



Programa de Pós-Graduação em Engenharia
de Recursos Naturais da Amazônia

**CONTROLE DE QUALIDADE DE BAUXITAS GIBBSÍTICAS: PREDIÇÃO DOS
PARÂMETROS $AvAl_2O_3$ E $RxSiO_2$ A PARTIR DE DADOS DIFRATOMÉTRICOS
POR REFLEXÃO E TRANSMISSÃO UTILIZANDO ESTATÍSTICA
MULTIVARIADA**

Caio César Amorim de Melo



**CONTROLE DE QUALIDADE DE BAUXITAS GIBBSÍTICAS: PREDIÇÃO DOS
PARÂMETROS $AvAl_2O_3$ E $RxSiO_2$ A PARTIR DE DADOS DIFRATOMÉTRICOS
POR REFLEXÃO E TRANSMISSÃO UTILIZANDO ESTATÍSTICA
MULTIVARIADA**

Caio César Amorim de Melo

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Engenharia de Recursos Naturais da Amazônia, ITEC, da Universidade Federal do Pará, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Recursos Naturais.

Orientadora: Simone Patrícia Aranha da Paz

Belém

Fevereiro de 2021

“CONTROLE DE QUALIDADE DE BAUXITAS GIBBSÍTICAS: PREDIÇÃO DOS PARÂMETROS $AvAl_2O_3$ E $RxSiO_2$ A PARTIR DE DADOS DIFRATOMÉTRICOS POR REFLEXÃO E TRANSMISSÃO UTILIZANDO ESTATÍSTICA MULTIVARIADA”

Caio César Amorim de Melo

TESE SUBMETIDA AO CORPO DOCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE RECURSOS NATURAIS DA AMAZÔNIA - PRODERNA/ITEC/UFPA COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA OBTENÇÃO DO GRAU DE DOUTOR EM ENGENHARIA DE RECURSOS NATURAIS.

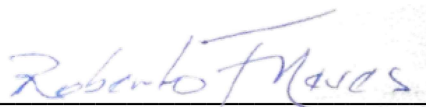
Aprovada por:



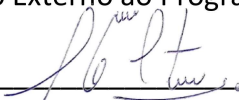
Prof.ª Dr.ª Simone Patrícia Aranha da Paz
(Orientadora – PRODERNA/UFPA)



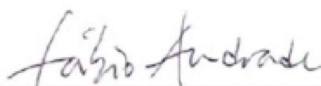
Prof. Dr. Marcos Allan Leite dos Reis
(Membro Interno – PRODERNA/UFPA)



Prof. Dr. Roberto de Freitas Neves
(Membro Externo ao Programa – UFPA)



Prof. Dr. Oscar Jesus Choque Fernandez
(Membro Externo à Instituição - IFPA)



Prof. Dr. Fábio Ramos Dias de Andrade
(Membro Externo à Instituição - USP)

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

M528c Melo, Caio César Amorim de.
 CONTROLE DE QUALIDADE DE BAUXITAS
 GIBBSÍTICAS: PREDIÇÃO DOS PARÂMETROS $AvAl_2O_3$ E
 R_xSiO_2 A PARTIR DE DADOS DIFRATOMÉTRICOS POR
 REFLEXÃO E TRANSMISSÃO UTILIZANDO ESTATÍSTICA
 MULTIVARIADA / Caio César Amorim de Melo. — 2021.
 xv, 65 f. : il. color.

Orientador(a): Prof^a. Dra. Simone Patrícia Aranha da Paz
Tese (Doutorado) - Universidade Federal do Pará, Instituto de
Tecnologia, Programa de Pós-Graduação em Engenharia de
Recursos Naturais na Amazônia, Belém, 2021.

1. Bauxita. 2. Alumina Aproveitável. 3. Sílica Reactiva. 4.
PLSR. 5. PCA. I. Título.

CDD 622

*Confie no Senhor de todo teu coração e não se apoie na sua própria inteligência.
Mas, se alguém tem falta de sabedoria, peça a Deus, e ele a dará porque é generoso e dá
com bondade a todos.*

Pv 3:5, Tg 1:6, Bíblia Sagrada

AGRADECIMENTOS

Agradeço primeiramente ao meu DEUS (a Ele toda honra e toda glória) por ter me concedido a dádiva da vida, a graça da saúde e o favor imerecido da sabedoria necessária para este completar desafio.

A minha esposa Tamires pelo inestimável apoio e compreensão ao longo de toda essa minha trajetória. Aos meus tesouros, Theo e Heloísa que me deram impagáveis momentos de alegria e porque são minha motivação de cada conquista.

Aos meus pais, Wagns e Georgina por serem meus exemplos de vida, e por todo sacrifício que a mim dispensaram até hoje.

A toda minha família! Amo todos vocês! Sem dúvidas essa minha caminhada até aqui foi graças ao amor, suporte e inspiração de vocês!

Aos professores Rômulo e Simone, meus orientadores em toda carreira acadêmica e amigos que tenho em grandíssima estima. Agradeço a Deus por ter colocado vocês na minha vida!

A todos os meus amigos e companheiros do LCM, em especial ao Aldemir, Alan, David, Andreia, Alexandre e Maurício que me ajudaram com as análises e discussões. Muito prazeroso o tempo que estive com todos vocês no laboratório.

Aos meus gestores na Hydro Erik e Marcelo que sempre me dispuseram o tempo necessário para a continuação / conclusão desta pesquisa.

A todos os meus amigos que acreditaram em mim, me ajudaram e pelos bons momentos de diversão e descontração.

Também agradeço a banca composta para avaliação do trabalho por aceitarem o convite.

DEDICATÓRIA

*Dedico este trabalho à minha família, Tamires, Theo, Heloisa e o pequeno Lucas, e aos
meus pais, Wagns e Gina.
Vocês são minha inspiração.*

Resumo da Tese apresentada ao PRODERNA/UFPA como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

**CONTROLE DE QUALIDADE DE BAUXITAS GIBBSÍTICAS: PREDIÇÃO DOS
PARÂMETROS $AvAl_2O_3$ E $RxSiO_2$ A PARTIR DE DADOS DIFRATOMÉTRICOS
POR REFLEXÃO E TRANSMISSÃO UTILIZANDO ESTATÍSTICA
MULTIVARIADA**

Caio César Amorim de Melo

Fevereiro/2021

Orientadora: Simone Patrícia Aranha da Paz

Área de Concentração: Transformação de Recursos Naturais

Atualmente, métodos tradicionais de química úmida são usados para controle de qualidade de bauxitas. Tais métodos quantificam indiretamente os teores de gibbsita e caulinita como alumina aproveitável ($AvAl_2O_3$) e sílica reativa ($RxSiO_2$), respectivamente, e são caros e demorados. Com o objetivo de obter um método rápido e confiável para estimar esses parâmetros como alternativa aos métodos atuais de química úmida, neste trabalho avaliou-se o uso de estatísticas multivariadas – Regressão por Mínimos Quadrados Parciais (PLSR) em dados de DRX de bauxitas brasileiras. Os difratogramas de raios-X foram coletados nos modos reflexão e transmissão, e os dados coletados por cada um desses tratamentos foram comparados com relação à qualidade dos modelos PLSR construídos. O método foi otimizado por meio da Análise de Componentes Principais (PCA) e Planejamento Fatorial de Experimentos (DOE), dos quais pode-se identificar *outliers*, agrupar amostras em três grupos com similaridades mineralógicas (C1, C2 e C3), e obter parâmetros otimizados de coleta e pré-tratamento dos difratogramas. Os melhores resultados foram obtidos utilizando o modo de coleta do difratograma por reflexão, reduzindo o range para $13^\circ - 34^\circ 2\theta$, aumentando o tamanho do passo de $0,026^\circ$ para $0,065^\circ$, e utilizando autoescalonamento dos dados. Essas condições de coleta, embora não ideais para a maioria das aplicações de DRX,

proporcionaram tanto uma melhor precisão dos modelos preditivos de $AvAl_2O_3$ e $RxSiO_2$ quanto uma redução no tempo de coleta (~ 40 segundos). Os resultados demonstraram que a precisão obtida ficou dentro dos limites aceitáveis industrialmente no controle de qualidade das bauxitas gibbissíticas ($AvAl_2O_3 = 0,49\%$ e $0,83\%$, e $RxSiO_2 = 0,32\%$ e $0,23\%$, respectivamente para amostras dos grupos C1 e C2). A predição não foi satisfatória apenas para amostras de bauxitas marginais (agrupadas majoritariamente em C3). A coleta de DRX por transmissão permitiu a eliminação do efeito de orientação preferencial, porém, a precisão do modelo foi aceitável apenas para amostras de C1. Em comparação com o método tradicional por química úmida, o método proposto se mostrou mais rápido, fácil de implementar e executar as análises, exige menos espaço e mão de obra de laboratório, além de não ser necessário reagentes químicos. Além disso, com a implementação da difratometria de raios-X em laboratórios de controle da indústria da bauxita e alumina, é possível acompanhar a mineralogia do minério que alimenta o processo Bayer e, portanto, estar ciente de como variações na composição mineralógica podem impactar o processo. Sendo válido ressaltar que tais informações são desconhecidas controlando apenas os parâmetros químicos.

Palavras-chave: Bauxita, Alumina Aproveitável, Sílica Reativa, PLSR, PCA

Abstract of Thesis presented to PRODERNA/UFPA as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.).

**QUALITY CONTROL OF GIBBSITIC BAUXITES: PREDICTION OF
PARAMETERS $AvAl_2O_3$ AND $RxSiO_2$ FROM REFLECTION AND
TRANSMISSION DIFFRATOMETRIC DATA BY USING MULTIVARIATE
STATISTICS**

Caio César Amorim de Melo

February/2021

Advisor: Simone Patrícia Aranha da Paz

Research Area: Transformation of Natural Resources

Currently, traditional wet chemistry methods are used for quality control of bauxites. Such methods indirectly quantify the gibbsite and kaolinite content as available alumina ($AvAl_2O_3$) and reactive silica ($RxSiO_2$), respectively, and they are very costly and time-consuming. In order to achieve a rapid and reliable method to estimate these parameters, as alternative to current wet chemistry methods, in this work it is evaluated the use of multivariate statistics – Partial Least Square Regression (PLSR) on XRD data of Brazilian bauxites. The X-ray diffractograms were collected in the reflection and transmission modes, and the data collected by each of these treatments were compared with respect to the quality of the PLSR models. The method was optimized through Principal Component Analysis (PCA) and Factorial Design of Experiments (DOE), from which it was possible to identify outliers, grouping samples with mineralogical similarities into three clusters (C1, C2 and C3), and obtain optimized parameters for the collection and pre-treatment of diffractograms. The best results were obtained using the reflection mode, reducing the 2θ range to $13^\circ - 34^\circ$ 2θ , increasing the step size from 0.026° to 0.065° , and using standardized data. These collection conditions, although not ideal for most XRD applications, provided both a better accuracy of the predictive models of $AvAl_2O_3$ and $RxSiO_2$ and a reduction in the collection time (~ 40 seconds). The results showed that the precision obtained was within the industrially

acceptable limits for the quality control of gibbissitic bauxites ($AvAl_2O_3 = 0.49\%$ and 0.83% , and $RxSiO_2 = 0.32\%$ and 0.23% , respectively for samples of the groups C1 and C2). The prediction was not satisfactory only for marginal bauxites samples (grouped mainly in C3). XRD by transmission allowed the elimination of the preferential orientation effect, however, the accuracy of the model was acceptable only for C1 samples. Compared to the traditional wet chemistry, the proposed method is significantly faster, easier to implement and perform the analyzes, requires less space and manpower, besides no chemical reagents are needed. In addition, with the implementation of X-ray diffraction in the laboratory of the bauxite and alumina industry, it is possible to follow the mineralogy of the ore that feeds the Bayer process and, therefore, to be aware of how variations in the mineralogical composition can impact the process. It is worth noting that such information is still unknown, controlling only the chemical parameters.

Keyword: Bauxite, Available Alumina, Reactive Silica, PLSR, PCA

SUMÁRIO

1. INTRODUÇÃO	1
1.1. MOTIVAÇÃO.....	2
1.2. OBJETIVOS.....	3
1.2.1. Geral	3
1.2.2. Específicos.....	3
2. REVISÃO DA LITERATURA.....	4
2.1. A BAUXITA	4
2.1.1. Formação dos depósitos de bauxitas lateríticas.....	6
2.1.2. Bauxitas com alta sílica reativa	7
2.1.3. Bauxitas da região de Paragominas-PA	8
2.1.4. Mineralogia das bauxitas lateríticas e Processo Bayer.....	10
2.2. CONTROLE DE QUALIDADE DE BAUXITAS	14
3. MATERIAIS E MÉTODOS	18
3.1. MATERIAIS	18
3.2. MÉTODOS E EQUIPAMENTOS	20
3.2.1. Preparação das amostras.....	20
3.2.2. Preparação das espécimes – DRX por reflexão e transmissão	21
3.2.3. Conjunto de dados	23
3.2.1. Análises de $AvAl_2O_3$ e $RxSiO_2$ por via úmida.....	25
3.2.2. Difração de raios-X (DRX)	26
3.2.3. Análise de Componentes Principais (PCA).....	28
3.2.4. Regressão por Mínimos Quadrados Parciais (PLSR).....	29
3.2.5. Planejamento Fatorial de Experimentos (DOE).....	31
4. RESULTADOS E DISCUSSÕES	33
4.1. CARACTERIZAÇÃO DAS BAUXITAS	33
4.2. AVALIAÇÃO PRELIMINAR DOS DADOS QUÍMICOS DAS BAUXITAS	35
4.3. AVALIAÇÃO DOS DADOS POR PCA.....	37
4.3.1. Dados químicos $AvAl_2O_3$ e $RxSiO_2$	37
4.3.2. Difratogramas obtidos pelo modo Reflexão.....	38

4.3.3. Difractogramas obtidos pelo modo Transmissão	41
4.4. AVALIAÇÃO DO PRÉ-PROCESSAMENTO DOS DADOS DE DRX.....	44
4.4.1. Predição por PLSR a partir dos dados de DRX por reflexão	46
4.4.2. Predição por PLSR a partir dos dados de DRX por transmissão	51
5. CONCLUSÕES.....	54
5.1. SUGESTÕES PARA TRABALHOS FUTUROS	56
6. REFERÊNCIAS	57
7. APÊNDICE – ARTIGOS DA TESE.....	65

LISTA DE FIGURAS

Figura 2.1: Estruturação do perfil laterítico-bauxítico típico do platô Miltônia, Paragominas-PA.....	8
Figura 2.2: Mineração Paragominas SA – NorksHydro: (A) usina; (B) mina; e (C) barragens de rejeito.....	9
Figura 2.3: Comparação da cristalinidade das caulinitas em cada litologia do perfil; b) DRX mostrando maior reatividade das litologias BN e BNC após a digestão em 1 h.....	13
Figura 3.1: a) Esquemático da malha de sondagem 25m x 25m; b) Perfil da seção destacada na malha de sondagem, Platô Miltonia 3. Adaptado de Picanço <i>et al.</i> (2013).....	19
Figura 3.2: Quantidade de amostras obtidas na campanha de sondagem e número de amostras selecionadas para realização do presente trabalho.....	20
Figura 3.3: Fluxograma da preparação física das amostras feitas pela empresa Hydro.....	21
Figura 3.4: a) Procedimento de preenchimento dos porta amostras para DRX pelo método <i>backloading</i> para coleta de dados por reflexão. Fonte: Oliveira (2017); e b) porta-amostra (esquerda) e filme Mylar® (direita) utilizados para coleta de dados por transmissão.....	23
Figura 3.5: Representação do dataset de amostras. Difrátogramas arranjados como matriz.....	24
Figura 3.6: Fluxograma das etapas do presente trabalho.....	25
Figura 3.7: Pré-processamento dos dados. Os dados de cada variável são apresentados pela barra de variação e seus valores centrais (médias). A) exemplo comum de dados brutos, B) dados centrados na média, C) dados padronizados pelo desvio-padrão, e D) dados centrados na média e padronizados – também conhecido como <i>autoescalamento</i>	29
Figura 4.1: Padrões difratométricos de todas as amostras de bauxita utilizadas neste estudo. Destaque para o <i>offset</i> do background entre as amostras.....	34
Figura 4.2: Padrões difratométricos de todas as amostras de bauxita utilizadas neste estudo obtidas por transmissão.....	34
Figura 4.3: Histograma das amostras de bauxita com relação aos dados químicos; a) $AvAl_2O_3$; b) $RxSiO_2$; e c) $RxSiO_2/AvAl_2O_3$	36
Figura 4.4: Gráfico de dispersão dos dados químicos de referência.....	37
Figura 4.5: Gráficos de dispersão da análise de PCA a partir dos dados químicos de referência; a) escores, e b) cargas.....	38

Figura 4.6: Gráficos de dispersão da análise de PCA a partir dos dados difratométricos; a) escores, e b) cargas.....	39
Figura 4.7: Representação linear das cargas (<i>loadings</i>) dos componentes principais PC1 e PC2 da análise de PCA a partir dos dados difratométricos.....	40
Figura 4.8: a) Gráfico de influência das amostras. Pontos pretos no quadrante inferior direito representam amostras com alto <i>leverage</i> ; pontos cinza no canto superior esquerdo são amostras com alto resíduo; b) comparação dos padrões difratométricos das amostras com alto <i>leverage</i> (região de reflexão da gibbsita d_{002}).....	41
Figura 4.9: Gráficos de dispersão da análise de PCA a partir dos dados difratométricos obtidos por transmissão; a) escores PC1xPC2, e b) escores PC1xPC3.....	42
Figura 4.10: a) Agrupamento das amostras por <i>K-means</i> representados no gráfico de dispersão da análise de PCA.....	43
Figura 4.11: Representação linear das cargas (<i>loadings</i>) dos componentes principais a) PC1 e PC2 e b) PC6.....	43
Figura 4.12: Comparação entre os métodos de pré-processamento dos dados - centrado na média (RMSECV = 0,845, RMSEP = 1,295, com 8 fatores) e autoescalado (RMSECV = 0,924 e RMSEP = e 0,804, com 5 fatores).....	46
Figura 4.13: Comparação entre estimativa e os valores de referência; a) $AvAl_2O_3$, b) $RxSiO_2$	47
Figura 4.14: Gráfico de resíduos médios para os modelos de cada grupo (reflexão).....	48
Figura 4.15: Representação linear das cargas do modelo PLSR para $RxSiO_2$ a partir das amostras de calibração do grupo C3.....	49
Figura 4.16: Padrão difratométrico das amostras do grupo C3 – mostrando o intervalo da reflexão principal da caulinita (d_{001}).....	50
Figura 4.17: Comparação dos valores de referência e estimados através dos modelos PLSR para cada grupo (DRX por transmissão). a) $AvAl_2O_3$, b) $RxSiO_2$	52
Figura 4.18: Gráfico de resíduos médios para os modelos de cada grupo.....	53

LISTA DE TABELAS

Tabela 2.1: Composição química das Bauxitas para uso metalúrgico e não metalúrgico.....	5
Tabela 2.2: Composição química típica de bauxitas lateríticas.....	10
Tabela 3.1: Fatores e níveis do DOE 2 ⁴	32
Tabela 3.2: Matriz de experimento do DOE 2 ⁴ com níveis codificados.....	32
Tabela 4.1: Resumo do critério de seleção de amostras para os modelos PLSR e n° de amostras em cada faixa de seleção.....	36
Tabela 4.2: Tempo estimado de coleta dos dados de difração de raios-X (baseado nas características do difratômetro utilizado neste trabalho).....	44
Tabela 4.3: ANOVA da estimativa de $RxSiO_2$ no planejamento fatorial 2 ⁴ . (RMSEP como variável resposta; R^2_{adj} = 0,935; nível de significância α = 0.05).....	46
Tabela 4.4: Valores de RMSEP para $AvAl_2O_3$ e $RxSiO_2$ para os modelos de cada grupo.....	48
Tabela 4.5: Resultados da predição de $AvAl_2O_3$ para os modelos de cada grupo.....	53
Tabela 4.6: Resultados da predição de $RxSiO_2$ para os modelos de cada grupo.....	53

LISTA DE SÍMBOLOS

AAS	<i>Atomic Absorption Spectroscopy</i> (espectroscopia por absorção atômica)
ADS	<i>Automatic Divergent Slit</i> (fenda divergente automática)
AvAl ₂ O ₃	<i>Available Alumina</i> (alumina disponível – medida indireta dos minerais de alumínio da bauxita – geralmente gibbsita, boehmita e/ou diásporo)
AMN	Agência Nacional de Mineração
BA	Bauxita Amorfa
BC	Bauxita Cristalizada
BCBA	Bauxita Cristalizada-Amorfa
BN	Bauxita Nodular
BNC	Bauxita Nodular-Cristalizada
DOE	<i>Design of Experiments</i> (planejamento experimental)
DRX	Difração de raios-X
DSP	<i>Dessilication Product</i> (produto da dessilicação Bayer)
FDS	<i>Fixed Divergent Slit</i> (fenda divergente fixa)
FRX	Fluorescência de raios-X
FWHM	<i>Full-Width-at-Half-Maximum</i> (largura a meia altura)
ICP	<i>Inductively Coupled Plasma</i> (espectroscopia por plasma acoplado indutivamente)
HTP	<i>High Temperature Digestion</i>
HYDRO	Norsk Hydro SA
LF	Laterita Ferruginosa
LOI	<i>Loss on Ignition</i> (perda ao fogo)
LTP	<i>Low Temperature Digestion</i>
MLR	Multiple Linear Regression (Regressão Múltipla)
MPSA	Mineração Paragominas SA
MRN	Mineração Rio do Norte SA
PCA	<i>Principal Component Analysis</i> (Análise de Componentes Principais)
PCR	<i>Principal Component Regression</i> (Regressão por Componentes Principais)

PLSR	<i>Partial Least Squares Regression</i> (Regressão por Mínimos Quadrados Parciais)
RMSE	<i>Root Mean Square Error</i>
RMSECV	<i>Root Mean Square Error of Cross-Validation</i>
RMSEP	<i>Root Mean Square Error of Prediction</i>
RSD	<i>Residual Standard Deviation</i> (desvio-padrão residual)
$RxSiO_2$	<i>Reactive Silica</i> (sílica reativa – medida indireta dos minerais de silício na bauxita – geralmente caulinita e/ou quartzo)
USGS	<i>U. S. Geological Survey</i> (Departamento americano de pesquisa geológica)

1. INTRODUÇÃO

Na indústria mineral como um todo, é comum ter parâmetros de controle de qualidade e processo baseados em dados químicos ao invés de mineralógicos. Isto ocorre principalmente devido à consolidação e disponibilidade dos métodos de análises químicas quantitativas por via úmida, enquanto que métodos para determinação mineralógicas ainda estão em desenvolvimento.

Na indústria do alumínio não é diferente. Tem-se tanto na mina quanto no processo Bayer, o controle de qualidade do minério realizado quase exclusivamente em termos da sua composição química. Assim, as amostras da pesquisa geológica, frente de lavra, planta de beneficiamento e minério final são, em geral, analisadas para determinação da composição química total (por fluorescência de raios X – FRX, plasma acoplado indutivamente – ICP ou absorção atômica – AAS), e dos teores de alumina disponível ($AvAl_2O_3$) e sílica reativa ($RxSiO_2$). Estes últimos, são determinação indireta dos minerais gibbsita/boehmita/diásporo e caulinita/quartzo, respectivamente – e são os dois principais parâmetros para o controle do processo, determinados tradicionalmente por métodos químicos de via úmida (Paz, 2016).

Nesse contexto, trabalhos recentes demonstraram que o teor de $RxSiO_2$ determinado por tais métodos pode ser subestimado, dependendo do conteúdo e do grau de cristalinidade da caulinita na bauxita, a qual muda significativamente ao longo do perfil bauxítico. Isto significa que o mineral de sílica presente pode ser mais reativo ao processo, apesar de sua concentração (Paz *et al.*, 2017; Melo *et al.*, 2017). Assim, não há garantia de que o simples conhecimento da composição química da bauxita permita um controle eficiente nos processos metalúrgicos. Diante disto, destaca-se o valor de conhecer a mineralogia e a textura dos minerais na bauxita, pois esses fatores determinam as interações físico-químicas que ocorrem nos processos (O'Connor, 1988; Paz *et al.*, 2017).

1.1. MOTIVAÇÃO

Como a determinação de $AvAl_2O_3$ e $RxSiO_2$ pelos métodos tradicionais de química úmida consome muito tempo (horas) e demanda uma grande quantidade de produtos químicos, mão de obra e espaço de laboratório, além de baixa reprodutibilidade (König *et al.*, 2012; Paz *et al.*, 2018), vários autores têm proposto métodos alternativos para obter estimativas confiáveis o mais prático, rápido (minutos) e mais barato possível. Geralmente, os métodos alternativos são baseados em dados de óxidos totais e perda ao fogo (*Loss on Ignition* – LOI), usando balanços estequiométricos e modelagem matemática (Bredell, 1983; Feret *et al.*, 1997; Najar, 2017; Carneiro *et al.*, 2017; Zhang *et al.*, 2019). A fraqueza de tais métodos se dá pela necessidade de se conhecer toda a mineralogia da bauxita, de forma que estes tendem à falhar quando alguma fase não é adequadamente identificada ou quando a característica da bauxita é alterada.

Acredita-se que a melhor forma para realizar o controle de qualidade de bauxitas é utilizando dados mineralógicos obtidos por difratometria de Raios-X (DRX). Tal método é reconhecidamente robusto, rápido e facilmente automatizado. Mas sua principal vantagem sobre os métodos por via úmida se dá por permitir uma compreensão mineralógica do minério e não uma estimativa indireta.

Estudos recentes apresentam bons avanços na classificação e quantificação mineralógica de bauxitas gibbsíticas com base no refinamento de Rietveld (Antoniassi, 2010, Paz *et al.*, 2018; Angélica *et al.*, 2018) e estatística multivariada (König *et al.*, 2012; Melo *et al.*, 2017). Neste contexto, o presente trabalho teve como foco desenvolver uma metodologia para controle de qualidade de bauxitas baseadas em dados de DRX e utilizando técnicas de estatística multivariada para classificação de amostras e predição dos parâmetros de controle. Assim, buscou-se desenvolver uma metodologia preditiva para $AvAl_2O_3$ e $RxSiO_2$ que atenda, com base em parâmetros estatísticos, os limites aceitáveis na indústria (obtidos pelos métodos tradicionais por via úmida), porém, com as vantagens de ter informações mineralógicas disponíveis, de forma rápida, mais barata e automatizável.

1.2. OBJETIVOS

1.2.1. Geral

Desenvolver uma metodologia otimizada para estimar os principais parâmetros do controle de qualidade de bauxita gibbsíticas ($AvAl_2O_3$ e $RxSiO_2$), utilizando Regressão por Mínimos Quadrados Parciais (*Partial Least Square Regression* – PLSR) a partir de dados de difração de raios-X (DRX).

1.2.2. Específicos

- Caracterizar mineralogicamente as bauxitas gibbsíticas da região de Paragominas-PA.
- Determinar dentre um banco de dados as amostras das litologias bauxíticas que são representativas para serem utilizadas no modelo.
- Classificar e agrupar amostras com similaridades mineralógicas.
- Determinar quais as condições ideais de aquisição dos difratogramas para predição utilizando PLSR.
- Avaliar técnicas de pré-processamento dos dados difratométricos para se obter condição otimizada de predição por PLSR.
- Determinar por modelos PLSR os parâmetros de controle de qualidade de bauxitas ($AvAl_2O_3$ e $RxSiO_2$) utilizando as condições definidas como ótimas.
- Comparar os modos de aquisição de DRX por reflexão e transmissão na predição por PLSR.
- Avaliar e validar estatisticamente os modelos PLSR obtidos.

2. REVISÃO DA LITERATURA

2.1. A BAUXITA

A bauxita é um material regolítico, produto do intemperismo de rochas aluminossilicatadas com elevada concentração de minerais de alumínio que perfaz reservas economicamente exploráveis (Freyssinet *et al.*, 2005). Foi descoberta em 1821 pelo químico francês Pierre Berthier, e é assim nomeada em homenagem a região de Les Baux no sul da França, onde foi primeiramente encontrada (Authier-Martin *et al.*, 2001 e Santos, 2011). Sua coloração é determinada pela quantidade de minerais de ferro presente, sendo levemente avermelhada em rochas com proporções menores que 4% e vermelha em rochas que contém acima de 25% (Sampaio *et al.*, 2005).

Existem várias formas de classificar as bauxitas, dependendo basicamente da sua composição química e mineralógica, da gênese dos depósitos e da finalidade tecnológica:

- Quanto à composição química e mineralógica, são classificadas em função do mineral óxi-hidróxido de alumínio presente majoritariamente no depósito. Dessa forma, as bauxitas podem ser: *gibbsíticas*, com a presença predominante do mineral gibbsita ($\text{Al}(\text{OH})_3$), *boehmítica* ou *diaspórica*, pela presença predominante dos minerais boehmita e diásporo ($\text{AlO}(\text{OH})$), respectivamente (Smith, 2009). Segundo Freyssinet *et al.* (2005), podem ainda ser classificadas como *High-grade* e *Low-grade*, isto em função dos teores de sílica reativa (R_xSiO_2), que está principalmente relacionada a caulinita. Assim bauxitas denominadas *High-grade* são aquelas que possuem pouca quantidade de caulinita, e a designação *Low-grade* aplica-se àquelas com altos teores de caulinita.
- Quanto à gênese dos depósitos, as bauxitas são classificadas como *cársticas* ou *lateríticas*. Os depósitos de bauxitas cársticas se formaram basicamente pelo acúmulo de minerais de alumínio ocasionado pela dissolução de carbonatos, e os depósitos de bauxitas lateríticas se desenvolveram a partir do processo de intemperismo de rochas aluminossilicatadas em condições de clima tropical e

subtropical. De modo geral, as atenções e estudos se voltam principalmente as bauxitas lateríticas, visto que, estas representam 90% das reservas globais de bauxita (Freyssinet *et al.*, 2005).

- Quanto à finalidade tecnológica, elas podem ser: *metalúrgica* e *não metalúrgica*. As metalúrgicas são destinadas a obtenção do alumínio metálico e as não metalúrgicas a fabricação de refratários, abrasivos e outros usos cerâmicos.

A Tabela 2.1 mostra, em resumo, as composições químicas que caracterizam as bauxitas em metalúrgicas e não metalúrgicas. Em geral, bauxitas metalúrgicas devem ter teor de alumina aproveitável ($AvAl_2O_3$ – medida indireta dos minerais de alumínio) de cerca de 50% e de sílica reativa abaixo de 5%. Bauxitas não metalúrgicas devem ter alto teor de alumina e pouca impureza, principalmente Fe_2O_3 (Sampaio *et al.*, 2005 e Santos, 2011).

Tabela 2.1: Composição química das Bauxitas para uso metalúrgico e não metalúrgico.

Bauxita metalúrgica		Bauxita não metalúrgica	
Constituintes	%	Constituintes	%
$AvAl_2O_3$	> 48	$AvAl_2O_3$	50
$RxSiO_2$	< 4	$RxSiO_2$	5 – 7
Fe_2O_3	11 – 12	Fe_2O_3	≤ 2,5

Fonte: Sampaio *et al.*, 2005 – adaptado.

Segundo dados mais recentes do US Geological Survey (USGS, 2021) e Agência Nacional de Mineração (ANM - Sumário Mineral 2018), as reservas mundiais de bauxita atualmente somam 30 bilhões de toneladas e em 2020 a produção foi de 310 Mton/ano. No Brasil, tem-se disponível como reservas lavráveis um total de 2,7 bilhões de toneladas (Brasil, 2019).

Em 2020 o Brasil alcançou uma produção de bauxita da ordem de 35 Mton, colocando o Brasil como 4º maior produtor de bauxita, atrás apenas da Austrália, China e Guiné, respectivamente (USGS, 2021). Segundo Brasil (2019), dentre as reservas brasileiras de bauxita, as mais expressivas estão localizadas no estado do Pará, e atualmente, três minas se destacam na produção de bauxita tipo metalúrgica nessa região, sendo estas responsáveis por mais de 90% de toda produção brasileira de bauxita metalúrgica:

a) Porto Trombetas, que está em operação desde o final dos anos 70, pertencente a Mineração Rio do Norte (MRN, Vale);

- b) Juruti, pertencente a ALCOA, iniciada em 2008; e
- c) Mina Miltônia 3, na região de Paragominas, iniciada em 2006 pela Vale e hoje pertencente a Norsk Hydro.

2.1.1. Formação dos depósitos de bauxitas lateríticas

Hoje, sabe-se que os depósitos lateríticos foram formados como resultado do intemperismo de uma grande variedade de rochas sedimentares, ígneas e metamórficas, em regiões de clima tropical e úmidas (Bocquier *et al.*, 1984).

A intensidade do processo de intemperismo varia de acordo com as composições químicas e mineralógicas das rochas matriz, e com as condições climáticas, de relevo e drenagem da região.

De forma geral, durante o processo todos os metais alcalinos e alcalino-terrosos dos minerais dissolvidos são completamente removidos, enquanto o Si pode ser completamente ou parcialmente lixiviada (mecanismo de alitização e monossilicização, respectivamente). Dessa forma, pode-se distinguir o processo de bauxitização das rochas matriz em dois tipos: direto e indireto. Na bauxitização direta, pela hidrólise dos minerais originais os únicos minerais neoformados são gibbsita e goethita. Enquanto Al e Fe acumulam-se nos contornos da estrutura dos grãos e nas fissuras e clivagens, todos os elementos maiores da rocha matriz inclusive Si são removidos. Durante a bauxitização indireta, a hidrólise dos minerais originais é incompleta, fazendo com que Fe, Al e parte da Si acumulem-se formando goethita e caulinita. A caulinita dissolve-se formando gibbsita e o Fe contido na goethita é então redistribuído e forma concreções de hematita nos microporos (Boulangé *et al.*, 1997; Toledo, 2000).

Os processos envolvidos durante a bauxitização direta e indireta são resumidamente: dissolução completa ou parcial dos minerais originais, perda da coesão das rochas matriz e aumento da porosidade, além da conservação de alguns minerais originais, que se fraturam, mas permanecem inalterados. Os horizontes que ainda não estão num estágio final de bauxitização, são chamados de saprólitos (Toledo, 2000).

Esses dois processos levam a formação das chamadas bauxitas originais, as quais ao longo do tempo podem passar por transformações químicas, estruturais e mineralógicas, formando bauxitas degradadas e novas unidades de acumulação, de acordo com o tipo de

transferência química. Por exemplo, crostas argilo-ferruginosas são formadas pela migração de Fe em um saprólito adjacente. Novas fáceis caulíníticas podem se formar pela ressilificação de acordo com as condições físico-químicas do ambiente de intemperismo (temperatura, humidade, vegetação) e topográfica (declive e drenagem). Os minerais neoformados podem estar associados à minerais originais da rocha matriz, assim, muitas unidades de intemperismo (horizontes, fáceis) podem ser definidas pelas relações textural e estrutural das fases mineralógicas presentes. Essas diferentes unidades de intemperismo estão diretamente associadas ao mecanismo (direto ou indireto) de bauxitização que os originaram (Boulangé *et al.*, 1997).

2.1.2. Bauxitas com alta sílica reativa

Alguns perfis de depósitos lateríticos-bauxíticos apresentam horizontes de bauxita com conteúdo expressivo de caulinita e em muitos casos são tratadas como minérios marginais ou gangas, e não são aproveitadas ainda que possam apresentar teor significativo de gibbsita. Estas bauxitas são frequentemente denominadas bauxitas com alta sílica reativa.

O limite percentual de sílica reativa presente nas bauxitas que constituem o minério que alimenta o processo Bayer é de forma geral, determinado pelas próprias empresas, e dependendo das características da refinaria (processo Bayer). No entanto, vários autores citam valores limites de caulinita que caracterizam as bauxitas com alta sílica reativa que, de forma geral, são aquelas onde o teor de caulinita está acima de até 5% em peso (Grubbs, 1987; O'Connor, 1988; McCormick *et al.*, 2002; Rayzman *et al.*, 2003; Smith, 2009).

A obtenção de alumina via processo Bayer, a partir dessas bauxitas é economicamente inviável, pois durante a etapa de lixiviação há significativa perda de hidróxido de sódio (agente lixiviante do processo Bayer) pela formação de um precipitado insolúvel (Sodalita Bayer) que, posteriormente, é descartado com os demais resíduos sólidos do processo. Devido à esta perda, estima-se que a quantidade soda cáustica necessária para processar bauxitas com alta sílica reativa acarretem um considerável aumento de cerca de 20% nos custos de produção de alumina metalúrgica (McCormick *et al.*, 2002 e Smith, 2009). Rayzman *et al.* (2003) apresentam um cálculo estequiométrico da sodalita formada durante o processo Bayer, onde esta fase insolúvel tem 0,5 mol de Al_2O_3 por mol de SiO_2 , e como resultado, para cada 1 kg de SiO_2 dissolvida no processo 0,666 kg de NaOH e 0,85 kg de

Al_2O_3 são perdidos irreversivelmente. Isto acarreta uma perda que pode chegar a 132 kg de NaOH e 170 kg de Al_2O_3 por tonelada de bauxita processada.

2.1.3. Bauxitas da região de Paragominas-PA

Os platôs do Distrito de Paragominas são capeados por um espesso latossolo amarelo conhecida como Argila de Belterra, que sobrepõem as crostas lateríticas ferruginosas e aluminosas. As aluminosas correspondem às bauxitas (Costa, 2010 e Mendes, 2012). A Figura 2.1 ilustra a estruturação de um perfil típico da região de Paragominas.

Uma característica peculiar das bauxitas de Paragominas diz respeito aos altos teores de caulinita. Por isso a necessidade de operações de desagregação e lavagem na usina de beneficiamento.

Paz *et al.* (2012) explicam que há uma problemática durante estas operações. Isto ocorre porque a liberação da gibbsita do mineral de ganga (caulinita) depende da relação textural entre esses dois minerais (intercrescimento). Entretanto, esta relação muda consideravelmente, no espaço de pequenas distâncias dentro do mesmo depósito, ou até mesmo dentro do mesmo nível ou horizonte de bauxita.



Figura 2.1: Estruturação do perfil laterítico-bauxítico típico do platô Miltônia, Paragominas-PA (Fonte: Silva *et al.*, 2008, adaptado).

Nem todos os níveis de bauxita são explorados, visto que a quantidade de sílica reativa nestas bauxitas mesmo após sua lavagem não atende às especificações. O desafio de reduzir a concentração de caulinita do minério e de se aproveitar essas bauxitas marginais (consideradas como estéril durante a lavra), tem impulsionado diversos estudos na área do beneficiamento dessas bauxitas, principalmente em um contexto onde minas contendo minério de alto teor vão se esgotando (Paz *et al.*, 2012, Novo *et al.*, 2015).

Atualmente, a ganga caulínica (rejeito da lavagem da bauxita) removida do minério durante o seu beneficiamento é depositada em grandes e onerosas áreas situadas perto da lavra, formando enormes lagoas de rejeito (Figura 2.2).

Análise mineralógica quantitativa desse rejeito mostra que sua composição média é: 48% gibbsita, 32% caulinita e 20% outros minerais (Nascimento *et al.*, 2013). Isto denota a forte relação textural desses minerais, de forma que, um *trade off* muito estreito entre recuperação mássica e recuperação metalúrgica aconteça durante o beneficiamento físico. Assim, embora o pretendido seja remover apenas caulinita, muita gibbsita é perdida no rejeito. Mas, ainda é mais economicamente viável perder gibbsita na remoção de caulinita do que soda cáustica no processo Bayer.



Figura 2.2: Mineração Paragominas SA – NorksHydro: (A) usina; (B) mina; e (C) barragens de rejeito (Fonte: Google Earth - 2020, adaptado).

2.1.4. Mineralogia das bauxitas lateríticas e Processo Bayer

Maior foco é geralmente dado para as bauxitas lateríticas devido sua ampla representatividade de reservas mundiais (Paz, 2016; Freyssinet et al., 2005). Segundo Authier-Martin *et al.* (2001), essas bauxitas apresentam mineralogia que pode variar de acordo com a geologia do depósito e sua localização geográfica. No entanto, estão sempre associadas à quatro elementos maiores: Al, Si, Fe e Ti, além de pelo menos outros 13 elementos menores. Esses elementos estão presentes como diversos minerais, conforme se observa na Tabela 2.2

Tabela 2.2: Composição química típica de bauxitas lateríticas.

Elementos	Mineral	Composição Química
Maiores		
Alumínio	Gibbsita	$\text{Al}(\text{OH})_3$
	Boehmita/Diásporo	$\text{AlO}(\text{OH})$
Silício	Caulinita/Halloysita	$\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$
	Quartzo	SiO_2
Ferro	Hematita	Fe_2O_3
	Goethita	$(\text{Fe},\text{Al})\text{OOH}$
Titânio	Anatásio	TiO_2
	Rutilo	TiO_2
Menores		
Carbono	Carbono orgânico	matéria húmica
Fósforo	Crandallita	$\text{CaAl}_3(\text{PO}_4)(\text{PO}_3\text{OH})(\text{OH})_6$
	Wavellita	$\text{Al}_3(\text{PO}_4)_2(\text{OH})_3 \cdot 5\text{H}_2\text{O}$
Cálcio	Calcita	CaCO_3
	Crandallita	$\text{CaAl}_3(\text{PO}_4)(\text{PO}_3\text{OH})(\text{OH})_6$
Potássio	Muscovita	$\text{KAl}_2(\text{Si}_3\text{AlO}_{10})(\text{OH})_2$
	Illita	
Manganês	Litoforita	$(\text{Al},\text{Li})\text{MnO}_2(\text{OH})_2$
Magnésio	Magnesita	MgCO_3
	Dolomita	$\text{CaMg}(\text{CO}_3)_2$
Sódio	Dawsonita	$\text{NaAlCO}_3(\text{OH})_2$
Estrôncio	Celestita	SrSO_4
Enxofre	Woodhouseita	$\text{CaAl}_3(\text{PO}_4)(\text{SO}_4)(\text{OH})_6$
	Pirita	FeS_2
Zinco	Gahnita	ZnAl_2O_4
Cromo	Cromita	FeCr_2O_4
Vanádio	Schubnelita	$\text{Fe}_2(\text{V}_2\text{O}_8) \cdot 2\text{H}_2\text{O}$
Zircônio	Zircão	ZrSiO_4

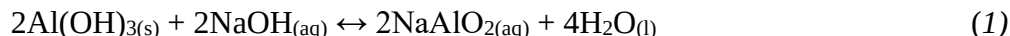
Fonte: Authier-Martin *et al.* (2001).

Vários autores destacam a importância de se conhecer a mineralogia das bauxitas, devido principalmente aos impactos no processo Bayer. Segundo Ostap (1984), paradoxalmente, as pessoas que estão mais cientes da origem, mineralogia e composição das bauxitas, são também aquelas menos informadas sobre a forma que esta pode afetar o desempenho e os custos de seu beneficiamento químico (processo Bayer). Neste trabalho, o autor se propôs a descrever com detalhes os impactos no processo causados pelos elementos químicos mais comuns nas bauxitas, de acordo com a fase mineralógica em que estão presentes.

Maior atenção é sempre dada às fases minerais que correspondem aos parâmetros de controle do processo – alumina aproveitável e sílica reativa. Nestes casos, o mineral de alumínio predominante na bauxita irá definir as condições do processo Bayer:

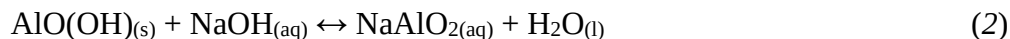
- *Low Temperature Digestion* (LTD), condições de operação da digestão menos dispendiosa energeticamente (temperatura entre 130 – 160 °C, menor pressão e concentração cáustica), é adequado exclusivamente para bauxitas gibbsíticas;

Reação de digestão da gibbsita ($\text{Al}(\text{OH})_3$):



- *High Temperature Digestion* (HTD), condições de digestão mais agressivas e com maior custo energético (temperatura entre 220 – 270 °C, pressões maiores e licor caustico mais concentrado), é utilizado para bauxitas boehmíticas (200 °C) e diaspóricas (270 °C) (Authier-Martin *et al.*, 2001).

Para a boehmita e diásporo ($\text{AlO}(\text{OH})$):

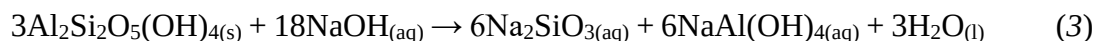


Problemas maiores ocorrem quando há misturas desses minerais de alumínio, como acontece, por exemplo, nas bauxitas gibbsíticas da Jamaica que contém entre 0,4 – 3,5% de

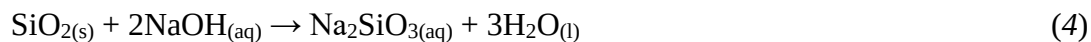
boehmita. Essa pequena quantidade de boehmita ocasiona uma perda significativa de alumina aproveitável devido à “reversão boehmíticas”, onde em condições LTD os cristais não dissolvidos dessa fase atuam como sementes para precipitar a alumina dissolvida no licor. Este precipitado é, por sua vez descartado no resíduo sólido do processo, conhecido como lama vermelha (Authier-Martin *et al.*, 2002).

Com relação à sílica reativa os impactos negativos se referem a perda da soda cáustica (agente lixiviante do processo). Conforme apresentado na Tabela 2.2, as fases minerais de sílica que são reativas as condições Bayer são caulinita/halloysita e quartzo. A caulinita e halloysita são facilmente solúveis em condições LTD, enquanto o quartzo, só é reativo em condições de digestão HTD (Ostap, 1984). Quando esses minerais são dissolvidos no licor Bayer se faz necessário realizar uma *dessilicação* do licor, uma vez que a sílica dissolvida no licor impactará na qualidade do produto final, bem como, ocasionará formação de *scales* em tubulações e trocadores de calor. Assim, promove-se a dessilicação aumentando o tempo de residência nos digestores. O produto da dessilicação Bayer (DSP, em inglês *Dessilication Product*) corresponde à fases zeolíticas (predominantemente sodalita) que são insolúveis ao processo e, portanto, são descartadas com o resíduo da bauxita (lama vermelha). Estas fases, contém uma quantidade significativa de soda na estrutura cristalina, a qual foi removida indesejavelmente do licor caustico (Ostap, 1984; McCormick *et al.*, 2002).

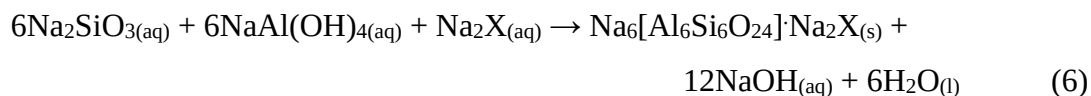
Reação de dissolução da caulinita ($\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$) em licor Bayer:



Para o quartzo (SiO_2):



Reação de dessilicação (precipitação do DSP):



Onde X representa uma variedade de ânions inorgânicos, mais comumente SO_4^{2-} , CO_3^{2-} , Cl^- e OH^- . Conforme essas reações, estima-se que cada tonelada de sílica reativa contida na bauxita que alimenta o processo Bayer, resulte em um consumo de aproximadamente 670 toneladas de soda na forma de DSP (Smith, 2009).

Uma avaliação mais detalhada conduzida por Melo *et al.* (2017) mostrou que além desta diferença entre os minerais correspondentes à sílica reativa, há dentro de um mesmo perfil laterítico-bauxítico, caulinitas com diferentes cristalinidades, cuja reatividade muda significativamente em função da evolução do perfil (Figura 2.3).

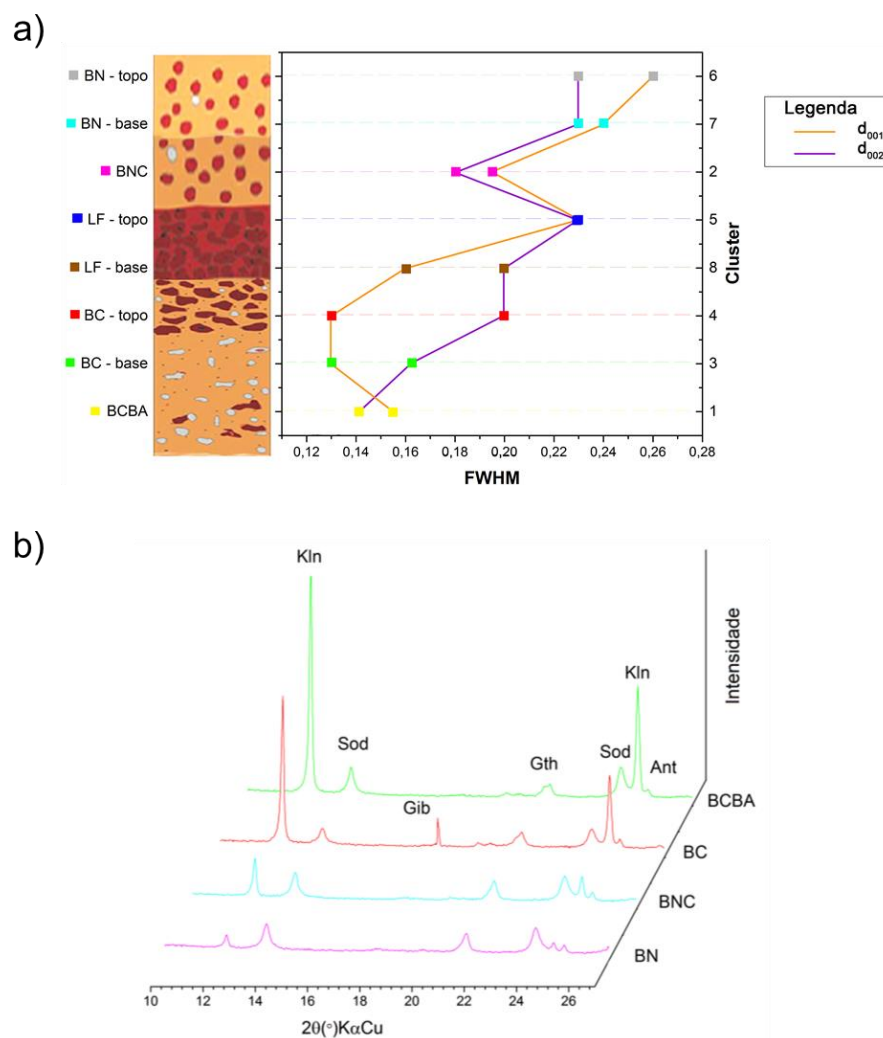


Figura 2.3: a) Comparação da cristalinidade das caulinitas em cada litologia do perfil; b) DRX mostrando maior reatividade das litologias BN e BNC após a digestão em 1 h. Fonte: Melo *et al.* (2017).

As demais fases presentes na bauxita são, em geral, insolúveis ao processo Bayer, sendo descartadas no resíduo da bauxita. No entanto, impactos significativos são observados nas etapas de decantação e lavagem da lama, quando há mudança na mineralogia da bauxita, principalmente devido aos minerais de ferro (quando há predominância de goethita ao invés de hematita). Na prática, o engenheiro de processos tem apenas a informação da composição química de ferro total das bauxitas, porém, a causa do problema de sedimentação da lama só pode ser conhecida após análise mineralógica (muitas vezes não disponível nas refinarias).

Outro exemplo interessante é reportado por Lawson *et al.* (2008). Os autores relatam que no período entre 2003 e 2006 a bauxita que alimenta a refinaria da West Indies Alumina Company (WINDALCO, Kirkvine-Jamaica) passou por uma mudança de composição que resultou em impactos severos a operação da planta. Foi observado neste período um aumento constante do fator de lama, da dificuldade de sedimentação, bem como da recuperação metalúrgica de alumina. Os autores destacaram que, surpreendentemente, não houve mudanças na composição química evidentes por análises de FRX (Al_2O_3 , Fe_2O_3 e LOI praticamente inalterados). A questão foi resolvida ao se observar a variação da composição mineralógica desta bauxita, onde se detectou a presença de goethita aluminosa e crandalita.

2.2. CONTROLE DE QUALIDADE DE BAUXITAS

Desenvolvido por Ostap na década de 80 e tendo sua última atualização em 1989 o método ALCAN é utilizado até os dias atuais em diversas refinarias e minas de bauxita para se realizar o controle de qualidade e de processo, determinando a alumina gibbsítica (Al_2O_3) e a sílica caulínica (SiO_2) em bauxitas de alto grau. Metodologias complementares adequadas a bauxitas de baixo grau, e contendo boehmita também foram desenvolvidas no mesmo período pela ALCAN. Outras empresas, como a ALCOA, também desenvolveram suas próprias metodologias para realizar este controle. No entanto, todos compartilham da mesma premissa e mecanismos, e consistem em apenas adaptações.

Estes métodos tradicionais, baseiam-se na química por via úmida. Assim, o procedimento consiste na digestão da bauxita utilizando licor cáustico sintético (solução NaOH a partir de reagente P.A.) em bombas Parr, simulando as condições de digestão do

processo Bayer. Em seguida, o licor é coletado para determinação do teor de $AvAl_2O_3$ por método titrimétrico e o material sólido é dissolvido em solução ácida concentrada (HCl P.A.), filtrado e, então, submetido à determinação do teor de $RxSiO_2$ por análise espectrofotométrica, gravimétrica, ICP ou AAS.

Abaixo seguem algumas características peculiares deste método, conforme descrições dadas nos próprios procedimentos:

- a) É exclusivamente desenvolvido para determinado tipo de bauxita (por exemplo, bauxitas gibbsíticas). Eventual presença de boehmita e/ou diásporo, bem como, quartzo não são determinados como alumina aproveitável e sílica reativa, respectivamente, e podem acarretar erros. Nos casos em que a presença desses minerais já é conhecida, pode-se usar os métodos complementares adaptados.
- b) Os procedimentos são ajustados para faixas estreitas de determinação de $AvAl_2O_3$ (ex.: entre 41 – 55%) e, para a $RxSiO_2$ é apresentada uma faixa mais ampla (ex.: 0,3 – 12%). Eventuais variações dos teores desses analitos na bauxita podem resultar em menor precisão do método, ou erros de determinação.
- c) Os demais minerais presentes na bauxita não interferem na determinação. No entanto, a reprodutibilidade é significativamente afetada pelas condições de digestão (temperatura do bloco digestor e tempo de digestão) e erros analíticos (referente ao operador e aos procedimentos de análise), por exemplo.
- d) É reportado que o tempo necessário para se determinar oito amostras é de aproximadamente 9 horas utilizando espectrofotometria, e cerca de 12 horas utilizando gravimetria. Isto, assumindo que os reagentes já estejam preparados. Além disso, a metodologia adequada para bauxitas de baixo grau requer um tempo adicional de 4 horas.
- e) A exatidão nas determinações tanto da $AvAl_2O_3$ como da $RxSiO_2$ não pode ser determinadas uma vez que não há um método de referência. Em geral, utiliza-se uma amostra de bauxita certificada como referência da exatidão do procedimento. Vale ressaltar que estas bauxitas certificadas são consumíveis caros.
- f) Em contrapartida, o método é reportado tendo uma excelente precisão. Para $AvAl_2O_3$ a reprodutibilidade é de 0,1% (mesmo analista, mesmo dia), 0,2% (diferente analista, longo período), 0,4% (interlaboratorial). Para $RxSiO_2$ a precisão varia conforme o

método de determinação e a concentração de caulinita na bauxita; variando entre 0,03 – 0,06% para % $RxSiO_2$ contendo 1 – 3%, e entre 0,1 – 0,3% para % $RxSiO_2 > 3\%$ (melhores resultados obtidos por espectrofotometria). Apesar da precisão reportada no descritivo do método, a prática em diversas empresas denota que estes valores são mais altos, sendo aceitáveis na indústria precisão de $< 1\%$ para $AvAl_2O_3$ e $< 0,5\%$ para $RxSiO_2$.

Paz *et al.* (2017) observaram que na Hydro Alunorte o consumo de soda cáustica é reportado como sendo significativamente superior ao esperado baseado nas medidas dos parâmetros de controle de qualidade pelo método tradicional. Curiosamente, os erros estatísticos dessas estimativas tendem a estar dentro do desvio esperado pelos valores certificados. Isto levantou a hipótese de que as condições físico-químicas dos métodos tradicionais não estavam adequadas para a determinação do teor de $RxSiO_2$ nas bauxitas da região de Paragominas-PA. Após um teste de significância estatística comparando os valores estimados pelo método tradicional e o método tradicional modificado, utilizando material de referência certificado, os autores concluíram que o teor de caulinita realmente era subestimado, o que justifica o maior consumo de soda no processo.

Além do tempo de análises ser longo e haver possibilidade de subestimação da caulinita, estes métodos tradicionais demandam grande espaço de laboratório e força de trabalho, uma vez que, é comum analisar cerca de uma centena de amostras por dia. Associado a isto, tem-se os riscos de trabalho envolvendo reagentes perigosos como NaOH e HCl P.A., erros analíticos inerentes às várias etapas do processo e ao operador sujeito à esta rotina, e ao alto custo de operação desse procedimento. A partir de uma estimativa preliminar realizada pelo autor, os custos anuais destas análises, considerando apenas os reagentes utilizados, podem ser superiores a R\$ 200.000,00. Acredita-se que este valor seja de ordem muito maior, se considerados custos com homem-hora, depreciação e manutenção de equipamentos, além de outros insumos não contabilizados na estimativa inicial, como gases e energia elétrica.

Diante das desvantagens na determinação de $AvAl_2O_3$ e $RxSiO_2$ pelos métodos tradicionais de química úmida citados (König *et al.*, 2012; Paz *et al.*, 2018), vários autores têm proposto métodos alternativos para obter estimativas confiáveis o mais prático, rápido e

mais barato possível. Geralmente, os métodos alternativos são baseados em dados de óxidos totais e perda ao fogo (LOI), a partir de balanços de massa estequiométricos e estatística multivariada (Bredell, 1983; Feret *et al.*, 1997; Najjar, 2017; Carneiro *et al.*, 2017; Zhang *et al.*, 2019). A fraqueza de tais métodos se dá pela necessidade de conhecer toda a mineralogia da bauxita, os quais tendem a falhar quando alguma fase não é adequadamente identificada ou quando a característica da bauxita é alterada.

Uma alternativa interessante para o controle de qualidade da bauxita é usar dados de DRX, devido à sua robustez, rapidez de análise, facilidade para ser totalmente automatizado e principalmente porque também permite uma compreensão mineralógica do minério. Estudos recentes apresentam bons avanços na classificação e quantificação mineralógica de bauxitas gibbsíticas com base no refinamento de Rietveld (Paz *et al.*, 2018; Angélica *et al.*, 2018) e estatística multivariada (König *et al.*, 2012; Melo *et al.*, 2017).

3. MATERIAIS E MÉTODOS

3.1. MATERIAIS

As amostras de bauxita utilizadas neste trabalho são de um estudo orientativo realizada pela Mineração Paragominas SA (Norsk Hydro) em 2012. Neste estudo foi realizada uma campanha de sondagem com furos dispostos em uma malha regular de 25 m x 25 m no platô Miltonia 3 (Figura 3.1), totalizando 174 furos e 1395 amostras correspondentes às litologias típicas do perfil laterítico de Paragominas (Picanço *et al.*, 2013).

Foi realizada, inicialmente, uma triagem a partir dos dados químicos de referência ($AvAl_2O_3$ e $RxSiO_2$) das amostras que compunham este banco de amostras da campanha de sondagem (vide seção 6.2). O objetivo dessa triagem foi selecionar para o presente estudo apenas amostra que fossem representativas do perfil bauxítico. Assim, do total de 1395 amostras foram preliminarmente selecionadas 142, correspondentes às litologias Bauxita Nodular (BN), Bauxita Nodular-Cristalizada (BNC), Bauxita Cristalizada (BC) e Bauxita Cristalizada-Amorfa (BCBA). Tais terminologias são amplamente utilizada na indústria e estão relacionadas aos aspectos geológicos do perfil laterítico de bauxita nessa região (Figura 3.3). Entre essas litologias, BC representa o horizonte principal do minério lavrado. As litologias LF e BA não foram consideradas no presente estudo, uma vez que, em nenhum aspecto correspondem à horizontes de minério – embora LF possa eventualmente estar presente no minério lavrado dependendo da estratégia da mina frente as discontinuidades do perfil.

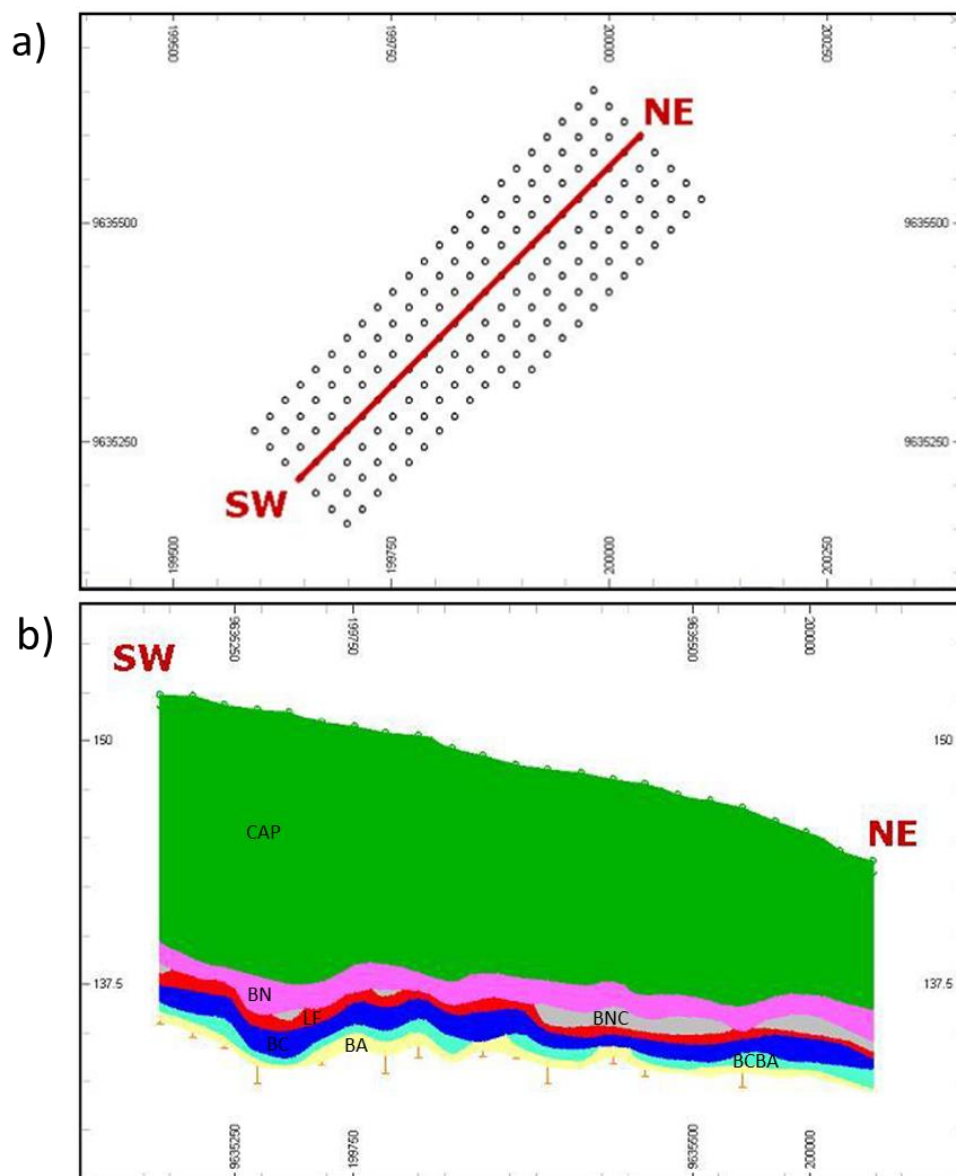


Figura 3.1: a) Esquemático da malha de sondagem 25m x 25m; b) Perfil da seção destacada na malha de sondagem, Platô Miltonia 3. Adaptado de Picanço *et al.* (2013).

Litologia		Nº amostras no banco de dados	Nº amostras selecionadas no trabalho	
	CAP	Argila	-	-
	BN	Bauxita Nodular	360	35
	BNC	Bauxita Nodular-Cristalizada	145	16
	LF	Laterita Ferruginosa	197	-
	BC	Bauxita Cristalizada	329	49
	BCBA	Bauxita Cristalizada/Amorfa	248	42
	BA	Bauxita Amorfa (Argila Mosqueada)	116	-
TOTAL		1395	142	

Figura 3.2: Nº de amostras obtidas na campanha de sondagem e de amostras selecionadas no presente trabalho. Detalhamento por cada litologia. Perfil adaptado de Pereira (2015).

3.2. MÉTODOS E EQUIPAMENTOS

3.2.1. Preparação das amostras

As amostras obtidas na campanha de sondagem foram preparadas pela equipe técnica da empresa MPSA/Hydro segundo os procedimentos de rotina da mina. O fluxograma da Figura 3.3 apresenta o procedimento realizado para a obtenção das amostras.

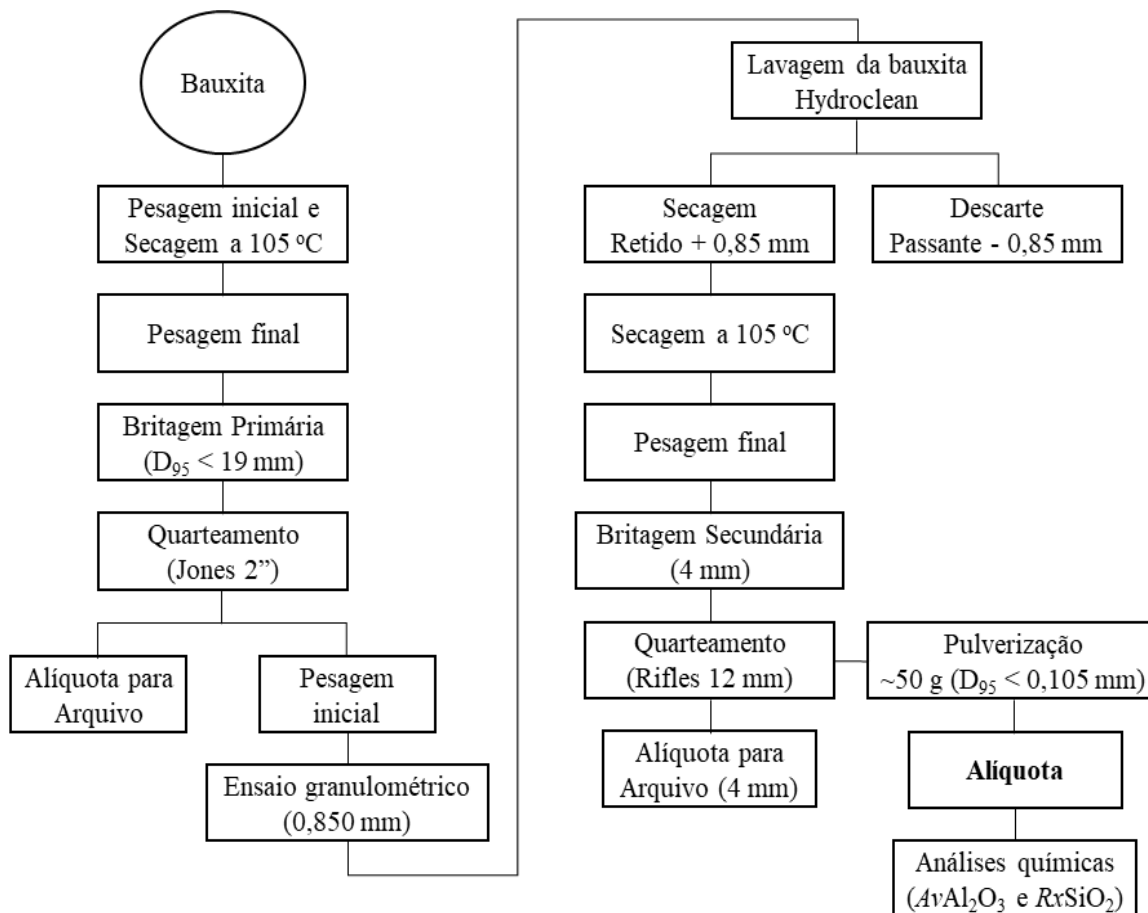


Figura 3.3: Fluxograma da preparação física das amostras feitas pela empresa Hydro. Adaptado de Oliveira (2017).

3.2.2. Preparação das espécimes – DRX por reflexão e transmissão

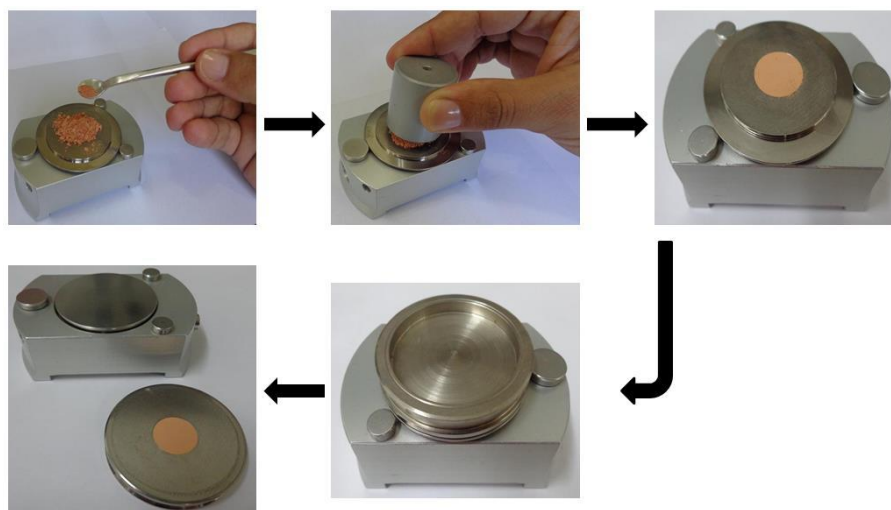
Observa-se que o material ao final da etapa de preparação apresenta distribuição granulométrica com $D_{90} < 0,105$ mm (passante em 140#). Esse tamanho de partícula já é adequado para as análises químicas de rotina por via úmida ($AvAl_2O_3$ e $RxSiO_2$), conforme procedimento descrito na seção 3.2.2. Para análises mineralógicas, na grande maioria dos casos, é recomendado pulverizar a amostra até obter distribuição de tamanho de partícula $< 20 \mu m$, devido principalmente aos efeitos relacionados à não-homogeneidade das partículas e estatística de contagem das partículas (Jenkins & Snyder, 1996). No entanto, com o objetivo de simplificar o procedimento para futura utilização industrial do método proposto neste trabalho, as análises de DRX por reflexão foram realizadas tal qual as amostras foram

fornecidas pela empresa. Já para as análises pelo modo transmissão as amostras foram maceradas em grau de ágata até a obtenção de um pó fino homogêneo. Em aspectos gerais, esse procedimento de maceração / homogeneização permite uma melhor estatística de partículas durante a coleta do difratograma, mas pode acarretar em alterações indesejadas na amostra, tais como, amorfização, decomposição, tensões residuais e contaminação do material. Nesse contexto, Jenkins & Snyder (1996) salienta como regra geral que “os melhores métodos de preparação de espécimes são aqueles que permitem o analista obter a informação desejada com a menor quantidade de tratamento da amostra”.

Maior cuidado, no entanto, foi tomado na preparação das espécimes para análises de DRX por reflexão, quanto à forma de preenchimento do porta-amostra. Utilizou-se preenchimento *backloading* (Figura 3.4a), com o objetivo de reduzir o efeito de orientação preferencial durante a aquisição dos dados difratométricos.

O problema gerado pelo efeito de orientação preferencial pode ser fortemente evitado quando a análise difratométrica é realizada pelo modo transmissão. Nesta geometria de coleta dos difratogramas, há um aumento da intensidade das reflexões ($hk0$) em detrimento aos planos basais ($00l$) de minerais lamelares como a caulinita e a gibbsita que tendem a se orientar durante a montagem do porta-amostra (Ermrich & Opper, 2013). Para estas análises é utilizado um porta-amostra específico (Figura 3.4b) no qual, uma pequena quantidade de amostra é posicionada entre duas folhas finas de um substrato orgânico (Mylar®, Figura 3.4b).

a)



b)

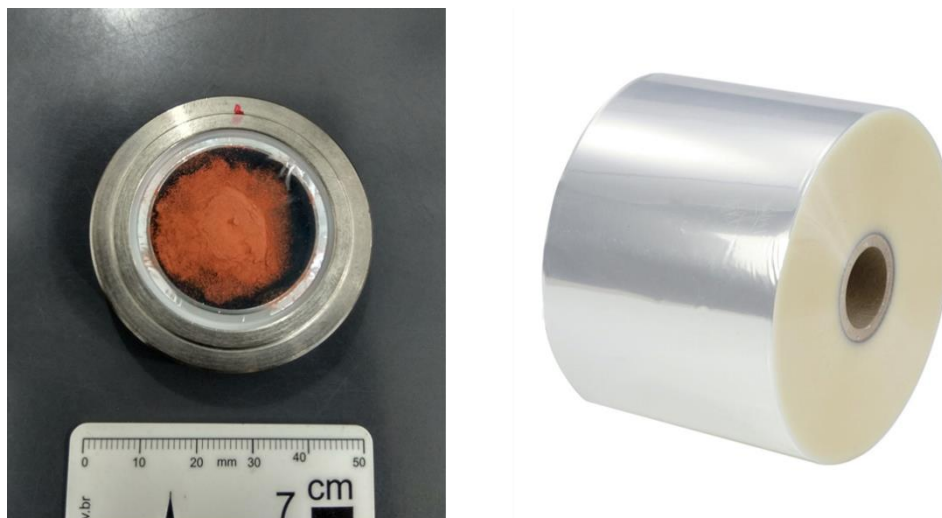


Figura 3.4: a) Procedimento de preenchimento dos porta amostras para DRX pelo método *backloading* para coleta de dados por reflexão. Fonte: Oliveira (2017); e b) porta-amostra (esquerda) e filme Mylar® (direita) utilizados para coleta de dados por transmissão.

3.2.3. Conjunto de dados

A técnica analítica fundamental utilizada para o desenvolvimento deste trabalho é a difratometria de raios-X. Os dados de difração (difratogramas) das amostras de bauxita foram inicialmente utilizados para identificação das fases e para a predição dos parâmetros $AvAl_2O_3$ e $RxSiO_2$. Neste último caso, como demonstrado na Figura 3.5, cada difratograma pode ser disposto em uma matriz $m \times n$, onde m corresponde ao número de amostras analisadas, e n corresponde aos valores de intensidade (contagem de fótons) para cada ponto (passo $^\circ 2\theta$ do goniômetro) – onde n depende da faixa de medida utilizada e tamanho do passo do goniômetro (neste trabalho estes dois parâmetros variaram conforme apresentados na Tabela 3.1 e explicados na seção 3.2.4).

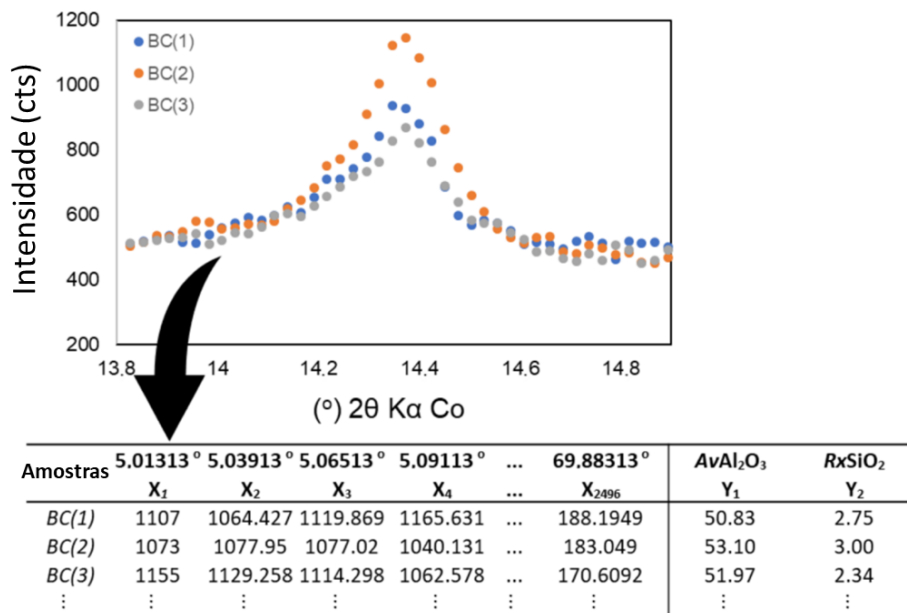


Figura 3.5: Representação do dataset de amostras. Difratomogramas arranjados como matriz.

As etapas seguidas na execução deste trabalho podem ser visualizadas na Figura 3.6. O procedimento basicamente consistiu em:

- obter as amostras de bauxita preparadas fisicamente e analisadas para $AvAl_2O_3$ e $RxSiO_2$;
- realizar triagem e seleção das amostras representativas;
- leitura de todas as amostras no DRX por reflexão e por transmissão;
- coleta e tratamento dos dados de DRX no software HighScore Plus 4.8; análise de PCA – classificação das amostras em *clusters*;
- otimização por DOE dos parâmetros utilizando apenas amostras de BC;
- calibração e teste dos modelos PLSR para cada *cluster* de amostras e utilizando as condições otimizadas; e
- análise dos resultados.

O conjunto de 142 amostras selecionadas foi subdividido entre amostras de calibração e de teste dos modelos PLSR. Assim, 120 amostras (~85% do dataset) foram selecionadas aleatoriamente para compor o conjunto de calibração e 22 amostras (~15% do dataset) o

conjunto de teste, respectivamente. Adicionalmente, algumas amostras do conjunto de testes foram analisadas em replicatas para avaliar a precisão dos modelos.

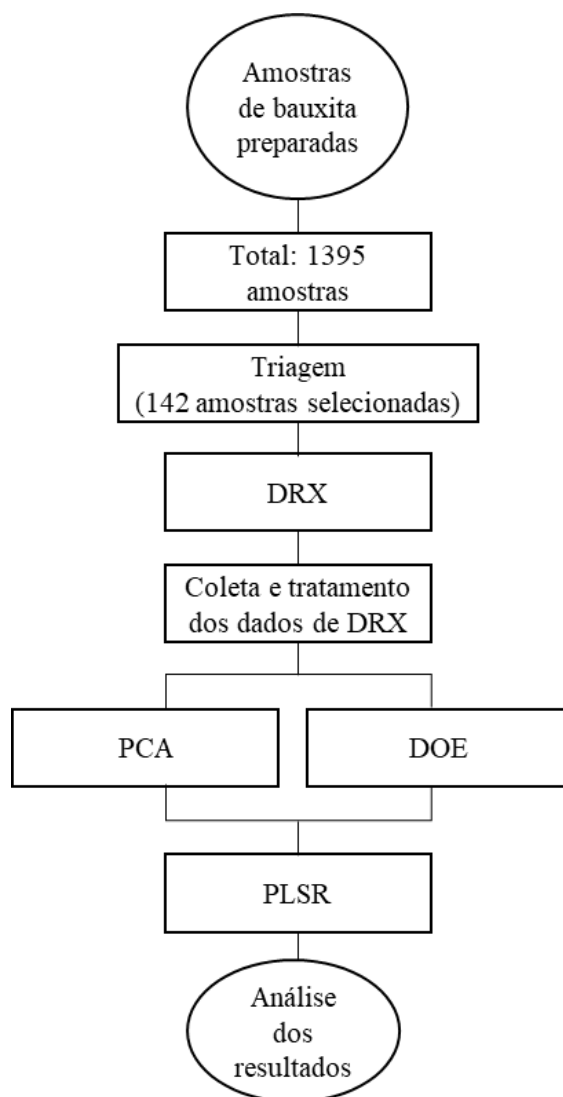


Figura 3.6: Fluxograma das etapas do presente trabalho.

3.2.1. Análises de $AvAl_2O_3$ e $RxSiO_2$ por via úmida

As análises para determinação dos parâmetros $AvAl_2O_3$ e $RxSiO_2$ foram realizadas no laboratório da MPSA/Hydro, em Paragominas, conforme método de rotina. Abaixo é descrito de forma sucinta as etapas de execução do método Alcan 1243-89, adotado na Hydro como metodologia padrão para a quantificação de Alumina Aproveitável e Sílica Reativa.

- a) secar amostra a 105°C por 2 h;
- b) pesar 3,25 g de material seco (duplicata);
- c) a massa é colocada na bomba parr;
- d) adiciona-se a solução de NaOH 10%;
- e) homogeneizar manualmente;
- f) colocar no bloco digestor para reação a 150 °C por 25 mim.
- g) tirar as bombas parr do bloco digestor e resfriar logo em seguida com água corrente (direto na pia-torneira com alto fluxo de água). Em média 10 mim.
- h) separar um balão de 500 mL, colocar 50 mL de NaCl 20% e, em seguida, colocar todo conteúdo do digestor no balão e aferi-lo até 500 ml com água deionizada, homogeneizar;
- i) filtrar em um Becker de 100 mL, ~ 50 mL do material do balão (filtragem lenta);
- j) o filtrado (solução incolor) será usado para analisar a $AvAl_2O_3$ por método titrimétrico com CDTA;
- k) recolocar o papel filtro dentro do balão de 500 ml e adicionar 30 mL de HCl concentrado, homogeneizar, esperar por 30 mim;
- l) acrescentar o floculante (10 mL de MAGNAFLUC 0,08%), esperar por 10 mim;
- m) filtrar;
- m) a partir da solução determinar a concentração de $RxSiO_2$ por absorção atômica ou ICP.

3.2.2. Difração de raios-X (DRX)

A difratometria de raios-X é um método analítico versátil e não destrutivo amplamente utilizado em diversos campos de pesquisa, desde a ciência dos materiais e geologia até farmacêutica, e representa o principal método para identificação de fases cristalinas em amostras de pó naturais e sintéticas. Apesar de ser utilizada majoritariamente para tal aplicação, a difratometria de raios-X compreende uma valiosa fonte de informações do material analisado. Intrinsecamente, é possível obter a partir dos padrões difratométricos:

- quantificação das fases presentes e estimativa de posições atômicas a partir da distribuição das intensidades dos picos de difração;

- determinação de parâmetros relacionados a estrutura cristalina (sistemas cristalinos, parâmetros de rede, etc.) a partir das posições dos picos de difração; e
- parâmetros relacionados ao comportamento microestrutural do material (tamanho de cristalitos, defeitos e deformação da rede cristalina) a partir da forma e largura dos picos.

Tais informações podem ser extraídas dos difratogramas a partir de métodos não triviais de refinamento das estruturas cristalinas e ajuste matemático do padrão difratométrico (Bish & Post, 1993). Neste trabalho, o foco foi direcionado na quantificação indireta dos minerais presentes na bauxita – gibbsita (relacionada à Al_2O_3) e caulinita (relacionada à SiO_2) utilizando modelagem estatística. Nesse contexto, decidiu-se investigar dois modos de coleta dos dados difratométricos: *reflexão* – os raios-X são espalhados pela superfície plana da amostra, e *transmissão* – os raios-X passam através das amostras relativamente transparentes a radiação.

O modo reflexão é o mais comumente utilizado e adequado para a maioria das aplicações de análises de DRX pelo método do pó. O modo de coleta por transmissão é uma alternativa ao modo reflexão, principalmente naqueles casos onde é necessária uma boa resolução dos picos em baixo ângulo, deseja-se evitar o efeito de orientação preferencial, e também quando a quantidade de amostra disponível é pequena para o preenchimento do porta-amostra. Especificamente neste trabalho, buscou-se avaliar os dois modos de coleta de dados de difração de raios-X no sentido de se mitigar o efeito de orientação preferencial e obter uma melhor fonte de dados para a modelagem estatística.

Os difratogramas foram coletados utilizando um difratômetro PANalytical Empyrean, equipado com um goniômetro θ - θ , tubo de raios-X de Cobalto ($K\alpha_1 = 1,78901 \text{ \AA}$), filtro Fe $K\beta$ e detector de área PIXel3D 2x2 (modo de varredura linear) com comprimento ativo de $3,3473^\circ 2\theta$ (255 canais), nas geometrias de reflexão e transmissão. As seguintes condições foram utilizadas na coleta dos dados difratométricos: 40 kV e 35 mA; tamanho de passo de $0,026^\circ 2\theta$ e faixa de leitura de $5 - 70^\circ 2\theta$; fenda *soller* de 0,04 rad; fendas fixas divergentes de $\frac{1}{4}^\circ$ e anti-espalhamento $\frac{1}{2}^\circ$, na coleta de dados por reflexão e $\frac{1}{8}^\circ$ para ambas as fendas na coleta por transmissão. Os difratogramas foram avaliados utilizando o software PANalytical HighScore Plus 4.8.

3.2.3. Análise de Componentes Principais (PCA)

A análise de componentes principais (*Principal Component Analysis* – PCA) é uma técnica multivariada amplamente usada como ferramenta em vários campos de aplicação – desde neurociências até computação gráfica e aprendizado de máquina. Basicamente, ela visa reduzir a dimensionalidade dos dados, de forma que as informações relevantes possam ser extraídas de conjuntos de dados extensos e confusos (Joliffe, 2002; Shlens, 2005). Uma descrição matemática abrangente, bem como exemplos práticos de aplicação de PCA em quimiometria podem ser encontrados em Cordella (2012) e Bro & Smilde (2014). A utilização de PCA em dados de DRX para identificar grupos de amostras e correlacioná-los com informações mineralógicas tem sido relatada nos últimos anos (Macchiarola *et al.*, 2007; König *et al.*, 2012; Gobbo *et al.*, 2013; König *et al.*, 2014; Melo *et al.*, 2017). Neste trabalho, o PCA foi aplicado como uma ferramenta de mineração de dados, a fim de encontrar grupos de amostras similares, identificar valores extremos (*outliers*) e o range 2 θ do difratograma que melhor descreve a variação entre as amostras de bauxita para ser usada nas análises subsequentes por PLSR (*Partial Least Square Regression*).

As análises de PCA foram feitas utilizando os valores químicos de referência ($AvAl_2O_3$ e $RxSiO_2$) e padrões de DRX das amostras. No primeiro caso, o conjunto de dados foi padronizado (valores de observações divididos pelos respectivos desvios-padrões), uma vez que as variáveis estão em diferentes intervalos ($AvAl_2O_3 = 31,4 - 59,3\%$ e $RxSiO_2 = 0,52 - 8,42\%$). Geladi e Kowalsky (1986) argumentam que dados de FTIR, por exemplo, não precisam ser padronizados, uma vez que todas as observações já estão na mesma escala. Baseado nisto, consideração semelhante foi aceita para os dados de DRX, os quais foram apenas centralizados na média para evitar principalmente erros relacionados a linha base do *background* dos difratogramas. A Figura 3.6 possibilita uma boa compreensão dos métodos de pré-processamento dos dados adotados neste trabalho.

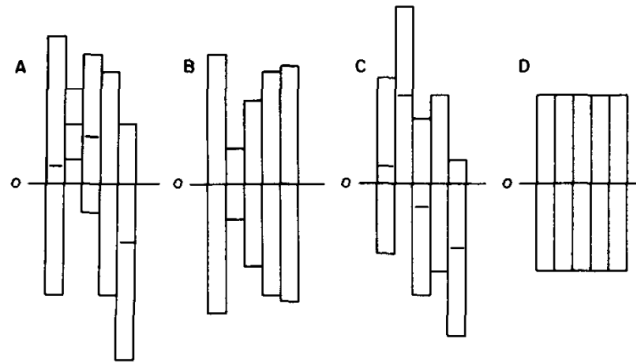


Figura 3.7: Pré-processamento dos dados. Os dados de cada variável são apresentados pela barra de variação e seus valores centrais (médias). A) exemplo comum de dados brutos, B) dados centrados na média, C) dados padronizados pelo desvio-padrão, e D) dados centrados na média e padronizados – também conhecido como *autoescalamento*. Fonte: Geladi e Kowalsky (1986).

3.2.4. Regressão por Mínimos Quadrados Parciais (PLSR)

A regressão por mínimos quadrados parciais (*Partial Least Square Regression* – PLSR, também conhecida como *Projection on Latent Structures*), é uma técnica multivariada robusta usada para prever informações ocultas de dados brutos (variáveis latentes), sendo uma boa alternativa para a regressão linear múltipla (*Multiple Linear Regression* – MLR) e a regressão por componentes principais (*Principal Component Regression* – PCR). O PLSR é amplamente utilizado em quimiometria, uma vez que é adequado para lidar com dados multicolineares (observações altamente correlacionadas) – como dados de espectroscopia no infravermelho, Raman e medidas de difração de raios-X (Wold *et al.*, 1984, Wold *et al.*, 2001).

Matematicamente, a análise consiste em inicialmente modelar por componentes principais as matrizes de dados **X** (variáveis independentes – preditores) e **Y** (variáveis dependentes – respostas).

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E} \quad (7)$$

$$\mathbf{Y} = \mathbf{UQ}' + \mathbf{F} \quad (8)$$

Em que **T** e **U** são escores, **P** e **Q** são “cargas” (*loadings*), e **E** e **F** são matrizes de resíduos). A particularidade do PLSR é que, simultaneamente, as matrizes de escores são relacionadas:

$$\hat{u} = b.t \quad (9)$$

Onde **b** são os coeficientes de regressão (assim como na MLR) determinados por mínimos quadrados. A relação resultante para **Y**, portanto, apresenta os coeficientes de regressão que maximizam a covariância com os escores da matriz preditora **X**:

$$\mathbf{Y} = \mathbf{TBQ}' + \mathbf{F} \quad (10)$$

De forma que **F** deve ser minimizado (Wold, 1966; Geladi & Kowalsky, 1986).

Em aplicações práticas, a propriedade de interesse contida em **Y** geralmente é obtida através de métodos ou medidas caras, difíceis e / ou demoradas, enquanto **X**, geralmente é uma medida barata, fácil de se obter e / ou rápida. Este método estatístico, portanto, permite facilmente prever a propriedade desejada (**Y**) através dos dados da matriz **X**. No presente trabalho, as propriedades de interesse contidas em **Y** são: alumina aproveitável ($AvAl_2O_3$) e sílica reativa ($RxSiO_2$), enquanto os preditores **X** são os difratogramas de raios-X.

O método de validação cruzada (“*leave-one-out*”) foi usada para encontrar os melhores números de fatores nos modelos e as estatísticas de resíduos (raiz do erro quadrático médio de validação cruzada (em inglês *Root Mean Square Error of Cross-Validation* – RMSECV), e raiz do erro quadrático médio da predição (*Root Mean Square Error of Prediction* – RMSEP) foram monitoradas para avaliar a qualidade da predição. Estes estimadores de erro médio expressam o quão acurado está a modelagem em relação aos valores de referência e são calculados através da equação:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{pred} - y_{ref})^2}{n}} \quad (11)$$

Onde y_{pred} e y_{ref} são os valores do parâmetro de interesse modelado e de referência, respectivamente, e n representa o número de amostras. A diferença, portanto, entre o RMSECV e RMSEP está no dataset utilizado – no primeiro caso, corresponde ao próprio dataset de calibração utilizando o método de validação cruzada; no segundo caso, utiliza-se um dataset com novas amostras as quais não foram utilizadas para estimar os parâmetros do modelo. Por não conter o viés de sobreajuste na estimativa do modelo, o RMSEP é um estimador mais realista do erro de predição em situações reais de aplicação do modelo.

Os melhores modelos de PLSR para determinar os parâmetros $AvAl_2O_3$ e $RxSiO_2$ foram encontrados após a definição dos parâmetros de pré-processamento otimizado dos dados de DRX e da definição dos grupos de amostras por PCA.

3.2.5. Planejamento Fatorial de Experimentos (DOE)

A fim de otimizar a qualidade e o tempo de coleta dos dados de DRX a serem utilizados no PLSR, alguns parâmetros foram avaliados usando um planejamento experimental fatorial ²⁴ (*Design of Experiments* – DOE). Nas Tabelas 3.1 e 3.2 estão apresentados os fatores e níveis adotados no planejamento fatorial e a matriz de experimento em ordem de planejamento.

Para este procedimento foram usadas 42 amostras da litologia BC (Bauxita Cristalizada, o qual representa o horizonte do minério) para compor o conjunto de amostras de calibração, e 4 amostras (com 5 repetições) para o conjunto de teste. Todas essas amostras foram extraídas do grupo total de amostras selecionadas para o desenvolvimento deste trabalho (142 amostras no total) descritas detalhadamente na seção 3.1 deste documento (vide Figura 3.2). Este estudo foi realizado utilizando apenas os difratogramas obtidos por reflexão.

A sequência de execução deste DOE compreende:

- a) Criação dos modelos PLSR para os parâmetros $AvAl_2O_3$ e $RxSiO_2$ utilizando o conjunto de amostras de calibração;
- b) Calibração dos modelos utilizando o método de validação cruzada “*leave-one-out*” e acompanhamento do RMSECV como critério para selecionar o número de fatores dos modelos;

- c) Determinação dos parâmetros $AvAl_2O_3$ e $RxSiO_2$ das amostras do conjunto de teste;
- d) Acompanhamento dos valores de RMSEP para determinar a qualidade dos modelos.

Assim, ao variar as condições experimentais do DOE os melhores modelos correspondem justamente àqueles cujo RMSEP é menor (menor erro entre os valores preditos e de referência) e, portanto, os valores de RMSEP foram utilizados como variável de resposta do DOE.

Tabela 3.1: Fatores e níveis do DOE 2⁴.

Fatores	Níveis	
	-1	+1
Tipo de fenda divergente ^a	FDS	ADS
Tamanho de passo (°)	0,026	0,065
Faixa de leitura DRX (° 2θ)	5-70	13-34
Modo de pré-processamento dos dados ^b	Centrado na média	Autoescalamento

^a Fenda fixa (*Fixed Divergent Slit* – FDS) e fenda automática (*Automatic Divergent Slit* – ADS);

^b O conjunto de dados é padronizado (dividindo cada variável pelo respectivo desvio-padrão) e centrado na média.

Tabela 3.2: Matriz de experimento do DOE 2⁴ com níveis codificados.

Tipo de fenda divergente	Tamanho de passo (°)	Faixa de leitura DRX (° 2θ)	Modo de pré-processamento dos dados
-1	-1	-1	-1
+1	-1	-1	-1
-1	+1	-1	-1
+1	+1	-1	-1
-1	-1	+1	-1
+1	-1	+1	-1
-1	+1	+1	-1
+1	+1	+1	-1
-1	-1	-1	+1
+1	-1	-1	+1
-1	+1	-1	+1
+1	+1	-1	+1
-1	-1	+1	+1
+1	-1	+1	+1
-1	+1	+1	+1
+1	+1	+1	+1

4. RESULTADOS E DISCUSSÕES

4.1. CARACTERIZAÇÃO DAS BAUXITAS

Os padrões difratométricos das amostras de cada litologia bauxítica é mostrado na Figura 4.3. Observa-se as mesmas fases mineralógicas presentes em todas as litologias de bauxita: caulinita ($d_{001} = 7,14 \text{ \AA}$ e $d_{002} = 3,56 \text{ \AA}$), gibbsita ($d_{002} = 4,84 \text{ \AA}$ e $d_{110} = 4,37 \text{ \AA}$), hematita ($d_{104} = 2,69 \text{ \AA}$), Al-goethita (pico d_{101} variando de $4,12 \text{ \AA}$ a $4,16 \text{ \AA}$ devido à substituição isomórfica de alumínio) (Neumann *et al.*, 2014) e anatásio ($d_{011} = 3,52 \text{ \AA}$).

Como observado, gibbsita é fase predominante em todas as amostras. Dependendo da sua concentração e do grau de cristalinidade, o pico principal (plano basal d_{002}) pode apresentar intensidade elevada devido ao efeito de orientação preferencial (Angélica *et al.*, 2018). Devido a isso, observou-se em várias amostras o pico referente ao $k\beta$ desse plano orientado. Situação semelhante pode eventualmente ocorrer com as caulinitas presentes, principalmente nas amostras da litologia BCBA, onde a concentração desta fase pode ser alta e contém cristais pouco alterados e mais ordenados (Melo *et al.*, 2017).

Pela região destacada na Figura 4.1, pode-se observar que há um *offset* da linha base nos difratogramas de algumas amostras, provavelmente devido ao efeito da fluorescência. Este fato ressalta a necessidade de pré-processar os dados (centralizar na média) para utilização destes nas análises de PCA e PLSR (Bro & Smilde, 2014).

A Figura 4.2 apresenta os difratogramas de todas as amostras desde trabalho obtidas pelo modo transmissão. Como já esperado, são observadas as mesmas fases mineralógicas em ambas as geometrias de análise. No entanto, é possível observar uma significativa redução da orientação preferencial da gibbsita ($d_{002} = 4,84 \text{ \AA}$) e comitantemente um aumento da intensidade relativa e resolução dos picos referentes as demais reflexões ($hk0$) das fases presentes. Interessante notar que devido a isto, as diferenças das intensidades entre as fases foram significativamente reduzidas, ou seja, os valores absolutos das observações em cada ponto 2θ são mais próximos e, com isso, estes apresentam também um desvio menor.

Por causa da pequena abertura das fendas utilizadas neste modo de coleta, todos os difratogramas apresentaram uma redução de aproximadamente 6x nas intensidades em relação às obtidas por reflexão. Adicionalmente, pode-se observar que em baixo ângulo

houve um aumento da radiação de fundo, embora, este modo de coleta permitisse uma melhor resolução de possíveis picos nesta região do difratograma.

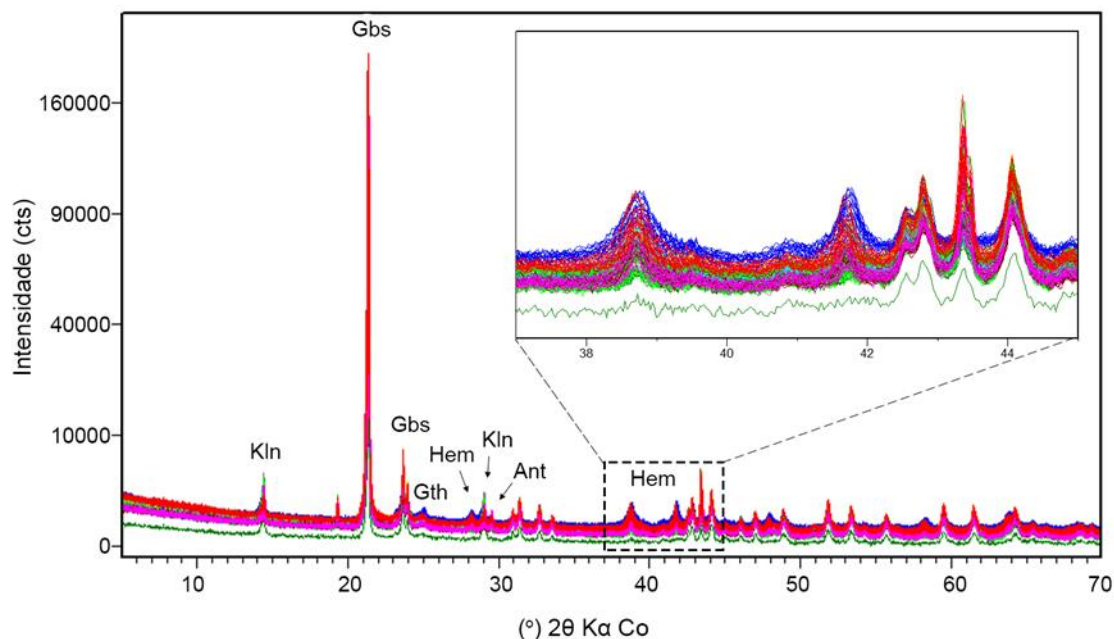


Figura 4.1: Padrões difratométricos de todas as amostras de bauxita utilizadas neste estudo obtidas *por reflexão*. Destaque para o *offset* do background entre as amostras. *Legenda:* Kln: caulinita; Gbs: gibbsita; Gth: Al-goethita; Ant: anatásio; Hem: hematita.

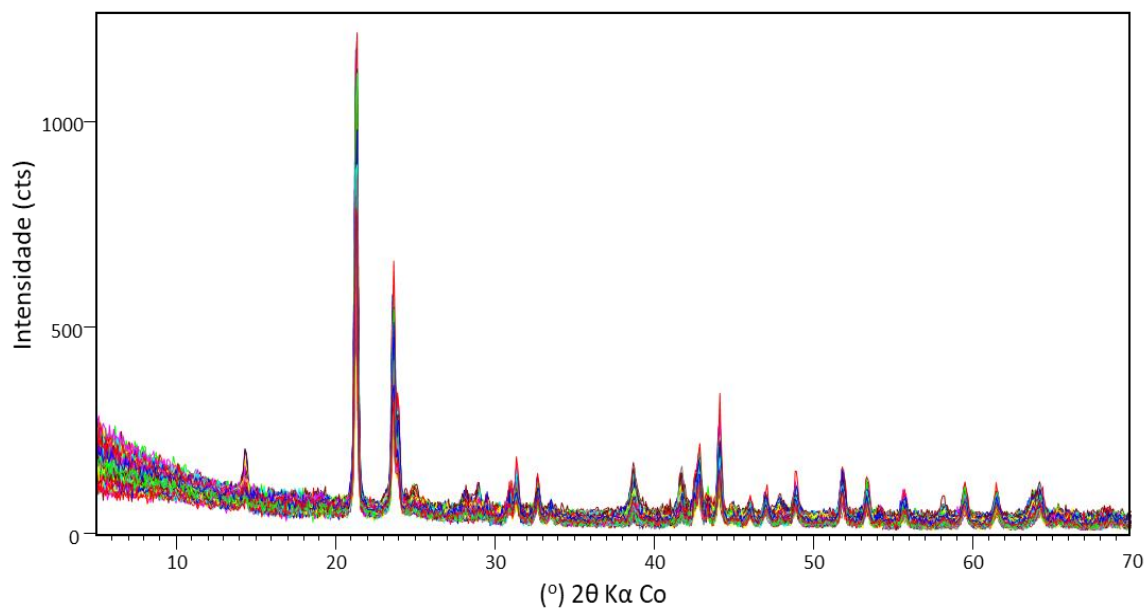


Figura 4.2: Padrões difratométricos de todas as amostras de bauxita utilizadas neste estudo obtidas *por transmissão*.

4.2. AVALIAÇÃO PRELIMINAR DOS DADOS QUÍMICOS DAS BAUXITAS

Do banco de dados de amostras de bauxita provenientes da campanha de sondagem em Miltônia 3 era necessário que todas as amostras selecionadas para o estudo de PLSR (analisadas por DRX para compor os conjuntos de amostras de calibração e de teste) fossem representativas do perfil laterítico da área de estudo. Dessa forma, foi realizado uma triagem a partir dos dados químicos de referência fornecidos.

As amostras foram avaliadas de acordo com os respectivos teores de $AvAl_2O_3$, $RxSiO_2$, bem como pela razão em peso $RxSiO_2/AvAl_2O_3$, baseado na correlação desses analitos em cada litologia. A Figura 4.3 apresenta os histogramas de frequência das amostras. Para $AvAl_2O_3$ (Figura 4.3a) pode-se observar que existem dois grupos de amostras distribuídas uniformemente com médias $\sim 39\%$ e $\sim 52\%$. O mesmo comportamento não é observado para $RxSiO_2$ (Figura 4.3b), na qual sua concentração aparentemente apresenta uma distribuição uniforme entre 2% e 8%. No entanto, a partir de uma avaliação mais detalhada por cada litologia se observa três grupos de amostras com médias em 2,7%, 4,7% e 7%.

Devido a diferença do comportamento químicos das amostras a seleção adequada destas não foi possível inicialmente. Verificou-se a necessidade de se correlacionar estes dois parâmetros a partir da relação arbitrária $RxSiO_2/AvAl_2O_3$ (Figura 4.3c). Com isso observou-se que a maioria das amostras apresentam razão entre 0,033 – 0,133, havendo ainda um grupo significativo de amostras na faixa de 0,134 – 0,293.

Como critério de seleção das amostras que comporiam o *dataset* para análise de PLSR, utilizou-se, portanto, o intervalo de 1 desvio-padrão dos valores médios das duas faixas de razão $RxSiO_2/AvAl_2O_3$ (Tabela 4.1). Assim, foram selecionadas aleatoriamente amostras de cada litologia que se encontram nessas faixas, admitindo que estas são representativas do perfil. Vale ressaltar que esta triagem se baseou exclusivamente nos dados químicos, podendo haver variações mineralógicas nessas amostras que só poderiam ser observadas após as análises de difração de raios-X. A correlação química utilizada como critério de seleção das amostras pode ser mais claramente observada no gráfico de dispersão apresentado na Figura 4.4.

Ressalta-se que dentre as 142 amostras selecionadas para este estudo, todas as amostras de BC, BCBA e BNC foram extraídas da Faixa 1, enquanto todas de BN foram

extraídas da Faixa 2. Todas as amostras de BCBA e BNC que foram discriminadas na Faixa 2 foram consideradas *outliers*, da mesma forma para todas as amostras de BN discriminadas na Faixa 1, e todas as amostras que apresentavam composição química fora das faixas 1 e 2 (> 1 desvio-padrão).

Tabela 4.1: Resumo do critério de seleção de amostras para os modelos PLSR e n° de amostras em cada faixa de seleção.

	Faixa 1	Faixa 2
Média $\pm 1\sigma$	$0,077 \pm 0,028$	$0,184 \pm 0,033$
Intervalo	$0,049 - 0,105$	$0,150 - 0,218$
BC	185	0
BCBA	139	42
BNC	120	1
BN	6	311
Total de amostras	450	354

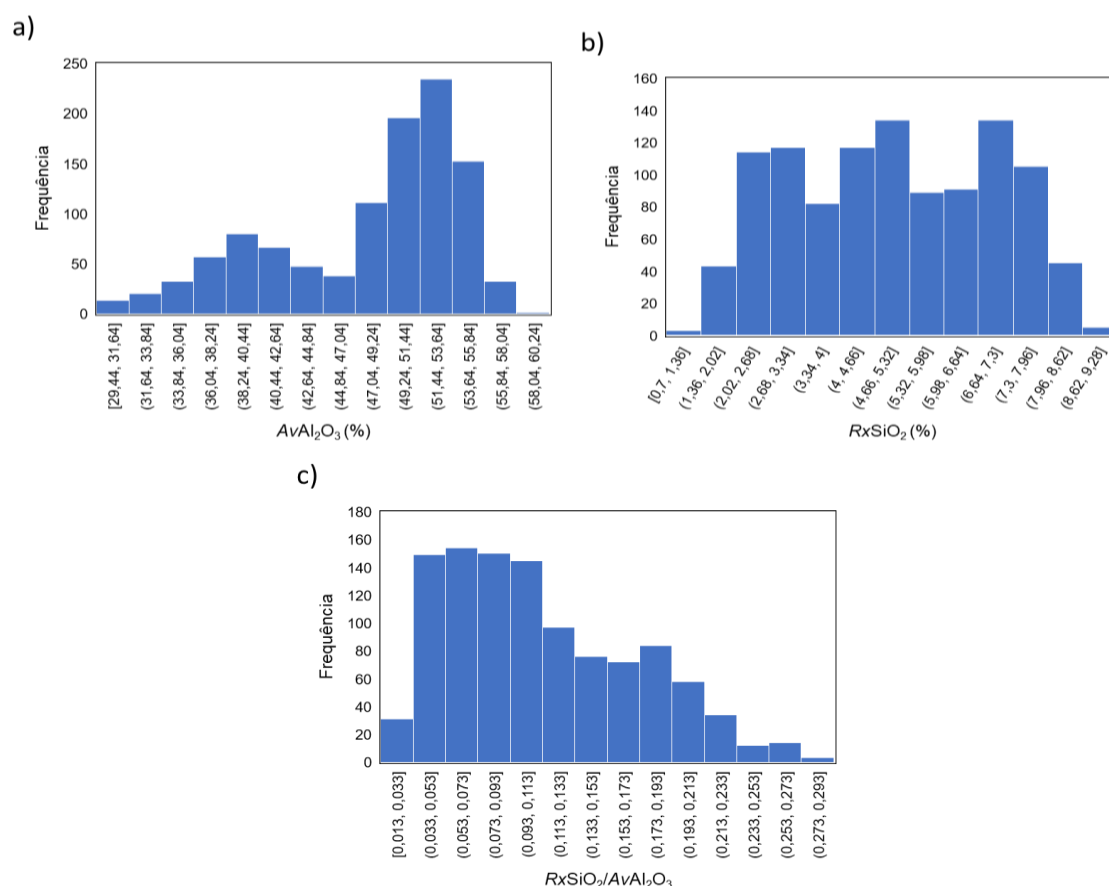


Figura 4.3: Histograma dos dados químicos das amostras de bauxita, a) $AvAl_2O_3$; b) $RxSiO_2$; e c) $RxSiO_2/AvAl_2O_3$.

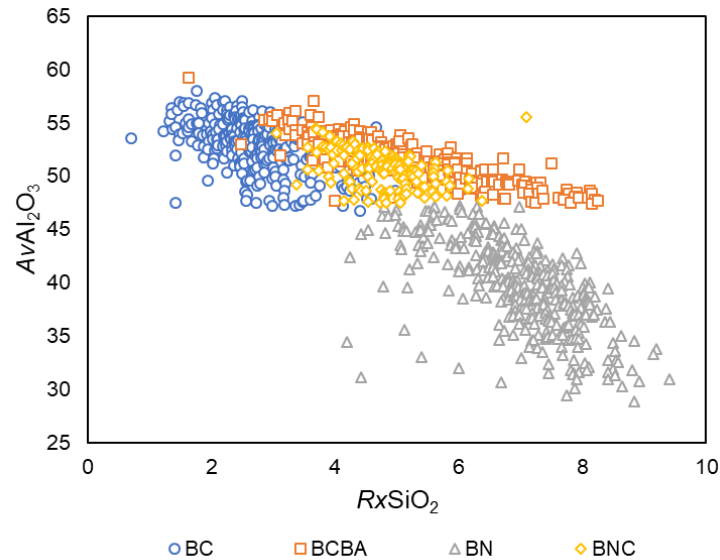


Figura 4.4: Gráfico de dispersão dos dados químicos de referência.

4.3. AVALIAÇÃO DOS DADOS POR PCA

4.3.1. Dados químicos $AvAl_2O_3$ e $RxSiO_2$

A análise de componentes principais foi realizada no conjunto de calibração (120 amostras) utilizando os valores químicos de referência e os padrões difratométricos. Os escores e cargas do PCA referente aos valores químicos são mostradas na Figura 4.5. O primeiro componente (PC1) explica 96% da variação nos dados e está negativamente relacionado ao teor de $AvAl_2O_3$ e positivamente relacionado ao $RxSiO_2$.

Comparando os dois gráficos, nota-se que entre as litologias, BN apresenta a maior dissimilaridade química – mais rica em $RxSiO_2$ e mais pobre $AvAl_2O_3$. Industrialmente adota-se a razão $AvAl_2O_3/RxSiO_2$ como parâmetro para classificar bauxitas de alto (> 10) e baixo grau (< 8). Devido à baixa razão $AvAl_2O_3/RxSiO_2$ em BN (em geral, < 7) esse horizonte superior é considerado estéril, sendo removido, e em seguida, depositado na cava da mina juntamente com o material de sobrecarga (Silva *et al.*, 2008). Eventualmente, uma certa quantidade de BN pode estar presente no minério como *blend* ou quando o horizonte subjacente BNC é lavrado e há descontinuidade horizontal do perfil e / ou seus limites não

estão bem definidos. Melo *et al.* (2017) demonstraram que a caulinita presente nesta litologia é significativamente mais reativa do que as presentes na BC (horizonte de minério).

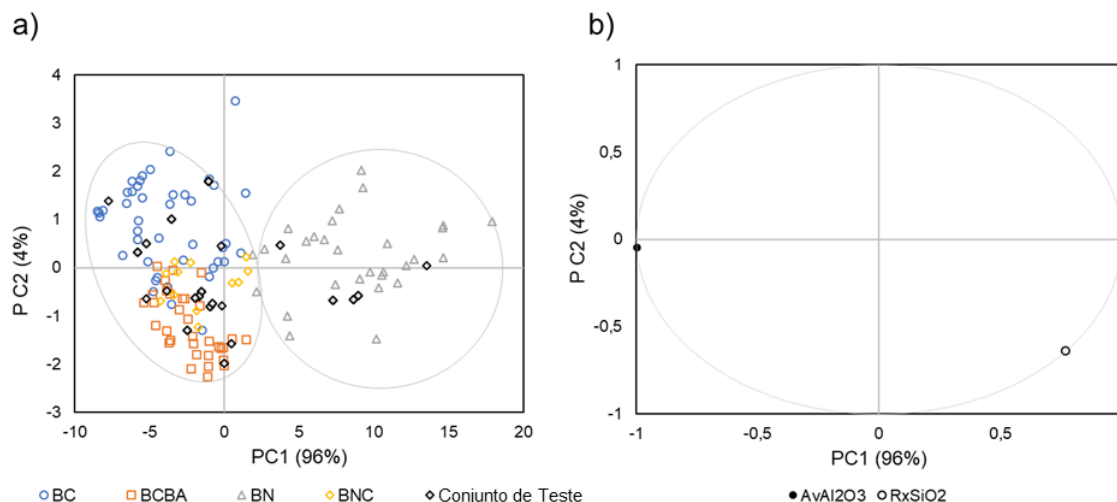


Figura 4.5: Gráficos de dispersão da análise de PCA a partir dos dados químicos de referência; a) escores, e b) cargas.

4.3.2. Difratomogramas obtidos pelo modo Reflexão

O gráfico de dispersão dos escores e cargas da análise de componentes principais realizado a partir dos difratogramas obtidos por reflexão é apresentado na Figura 4.6. Pode-se observar que esses dados podem ser descritos pelos dois primeiros componentes, que representam 98% da variação. No gráfico de escores (Figura 4.6a), três grupos de amostras (*clusters*) são destacados (denominados C1, C2 e C3), onde se observa que C1 e C2 sistematicamente tendem a se afastar um do outro. Isso significa que a parte dos padrões difratométricos (observações) que majoritariamente explicam a variação dos dados tem uma distribuição assimétrica, conforme observado no gráfico de cargas (Figura 4.6b).

Curiosamente, esses *clusters* não apresentam uma discriminação das litologias, conforme descrito por Melo *et al.* (2017). Em vez disso, o gráfico de escores mostra que há algum aspecto nos difratogramas dessas amostras que descreve uma heterogeneidade de cada litologia no perfil de bauxita. De fato, isso é explicitamente verdade para as litologias BCBA e BC, e esse comportamento mineralógico não é detectado apenas pela composição química. Alguns autores demonstraram a presença de várias gerações de gibbsita e caulinita ao longo

desses horizontes de minério (Kotschoubey *et al.*, 2005; Pereira, 2015). Como o PCA utiliza o todo o padrão difratométrico, não é levado em consideração apenas o teor de cada fase para discriminar as amostras, mas também suas características cristalinas. Obviamente, pode existir alguma influência não contabilizada devido à preparação da amostra – como orientação preferencial, no entanto, até o momento essa influência está desconsiderada, principalmente pelo fato de este ser um erro sistemático, o qual afetaria praticamente todas as amostras de bauxita.

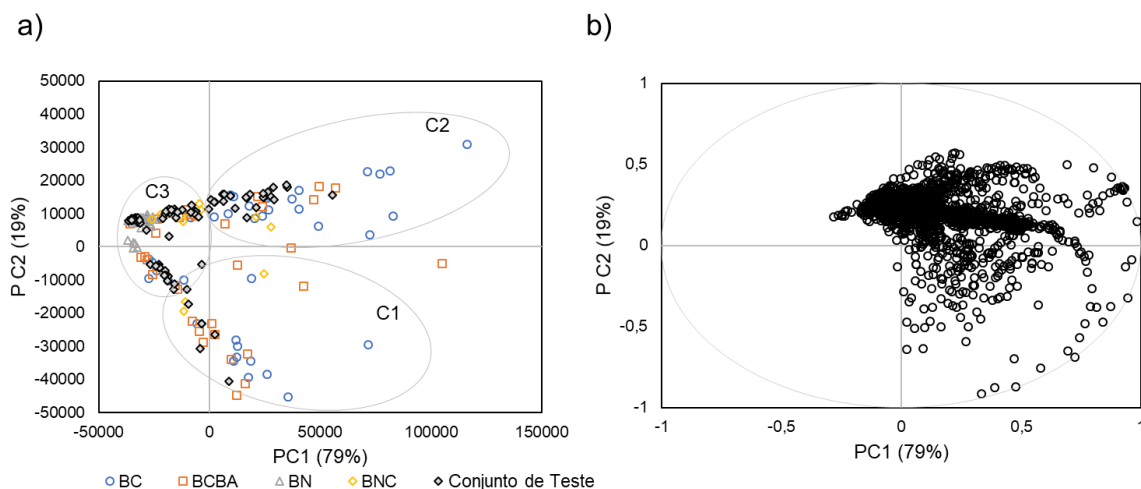


Figura 4.6: Gráficos de dispersão da análise de PCA a partir dos dados difratométricos obtidos por reflexão; a) escores, e b) cargas.

A Figura 4.7 mostra as cargas referentes aos componentes principais PC1 e PC2. Essa representação linear ajuda a entender a região do difratograma que descreve cada componente principal. Observa-se que a variação em PC1 é afetada principalmente pelas reflexões basais da gibbsita d_{002} e d_{004} (localizadas em $21,25^\circ$ e $43,38^\circ$ 2θ $K\alpha Co$, respectivamente). Observa-se também que partes indesejadas do difratograma, como o pico de $k\beta$ da gibbsita que aparece em $19,28^\circ$ (2θ $K\alpha Co$) e o ruído do *background* em ângulos abaixo de 12° (2θ), estão explicando uma parte significativa da variação em PC1. PC2 correlaciona negativamente os picos da caulinita (d_{001} e d_{002}), e de forma assimétrica, o pico principal da gibbsita (d_{002}), o que pode estar associado à duas gibbsitas com características cristalinas diferentes.

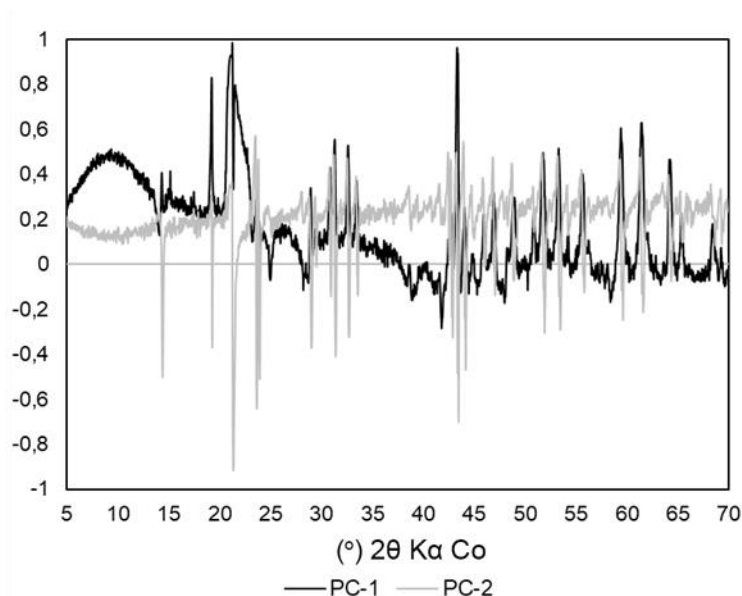


Figura 4.7: Representação linear das cargas (*loadings*) dos componentes principais PC1 e PC2 da análise de PCA a partir dos dados difratométricos obtidos por reflexão.

Através do gráfico de influência das amostras (Figura 4.8a) é possível detectar os *outliers*. Ao comparar esse gráfico com os escores e cargas (Figura 4.6), nota-se que a maior variação explicada pelo PC1 está relacionada principalmente a duas amostras com alto *leverage* (BC(10) e BCBA(22)) nas quais a intensidade das reflexões basais da gibbsita, d_{002} e d_{004} , é muito alta em comparação com a média (Figura 4.8b). Nesse caso, as duas amostras foram ser removidas do conjunto de calibração para evitar erros de predição. Decidiu-se manter as demais amostras com *leverage* e resíduos baixos no conjunto de calibração, pois pressupõe-se que o erro dessas amostras seja inerente à mineralogia das bauxitas e à preparação da amostra para análise de DRX.

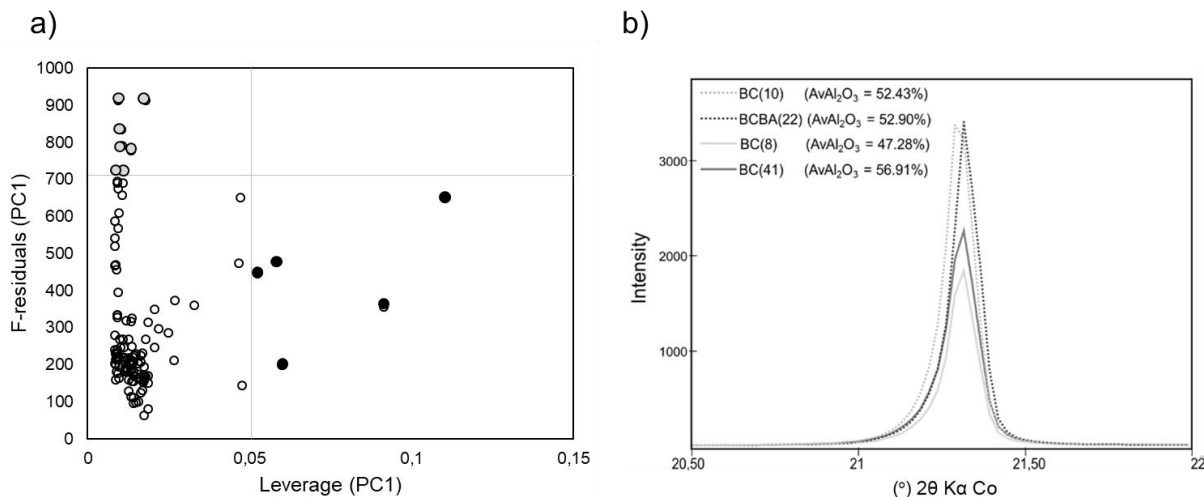


Figura 4.8: a) Gráfico de influência das amostras. Pontos pretos no quadrante inferior direito representam amostras com alto *leverage*; pontos cinza no canto superior esquerdo são amostras com alto resíduos; b) comparação dos padrões difratométricos das amostras com alto *leverage* (região de reflexão da gibbsita d_{002}).

4.3.3. Difrátogramas obtidos pelo modo Transmissão

A Figura 4.9 apresenta os gráficos de escores para os componentes principais PC1, PC2 (Figura 4.9a) e PC3 (Figura 4.9b). Como observado, os componentes principais PC1, PC2 e PC3 explicam, respectivamente, 35%, 22% e 12% da variabilidade dos dados (totalizando apenas 69% da variância explicada). Mesmo considerando 8 componentes principais não é possível obter um percentual maior que 75%. Isto significa que a partir destes dados não é possível obter uma projeção que resulte em um bom agrupamento de amostras correlacionadas. Devido a isto, utilizou-se o método de agrupamento *K-means*, com medida Euclidiana das distâncias para classificação das amostras em 3 grupos (Figura 4.10). Estes grupos foram utilizados posteriormente na criação dos modelos PLSR.

Como observado e discutido anteriormente, neste modo de coleta de dados difratométricos todas as observações (pontos de coleta ao longo do range 2θ) apresentam variâncias significativamente menores que na condição analisada por reflexão. Este fato contribui para o entendimento de que, para utilização de padrões difratométricos em qualquer tratamento e modelagem estatística em que haja redução de dimensionalidade e projeção dos dados por componentes (ou fatores) principais, é necessário que haja naturalmente uma

variabilidade considerável entre alguns picos (em geral os picos principais das fases), ou que esta variabilidade seja promovida durante a coleta (modificando o tempo de contagem, abertura das fendas). É sabido que neste último caso, há um aumento proporcional do ruído, no entanto, apenas uma pequena parcela deste ruído será absorvida pelos componentes principais utilizados para descrever a amostra. A outra possibilidade de promover um aumento na variabilidade das amostras é durante a sua preparação. Neste caso, grandes variações podem ser inseridas nos difratogramas relacionadas à montagem dos porta-amostras – orientação preferencial, deslocamento de picos devido à altura da amostra; e à cominuição/desagregação em grau – microdeformação, decomposição, alteração de polimorfos e amorfização da amostra (Jenkins & Snyder, 1996). Em todos estes caso, isto representa uma variabilidade não controlada e resulta em erros de classificação e predição das amostras, devendo ser evitadas ao máximo para tais aplicações.

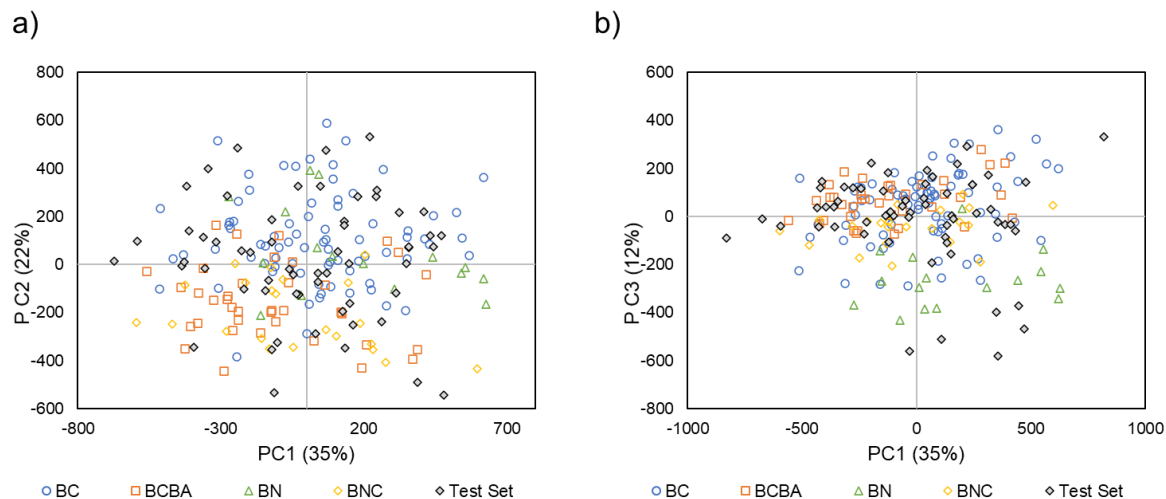


Figura 4.9: Gráficos de dispersão da análise de PCA a partir dos dados difratométricos obtidos por transmissão; a) escores PC1xPC2, e b) escores PC1xPC3.

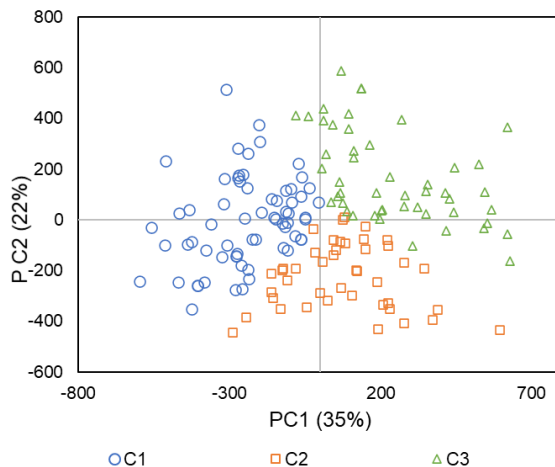


Figura 4.10: a) Agrupamento das amostras por *K-means* representados no gráfico de dispersão da análise de PCA.

Interessante notar no gráfico linear das cargas (*loadings*) do PCA para amostras obtidas por transmissão (Figura 4.11), que para o componente principal PC1 não há destaque do pico principal da gibbsita (>70%, no gráfico apresentado pela linha tracejada), pelo contrário, praticamente todos picos desta fase, inclusive das fases em menor teor como hematita, goethita e anatásio tem pesos proporcionais ao da fase majoritária (gibbsita). E para o PC2 o maior peso (>70%) de contribuição para a explicação da variância dos dados está relacionado apenas ao background em baixo ângulo. Também se nota que apenas o sexto componente principal (PC6) conseguiu extrair a variabilidade relacionada à caulinita, este, no entanto, explica apenas 1% da variância total dos dados.

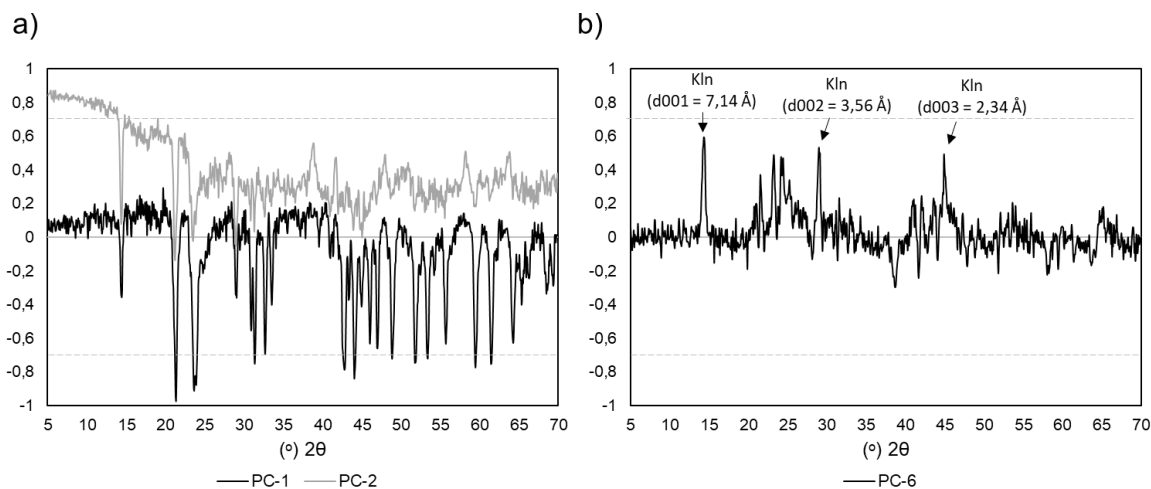


Figura 4.11: Representação linear das cargas (*loadings*) dos componentes principais a) PC1 e PC2 e b) PC6.

4.4. AVALIAÇÃO DO PRÉ-PROCESSAMENTO DOS DADOS DE DRX

Uma análise de PLSR foi realizado para definir as melhores condições de pré-processamento dos dados de DRX, seguindo um planejamento experimental fatorial 2^4 . Para o controle de qualidade da bauxita, é necessário que a metodologia de predição utilizando do PLSR com dados de DRX atenda aos requisitos estatísticos de exatidão e precisão, mas também que o tempo de análise seja rápido.

Em termos dos parâmetros de aquisição de dados de DRX, ao reduzir o intervalo 2θ e aumentar o tamanho do passo do goniômetro, o tempo de coleta de dados é significativamente reduzido, neste caso específico, 40 segundos (Tabela 4.2). Isso significa que toda a determinação dos parâmetros do controle de qualidade da bauxita pode ser obtida rapidamente (menos de cinco minutos por amostra – considerando o tempo de preparação dos espécimes para leitura).

Tabela 4.2: Tempo estimado de coleta dos dados de difração de raios-X (baseado nas características do difratômetro utilizado neste trabalho).

Tamanho de passo (2θ)	$5^\circ - 70^\circ (2\theta)$	$13^\circ - 34^\circ (2\theta)$
0,026°	4 min 54 s	1 min 35 s
0,065°	1 min 58 s	40 s

Para evitar a perda de intensidade em altos ângulos 2θ e reduzir a alta intensidade do *background* em ângulos baixos, é comumente usada uma fenda divergente automática (*automatic divergent slit* – ADS), que garante uma área irradiada constante na amostra durante a aquisição do difratograma. Neste estudo, uma transformação matemática foi aplicada nos dados de DRX coletados usando fenda divergente fixa (*Fixed divergent slit* – FDS). De acordo com Jenkins & Snyder (1996) o comprimento da área irradiada (L) é dado por:

$$L = \frac{2R \tan(\alpha/2)}{\sin \theta} \quad (11)$$

Onde R é o raio do goniômetro e α é o ângulo divergente. Portanto, se obtém um valor constante de L para todo o intervalo θ da medida, através da variação de α . Vale ressaltar que esses três pré-tratamentos de dados foram definidos após discussão dos resultados do PCA.

A centralização dos dados na média é o pré-processamento mais comum aplicado. Em geral, ele facilita os cálculos e torna o modelo menos complexo, eliminando a necessidade de ajustar um intercepto diferente de zero (Geladi & Kowalsky, 1986; Ferraro *et al.*, 2011). Escalonar os dados pelo desvio padrão (padronização) é um requisito personalizado geralmente aplicado quando as variáveis estão em escalas diferentes (por exemplo, %, ppm, °C). Neste caso, o peso de cada variável torna-se igual a 1, o que significa que cada uma têm a mesma oportunidade de ser modelada (Bro & Smilde, 2014). Comumente os dados utilizados em estatísticas multivariadas são *autoescalonados*, que consiste em padronizar e centralizar o conjunto de dados (Wold *et al.*, 2001).

Os resultados do DOE para a estimativa de $RxSiO_2$ estão resumidos na tabela de ANOVA (Tabela 4.3). Os melhores resultados foram alcançados utilizando autoescalonamento dos dados, onde geralmente 5 fatores são suficientes para o modelo preditivo com menor RMSEP. Interessante notar que o método de centralização na média geralmente resultou em modelos com maior número de fatores (8 fatores), os quais, apesar de apresentarem RMSECV mais baixos, ao serem utilizados nas amostras do conjunto de testes, os RMSEP obtidos foram maiores devido ao efeito de sobreajuste (Figura 4.12) (Faber *et al.*, 2003; Ferraro *et al.* a., 2011).

Adicionalmente, observou-se que reduzindo o intervalo de XRD 2θ para apenas 12° a 34° resultou em medições mais rápidas e predições mais precisas. A utilização de fenda automática não resultou em benefícios para a modelagem, ao contrário, conforme observado no caso dos difratogramas obtidos por transmissão, o realce de reflexões deletérias promove a incorporação destes nos componentes (fatores) principais os quais apresentam uma variância mais baixa o que resulta em modelos com baixo poder de predição e/ou sobreajustados.

Tabela 4.3: ANOVA da estimativa de $RxSiO_2$ no planejamento fatorial 2^4 . (RMSEP como variável resposta; $R^2_{adj} = 0,935$; nível de significância $\alpha = 0.05$).

Fatores	SS	df	MS	F	p-value
(1) Fenda Divergente	0,000052	1	0,000052	0,1636	0,702558
(2) Tamanho do passo	0,000199	1	0,000199	0,6244	0,465226
(3) Intervalo 2θ	0,008892	1	0,008892	27,9375	0,003230
(4) Modo de pré-processamento	0,060896	1	0,060896	191,3330	0,000035
1*2	0,001098	1	0,001098	3,4498	0,122386
1*3	0,000352	1	0,000352	1,1073	0,340848
1*4	0,000052	1	0,000052	0,1629	0,703140
2*3	0,002428	1	0,002428	7,6301	0,039728
2*4	0,000046	1	0,000046	0,1430	0,720858
3*4	0,003303	1	0,003303	10,3790	0,023423
Erro	0,001591	5	0,000318		
Total SS	0,078909	15			

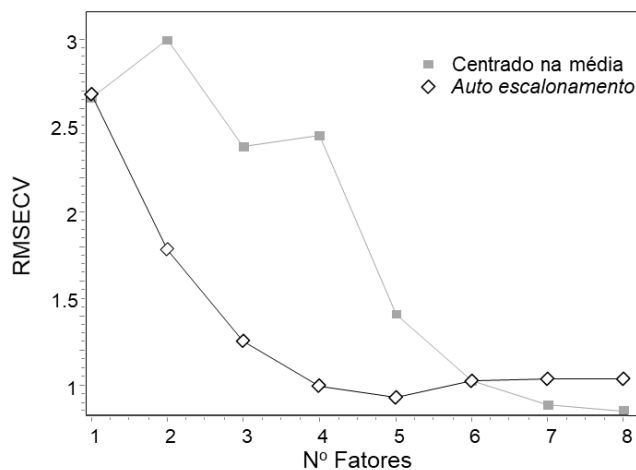


Figura 4.12: Comparação entre os métodos de pré-processamento dos dados - centrado na média (RMSECV = 0,845, RMSEP = 1,295, com 8 fatores) e autoescalonado (RMSECV = 0,924 e RMSEP = e 0,804, com 5 fatores)

4.4.1. Predição por PLSR a partir dos dados de DRX por reflexão

Com base nos grupos (*clusters*) identificados na análise de PCA, o conjunto de dados foi subdividido para gerar os modelos de PLSR e avaliar a qualidade das predições de $AvAl_2O_3$ e $RxSiO_2$. O conjunto de dados também foi pré-processado de acordo com as

melhores condições encontradas através do DOE – intervalo de medida: $13^{\circ} - 34^{\circ}$ (20), tamanho do passo: $0,065^{\circ}$, fenda divergente fixa e autoescalonamento dos difratogramas. A Figura 4.13 mostra a correlação entre os valores de referência (química úmida) e os estimados usando os modelos PLSR. Como observado, as amostras agrupadas em C1 e C2 que, em geral, têm o conteúdo de $AvAl_2O_3$ e $RxSiO_2$ na faixa de 49 – 55% e 2,5 – 5,5%, respectivamente, são bem modeladas, resultando em uma excelente predição. Por outro lado, a maioria das amostras em C3 não foi predita de maneira satisfatória.

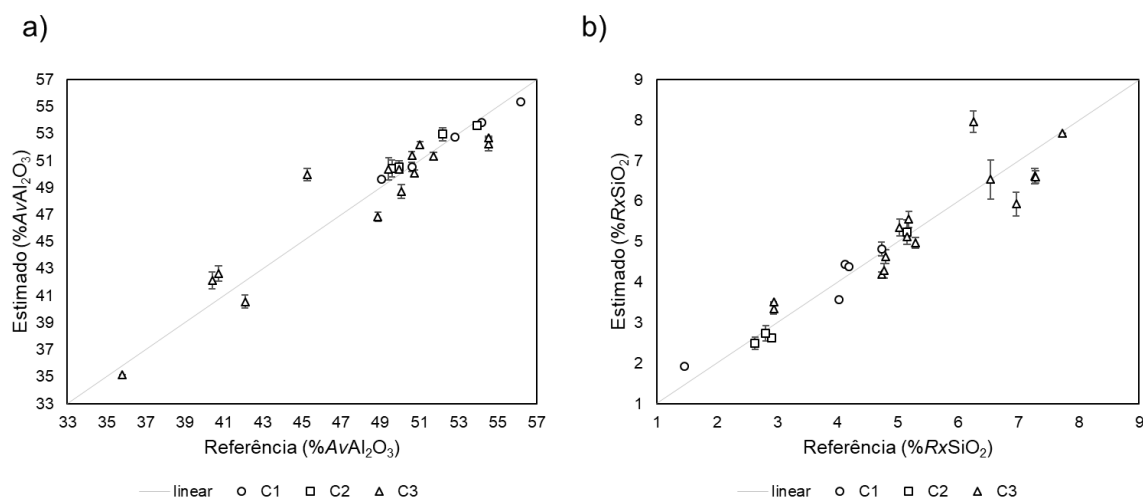


Figura 4.13: Comparação entre os valores de referência e estimados através dos modelos PLSR para cada grupo a partir de DRX por reflexão a) $AvAl_2O_3$, b) $RxSiO_2$. Barras indicando o RSD (*Residual Standard Deviation* - desvio padrão residual).

A precisão de cada modelo foi verificada pelo respectivo RMSEP indicado na Tabela 4.4 e a média dos resíduos na Figura 4.14. Os resultados mostraram uma precisão satisfatória para ambos os parâmetros quando preditos pelos modelos C1 e C2 ($RMSEP < 0,85$). Os resíduos próximos a zero para todos os modelos significam que eles não são estatisticamente enviesados (*bias*), e o RSD menor que 1% mostra uma boa precisão. Para as amostras C3 também foi obtida uma precisão satisfatória para $RxSiO_2$, mas observa-se um grande desvio ao tentar prever $AvAl_2O_3$. Industrialmente, para o controle de qualidade da bauxita geralmente é adotado uma precisão de $< 0,5\%$ para sílica reativa e de $1 - 1,5\%$ para alumina disponível, quando determinado pelos métodos tradicionais de química úmida. Feret (2018) argumenta que a precisão real não pode ser determinada pelos métodos de química úmida, uma vez que são empíricos, e na prática, muitas vezes é difícil atingir esse limite de precisão. Faber *et al.* (2003) demonstraram que é possível obter um modelo de PLSR mais preciso do

que a referência usada na calibração, e o erro de predição (RMSEP) em muitos casos pode ser tendenciosamente mais alto quando a predição é baseada em valores de referência relativamente imprecisos, uma vez que essa incerteza é incorporada na valores medidos.

Tabela 4.4: Valores de RMSEP para $AvAl_2O_3$ e $RxSiO_2$ para os modelos de cada grupo.

Modelos	RMSEP	
	$AvAl_2O_3$	$RxSiO_2$
C1	0.493	0.326
C2	0.834	0.230
C3	1.704	0.641

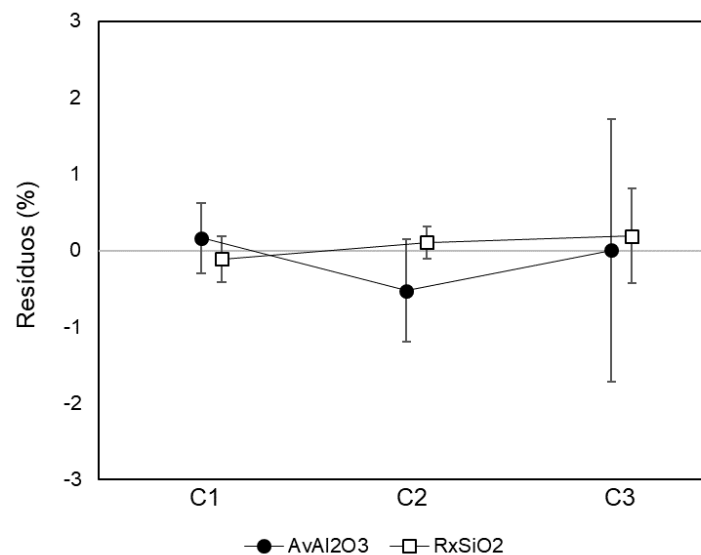


Figura 4.14: Gráfico de resíduos médios para os modelos de cada grupo. Barras indicando o RSD (desvio-padrão residual).

De forma semelhante a afirmação apresentada por Wold *et al.* (2001), os difratogramas podem ser vistos como a soma das reflexões das fases presentes multiplicadas pelas respectivas concentrações numa dada amostra. Isso significa que, em uma modelagem de PLSR, as variáveis latentes contidas em $\mathbf{X}$ a serem determinadas (neste caso, para prever $AvAl_2O_3$ e $RxSiO_2$) são explicadas pelo padrão difratométrico característico (identificado como $\mathbf{p}$ -cargas) e pela concentração da fase mineral relacionada a essa variável (identificada como escores $\mathbf{t}$) $\mathbf{X} = \mathbf{t}_1\mathbf{p}_1' + \mathbf{t}_2\mathbf{p}_2' + \dots + \text{ruído}$. Como exemplo, a Figura 4.15 mostra a representação linear das cargas para $RxSiO_2$ no modelo C3. Pode-se observar o quanto as intensidades de cada fase, bem como a variação da mineralogia, são responsáveis pelos

preditores. Neste último caso, pode-se observar que duas gerações de gibbsita foram distinguidas pelas cargas do modelo. König *et al.* (2012), Pereira (2015) e Negrão *et al.* (2018) descrevem essas duas gibbsitas como uma fase grosseira de alta cristalinidade preenchendo os microporos e uma fase dispersa na matriz de granulometria fina e baixa cristalinidade.

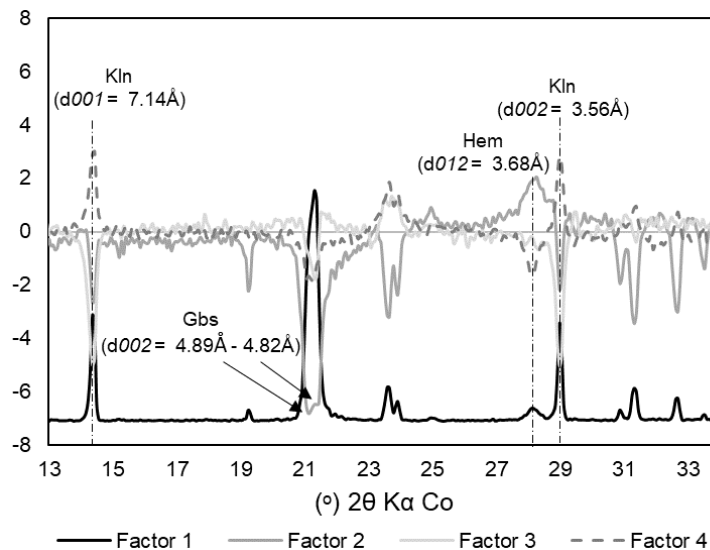


Figura 4.15: Representação linear das cargas do modelo PLSR para $RxSiO_2$ a partir das amostras de calibração do grupo C3. *Legenda:* Kln = caulinita; Gbs = gibbsita; Hem = hematita.

De fato, as intensidades de pico (I_{hkl}) de uma fase α em um difratograma estão relacionadas à sua fração de volume (v_α) de acordo com a equação simplificada:

$$I_{(hkl)\alpha} = \frac{K_e \cdot K_{(hkl)\alpha} \cdot v_\alpha}{\mu_s} \quad (12)$$

Onde K_e é uma constante para um sistema experimental específico; $K_{(hkl)\alpha}$ são constantes para cada reflexão de difração hkl da fase α ; e μ_s é o coeficiente de atenuação (Jenkins & Snyder, 1996). Baseado nisso, vários métodos podem ser usados para estimar a composição mineralógica da amostra, onde o principal é o refinamento de Rietveld. A aplicação deste método no controle de qualidade da bauxita está sendo constantemente desenvolvida, onde as imprecisões ocasionadas pelo teor de amorfo e substituições isomórficas foram substancialmente melhoradas (Neumann *et al.*, 2014; Paz *et al.*, 2018; Angélica *et al.*, 2018). Apesar disso, esse método ainda não é aplicado na indústria da

bauxita, principalmente devido à sua dificuldade de implementação – exigindo um conhecimento cristalográfico avançado e uma preparação de amostra personalizada. Além disso, é essencial uma boa coleta de dados de DRX que resultaria em uma medição demorada.

O PLSR baseado em dados de DRX, por sua vez, pode ser facilmente implementado, obtendo uma estimativa rápida e confiável dos parâmetros do controle do processo de bauxita. Um ponto de atenção ainda é a preparação da amostra. Para demonstrar isso, a Figura 4.16 mostra os padrões difratométricos das amostras do conjunto de teste C3, no intervalo da reflexão de caulinita d_{001} . As amostras duplicadas de BCBA destacadas foram aquelas que apresentaram pior predição (réplica 1 = 8,14%, réplica 2 = 7,76%, valor de referência = 6,25%, veja a Figura 4.9b). Nesse caso, esta amostra de BCBA possui uma caulinita de maior cristalinidade que as demais (medida pela largura à meia altura – $FWHM \sim 0,200$; $FWHM$ médio das demais amostras $\sim 0,300$). Isso resultou numa forte orientação nos planos $00l$, e por consequência, uma correlação falsa entre intensidade e teor de caulinita. De acordo com König & Norberg (2019), o PLSR é menos sensível que o refinamento Rietveld para a preparação de amostras não-ideais, desde que as condições sejam constantes (tamanho de partículas e pressão aplicada para compactar as amostras). Considerando que as amostras deste trabalho foram preparadas manualmente, acredita-se que grande parte do erro de predição das amostras C3 foi em decorrência de uma condição de preparação não constante. Conforme sustentado por Jenkins & Snyder (1996), neste trabalho também se observou que o maior problema relacionado à preparação de amostras para DRX é devido ao efeito de orientação preferencial.

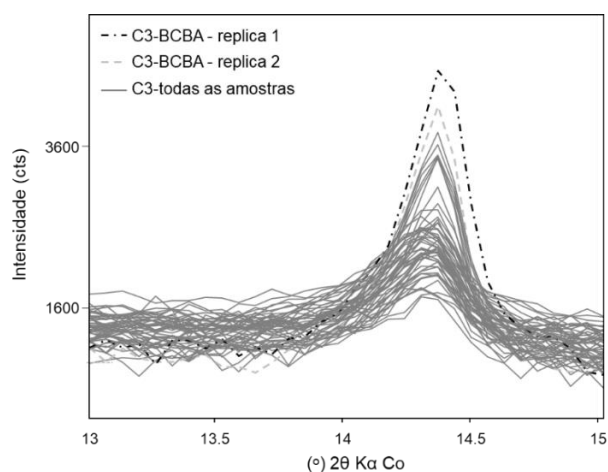


Figura 4.16: Padrão difratométricos das amostras do grupo C3 – mostrando o intervalo da reflexão principal da caulinita (d_{001}).

4.4.2. Predição por PLSR a partir dos dados de DRX por transmissão

Seguindo um procedimento semelhante ao adotado para os dados difratométricos coletados no modo reflexão, as amostras foram agrupadas utilizando a análise de PCA em três grupos (C1, C2 e C3) seguindo o método heurístico de agrupamento por *K-means*.

Estas amostras foram então utilizadas para calibração dos modelos PLSR, adotando os parâmetros obtidos pelo DOE – tamanho do passo: $0,065^\circ$ e fenda divergente fixa. Neste caso não se utilizou o intervalo de medida dos difratogramas otimizado pelo DOE ($13^\circ - 34^\circ$). Como discutido anteriormente, esse range otimizado para a condição de reflexão abrange os picos principais da gibbsita e caulinita, os quais como observado pelo PCA tem peso significativo nos fatores (componentes principais). Em contrapartida, na análise de PCA dos dados por transmissão mostrou que os efeitos principais dependem dos picos ao longo de todo padrão difratométrico, tendo apenas a região de baixo ângulo ($< 13^\circ$) um efeito deletério ocasionado pelo *background*. Assim, para esta condição de coleta de dados decidiu-se utilizar o range de 13° a 65° (2θ).

Adicionalmente, não se estabeleceu um método de pré-processamento dos difratogramas como observado na condição otimizada, uma vez que, em testes preliminares se observou que para este modo de coleta de dados não é possível fixar este parâmetro e obter uma modelagem satisfatória.

Os resultados de predição da $AvAl_2O_3$ e $RxSiO_2$ correlacionados aos valores químicos de referência são apresentados na Figura 4.17. Pode-se observar que há um bom ajuste dos valores preditos para ambos os parâmetros de controle de qualidade, principalmente naquelas amostras agrupadas em C1. Embora os valores médios preditos sejam próximo dos valores de referência, as amostras dos grupos C2 e C3 apresentaram uma precisão (RMSEP) acima dos limites aceitáveis para o controle de qualidade das bauxitas (Feret, 2018).

As Tabelas 4.5 e 4.6 apresentam os valores de qualidade do modelo RMSEP e média dos resíduos, bem como o método de pré-processamento e o número de fatores adotados em cada um dos grupos para predições dos parâmetros $AvAl_2O_3$ e $RxSiO_2$, respectivamente. Estes resultados podem ser visualizados na Figura 4.18.

Em comparação com os resultados obtidos nos modelos com difratogramas coletados por reflexão, pode-se observar que o tratamento utilizando transmissão foi menos favorável. Conforme observado nos difratogramas, é evidente a redução da orientação preferencial, a

qual possivelmente, foi o principal efeito deletério na qualidade de predição no tratamento por reflexão. No entanto, este modo de coleta de dados favorece a incorporação de muitas observações (picos de difração) nos componentes principais, o que significa que esta fonte de dados não possibilitou a detecção de nuances do difratograma que são características à sua composição química. De forma geral, os horizontes de bauxita referentes às bauxitas marginais / gangas (BN e BCBA) que apresentam uma grande variação de concentração e cristalinidade dos minerais, principalmente gibbsita e caulinita, ao longo do perfil, apresentam dificuldades em se obter uma predição satisfatória pelo método desenvolvido. Esta observação ressalta uma possível limitação do método para aplicações na mineração de bauxita relacionadas à pesquisa geológica e plano de mina.

O tratamento utilizando coleta de dados por transmissão demonstrou que não é necessariamente requerido a eliminação do efeito de orientação preferencial durante a preparação, mas sim que esta preparação seja feita de forma mais homogênea possível.

Em contrapartida, a predição por PLSR de amostras correspondente ao horizonte do minério (BC) se mostrou satisfatória em ambos os modos de coleta de dados. Isto demonstra que o método pode ser aceitavelmente utilizado para o controle de qualidade das bauxitas que alimentam o processo. Embora na prática não se observe uma variação brusca na composição química deste horizonte que acarretassem em predições erradas, à medida que a mineração vai avançando para novas áreas é recomendável uma reclassificação / recalibração dos modelos.

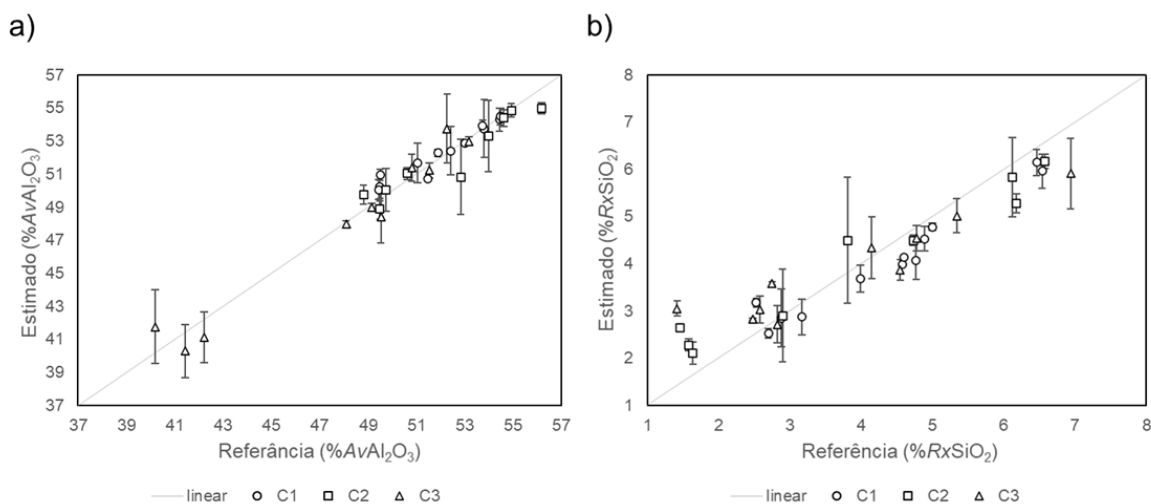


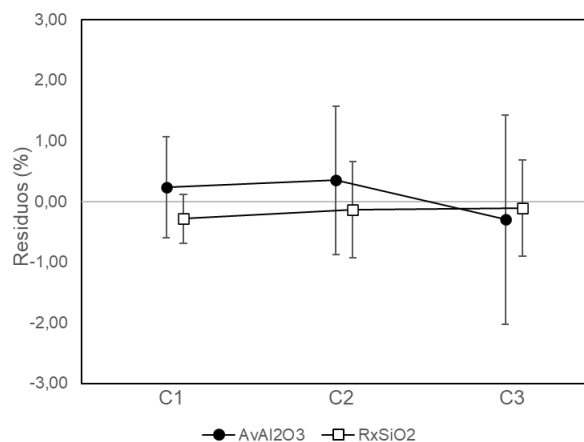
Figura 4.17: Comparação dos valores de referência e estimados através dos modelos PLSR para cada grupo (DRX por transmissão). a) $AvAl_2O_3$, b) $RxSiO_2$. Barras indicando o RSD.

Tabela 4.5: Resultados da predição de AvAl_2O_3 para os modelos de cada grupo.

Modelos	RMSEP	Média	Método de pré-processamento (nº de fatores do modelo)
C1	0,85	0,24	Centrado na média (5)
C2	1,24	0,35	Centrado na média (4)
C3	1,71	-0,30	Centrado na média (5)

Tabela 4.6: Resultados da predição de RxSiO_2 para os modelos de cada grupo.

Modelos	RMSEP	Média	Método de pré-processamento (nº de fatores do modelo)
C1	0,49	-0,28	Autoescalonado (3)
C2	0,78	-0,14	Centrado na média (5)
C3	0,78	-0,11	Autoescalonado (4)

**Figura 4.18:** Gráfico de resíduos médios para os modelos de cada grupo. Barras indicando RSD.

5. CONCLUSÕES

Neste trabalho buscou-se desenvolver um método para estimar os parâmetros de controle de qualidade de bauxita – $AvAl_2O_3$ e $RxSiO_2$, como alternativa aos métodos atuais de química úmida. Para tal, investigou-se o uso de PLSR com dados de difração de raios-X por reflexão e transmissão em bauxitas do platô Miltônia 3, Paragominas-PA.

A caracterização realizada das amostras de bauxita provenientes desta região demonstrou que todas apresentam a mesma composição mineralógica, variando apenas a composição e a cristalinidades dos minerais presentes. Com efeito de se estimar os parâmetros químicos de qualidade destas bauxitas, se observou que a $AvAl_2O_3$ corresponde exclusivamente ao mineral gibbsita, enquanto a $RxSiO_2$ à caulinita.

Através de análise por PCA foi possível identificar amostras *outliers* e agrupamento de amostras com similaridades mineralógicas. O PCA realizado com difratogramas obtidos por transmissão apresentou baixa variância explicada e as amostras foram agrupadas utilizando o método *K-means*. Em ambos os tratamentos (coleta de dados por reflexão e transmissão) as amostras foram agrupadas em 3 grupos.

Utilizando planejamento experimental fatorial (DOE) foi possível obter uma melhor condição de coleta e pré-tratamento dos difratogramas. Neste contexto, foi possível otimizar o tempo de coleta dos difratogramas através da redução do intervalo 2θ e do tamanho do passo, partindo de um tempo de ~5 minutos (para o intervalo de coleta: $5^\circ - 70^\circ$ e tamanho de passo: $0,026^\circ$), chegando a ~ 40 segundos reduzindo o intervalo para $13^\circ - 34^\circ$ e aumentando o tamanho de passo para $0,065^\circ$. Esta condição de análise por difratometria de raios-X, embora peculiar e insatisfatória para a maioria das aplicações da técnica, proporcionou uma melhoria relacionada à predição por PLSR.

Baseado nos resultados de predição se observou que a coleta de dados por reflexão é mais vantajosa que por transmissão. Embora este último possibilite a redução do efeito de orientação preferencial, o destacamento proporcionado dos demais planos $hk0$ das fases majoritárias e dos picos principais das fases deletérias para a predição, tornou esta fonte pobre em fornecer características latentes correlacionadas aos parâmetros preditos. Um ponto de destaque observado é que, em ambos os tratamentos, o método desenvolvido apresenta baixa precisão para bauxitas marginais / gangas, podendo ser esta uma limitação do método ou uma

oportunidade de otimização. Outro ponto de atenção para o método, considerando a coleta por reflexão, é em relação a preparação das amostras. Conforme observado nos resultados, não é necessário a eliminação do efeito de orientação preferencial durante a preparação, mas é fundamental que esta seja o mais homogêneo possível. Felizmente, em indústrias é possível automatizar a preparação e montagem dos porta-amostras, obtendo simultaneamente um ganho de produtividade e redução do erro de predição.

A qualidade dos modelos preditivos foi avaliada com base no RMSEP o qual corresponde a dimensão do erro de predição ao se utilizar amostras diferentes daquelas usadas para calibração, além da avaliação da média dos resíduos o qual corresponde ao quanto o modelo é enviesado para sub ou sobrestimar os parâmetros. Para tal avaliação, se considerou como referência o erro admitido para quantificação dos parâmetros $AvAl_2O_3$ e $RxSiO_2$ na indústria da bauxita e alumina. Nesse contexto, todas as amostras correspondentes ao horizonte do minério (BC) e amostras de bauxitas marginais que fazem interface (BNC e BCBA superior) atenderam satisfatoriamente os limites de $AvAl_2O_3 \pm 1\%$ e $RxSiO_2 \pm 0,5\%$ requeridos pela indústria.

Estes resultados demonstraram que o método proposto tem potencial para ser utilizado industrialmente no controle de qualidade de bauxitas gibbsíticas, com precisão dentro da faixa esperada para o minério. Em comparação com o método tradicional, o método proposto é significativamente mais rápido, fácil de implementar e executar as análises, exige menos espaço e mão-de-obra no laboratório, pode ser totalmente automatizado, e não requer uso de reagentes químicos. Além disso, com a implementação da difratometria de raios-X em laboratórios de controle da indústria da bauxita e alumina, é possível acompanhar a mineralogia do minério que alimenta o processo Bayer e, portanto, estar ciente de como variações na composição mineralógica podem impactar o processo. Sendo válido ressaltar que tais informações são desconhecidas controlando apenas os parâmetros químicos.

5.1. SUGESTÕES PARA TRABALHOS FUTUROS

Aumentar o número de amostras analisadas com o objetivo de dar maior confiabilidade estatística referente à qualidade dos modelos preditivos.

Realizar o preenchimento do porta amostras para leitura no difratômetro de forma automatizada, com o objetivo de padronizar a carga aplicada e eliminar / reduzir os erros de predição relacionados a não uniformidade da preparação de amostra.

Realizar análises complementares de MEV e PSD em amostras mais representativas e *outliers* para melhorar a discussão sobre as gerações de gibbsita e de caulinita, bem como da relação textural desses minerais.

Analisar por difratometria de raios-X bauxitas de diferentes localidades, em especial bauxitas contendo outras espécies de $AvAl_2O_3$ como boehmita e diásporo, e avaliar o desempenho da metodologia proposta neste trabalho para predizer os teores de alumina aproveitável e sílica reativa a partir destas amostras.

Desenvolver código de programação integralizando as operações unitárias da metodologia proposta: classificação e agrupamento de amostras por PCA; calibração dos modelos PLSR de $AvAl_2O_3$ e $RxSiO_2$ para cada grupo de amostras similares; automatização da classificação de novas amostras de acordo com os grupos pré-configurados e predição dos parâmetros também de acordo com o modelo PLSR referente ao respectivo grupo. Codificar uma interface GUI intuitiva para que o usuário (ex.: técnico ou analista de laboratório de controle de processo) possa utilizar o método de forma simples e objetiva a fornecer os parâmetros de controle de qualidade das bauxitas.

6. REFERÊNCIAS

Alcan, 1989a. Determination of gibbsitic alumina and kaolinitic silica in high-grade bauxites ($>41\%$ g. Al_2O_3) using a single low-temperature caustic digest. Method No 1243-89. Alcan International Limited.

Alcan, 1989b. Chemical determination of gibbsitic alumina in low-grade bauxites ($>41\%$ g. Al_2O_3) using a low-temperature caustic digest. Method No 1253-89. Alcan International Limited.

Angélica, R.S., Kahn, H., Paz, S.P.A., 2018. A proposal for bauxite quality control using the combined Rietveld – Le Bail – Internal Standard PXRD method – Part 2: Application to a gibbsitic bauxite from the Paragominas region, northern Brazil. *Minerals Engineering*, 122, 148–155.

Antoniassi, J. L., 2010. A difração de raios X com o método de Rietveld aplicada a bauxitas de Porto Trombetas, PA. Dissertação (Mestrado em Engenharia Mineral) - Escola Politécnica, Universidade de São Paulo, São Paulo. doi:10.11606/D.3.2010.tde-23112010-090249.

Authier-Martin, M., Forte, G., Ostap, S., See, J., 2001. The mineralogy of bauxite for producing smelter-grade alumina. *Journal of Metals*, December 2001, 36-40.

Authier-Martin, M., Girard, R., Risto, G., Verville, D., 2002. Boehmite reversion: Predictive test and critical parameters for bauxites from different geographical origins. In: *Proceedings of the 6th International Alumina Quality Workshop*, 109-114.

Bish, D.L., Post, J.E. 1993. Quantitative mineralogical analysis using the Rietveld full-pattern fitting method. *American Mineralogist*, 78, 932-940.

Bocquier, G., Muller, J. P., Boulange, B., 1984. Les laterites: connaissances et perspectives actualles sur les mecanismes de leur differenciation. In : *Livre jubilaire du Cinquantenaire de I'AFES*, 123 – 138.

Boulangé, B., Ambrosi, J. P., e Nahon, D. *Laterites and Bauxites*. Cap. 3. 49-55, 1997.

Brasil. Agência Nacional de Mineração. *Sumário Mineral 2017*. Coordenação Geral Marina Marques Dalla Costa, Karina Andrade Medeiros e Thiers Muniz Lima. Brasília: ANM, 2019.

Bredell, J. H., 1983. Calculation of Available Alumina in Bauxite during Reconnaissance Exploration. *Economic Geology*, 78, 319-325.

Bro, R., Smilde, A. K., 2014. Principal Component Analysis. *Anal. Methods*, 6, 2812-2831. DOI: 10.1039/c3ay41907j

Carneiro, C. C., Yanez, D. N. D. V. S., Ulsen, C., Fraser, S. J., Antoniassi, J. L., Paz, S. P. A., Angélica, R. S., Kahn, H., 2017. Imputation of reactive silica and available alumina in bauxites by self-organizing-maps. In: *Anais. Nancy: IEEE*, 2017. Disponível em: DOI: 10.1109/WSOM.2017.8020008.

Cordella, C. B. Y., 2010. PCA: The Basic Building Block of Chemometrics. Analytical Chemistry. *IntechOpen*. Nov 7. DOI: 10.5772/51429

Costa, M.L., 1997. Lateritization as a major process of ore deposit formation in the Amazon region. *Exploration and Mining Geology* 6, 79-104.

Ermrich, M., Oppen, D. *XRD for the analyst – Getting acquainted with principles*, PANalytical B.V., Almelo, 2013.

Faber, N.M., Song, X.-H., Hopke, P.K., 2003. Sample-specific standard error of prediction for partial least squares regression. *Trends in Analytical Chemistry*, 22 (5), 330-334.

Feret, F. R., Authier-Martin, M., Sajó, I., 1997. Quantitative phase analysis of Bidi-Koum bauxites (Guinea). *Clays and Clay Minerals*, 45-3, 418-427.

Feret, F. R., 2018. Inaccuracies in Estimation of Bauxite Extractable and Mineralogical Constituents. In: *Proceedings of the 36th International ICSOBA Conference, Belem, Brazil, 29 October - 1 November, 2018*, 155-168.

Ferraro, M. C. F., Castellano, P. M., Kaufman, T. S., 2001. A spectrophotometric - partial least squares (PLS-1) method for the simultaneous determination of furosemide and amiloride hydrochloride in pharmaceutical formulations. *Journal of Pharmaceutical and Biomedical Analysis*. Vol. 26, 43–451.

Freyssinet, Ph., Butt, C. R. M., Morris, R. C., Piantone, P., 2005. Ore-forming process related to lateritic weathering. *Economic Geology*. 100, 681-722.

Geladi, p., Kowalski, B. R., 1986. Partial least-squares regression: a tutorial. *Analytica Chimica Acta*. 185, 1–17.

Gobbo, L. A., Montanheiro, T. J., Montanheiro, F., Sant'Agostino, L. M., 2013. Avaliação de cimento portland a partir da difração de raios X associada à análise por agrupamento. *Revista do Instituto Geológico, São Paulo*, 34 (2), 41-49, DOI: 10.5935/0100-929X.20130008.

Grubbs, D.K. ,1987. Reduction of fixed soda losses in the Bayer process by low temperature processing of high silica bauxites. *Light Metals TMS*, Warrendale, 19–25.

Jenkins, R., Snyder, R. L. *Introduction to X-Ray Powder Diffractometry*. John Wiley and Sons, New York, 1996, 403+XXIII pp.

Jolliffe, I. T., *Principal Component Analysis*, 2nd ed. Springer-Verlag, New York, 2002.

König, U., Angélica, R. S., Norberg, N., Gobbo, L., 2012. Rapid X-ray diffraction (XRD) for grade control of bauxites. In: *International Committee for Study of Bauxite, Alumina & Aluminium – ICSOBA, November 2012*, Belém. Papers, Bauxite Program BX14-T, 11 p., CD-ROM.

König, U., Degen, T., Norberg, N., 2014. PLSR as new XRD method for downstream processing of ores – case study: Fe^{2+} determination in iron ore sinter. *Powder Diffraction*, 29, 78 – 83.

König, U., Degen, T., Norberg, N., 2019. New tools for process control in aluminium industries – PLSR on XRD raw data. In: AQW Conference. Available at: <<http://www.aqw.com.au/papers/item/new-tools-for-process-control-in-aluminium-industries-plsr-on-xrd-raw-data.html>>

Kotschoubey B., Calaf J.M.C., Lobato A.C.C., Leite A.S. Azevedo C.H.D., 2005. Caracterização e gênese dos depósitos de bauxita da Província Bauxitífera de Paragominas, noroeste da bacia do Grajaú, nordeste do Pará/ oeste do Maranhão. In: *MARINI, O. (Org.). 2005. Caracterização em depósitos minerais em distritos mineiros da Amazônia*. Brasília, DF: DNPM-CT/MINERAL-ADIMB. p. 687–782.

Lawson, D., Rijkeboer, A., Andermann Jr., L. J., Mooney, A., 2008. Impact of Jamaican Bauxite Mineralogy on plant operations. *Light Metals TMS 2008*, 48 – 53.

McCormick, P.G., Picaro, T., e Smith, P.A.I., 2002. Mechanochemical treatment of high silica bauxite. *Minerals Engineering*, 15, 211-214.

Melo, C.C.A., Oliveira, K. S., Angélica, R. S., Paz, S. P. A., 2017. Análise de agrupamentos associada a difratometria de raios-X: uma classificação mineralógica prática de bauxitas e seus produtos de digestão Bayer. *Holos*, ano 33, Vol. 06, 32-42.

Mendes, F.A., 2012. Operating experience with Paragominas bauxitas wash plant. In: *International Committee for Study of Bauxite, Alumina & Aluminium – ICSOBA, 2012*, Belém. Papers, Bauxite Program, BX 20-T, CD-ROM.

Molai, H., 2011. Role of Mineralogy and Geochemistry in the Beneficiation of Jajarm Bauxite from North East Iran: Comparison with some other Bauxite Deposits of the World. *Iranian Journal of Earth Sciences*, 3, 134-167.

Nascimento, L.G., Paz, S.P.A., Angélica, R.S., e Kahn, H., 2013. Síntese de zeólitas a partir do rejeito gibbsítico/caulinítico do beneficiamento de bauxita de Paragominas-PA. In: *57º Congresso Brasileiro de Cerâmica & 5º Congresso Iberoamericano de Cerâmica*, Natal-RN

Najar, M. P. A., 2017. Rapid Analysis of Available Alumina and Reactive Silica in Bauxite. *Minerals & Metals Review* - January 2017, 99-100.

Negrão, L. B. A., Costa, M. L., Pöllmann, H., Horn, A., 2018. An application of the Rietveld refinement method to the mineralogy of bauxite-bearing regolith in Lower Amazon. *Mineralogical Magazine*, 82 (2), 413-431.

Neumann, R., Avelar, A. N., Costa, G. M., 2014. Refinement of the isomorphic substitutions in goethite and hematite by the Rietveld method, and relevance to bauxite characterization and processing, *Minerals Engineering*, 55, 80–86.

Novo, B. L., Sampaio, J. A., Silva, F. A. N. G., 2015. Estudo de caracterização e beneficiamento de bauxita com alto teor de ferro e sílica total. In: *jornada de iniciação científica*, 23, Rio de Janeiro. Anais. Rio de Janeiro: CETEM/MCT.

O'Connor, D. J., 1988. *Alumina Extraction from Non Bauxitic Materials*. Sydney: Aluminium-Verlag, 370 p.

Oliveira, K. S., 2017. Análise mineralógica por difratometria de raios-X e método de análise de agrupamento (cluster analysis) como critério para individualização de horizontes bauxíticos. Dissertação (Mestrado), PPGG/UFPA, Belém, PA, Brasil.

Ostap, S. 1984. Effect of bauxite mineralogy on its processing characteristics. In: *Bauxite, Proceedings of the 1984 Bauxite Symposium*, Los Angeles, California, Feb. 27 – March 1, 1984, Chapter 30, p. 651.

Paz, S.P.A., Angélica, R.S., Scheller, T., 2012. X-ray diffraction (XRD) studies of kaolinites to support mineralogical quantification of high silica bauxites from the Brazilian Amazon region. In: *International Committee for Study of Bauxite, Alumina & Aluminium – ICSOBA, 2012*, Belém. Papers, Bauxite Program, BX 16-T, 7 p., CD-ROM.

Paz, S. P. A., Angélica, R. S., Kahn, H., 2017. Optimization of the reactive silica quantification method applied to Paragominas-type gibbsitic bauxites. *International Journal of Mineral Processing*. 162, 48-57.

Paz, S.P.A., Kahn, H., Angélica, R.S., 2018. A proposal for bauxite quality control using the combined Rietveld – Le Bail – Internal Standard PXRD method – Part 1: hkl model developed for kaolinite. *Minerals Engineering*, 118, 52–61.

Paz, S. P. A., 2016. Desenvolvimento e otimização de métodos de controle de qualidade e de processo de beneficiamento para bauxitas gibbsíticas tipo-Paragominas. Tese (Doutorado em Engenharia Mineral) - Escola Politécnica, Universidade de São Paulo, São Paulo, 2016. doi:10.11606/T.3.2016.tde-22092016-100712

Pereira, C. B., 2015. Mineralogia e geoquímica dos perfis bauxíticos da mina Miltônia 3, região de Paragominas/PA. Dissertação (Mestrado), PPGG/UFPA, Belém, PA, Brasil.

Picanço, F. E., Pinheiro, D. C., Calado, W. M., Almeida, K. R., Quadros, C. R. R., Rabelo, S. A., 2013. Amostragem em diferentes malhas de sondagem para minério de bauxita – mina de Miltônia 03 – Paragominas-PA. In: *Anais do 13º Simpósio de Geologia da Amazônia*, setembro, 2013, 645-647.

Rayzman, V. L., Pevzner, I. Z., Sizyakov, V. M., Ni, L. P., Filipovich, I. K., Aturin, A. V., 2003. Extracting Silica and Alumina from Low-Grade Bauxite. *Extractive Metallurgy*, 47 – 50.

Sampaio, J.A., Andrade, M.C., e Dutra, A.J.B., 2005. Bauxitas. In: *Rochas & Minerais Industriais: Usos e Especificações*, 279-304. Rio de Janeiro: CETEM.

Santos, W.M., 2011. *Mineralogia e geoquímica da bauxita derivada do anortosito Barro Alto, Goiás*. Dissertação de Mestrado. Brasília, DF: UNIVERSIDADE DE BRASÍLIA - UnB; (Instituto de Geociências - IG).

Shlens, J., 2005. A Tutorial on Principal Component Analysis. *arXiv:1404.1100*

Silva, H. M., Picanço, E., Maurity, C., Morais, W., Santos, H. C., Guimarães, O., 2008. Geology, mining operation and scheduling of the Paragominas bauxite mine. In: *Proceedings of the 8th International Alumina Quality Workshop*, 11-16.

Smith, P., 2009. The processing of high silica bauxites - review of existing and potential processes. *Hydrometallurgy*, 98, 162–176.

Toledo, M.C.M., Oliveira, S.M.B., e Melfi, A.J., 2000. Intemperismo e formação do solo. In: *Decifrando a Terra*, por W. Teixeira, M.C.M. Toledo, T.R. Fairchild e F. Taioli, 139-166. São Paulo: Oficina de Texto.

U.S. Geological Survey, 2020, Mineral commodity summaries 2020: U.S. Geological Survey, 200 p., <https://doi.org/10.3133/mcs2020>.

Wold, H., 1966. Estimation of principal components and related models by iterative least squares. In: *Multivariate Analysis*, Academic Press, New York, 1966, 391-420.

Wold, S. et al., 1984, The collinearity problem in linear regression, the partial least squares approach to generalized inverses, *SIAM J. Sci. Stat. Comput.* Vol.5, 735–743.

Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemometr. Intell. Lab. Syst.* 58, 109–130.

Zhang, R., Gong, E., Wang, G., Yang, Z., 2019. An analytical shortcut to estimate alumina content by LOI in lateritic gibbsite bauxite prospecting. In: *AQW Conference*. Disponível em: <<http://www.aqw.com.au/papers/item/an-analytical-shortcut-to-estimate-alumina-content-by-loi-in-lateritic-gibbsite-bauxite-prospecting.html>>, acessado em 05-01-2020.

7. APÊNDICE – ARTIGOS DA TESE



A proposal for rapid grade control of gibbsitic bauxites using multivariate statistics on XRD data



Caio César Amorim Melo^{a,*}, Rômulo Simões Angélica^b, Simone Patrícia Aranha Paz^{a,b}

^a UFPA - Universidade Federal do Pará, Programa de Pós-Graduação em Engenharia de Recursos Naturais (PRODERNA), 66075-110 Belém, PA, Brazil

^b UFPA - Universidade Federal do Pará, Instituto de Geociências, Laboratório de Caracterização Mineral, 66075-110 Belém, PA, Brazil

ARTICLE INFO

Keywords:

Bauxite
Quality control
PLSR

ABSTRACT

Currently, traditional wet chemistry methods are used for quality control of bauxites. Such methods indirectly quantify the gibbsite and kaolinite content as available alumina ($AvAl_2O_3$) and reactive silica ($RxSiO_2$), respectively, and it is very costly and time-consuming. In order to achieve a rapid and reliable method to estimate these parameters, as alternative to current wet chemistry methods, this study, evaluates the use of multivariate statistics – Partial Least Square Regression (PLSR) on XRD data of gibbsitic bauxites (Miltonia 3 plateau). The method was optimized through Principal Component Analysis (PCA) and factorial Design of Experiments (DOE). The results demonstrated that such method has a potential to be industrially applied to quality control of bauxites. In comparison with traditional method, the same precision is achieved for the ore samples, but faster, easier to implement and carry out the analyses, demands less laboratory space and manpower, and no reagent is necessary. In addition, the use of XRD for quality control makes it possible to track the mineralogy of the bauxite that feeds the Bayer process, and therefore, be aware of how it could impact the process.

1. Introduction

Lateritic bauxite is the main ore to produce alumina, representing over than 90% of the deposits in the world. It consists in weathered bodies that accumulated economical amount of aluminum-bearing minerals to be explored and further extracted in Bayer Process (Freysinet et al., 2005). Mineralogically, the extractable sources of alumina are gibbsite ($Al(OH)_3$), boehmite and diaspore ($AlO(OH)$), and gangue mineral assembly is commonly kaolinite, quartz, hematite, goethite, anatase (minor Al, Fe, Si, Ti and P minerals are eventually present depending on the deposit). The content of these minerals, their crystallinity, and extent of isomorphic substitution can significantly impact the ore grade and performance during its processing (Ostap, 1984; Authier-Martin et al., 2001; Smith, 2009; Molai, 2011).

Interestingly, both in the mine and refinery, the ore quality control is taken almost exclusively in terms of the bauxite chemical composition. Thus, the material from geological survey, exploration, beneficiation plant and final ore are usually analyzed for total chemistry (generally by X-ray fluorescence spectroscopy – XRF), and content of available alumina ($AvAl_2O_3$) and reactive silica ($RxSiO_2$) as an indirect measurement of gibbsite/boehmite/diaspore and kaolinite/quartz, respectively – the two main parameters to process control, traditionally determined by wet chemistry. In this context, Paz et al. (2017)

demonstrated that the $RxSiO_2$ content can be underestimated by standard wet chemistry methods, depending on the content and crystallinity of kaolinite in bauxite. Similarly, Melo et al. (2017) verified that the structural order of kaolinite changes significantly along the bauxite profile, which means that silica may be more reactive despite their content. Thus, there are no guarantee that only knowing the bauxite chemical composition will allow an efficient control in the metallurgical processes. Therefore, it is valuable to know the mineralogy and texture of the minerals in the bauxite, since these factors determine the physicochemical interactions occurring in the processes (O'Connor, 1988; Paz et al., 2017).

Since the determination of $AvAl_2O_3$ and $RxSiO_2$ by standard wet chemistry is significantly time-consuming and demand a large quantity of chemicals, manpower and laboratory space (König et al., 2012; Paz et al., 2018), several authors have been proposing alternative methods to achieve good and reliable estimations as practical, rapid and economy as possible. Commonly, the methods are based on total oxides and loss of ignition (LOI) data using stoichiometric mass balance and multivariate technics (Bredell, 1983; Feret et al., 1997; Najjar, 2017; Carneiro et al., 2017; Zhang et al., 2019). The weakness of such methods is the need to know the entire bauxite mineralogy, and they tend to fail when some phase is not properly identified or when the characteristic of the bauxite is altered. An interesting alternative for

* Corresponding author.

E-mail address: eng.caiomelo@hotmail.com (C.C.A. Melo).

bauxite quality control is using X-ray diffraction (XRD) data, due to its robustness, rapidness of analysis, ease to be fully automated, and mainly because it also allows a mineralogical understanding of ore. Good advances have been demonstrated in mineralogical quantification of gibbsitic bauxites based on Rietveld refinement (Paz et al., 2018; Angélica et al., 2018) and multivariate statistics (König et al., 2012; Melo et al., 2017).

In this context, the present work aimed to develop an optimized methodology to estimate the main parameters of bauxite quality control ($AvAl_2O_3$ and $RxSiO_2$) using the multivariate statistics method Partial Least Square Regression (PLSR) on the XRD data, preprocessing the data and building adjusted models based on the best predictors for bauxites with different mineralogical characteristics. Therefore, the purpose is to present a faster and cheaper alternative to traditional wet chemistry methods, ensuring the reproducibility and accuracy of prediction even when the ore feeding the plant is changed.

2. Materials and methods

2.1. Materials

The bauxite samples used in this work are from a survey conducted by Mineração Paragominas SA (Norsk Hydro) in 2012, in which a drilling campaign with holes arranged in a regular grid of 25 m × 25 m on the Miltonia 3 plateau was carried out, in a total of 174 holes and 1396 samples corresponding to typical lithologies of the Paragominas lateritic profile (Picanço et al., 2013). An aliquot of 10 g of each sample was provided by Hydro, where all sample preparation (comminution, homogenization / quartering) and chemical analysis ($AvAl_2O_3$ and $RxSiO_2$) were performed by the company itself, following the routine wet chemistry method. From this sampling, 142 samples were selected for this study corresponding to the lithologies Nodular Bauxite (BN), Nodular-Crystallized Bauxite (BNC), Crystallized Bauxite (BC) and Crystallized-Amorphous Bauxite (BCBA). Such terminology is widely used in the industry and is related to the geological aspects of the lateritic bauxite profile in this region, which is depicted in Fig. 1. Among these lithologies, BC represents the main ore horizon.

In this work, the dataset was divided to calibrate and test the PLSR models. In total, 120 samples were randomly selected for the calibration set and 22 for the test set. In addition, some samples from the test set were analyzed in replicates to assess model precision.

2.2. XRD data

The fundamental analytical technique used for the development of this work is X-ray diffraction. Sample preparation is an important factor in achieving representative results. Since the main phases in the bauxite – gibbsite and kaolinite, present a plate-like habit, the basal planes (001) are easily oriented as consequence of the sample preparation (known as preferred orientation). In this case, the XRD pattern obtained has uneven intensities, with main reflections of these phases (gibbsite d_{002} and kaolinite d_{001}) overestimated (Jenkins & Snyder, 1996). To reduce this preferential orientation effect, backloading preparation was used (McCusker et al., 1999; Paz et al., 2018). The samples were prepared manually trying to be as constant as possible to reduce the non-uniform preparation error.

The powder XRD data were collected using a PANalytical Empyrean diffractometer, equipped with a θ - θ goniometer, Co X-ray tube ($K\alpha_1 = 1.78901 \text{ \AA}$), Fe K β filter and PIXel3D 2x2 area detector (linear scanning mode) with active length of $3.3473^\circ 2\theta$ (255 channels). The diffractograms were collected under the following conditions: 40 kV and 35 mA; soller slit of 0.04 rad; fixed divergent slit of $\frac{1}{4}^\circ$ and $\frac{1}{2}^\circ$ anti-scattering slit; 0.026° 2 θ step size; scanning range from 5° to 70° 2 θ . Diffractograms were evaluated using the PANalytical software HighScore Plus 4.8.

The XRD data were initially used to identify the mineralogical phases and then to predict the parameters $AvAl_2O_3$ and $RxSiO_2$ using PCA and PLSR. In the latter case, all diffractograms were arranged in an $m \times n$ matrix, where m are the 142 samples analyzed, and n is taken as the dependent variables which store the values of intensity for each 2 θ step (Fig. 2). Thus, the total of n variables depends on the measurement 2 θ range analyzed and step-size (in this work these two parameters varied as shown in Table 1).

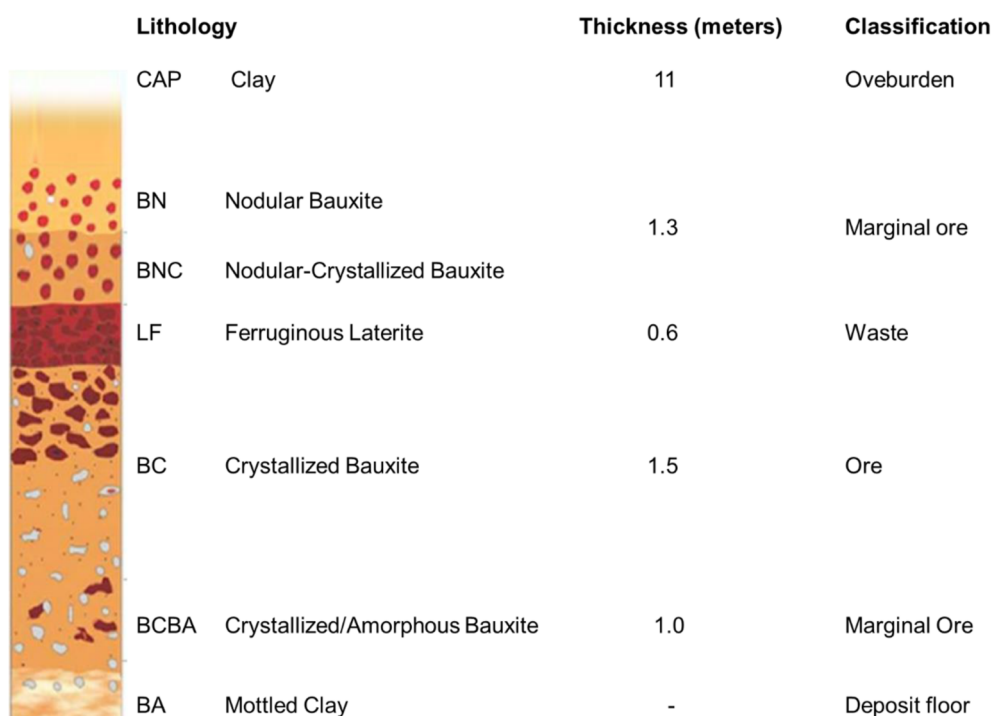


Fig. 1. Typical geological profile of Miltonia Plateau. Adapted of Silva et al., 2008.

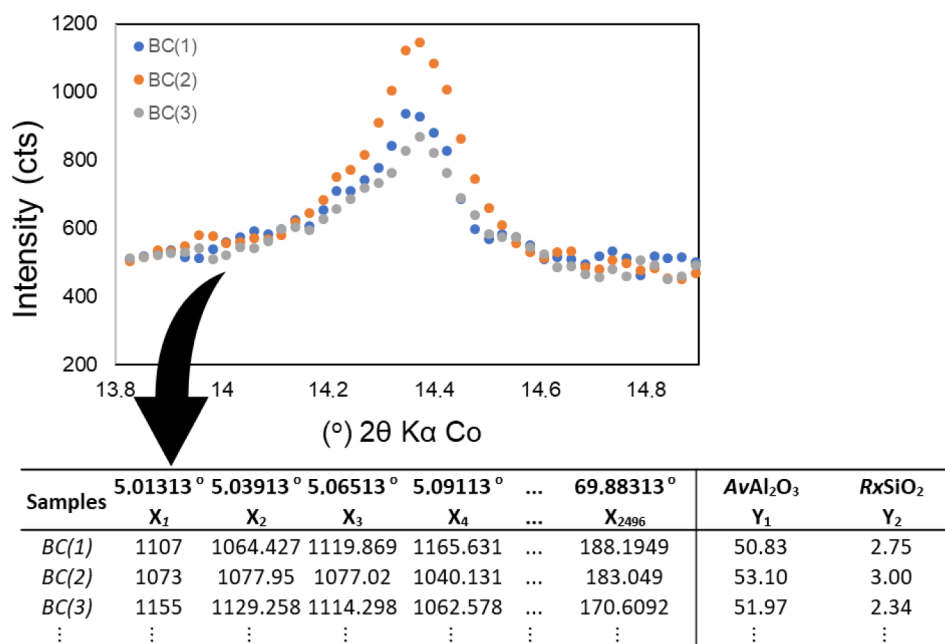


Fig. 2. Representation of the dataset. Diffractograms of the samples as arranged as matrix.

2.3. Principal component analysis (PCA)

Principal component analysis is a multivariate technique widely used in several fields of application – from neuroscience to computer graphic. Basically, it aims the reduction of data dimensionality in such a way that relevant information can be extracted from extensive and confusing datasets (Jolliffe, 2002; Shlens, 2005). A comprehensive mathematical description, as well as practical examples of PCA application in chemometrics can be found in Cordella (2010) and Bro & Smilde (2014). The utilization of PCA on XRD data to identify clusters of samples and correlate them with mineralogical information have been reported in recent years (Macchiarola et al., 2007; König et al., 2012; Gobbo et al., 2013; König et al., 2014; Melo et al., 2017). In this work, PCA was applied as a data mining tool, in order to find sample clusters, identify outliers and the diffractogram 2θ range that mostly describes the variance among the bauxite samples.

PCA analyzes were performed using the chemical reference values (AvAl₂O₃ and RxSiO₂) and XRD patterns of the samples. In the first case, the dataset was previously scaled since the variables are in different ranges (AvAl₂O₃: 31.4–59.3% and RxSiO₂: 0.52–8.42%). Geladi & Kowalsky (1986) comment that spectroscopy data do not need to be scaled since all observations are already at the same scale, thus, in this case similar consideration was made for the XRD data, which were only mean-centered.

2.4. Partial Least Squares Regression (PLSR)

Partial Least Square Regression (PLSR) is a robust multivariate technique used to predict hidden information from raw data, being a good alternative to Multiple Linear Regression (MLR) and Principal Component Regression (PCR). PLSR is widely used in chemometrics since it is a well-suited method for handling with multicollinear data (high correlated observations) – such as Infrared spectroscopy, Raman and X-ray diffraction measurements (Wold et al., 1984; Wold et al., 2001).

Mathematically, the analysis consists in initially modelling by principal components the data matrices X (independent variables – predictors) and Y (dependent variables – responses) as outer relations:

$$X = TP' + E = \sum t_h p'_h + E$$

and

$$Y = UQ' + F = \sum u_h q'_h + F$$

where T and U are scores, P and Q are loadings, and E and F are residuals matrices, and the subscript h represent the number of components. The particularity of PLSR is that, an inner relation between the X -score (t) and the Y -score (u) for every component, simultaneously take place:

$$u_h = b_h t_h$$

where b_h represents the PLSR coefficients for each component and its solution is given by least-squares $b_h = u'_h t_h / t'_h t_h$. It means that for each component, t has maximum covariance with u . Hence, it gives the mixed relation:

$$Y = TBQ' + F = \sum b_h t_h q'_h$$

In a way that F is to be iteratively minimized for every component considered in the model (Wold, 1966; Geladi & Kowalsky, 1986).

In practical applications, the property of interest Y usually relies on expensive, difficult and/or time-consuming methods or measures, while X is generally a cheap and/or fast measure. Thus, one can easily predict the desired property through the X data.

A “leave-one-out” cross-validation was used to find the best numbers of factors (components) in the models and the statistics of residuals (Root Mean Square Error of Cross-Validation – RMSECV, and Root Mean Square Error of Prediction – RMSEP) were monitored to assess the prediction quality. The best PLSR models to predict AvAl₂O₃ and RxSiO₂ were found after defining the optimized pre-processing of XRD data through DOE and clustering the samples by PCA.

2.5. 2⁴ Design of Experiments (DOE)

In order to optimize the quality and the time of collection of XRD data to be used in the PLSR, some XRD data parameters were evaluated using a factorial DOE 2⁴. Tables 1 and 2 show the factors and levels utilized and the design matrix in planning order. For this optimization, 42 samples were used to calibrate the PLSR models and 4 samples (with 5 replicates) to test them. These samples were extracted from the dataset (142 samples) and correspond only to the BC lithology samples. Thus, for each designed condition a PLSR model for AvAl₂O₃ and RxSiO₂ were constructed. The best number of PLSR factors were

Table 1
DOE 2⁴ factors and their levels.

Factors	Levels	
	−1	+1
Type of Divergent Slit ^a	FDS	ADS
Step-size (°)	0.026	0.065
XRD range (° 2θ)	5 – 70	13 – 34
Scaling mode ^b	Mean-Centered	Standardized

^a Fixed Divergent Slit – FDS and Automatic Divergent Slit – ADS.

^b Standardized by dividing each variable for its standard deviation. In this case the dataset is also mean-centered - *autoscaling*.

Table 2
DOE 2⁴ design matrix with coded factors.

Divergent slit ^a	Step-size (°)	XRD range (° 2θ)	Scaling mode ^b
−1	−1	−1	−1
+1	−1	−1	−1
−1	+1	−1	−1
+1	+1	−1	−1
−1	−1	+1	−1
+1	−1	+1	−1
−1	+1	+1	−1
+1	+1	+1	−1
−1	−1	−1	+1
+1	−1	−1	+1
−1	+1	−1	+1
+1	+1	−1	+1
−1	−1	+1	+1
+1	−1	+1	+1
−1	+1	+1	+1
+1	+1	+1	+1

selected by “leave-one-out” cross-validation and after testing the models, the RMSEP was taken as DOE response variable. The optimized experimental conditions, which resulted in the lowest RMSEP, were fixed to perform the PLSR analyses on the entire data set.

3. Results and discussion

The XRD pattern of the bauxite samples is shown in Fig. 3. It is observed the same mineralogical phases present in all bauxite lithologies: gibbsite as the major phase ($d_{002} = 4.84 \text{ \AA}$ and $d_{110} = 4.37 \text{ \AA}$), kaolinite ($d_{001} = 7.14 \text{ \AA}$ and $d_{002} = 3.56 \text{ \AA}$), hematite ($d_{104} = 2.69 \text{ \AA}$),

Al-goethite (d_{101} peak varying from $4.12 \text{ \AA}$ to $4.16 \text{ \AA}$ due to aluminium isomorphic substitution) (Neumann et al., 2014) and anatase ($d_{011} = 3.52 \text{ \AA}$).

3.1. Data evaluation by PCA

Principal Component Analysis was carried out on the training set (120 samples) for both chemical reference values and XRD patterns. The PCA scores and loadings from the reference values are shown in Fig. 4. The first component (PC1) explains 96% of variance in the data and it is negatively related to the Al_2O_3 content and positively related to R_xSiO_2 . Comparing both plots one can note that among the lithologies BN has the largest chemical dissimilarity. Due to this low $\text{Al}_2\text{O}_3/\text{SiO}_2$ ratio, this upper horizon is considered mining waste material, which is removed and then backfilled in the mining site together with the overburden clay material (Silva et al., 2008). Eventually, some amount of BN may be present in the ore by blending or when the underlying lithology BNC is mined and there is horizontal profile discontinuity and/or their limits are not well defined. Melo et al. (2017) demonstrated that the kaolinite present in this lithology is significantly more reactive than those present in BC (ore horizon).

The scatter plot of scores and loadings of PCA on XRD data are presented in Fig. 5, and the scree plot of the eigenvalues for each principal component is depicted in Fig. 6. As observed, this data can be described by the two first components, which account for 98% of the variance. In the score plot (Fig. 5a) three clusters of samples are highlighted (named C1, C2 and C3), where C1 and C2 are systematically spread from each other. It means that the part of the XRD patterns (observations) that mostly explain the variation in data has an asymmetrical distribution, as observed in loadings plot (Fig. 5b).

Interestingly, these clusters are not discriminating the lithologies as described by Melo et al. (2017). Instead, the scores plot shows that there is some aspect in the diffractograms of these samples that describes a heterogeneity of each lithology in the bauxite profile. In fact, it is explicitly true for BCBA and BC, and such mineralogical behavior is not detected by the chemical composition. Some authors have demonstrated the presence of several generations of gibbsite and kaolinite along these ore horizons (Kotschoubey et al., 2005; Pereira, 2015). Since the PCA uses the full XRD pattern, not only the content of each phase is taken into account to discriminate the samples, but also their crystalline characteristics.

Fig. 7 depicts the loadings accounted for PC1 and PC2. This linear representation helps to understand the region of the XRD pattern which

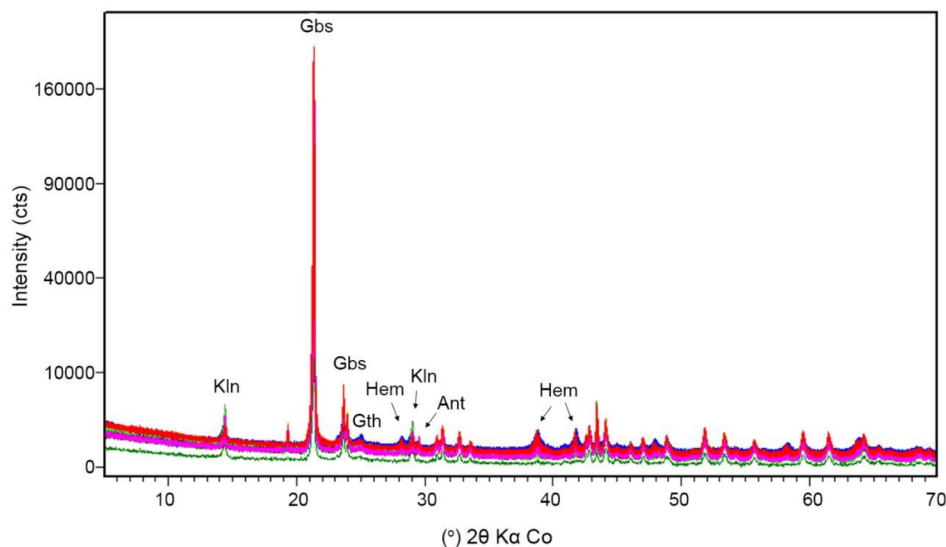


Fig. 3. XRD pattern of all bauxite samples. Legends: Kln: kaolinite; Gbs: gibbsite; Gth: Al-goethite; Ant: anatase; Hem: hematite.

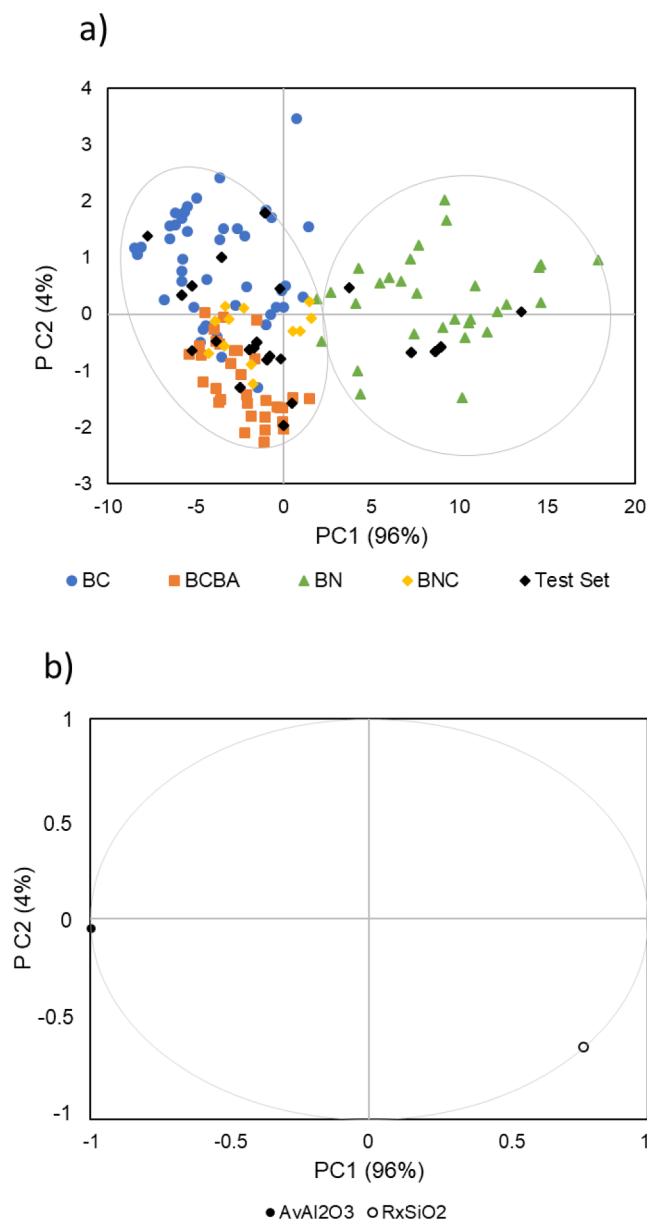


Fig. 4. Scatter plots of PCA a) Scores and b) Loadings of the chemical reference dataset.

describes the principal components. The variation in the PC1 is mainly affected by the basal reflections of gibbsite $d_{002} = 21.25$ and $d_{004} = 43.38$ ($^{\circ} 2\theta$). It is also observed that undesired parts of the diffractogram, such as gibbsite k_{β} peak appearing in 19.28 ($^{\circ} 2\theta$) and the background noise in angles below to 12° (2θ) are explaining a significant part of variance in PC1.

The influence plot (Fig. 8a) shows the outlier samples. By comparing this plot with scores and loadings, one can note that the largest variance accounted in PC1 is mainly related to two high leverage samples (BC(10) and BCBA(22)). In this samples the intensities of gibbsite basal reflections, d_{002} and d_{004} , are too high comparing to the average (Fig. 8b). In this case, both samples need to be removed from the calibration set to avoid further error of prediction. The other samples with small leverage and residuals were kept in the calibration set, since it is assumed that such deviation is inherent to the bauxite mineralogy and sample preparation.

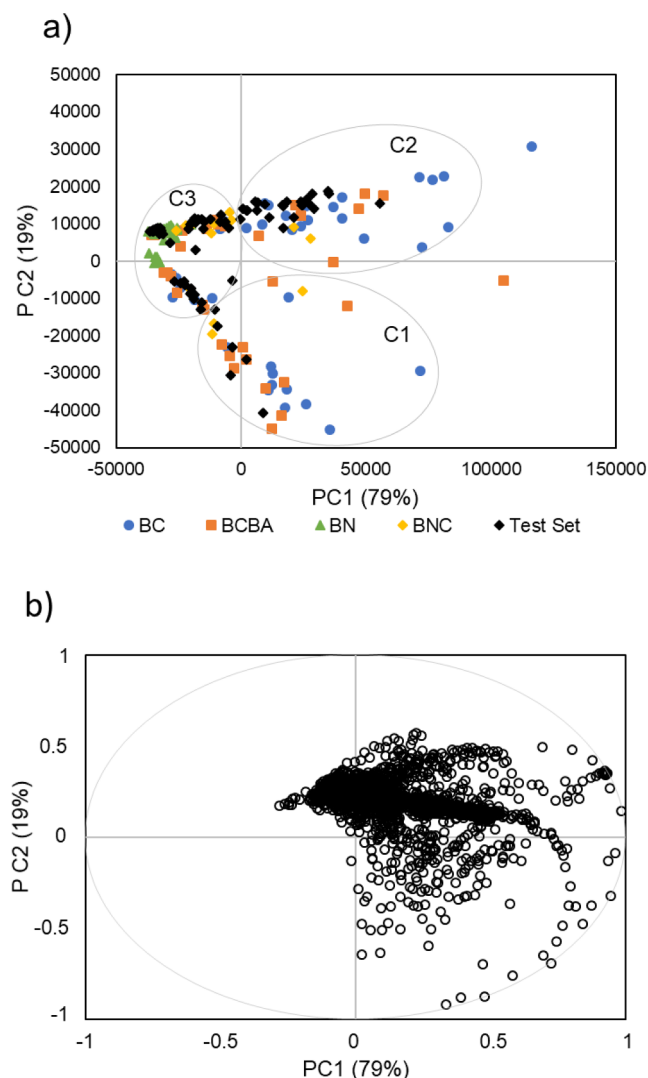


Fig. 5. Scatter plots of PCA a) scores (scores of the test set is also depicted), and b) loadings of the XRD dataset.

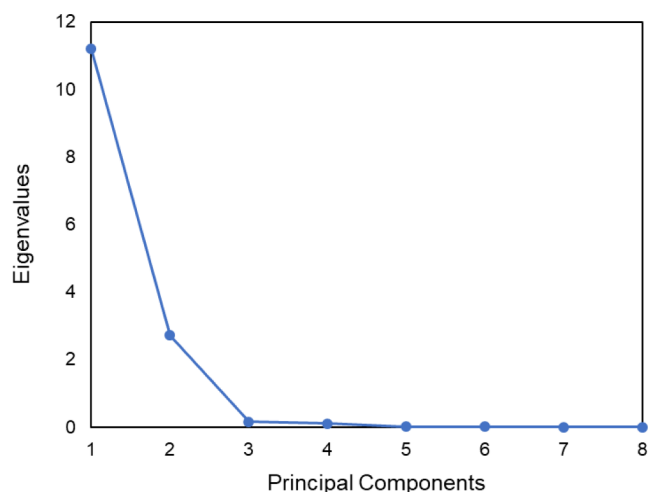


Fig. 6. Scree Plot of the eigenvalues. The first two principal components explain 98% of the data variance.

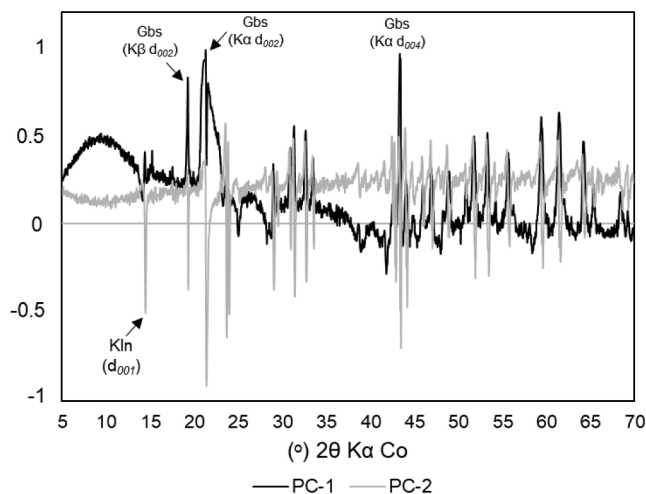


Fig. 7. Linear representation of PCA loadings of the XRD dataset. Legend: Kln = kaolinite; Gbs = gibbsite.

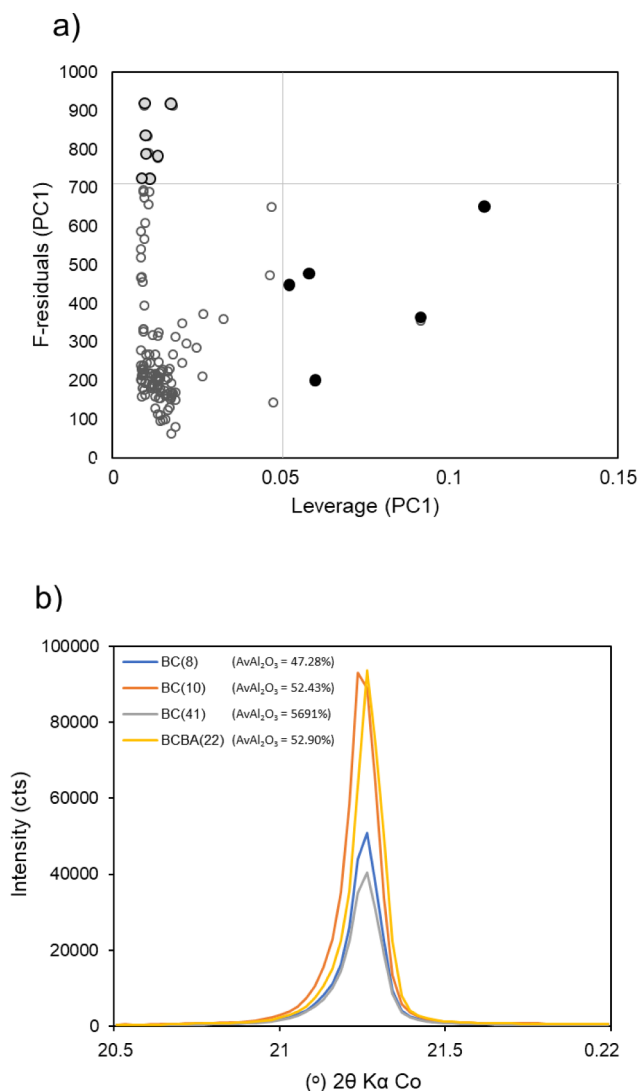


Fig. 8. a) Influence plot. Black dots in the lower-right quadrant represents samples with high leverage; gray dots in upper-left are samples with large residuals; b) comparison of XRD patterns of high leverage samples (diffractogram region of gibbsite d_{002} reflection).

Table 3

Time estimate for XRD measurements. (based on the present diffractometer features).

Step-size	5°–70° (2θ)	13°–34° (2θ)
0.026°	4 min 30 s	1 min 30 s
0.065°	1 min 50 s	40 s

3.2. Evaluation of pre-processing the XRD data

A PLSR was carried out to define the best conditions of pre-processing the XRD data following a factorial DOE. In order to use PLSR and XRD in the bauxite quality control it is expected that this methodology meets the statistical requirements of accuracy and precision, but also the time of analysis is fast.

In terms of XRD data acquisition, by reducing the 2θ interval and increasing the goniometer step-size, the data collection time is significantly reduced, in this particular case, 40 s (Table 3). Thus, the entire procedure for determining the bauxite quality control parameters could take less than five minutes per sample. This means a substantial improvement for the industry, which currently takes more than one hour in these analyses by the traditional method.

To avoid intensity loss at high 2θ angles and to reduce the high background intensity at low angles, it is commonly used an Automatic Divergent Slit (ADS) which ensure a constant irradiated area on the sample during the data acquisition. In this study a mathematical transformation was applied instead in the XRD data collected using a Fixed Divergent Slit (FDS), as follows (Jenkins & Snyder, 1996):

$$L = \frac{2RTan(\alpha/2)}{\sin\theta}$$

where L is the irradiation length, R is the goniometer radius and α is the divergence angle. It is worthy to note that these three pretreatments of data were defined after discussion of the PCA results.

Mean-centering the data is the most common preprocess applied. It makes the calculations easier and the model less complex by eliminating the need to fit a non-zero intercept (Geladi & Kowalsky, 1986; Ferraro et al., 2001). Scaling the data by the standard deviation is a tailor requirement usually applied when the variables are in different scales (e.g. %, ppm, °C). It turns the weights of each variable equals 1, which means that they have same opportunity to be modelled (Bro & Smilde, 2014).

The DOE results for $RxSiO_2$ estimation are summarized in the ANOVA table (Table 4). Interestingly, the best results were achieved autoscaling the data, where usually 5 factors are enough to predict with lower RMSEP. In contrast, mean-centering usually needs 8 factors which, despite obtaining a lower RMSECV, when predicting the samples of the test set, the RMSEP tends to be higher due to overfitting (Fig. 9) (Faber et al., 2003; Ferraro et al., 2001).

Table 4

ANOVA of the $RxSiO_2$ estimation in the 2^4 DOE. (RMSEP as response variable; $R^2_{adj} = 0.935$; level of significance $\alpha = 0.05$).

Factors	SS	df	MS	F	p-value
(1) Divergent Slit	0.000052	1	0.000052	0.1636	0.702558
(2) Step-size	0.000199	1	0.000199	0.6244	0.465226
(3) XRD 2θ range	0.008892	1	0.008892	27.9375	0.003230
(4) Scaling mode	0.060896	1	0.060896	191.3330	0.000035
1*2	0.001098	1	0.001098	3.4498	0.122386
1*3	0.000352	1	0.000352	1.1073	0.340848
1*4	0.000052	1	0.000052	0.1629	0.703140
2*3	0.002428	1	0.002428	7.6301	0.039728
2*4	0.000046	1	0.000046	0.1430	0.720858
3*4	0.003303	1	0.003303	10.3790	0.023423
Error	0.001591	5	0.000318		
Total SS	0.078909	15			

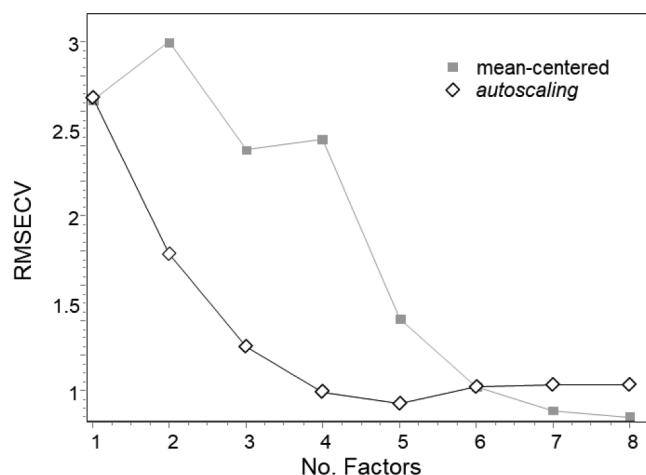


Fig. 9. Comparison between data pre-processing methods – mean-centered (RMSECV = 0.845, RMSEP = 1.295, with 8 factors) and autoscaling (RMSECV = 0.924 and RMSEP = 0.804, with 5 factors).

In addition, it was observed that reducing the XRD 2θ range to only $13\text{--}34^\circ$ leads to faster measurement and more accurate prediction. This pre-processing reduced the number of independent variables in the PLSR models from 2496 to 807, which means that the largest amount of information in the XRD pattern accounts only for the noise.

3.3. Estimation of AvAl_2O_3 and RxSiO_2 by PLSR

Based on the clusters identified in PCA the dataset was divided to generate the PLSR models and evaluate the quality of the AvAl_2O_3 and RxSiO_2 predictions. The dataset was also pre-processed according to the best conditions found in DOE: XRD 2θ range: $13\text{--}34^\circ$, step-size: 0.065° , divergent slit and autoscaling. Fig. 10 depicts the correlation between the reference values (standard wet chemistry) and the estimated using PLSR models. As observed the samples clustered in C1 and C2 that, in general, have the content of AvAl_2O_3 and RxSiO_2 in a range of 49–55% and 2.5–5.5%, respectively, are good modelled. In contrast, most samples in C3 were not satisfactory satisfactorily predicted.

The accuracy of each model was verified by their RMSEP, given in the Table 5, and the average of residuals in Fig. 11. The results showed good accuracy for both parameters when predicted by models C1 and C2 (RMSEP < 0.85). The residuals close to zero for all models means that they are not statistically biased, and the residual standard deviation (RSD) lower than 1% shows a good precision. The prediction of RxSiO_2 for C3 samples were reasonably accurate. However, a large deviation is observed when trying to predict AvAl_2O_3 . Industrially, bauxite quality control requires a precision of < 0.5% for reactive silica and 1–1.5% for available alumina. Feret (2018) argue that the real accuracy of wet chemistry methods cannot be determined since they are empirical and there is no independent reference method, and it is often difficult to attain this precision limit in the industrial routine.

Following the argument of Wold et al. (2001), the diffractograms of each sample can be seen as the sum of the reflections of the phases present multiplied by their concentrations in the sample. This means that in PLSR modelling, the latent variable hidden in $\mathbf{X}$ (in this case, to predict AvAl_2O_3 and RxSiO_2) is explained by the crystalline characteristics of the phases in the XRD pattern (identified as $\mathbf{p}$ – loadings) and the concentration of each mineral phase related to this variable (identified as $\mathbf{t}$ – scores): $\mathbf{X} = \mathbf{t}_1\mathbf{p}_1' + \mathbf{t}_2\mathbf{p}_2' + \dots = \mathbf{TP}' + \text{noise}$. As example, the Fig. 12 shows the linear representation of the loadings for RxSiO_2 in C3 model. It can be observed how much the intensities of each phase, as well as the variation of mineralogy are accounting for predictors. In the latter case, two generations of gibbsite were distinguished by the loadings of the model. König et al. (2012) and Negrão et al. (2018)

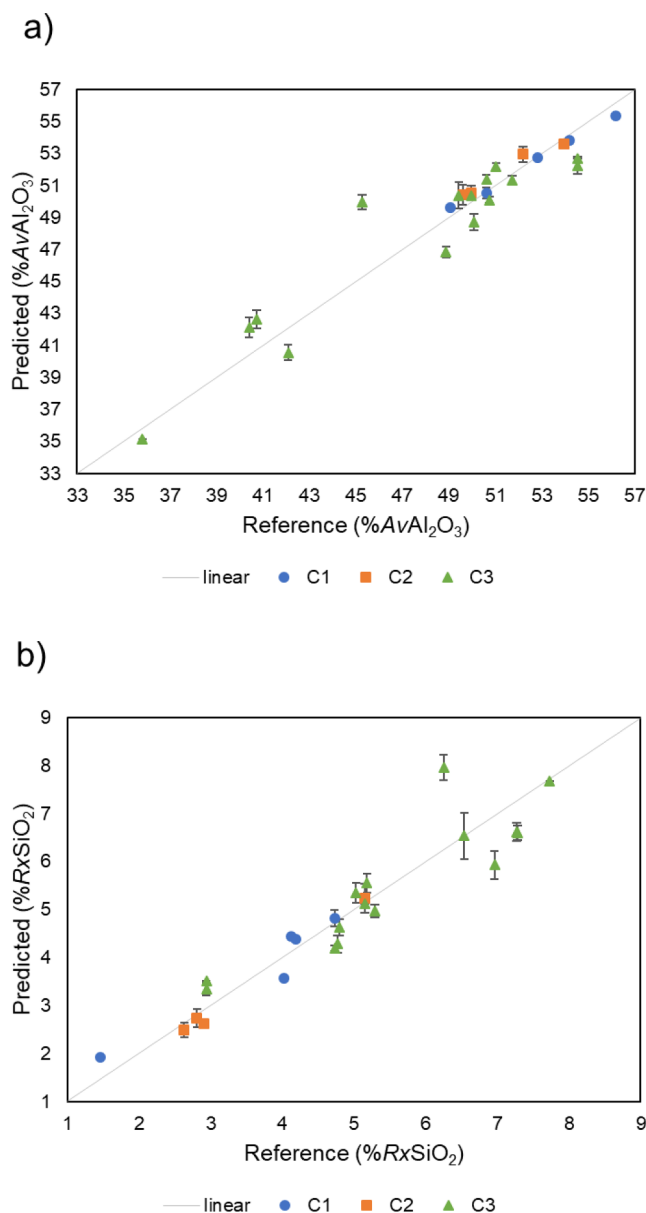


Fig. 10. Comparison between estimated and reference a) AvAl_2O_3 , b) RxSiO_2 . Bars indicating the RSD (residual standard deviation).

Table 5
RMSEP of AvAl_2O_3 and RxSiO_2 for the cluster models.

Model	RMSEP	
	AvAl_2O_3	RxSiO_2
C1	0.493	0.326
C2	0.834	0.230
C3	1.704	0.641

described these phases as a coarse high crystalline phase filling the bauxite microvoids and a fine-grained and low crystalline disperse in the matrix.

In fact, the peak intensities (I_{hkl}) of a phase α in a diffractogram is related to its volume fraction (v_α) according to the simplified equation:

$$I_{(hkl)\alpha} = \frac{K_e \cdot K_{(hkl)\alpha} \cdot v_\alpha}{\mu_s}$$

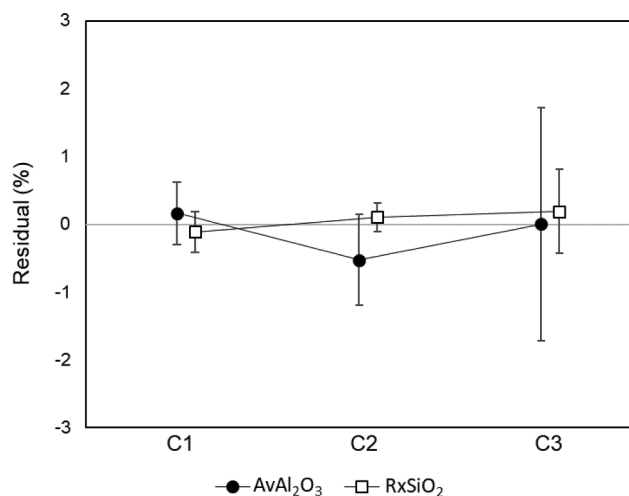


Fig. 11. Plot of average of the residuals for each cluster model. Bars indicating the RSD.

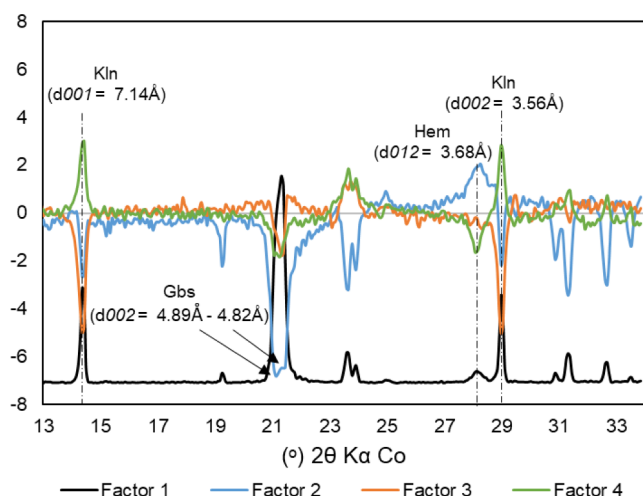


Fig. 12. Linear representation of PLSR loadings of the C3 calibration set for $RxSiO_2$ model. Legend: Kln = kaolinite; Gbs = gibbsite; Hem = hematite.

where K_e is a constant for a particular experimental system; $K_{(hkl)\alpha}$ are constants for each diffraction reflection hkl of phase α ; and μ_s is the attenuation coefficient (Jenkins & Snyder, 1996). Based on that, several methods can be used to estimate the mineralogical composition of the sample, where the principal is Rietveld refinement. The application of this method in the bauxite quality control is being constantly developed, where inaccuracies due to amorphous content and isomorphous substitution was substantially improved (Neumann et al., 2014; Paz et al., 2018; Angélica et al., 2018). In spite of that, this method is not yet applied in the bauxite industry, mainly due its difficulty of implementation – requiring an advanced crystallographic knowledge and a tailored sample preparation. In addition, it is essential a good XRD data collection which would result in a time-consuming measurement.

In contrast, using the PLSR method based on XRD data it is possible to obtain a quick and reliable estimate of bauxite quality control parameters. This can be easily implemented industry (fully automated), there is no need to know the crystallography of the sample, and robust XRD data collection is not required – diffractograms with low peak resolution and very small measurement range are enough.

A point of attention is still the sample preparation. To demonstrate it, Fig. 13 shows XRD patterns of C3 test samples in the range of the

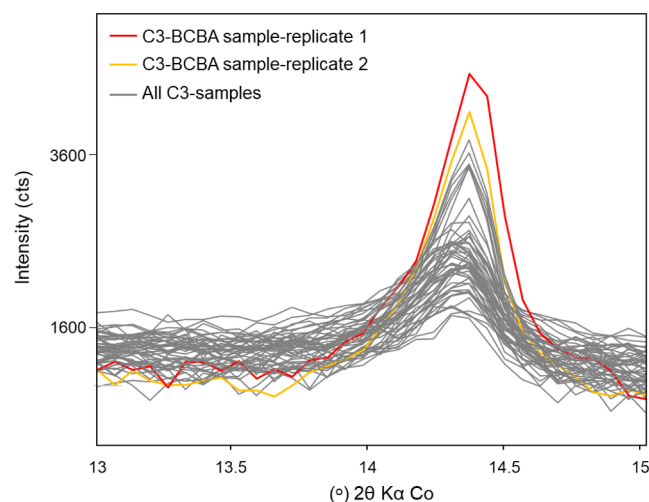


Fig. 13. XRD pattern of C3 samples – depicting the range of kaolinite d_{001} reflection.

kaolinite d_{001} reflection. The highlighted duplicate BCBA samples presented worst prediction (replicate 1 = 8.14%, replicate 2 = 7.76%, reference value = 6.25%, see Fig. 8b). In this case, the BCBA sample has a higher crystalline kaolinite (measured by the *Full-Width-at-Half-Maximum* – FWHM ~ 0.200 ; for comparison, the average FWHM for all C3 samples is ~ 0.300). This resulted in a strong orientation in the $00l$ planes, and consequently, a false correlation between intensities and kaolinite content. According to König et al. (2019), PLSR is less sensitive to non-ideal sample preparation compared to Rietveld refinement, however, the conditions should be constant – mainly the sample particle size and pressure applied. Considering that the samples in this work were prepared manually, it is believed that a large part of the prediction error of C3 samples was due to a non-uniform preparation condition.

4. Conclusions

In order to achieve a rapid and reliable method to estimate the parameters of bauxite quality control ($AvAl_2O_3$ and $RxSiO_2$), The use of PLSR on XRD data on gibbsitic bauxites from Miltonia 3 plateau (Paragominas region, northern Brazil) was investigated as an alternative to current wet chemistry method. An optimization of the dataset pre-processing was applied. Through PCA it was possible to identify outliers and clusters of samples with similar mineralogy. The best condition of XRD data collection and scaling the dataset was found using DOE. In this case, fast measurement (~ 40 s) was attained by reducing the 2θ range ($13\text{--}34^\circ$) and step-size (0.065°).

These conditions applied to test samples resulted in good predictions in terms of accuracy and precision for samples classified in the clusters C1 and C2 (RMSEP < 0.85 and very low bias). For samples in cluster C3 the prediction was hampered, probably due to their non-uniform sample preparation, associated with variations in the degree of crystallinity of kaolinite and gibbsite present.

The results demonstrated that the proposed method has the potential to be used industrially for bauxite quality control. Compared to the traditional method, it is faster, easier to implement and perform the analyses, requires less laboratory space and manpower, and no reagent is needed. In addition, the use of XRD for quality control makes it possible to track the mineralogy of the bauxite that feeds the Bayer process, and therefore, be aware of how it could impact the process. Such impact cannot be known by controlling the chemical composition alone.

CRediT authorship contribution statement

Caio César Amorim Melo: Conceptualization, Methodology, Validation, Investigation, Formal analysis, Data curation, Writing - original draft, Writing - review & editing. **Rômulo Simões Angélica:** Conceptualization, Methodology, Resources. **Simone Patrícia Aranha Paz:** Conceptualization, Validation, Resources, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors thank the Brazilian agencies: CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for a PhD scholarship to the first author; CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for financial support (EditalMCT/CT-Mineral/VALE/CNPq N° 12/2009, 550.297/2010-3); and FAPESPA (Fundação Amazônia de Amparo a Estudos e Pesquisas), Edital 01/2010, ICAAF No. 027/2011, also for financial support. We also acknowledge the Mineração Paragominas SA Company (Norsk Hydro) for the support.

References

- Angélica, R.S., Kahn, H., Paz, S.P.A., 2018. A proposal for bauxite quality control using the combined Rietveld – Le Bail – Internal Standard PXRD method – Part 2: Application to a gibbsitic bauxite from the Paragominas region, northern Brazil. *Miner. Eng.* 122, 148–155.
- Authier-Martin, M., Forte, G., Ostap, S., See, J., 2001. The mineralogy of bauxite for producing smelter-grade alumina. *J. Met.*, 2001, 36–40.
- Bredell, J.H., 1983. Calculation of available alumina in bauxite during reconnaissance exploration. *Econ. Geol.* 78, 319–325.
- Bro, R., Smilde, A.K., 2014. Principal component analysis. *Anal. Methods* 6, 2812–2831. <https://doi.org/10.1039/c3ay41907j>.
- Carneiro, C. C., Yanez, D. N. D. V. S., Ulsen, C., Fraser, S. J., Antoniassi, J. L., Paz, S. P. A., Angélica, R. S., Kahn, H., 2017. Imputation of reactive silica and available alumina in bauxites by self-organizing-maps. In: *Anais. Nancy: IEEE*, 2017. Disponível em: DOI: 10.1109/WSOM.2017.8020008.
- Cordella, C. B. Y., 2010. PCA: The Basic Building Block of Chemometrics. *Analytical Chemistry*. IntechOpen. Nov 7. DOI: 10.5772/51429.
- Faber, N.M., Song, X.-H., Hopke, P.K., 2003. Sample-specific standard error of prediction for partial least squares regression. *Trends Anal. Chem.* 22 (5), 330–334.
- Feret, F.R., Authier-Martin, M., Sajó, I., 1997. Quantitative phase analysis of Bidi-Koum bauxites (Guinea). *Clays Clay Miner.* 45–3, 418–427.
- Feret, F. R., 2018. Inaccuracies in Estimation of Bauxite Extractable and Mineralogical Constituents. *Proceedings of the 36th International ICSOBA Conference*, Belém, Brazil, 29 October - 1 November, 2018, 155–168.
- Ferraro, M.C.F., Castellano, P.M., Kaufman, T.S., 2001. A spectrophotometric - partial least squares (PLS-1) method for the simultaneous determination of furosemide and amiloride hydrochloride in pharmaceutical formulations. *J. Pharm. Biomed. Anal.* 26, 43–451.
- Freyssinet, Ph, Butt, C.R.M., Morris, R.C., Piantone, P., 2005. Ore-forming process related to lateritic weathering. *Econ. Geol.* 100, 681–722.
- Geladi, p., Kowalski, B. R., 1986. Partial least-squares regression: a tutorial. *Analytica Chimica Acta*. 185, 1–17.
- Gobbo, L. A., Montanheiro, T. J., Montanheiro, F., Sant'Agostino, L. M., 2013. Avaliação de cimento portland a partir da difração de raios X associada à análise por agrupamento. *Revista do Instituto Geológico*, São Paulo, 34 (2), 41–49, DOI: 10.5935/0100-929X.20130008.
- Jenkins, R., Snyder, R. L. *Introduction to X-Ray Powder Diffractometry*. John Wiley and Sons, New York, 1996, 403 + XXIII pp.
- Jolliffe, I.T., 2002. *Principal Component Analysis*, 2nd ed. Springer-Verlag, New York.
- König, U., Angélica, R. S., Norberg, N., Gobbo, L., 2012. Rapid X-ray diffraction (XRD) for grade control of bauxites. *International Committee for Study of Bauxite, Alumina & Aluminium – ICSOBA*, November 2012, Belém. *Papers, Bauxite Program BX14-T*, 11 p., CD-ROM.
- König, U., Degen, T., Norberg, N., 2014. PLSR as new XRD method for downstream processing of ores – case study: Fe2+ determination in iron ore sinter. *Powder Diffraction*, 29, 78–83.
- König, U., Degen, T., Norberg, N., 2019. New tools for process control in aluminium industries - PLSR on XRD raw data. In: *AQW Conference*. Available at: < <http://www.aqw.com.au/papers/item/new-tools-for-process-control-in-aluminium-industries-plsr-on-xrd-raw-data.html> > .
- Kotschoubey B., Calaf J.M.C., Lobato A.C.C., Leite A.S. Azevedo C.H.D., 2005. Caracterização e gênese dos depósitos de bauxita da Província Bauxítica de Paragominas, noroeste da bacia do Grajaú, nordeste do Pará/ oeste do Maranhão. In: *MARINI, O. (Org.). 2005. Caracterização em depósitos minerais em distritos mineiros da Amazônia*. Brasília, DF: DNPM-CT/MINERAL-ADIMB. p. 687–782.
- Macchiarola, K., König, U., Gobbo, L., Campbell, I., McDonald, A.M., Cirelli, J., 2007. Modern X-ray Diffraction Techniques for Exploration and Analysis of Ore Bodies. In: *Proceedings of Exploration 07: Fifth Decennial International Conference on Mineral Exploration*, p. 1007–1011.
- McCusker, L.B., Von Dreele, R.B., Cox, D.E., Louër, D., Scardi, P. (1999). Rietveld refinement guidelines. *J. Appl. Cryst.* 32, 36–50.
- Melo, C.C.A., Oliveira, K. S., Angélica, R. S., Paz, S. P. A., 2017. Análise de agrupamentos associada a difratometria de raios-X: uma classificação mineralógica prática de bauxitas e seus produtos de digestão Bayer. *Holos*, ano 33, Vol. 06, 32–42.
- Molai, H., 2011. Role of mineralogy and geochemistry in the beneficiation of Jajarm Bauxite from North East Iran: comparison with some other bauxite deposits of the world. *Iran. J. Earth Sci.* 3, 134–167.
- Najar, M. P. A., 2017. Rapid Analysis of Available Alumina and Reactive Silica in Bauxite. *Minerals & Metals Review - January* 2017, 99–100.
- Negrão, L.B.A., Costa, M.L., Pöhlmann, H., Horn, A., 2018. Na application of the Rietveld refinement method to the mineralogy of bauxite-bearing regolith in Lower Amazon. *Mineral. Mag.* 82 (2), 413–431.
- Neumann, R., Avelar, A.N., Costa, G.M., 2014. Refinement of the isomorphic substitutions in goethite and hematite by the Rietveld method, and relevance to bauxite characterization and processing. *Miner. Eng.* 55, 80–86.
- O'Connor, D.J., 1988. *Alumina Extraction from Non Bauxitic Materials*. Sydney: Aluminium-Verlag, 370 p.
- Ostap, S. 1984. Effect of bauxite mineralogy on its processing characteristics, Bauxite, *Proceedings of the 1984 Bauxite Symposium*, Los Angeles, California, Feb. 27 – March 1, 1984, Chapter 30, p. 651.
- Paz, S.P.A., Angélica, R.S., Kahn, H., 2017. Optimization of the reactive silica quantification method applied to Paragominas-type gibbsitic bauxites. *Int. J. Miner. Process.* 162, 48–57.
- Paz, S.P.A., Kahn, H., Angélica, R.S., 2018. A proposal for bauxite quality control using the combined Rietveld – Le Bail – Internal Standard PXRD method – Part 1: hkl model developed for kaolinite. *Miner. Eng.* 118, 52–61.
- Picanço, F. E., Pinheiro, D. C., Calado, W. M., Almeida, K. R., Quadros, C. R. R., Rabelo, S. A., 2013. Amostragem em diferentes malhas de sondagem para minério de bauxita – mina de Miltonia 03 – Paragominas-PA. *Anais do 13º Simpósio de Geologia da Amazônia*, September, 2013, 645–647.
- Shlens, J., 2005. A Tutorial on Principal Component Analysis. *arXiv:1404.1100*.
- Silva, H. M., Picanço, E., Maurity, C., Morais, W., Santos, H. C., Guimarães, O., 2008. Geology, mining operation and scheduling of the Paragominas bauxite mine. *Proceedings of the 8th International Alumina Quality Workshop*, 11–16.
- Pereira, C.B., 2015. *Mineralogia e geoquímica dos perfis bauxíticos da mina Miltonia 3, região de Paragominas-PA*. Master Dissertation. Geoscience Institute, Federal University of Pará.
- Smith, P., 2009. The processing of high silica bauxites - review of existing and potential processes. *Hydrometallurgy* 98, 162–176.
- Wold, H., 1966. 1966. Estimation of principal components and related models by iterative least squares. In: *Krishnaiah, P.R. (Ed.), Multivariate Analysis*. Academic Press, New York, pp. 391–420.
- Wold, S., et al., 1984. The collinearity problem in linear regression, the partial least squares approach to generalized inverses. *SIAM. J. Sci. Stat. Comput.* 5, 735–743.
- Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. *Chemometr. Intell. Lab. Syst.* 58, 109–130.
- Zhang, R., Gong, E., Wang, G., Yang, Z., 2019. An analytical shortcut to estimate alumina content by LOI in lateritic gibbsite bauxite prospecting. In: *AQW Conference*. Available at: < <http://www.aqw.com.au/papers/item/an-analytical-shortcut-to-estimate-alumina-content-by-loi-in-lateritic-gibbsite-bauxite-prospecting.html> > .