.03	66.34	...	92.92	747.0	110.0	194.00
2	2017-12-03	512.39	68.00	79.83	...	92.26	793.0	110.0	194.00

[3 rows x 19 columns]

Une as colunas que fara um estudo

X=df.drop(['Data'], axis=1)

```
print(X.head(3)) # shows headers with top 5 rows
```

	Patm\									
→	TC	n	Vaz_Vap	Vaz_agua	...	Cinzas	PCI_Cinzas	Temp_Desc	TC_Gases	
0	505.19	67.88	71.66	75.45	...	94.04	791.0	110.0	194.97	
1	503.29	68.03	66.34	69.80	...	92.92	747.0	110.0	194.00	
2	512.39	68.00	79.83	83.14	...	92.26	793.0	110.0	194.00	

[3 rows x 18 columns]

Busca de dados sujos

```
print(X.isnull().any())
```

→TC	True
Patm\n	True
Vaz_Vap	True
Vaz_agua	True
Agua_Vapor	True
Perc_Descarga	True
Pressao_01\n	True
Pressao_02\n	True
Perc_Abertura	True
TC_Fornalha	True
Oxigenio	True
H20_Biom	True
PCI_Biomas	True
Escesso_Ar	True
Cinzas	True
PCI_Cinzas	True
Temp_Desc	True
TC_Gases	True
dtype: bool	

Limpeza dos dados

```
X=X.dropna()
#print(X.isnull().any())
```

...

ENTRADA:

H20_Biom	True
PCI_Biomas	True
Escesso_Ar	True
Vaz_agua	True
Temp_Desc	True
BIOMASSA_M3_DIA	

SAIDA:

```
TC           True
Patm\n      True
Vaz_Vap     True
Agua_Vapor  True
Perc_Descarga True
Pressao_01\n True
Pressao_02\n True
Perc_Abertura True
TC_Fornalha  True
PCI_Cinzas   True
TC_Gases    True
Cinzas       True
Oxigenio    True
'''
```

```
↳ \nENTRADA\nH20_Biom  True\nPCI_Biomas  True\nEscesso_Ar  True\nVaz_agua
```

```
Vapor=[]
Vapor=np.copy(X['Vaz_Vap'])
Vapor=pd.DataFrame({'Vaz_Vap':Vapor})
#print(Vapor)

#print(X.head())
```

```
#sortedPoints = sorted(Vapor)
#mid = len(sortedPoints)
#print(mid)
```

```
print(Vapor.describe())
```

```
↳      Vaz_Vap
count  31.000000
mean   70.799032
std    5.839929
min    60.830000
25%   67.745000
50%   71.160000
75%   73.035000
max    86.33000
```

```

Num=len(Vapor)
Vapor_x=np.copy(Vapor)
for i in range(Num):
    if(Vapor_x[i]<67.75):
        Vapor_x[i]=1 # Vapor=1 é baixo
    else:
        if(Vapor_x[i]>73.03):
            Vapor_x[i]=3 # Vapor=3 é alto
        else:
            Vapor_x[i]=2 # Vapor=1 é médio
#Vapor_x=np.array(Vapor_x)
#print(Vapor_x)
#print(X['Vaz_Vap'])

#Nomes=['H20_Biom','PCI_Biomas','Escesso_Ar','Vaz_agua','BIOMASSA_M3_DIA']
Nomes=['H20_Biom','PCI_Biomas','Escesso_Ar','Vaz_agua'] # ???BIOMASSA_M3_DIA'
# 'Temp_Desc' NAO ESTA VARIANDO
x=X[Nomes]
print(x)

```

y = Vapor_x

	H20_Biom	PCI_Biomas	Escesso_Ar	Vaz_agua
0	30.350	3060.0	24.63	75.45
1	35.320	2905.0	30.03	69.80
2	34.460	2932.0	27.04	83.14
3	35.260	2907.0	28.76	89.18
4	43.200	2658.0	31.25	87.13
5	42.945	3192.0	30.19	71.57
6	35.400	2902.0	25.90	68.98
7	33.270	2970.0	26.74	63.12
8	29.880	3075.0	31.25	71.09
9	33.660	2957.0	32.08	71.08
10	35.740	2892.0	29.63	71.36
11	37.420	2667.0	31.25	77.83
12	39.380	2800.0	32.92	74.96
13	36.620	2864.0	30.03	63.08
14	38.800	2777.0	33.16	68.52
15	37.140	2848.0	29.39	76.33
16	38.460	3022.0	32.58	77.10
17	43.960	2484.0	32.91	73.43
18	37.560	3129.0	32.80	75.38
19	37.200	3040.0	32.33	63.11
20	35.020	3107.0	31.50	73.77
21	35.540	2962.0	33.76	74.87
22	40.400	2571.0	28.83	74.66
23	33.320	2967.0	32.08	74.55
24	29.340	3092.0	30.43	72.79
25	25.520	3010.0	32.91	70.80
26	30.340	3060.0	33.76	73.25
27	38.480	2806.0	32.49	72.92
28	39.100	2787.0	40.00	68.12
29	37.340	2842.0	33.76	68.18

```

# plot feature importance using built-in function
from numpy import loadtxt
from xgboost import XGBClassifier
from xgboost import plot_importance
from matplotlib import pyplot

# split data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(x,y, test_size=0.33, random_state=7)
#fit model no training data
model = XGBClassifier()
model.fit(X_train, y_train)
# make predictions for test data
y_pred = model.predict(X_test)
predictions = [round(value) for value in y_pred]
#evaluate predictions
accuracy = accuracy_score(y_test, predictions)
print("Accuracy: %.2f%%" % (accuracy * 100.0))
#fit model no training data
model = XGBClassifier()
model.fit(x, y)
kfold = KFold(n_splits=10, random_state=7)
results = cross_val_score(model, x, y, cv=kfold)
print("Accuracy cross: %.2f%% (%.2f%%)" % (results.mean()*100, results.std()*100))

```

↳ Accuracy: 100.00%

Accuracy cross: 87.50% (15.48%)

/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversion y = column_or_1d(y, warn=True)

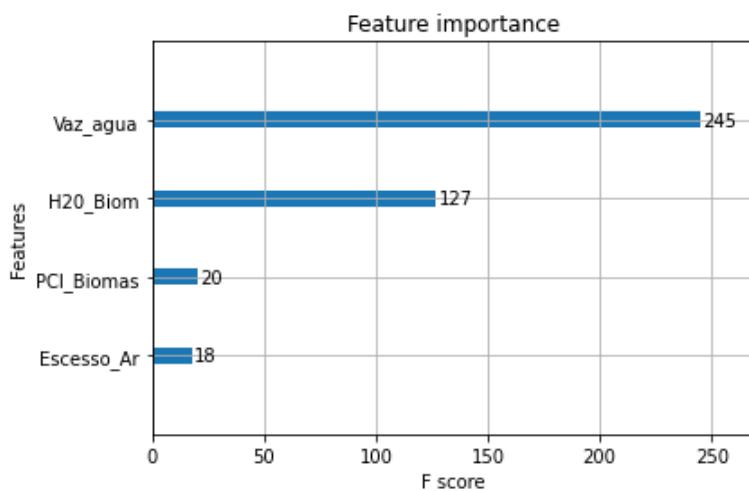
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversion y = column_or_1d(y, warn=True)

/usr/local/lib/python3.6/dist-packages/sklearn/model_selection/_split.py:296: FutureWarning

```

#plot feature importance
plot_importance(model)
pyplot.show()

```

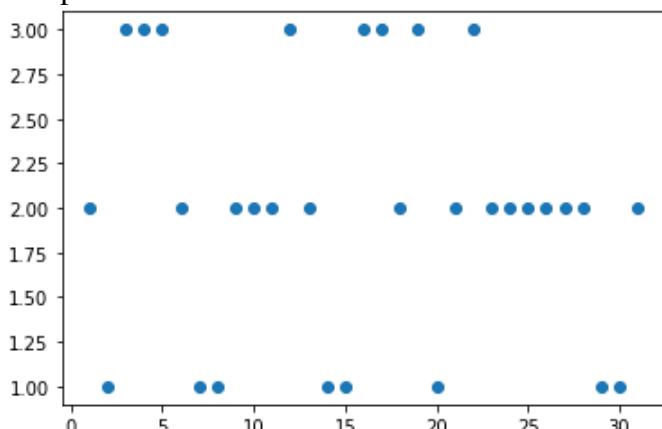


```
# split data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(x,y, test_size=0.33, random_state=7)
# fit model no training data
model = XGBClassifier()
model.fit(X_train, y_train)
# make predictions for test data
y_pred = model.predict(X_test)
predictions = [round(value) for value in y_pred]
# evaluate predictions
accuracy = accuracy_score(y_test, predictions)
print("Accuracy: %.2f%%" % (accuracy * 100.0))
# fit model no training data
```

↪ Accuracy: 100.00%
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversion y = column_or_1d(y, warn=True)
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversion y = column_or_1d(y, warn=True)

```
xw=range(1,(len(y)+1))  
plt.scatter(xw,y)
```

→ <matplotlib.collections.PathCollection at 0x7effc0db1ef0>



```
TC_Fornalha=[]
TC_Fornalha=np.copy(X['TC_Fornalha'])
TC_Fornalha=pd.DataFrame({'TC_Fornalha':TC_Fornalha})
#print(TC_Fornalha)
```

```
print(TC_Fornalha.describe())
```

```
↳ TC_Fornalha
count    31.000000
mean    738.951935
std      37.789074
min     670.120000
25%    720.185000
50%    741.290000
75%    768.625000
max     812.730000
```

```
Num=len(TC_Fornalha)
TC_Fornalha_x=np.copy(TC_Fornalha)
for i in range(Num):
    if(TC_Fornalha_x[i]<720.18):
        TC_Fornalha_x[i]=1 # TC_Fornalha=1 é baixo
    else:
        if(TC_Fornalha_x[i]>768.62):
            TC_Fornalha_x[i]=3 # TC_Fornalha=3 é alto
        else:
            TC_Fornalha_x[i]=2 # TC_Fornalha=1 é médio
```

```
print(TC_Fornalha)
print(TC_Fornalha_x)
```

```
↳
```

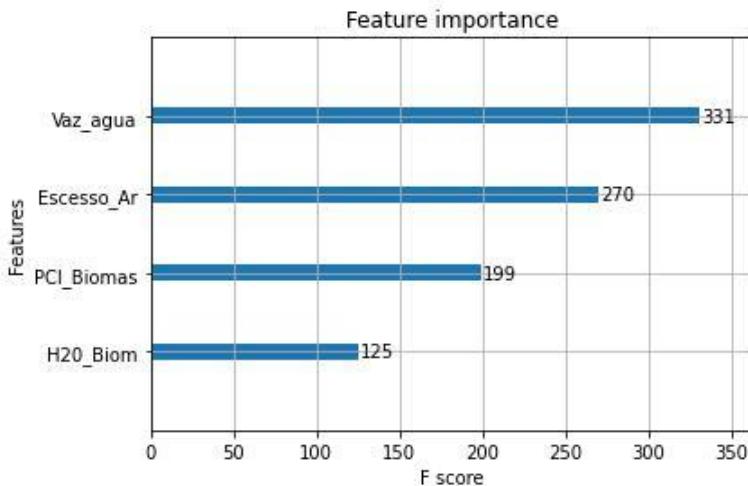
TC_Fornalha

0	730.92
1	730.86
2	730.99
3	783.20
4	730.74
5	741.29
6	696.92
7	670.12
8	694.93
9	686.89
10	694.24
11	771.72
12	812.73
13	670.12
14	756.73
15	791.56
16	765.53
17	779.48
18	777.62
19	670.12
20	736.68
21	739.04
22	741.91
23	779.80
24	761.22
25	746.79
26	755.39
27	745.96
28	713.98
29	726.39
30	773.64

[[2.]
[2.]
[2.]
[3.]
[2.]
[2.]
[1.]
[1.]
[1.]
[1.]
[3.]
[3.]
[1.]
[2.]
[3.]
[2.]
[3.]
[3.]

```
[1.]  
[2.]  
[2.]  
[2.]  
y = TC_Fornalha_x  
[3.]  
[2.]  
[2.]  
# fit model no training data  
[2.]  
model = XGBClassifier()  
[2.]  
model.fit(x, y)  
[1.]  
kfold = KFold(n_splits=10, random_state=7)  
results = cross_val_score(model, x, y, cv=kfold)  
print("Accuracy cross: %.2f%% (%.2f%%)" % (results.mean()*100, results.std()*100))  
#     plot feature importance  
#     plot_importance(model)  
#     pyplot.show()
```

↳ /usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversion y = column_or_1d(y, warn=True)
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversion y = column_or_1d(y, warn=True)
/usr/local/lib/python3.6/dist-packages/sklearn/model_selection/_split.py:296: FutureWarning
FutureWarning
Accuracy cross: 35.83% (10.57%)



```

# split data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(x,y, test_size=0.33, random_state=7)
# fit model no training data
model = XGBClassifier()
model.fit(X_train, y_train)
# make predictions for test data
y_pred = model.predict(X_test)
predictions = [round(value) for value in y_pred]
# evaluate predictions
accuracy = accuracy_score(y_test, predictions)
print("Accuracy: %.2f%%" % (accuracy * 100.0))
# fit model no training data

kfold = KFold(n_splits=10, random_state=7)
results = cross_val_score(model, x, y, cv=kfold)
print("Accuracy cross: %.2f%% (%.2f%%)" % (results.mean()*100, results.std()*100))

```

↳

```

Accuracy: 45.45%
Accuracy cross: 35.83% (10.57%)
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversionWarning
  y = column_or_1d(y, warn=True)

```

```

# plt.scatter(X['Vaz_agua'],y)
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversionWarning
  y = column_or_1d(y, warn=True)

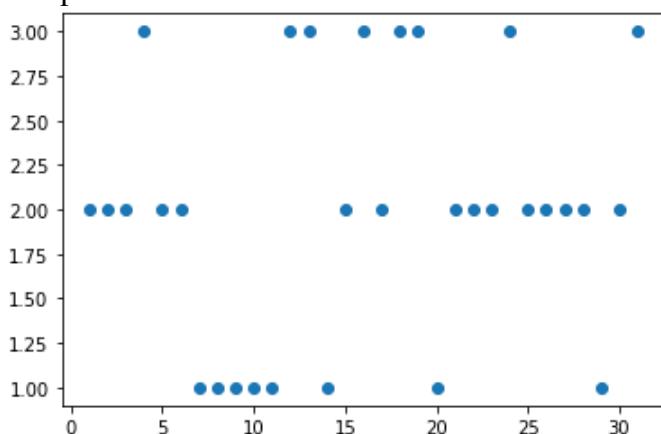
```

```

/usr/local/lib/python3.6/dist-packages/sklearn/model_selection/_split.py:296: FutureWarning
xw=range(1,(len(y)+1))
FutureWarning
plt.scatter(xw,y)

```

↳ <matplotlib.collections.PathCollection at 0x7effbe4a87b8>



```
#TC_Gases=[]
#TC_Gases=np.copy(X['TC_Gases'])
#TC_Gases=pd.DataFrame({'TC_Gases':TC_Gases})
#TC_Gases=pd.DataFrame({'TC_Gases':X['TC_Gases']})
#print(TC_Gases)

print(TC_Gases.describe())

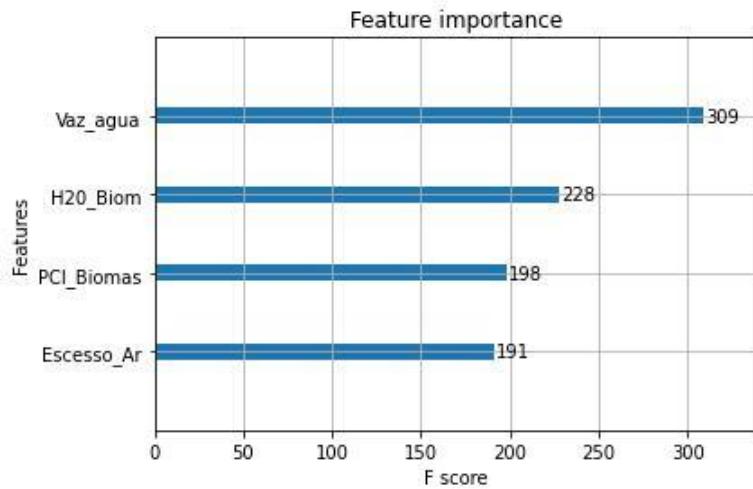
[TC_Gases]
count    31.000000
mean     194.298387
std      4.712509
min     185.410000
25%    192.200000
50%    194.630000
75%    196.750000
max    211.240000

Num=len(TC_Gases)
TC_Gases_x=np.copy(TC_Gases)
for i in range(Num):
    if(TC_Gases_x[i]<192.2):
        TC_Gases_x[i]=1 # TC_Gases=1 é baixo
    else:
        if(TC_Gases_x[i]>196):
            TC_Gases_x[i]=3 # TC_Gases=3 é alto
        else:
            TC_Gases_x[i]=2 # TC_Gases=1 é médio

y = TC_Gases_x
# fit model no training data
model = XGBClassifier()
model.fit(x, y)
# plot feature importance
plot_importance(model)
pyplot.show()

[usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversion y = column_or_1d(y, warn=True)

/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversion y = column_or_1d(y, warn=True)
```



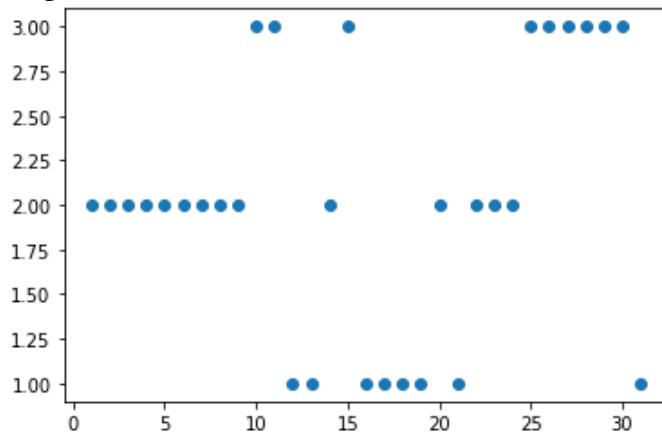
```
# split data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(x,y, test_size=0.33, random_state=7)
# fit model no training data
model = XGBClassifier()
model.fit(X_train, y_train)
# make predictions for test data
y_pred = model.predict(X_test)
predictions = [round(value) for value in y_pred]
# evaluate predictions
accuracy = accuracy_score(y_test, predictions)
print("Accuracy: %.2f%%" % (accuracy * 100.0))
# fit model no training data

kfold = KFold(n_splits=10, random_state=7)
results = cross_val_score(model, x, y, cv=kfold)
print("Accuracy cross: %.2f%% (%.2f%%)" % (results.mean()*100, results.std()*100))

→ Accuracy: 45.45%
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversion y =
column_or_1d(y, warn=True)
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversion y =
column_or_1d(y, warn=True)
/usr/local/lib/python3.6/dist-packages/sklearn/model_selection/_split.py:296: FutureWarning
Accuracy cross: 38.33% (23.63%)

xw=range(1 (len(y)+1))
plt.scatter(xw,y)
```

↳ <matplotlib.collections.PathCollection at 0x7effbe3894a8>



```
Oxigenio=pd.DataFrame({'Oxigenio':X['Oxigenio']})  
#print(Oxigenio)
```

```
print(Oxygenio.describe())
```

```

    ↗ Oxigenio
count 31.000000
mean 4.960968
std 0.360595
min 4.150000
25% 4.785000
50% 5.000000
75% 5.155000
max 6.000000

```

Num=len(Oxigenio)

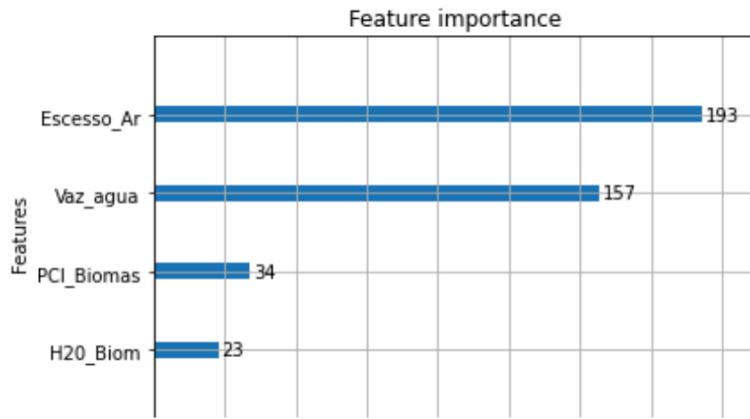
```
Oxigenio_x=np.copy(Oxigenio)
for i in range(Num):
    if(Oxigenio_x[i]<4.785):
        Oxigenio_x[i]=1 # Oxigenio=1 é baixo
    else:
        if(Oxigenio_x[i]>5.155):
            Oxigenio_x[i]=3 # Oxigenio=3 é alto
        else:
            Oxigenio_x[i]=2 # Oxigenio=1 é médio
```

```
y = Oxigenio_x  
# fit model no training data  
model = XGBClassifier()  
model.fit(x, y)  
# plot feature importance  
plot_importance(model)  
pyplot.show()
```

↳

```
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversion y =
column_or_1d(y, warn=True)
```

```
/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversion y =
column_or_1d(y, warn=True)
```



```
# split data into train and test sets
X_train, X_test, y_train, y_test = train_test_split(x,y, test_size=0.33, random_state=7)
# fit model no training data
model = XGBClassifier()
model.fit(X_train, y_train)
# make predictions for test data
y_pred = model.predict(X_test)
predictions = [round(value) for value in y_pred]
# evaluate predictions
accuracy = accuracy_score(y_test, predictions)
print("Accuracy: %.2f%%" % (accuracy * 100.0))
# fit model no training data

kfold = KFold(n_splits=10, random_state=7)
results = cross_val_score(model, x, y, cv=kfold)
print("Accuracy cross: %.2f%% (%.2f%%)" % (results.mean()*100, results.std()*100))
```

↳ /usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:235: DataConversion y =
column_or_1d(y, warn=True)

/usr/local/lib/python3.6/dist-packages/sklearn/preprocessing/_label.py:268: DataConversion y =
column_or_1d(y, warn=True)

/usr/local/lib/python3.6/dist-packages/sklearn/model_selection/_split.py:296: FutureWarning

FutureWarning:

Accuracy: 90.91%

Accuracy cross: 76.67% (26.03%)

```
xw=range(1,(len(y)+1))
```

```
plt.scatter(xw,y)
```

⇨

<matplotlib.collections.PathCollection at 0x7effbe2e1fd0>

