

Universidade Federal do Rio Grande do Sul
Programa de Pós-Graduação em Ciência dos Materiais

**OTIMIZAÇÃO DE PROPRIEDADES FÍSICAS DE PONTOS QUÂNTICOS E
NANOPARTÍCULAS DE SEMICONDUTORES POR INTELIGÊNCIA
ARTIFICIAL E PLANEJAMENTO EXPERIMENTAL**

Fábio Baum

Defesa de doutorado

Porto Alegre, setembro de 2020.

Universidade Federal do Rio Grande do Sul
Programa de Pós-Graduação em Ciência dos Materiais

**OTIMIZAÇÃO DE PROPRIEDADES FÍSICAS DE PONTOS QUÂNTICOS E
NANOPARTÍCULAS DE SEMICONDUTORES POR INTELIGÊNCIA
ARTIFICIAL E PLANEJAMENTO EXPERIMENTAL**

Fábio Baum

Tese realizada sob a orientação do Prof. Dr. Marcos José Leite Santos, apresentada ao Programa de Pós-Graduação em Ciência dos Materiais da Universidade Federal do Rio Grande do Sul em preenchimento parcial dos requisitos para a obtenção do título de Doutor em Ciência dos Materiais.

Defesa de doutorado

Porto Alegre, setembro de 2020.

CIP - Catalogação na Publicação

Baum, Fábio

Otimização de propriedades físicas de pontos quânticos e nanopartículas de semicondutores por inteligência artificial e planejamento experimental / Fábio Baum. -- 2020.

143 f.

Orientador: Marcos José Leite Santos.

Tese (Doutorado) -- Universidade Federal do Rio Grande do Sul, Instituto de Química, Programa de Pós-Graduação em Ciência dos Materiais, Porto Alegre, BR-RS, 2020.

1. nanopartículas. 2. pontos quânticos. 3. machine learning. 4. síntese coloidal. 5. planejamento experimental. I. Leite Santos, Marcos José, orient. II. Título.

“My son, ask for thyself another kingdom,
For that which I leave is too small for thee...”

Philip II of Macedon

(in Alexander, the Great – Iron Maiden)

AGRADECIMENTOS

Em primeiro lugar à minha mãe, Angelina Baum, sem a qual eu não estaria aqui. Não biologicamente, porque sou adotado, mas em termos éticos, morais e intelectuais. Ela tem uma sabedoria e um espírito que eu não encontraria em nenhuma escola ou laboratório do mundo. No futuro, quem começar a ler essa tese pelos agradecimentos, já vai ter contato com algo pessoal como saber que sou adotado na segunda linha, interessante.

Agradeço ao meu orientador, Professor Marcos José Leite Santos, por ter me recebido em seu grupo, me orientado e me aguentado desde 2014, no mestrado. Acho que não é muito fácil me orientar, pois tenho muitas ideias atípicas (talvez seja possível perceber por essa seção de agradecimentos), mas também acredito que conseguir colaborar para o crescimento do grupo.

Também só tenho a agradecer à minha colega Tatiane Pretto, cuja participação é de valor imensurável neste e em outros projetos. No primeiro dia que ela chegou, eu estava falando sozinho no laboratório (xingando um computador), ainda bem que ela não viu e não pensou que eu era louco.

Karol e Luísa, muito obrigado por terem feito parte deste projeto, vocês foram colegas maravilhosas e moram no meu coração.

A Ariadne participou da coleta de dados para implantação dos modelos de machine learning, muito obrigado por acreditar nesse projeto.

Agradeço em geral aos colegas do LAMAI, pelo companheirismo, amizade, paciência e café.

Também preciso agradecer à minha terapeuta, Tainá, que tem me ajudado a entender eu mesmo. Se você está lendo esse trabalho, recomendo fazer terapia. Fazer ciência é muito recompensador, mas também desgastante mentalmente. São muitas perguntas, parte delas sem resposta e o trabalho está sob constante e necessário questionamento. Mentes saudáveis fazem ciência melhor.

Outro agradecimento mais que necessário é ao meu amigo Guilherme, sem o qual o trabalho não aconteceria (ou teria sido feito à mão), já que ele conseguiu consertar o meu notebook em um momento crucial. Já pensou esse monte de gráfico desenhado à caneta? Já aviso que desenho mal...

Vou precisar de outra página, desculpa, mas acho que a biblioteca vai entender...

Agradeço ao pessoal do LabNano, em especial ao Linus, à Gabriela, à Flávia e ao Professor Gustavo Andrade pela ajuda para usar o espectrômetro Raman e pelos passeios em Juiz de Fora.

Agradeço à Professora Giovanna Machado, do CETENE, pela recepção durante a minha visita e possibilidade de usar o DRX e o MET.

Obrigado ao CNPEM e ao CBPF por todos os cursos que participei e à CAPES e ao CNPq pela bolsa de doutorado.

Quero agradecer também aos técnicos do CMM da UFRGS, Cris, Raquel, Ronei, Matheus e Rodolfo, pelo trabalho que permitiu aquelas belas imagens de microscopia eletrônica de transmissão.

Agradeço também ao CNANO da UFRGS e ao Matheus, que realizou as medidas de difratometria de raios-x.

Agradeço ao CEOMAT, em especial, Prof. Marcelo Barbalho e ao Rodrigo, onde realizamos as medidas de espectroscopia de absorção no UV-Vis-NIR.

Sou grato (pra variar a palavra) à Central Analítica do IQ UFRGS e aos seus técnicos, pelo uso do Fluorímetro e desculpa pela vez que levei a chave embora, eu realmente não vi que ela estava no meu bolso, só vi em casa.

Obrigado à dona Zenaide e ao Lucas, do LAMOCA, pela convivência diária, ajuda e conversas que tivemos ao longo desse tempo.

Obrigado ao Seu Zé, motorista que nos levou inúmeras vezes para pegar solvente e nitrogênio líquido e muitas vezes nos ajudou a carregar o dewar.

Em geral, um agradecimento a todos os servidores e funcionários terceirizados da universidade. Nós chegamos cedo e só saímos ao final da tarde e no meio disso pegamos o Circular (quando tem, né), almoçamos, usamos o banheiro, etc e tudo isso só é possível graças a eles. E sem isso não tem como fazer pesquisa.

Agradeço também ao Tio e à Tia dos doces, muitas vezes eu ia para a universidade pela manhã e ficava só esperando a hora de comer um doce depois do almoço. Ai, que saudade!

Sumário

LISTA DE FIGURAS	IX
LISTA DE TABELAS	XV
CAPÍTULO 1: INTRODUÇÃO	18
CAPÍTULO 2: REVISÃO BIBLIOGRÁFICA	21
2.1. CLASSIFICAÇÃO DOS MATERIAIS SÓLIDOS.....	21
2.1.1. Semicondutores intrínsecos e extrínsecos	21
2.1.2. Banda proibida direta e indireta	22
2.2. PONTOS QUÂNTICOS E NANOPARTÍCULAS DE SEMICONDUCTORES.....	23
2.2.1. Aplicações dos pontos quânticos em LEDs e telas para televisores	24
2.2.2. Aplicações dos pontos quânticos em células solares	28
2.2.3. Sensores baseados em pontos quânticos	30
2.2.4. Aplicações biológicas dos pontos quânticos.....	31
2.2.5. Pontos quânticos aplicados em fotocatalise	32
2.3. PONTOS QUÂNTICOS E NANOPARTÍCULAS DE CALCOGENETOS DE COBRE	34
2.3.1. Sulfeto de cobre e alumínio	35
2.4. INTELIGÊNCIA ARTIFICIAL E <i>MACHINE LEARNING</i>	36
2.4.1 Métricas de avaliação de inteligência artificial	37
2.4.2. Regressão Linear.....	39
2.4.3. Regressão multilinear	39
2.4.4. Métodos de reamostragem	40
2.4.5. Regressão Polinomial	41
2.4.6. Regressão segmentada e smoothing splines	41
2.4.7. Métodos Baseados em Árvores de Decisão	43
2.4.8. Bagging.....	44
2.4.9. Random Forest	44
2.4.10. Boosting	45
2.4.11. Deep Learning	45
2.5 LINGUAGEM R	46
CAPÍTULO 3: INTELIGÊNCIA ARTIFICIAL	48
3.1. MÉTODOS.....	48
3.1.1. Extração de dados.....	48

3.1.2. Feature engineering.....	50
3.1.3. Regressão multilinear	50
3.1.4. Smoothing splines.....	50
3.1.5. Árvores de decisão	51
3.1.6. Random forests	51
3.1.7. <i>Deep learning</i>	51
3.2. RESULTADOS E DISCUSSÃO	53
3.2.1. Regressão multilinear	56
3.2.2. Regressão não-linear.....	63
3.2.3. Árvores de regressão.....	66
3.2.4. Random Forest	72
3.2.5. Abordagens usando <i>machine learning</i> automatizado	78
3.2.6. Deep learning	81
3.2.7. Simulando e sugerindo novas rotas de síntese	84
3.3. CONCLUSÕES PARCIAIS	87
CAPÍTULO 4: PLANEJAMENTO EXPERIMENTAL	89
4.1. PARTE EXPERIMENTAL	89
4.1.1. Materiais	92
4.1.2. Síntese das nanopartículas via heat-up.....	92
4.1.3. Preparação de amostras para microscopia confocal de fluorescência	93
4.1.4. Caracterização óptica	93
4.1.5. Caracterização estrutural e morfológica	93
4.2. RESULTADOS E DISCUSSÃO	95
4.2.1. Segundo Planejamento Experimental.....	110
4.2.2. Aplicação em Microscopia Confocal de Fluorescência	113
5. CONCLUSÕES PARCIAIS.....	115
CAPÍTULO 5: CONCLUSÃO.....	116
REFERÊNCIAS.....	117
ANEXO 1.....	127
ANEXO 2.....	129
ANEXO 3.....	134
ANEXO 4.....	139

LISTA DE FIGURAS

Figura 1: Representação esquemática da teoria de bandas para materiais condutores (a, b), isolantes (c) e semicondutores (d). ^[34]	21
Figura 2: Representação esquemáticas de bandas proibidas com transição direta (a) e indireta (b). ^[35]	23
Figura 3: Propriedades ópticas de pontos quânticos coloidais de CdTe de diferentes tamanhos. a) Diferentes cores da suspensão coloidal; b) Espectros de absorção UV-visível; e c) Espectros de fotoluminescência das amostras. ^[37] ...	24
Figura 4: Diagrama de cromaticidade, adaptado de [1].	25
Figura 5: Representação esquemática das diferentes tecnologias de construção de televisores com telas de pontos quânticos. (Adaptado de [43])	27
Figura 6: Células solares baseadas em pontos quânticos (adaptado de [1]).	29
Figura 7: Detecção de ions por leitura óptica (adaptado de [55]).	31
Figura 8: Representação esquemática da reação de produção de hidrogênio fotocatalisada, na presença de sulfito de sódio (Adaptado de [74]).	33
Figura 9: Relação entre inteligência artificial, machine learning e deep learning (a) e comparação do funcionamento do machine learning em comparação com a programação clássica (b). ^[90]	46
Figura 10: (a) Histograma dos diâmetros de pontos quânticos de CdS da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de CdS sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.	53
Figura 11: (a) Histograma dos diâmetros de pontos quânticos de CdSe da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de CdSe sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.	54
Figura 12: (a) Histograma dos diâmetros de pontos quânticos de PbS da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de PbS sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.	54
Figura 13: (a) Histograma dos diâmetros de pontos quânticos de PbSe da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de PbSe sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.	54

Figura 14: (a) Histograma dos diâmetros de pontos quânticos de ZnSe da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de ZnSe sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.	55
Figura 15: Gráfico de violino das temperaturas de injeção e crescimento para cada composto.	55
Figura 16: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de CdS.	57
Figura 17: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de CdSe.	58
Figura 18: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de PbS.	58
Figura 19: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de PbSe.	59
Figura 20: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de ZnSe.	60
Figura 21: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de CdS.	61
Figura 22: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de CdSe.	61
Figura 23: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de PbS.	62
Figura 24: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de PbSe.	62
Figura 25: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de ZnSe.	63

Figura 26: (a and b) Gráficos das observações vs predições pelo modelo smoothing splines com interações entre as variáveis para pontos quânticos de CdS.	64
Figura 27: (a and b) Gráficos das observações vs predições pelo modelo smoothing splines com interações entre as variáveis para pontos quânticos de CdSe.	64
Figura 28: (a and b) Gráficos das observações vs predições pelo modelo smoothing splines com interações entre as variáveis para pontos quânticos de PbS.	65
Figura 29: (a and b) Gráficos das observações vs predições pelo modelo smoothing splines com interações entre as variáveis para pontos quânticos de PbSe.	66
Figura 30: (a and b) Gráficos das observações vs predições pelo modelo smoothing splines com interações entre as variáveis para pontos quânticos de ZnSe.	66
Figura 31: Árvore de regressão para o diâmetro de pontos quânticos de CdS como função de todas as variáveis da base de dados.	67
Figura 32: Árvore de regressão para o diâmetro de pontos quânticos de CdSe como função de todas as variáveis da base de dados.	68
Figura 33: Árvore de regressão para o diâmetro de pontos quânticos de PbS como função de todas as variáveis da base de dados.	68
Figura 34: Árvore de regressão para o diâmetro de pontos quânticos de PbSe como função de todas as variáveis da base de dados.	69
Figura 35: Árvore de regressão para o diâmetro de pontos quânticos de ZnSe como função de todas as variáveis da base de dados.	69
Figura 36: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de CdS.	70
Figura 37: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de CdSe.	70
Figura 38: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de PbS.	71
Figura 39: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de PbSe.	72
Figura 40: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de ZnSe.	72

Figura 41: Gráficos das observações vs predições pelo modelo random forest para pontos quânticos de CdS.	73
Figura 42: Gráficos das observações vs predições pelo modelo random forest para pontos quânticos de CdSe.	73
Figura 43: Gráficos das observações vs predições pelo modelo random forest para pontos quânticos de PbS.	74
Figura 44: Gráficos das observações vs predições pelo modelo random forest para pontos quânticos de PbSe.	75
Figura 45: Gráficos das observações vs predições pelo modelo random forest para pontos quânticos de ZnSe.....	75
Figura 46: Importância das variáveis no modelo random forest para pontos quânticos de CdS.	76
Figura 47: Importância das variáveis no modelo random forest para pontos quânticos de CdSe.	76
Figura 48: Importância das variáveis no modelo random forest para pontos quânticos de PbS.	77
Figura 49: Importância das variáveis no modelo random forest para pontos quânticos de PbSe.	77
Figura 50: Importância das variáveis no modelo random forest para pontos quânticos de ZnSe.	78
Figura 51: Gráficos das observações vs predições pelo modelo de gradient boosting machine para pontos quânticos de CdS.	79
Figura 52: Gráficos das observações vs predições pelo modelo de gradient boosting machine para pontos quânticos de CdSe.	79
Figura 53: Gráficos das observações vs predições pelo modelo de gradient boosting machine para pontos quânticos de PbS.	80
Figura 54: Gráficos das observações vs predições pelo modelo de gradient boosting machine para pontos quânticos de PbSe.	80
Figura 55: Gráficos das observações vs predições pelo modelo de gradient boosting machine para pontos quânticos de ZnSe.....	81
Figura 56: Gráficos das observações vs predições pelo modelo de deep learning para pontos quânticos de CdS.	81
Figura 57: Gráficos das observações vs predições pelo modelo de deep learning para pontos quânticos de CdSe.	82

Figura 58: Gráficos das observações vs predições pelo modelo de deep learning para pontos quânticos de PbS.	82
Figura 59: Gráficos das observações vs predições pelo modelo de deep learning para pontos quânticos de PbSe.	82
Figura 60: Gráficos das observações vs predições pelo modelo de deep learning para pontos quânticos de ZnSe.....	83
Figura 61: Mapa das condições de reação para pontos quânticos de CdSe com tamanho alvo entre 1.8 a 2.2 nm pelos modelos Random Forest (a), GBM (b), and Deep Learning (c). Mapa das condições de reação para pontos quânticos de CdSe com tamanho alvo entre 4.8 a 5.2 nm pelos modelos Random Forest (d), GBM (e), and Deep Learning (f). Mapa das condições de reação para pontos quânticos de CdSe com tamanho alvo entre 7.8 a 8.2 nm pelos modelos Random Forest (g), GBM (h), and Deep Learning (i).....	86
Figura 62: Imagens de MET das nanopartículas sintetizadas pelo método heat up sob 16 diferentes condições de síntese, de acordo com o planejamento experimental. Barra de escala: 50 nm.....	95
Figura 63: Histograma dos lados das placas hexagonais da amostra HU-001.	96
Figura 64: Histograma dos lados das placas hexagonais da amostra HU-002.	96
Figura 65: Histograma dos lados das placas hexagonais da amostra HU-003.	97
Figura 66: Histograma dos lados das placas hexagonais da amostra HU-005.	97
Figura 67: Histograma dos diâmetros das amostras HU-001 (a), HU-004 (b), HU-006 (c), HU-007 (d), HU-008 (e), HU-009 (f), HU-010 (g), HU-011 (h), HU-012 (i), HU-013 (j), HU-014 (k), HU-015 (l) e HU-016 (m).	98
Figura 68: Difrátogramas de raios-x das amostras sintetizadas pelo método <i>heat up</i>	99
Figura 69: Espectros Raman das amostras sintetizadas pelo método <i>heat up</i>	100
Figura 70: Espectros de absorção UV-Vis-NIR das nanopartículas de sulfeto de cobre e alumínio sintetizadas pelo método <i>heat up</i>	101
Figura 71: Histograma dos valores de bandgap indireto (vermelho) e direto (azul) das amostras sintetizadas pelo método <i>heat up</i> (a) e diagrama de caixas comparando os valores de <i>bandgap</i> direto da literatura com aqueles encontrados neste trabalho (b).	103

Figura 72: Gráfico de Pareto dos efeitos dos parâmetros de síntese no <i>bandgap</i> direto (esquerda) e indireto (direita) das amostras sintetizadas pelo método <i>heat up</i>	103
Figura 73: Espectros de fotoluminescência das nanopartículas sintetizadas pelo método <i>heat up</i> sob as condições HU-001, HU-002, HU-003 e HU-004.....	104
Figura 74: Espectros de fotoluminescência sintetizadas pelo método <i>heat up</i> sob as condições HU-005, HU-006, HU-007 e HU-008.	105
Figura 75: Espectros de fotoluminescência das nanopartículas sintetizadas pelo método <i>heat up</i> sob as condições HU-009, HU-010, HU-011 e HU-012.....	105
Figura 76: Espectros de emissão de fluorescência das nanopartículas sintetizadas pelo método <i>heat up</i> sob as condições HU-013, HU-014, HU-015 e HU-016.....	106
Figura 77: Histograma dos valores de energia de fotoluminescência das amostras (a) e diagrama de caixas comparando os valores da literatura com aqueles encontrados neste trabalho (b).	108
Figura 78: Gráfico de Pareto dos efeitos dos parâmetros da síntese <i>heat up</i> na intensidade da fluorescência das nanopartículas quando excitadas por um feixe com comprimento de onda de (a) 350, (b) 400, (c) 450 e (d) 500 nanômetros.	109
Figura 79: Espectros de absorção na região do UV-Vis das nanopartículas de sulfeto de cobre e alumínio sintetizadas nas condições descritas na Tabela 7.	110
Figura 80: Espectros de fotoluminescência das nanopartículas obtida por método <i>heat up</i> , amostras HU-101, HU-102, HU-103 e HU-104.	111
Figura 81: Espectros de fotoluminescência das nanopartículas obtida por método <i>heat up</i> , amostras HU-105, HU-106, HU-107 e HU-108.	111
Figura 82: Gráfico de Pareto dos efeitos dos parâmetros da síntese <i>heat up</i> na intensidade da fluorescência das nanopartículas quando excitadas por um feixe com comprimento de onda de (a) 350, (b) 400, (c) 450 e (d) 500 nanômetros, sob as condições constantes nas Tabelas 7 e 8.....	113
Figura 83: Imagens de microscopia confocal de fluorescência das nanopartículas sintetizadas pelo método <i>heat up</i> , depositadas sobre lâminas de vidro, sem excitação (superior esquerda), excitada por laser de 405 nm (nos quadros, na posição superior direita), 473 nm (inferior esquerda) e 559 nm (inferior direita).	114

LISTA DE TABELAS

Tabela 1: Resumo das variáveis usadas nos modelos.	48
Tabela 2: Métricas de performance dos modelos.....	83
Tabela 3: Condições de reação sugeridas pelo algoritmo para obtenção de pontos quânticos de CdSe, com tamanhos na faixa de 1.8 a 2.2, 4.8 a 5.2, e 7.8 a 8.2 nm.	85
Tabela 4: Parâmetros experimentais avaliados na síntese de nanopartículas de sulfeto de cobre e alumínio.	89
Tabela 5: Condições experimentais das sínteses	90
Tabela 6: Nomes das amostras após aleatorização.....	90
Tabela 7: Parâmetros experimentais avaliados no segundo planejamento experimental da síntese de nanopartículas de sulfeto de cobre e alumínio.	91
Tabela 8: Condições experimentais das sínteses do segundo planejamento experimental.....	91
Tabela 9: Nomes das amostras do segundo planejamento experimental após aleatorização.	91
Tabela 10: Valores de <i>bandgap</i> direto e indireto calculado a partir do gráfico de Tauc, das amostras sintetizadas pelo método <i>heat up</i>	102
Tabela 11: Intensidades de fluorescência das nanopartículas.	107
Tabela 12: Energia de emissão das amostras.	107
Tabela 13: Comparação da energia de emissão das amostras sintetizadas pelo método <i>heat up</i> com os valores encontrados na literatura.....	108
Tabela 14: Intensidades de fluorescência das nanopartículas sintetizadas pelo método <i>heat up</i> sob as condições de síntese das Tabelas 7 e 8.....	112
Tabela 15: Energia de emissão das amostras sintetizadas pelo método <i>heat up</i> sob as condições de síntese das Tabelas 7 e 8.....	112

Resumo

Título: Otimização de propriedades físicas de pontos quânticos e nanopartículas de semicondutores por inteligência artificial e planejamento experimental

Autor: Fábio Baum

Orientador: Prof. Marcos José Leite Santos

Pontos quânticos são materiais de grande interesse tecnológico, com propriedades ópticas e elétricas aplicáveis em catálise, células solares, biosensores, diodos emissores de luz (LEDs) e *displays*. A produção em larga escala desse material cresceu nos últimos anos, porém a manufatura de pontos quânticos de alta qualidade e com reprodutibilidade batelada após batelada continua um dos objetivos mais desafiadores para a indústria. Nesse contexto, a inteligência artificial tem se revelado uma excelente ferramenta para entender e controlar sistemas químicos. Na primeira parte deste trabalho, foi construída uma base de dados de sínteses *hot injection* de pontos quânticos dos tipos II-VI e IV-VI (CdS, CdSe, PbS, PbSe e ZnSe). Essa base de dados foi usada para construir modelos de regressão multilinear, *smoothing splines*, árvores de regressão, *random forest*, *gradient boosting machine* e *deep learning*. Para a construção de sugestões de novas rotas de síntese, os modelos selecionados foram *random forest*, *gradient boosting machine* e *deep learning*. A partir desses modelos, mapas de calor foram gerados, variando-se temperatura de injeção e tempo de reação, possibilitando a seleção da região com menor variação de diâmetro final. Na segunda parte deste trabalho, o foco foi a busca por alternativas ao uso de semicondutores baseados em cádmio e chumbo. Foram sintetizadas nanopartículas coloidais de sulfeto de cobre e alumínio através do método *heat up*. As amostras foram analisadas por microscopia eletrônica de transmissão, difratometria de raios-x, espectroscopia Raman, espectroscopia de absorção no UV-Vis-NIR e espectroscopia de fotoluminescência. O efeito dos parâmetros de síntese na intensidade de emissão de fotoluminescência foi avaliado, onde foi possível concluir que o principal parâmetro de síntese é a concentração do precursor de alumínio. As amostras foram aplicadas em microscopia confocal de fluorescência, demonstrando potencial para aplicações como marcadores biológicos.

Abstract

Title: Optimization of physical properties of quantum dots and semiconductor nanoparticles by artificial intelligence and design of experiments

Author: Fábio Baum

Advisor: Prof. Marcos José Leite Santos

Quantum dots are materials of great technological interest, with optical and electrical properties applicable in catalysis, solar cells, biosensors, light emitting diodes and displays. Production in large-scale of these materials grown in the last years, but the producing high quality and batch to batch reproducible quantum dots are still a challenge for the industry. In this context, artificial intelligence has been an interesting tool to understand and control chemical systems. In the first part of this work, a database of hot injection synthesis of II-VI and IV-VI (CdS, CdSe, PbS, PbSe, and ZnSe) quantum dots. This database was used to build multilinear regression, smoothing splines, regression trees, random forest, gradient boosting machine, and deep learning models. To find suggestions of new synthesis routes, the selected models were random forest, gradient boosting machine, and deep learning. Then, heatmaps were generated, with the growth temperature and the reaction time, allowing one to select the region with the lowest variability on the final diameter. In the second part of this work, the focus was the search for alternatives to use of cadmium and lead-based semiconductors. Colloidal copper aluminum sulfide nanoparticles were synthesized by the heat up method. The samples were analyzed by transmission electron microscopy, x-ray diffraction, Raman spectroscopy, UV-Vis-NIR absorption spectroscopy and photoluminescence spectroscopy. The effect of the synthetic parameters on the photoluminescence intensity was evaluated. It was possible to conclude that the main parameter affecting the photoluminescence intensity was aluminum precursor concentration. The samples were applied in fluorescence confocal microscopy, showing potential for applications as biological labels.

CAPÍTULO 1: INTRODUÇÃO

Pontos quânticos têm sido intensamente estudados durante as últimas três décadas, devido às suas propriedades ópticas e elétricas e uma ampla gama de aplicações tecnológicas, como em catálise, células solares, biosensores, diodos emissores de luz (LEDs) e *displays*.^[1-5] A pesquisa em pontos quânticos teve início na década de 80^[6] e se intensificou a partir da publicação de um método coloidal simplificado de síntese, descrito por Bawendi e colaboradores em 1993,^[7] que ficou conhecido como injeção à quente (em inglês, *hot injection*). De acordo com relatórios de mercados globais, o mercado de pontos quânticos alcançou o valor de 3,5 bilhões de dólares em 2019 e é esperado que atinja o valor de 35 bilhões de dólares até 2030.^[8,9]

Apesar da produção em larga escala ter crescido nos últimos anos, a manufatura de pontos quânticos de alta qualidade e com reprodutibilidade batelada após batelada continua um dos objetivos mais desafiadores para a indústria. Sínteses de pontos quânticos têm sido baseadas em tentativa e erro, o que pode reduzir a velocidade dos avanços na área, considerando a escala, velocidade e volatilidade dos dados adquiridos manualmente.^[10] A rota mais comumente utilizada para obtenção de pontos quânticos é o método *hot injection*.^[11] Assim, milhares de artigos têm sido publicados mostrando sínteses de diversos compostos, gerando uma enorme base de dados de condições sintéticas, precursores químicos e propriedades de materiais. Dentro deste contexto, a inteligência artificial se apresenta como uma ferramenta poderosa para facilmente detectar padrões e fazer rápidas e precisas previsões usando uma grande quantidade de dados que podem ser gerados por experimentos e por simulações.^[12,13]

O uso de inteligência artificial para entender e controlar sistemas químicos e para prever a relação entre estrutura e propriedades tem sido cada vez mais explorado.^[14,15] O grupo da Professora Elsa Olivetti, do *Massachusetts Institute of Technology*, usou Natural Language Processing (NLP) para extrair informações de síntese de artigos científicos. *Machine learning* foi utilizado para prever os parâmetros críticos na síntese de nanotubos de titânia^[16] e a relação entre a estrutura de zeólitas e as condições de síntese.^[17] Coley e colaboradores

desenvolveram uma plataforma robótica para sintetizar compostos orgânicos, baseada em modelos de *machine learning*, capaz de planejar sínteses quando alimentados por publicações descrevendo reações químicas.^[18] Na área de pontos quânticos, Voznyy e colaboradores aplicaram modelos de *machine learning* à dados experimentais de 6 anos de notas de seu laboratório para encontrar regiões no espaço paramétrico de síntese que levam a melhores monodispersividades de PbS.^[19]

Em outra perspectiva, para gerar e explorar dados experimentais, o planejamento experimental é uma ferramenta poderosa pra investigar como um determinado fenômeno se comporta quando seus parâmetros são mudados. Essa técnica nos permite extrair o máximo de informação possível com poucos experimentos, economizando tempo e materiais.^[20] Em trabalhos anteriores, foi demonstrado como o planejamento experimental pode ser usado para encontrar parâmetros de síntese que controlam o tamanho de nanopartículas,^[21,22] assim como fatores que influenciam a estrutura do material. Portanto, a técnica de síntese pode ser otimizada para alcançar uma determinada composição, tamanho ou direcionar suas propriedades físicas e químicas.^[23-25,26]

Na primeira parte deste trabalho, buscou-se na literatura por artigos que relatassem a síntese *hot injection* de pontos quânticos dos tipos II-VI e IV-VI (CdS, CdSe, PbS, PbSe e ZnSe). A base de dados foi criada do zero, extraíndo informação das temperaturas de injeção e crescimento, tempo de reação, precursores metálicos, solventes, ligantes e suas concentrações. Uma vez que a técnica *hot injection* é mais complexa que as técnicas hidrotermais usadas no trabalho de Olivetti, os parâmetros tiveram de ser extraídos manualmente. Métodos de regressão linear e não linear foram usados para modelar os dados, com seis algoritmos de regressão: multilinear, *smoothing splines*, árvore de regressão, *random forest*, *gradient boosting machines* e *deep learning sequential model*. Os melhores modelos foram selecionados e o método de Monte Carlo foi aplicado para mapear quais as condições sintéticas levariam a um tamanho desejado de pontos quânticos.

Na segunda parte deste trabalho, buscou-se por alternativas ao uso de semicondutores baseados em cádmio e chumbo. Historicamente, a alternativa mais pesquisada na literatura é o sulfeto de cobre e índio (CuInS₂).^[27] Com

propriedades ópticas semelhantes ao CdSe e ampla faixa de absorção e emissão ao longo da região visível, o CuInS₂ encontrou aplicações similares ao CdSe, porém com menor impacto ambiental. Contudo, devido à sua baixa abundância na crosta terrestre, o elemento índio apresenta elevado custo. Com isso, novos trabalhos surgiram, pesquisando a substituição do elemento índio por alternativas como antimônio.^[28] Recentemente, também passou-se a explorar as propriedades do sulfeto de cobre e alumínio (CuAlS₂). Assim como seu análogo de índio, apresenta estrutura do tipo calcopirita^[29] e propriedades ópticas promissoras.^[30] O uso deste material em aplicações tecnológicas ainda é bastante incipiente, com exemplos de emprego de CuAlS₂ como marcador biológico,^[30] camada ativa em LEDs^[31] e folhas artificiais para realização de fotossíntese.^[32] O objetivo desta etapa do trabalho é sintetizar nanopartículas coloidais de sulfeto de cobre e alumínio através do método *heat up*. Em seguida, avaliar o efeito dos parâmetros de síntese, como temperatura, tempo de reação e concentração dos precursores, nas propriedades ópticas do material, através de medidas de absorção de luz e fotoluminescência.

CAPÍTULO 2: REVISÃO BIBLIOGRÁFICA

2.1. CLASSIFICAÇÃO DOS MATERIAIS SÓLIDOS

Os materiais sólidos podem ser classificados, dentre outras formas, de acordo com a sua capacidade de conduzir corrente elétrica e suas propriedades elétricas podem ser explicadas através da Teoria de Bandas. A Figura 1 apresenta uma representação esquemática da estrutura de bandas dos materiais, onde E_f representa o nível de Fermi do material.^[34] Desse modo, existem três grupos: condutores, semicondutores e isolantes. Os materiais condutores podem possuir a banda de condução parcialmente preenchida, como mostrado na Figura 1(a), ou sobreposição das bandas de valência e condução, sem apresentar *bandgap*, como mostrado na Figura 1(b). Estes materiais apresentam alta condutividade na ordem de 10^7 S.m⁻¹. Por outro lado, os materiais isolantes (Figura 1(c)) apresentam uma grande diferença energética que separa as extremidades das bandas de valência e de condução, apresentando baixas condutividades, entre 10^{-10} e 10^{-20} S.m⁻¹, e largura da banda proibida geralmente superior a 4 eV.^[33] Os semicondutores, mostrados na Figura 1(d), apresentam propriedades intermediárias entre isolantes e condutores. Estes materiais apresentam *bandgap*, mas não tão grandes como em isolantes e apresentam condutividade elétrica intermediária, entre 10^{-6} e 10^4 S.m⁻¹, e largura da banda proibida menor que 4 eV.^[33,34]

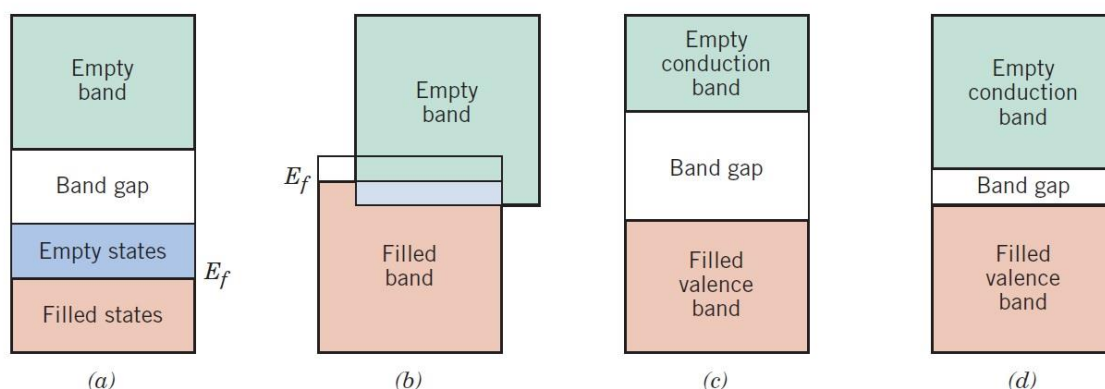


Figura 1: Representação esquemática da teoria de bandas para materiais condutores (a, b), isolantes (c) e semicondutores (d).^[34]

2.1.1. Semicondutores intrínsecos e extrínsecos

Os semicondutores podem ser classificados em dois tipos: intrínsecos e extrínsecos. Nos semicondutores intrínsecos, o comportamento elétrico depende

apenas da estrutura do material puro, enquanto nos semicondutores extrínsecos, as propriedades elétricas são determinadas por átomos de impurezas na rede cristalina.^[34] Nos semicondutores intrínsecos, que são o foco deste trabalho, quando a temperatura é igual a 0 K, a banda de valência, que encontra-se completamente preenchida, é separada da banda de condução, totalmente vazia, pela banda proibida. Além dos semicondutores elementares, silício e germânio, existe uma variedade de semicondutores compostos que exibem comportamento intrínseco.^[34] Nos semicondutores intrínsecos, quando um elétron é promovido para a banda de condução, ocorre a formação de uma vacância deste elétron na banda de valência. Essa vacância é chamada de buraco, e considera-se que possui carga de mesma magnitude do elétron, entretanto, com sinal positivo, de forma que, na presença de um campo elétrico, elétrons e buracos se movem em direções opostas.^[34]

2.1.2. Banda proibida direta e indireta

De acordo com os processos de transição eletrônica da banda de valência para a banda de condução, a banda proibida, presente nos materiais semicondutores e isolantes, pode ser classificada em direta ou indireta. Na transição direta, um fóton é absorvido por um elétron na banda de valência do cristal e é promovido diretamente à banda de condução, gerando de um buraco, de mesma carga e sinal oposto, na sua banda de origem. No processo indireto, existe um desalinhamento entre o ponto de mínima energia da banda de condução e a máxima energia na banda de valência, como mostrado na Figura 2.^[35] Esse desalinhamento é representado pelo vetor de onda k_c . Neste caso, a transição direta do fóton no ponto de energia mínima entre as bandas não satisfaz a necessidade de conservação do vetor de onda, uma vez que os vetores de onda dos fótons são desprezíveis no intervalo de energia de interesse. Entretanto, se um fônon com vetor de onda K e frequência (Ω) for criado no processo, é possível satisfazer as leis de conservação.^[35] A energia do fônon $h(\Omega)$ será geralmente muito menor do que E_g : um fônon, mesmo possuindo um vetor de onda elevado, é uma fonte não dissipativa na quantidade de movimento do cristal, porque as energias dos fônons são caracteristicamente pequenas ($\sim 0,01$ até $0,03$ eV) em comparação com a largura da banda proibida. Se a temperatura for tão elevada que o fônon necessário já esteja termicamente

excitado no cristal, é possível que haja um processo de absorção de fóton no qual o fônon também seja absorvido.^[35]

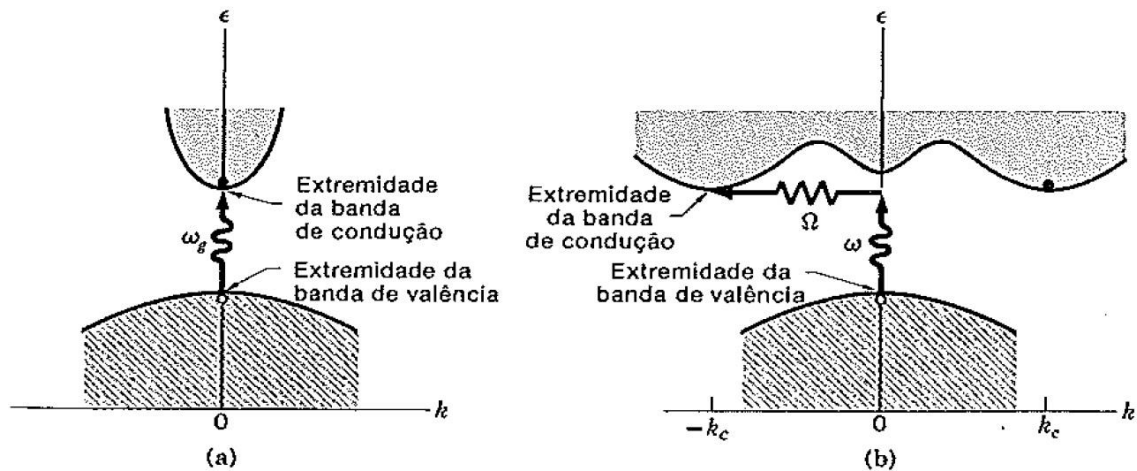


Figura 2: Representação esquemática de bandas proibidas com transição direta (a) e indireta (b).^[35]

2.2. PONTOS QUÂNTICOS E NANOPARTÍCULAS DE SEMICONDUTORES

Nanopartículas, de acordo com a Organização Internacional de Normalização (em inglês: *International Organization for Standardization*, ISO), são todas estruturas que apresentam tamanho entre 1 e 100 nm nas três dimensões espaciais e não possuem nenhum eixo significativamente maior que os outros (tipicamente, mais de 3 vezes).^[36] Pontos quânticos, por sua vez, são nanopartículas de semicondutores cujo raio é igual ou inferior ao raio de Bohr do composto. O raio de Bohr (a_B) é a distância entre o elétron e o buraco na formação de um éxciton (par elétron-buraco), que ocorre quando um elétron é excitado da banda de valência para a banda de condução, definido pela equação 1:

$$a_B = \frac{4\pi\hbar^2\epsilon}{e^2} \left(\frac{1}{m_h} + \frac{1}{m_e} \right) \quad \text{Eq. 1.}$$

Onde \hbar é a constante reduzida de Planck, ϵ é a constante dielétrica do material e m_h e m_e , são, respectivamente, a massa efetiva do buraco e do elétron no semicondutor. Em semicondutores, esse valor normalmente é de poucos nanômetros.^[37] Nos semicondutores macroscópicos, o éxciton pode se mover livremente em todas as direções, mas, quando o tamanho da partícula é da mesma ordem do raio de Bohr, ocorre o Efeito de Confinamento Quântico. Nessas condições, o comportamento do éxciton sofre alterações, o que confere

aos semicondutores novas propriedades ópticas, que podem ser moduladas através do controle do tamanho da partícula,^[37] conforme mostrado na Figura 3.

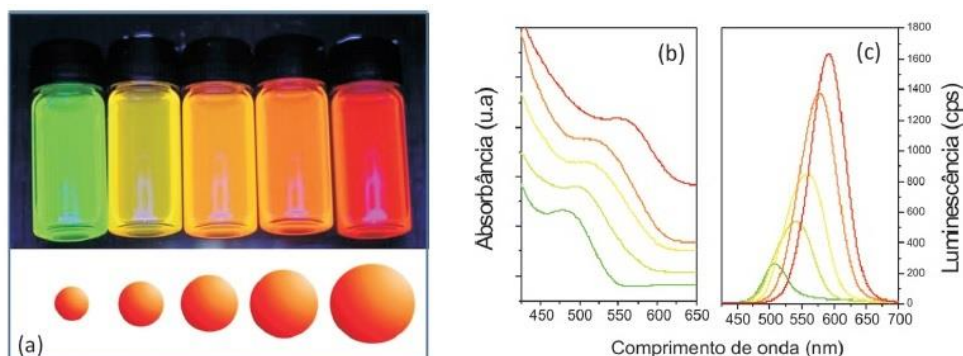


Figura 3: Propriedades ópticas de pontos quânticos coloidais de CdTe de diferentes tamanhos. a) Diferentes cores da suspensão coloidal; b) Espectros de absorção UV-visível; e c) Espectros de fotoluminescência das amostras.^[37]

A diminuição do tamanho de um ponto quântico causa uma intensificação no efeito de confinamento quântico sofrido pelos éxcitons, aumentando a banda proibida.^[37] Uma vez que cada material possui um valor característico de raio de Bohr, o tamanho exato no qual cada semiconductor passa a apresentar confinamento quântico pode variar significativamente. Quando esse material recebe incidência de luz com energia maior que a largura da banda proibida, os elétrons na banda de valência são excitados e passam para a banda de condução, gerando buracos na banda de valência. Após um intervalo de tempo na ordem dos nanossegundos (10^{-9} s), ocorre a recombinação do elétron na banda de condução com o buraco na banda de valência, gerando um fóton (luz). Esse efeito é chamado fluorescência, e é um dos grandes responsáveis pelas aplicações tecnológicas dos pontos quânticos.^[37]

2.2.1. Aplicações dos pontos quânticos em LEDs e telas para televisores

As propriedades ópticas dos pontos quânticos permitem que estes sejam utilizados em diversas aplicações, como marcadores biológicos,^[38] catalisadores,^[39] sensores,^[40] diodos emissores de luz (LEDs),^[41] agentes sensibilizadores em células solares,^[42] camadas ativas em telas de televisores^[43] e *tablets*.^[44] Os LEDs baseados em pontos quânticos (QLEDs) têm sido objeto de intensas pesquisas nas últimas décadas.^[45] O grande interesse em dispositivos montados a partir de pontos quânticos coloidais se deve à possibilidade de processamento em solução, compatível com as técnicas de

impressão de LEDs orgânicos (OLEDs), de tecnologia conhecida e de baixo custo.^[46] A Figura 4 demonstra o diagrama de cromaticidade, mostrando o potencial dos pontos quânticos na representação das cores em relação às atuais tecnologias. Neste diagrama, o ponto branco central corresponde à luz branca emitida pelo Sol. Esse ponto de referência possui CRI (Color Rendering Index, ou Índice de Reprodução de Cor) igual a 100. As cores são definidas em relação a esta referência usando coordenadas polares, onde o matiz é o componente angular e a saturação é o componente radial, normalizado pelo raio máximo possível para aquele matiz. Os limites externos do diagrama representam as cores que podem ser percebidas pelo olho de um humano médio, enquanto que os pontos dentro do diagrama representam as cores que podem ser emitidas pelas telas das diferentes tecnologias.

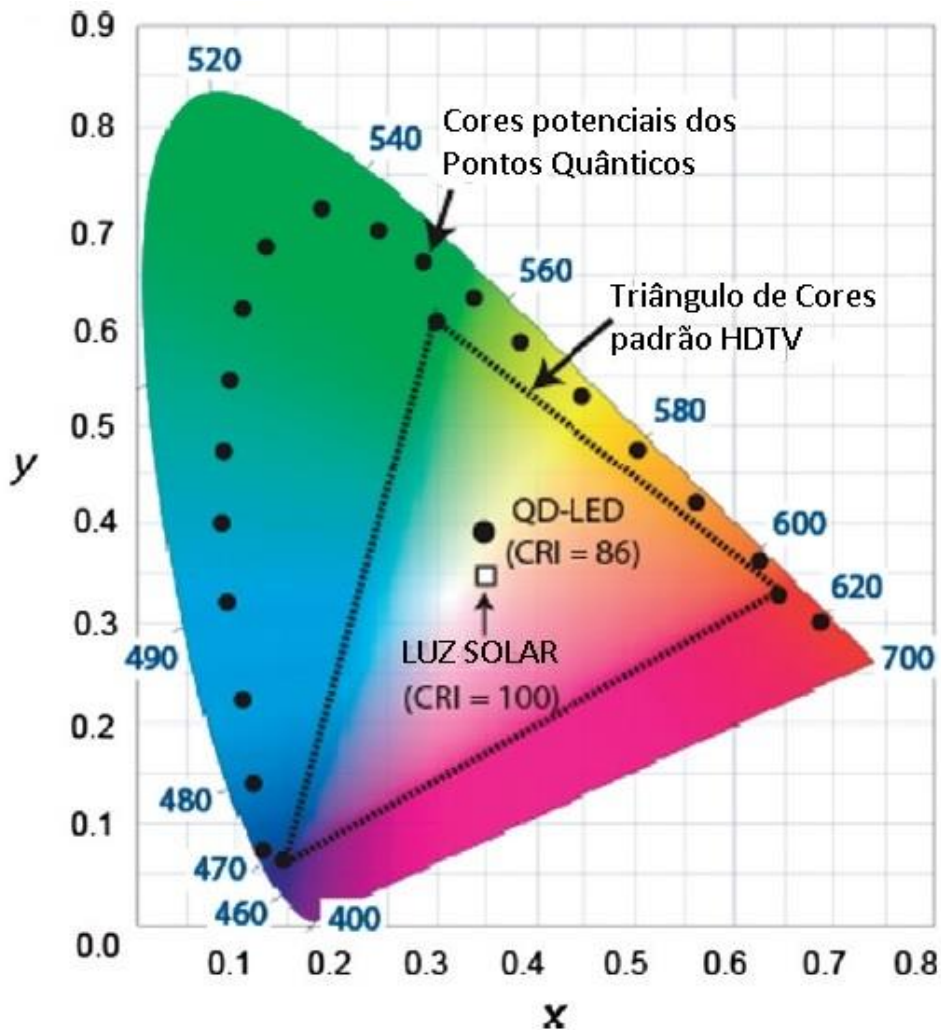


Figura 4: Diagrama de cromaticidade, adaptado de ^[47].

Em comparação com outros materiais, os pontos quânticos possuem uma capacidade única de emitir luz em diferentes comprimentos de onda, de acordo com o tamanho da nanopartícula. Essa propriedade é explorada por diversas tecnologias em telas de televisores.^[43] A Figura 5 ilustra diferentes tecnologias desenvolvidas e em desenvolvimento, onde pontos quânticos são aplicados, mostrando suas principais diferenças e semelhanças. A tecnologia de telas de cristal líquido (LCD) retroiluminadas por LED, mostrada na Figura 5(a), emprega os pontos quânticos para aprimorar a qualidade as cores produzidas pelo dispositivo. Entre as principais vantagens estão cores mais profundas sob altos picos de luminância, menor quantidade de pixels defeituosos na tela, compatibilidade com os atuais processos de fabricação de telas LCD e baixo custo em comparação com as telas de LED orgânico (OLED).^[43] Entre as desvantagens, estão o estreito ângulo de visão, cores escuras menos profundas, baixas velocidades de resposta e impossibilidade de serem dobradas.^[43]

Em outra tecnologia, de pontos quânticos fotoemissivos, mostrada na Figura 5(b), os pontos quânticos substituem os filtros e se tornam os próprios subpixels. A retroiluminação é fornecida por LEDs azuis, fazendo com que os subpixels dessa cor sejam pontos transparentes no arranjo e a luz passe por eles sem sofrer mudanças. Os subpixels vermelho e verde são feitos de pontos quânticos que absorvem a luz azul e emitem comprimentos de onda precisos de luz vermelha e verde. Combinando a remoção dos filtros com a alta eficiência dos pontos quânticos (~99%), a tela pode criar imagens com o dobro do brilho e da eficiência das telas de LCD. Essas telas têm ainda um alto ângulo de visão, uma vez que os pontos quânticos estão na parte frontal e emitem luz em todas as direções. Televisores com tecnologia de pontos quânticos fotoemissivos, inicialmente previstos para entrar no mercado a partir de 2019, tiveram sua estreia adiada para depois de 2020. A tecnologia de pontos quânticos fotoemissivos torna possível a fabricação de televisores flexíveis e dobráveis.^[43]

No sistema de telas de pontos quânticos eletroemissivos, mostrado na Figura 5(d), os pontos quânticos são estimulados a emitir fótons através de corrente elétrica ao invés de feixe luminoso. Assim como na tecnologia fotoemissiva, não há a necessidade de filtros de cores, nem de, nesse caso, retroiluminação. Nesse arranjo, cada subpixel possui um endereço, que pode ser ligado a partir de corrente elétrica, evitando o desperdício de energia com a

retroiluminação. As telas de pontos quânticos eletroemissivos podem proporcionar a espessura fina e flexibilidade das telas OLED, mas com o custo, qualidade de cor, brilho e durabilidade dos pontos quânticos. Esta tecnologia ainda está em fase de desenvolvimento, sendo prevista sua entrada no mercado nos próximos 5 anos.^[43]

A combinação da capacidade de emissão dos pontos quânticos com a tecnologia micro-LED, faz com que cada subpixel vermelho, verde ou azul seja um LED, como mostrado na Figura 5(e). Essa tecnologia permite a produção de telas com imagens mais brilhantes, cores escuras mais profundas, com menor tempo de resposta e consumo de energia, porém o custo ainda é um desafio.^[43]

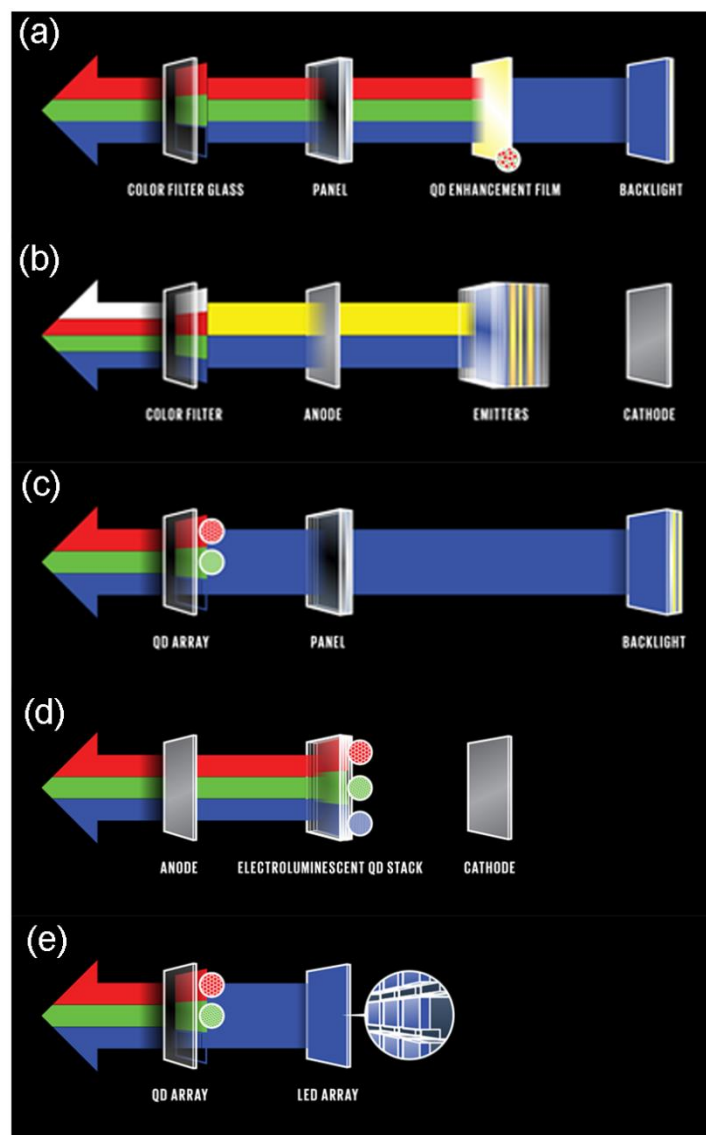


Figura 5: Representação esquemática das diferentes tecnologias de construção de televisores com telas de pontos quânticos. (Adaptado de [43])

2.2.2. Aplicações dos pontos quânticos em células solares

A energia solar é uma importante alternativa ao uso de combustíveis fósseis e nucleares que tem surgido em maior evidência nos últimos anos. Nesse contexto, diferentes tipos de tecnologias têm sido desenvolvidas para explorar esta energia.^[48] As tecnologias em células solares podem ser divididas cronologicamente em três gerações. A primeira geração de células fotovoltaicas é baseada em uma pastilha de um único material semiconductor monocristalino.^[48] Sua principal desvantagem é seu alto custo de fabricação.^[49] A segunda geração utiliza um filme fino de um semiconductor policristalino na montagem da célula. Apesar de possuir um custo de produção menor que suas antecessoras, apresentavam um rendimento inferior (menos de 14%), em comparação com células da primeira geração, que podem alcançar rendimentos de até 27%, em laboratório. A terceira geração de células solares surgiu com a possibilidade de superar o limite termodinâmico de eficiência de Shockley-Queisser,^[48] estabelecido em 32% para células solares baseadas em silício, com um menor custo de produção.^[49] Dentre as células solares de terceira geração, três principais tipos de células dominaram as pesquisas nos últimos anos: (I) células solares sensibilizadas por pigmentos (DSSC), (II) células fotovoltaicas de heterojunção distribuída ou células orgânicas fotovoltaicas, e (III) células solares de pontos quânticos (QDSC).^[50] A Figura 6 apresenta uma representação esquemática de três tipos de células solares que usam semicondutores para absorção de luz solar.

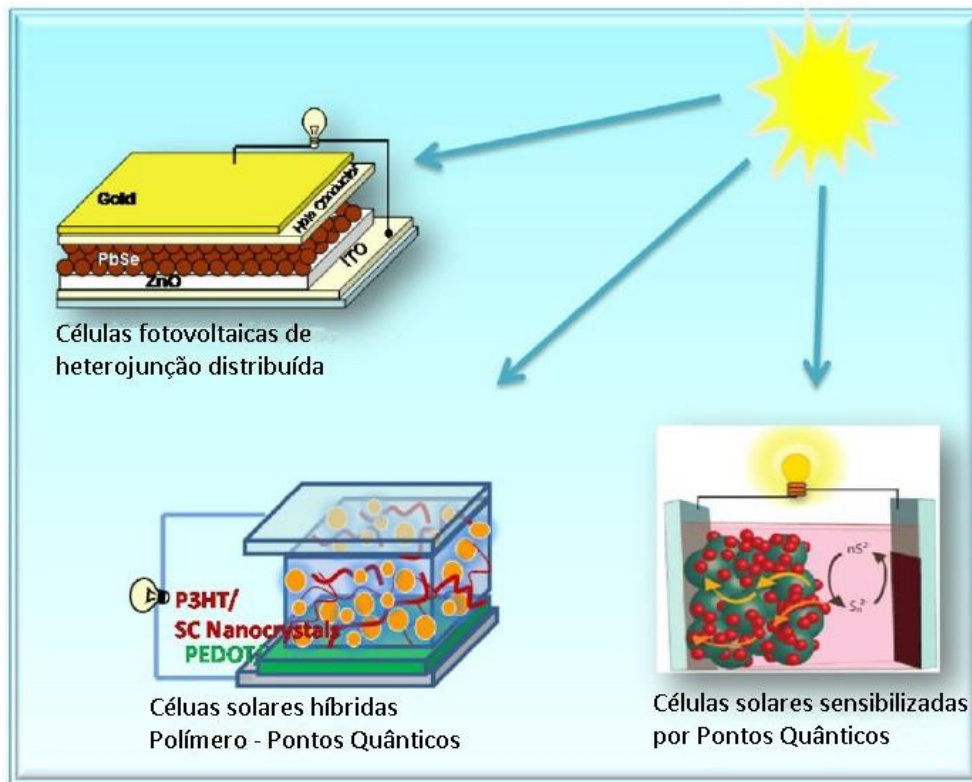


Figura 6: Células solares baseadas em pontos quânticos (adaptado de [50]).

As células solares de terceira geração apresentam potencial para alta eficiência na conversão da energia solar e custo de fabricação menor do que as atuais tecnologias.^[49] Os procedimentos de síntese simplificados, a capacidade de ajuste da absorção de luz, a sensibilidade à luz difusa, e a capacidade de utilização em painéis solares flexíveis são as principais vantagens das nanoestruturas semicondutoras como candidatas à agentes de absorção de luz em células fotovoltaicas.^[50] Foram desenvolvidas três formas de utilização de pontos quânticos em células solares: metal-semicondutor, ou Junção de Schottky; célula solar híbrida de polímero-semicondutor; e célula solar sensibilizada com pontos quânticos.^[49]

As principais vantagens do uso de pontos quânticos na confecção de células solares são: a possibilidade, através de um dispositivo de multi-junção, de usar o mesmo composto para absorver um amplo espectro da radiação solar, desde UV até infravermelho; maior estabilidade e maior resistência ao oxigênio em comparação com polímeros;^[51] e a possibilidade de se alcançar rendimentos maiores do que os previstos para células solares baseadas em silício.^[49]

2.2.3. Sensores baseados em pontos quânticos

Sensores são dispositivos que sofrem alterações em suas propriedades físicas (condutividade, absorção e emissão de luz, etc.) em resposta a alterações físicas e químicas no meio onde estão presentes. Os sensores de temperatura baseiam-se na resposta de fluorescência dos pontos quânticos em função de mudanças na temperatura.^[52] Exemplos de aplicações de pontos quânticos como sensores de temperatura incluem: tintas sensíveis à temperatura (TSP), usadas na indústria aeroespacial;^[53] nanotermômetros, capazes de indicar variações e gradientes de temperatura no interior de uma célula viva;^[54] e na medição de temperatura em micro e nanoestruturas, onde as técnicas convencionais não podem dar uma boa resolução espacial na medição.^[54]

Entre os sensores químicos, uma das principais aplicações é a detecção de íons e determinação de sua concentração. A leitura óptica é uma alternativa para análises em ambiente intracelular, uma vez que os métodos eletroquímicos são invasivos. Os principais processos pelos quais os pontos quânticos atuam na detecção de íons são a Transferência Ressonante de Energia por Fluorescência (FRET), onde os nanocristais são conjugados a fluoróforos orgânicos sensíveis à presença de íons específicos; e a ligação direta entre os íons presentes no meio e os pontos quânticos, que pode causar um aumento ou diminuição da fluorescência, gerando uma leitura de saída.^[55]

A seletividade da resposta à presença de íons depende da natureza química dos ligantes usados nos nanocristais.^[56] A resposta obtida por pontos quânticos cobertos com diferentes agentes ligantes foi estudada para uma grande diversidade de íons,^[56-59] moléculas orgânicas^[60-62] e como sensores de pH.^[63-66] A Figura 7 mostra o esquema de funcionamento genérico de sensores químicos utilizando nanopartículas e suas propriedades ópticas.

Ion sensing with colloidal nanoparticles

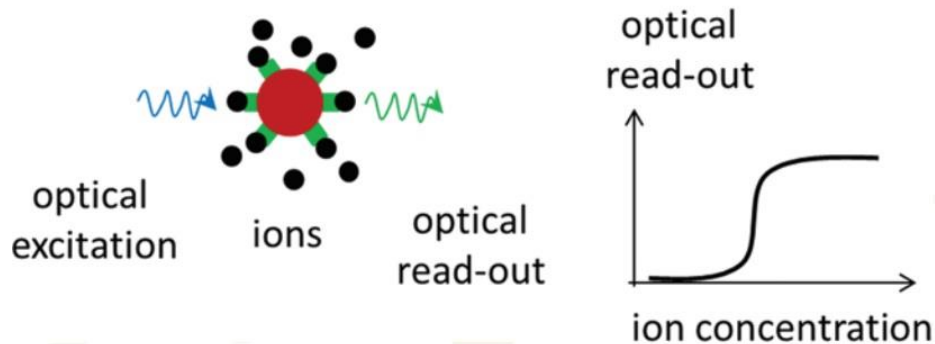


Figura 7: Detecção de ions por leitura óptica (adaptado de [55]).

2.2.4. Aplicações biológicas dos pontos quânticos

O potencial da utilização de pontos quânticos em aplicações envolvendo marcadores biológicos foi demonstrado pela primeira vez em 1998.^[68] Esses materiais são adequados para utilização em marcação imunológica, análise de mobilidade celular, hibridização *in situ*, como marcadores de células vivas, bem como apenas um corante fluorescente não específico. Os principais métodos de rastreamento e detecção baseados em pontos quânticos são a microscopia confocal, microscopia de reflexão interna total, microscopia de epifluorescência de campo largo e fluorometria.^[67]

A capacidade de alguns pontos quânticos de emitir radiação eletromagnética na faixa do infravermelho e infravermelho próximo permite a aplicação desses materiais em imageamento e diagnóstico de tecidos profundos, uma vez que a absorção destes tecidos é mínima nestas regiões do espectro.^[67] Sua fotoestabilidade e resistência a degradação metabólica em células vivas tornam possível o uso desses materiais em estudos de imageamento por longos períodos.^[68]

Resultados de pesquisas mostram que o uso de pontos quânticos como marcadores possibilita uma sensibilidade maior em imuno-histoquímica do que os corantes convencionais, apresentando um sinal melhor nas seções marcadas com pontos quânticos, se comparado aos métodos de amplificação enzimática de sinal.^[67,69]

Esses materiais ainda podem ser usados na detecção de ácidos nucléicos, onde se apresentam como alternativa aos clássicos corantes orgânicos fluorescentes, que possuem problemas como a clivagem da molécula

de DNA e a formação de radicais livres que perturbam as interações DNA-proteína.^[67]

Quando funcionalizados adequadamente, os pontos quânticos podem ser cobertos com peptídeos e são internalizados após se ligarem à superfície celular e podem ser utilizados na administração de medicamentos, além de evitar o vazamento prematuro de moléculas de medicamentos de um determinado tamanho para fora das esferas.^[67]

Entre as dificuldades para a utilização dos pontos quânticos em terapias e análises biológicas está o tamanho do complexo envolvendo o nanocristal, que incluindo os surfactantes, pode impedir a penetração através da membrana celular, além da citotoxicidade e de alterações das funções celulares.^[70]

Os pontos quânticos possuem propriedades ópticas superiores em comparação com as atuais moléculas usadas em diagnóstico por imagem e análises biológicas, mas aplicações desses materiais para estudos *in vivo* ainda têm um longo caminho a percorrer antes de atingirem o mercado.^[67,70]

2.2.5. Pontos quânticos aplicados em fotocatalise

A fotocatalise por semicondutores tem-se mostrado útil em diversas aplicações, como a destruição de microorganismos nocivos (bactérias e vírus), inativação de células cancerosas, controle de odor, quebra da molécula da água para produção de hidrogênio, fixação de nitrogênio e degradação de resíduos oleosos.^[71]

Os fotocatalisadores, normalmente, são semicondutores sólidos que possuem como características: (I) a capacidade de absorver fótons na faixa da luz visível e ultravioleta, (II) fotoestabilidade e inércia química e biológica, (III) baixo custo e (IV) atoxicidade.^[72] Algumas características fazem dos semicondutores nanocristalinos coloidais do tipo II-VI materiais de interesse em fotocatalise: (I) sua capacidade de absorção na faixa da luz visível, absorvendo uma parcela significativa da luz solar; (II) a grande capacidade de absorção de luz dos nanocristais de calcogenetos, com absorvidade na ordem de $10^5 - 10^7 \text{ M}^{-1}\text{cm}^{-1}$; (III) a capacidade de modulação de suas fronteiras de bandas, potencial redox e espectro de absorção, através de mudanças de composição, tamanho e forma; (IV) o acesso facilitado dos portadores de carga à superfície dos nanocristais; (V) a grande diversidade de surfactantes que podem ser utilizados

na estabilização das nanopartículas, possibilitando a escolha entre dispersividade em meio aquoso ou orgânico.^[73]

As propriedades dos nanocristais que mais influenciam a cinética da reação fotocatalítica são a taxa de transferência de elétrons (ET), que pode ser medida usando-se receptores com fortes assinaturas espectroscópicas, como rodamina B, antraquinona e azul de metileno, e a taxa de transferência de buracos (HT), que pode ser medida através da verificação da assinatura espectroscópica das espécies oxidadas de fenotiazina e p-fenilenodiamina.^[73]

Uma recente abordagem no uso da fotocatalise é a produção de combustíveis por reações químicas catalisadas pela luz solar. As propriedades altamente moduláveis dos semicondutores nanocristalinos os faz potenciais candidatos à absorventes de luz para sistemas de geração de combustíveis. Recentemente, uma série de estudos tem sido realizados utilizando semicondutores nanocristalinos como parte de sistemas catalíticos que promovem reações como a geração de hidrogênio e a redução do dióxido de carbono.^[73] A Figura 8 mostra o esquema da reação fotocatalisada por pontos quânticos de CdSe para produção de hidrogênio na presença de sulfito de sódio.^[74]



Figura 8: Representação esquemática da reação de produção de hidrogênio fotocatalisada, na presença de sulfito de sódio (Adaptado de [74]).

Outro desafio no campo da fotocatalise é o aproveitamento dos fótons de baixa energia, na faixa do infravermelho e infravermelho próximo, pois cerca de 40% da energia solar que chega à superfície terrestre encontra-se nessa faixa.^[39] Diversas pesquisas têm sido realizadas, com o objetivo de aumentar a eficiência

na geração e na transmissão das cargas a seus respectivos receptores, aumentando a capacidade dos sistemas fotocatalíticos.^[39,73]

2.3. PONTOS QUÂNTICOS E NANOPARTÍCULAS DE CALCOGENETOS DE COBRE

Pontos quânticos e nanopartículas de calcogenetos de cobre e seus derivados multinários, que incluem íons de metais de transição adicionais (Zn, Fe, Cr, Cd, Hg), elementos representativos (In, Ga, Al, Sn, Bi, Sb) e semimetais (Ge), são uma classe extraordinariamente interessante de materiais com uma ampla gama de fases cristalinas estáveis. O grande interesse nesses materiais se deve a três fatores principais: (i) abundância na crosta terrestre, baixo custo, menor impacto ambiental e à saúde, quando comparado com compostos baseados em cádmio e chumbo; (ii) excelentes propriedades funcionais intrínsecas, como *bandgap* direto, adequado para dispositivos fotovoltaicos, propriedades plasmônicas, grande mobilidade dos portadores de cargas, potencial alta concentração dos portadores e baixa condutividade térmica; (iii) versatilidade composicional, estequiométrica, incluindo abundantes fases não-estequiométricas, uma ampla gama de soluções sólidas e energia de formação de defeitos relativamente baixa. Mesmo o mais simples dos compostos binários, Cu_{2-x}A (onde $\text{A} = \text{S}, \text{Se}$ ou Te) apresenta mais de 20 diferentes estequiometrias, polimorfismos e fases de defeitos relacionadas, além de diversos compostos ternários e multinários.^[75]

A introdução de metais e semimetais do bloco p (Ga, In, Sn, Ge) ou metais de transição (Zn, Fe, Cd) para formar compostos ternários e quaternários de calcogenetos de cobre adiciona um alto grau de complexidade na síntese dos nanocristais, porém aumenta a gama de aplicações tecnológicas. As numerosas possibilidades de substituições químicas e modificações estruturais em compostos ternários e quaternários provém um escopo significativo para modular as propriedades fundamentais do material para a aplicação desejada. Por exemplo, dissulfeto de cobre e índio (CuInS_2 , CIS), disseleneto de cobre e índio (CuInSe_2 , CISE) e disseleneto de cobre, índio e gálio (CuInGaSe_2 , CIGSe) são importantes compostos semicondutores que têm sido estudados por mais de duas décadas como absorvedores do tipo p em fotovoltaicos de filme fino. Esses materiais são conhecidos por sua alta eficiência em conversão de energia,

alto coeficiente óptico de absorção, boa fotoestabilidade sob radiação a longo prazo, toxicidade relativamente baixa em comparação com os materiais convencionais baseados em cádmio.^[75]

Um grupo particular dentre os calcogenetos de cobre, que apresentam grande interesse tecnológico são as calcopiritas. Estes compostos têm uma estequiometria geral Cu-III-VI₂ (onde III = Al, Ga, In; e VI = S, Se ou Te), pertencem ao sistema cristalino escalenoedro tetragonal e formam um grande grupo de materiais semicondutores, com diversas propriedades ópticas, elétricas e estruturais, apropriadas para diversas aplicações. Os efeitos de desvios estequiométricos podem ser usados para influenciar a condutividade,^[76] *bandgap*,^[77] e luminescência.^[78,79]

2.3.1. Sulfeto de cobre e alumínio

O sulfeto de cobre e alumínio é um membro atrativo das calcopiritas, devido à abundância de seus elementos constituintes na crosta terrestre e baixa toxicidade em comparação com pontos quânticos tradicionais.^[32] Esse material é conhecido por ser um semicondutor do tipo p, com alta energia de *bandgap* (> 3.2 eV) e alta condutividade de buracos.^[80] Entre as aplicações já relatadas incluem-se a marcação *in vitro* de células cancerosas para análise por microscopia de fluorescência,^[30] a fotocatalise para redução de CO₂,^[32] catalise na pirólise de celulose^[81] e atuação como camada transportadora de buracos em células solares híbridas.^[80]

Quando usado em estruturas do tipo caroço/casca, pontos quânticos de sulfeto de cobre e alumínio apresentaram alto rendimento quântico (63%) e longo tempo de vida no estado excitado (~1500 ns), com potencial para aplicações em LEDs.^[31] Os autores também verificaram que o sulfeto de cobre e alumínio é sensível ao oxigênio atmosférico, causando degradação em óxido de alumínio e sulfeto de cobre. O recobrimento com uma casca inorgânica aumenta a estabilidade do núcleo de sulfeto de cobre e alumínio.^[31]

Vários métodos de síntese já foram empregados para obtenção de sulfeto de cobre e alumínio. Para a obtenção de filmes finos deste material, predominam as rotas a vácuo.^[82-84] No entanto, para a síntese de nanopartículas ou pontos quânticos, predominam as rotas químicas, como método do poliol,^[85] hidrotermal^[86] e injeção à quente.^[32,76,78] Destes métodos, a injeção à quente se

destacou por permitir a obtenção de materiais com propriedades ópticas adequadas para aplicação como marcadores biológicos,^[30] LEDs^[31] e fotocatalise.^[32]

2.4. INTELIGÊNCIA ARTIFICIAL E *MACHINE LEARNING*

Inteligência Artificial (IA) é comumente interpretada como sendo a incorporação de inteligência humana em máquinas. IA é um conceito amplo, que engloba desde os primeiros algoritmos, na década de 1950, que envolviam a codificação de regras rígidas para a resolução de problemas específicos, como jogar xadrez, até os algoritmos modernos de *machine learning*, mais flexíveis, como *deep learning* e aprendizado por reforço, capazes de reconhecer fala e dirigir carros autônomos.^[87,88,89,90]

Machine learning é uma técnica de implementação de IA em computadores para que sejam capazes de desempenhar uma tarefa a partir de um conjunto de dados, sem que haja necessidade de programação explícita para cada caso. As técnicas de *machine learning* têm-se mostrado ótimas para resolver problemas que exigem muitos ajustes manuais ou longas listas de regras, onde atuam simplificando o processo e melhorando a performance. Além disso, os algoritmos de *machine learning* têm mostrado excelente capacidade para resolução de problemas complexos para os quais não existe uma boa solução usando os métodos tradicionais e têm demonstrado capacidade de adaptação a novos dados e a grandes quantidades de dados.^[87,88,89,90]

Existem diversos algoritmos de *machine learning* para uma infinidade de tarefas. Uma das formas mais convenientes de classificação dos diferentes tipos de algoritmos é quanto ao tipo de supervisão que recebem durante o treinamento. Nesse contexto, há quatro grandes categorias: aprendizado supervisionado, não supervisionado, semi-supervisionado e aprendizado por reforço. Neste trabalho, pela característica dos dados de entrada e pelos dados de saída esperados dos modelos, são usados algoritmos do tipo supervisionado.^[87,88,89,90]

No aprendizado supervisionado, os dados fornecidos para treinamento dos modelos incluem os dados de entrada e as saídas resultantes dessas entradas, que são conhecidas como “*labels*”. Os algoritmos supervisionados

podem ser usados para classificação, quando recebem dados com exemplos de classes, como *e-mails* classificados em *spam* ou não *spam* e precisam aprender a classificar novos *e-mails* a partir desses dados; ou também podem ser usados para regressão, quando sua função é prever um valor numérico de saída, como o preço de automóveis a partir de um conjunto de *features* (quilometragem, idade, marca, etc), chamados preditores. Para que o algoritmo seja devidamente treinado, é necessário fornecer um grande conjunto de dados que incluam os preditores e suas *labels* (nesse exemplo, o preço dos automóveis).^[87,88,89,90]

Para medir o desempenho de um algoritmo de *machine learning*, é necessário determinar a distância entre a saída de dados das predições fornecidas pelo algoritmo e a saída de dados reais, constante nos dados de treinamento. Essas medidas são usadas para ajustar os parâmetros que determinam o funcionamento do algoritmo. Essa etapa de ajuste é o que se chama de aprendizado.^[87,88,89,90]

2.4.1 Métricas de avaliação de inteligência artificial

Para avaliar a performance de um algoritmo em uma base de dados, é necessário medir o quão bem suas predições correspondem aos dados observados. Em outras palavras, é necessário quantificar a extensão a qual os valores preditos para uma dada observação é próxima do valor verdadeiro daquela observação.^[91,92,93]

O R-quadrado, também conhecido como coeficiente de determinação, representa a proporção da variância explicada pelo modelo. R-quadrado é uma métrica relativa, o que torna possível usá-la para comparar com outros modelos treinados nos mesmos dados. O valor de R-quadrado é obtido através da relação entre a soma de quadrados total (SQT) e a soma dos quadrados dos resíduos (SQR), de acordo com as equações 2 a 4:^[91,92,93]

$$SQT = \sum_{i=0}^n (y_i - \bar{y})^2 \quad \text{Eq. 2,}$$

$$SQR = \sum_{i=0}^n (y_i - \hat{f}(x_i))^2 \quad \text{Eq. 3,}$$

$$R^2 = 1 - \frac{SQT}{SQR} \quad \text{Eq. 4.}$$

Para métodos de regressão, a medida mais comum é o erro quadrático médio (mean squared error, MSE), dado por:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad \text{Eq. 5.}$$

Onde $\hat{f}(x_i)$ é a predição que \hat{f} faz para a i -ésima observação. O valor de MSE será pequeno se as respostas preditas forem muito próximas às respostas verdadeiras e será alto se em algumas das observações, a predição e a resposta forem substancialmente diferentes.^[91,92,93]

A fim de se obter o valor do erro na mesma unidade das predições e observações, é pode-se efetuar a radiciação no valor de MSE, obtendo-se a Raiz do Erro Quadrático Médio (Root Mean Square Error, RMSE):

$$RMSE = \sqrt{MSE} \quad \text{Eq. 6.}$$

O Erro Médio Absoluto (Mean Absolute Error, MAE) é a métrica de interpretação mais simples em problemas de regressão. Ao contrário do RMSE, predições ruins não resultam em valores de MAE desproporcionalmente altos. A principal razão disso é que no RMSE os erros são elevados ao quadrado, fazendo com que a presença de valores de erros mais elevados tenham impacto maior na métrica.^[91,92,93] O MAE é dado por:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{f}(x_i)| \quad \text{Eq. 7.}$$

Tanto o MAE quanto o RMSE expressam um erro médio dos modelos de predição na variável de interesse. Ambas as métricas podem variar de zero ao

infinito e são indiferentes à direção dos erros. Quanto menores forem os valores de MAE e RMSE, melhor é a performance do modelo para um dado grupo de dados.^[91,92,93]

2.4.2. Regressão Linear

A regressão linear é um método amplamente usado para prever uma resposta quantitativa, que tem sido usada por um longo tempo e é tópico de inúmeros livros-textos. Esse método serve como ponto de partida para muitas técnicas sofisticadas de aprendizado estatístico, que muitas vezes podem ser vistas como generalizações ou extensões da regressão linear. A regressão linear é uma abordagem direta para prever uma resposta quantitativa Y com base em um único preditor X , assumindo que a relação entre X e Y é aproximadamente linear.^[88] Matematicamente, descreve-se como:

$$Y \approx \beta_0 + \beta_1 X \quad \text{Eq. 8,}$$

onde β_0 e β_1 são duas constantes desconhecidas que representam a interceptação da reta com o eixo vertical e a inclinação da reta em relação ao eixo horizontal. Esses valores são conhecidos como coeficientes ou parâmetros do modelo. Os coeficientes obtidos a partir do conjunto de dados de treinamento são denominados $\hat{\beta}_0$ e $\hat{\beta}_1$ e a resposta quantitativa prevista pelo modelo pode ser obtida a partir da equação:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X \quad \text{Eq. 9,}$$

onde \hat{y} indica uma predição de Y com base em $X = x$. O acento circunflexo (^) é usado para denotar o valor estimado de um coeficiente ou parâmetro ou o valor predito para uma resposta quantitativa.

2.4.3. Regressão multilinear

Quando há mais de um preditor para uma resposta quantitativa, ao invés de ajustar modelos de regressão linear simples para cada um, é mais eficaz estender o modelo de regressão de modo que possa acomodar simultaneamente múltiplos preditores. Assim, cada preditor assume o valor de um coeficiente de inclinação da reta em relação ao eixo horizontal em um único modelo.^[88] Esse modelo de regressão linear múltipla assume a forma:

$$Y = \beta_0 X_0 + \beta_1 X_1 + \dots + \beta_p X_p + \epsilon \quad \text{Eq. 10,}$$

onde X_p representa o p -ésimo preditor e β_p quantifica a associação entre a variável e a resposta quantitativa. O valor de β_p é interpretado como sendo o efeito médio em Y do aumento de uma unidade em X_p , mantendo todos os outros preditores fixos.

2.4.4. Métodos de reamostragem

Métodos de reamostragem são ferramentas indispensáveis para o aprendizado estatístico. Esses métodos permitem que amostras sejam retiradas repetidamente de um conjunto de dados de treinamento e que seja feito um ajuste para cada amostra, com o objetivo de obter informações adicionais sobre o modelo ajustado. Pode-se por exemplo, estimar a variabilidade de um ajuste de regressão linear retirando-se repetidas amostras dos dados de treinamento, ajustando uma regressão para cada amostra e avaliando a divergência entre os ajustes. Esse tipo de informação não seria obtida caso fosse realizado um único ajuste usando a totalidade do conjunto de dados de treinamento.^[88]

O custo computacional dos métodos de reamostragem pode ser alto, principalmente porque estes envolvem ajustar o mesmo método múltiplas vezes usando subconjuntos dos dados de treinamento. Entretanto, com os avanços de hardware em poder computacional, esses custos deixaram de ser proibitivos.^[88]

Em um modelo de aprendizado estatístico, o erro de teste é o erro médio resultante do uso deste método para prever uma resposta em uma nova observação, enquanto o erro de treino é resultante do erro médio deste método quando usado para estimar as respostas oriundas do conjunto de dados de treinamento. Para estimar o erro de teste associado com um dado modelo de aprendizado estatístico a fim de avaliar sua performance ou selecionar o nível de flexibilidade, pode-se usar um método de validação cruzada. Neste método, um conjunto de dados é separado em dados e mantido fora do processo de treinamento do algoritmo que, após treinado, é aplicado nessas observações para se estimar o erro de teste.^[88]

2.4.5. Regressão Polinomial

Para estender as configurações da regressão linear para conjuntos de dados nos quais a relação entre os preditores e a resposta é não-linear, usa-se uma função polinomial:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \dots + \beta_d x_i^d + \epsilon_i \quad \text{Eq. 11,}$$

onde ϵ_i é o termo do erro. Para valores suficientemente grandes de d é possível produzir curvas extremamente não-lineares. Os coeficientes β podem ser estimados a partir do método dos mínimos quadrados. Entretanto, é incomum que se use um valor de d maior que 3 ou 4, pois para grandes valores de d a curva polinomial se torna flexível demais e pode gerar overfit.^[88]

2.4.6. Regressão segmentada e smoothing splines

Ao invés de usar uma função polinomial de alto grau sobre todo o eixo de X , a regressão segmentada ajusta polinomiais de baixo grau, separadas em diferentes regiões de X .^[88] Por exemplo, uma regressão polinomial segmentada cúbica funciona ajustando uma função cúbica:

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \epsilon_i \quad \text{Eq. 12,}$$

onde os coeficientes β_0 , β_1 , β_2 , e β_3 são diferentes nas diferentes partes de X . Os pontos onde esses coeficientes mudam são chamados de nós. Uma regressão polinomial cúbica segmentada com um único nó no ponto c assume a forma de:

$$y_i = \begin{cases} \beta_{01} + \beta_{11}x_i + \beta_{21}x_i^2 + \beta_{31}x_i^3 + \epsilon_i & \text{if } x_i < c; \\ \beta_{02} + \beta_{12}x_i + \beta_{22}x_i^2 + \beta_{32}x_i^3 + \epsilon_i & \text{if } x_i > c. \end{cases} \quad \text{Eq. 13.}$$

Em resumo, são ajustadas duas funções polinomiais diferentes a cada conjunto de dados, para $x_i < c$ ou para $x_i > c$. A primeira função polinomial tem coeficientes β_{01} , β_{11} , β_{21} , β_{31} e a segunda tem coeficientes β_{02} , β_{12} , β_{22} , β_{32} . Cada uma dessas funções polinomiais pode ser ajustada usando o método dos mínimos quadrados aplicados aos preditores originais. Aumentar o número de nós aumenta a flexibilidade das polinomiais segmentadas, sendo que para um valor K de nós, são ajustadas $K+1$ diferentes funções polinomiais.^[88]

É possível impor condições ao ajuste fornecido por uma polinomial segmentada. Uma dessas restrições é que a curva resultante deve ser contínua, sem que haja um “pulo” nos valores de y obtido nos nós. Porém, essa restrição pode levar à formação de um “V” nos pontos dos nós, gerando um aspecto não natural à curva de regressão. Para suavizar essa região, duas restrições adicionais devem ser adicionadas, fazendo com que a primeira e a segunda derivada da curva de regressão também devam ser contínuas. Dessa forma, a curva de regressão não somente é contínua, mas também suave (*smooth*).^[88]

Uma *spline* de grau d pode ser definida como sendo uma polinomial segmentada de grau d , com continuidade nas derivadas até $d-1$ em cada nó. Assim, uma *spline* cúbica é obtida pelo ajuste de uma polinomial cúbica com continuidade em cada nó até a segunda derivada.^[88] Essa *spline* cúbica pode ser modelada por:

$$y_i = \beta_0 + \beta_1 b_1(x_i) + \beta_2 b_2(x_i) + \dots + \beta_{K+3} b_{K+3}(x_i) + \epsilon_i \quad \text{Eq. 14.}$$

A seleção das funções base b_1, b_2, \dots, b_{K+3} pode ser realizada por ajuste de mínimos quadrados. A forma mais direta de representar uma *spline* cúbica é começar com uma base para uma função polinomial cúbica, como x, x^2, x^3 e então adicionar uma função potência base truncada pelo nó.^[88] Uma função potência base truncada é definida por:

$$h(x, \xi) = (x - \xi)_+^3 = \begin{cases} (x - \xi)^3 & \text{if } x > \xi \\ 0 & \text{otherwise} \end{cases} \quad \text{Eq. 15,}$$

onde ξ é o nó. Outra forma de produzir uma *spline* é a partir da minimização do somatório:

$$\sum_{i=1}^n (y_i - g(x_i))^2 + \lambda \int g''(t)^2 dt \quad \text{Eq. 16,}$$

onde λ é um parâmetro de ajuste não negativo, o termo $\sum_{i=1}^n (y_i - g(x_i))^2$ é uma *loss function* que encoraja a função a se ajustar aos dados, enquanto o termo $\lambda \int g''(t)^2 dt$ é um termo que penaliza a variabilidade da função. A função que minimiza esse valor é conhecida como *smoothing spline*.^[88]

2.4.7. Métodos Baseados em Árvores de Decisão

Os métodos baseados em árvores de decisão consistem em segmentar o espaço dos preditores em um número de regiões simples. A predição para uma dada observação é feita a partir da média ou da moda das observações de treinamento na região à qual ela pertence.^[88]

O processo de construção de uma árvore de regressão envolve duas etapas:^[88]

1. Divisão do espaço dos preditores, ou seja, o conjunto de valores possíveis para X_1, X_2, \dots, X_p em J regiões diferentes e não sobrepostas: R_1, R_2, \dots, R_J ;
2. A mesma predição, que é a média dos valores de resposta para as observações de treinamento, é feita para todas as observações que caem numa região R_j .

Na Etapa 1, o objetivo é encontrar regiões R_1, \dots, R_J que minimizem o valor de RSS dado por:

$$\sum_{j=1}^J \sum_{i \in R_j} (y_i - \widehat{y}_{R_j})^2 \quad \text{Eq. 17,}$$

onde \widehat{y}_{R_j} é a resposta média das observações de treinamento dentro da j -ésima caixa. A divisão do espaço amostral é feita por uma abordagem *top-down*, onde o espaço dos preditores começa com todas as possibilidades e vai sendo dividido em J caixas. No processo de construção da árvore, a cada divisão, o objetivo é minimizar o valor de RSS para cada etapa em particular e não necessariamente da árvore como um todo.^[88]

Apesar dos métodos baseados em árvores de decisão serem de simples interpretação, eles não são competitivos em termos de acurácia de predição. Entretanto, existem métodos de combinação de um grande número de árvores que resultam em uma melhora dramática na capacidade de predição, porém com perdas na interpretação dos modelos. Métodos como *bagging*, *random forests* e *boosting* usam árvores como blocos para construir modelos de maior poder preditivo.^[88]

2.4.8. Bagging

Um dos problemas das árvores de decisão é que elas possuem uma alta variância, o que significa que se dividirmos os dados de treinamento aleatoriamente em duas partes e ajustar uma árvore de decisão para cada metade, os resultados podem ser bem diferentes. Um bom método possui baixa variância, em contrapartida, tem resultados semelhantes quando aplicado repetidamente às divisões dos dados de treinamento.^[88]

Para reduzir a variância das árvores de regressão e conseqüentemente melhorar a sua performance, pode-se aplicar um procedimento conhecido como *bootstrap*. *Bootstrap* (em inglês, alça de bota) é um termo derivado de outro, “levantar a si próprio pela alça da bota”, que, no século XIX era usado para denominar tarefas impossíveis. Em *machine learning*, *bootstrap* envolve retirar repetidas amostras de um único conjunto de dados de treinamento, gerando, na prática, vários conjuntos de dados a partir de um só. Esses conjuntos são usados para se construir várias árvores de regressão separadamente para cada conjunto e se fazer uma média entre as predições resultantes. Esse procedimento é conhecido como *bagging*.^[88] Assim, podemos calcular $\widehat{f}^1(x)$, $\widehat{f}^2(x)$, ..., $\widehat{f}^B(x)$ usando B conjuntos de treino separados e fazendo a média entre eles para obter um único modelo de aprendizado estatístico de baixa variância, dado por:

$$\widehat{f}_{avg}(x) = \frac{1}{B} \sum_{b=1}^B \widehat{f}^b(x) \quad \text{Eq. 18.}$$

O *bagging* pode ser usado em diversos métodos de regressão, mas é particularmente útil para árvores de decisão. Individualmente, cada árvore possui alta variância, mas baixo viés. Ao usar a média de centenas ou milhares de árvores, essa variância é reduzida, melhorando o desempenho do conjunto como um todo.^[88]

2.4.9. Random Forest

Assim como no *bagging*, a *random forest* constrói um grande número de árvores de regressão com amostras dos dados de treinamento produzidas por *bootstrap*. Entretanto, durante a construção das árvores, na etapa de divisão, uma amostra aleatória de m preditores é escolhida para o espaço dos preditores

a partir de um conjunto completo de p preditores. Essa divisão pode usar somente um desses m preditores. A cada nova divisão formada pela árvore, uma nova amostra de tamanho m é tomada, onde normalmente $m = \sqrt{p}$. Assim, durante a construção de uma *random forest*, o algoritmo não pode considerar a maioria dos preditores. Esse tipo de amostragem faz com que seja possível descorrelacionar as árvores, principalmente quando um dos preditores é muito mais influente que os demais. Assim, a principal diferença entre *bagging* e *random forests* é no tamanho do subconjunto de preditores. Quando $m = p$, o método é conhecido como *bagging*. Por outro lado, quando $m < p$, trata-se de uma *random forest*, sendo que estas assumem tipicamente o valor de $m = \sqrt{p}$.^[88]

2.4.10. Boosting

Outra abordagem para melhorar as previsões das árvores de decisão é o *boosting*. Nesta técnica, as árvores são construídas sequencialmente, onde cada árvore é construída usando informação das árvores anteriores. *Boosting* não envolve amostragem *bootstrap* e cada árvore é um ajuste sobre uma versão modificada do conjunto de dados original. Neste modelo, as árvores de decisão vão sendo ajustadas aos resíduos das árvores anteriores ao invés da saída Y como resposta. Cada uma dessas árvores pode ter apenas alguns ramos finais, determinados pelo parâmetro d no algoritmo. Usar árvores pequenas nos resíduos pode melhorar \hat{f} em áreas de baixo desempenho. O parâmetro de enxugamento λ diminui a velocidade do processo ainda mais, permitindo que árvores de diferentes formatos ataquem os resíduos. Em geral, abordagens de aprendizado estatístico que aprendem lentamente apresentam bom desempenho. Ao contrário do *bagging*, no *boosting*, a construção de cada árvore depende fortemente das árvores anteriores.^[88]

2.4.11. Deep Learning

Deep learning é uma subárea de *machine learning*, uma técnica de aprendizado por representação de dados que usa sucessivas camadas de representação dos dados. Essa representação em camadas é aprendida através de modelos chamados redes neurais, estruturadas em camadas, uma após a outra. A especificação do tipo de transformação implementada por cada camada é armazenada nos pesos, que também são conhecidos como parâmetros de

uma camada. A *loss function*, também chamada de função objetiva, controla a distância entre as previsões feitas pela rede neural e o valor verdadeiro da observação nos dados de treino, expressando o desempenho da rede para esse conjunto de dados. A pontuação gerada pela *loss function* serve como um feedback para ajustar o valor dos pesos em uma direção que reduza o valor da *loss function* na próxima iteração. Esse ajuste é feito por um otimizador, que implementa o algoritmo de retropropagação, algoritmo central no *deep learning*.^[90]

Inicialmente, os pesos das camadas da rede neural recebem valores aleatórios, então a rede implementa uma série de transformações aleatórias. Assim, seus valores de saída iniciais estão longe do ideal e a pontuação da *loss function* é alta. Entretanto, a partir de cada exemplo que a rede processa, os pesos vão sendo ajustados cada vez mais na direção correta e a pontuação da *loss function* diminui. Este é o *loop* de treinamento, que se repetido um número suficiente de vezes (tipicamente dezenas de iterações com milhares de exemplos), resulta em pesos que minimizam a *loss function*. Uma rede com um *loss* mínimo é aquela que para cada saída chega o mais perto possível dos alvos: uma rede treinada.^[90] A Figura 9 mostra a relação entre inteligência artificial, *machine learning* e *deep learning*, assim como uma comparação entre os métodos clássicos de programação e *machine learning*.

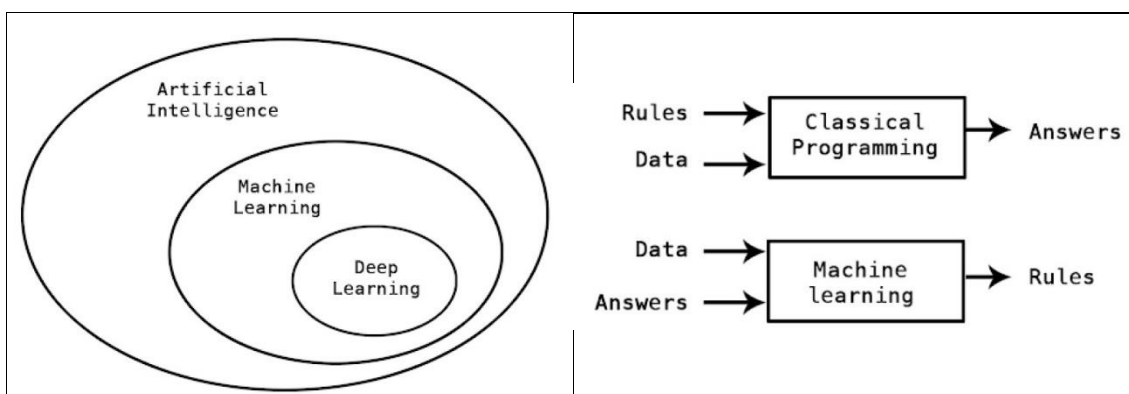


Figura 9: Relação entre inteligência artificial, *machine learning* e *deep learning* (a) e comparação do funcionamento do *machine learning* em comparação com a programação clássica (b).^[90]

2.5 LINGUAGEM R

A linguagem R é uma linguagem de programação e um ambiente para cálculos estatísticos e gráficos. Criada por Ross Ihaka e Robert Gentleman, da

University of Auckland (Nova Zelândia), iniciou como uma implementação da linguagem e ambiente S. Atualmente, é desenvolvida e atualizada pelo Núcleo de Desenvolvimento do R (*R Development Core Team*). Seu nome deriva das iniciais de seus dois criadores e de uma brincadeira com o nome da linguagem S. O projeto inicial foi concebido em 1992, com uma versão inicial lançada em 1995 e sua primeira versão beta disponibilizada em 2000. R está disponível como um software livre sob os termos da *Free Software Foundation* da GNU *General Public License* na forma de código fonte. Ele compila e roda em uma ampla variedade de plataformas UNIX e sistemas similares, Windows e MacOS.^[94]

As capacidades do R podem ser estendidas através de pacotes criados por usuários, que permitem que ferramentas estatísticas especializadas, algoritmos de inteligência artificial, além de ferramentas para importação e exportação de arquivos sejam utilizadas. Os pacotes também podem ser desenvolvidos por pesquisadores, criando um compêndio que organiza dados de pesquisa, códigos e arquivos de maneira sistemática para compartilhamento e arquivo público. Um pacote básico está incluído na instalação do R, porém existem mais de 15000 pacotes adicionais disponíveis na Comprehensive R Archive Network (CRAN), Bioconductor, Omegahat, GitHub e outros repositórios.^[94]

CAPÍTULO 3: INTELIGÊNCIA ARTIFICIAL

3.1. MÉTODOS

3.1.1. Extração de dados

Primeiramente, foi realizada uma pesquisa em motores de busca específicos e generalizados na internet, procurando por termos como “cadmium selenide” e “hot injection” para obter dados de literatura sobre CdSe. É importante ressaltar que apesar de esta busca ter encontrado muitos resultados, somente uma fração era completa o suficiente para aplicar nos modelos de *machine learning*. Procedimento análogo foi usado para pontos quânticos de CdS, PbS, PbSe e ZnSe. Em seguida, os arquivos .pdf foram baixados em uma pasta e as condições experimentais extraídas da seção experimental e os resultados de cada artigo.

A busca resultou em uma base de dados contendo 107 artigos, totalizando 842 sínteses entre todos os pontos quânticos estudados neste trabalho. Destes, foram 17 artigos de CdS, com 162 sínteses, 36 artigos de CdSe, com 358 sínteses, 25 artigos de PbS, com 115 sínteses, 17 artigos de PbSe, com 85 sínteses e finalmente 13 artigos de ZnSe, com 122 sínteses. O diâmetro dos pontos quânticos apresentados nesses artigos foram obtidos por medida direta usando Microscopia Eletrônica de Transmissão ou pela aproximação da massa efetiva usando o *bandgap* óptico a partir do espectro de absorção no UV-Visível.

Os parâmetros de reação são mostrados na Tabela 1 e um sumário de cada variável das bases de dados é dado no Anexo 1.

Tabela 1: Resumo das variáveis usadas nos modelos.

Variável	Símbolo	Unidade	Breve Definição
Temperatura de injeção	T_{inj}	°C	Temperatura da solução no balão de três bocas na hora da injeção.
Temperatura de crescimento	T_{grw}	°C	Temperatura da reação mantida após a injeção.
Tempo de reação	t	min	Tempo entre a injeção e o término da reação.
Precursor metálico	M	-	Precursor químico do metal, tipicamente, um sal.
Massa de precursor metálico	M_g	g	Quantidade, em gramas, do precursor metálico usado na reação.

Variável	Símbolo	Unidade	Breve Definição
Quantidade de íon metálico	M_{mmol}	mmol	Quantidade, em mmol, do precursor metálico usado na reação.
Concentração de íon metálico	[M]	mmol/g	Concentração do precursor metálico na reação, dado pela quantidade de íon metálico, em mmol, dividida pela massa total.
Precursor de calcogênio	C	-	Precursor químico de calcogênio.
Massa de precursor de calcogênio	C_g	g	Quantidade, em gramas, do precursor de calcogênio usado na reação.
Quantidade de calcogênio	C_{mmol}	mmol	Quantidade, em mmol, do precursor de calcogênio usado na reação.
Concentração de calcogênio	[C]	mmol/g	Concentração do precursor de calcogênio na reação, dada pela quantidade de calcogênio, em mmol, dividida pela massa total.
Razão molar metal/calcogênio	M / C	-	Razão molar entre os precursores metálico e de calcogênio, dada pela Quantidade de íon metálico, em mmol, dividida pela quantidade de calcogênio, em mmol.
Ácido carboxílico	CA	-	Ácido carboxílico usado na reação. No caso da ausência de ácido carboxílico, um "0" é usado em seu lugar na base de dados.
Massa de ácido carboxílico	CA_g	g	Quantidade, em gramas, de ácido carboxílico usado na reação.
Quantidade de ácido carboxílico	CA_{mmol}	mmol	Quantidade, em mmol, de ácido carboxílico usado na reação.
Razão molar metal/ácido carboxílico	M / CA	-	Razão molar entre íon metálico e ácido carboxílico, dada pela divisão da Quantidade de íon metálico, em mmol, dividida pela quantidade de ácido carboxílico, em mmol.
Aminas	A	-	Amina usada na reação. No caso de ausência de aminas, um "0" é usado em seu lugar na base de dados.
Massa de aminas	A_g	g	Quantidade, em gramas, de amina usada na reação.
Fosfinas	P	-	Fosfina usada na reação. No caso de ausência de fosfinas, um "0" é usado em seu lugar na base de dados.
Massa de fosfinas	P_g	g	Quantidade, em gramas, de fosfina usada na reação.
Quantidade de fosfinas	P_{mmol}	mmol	Quantidade, em mmol, de fosfina usada na reação.
Razão molar calcogênio/fosfinas	C / P	-	Razão molar entre a quantidade de calcogênio e fosfina, dada pela divisão da quantidade de calcogênio, em mmol, dividida pela quantidade de fosfinas, em mmol.

Solvente 1	S ¹	-	Solvente usado na reação. No caso de ausência de solventes, um “0” é usado em seu lugar na base de dados.
Massa do solvente 1	S ¹ _g	g	Quantidade, em gramas, de solvente usado na reação.
Solvente 2	S ²	-	Segundo solvente usado na reação, quando uma combinação de solventes foi usada. No caso de ausência deste solvente, um “0” é usado em seu lugar na base de dados.
Massa do solvente 2	S ² _g	g	Quantidade, em gramas, do segundo solvente usado na reação.
Massa total	T	g	A soma das massas de todos os componentes, em gramas.
Diâmetro	D	nm	O diâmetro final dos pontos quânticos obtidos em cada reação.

3.1.2. Feature engineering

Duas novas variáveis foram criadas, uma representando a quantidade total de pontos quânticos (CdS, CdSe, PbS, PbSe e ZnSe) que poderiam ser formados, baseados na estequiometria da reação, em mmol, e outra baseada na sua concentração, em mmol por grama, de acordo com as equações 19 e 20:

$$QD_{total} = \frac{M_{mmol} + C_{mmol}}{2} - \frac{|M_{mmol} - C_{mmol}|}{2} \quad \text{Eq. 19,}$$

$$[QD] = \frac{QD_{total}}{T} \quad \text{Eq. 20.}$$

3.1.3. Regressão multilinear

O modelo de regressão multilinear foi treinado em 50% das amostras e a validação cruzada foi feita nos outros 50%. A função `lm`, do pacote básico da linguagem R foi usado. As variáveis quantitativas foram usadas para construir o modelo: T_{inj} , T_{grw} , t , $[M]$ e $[C]$, razões entre M/C , M/CA , C/P e a concentração de pontos quânticos ($[QD]$) formados. Nenhum processo de pré-tratamento foi aplicado à base de dados.

3.1.4. Smoothing splines

A regressão por *smoothing splines* foi treinada usando a função “`smooth.spline`”, da biblioteca “`splines`” da linguagem R. Foram usadas as configurações padrão da função “`smooth.spline`” do pacote “`splines`” para a execução da regressão. Os parâmetros usados para construir o modelo foram:

T_{inj} , T_{grw} , t , razão entre M/C, quantidade de pontos quânticos formados (QD) e concentração de pontos quânticos formados ([QD]). As razões entre M/CA e C/P não foram usadas, pois a função não comporta a presença de valores faltantes (NAs).

3.1.5. Árvores de decisão

A regressão por árvores foi realizada com a função “rpart”, da biblioteca “rpart” da linguagem R. Para executar as árvores de regressão, o primeiro passo foi remover todas as instâncias da base de dados que continham valores faltantes. A variável aminas (A) foi colapsada para dois níveis: com amina e sem amina presente na reação, devido ao baixo número de observações para algumas aminas específicas. Então, a árvore de regressão foi aplicada, usando todos os preditores (qualitativos e quantitativos) para fazer a regressão. As configurações padrão da função “rpart” foram aplicadas.

3.1.6. Random forests

O mesmo pré tratamento aplicado à base de dados usado na árvore de decisão foi usada para treinar os modelos de regressão via *random forest*. A função usada para realizar esta regressão foi a “randomForest”, da biblioteca “randomForest”. O valor de m usado na regressão foi 21, os demais parâmetros foram deixados na configuração padrão. Para obtenção da importância das variáveis na regressão, a função “varImpPlot”, do pacote “randomForest” foi usada.

3.1.7. Deep learning

Para executar o modelo de *deep learning*, todos os preditores que continham valores faltantes foram removidos (razões entre M/CA e C/P). A variável precursor de calcogênio (C) foi removida da base de dados de CdSe por conter apenas um nível (Selênio). A base de dados foi dividida em 80% para dados de treinamento e 20% para dados de teste. A variável tempo de reação (t) passou por uma transformação logarítmica nas bases de dados de CdS e CdSe para aumentar a performance do modelo. Variáveis binárias fictícias foram atribuídas às variáveis qualitativas usando a função “step_dummy”. Na variável ácido carboxílico da base de dados de PbSe, por exemplo, que possui três níveis (Sem ligante, Ácido ascórbico, Ácido oleico), a função irá adicionar duas colunas

de 0/1 para dois dos três níveis, removendo a coluna original. Finalmente, as variáveis foram normalizadas para apresentar média 0, usando a função “step_scale”.

Um modelo sequencial de *deep learning* foi criado usando o pacote Keras, com 3 camadas densas. As duas primeiras camadas possuíam 64 nós e a terceira camada possuía um nó de saída (diâmetro). A função de ativação escolhida foi a “relu”, ativação linear retificada (*Rectified Linear Activation*). A *loss function* escolhida foi a MSE (*Mean Squared Error*) e o otimizador foi RMSprop.

3.2. RESULTADOS E DISCUSSÃO

As Figuras 10 a 14 mostram os histogramas com diâmetros dos pontos quânticos de CdS, CdSe, PbS, PbSe e ZnSe da base de dados. As Figuras 10(a), 11(a), 12(a), 13(a) e 14(a) mostram a distribuição total dos diâmetros, enquanto as Figuras 10(b), 11(b), 12(b), 13(b) e 14(b) mostram a distribuição de diâmetros separada por tipo de ligantes. A distribuição assimétrica possui duas diferentes explicações: pontos quânticos possuem, tipicamente, poucos nanômetros além do raio crítico de um núcleo, que determina o diâmetro mínimo para uma partícula estável;^[95] e, como podemos ver, o tamanho dos pontos quânticos está distribuído ao longo do dobro do raio de Bohr, então compostos baseados em chumbo possuem uma distribuição mais ampla, enquanto os compostos baseados em zinco ou cádmio são tipicamente menores que 10 nm. Como esperado pelos estudos de nucleação e crescimento de nanopartículas,^[96] as sínteses sem o uso de ácido carboxílico resultaram em diâmetros espalhados pelo intervalo, demonstrando o papel dos ácidos carboxílicos em melhorar o controle de tamanho.

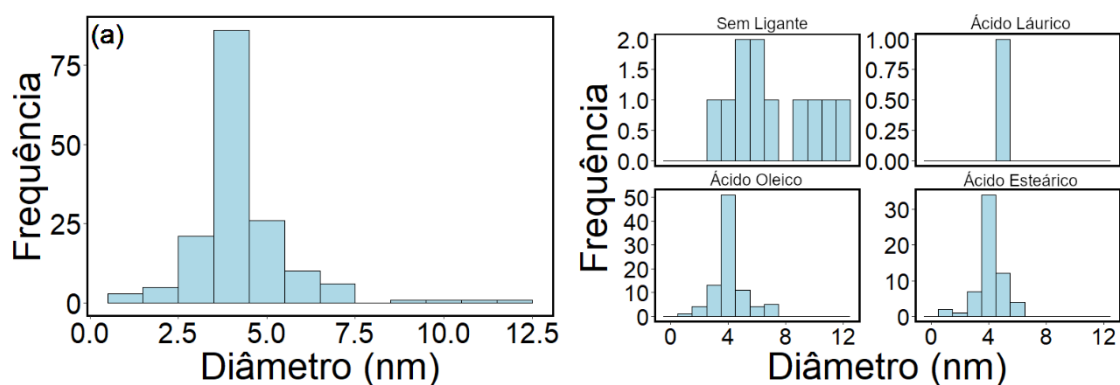


Figura 10: (a) Histograma dos diâmetros de pontos quânticos de CdS da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de CdS sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.

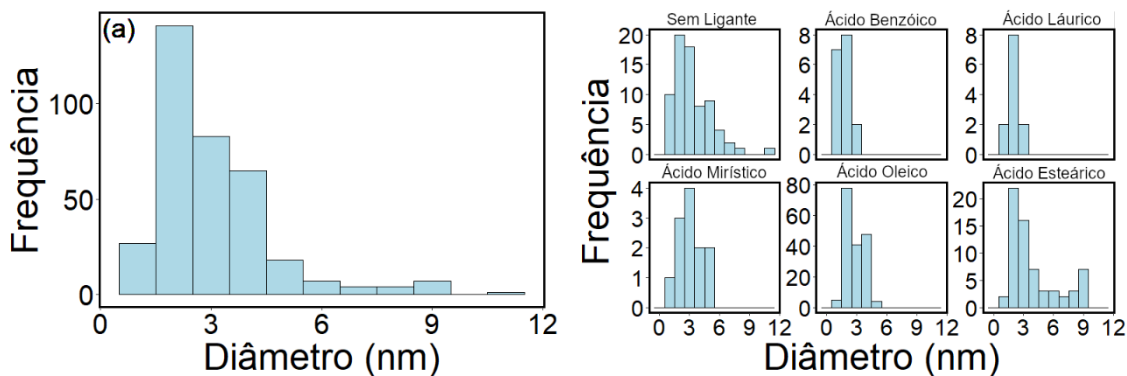


Figura 11: (a) Histograma dos diâmetros de pontos quânticos de CdSe da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de CdSe sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.

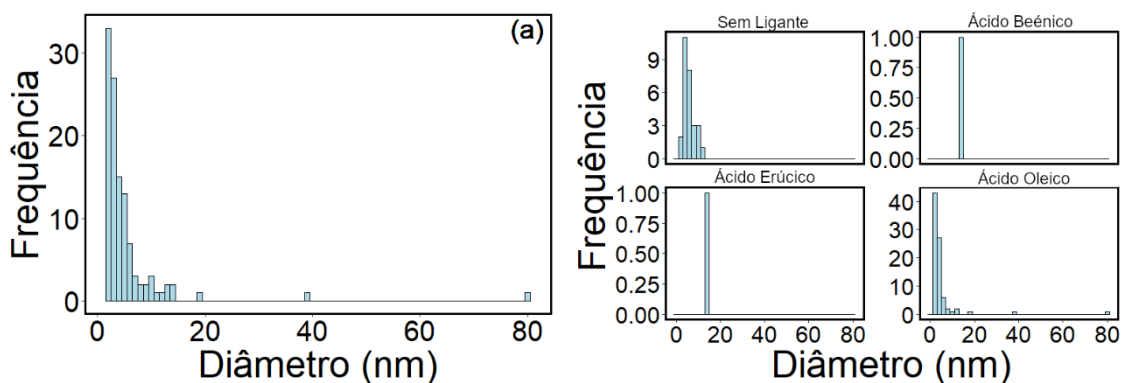


Figura 12: (a) Histograma dos diâmetros de pontos quânticos de PbS da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de PbS sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.

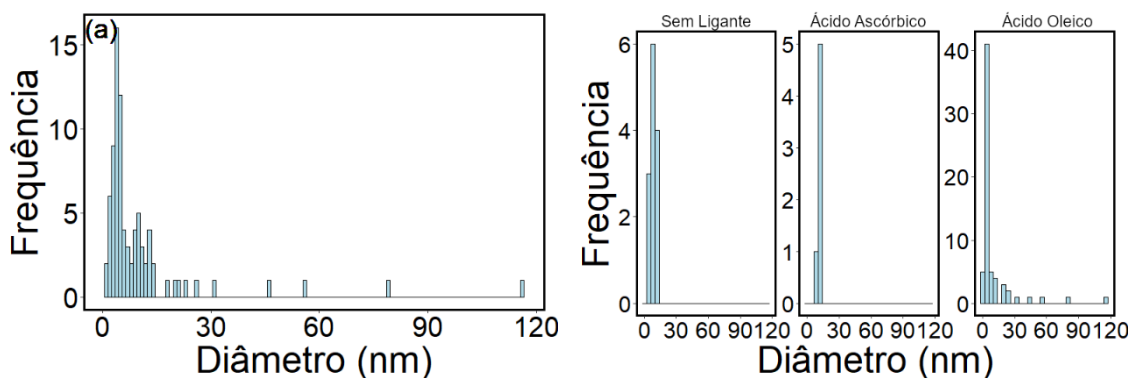


Figura 13: (a) Histograma dos diâmetros de pontos quânticos de PbSe da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de PbSe sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.

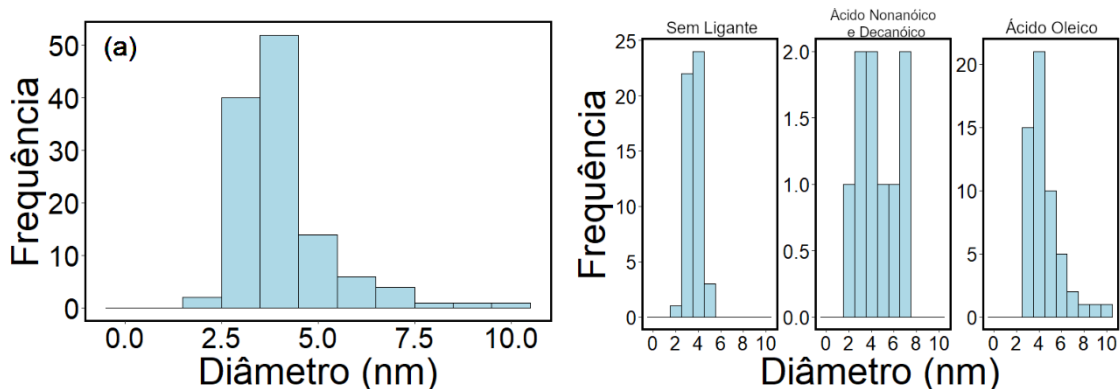


Figura 14: (a) Histograma dos diâmetros de pontos quânticos de ZnSe da literatura, incluídos nesta base de dados e (b) histogramas de pontos quânticos de ZnSe sintetizados com diferentes ácidos carboxílicos e livre de ácidos carboxílicos como ligantes.

A Figura 15 apresenta um gráfico de violino, que mostra a densidade de probabilidades dos dados apresentarem uma determinada temperatura de injeção (a) ou temperatura de crescimento (b). É possível observar que os calcogenetos de chumbo apresentam maior densidade de probabilidades em temperaturas inferiores aos compostos de cádmio e zinco. As caudas dos gráficos, com temperaturas de reação mais baixas para CdSe e ZnSe estão tipicamente relacionadas com síntese de pontos quânticos de tamanhos mágicos.^[97] Pontos quânticos de tamanhos mágicos (*magic-size quantum dots* ou *magic-size clusters*) são materiais que possuem emissão de luz de banda mais estreita do que seus análogos convencionais, com processos de nucleação e crescimento são mais complexos, ainda não bem compreendidos na comunidade científica.^[98]

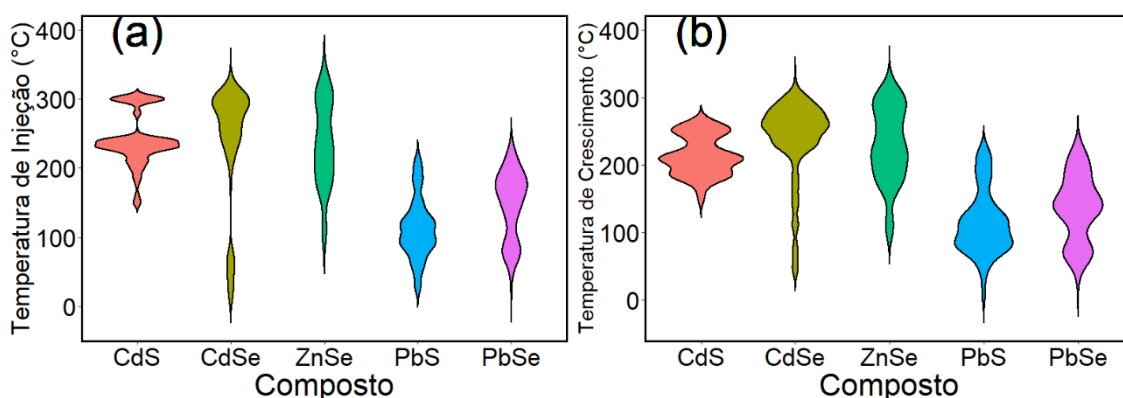


Figura 15: Gráfico de violino das temperaturas de injeção e crescimento para cada composto.

3.2.1. Regressão multilinear

Como uma ferramenta inicial para encontrar um modelo que descreva adequadamente a síntese dos pontos quânticos, a regressão multilinear foi escolhida pela sua simplicidade. A base de dados foi dividida em duas partes, 75% para o treinamento do modelo e 25% para teste. Nesse modelo, somente as variáveis contínuas foram usadas: tempo de reação, temperatura de injeção, temperatura de crescimento, concentração do precursor metálico, concentração do calcogênio, razão molar metal/calcogênio, razão molar metal/ácido carboxílico (exceto para PbS) e concentração de pontos quânticos. Neste modelo, a interação entre as variáveis não foi considerada. Isso gerou um modelo pouco flexível, em outras palavras, menos capaz de se adaptar aos dados.

Essa falta de flexibilidade pode ser vista na Figura 16, que compara os valores observados e preditos pelo modelo para pontos quânticos de CdS. Para pontos quânticos de CdS, o modelo de regressão multilinear sem interação entre os preditores apresentou valor de R-quadrado de 0,166 e 0,191 para o conjunto de dados de treino e de teste, respectivamente. O valor de R-quadrado mostra o percentual da variância dos dados que é explicada pelo modelo. Assim, os valores do modelo multilinear indicam baixa capacidade de explicação para os dados referentes a pontos quânticos de CdS. O valor de RMSE foi de 0,838 nm para dados de treinamento e 1,078 nm para dados de teste. Como mostrado na seção 2.4.1, o RMSE tende a penalizar mais modelos que apresentam uma menor quantidade de erros de grande magnitude do que modelos que apresentam uma maior quantidade de erros de menor magnitude. Observando a Figura 16(b) é possível localizar a presença de pontos distantes da linha central, apresentados em cor azul mais clara. Esses pontos representam amostras para as quais o modelo previu valores maiores do que os valores reais (à esquerda e acima da linha central) e amostras onde o modelo previu valores menores do que os valores reais (à direita e abaixo da linha central). O valor do MAE da regressão multilinear foi de 0,569 e 0,704 nm para dados de treinamento e teste, respectivamente. Esse valor é interpretado de maneira mais simples, correspondendo à média absoluta da diferença entre os valores preditos e os valores observados. Embora tanto o RMSE e o MAE possam parecer baixos, as

mudanças nas propriedades físicas dos pontos quânticos ocorrem com a variação de poucos nanômetros, o que torna esse modelo insuficiente para realizar previsões mais acuradas exigidas para o controle de tamanho do material.

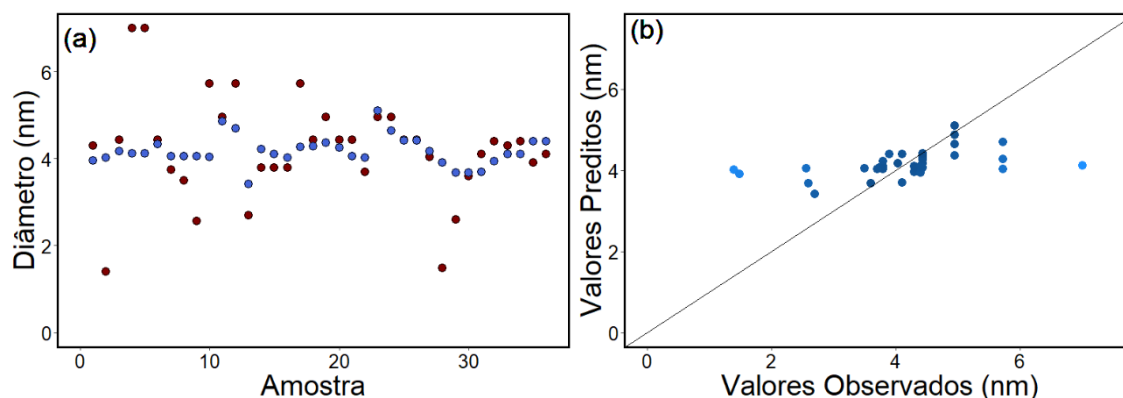


Figura 16: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de CdS.

A Figura 17 mostra a comparação entre os valores observados e preditos pelo modelo multilinear sem interações para os pontos quânticos de CdSe. O R-quadrado dos dados de teste e treinamento do modelo multilinear sem interação entre os preditores aplicado aos pontos quânticos de CdSe foi de 0,389 e 0,243, respectivamente. Assim como no caso do CdS, o modelo apresentou baixa capacidade preditiva. Isso se reflete nos valores de RMSE (1,285 nm para dados de treinamento e 1,599 nm para dados de teste) e de MAE (0,894 e 1,047 nm, para dados de treino e teste, respectivamente). Observando a Figura 17, é possível constatar que o modelo apresenta performance especialmente baixa para tamanhos de partícula superiores a 5 nm. Considerando a relação entre tamanho de partícula e propriedades ópticas dos pontos quânticos, erros dessa magnitude resultam em uma grande diferença no comprimento de onda de absorção e emissão do material.

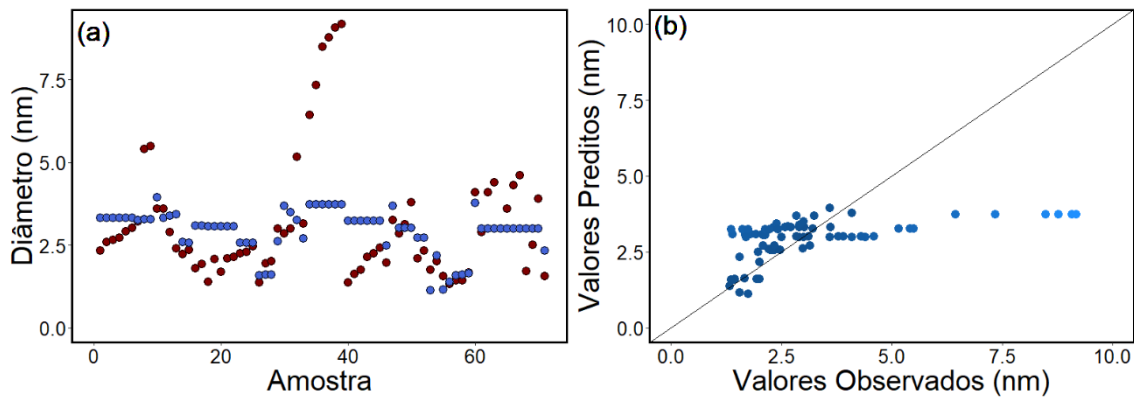


Figura 17: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de CdSe.

A comparação entre os valores observados na base de dados de pontos quânticos de PbS e as previsões por regressão multilinear sem interação entre as variáveis é mostrada na Figura 18. O modelo resultou em valores de R-quadrado de 0,272 e 0,250, para dados de treino e teste, respectivamente. Assim como nos casos de CdS e CdSe, o modelo apresentou baixa capacidade de explicar a variância dos dados. Os valores de RMSE, 1,57 e 11,96 nm, e MAE, 1,20 e 3,81 nm, para dados de treinamento e teste, respectivamente, confirmam essa tendência. O alto valor de RMSE para dados de teste (11,96 nm) se explica pela presença de *outliers* no conjunto de dados de PbS. Como sulfeto de chumbo apresenta um valor mais alto do raio de Bohr (20 nm),^[99] isso permite uma variabilidade maior de diâmetros encontrados na literatura. Aqui é possível perceber a maior robustez das métricas de desempenho RMSE e MAE, que conseguem apontar o mau desempenho do modelo em relação a esses *outliers*.

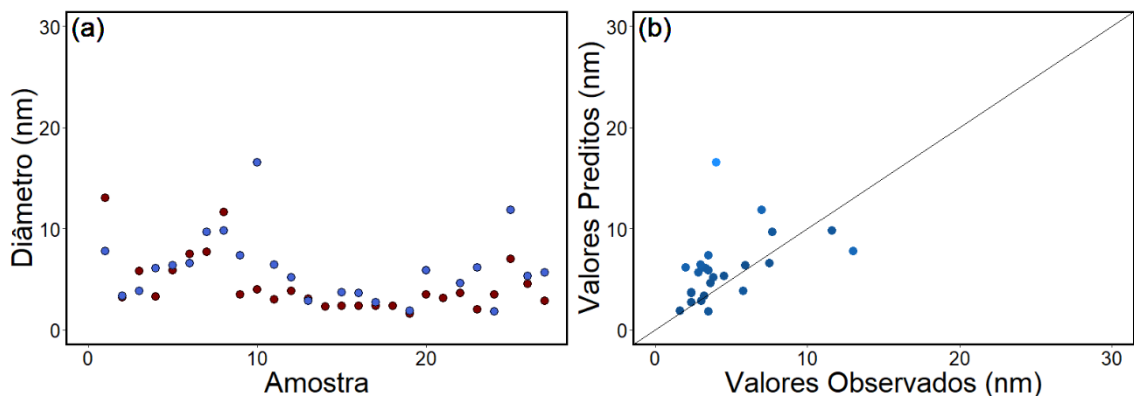


Figura 18: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de PbS.

A Figura 19 mostra a comparação entre os valores observados e preditos pelo modelo para os pontos quânticos de PbSe. Assim como a base de dados

de pontos quânticos de PbS, o conjunto de dados de PbSe também apresenta a característica de possuir *outliers* devido à maior possibilidade de tamanhos de pontos quânticos, uma vez que o raio de Bohr do PbSe é de 46 nm.^[99] O desempenho do modelo de regressão multilinear sem interação entre os preditores aplicado ao conjunto de dados de pontos quânticos de PbSe apresenta um desempenho muito semelhante ao caso de PbS. Os valores de R-quadrado ficaram em 0,411 para dados de treino e 0,332 para dados de teste, enquanto o RMSE foi de 10,34 e 23,21 nm e o MAE foi de 6,308 e 10,01 nm, para dados de treinamento e teste, respectivamente. O desempenho inferior em relação aos dados de PbS pode ser atribuído ao tamanho menor da base de dados de PbSe e variabilidade maior da base de dados ocasionada pelo maior raio de Bohr.

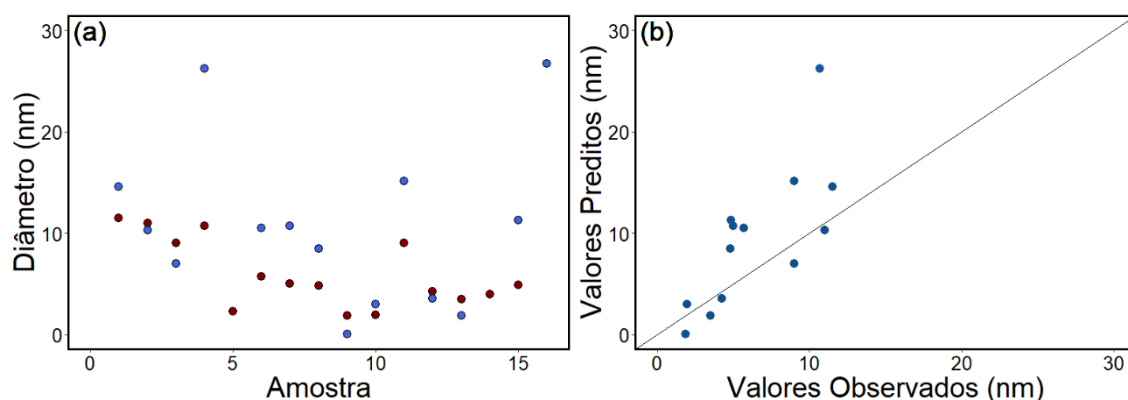


Figura 19: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de PbSe.

Por fim, o modelo multilinear sem interação entre as variáveis foi aplicado ao conjunto de dados de pontos quânticos de ZnSe. O modelo apresentou valores de R-quadrado de 0,396 e 0,257 para dados de treinamento e teste, respectivamente. Os valores de RMSE (1,026 e 1,086 nm) e MAE (0,684 e 0,787 nm) para dados de treinamento e teste, mostram um desempenho comparável com os modelos usados nos dados de CdS e CdSe. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de ZnSe é mostrada na Figura 20. Assim como no caso do modelo multilinear aplicado aos dados de CdSe, o modelo não foi capaz de realizar previsões de pontos quânticos de ZnSe de tamanhos superiores a 5 nm de diâmetro.

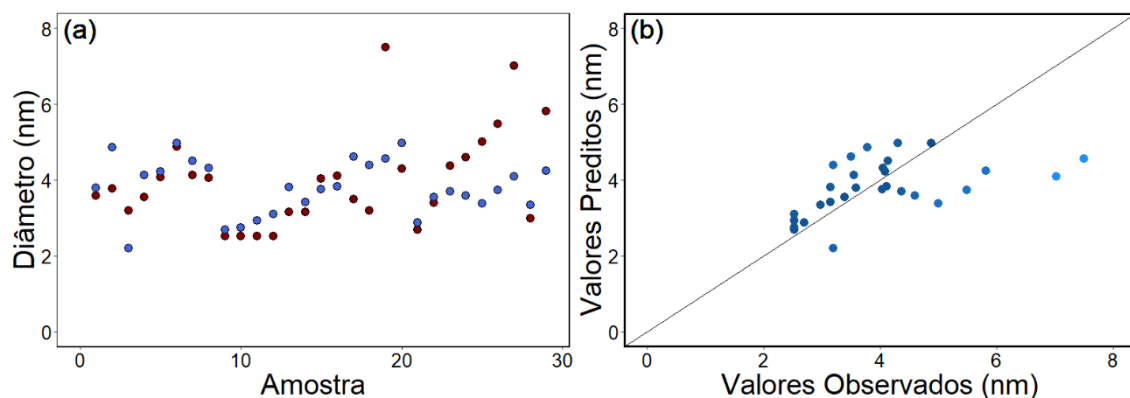


Figura 20: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear sem interações entre as variáveis para pontos quânticos de ZnSe.

O modelo de regressão multilinear com interação entre as variáveis possui uma flexibilidade muito maior que o mesmo modelo sem as interações. Assim como no modelo anterior, somente as variáveis contínuas foram usadas: tempo de reação, temperatura de injeção, temperatura de crescimento, concentração do precursor metálico, concentração do calcogênio, razão molar metal/calcogênio (exceto CdS e PbS), razão molar metal/ácido carboxílico (exceto para PbS), razão molar entre calcogênio/fosfina (somente para CdSe) e concentração de pontos quânticos (exceto CdS e PbS).

A Figura 21 mostra uma comparação entre os diâmetros dos pontos quânticos da base de dados de CdS e os valores preditos pelo modelo multilinear, considerando a interação entre as variáveis. Os valores de R-quadrado para o modelo foram 0,575 e 0,001, para os dados de treino e teste, respectivamente. Essa grande diferença é característica de *overfit*, ou seja, devido à flexibilidade do modelo, ele se ajustou bem ao conjunto de dados de treinamento, mas perde desempenho diante dos dados de teste. Os valores de RMSE (0,599 e 123,9 nm) e MAE (0,372 e 21,31 nm) confirmam que o modelo se adaptou muito bem aos dados de treinamento, mas é incapaz de generalizar o aprendizado para os dados de teste.

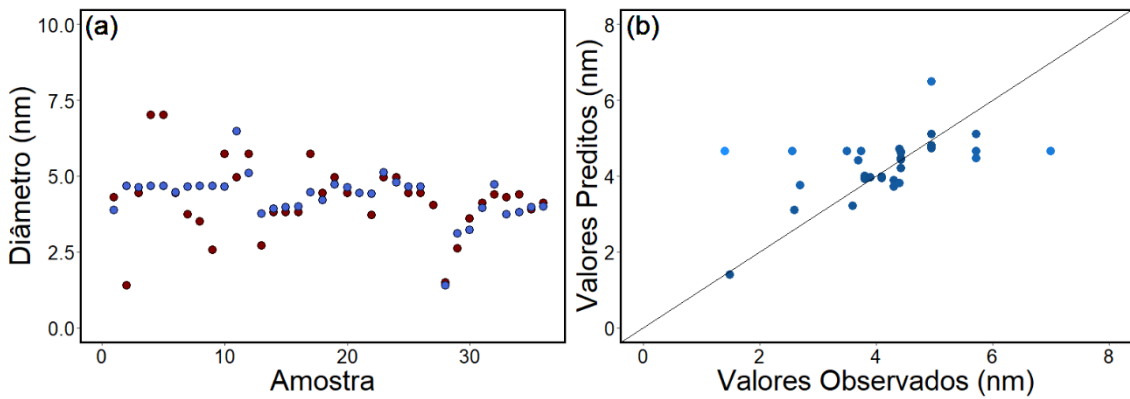


Figura 21: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de CdS.

A Figura 22 mostra a comparação entre os valores de diâmetro dos pontos quânticos de CdSe contidos na base de dados e os valores preditos pelo modelo de regressão multilinear com interação entre os preditores. O modelo mostra o seu poder de adaptação aos dados de treino, apresentando valores de R-quadrado de 0,957, RMSE de 0,299 nm e MAE de 0,187 nm. Entretanto, quando olhamos para os dados de teste, vemos um R-quadrado de 0,007, RMSE de 160,6 nm e MAE de 22,75 nm. Novamente, um típico caso de *overfit*, com um modelo perfeitamente adaptado aos dados de treinamento, porém incapaz de realizar quaisquer previsões fora deste conjunto.

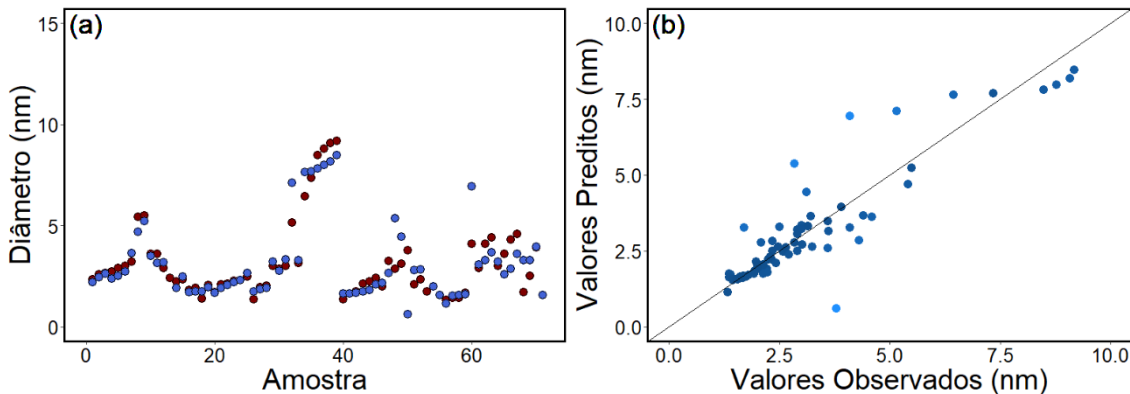


Figura 22: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de CdSe.

A regressão multilinear com interação entre as variáveis apresentou um desempenho semelhante ao do CdSe quando aplicada ao conjunto de dados de PbS. O valor de R-quadrado para os dados de treinamento foi de 0,522, com RMSE de 6,451 nm e MAE de 2,401 nm. Todavia, ao observar as métricas de desempenho do modelo aplicado ao conjunto de dados de teste (R-quadrado:

0,004, RMSE: 10,61 nm e MAE: 5,436 nm), é possível perceber que se trata de mais um caso de *overfit*. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de PbS é mostrada na Figura 23.

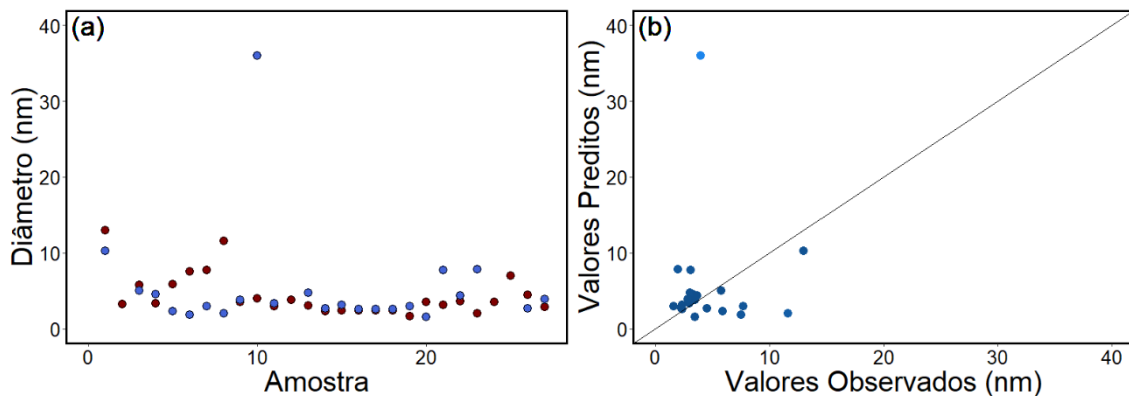


Figura 23: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de PbS.

A Figura 24 mostra a comparação entre os diâmetros dos pontos quânticos de PbSe da base de dados e as predições realizadas pelo modelo de regressão multilinear com interação entre as variáveis predictoras. Da mesma forma que nos casos anteriores, o modelo tem um desempenho fortemente degradado quando retirado dos dados de treinamento e aplicado aos dados de teste. Para os dados de treinamento, os valores de R-quadrado, RMSE e MAE foram de 0,999, 0,237 nm e 0,106 nm, enquanto para os dados de teste foram de 0,012, 3730 nm, e 1165 nm.

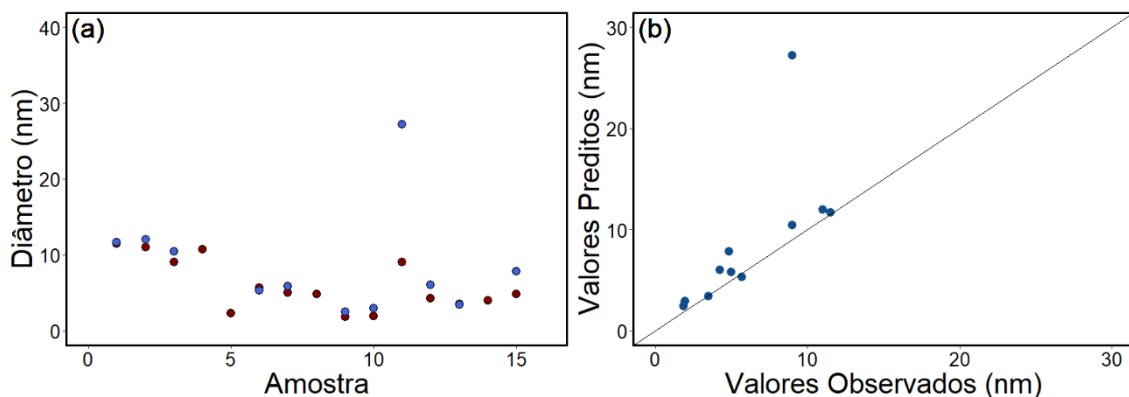


Figura 24: (a and b) Gráficos das observações vs predições pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de PbSe.

Assim como nos casos anteriores, a regressão multilinear com interação entre as variáveis resultou em *overfit* também para os dados dos pontos quânticos de ZnSe. Os valores encontrados para os dados de treinamento foram

de 0,932, 0,343 nm e 0,198 nm para R-quadrado, RMSE e MAE, respectivamente. Em contrapartida, para os dados de teste, foram encontrados valores de 0,005, 58625 nm e 14368 nm para R-quadrado, RMSE e MAE, respectivamente. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de ZnSe é mostrada na Figura 25.

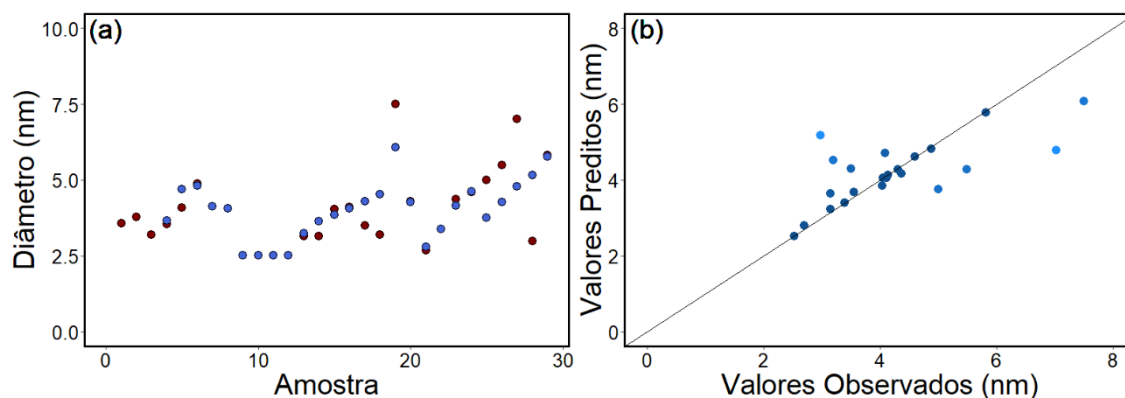


Figura 25: (a and b) Gráficos das observações vs previsões pelo modelo de regressão multilinear com interações entre as variáveis para pontos quânticos de ZnSe.

3.2.2. Regressão não-linear

Após os modelos multilineares, foi decidido aplicar um modelo não-linear. Uma vez que modelos polinomiais falharam em convergir, foi escolhido o método *smoothing splines*. Assim como os modelos de regressão multilinear, apenas variáveis contínuas foram usadas como preditores no modelo de *smoothing splines*, que também é pouco tolerante com variáveis correlacionadas. Assim, as variáveis escolhidas foram: tempo de reação, temperatura de injeção, temperatura de crescimento, razão molar entre metal/calco gênio e concentração de pontos quânticos.

O resultado da aplicação do modelo *smoothing splines* à base de dados de CdS é mostrado na Figura 26. Os valores de R-quadrado para o modelo foram 0,391 e 0,148, para os dados de treino e teste, respectivamente. Os valores de RMSE (0,727 e 1,087 nm) e MAE (0,446 e 0,774 nm) mostram um desempenho muito semelhante ao modelo de regressão multilinear sem interações entre as variáveis, inclusive apresentando a mesma dificuldade em realizar previsões de pontos quânticos com diâmetro superior a 5 nm.

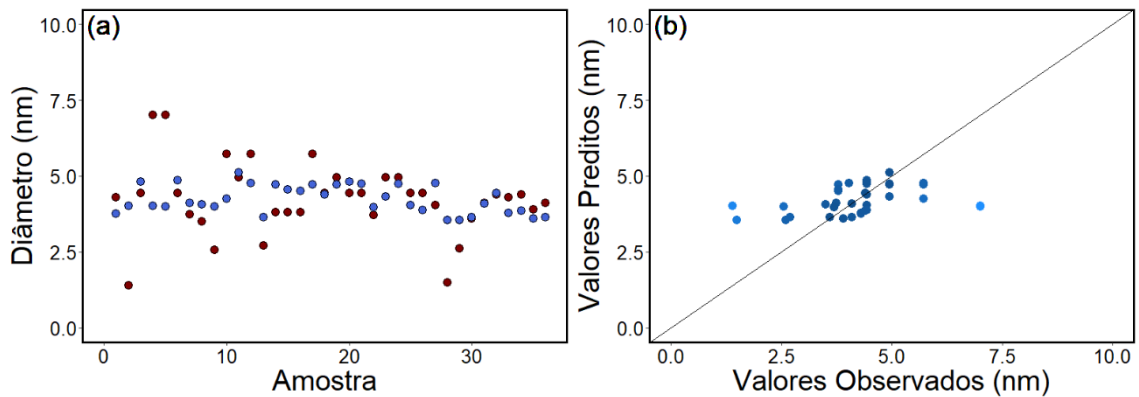


Figura 26: (a and b) Gráficos das observações vs previsões pelo modelo *smoothing splines* com interações entre as variáveis para pontos quânticos de CdS.

A Figura 27 mostra a comparação entre os valores observados e preditos pelo modelo *smoothing splines* para os pontos quânticos de CdSe. O valor de R-quadrado dos dados de teste e treinamento do modelo multilinear sem interação entre os preditores aplicado aos pontos quânticos de CdSe foi de 0,957 e 0,771, respectivamente. O modelo de *smoothing splines* apresentou um desempenho bastante superior aos modelos de regressão multilinear para o conjunto de dados de CdSe. Isso se confirma ao verificar os valores de RMSE (0,300 nm para dados de treinamento e 0,890 nm para dados de teste) e de MAE (0,201 e 0,565 nm, para dados de treino e teste, respectivamente). A Figura 27 também mostra que houve melhora no desempenho para tamanhos de partícula superiores a 5 nm.

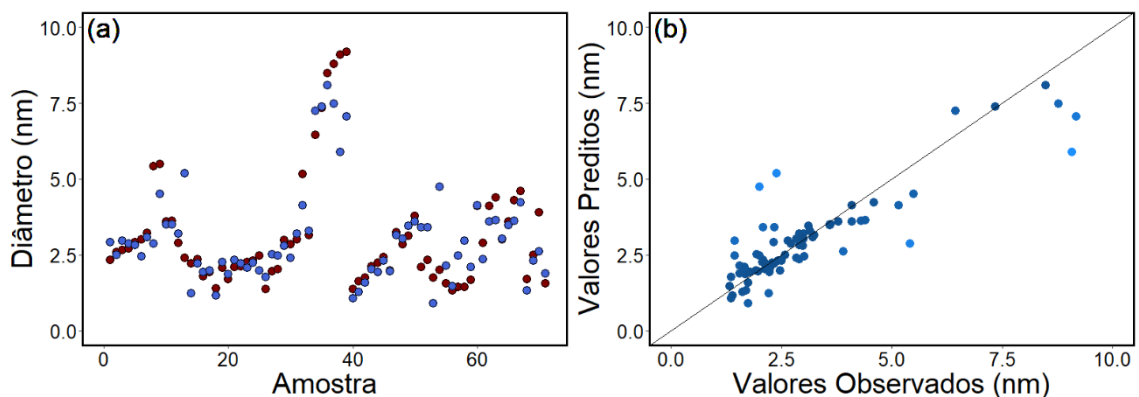


Figura 27: (a and b) Gráficos das observações vs previsões pelo modelo *smoothing splines* com interações entre as variáveis para pontos quânticos de CdSe.

A Figura 28 mostra o desempenho do modelo *smoothing splines* aplicado à base de dados de pontos quânticos de PbS. O modelo resultou em valores de R-quadrado de 0,570 e 0,094, para dados de treino e teste, respectivamente. Assim como nos casos de CdS e CdSe o modelo apresentou baixa capacidade

de explicara variância dos dados. Os valores de RMSE, 6,151 e 3,117 nm, e MAE, 2,178 e 1,955 nm, para dados de treinamento e teste, respectivamente, confirmam essa tendência. O fato dos valores de RMSE e MAE serem maiores para os dados de treinamento do que para os dados de teste é interessante. Uma das causas pode ser na divisão dos dados, onde os outliers podem ter se concentrado na fração de treinamento, fazendo com que os dados de teste fossem mais homogêneos. Essa homogeneidade pode ter permitido que o modelo apresentasse um desempenho melhor nos dados de teste em comparação com os dados de treinamento.

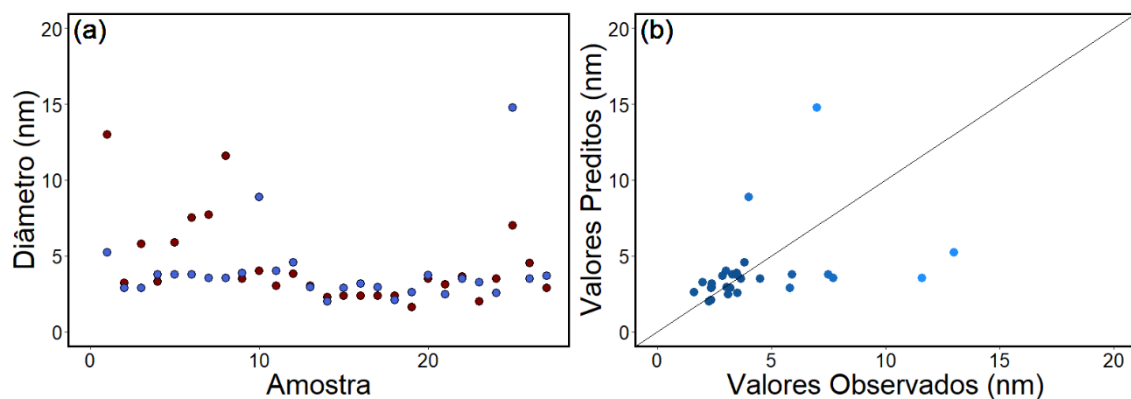


Figura 28: (a and b) Gráficos das observações vs previsões pelo modelo *smoothing splines* com interações entre as variáveis para pontos quânticos de PbS.

A Figura 29 mostra a comparação entre os valores observados e preditos pelo modelo para os pontos quânticos de PbSe. O desempenho do modelo de *smoothing splines* aplicado ao conjunto de dados de pontos quânticos de PbSe apresenta um desempenho muito semelhante ao caso de PbS. Os valores de R-quadrado ficaram em 0,553 para dados de treino e 0,109 para dados de teste, enquanto o RMSE foi de 9,208 e 26,00 nm e o MAE foi de 4,655 e 12,27 nm, para dados de treinamento e teste, respectivamente. Da mesma forma que o caso anterior, a presença de outliers na base de dados de PbSe pode ser uma das causas da melhor performance do algoritmo nos dados de teste em relação aos dados de treinamento.

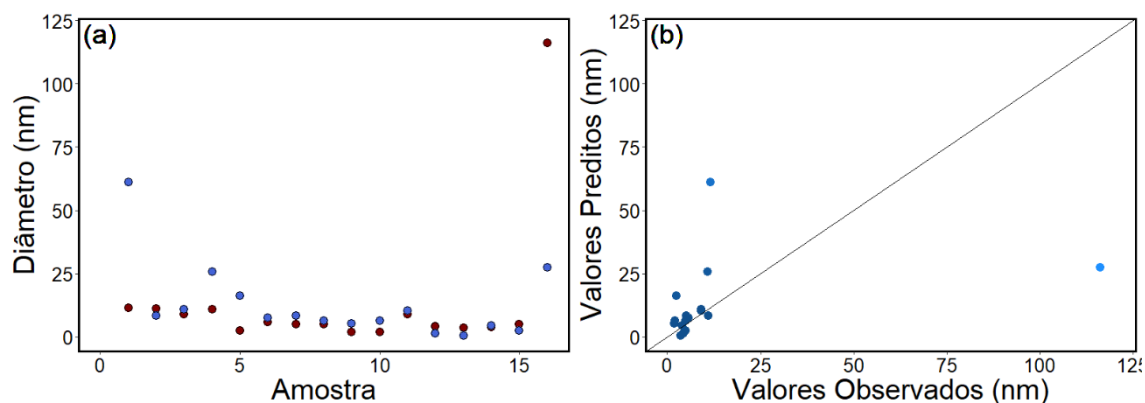


Figura 29: (a and b) Gráficos das observações vs previsões pelo modelo *smoothing splines* com interações entre as variáveis para pontos quânticos de PbSe.

Finalmente, o modelo *smoothing splines* foi aplicado aos dados de pontos quânticos de ZnSe. O modelo apresentou valores de R-quadrado de 0,664 e 0,534 para dados de treinamento e teste, respectivamente. Os valores de RMSE (0,810 e 0,860 nm) e MAE (0,539 e 0,589 nm) para dados de treinamento e teste, mostram um desempenho comparável com os modelos usados nos dados de CdSe. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de ZnSe é mostrada na Figura 20.

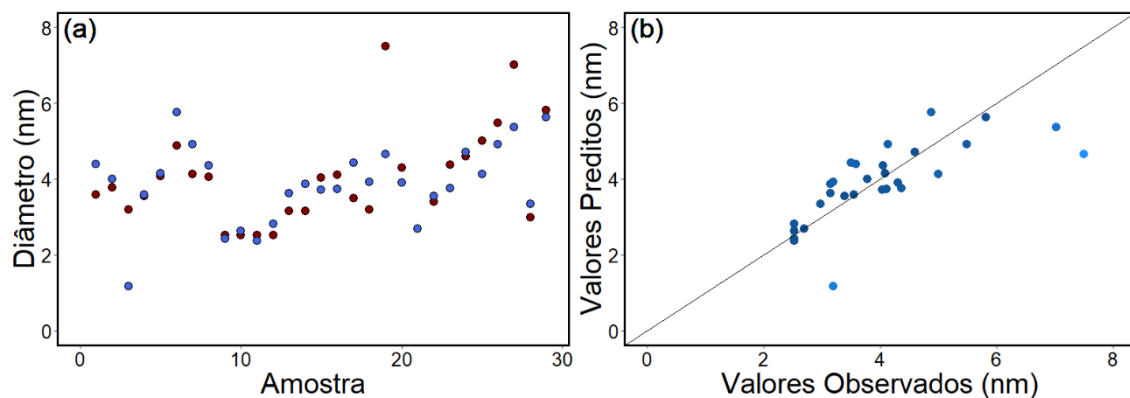


Figura 30: (a and b) Gráficos das observações vs previsões pelo modelo *smoothing splines* com interações entre as variáveis para pontos quânticos de ZnSe.

3.2.3. Árvores de regressão

As árvores de regressão, construídas a partir de todos os preditores da base de dados dos pontos quânticos, são mostradas nas Figuras 31 a 35. Os preditores que aparecem no topo da árvore podem ser interpretados como os mais fortes desse modelo. Para pontos quânticos de CdS (Figura 31), esse preditor é a concentração de calcogênio, seguido da massa total da reação. Nos pontos quânticos de CdSe (Figura 32), a temperatura de crescimento tem a

maior influência no tamanho final das partículas. Nesse segmento, amostras com temperaturas de crescimento inferiores a 295 °C passam através de mais ramos da árvore, enquanto aqueles com temperaturas superiores são agrupados em um único grupo. A quantidade de calcogênio usado na reação é o preditor mais forte para PbS, enquanto o tipo de precursor metálico desempenha o papel mais importante para PbSe. Por fim, pontos quânticos de ZnSe tiveram o seu tamanho fortemente influenciado pela quantidade de solvente, de acordo com o método da árvore de regressão.

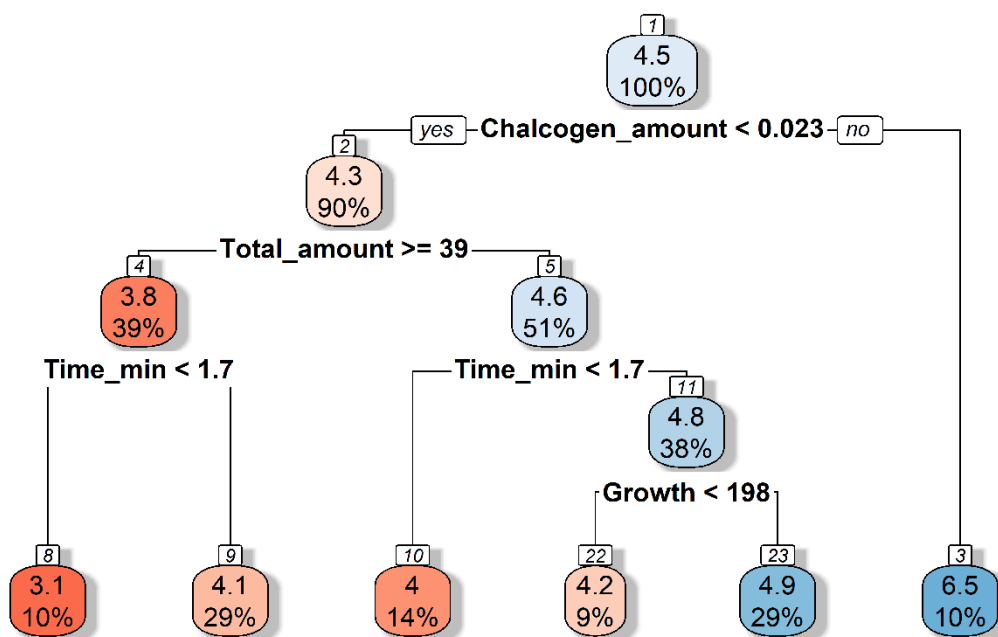


Figura 31: Árvore de regressão para o diâmetro de pontos quânticos de CdS como função de todas as variáveis da base de dados.

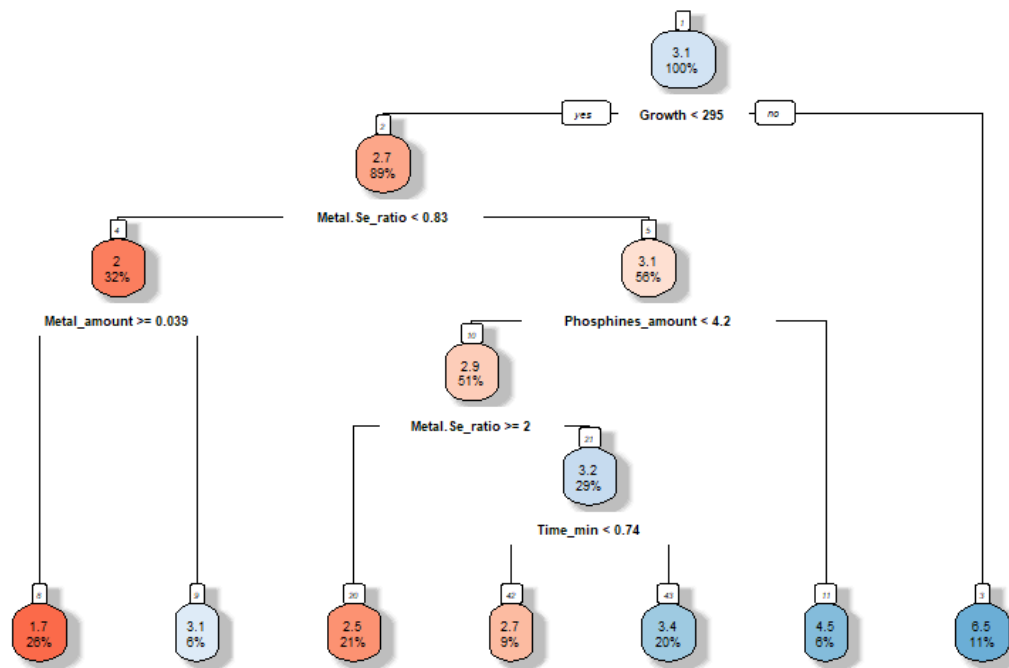


Figura 32: Árvore de regressão para o diâmetro de pontos quânticos de CdSe como função de todas as variáveis da base de dados.

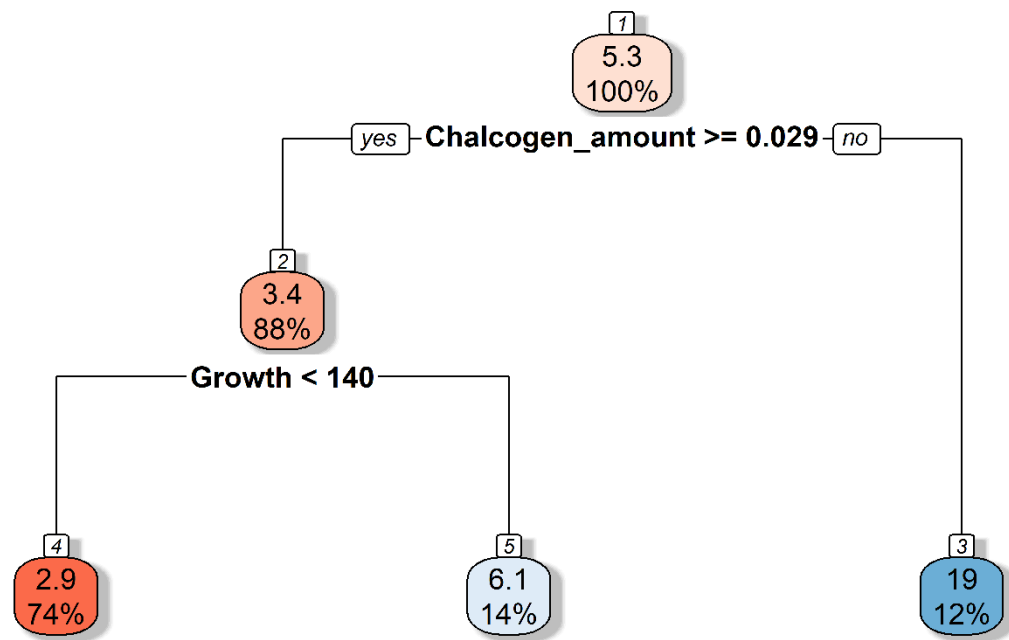


Figura 33: Árvore de regressão para o diâmetro de pontos quânticos de PbS como função de todas as variáveis da base de dados.

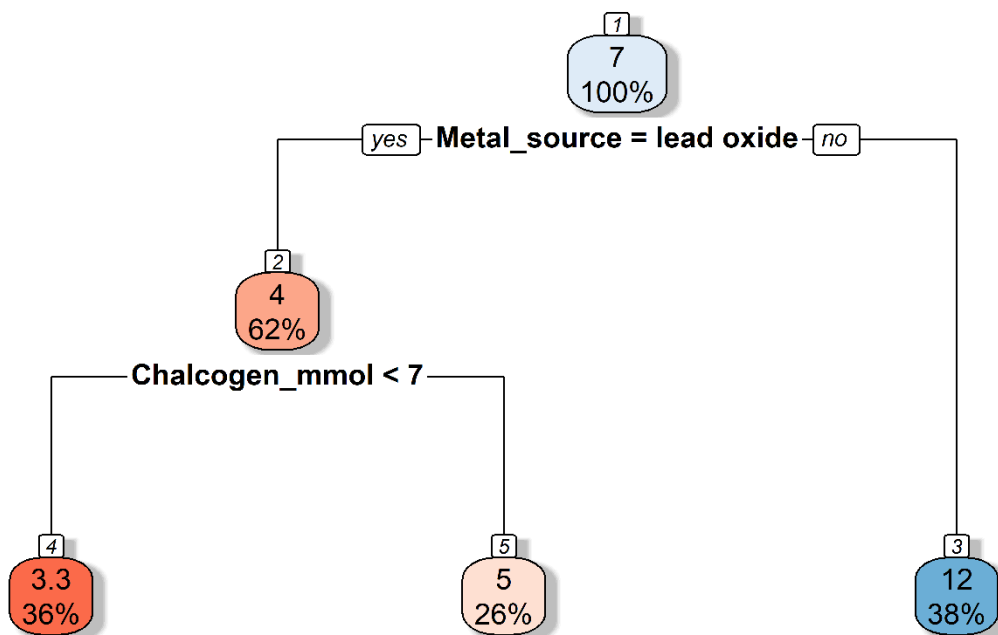


Figura 34: Árvore de regressão para o diâmetro de pontos quânticos de PbSe como função de todas as variáveis da base de dados.

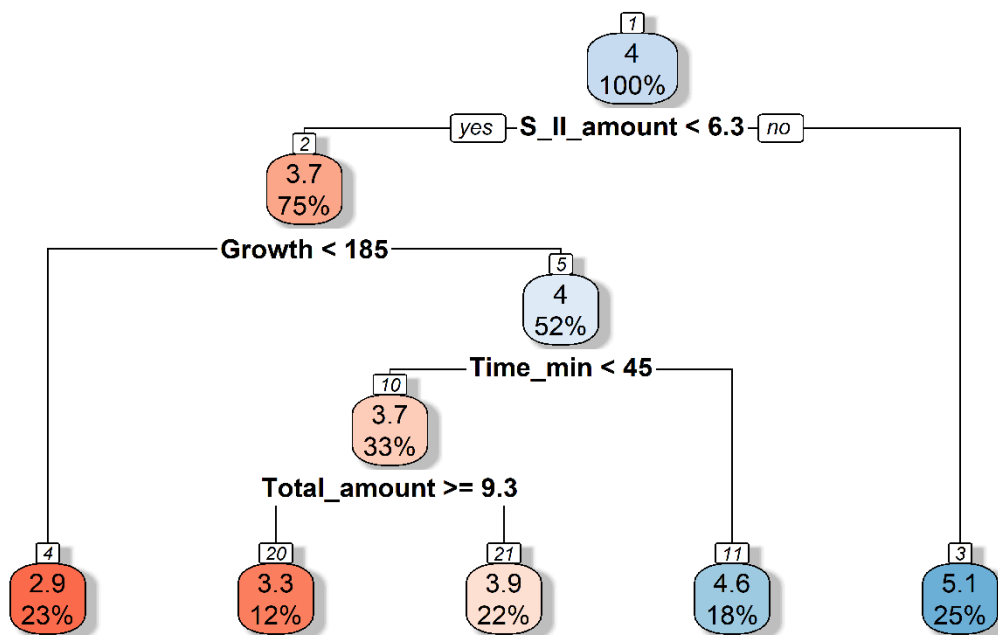


Figura 35: Árvore de regressão para o diâmetro de pontos quânticos de ZnSe como função de todas as variáveis da base de dados.

A Figura 36 apresenta o gráfico comparando os valores reais dos diâmetros dos pontos quânticos de CdS versus os valores preditos pelo modelo de árvore de regressão. Esse modelo apresentou valor de R-quadrado de 0,570,

RMSE de 0,947 nm e MAE de 0,563 nm para os dados de treinamento. Para os dados de teste, os valores foram de 0,360 para R-quadrado, 1,059 nm para RMSE e 0,665 nm para MAE. Apesar de apresentar melhor desempenho que os modelos de regressão multilinear e *smoothing splines* para os dados de teste, a performance ainda é insuficiente para realizar uma predição acurada dos diâmetros dos pontos quânticos de forma que viabilize o controle das propriedades ópticas.

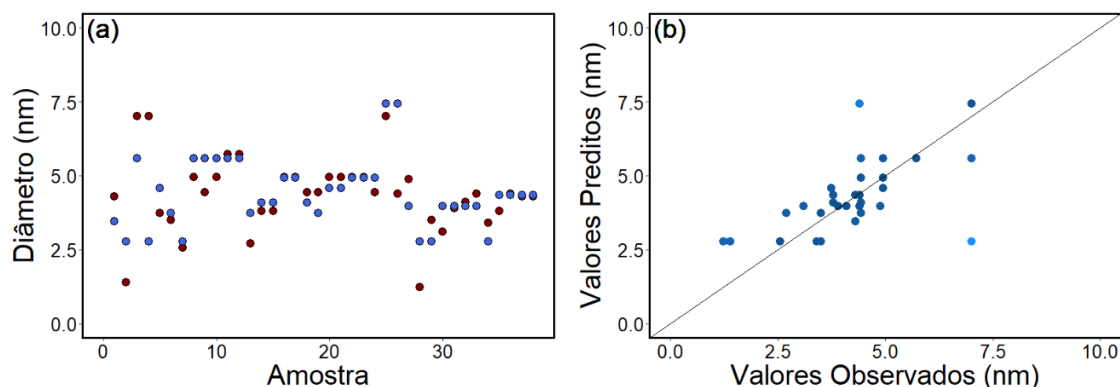


Figura 36: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de CdS.

A Figura 37 mostra a comparação entre os valores de observação e predições para os pontos quânticos de CdSe. O R-quadrado dos dados de teste e treinamento do modelo de árvore de regressão aplicado aos pontos quânticos de CdSe foi de 0,735 e 0,607, respectivamente. O modelo apresentou um desempenho semelhante ao conjunto de dados de CdS, com valores de RMSE (0,893 nm para dados de treinamento e 0,816 nm para dados de teste) e de MAE (0,570 nm, para dados de treino e teste). Novamente, erros dessa magnitude resultam em uma grande diferença no comprimento de onda de absorção e emissão do material.

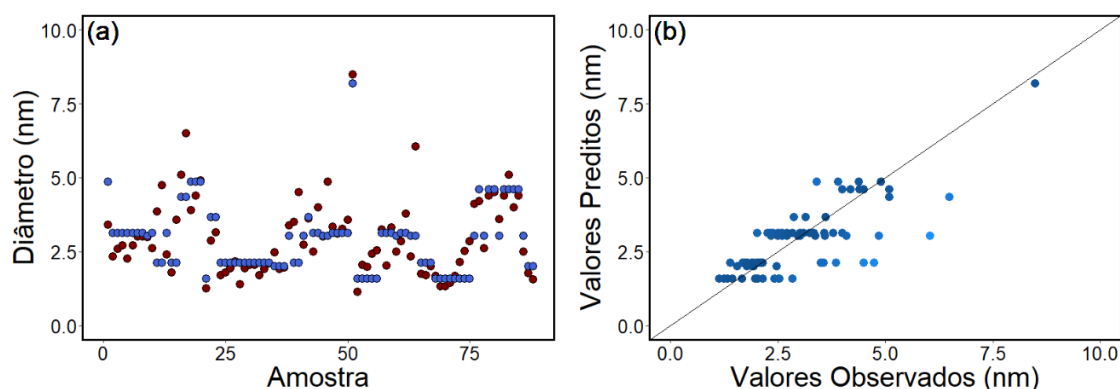


Figura 37: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de CdSe.

A árvore de regressão apresentou um desempenho inferior ao do CdSe quando aplicada ao conjunto de dados de PbS. A explicação para essa perda de desempenho é o volume de dados menor do conjunto de dados de PbS em relação ao de CdSe. O valor de R-quadrado para os dados de treinamento foi de 0,401, com RMSE de 7,221 nm e MAE de 2,435 nm. É interessante notar que as métricas de desempenho do modelo aplicado ao conjunto de dados de teste (R-quadrado: 0,582, RMSE: 1,893 nm e MAE: 1,134 nm) são, assim como o modelo de *smoothing splines*, superiores às dos dados de treinamento. Esse comportamento atípico pode ser explicado por uma eventual presença de outliers nos dados de treinamento e ausência destes nos dados de teste, assim como ocorreu no algoritmo *smoothing splines*. Essa hipótese ganha força com o fato de que a divisão de dados de treino e teste é a mesma para os algoritmos de regressão multilinear, *smoothing splines* e árvores de regressão.

A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de PbS é mostrada na Figura 38.

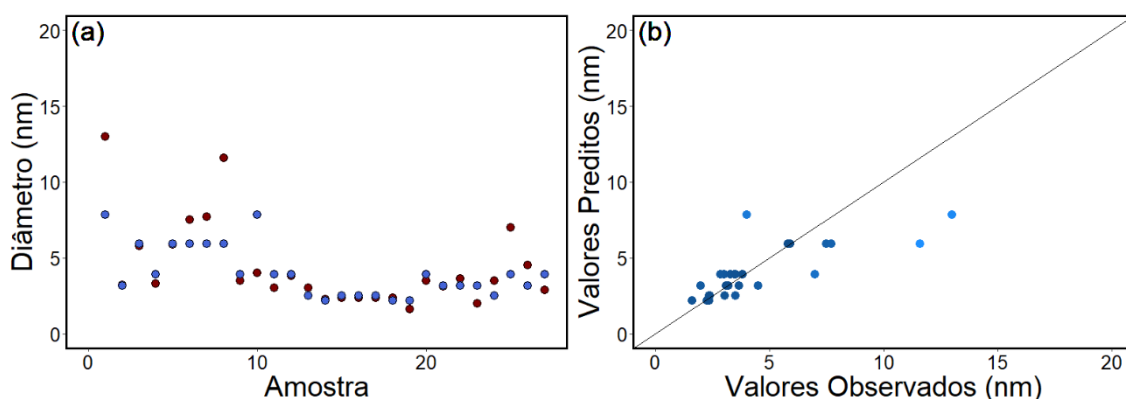


Figura 38: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de PbS.

A árvore de regressão apresentou um desempenho semelhante ao do PbS quando aplicada ao conjunto de dados de PbSe. O valor de R-quadrado para os dados de treinamento foi de 0,396, com RMSE de 13,94 nm e MAE de 5,504 nm. Da mesma forma que na árvore de regressão aplicadas aos dados de PbS, as métricas de desempenho do modelo aplicado ao conjunto de dados de teste (R-quadrado: 0,290, RMSE: 9,089 nm e MAE: 4,396 nm) apresentaram desempenho superior às dos dados de treinamento. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de PbS é mostrada na Figura 39.

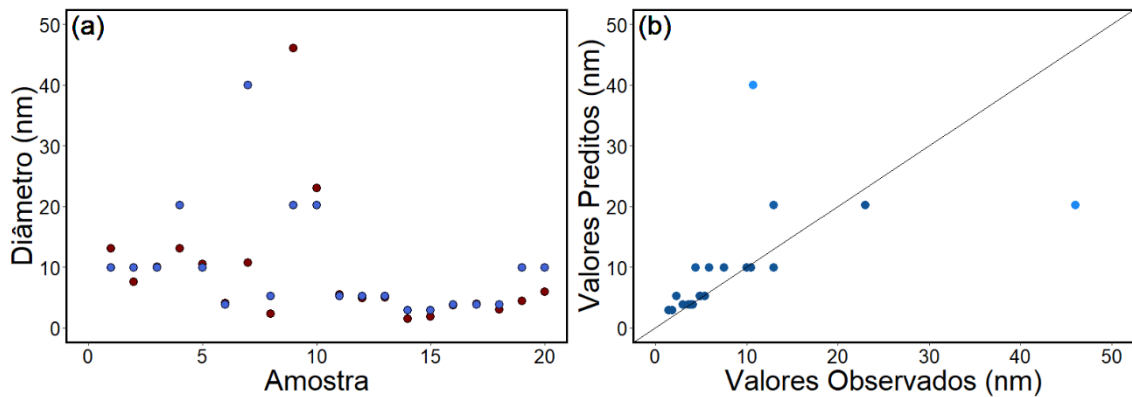


Figura 39: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de PbSe.

Finalmente, o modelo de árvore de regressão foi aplicado ao conjunto de dados de pontos quânticos de ZnSe. O modelo apresentou valores de R-quadrado de 0,677 e 0,362 para dados de treinamento e teste, respectivamente. Os valores de RMSE (0,750 e 1,008 nm) e MAE (0,532 e 0,694 nm) para dados de treinamento e teste, mostram um desempenho comparável com os modelos usados nos dados de CdS e CdSe. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de ZnSe é mostrada na Figura 40.

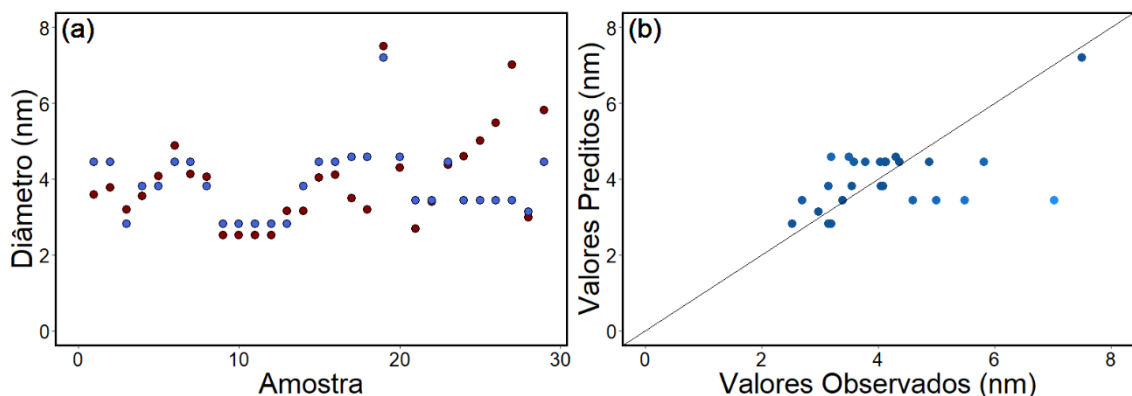


Figura 40: Gráficos das observações vs predições pelo modelo de árvore de regressão para pontos quânticos de ZnSe.

3.2.4. Random Forest

O modelo *random forest* adiciona maior flexibilidade às árvores de regressão pela adição de centenas de árvores de regressão que avaliam os diferentes preditores da base de dados. Assim como as árvores de regressão da seção anterior, os modelos de random forest desta seção foram construídos usando-se todos os preditores das bases de dados.

A Figura 41 mostra uma comparação entre os diâmetros dos pontos quânticos da base de dados de CdS e os valores preditos pelo modelo de

random forest. Os valores de R-quadrado para o modelo foram 0,413 e 0,222, para os dados de treino e teste, respectivamente. Essa diferença pode ser um indicativo de *overfit*. Os valores de RMSE (1,267 e 1,133 nm) e MAE (0,748 e 0,771 nm) mostram que o modelo se adaptou muito bem aos dados de treinamento, mas ainda não tem a performance ideal nos dados de teste.

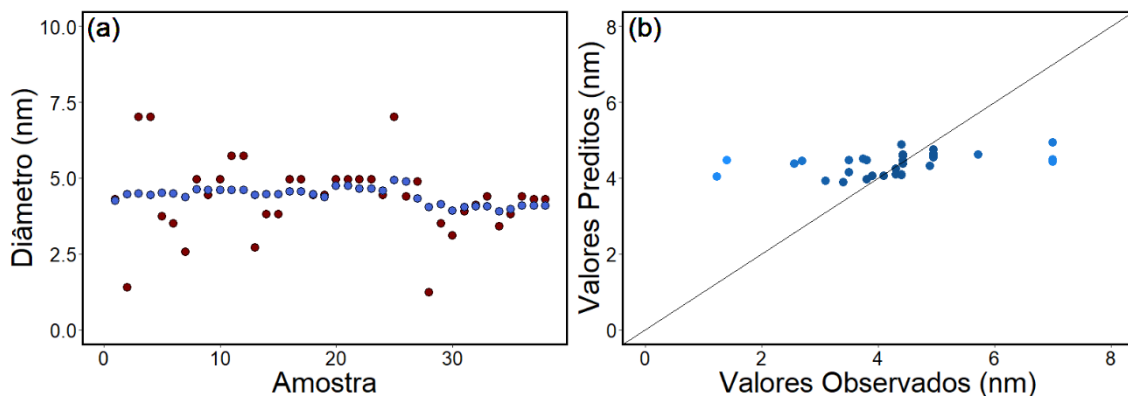


Figura 41: Gráficos das observações vs previsões pelo modelo *random forest* para pontos quânticos de CdS.

A Figura 42 mostra a comparação entre os valores de observação e previsões para os pontos quânticos de CdSe. O R-quadrado dos dados de treinamento e teste do modelo *random forest*, aplicado aos pontos quânticos de CdSe, foi de 0,759 e 0,604, respectivamente. Diferentemente do caso do CdS, o modelo apresentou uma performance mais robusta. Isso se reflete nos valores de RMSE (0,887 nm para dados de treinamento e 0,807 nm para dados de teste) e de MAE (0,528 e 0,586 nm, para dados de treino e teste, respectivamente). Considerando a relação entre tamanho de partícula e propriedades ópticas dos pontos quânticos, erros dessa magnitude resultam em uma grande diferença no comprimento de onda de absorção e emissão do material.

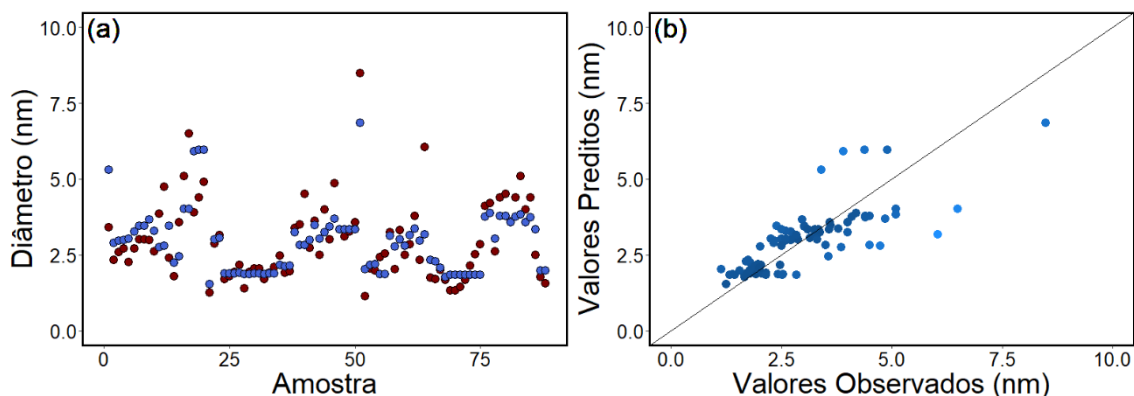


Figura 42: Gráficos das observações vs previsões pelo modelo *random forest* para pontos quânticos de CdSe.

O modelo *random forest* também apresentou um desempenho inferior ao do CdSe quando aplicada ao conjunto de dados de PbS. Da mesma forma que a árvore de regressão, a explicação para essa perda de desempenho é o volume de dados menor do conjunto de dados de PbS em relação ao de CdSe. O valor de R-quadrado para os dados de treinamento foi de 0,392, com RMSE de 7,629 nm e MAE de 2,414 nm. As métricas de desempenho do modelo aplicado ao conjunto de dados de teste (R-quadrado: 0,511, RMSE: 1,947 nm e MAE: 1,358 nm) mostram, novamente, um desempenho melhor no teste do que no treinamento. Dadas as condições vistas nos modelos anteriores, esse resultado é esperado, uma vez que o conjunto de dados foi dividido da mesma forma para o modelo de *random forest*. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de PbS é mostrada na Figura 43.

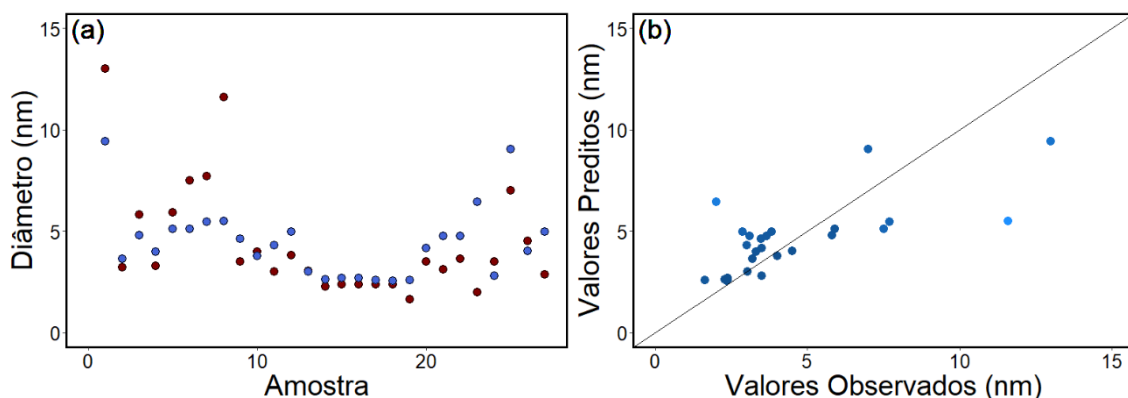


Figura 43: Gráficos das observações vs previsões pelo modelo *random forest* para pontos quânticos de PbS.

A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de PbSe é mostrada na Figura 44. O modelo *random forest* apresentou um desempenho semelhante ao do PbS quando aplicada ao conjunto de dados de PbSe. O valor de R-quadrado para os dados de treinamento foi de 0,390, com RMSE de 15,12 nm e MAE de 6,693 nm. Da mesma forma, as métricas de desempenho do modelo aplicado ao conjunto de dados de teste (R-quadrado: 0,403, RMSE: 8,125 nm e MAE: 5,882 nm) são superiores às dos dados de treinamento pelos motivos mencionados anteriormente.

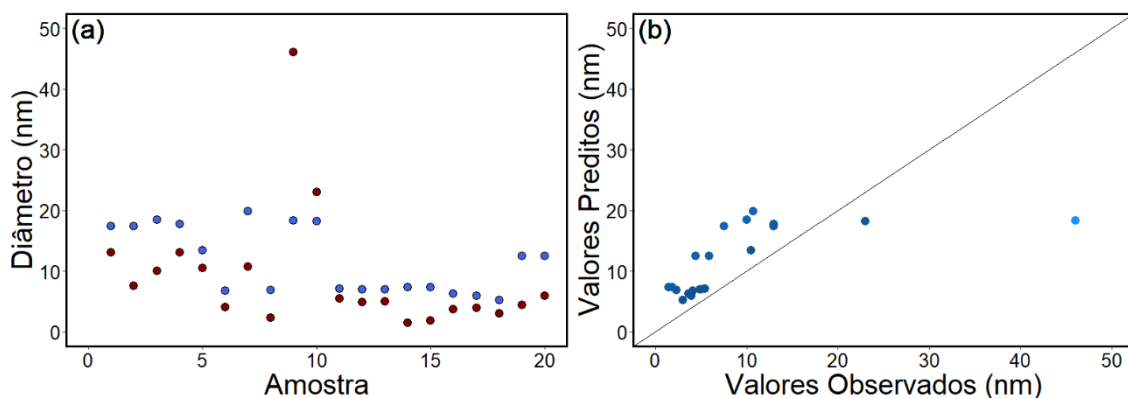


Figura 44: Gráficos das observações vs previsões pelo modelo *random forest* para pontos quânticos de PbSe.

Finalmente, o modelo *random forest* foi aplicado ao conjunto de dados de pontos quânticos de ZnSe. O modelo apresentou valores de R-quadrado de 0,464 e 0,153 para dados de treinamento e teste, respectivamente. Os valores de RMSE (0,985 e 1,151 nm) e MAE (0,654 e 0,836 nm) para dados de treinamento e teste, mostram um desempenho comparável com os modelos usados nos dados de CdSe. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de ZnSe é mostrada na Figura 45.

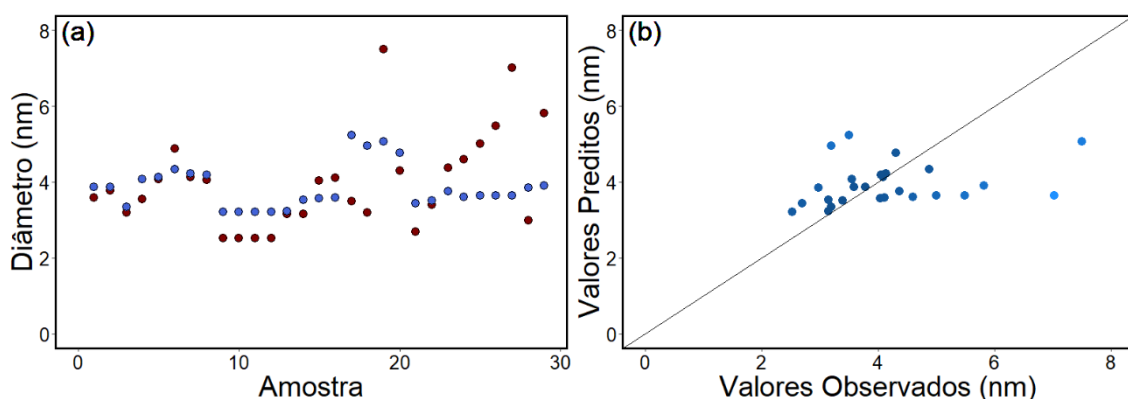


Figura 45: Gráficos das observações vs previsões pelo modelo *random forest* para pontos quânticos de ZnSe.

As Figuras 46 a 50 mostram os gráficos de influência dos preditores no modelo de *random forest*. Para pontos quânticos de CdS, os preditores mais importantes foram o tempo de reação, massa total, concentração de calcogênio e quantidade de ácido carboxílico. Para CdSe, a temperatura de crescimento e o tempo de reação são as variáveis mais influentes. Já para PbS, as concentrações de calcogênio, de precursor metálico, e de pontos quânticos formados, além do tempo de reação são os fatores que mais influenciam no tamanho dos pontos quânticos. Os pontos quânticos de PbSe possuem como

maiores preditores de seu tamanho final, o precursor metálico e a temperatura de crescimento. Finalmente, os pontos quânticos de ZnSe possuem como preditor mais forte o tempo de reação.

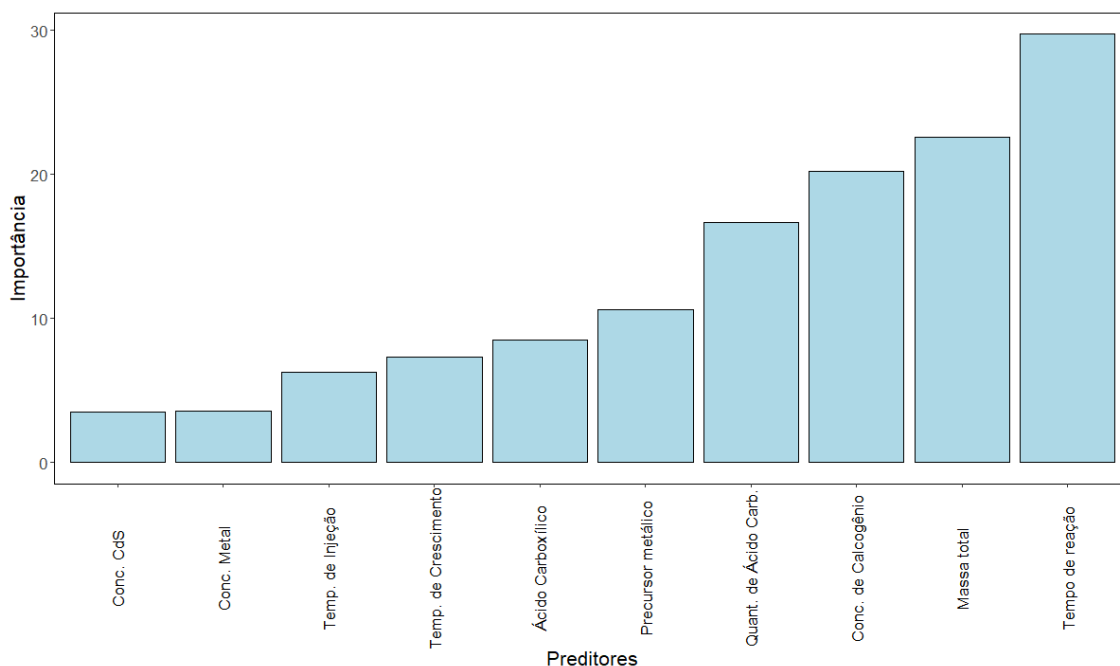


Figura 46: Importância das variáveis no modelo *random forest* para pontos quânticos de CdS.

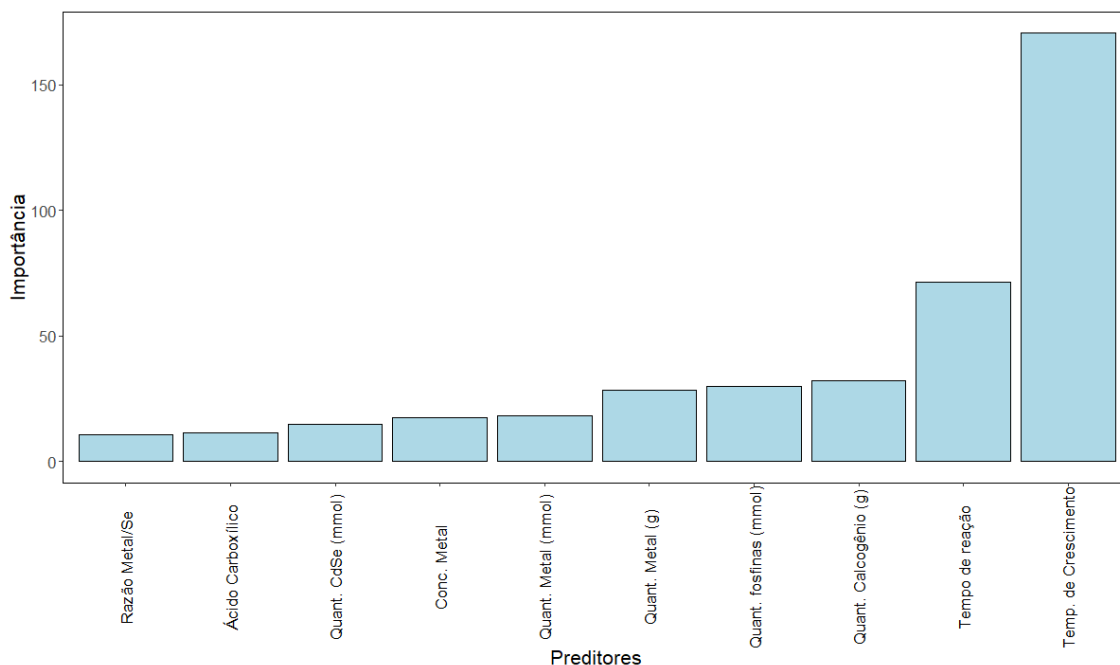


Figura 47: Importância das variáveis no modelo *random forest* para pontos quânticos de CdSe.

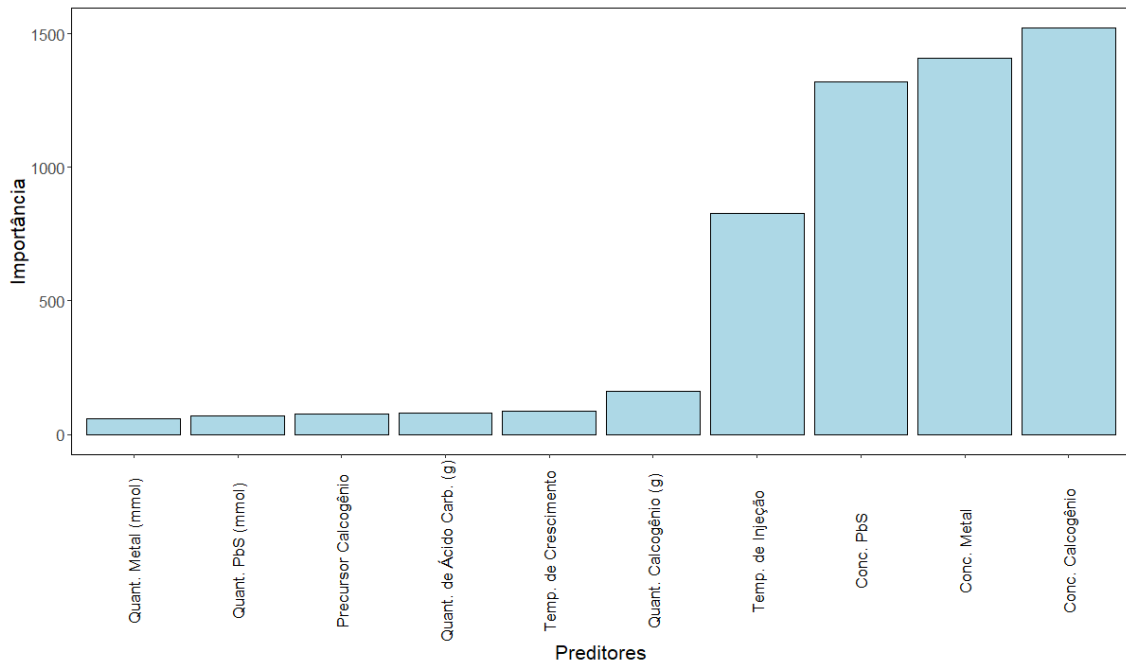


Figura 48: Importância das variáveis no modelo *random forest* para pontos quânticos de PbS.

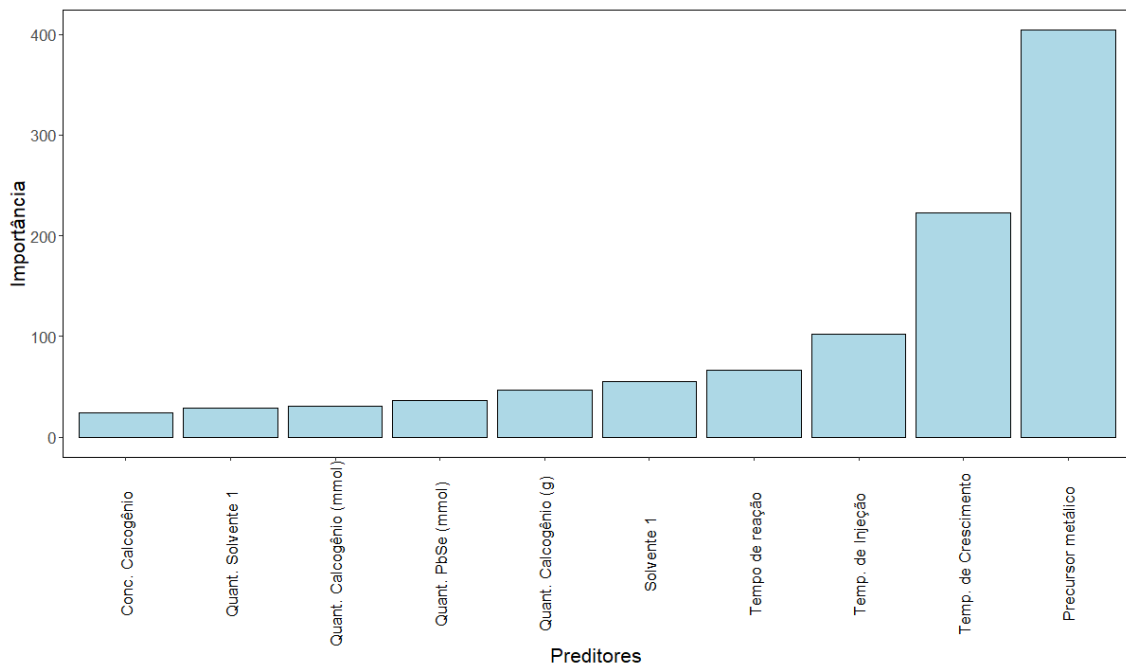


Figura 49: Importância das variáveis no modelo *random forest* para pontos quânticos de PbSe.

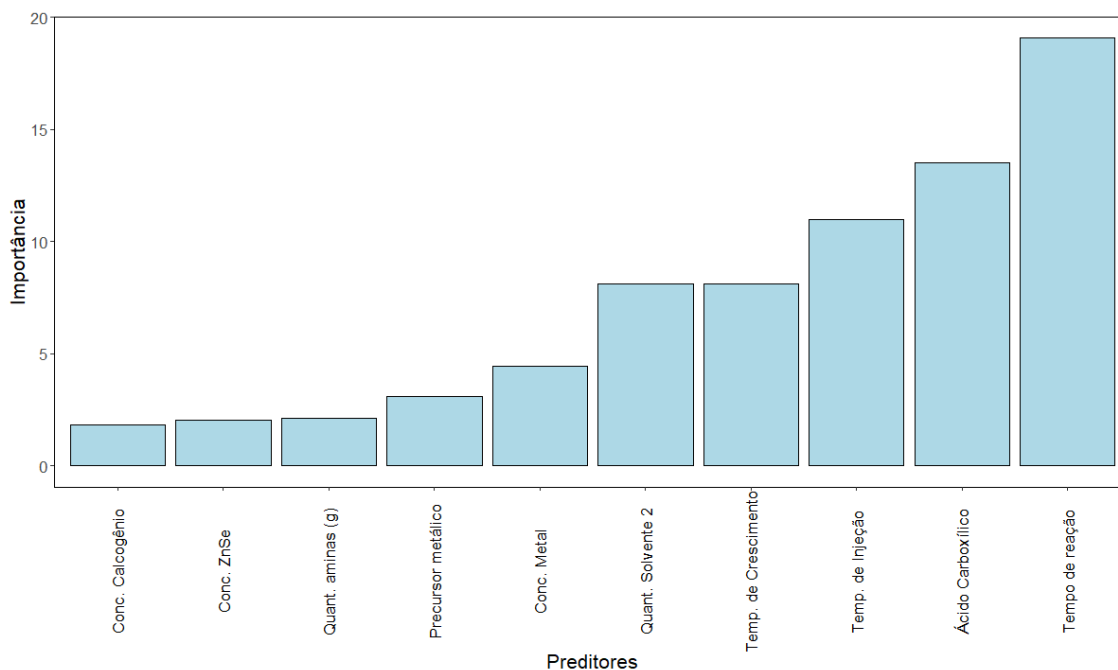


Figura 50: Importância das variáveis no modelo *random forest* para pontos quânticos de ZnSe.

3.2.5. Abordagens usando *machine learning* automatizado

Foram realizadas regressões usando a aplicação h2o^[100] para *machine learning* automatizado, a fim de aumentar a complexidade dos modelos. Através desta aplicação, é possível olhar para vários diferentes algoritmos de *machine learning* e selecionar a melhor performance de regressão. A aplicação h2o selecionou como melhor algoritmo um modelo *Stacked Ensemble*, que apresentou a menor média residual de desvio e RMSE. Entretanto, modelos do tipo *Stacked Ensemble* tem problemas de interpretabilidade, não sendo possível identificar a importância de cada variável que constrói o modelo. Portanto, foi escolhido um segundo modelo, um *gradient boosting machine* (GBM), que teve um desempenho similar ao *Stacked Ensemble*, porém com uma melhor interpretabilidade.

A Figura 51 mostra a comparação entre os valores de observação e predições para os pontos quânticos de CdS. O R-quadrado dos dados de treinamento e teste do modelo multilinear sem interação entre os preditores aplicado aos pontos quânticos de CdS foi de 0,992 e 0,766, respectivamente. Os valores de RMSE (0,140 nm para dados de treinamento e 0,611 nm para dados de teste) e de MAE (0,090 e 0,389 nm, para dados de treino e teste,

respectivamente) mostram que o modelo GBM apresentou uma performance superior aos modelos anteriores para os dados de CdS.

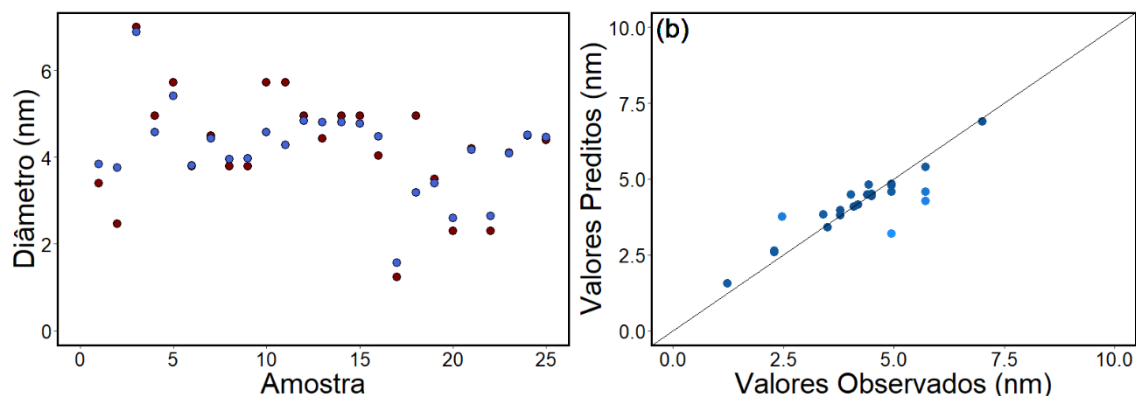


Figura 51: Gráficos das observações vs previsões pelo modelo de *gradient boosting machine* para pontos quânticos de CdS.

A Figura 52 mostra a comparação entre os valores de observação e previsões para os pontos quânticos de CdSe. O R-quadrado dos dados de teste e treinamento do modelo GBM aplicado aos pontos quânticos de CdSe foi de 0,998 e 0,933, respectivamente. A performance robusta do modelo GBM se reflete nos valores de RMSE (0,080 nm para dados de treinamento e 0,447 nm para dados de teste) e de MAE (0,053 e 0,240 nm, para dados de treino e teste, respectivamente).

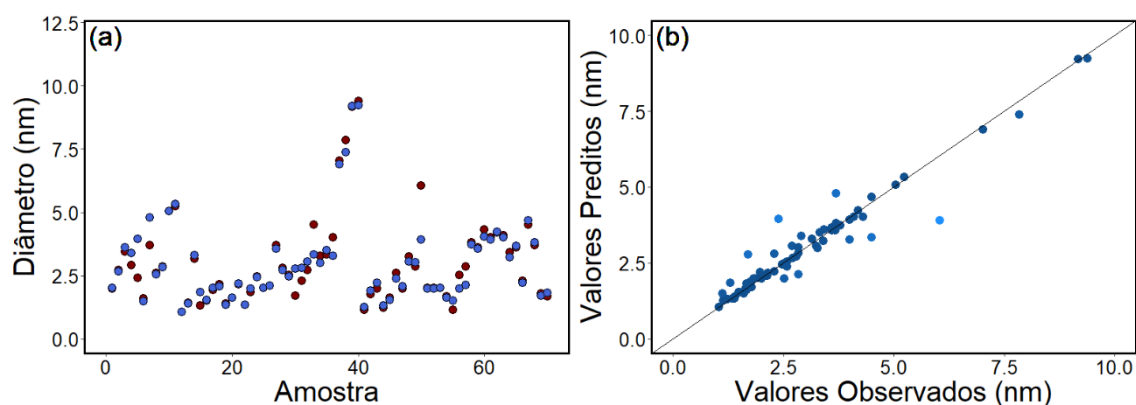


Figura 52: Gráficos das observações vs previsões pelo modelo de *gradient boosting machine* para pontos quânticos de CdSe.

A regressão por GBM apresentou um desempenho semelhante ao do CdS quando aplicada ao conjunto de dados de PbS. O valor de R-quadrado para os dados de treinamento foi de 0,749, com RMSE de 4,581 nm e MAE de 1,351 nm. Os valores de RMSE para dados de teste (1,297 nm) e de MAE (0,857 nm) mostram que o modelo GBM apresentou uma performance superior aos modelos

anteriores para os dados de PbS. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de PbS é mostrada na Figura 53.

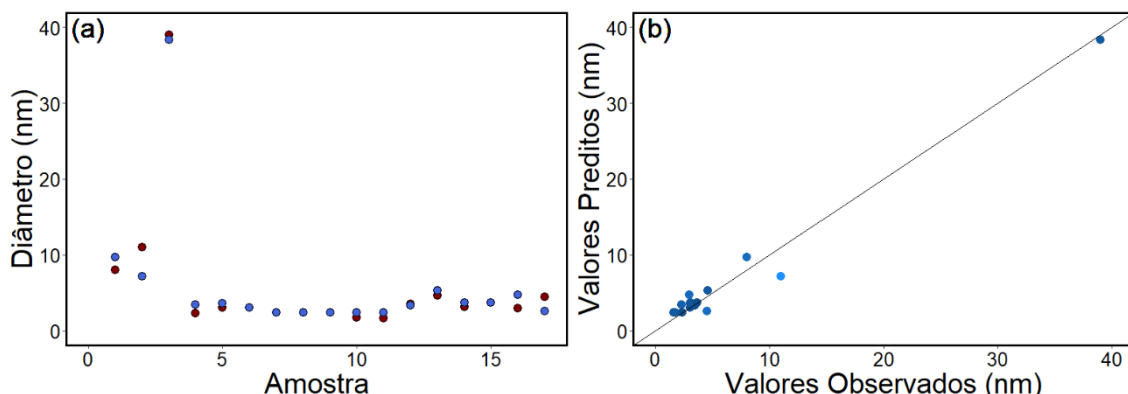


Figura 53: Gráficos das observações vs previsões pelo modelo de *gradient boosting machine* para pontos quânticos de PbS.

O modelo GBM apresentou um desempenho bastante inferior ao dos demais compostos quando aplicado ao conjunto de dados de PbSe. O valor de R-quadrado para os dados de treinamento foi de 0,582, com RMSE de 16,8 nm e MAE de 9,2 nm. Por outro lado, as métricas de desempenho do modelo aplicado ao conjunto de dados de teste (RMSE: 5,6 nm e MAE: 5,2 nm) são superiores às dos dados de treinamento. A explicação para essa perda de desempenho é o volume de dados menor do conjunto de dados de PbSe em relação ao de CdSe. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de PbSe é mostrada na Figura 54.

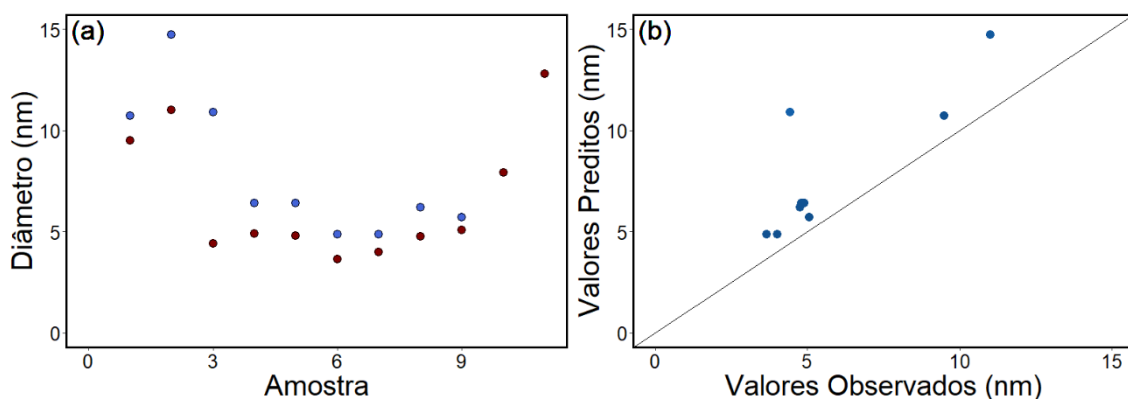


Figura 54: Gráficos das observações vs previsões pelo modelo de *gradient boosting machine* para pontos quânticos de PbSe.

Por fim, o modelo GBM foi aplicado ao conjunto de dados de pontos quânticos de ZnSe. O modelo apresentou valores de R-quadrado de 0,99 e 0,81 para dados de treinamento e teste, respectivamente. Os valores de RMSE (0,35 e 0,09 nm) e MAE (0,28 e 0,07 nm) para dados de treinamento e teste, mostram

um desempenho superior ao dos modelos anteriores para ZnSe. A comparação entre os valores observados e preditos pelo modelo para pontos quânticos de ZnSe é mostrada na Figura 55.

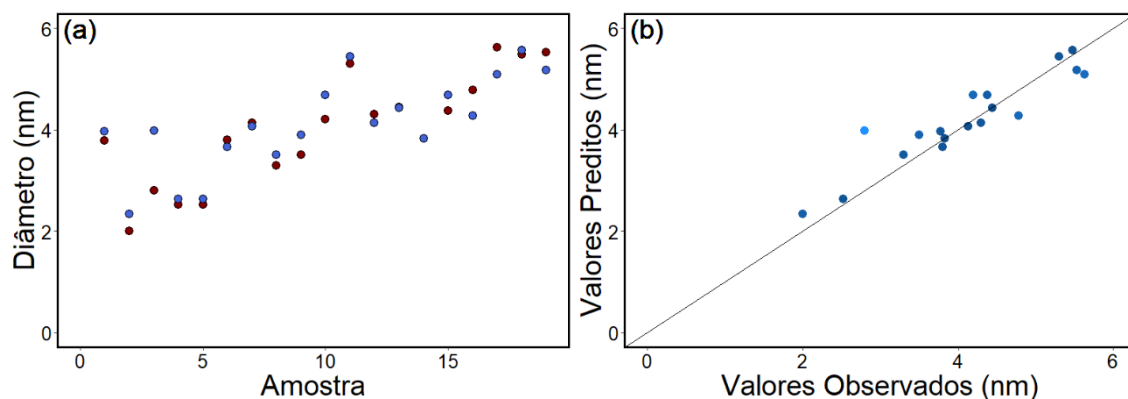


Figura 55: Gráficos das observações vs predições pelo modelo de *gradient boosting machine* para pontos quânticos de ZnSe.

3.2.6. Deep learning

O gráfico de observações versus predições do modelo sequencial de *deep learning* é mostrado nas Figuras 56 a 60. Considerando os valores de RMSE na Tabela 2, a performance do *deep learning* é similar ao *random forest* para os pontos quânticos de CdSe. Para o CdS, o *deep learning* é o melhor modelo, enquanto para PbSe e ZnSe, ele apresenta uma performance inferior.

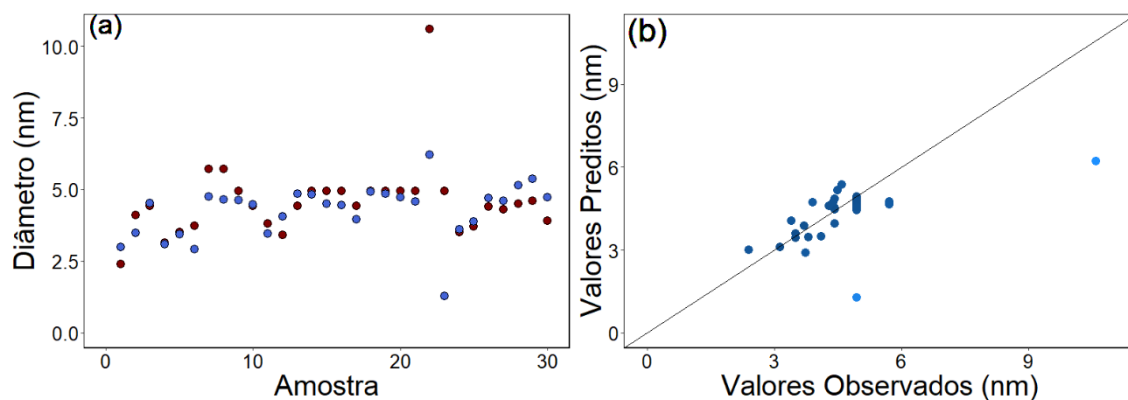


Figura 56: Gráficos das observações vs predições pelo modelo de *deep learning* para pontos quânticos de CdS.

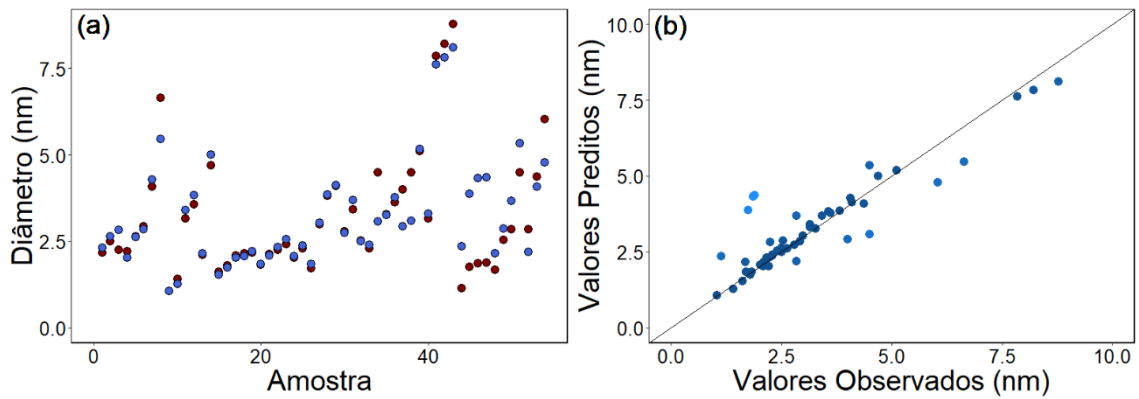


Figura 57: Gráficos das observações vs previsões pelo modelo de *deep learning* para pontos quânticos de CdSe.

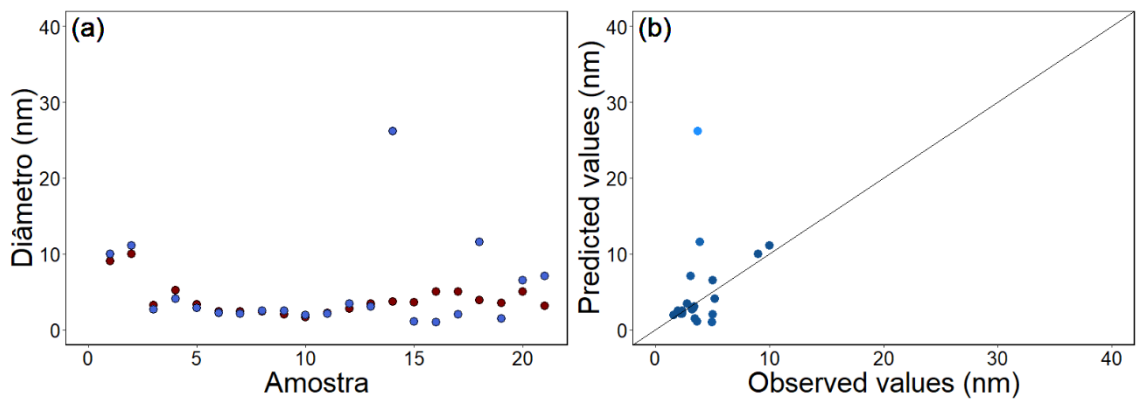


Figura 58: Gráficos das observações vs previsões pelo modelo de *deep learning* para pontos quânticos de PbS.

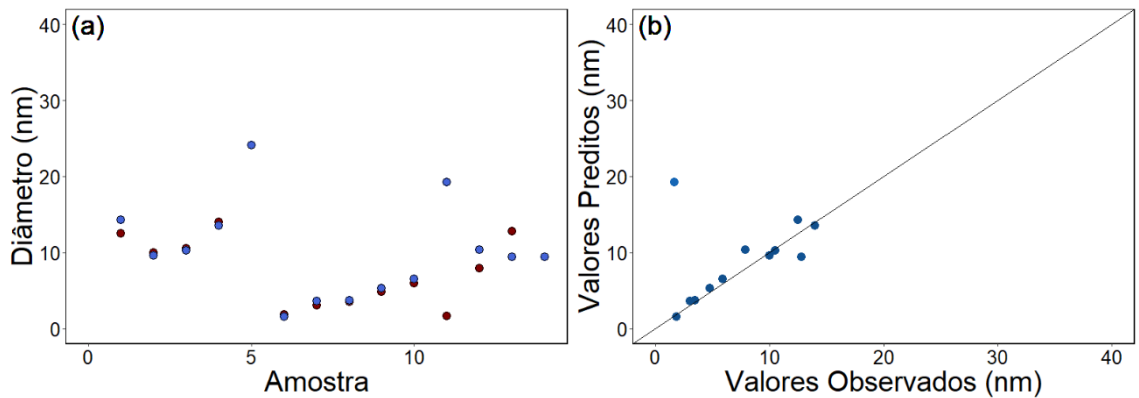


Figura 59: Gráficos das observações vs previsões pelo modelo de *deep learning* para pontos quânticos de PbSe.

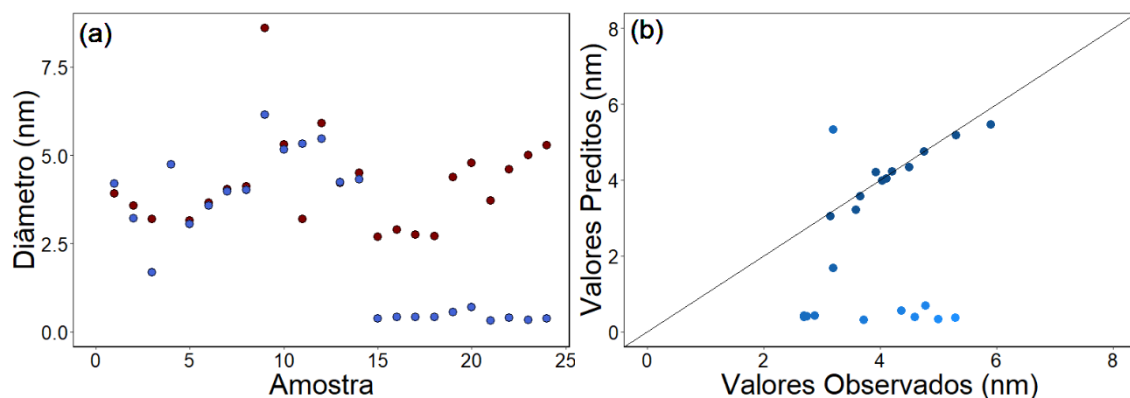


Figura 60: Gráficos das observações vs previsões pelo modelo de *deep learning* para pontos quânticos de ZnSe.

A Tabela 2 apresenta as métricas de performance dos modelos para cada composto. Valores de R^2 próximos de 1 significam que o modelo se adapta melhor aos dados. Valores menores de RMSE e MAE representam melhor performance do modelo. Modelos com alta performance nos dados de treinamento e baixa performance nos dados de teste são típicos casos de *overfit*. Esse é o caso da regressão multilinear com interação entre os preditores, que apresenta altos valores de R^2 para CdSe, PbSe e ZnSe, mas esses valores são completamente degradados nos dados de teste. No extremo oposto, a regressão multilinear sem interações entre os preditores apresenta uma baixa performance, tanto nos dados de treinamento, quanto nos dados de teste, o que é típico de *underfit*. Os modelos baseados em árvores de decisão, árvore de regressão, *Random Forest* e GBM apresentam métricas mais balanceadas. Entre esses modelos, GBM e *Random Forest* apresentam as melhores performances. A regressão por *deep learning* apresentou uma performance promissora com a maior base de dados (CdSe), mas perdeu performance rapidamente com os conjuntos de dados menores. Portanto, os modelos escolhidos como base para a simulação de novas condições de síntese de pontos quânticos foram *random forest*, GBM e *deep learning*.

Tabela 2: Métricas de performance dos modelos.

Composto	Modelo	Dados de treino			Dados de teste		
		R^2	RMSE	MAE	R^2	RMSE	MAE
CdS	Multilinear S. I.	0,166	0,838	0,569	0,191	1,078	0,704
CdS	Multilinear C. I.	0,575	0,599	0,372	0,001	123,9	21,31
CdS	Smoothing Splines	0,391	0,727	0,446	0,148	1,087	0,774
CdS	Árvores de regressão	0,570	0,947	0,563	0,360	1,059	0,665
CdS	Random Forest	0,413	1,267	0,748	0,222	1,133	0,771
CdS	GBM	0,992	0,140	0,090	0,766	0,611	0,389
CdS	Deep Learning	0,780	0,666	0,459	0,328	1,151	0,649
CdSe	Multilinear S. I.	0,202	1,285	0,894	0,297	1,599	1,047

CdSe	Multilinear C. I.	0,957	0,299	0,187	0,007	160,6	22,75
CdSe	Smoothing Splines	0,957	0,300	0,201	0,771	0,890	0,565
CdSe	Árvores de regressão	0,735	0,893	0,570	0,607	0,816	0,570
CdSe	Random Forest	0,759	0,887	0,528	0,604	0,807	0,586
CdSe	GBM	0,998	0,080	0,053	0,933	0,447	0,240
CdSe	Deep Learning	0,623	1,108	0,481	0,811	0,747	0,436
PbS	Multilinear S. I.	0,272	7,966	3,564	0,250	3,521	2,530
PbS	Multilinear C. I.	0,522	6,451	2,401	0,004	10,61	5,436
PbS	Smoothing Splines	0,570	6,151	2,178	0,094	3,117	1,955
PbS	Árvores de regressão	0,401	7,221	2,435	0,582	1,893	1,134
PbS	Random Forest	0,392	7,629	2,414	0,511	1,947	1,358
PbS	GBM	0,749	4,581	1,351	0,978	1,297	0,857
PbS	Deep Learning	0,741	4,637	1,778	0,126	5,428	2,565
PbSe	Multilinear S. I.	0,411	10,34	6,308	0,332	23,21	10,01
PbSe	Multilinear C. I.	0,999	0,237	0,106	0,012	3730	1165
PbSe	Smoothing Splines	0,553	9,208	4,655	0,109	26,00	12,27
PbSe	Árvores de regressão	0,396	13,94	5,504	0,290	9,089	4,396
PbSe	Random Forest	0,390	15,12	6,693	0,403	8,125	5,882
PbSe	GBM	0,582	12,32	4,934	0,589	9,493	5,605
PbSe	Deep Learning	0,154	16,28	3,686	0,214	14,59	6,921
ZnSe	Multilinear S. I.	0,396	1,026	0,684	0,257	1,086	0,787
ZnSe	Multilinear C. I.	0,932	0,343	0,198	0,005	58625	14368
ZnSe	Smoothing Splines	0,664	0,810	0,539	0,534	0,860	0,589
ZnSe	Árvores de regressão	0,677	0,750	0,532	0,362	1,008	0,694
ZnSe	Random Forest	0,464	0,985	0,654	0,153	1,151	0,836
ZnSe	GBM	0,989	0,145	0,104	0,879	0,385	0,279
ZnSe	Deep Learning	0,152	1,668	0,893	0,204	2,433	1,765

3.2.7. Simulando e sugerindo novas rotas de síntese

O método Monte Carlo foi usado para simular 5000 condições de reação com restrições baseadas na disponibilidade dos reagentes em nosso laboratório: óxido de cádmio, selênio em pó, ácido oleico, oleilamina, trioctilfosfina, octadeceno, parafina líquida, ácido fosfônico e óxido de tributilfosfina. Baseado na performance de cada algoritmo (Tabela 2), foram aplicados os modelos de *random forest*, GBM e *deep learning* para encontrar quais as condições levariam a três diferentes faixas de tamanhos de pontos quânticos de CdSe: pequenos (1,8 a 2,2 nm), médios (4,8 a 5,2 nm) e grandes (7,8 a 8,2 nm). Para comparar o comportamento dos modelos, foram selecionadas condições reacionais comuns aos três modelos que levariam ao tamanho-alvo desejado (Tabela 3). Foram aplicadas variações em duas condições experimentais, tempo de reação (1 a 200 min) e temperatura de crescimento (150 a 350 °C). Para encontrar as condições mais robustas a pequenas variações no setup experimental, foram

selecionadas aquelas que apresentaram as maiores áreas adjacentes dentro do tamanho-alvo desejado.

Tabela 3: Condições de reação sugeridas pelo algoritmo para obtenção de pontos quânticos de CdSe, com tamanhos na faixa de 1.8 a 2.2, 4.8 a 5.2, e 7.8 a 8.2 nm.

Alvo	T _{inj}	T _{grw}	M	M _g	C	C _g	CA	CA _g	P	P _g	S ¹	S _g ¹	S ²	S _g ²	t
1.8 – 2.2	200	200	CdO	0.11	Se	1.92	OA	1.885	-	-	ODE	3.71	PA	3.31	92.27
4.8 – 5.2	255	325	CdO	0.93	Se	1.09	-	-	TOP	2.045	LP	2.45	TBPO	4.57	94.64
7.8 – 8.2	220	230	CdO	1.56	Se	0.74	OA	5.43	TOP	4.62	-	-	-	-	79.57

OA – ácido oleico; TOP – trioctilfosfina; TBPO – óxido de tributilfosfina; PA – ácido fosfônico; ODE – octadeceno; LP – parafina líquida.

Mapas de calor foram gerados para cada tamanho alvo e podem ser vistos na Figura 61. Os modelos baseados em árvores apresentam uma divisão pronunciada perto da temperatura de crescimento de 300 °C. Este comportamento é devido à sua estrutura de árvore, como podemos ver nas Figuras 31 a 35, com o primeiro nó da árvore sendo temperaturas de crescimento maiores ou menores que 295 °C. Para pontos quânticos de tamanhos maiores (Figuras 61 (g, h e i)), os modelos baseados em árvores se tornam mais instáveis. Se olharmos para os gráficos de predições versus observações para *random forest* (Figura 42) e GBM (Figura 52), de fato podemos ver uma performance menor para maiores tamanhos de pontos quânticos. O modelo de *deep learning* (Figura 57), por outro lado, possui uma performance mais balanceada através de todos os tamanhos de pontos quânticos.

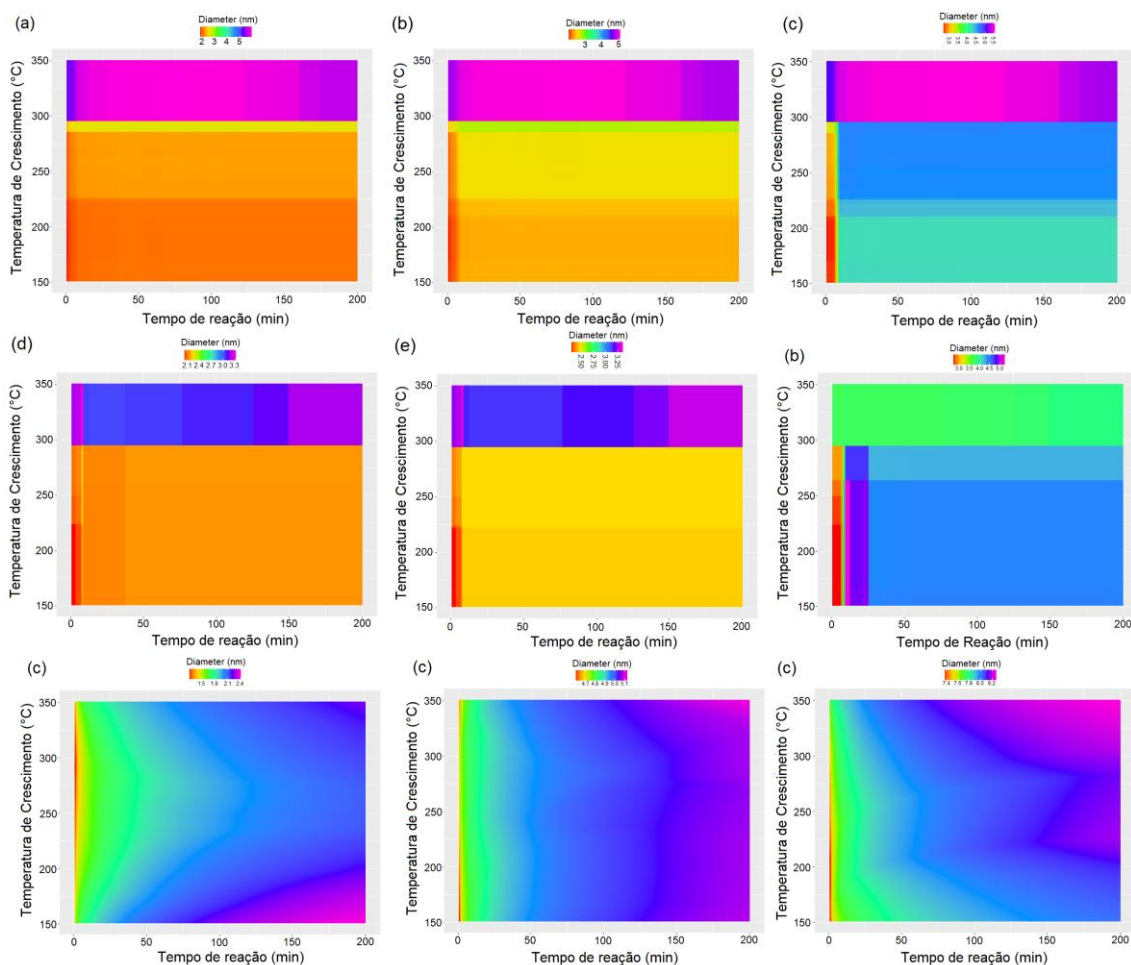


Figura 61: Mapa das condições de reação para pontos quânticos de CdSe com tamanho alvo entre 1.8 a 2.2 nm pelos modelos *Random Forest* (a), *GBM* (b), and *Deep Learning* (c). Mapa das condições de reação para pontos quânticos de CdSe com tamanho alvo entre 4.8 a 5.2 nm pelos modelos *Random Forest* (d), *GBM* (e), and *Deep Learning* (f). Mapa das condições de reação para pontos quânticos de CdSe com tamanho alvo entre 7.8 a 8.2 nm pelos modelos *Random Forest* (g), *GBM* (h), and *Deep Learning* (i).

Usar os mapas de calor pode ajudar o projetista de síntese a escolher os parâmetros de reação esperados para serem os mais estáveis e que resultem em pontos quânticos com tamanhos escolhidos. Ademais, é possível aplicar este modelo a dados de pontos quânticos do tipo III-V, ternários e multinários, desde que se consiga obter uma base de dados de treinamento. Esta abordagem também pode ser usada com ferramentas automatizadas, como a demonstrada por Coley e colaboradores^[18], para otimizar e escalar a síntese de diferentes tipos de pontos quânticos. Além disso, é possível escolher uma dada fonte de metal, ligante ou solvente, como uma necessidade do processo de fabricação, de produto ou disponibilidade de reagentes no laboratório.

3.3. CONCLUSÕES PARCIAIS

Algoritmos de *machine learning* foram aplicados à base de dados contendo os parâmetros de síntese de pontos quânticos de CdS, CdSe, PbS, PbSe e ZnSe para encontrar as variáveis que influenciam no tamanho médio final destas partículas.

O modelo de regressão multilinear foi escolhido como modelo inicial por sua simplicidade. Quando aplicado sem interação entre as variáveis, o modelo não foi capaz de realizar boas previsões dos dados de teste, obtendo valores de MAE superiores a 0,7 nm para todas as bases de dados. Usando interação entre as variáveis, entretanto, o modelo se torna tão flexível que se adapta de maneira quase perfeita aos dados de treinamento, causando um fenômeno conhecido como overfit. Assim, os modelos de regressão multilíneares não atenderam às necessidades de controle fino do tamanho dos pontos quânticos.

Após os modelos multilíneares, foi aplicado um algoritmo de *smoothing splines* às bases de dados dos pontos quânticos. Esse modelo apresentou um desempenho superior aos modelos multilíneares, principalmente por melhorar a capacidade de previsão de tamanhos de pontos quânticos nos quartis superiores das bases de dados.

Subsequentemente, foram aplicados modelos de árvore de regressão e seus derivados, *random forest* e *gradient boosting machine*. Nesse conjunto, os modelos que apresentaram melhor desempenho preditivo foram aqueles baseados em *gradient boosting machines*. As variáveis mais influentes no tamanho dos pontos quânticos foram o tempo de reação (CdS e ZnSe), a temperatura de crescimento (CdSe), as concentrações dos precursores (PbS) e o tipo de precursor (PbSe).

Por fim, foram aplicados modelos de deep learning que, para as bases de dados maiores (CdS e CdSe) apresentou um bom desempenho, com erro médio de previsão ao redor de 0,4 nanômetros.

Após verificar o desempenho de todos os algoritmos, foram selecionados os modelos de random forest, gradient boosting machine e deep learning para

realizar uma simulação através do método de Monte Carlo. Essas simulações tiveram como objetivo demonstrar a capacidade dos modelos de sugerir novas condições de síntese de acordo com as necessidades do pesquisador. As sugestões tomaram a forma das tabelas de condições de síntese e dos mapas de calor contendo tempo e temperatura de reação.

A importância deste trabalho se dá, entre outras, por algumas razões: (i) a criação de uma base de dados, disponível para a comunidade científica criar seus próprios modelos,^[101] (ii) abrir uma nova perspectiva no uso de algoritmo de machine learning para auxiliar no design de sínteses de nanomateriais, sendo, até o momento da publicação deste trabalho, a primeira tese de doutorado a tratar desse tema no Brasil e (iii) auxiliar na facilitação do acesso a pontos quânticos com propriedades customizadas por parte de pesquisadores de fora da área de química e ciência dos materiais, que podem aplicar essas propriedades em atividades interdisciplinares.

Como perspectivas futuras ficam as possibilidades de aplicação deste método para outros compostos semicondutores, como compostos do tipo III-V, compostos ternários e multinários. Para tanto, será necessária a construção de uma base de dados, nos moldes já sugeridos neste trabalho, envolvendo a síntese destes compostos com base nos trabalhos publicados na literatura. Adicionalmente, também é possível incluir características físicas e químicas dos precursores e dos pontos quânticos estudados neste trabalho a fim de sugerir rotas de síntese para compostos ainda não vistos nas bases de dados.

CAPÍTULO 4: PLANEJAMENTO EXPERIMENTAL

Para que se possa criar uma base de dados de sínteses, realizar experimentos é uma etapa indispensável. Nesta etapa trabalho, buscou-se por alternativas não somente ao uso de semicondutores baseados em cádmio e chumbo, mas também ao uso de elementos em risco de escassez, como o índio. Nesse contexto, o material escolhido foi o sulfeto de cobre e alumínio. Para explorar a forma com que alterações nas condições experimentais influenciam nas propriedades físicas do material, a técnica de planejamento experimental foi empregada neste capítulo. O objetivo foi sintetizar nanopartículas de sulfeto de cobre e alumínio através do método *heat up* e avaliar os efeitos dos parâmetros de síntese como temperatura, tempo de reação e concentração dos precursores, nas propriedades ópticas do material.

4.1. PARTE EXPERIMENTAL

Para a triagem das variáveis de síntese que mais afetam as propriedades ópticas das nanopartículas de sulfeto de cobre e alumínio, foram selecionados os 6 parâmetros experimentais mais convenientes para variação, de acordo com a Tabela 4.

Tabela 4: Parâmetros experimentais avaliados na síntese de nanopartículas de sulfeto de cobre e alumínio.

Parâmetro	Símbolo	Nível (-)	Nível (+)
Temperatura	T	200 °C	250 °C
Tempo	t	2 min	15 min
Conc. de CuCl	C	0,2	1,0
Conc. de AlCl ₃	A	0,2	1,0
Conc. de S	S	0,4	2,0
Razão OLA/ODE	O	3/7	7/3

Considerando os parâmetros mostrados na Tabela 4, caso fosse realizado um planejamento fatorial 2^k completo, o número total de experimentos seria 2^6 , ou seja, 64 experimentos. Entretanto, a fim de manter a economicidade e a razoabilidade, optou-se por realizar o planejamento fatorial 2^{6-2} com resolução 3, resultando em 16 experimentos, onde o fator 5 (S) é igual a 1234 (T x t x C x A)

e o fator 6 (O) é igual a 123 (T x t x C). A lista completa das condições experimentais é apresentada na Tabela 5.

Tabela 5: Condições experimentais das sínteses

Experimento	T	t	C	A	S	O
1	-	-	-	-	+	-
2	+	-	-	-	-	+
3	-	+	-	-	-	+
4	+	+	-	-	+	-
5	-	-	+	-	-	+
6	+	-	+	-	+	-
7	-	+	+	-	+	-
8	+	+	+	-	-	+
9	-	-	-	+	-	-
10	+	-	-	+	+	+
11	-	+	-	+	+	+
12	+	+	-	+	-	-
13	-	-	+	+	+	+
14	+	-	+	+	-	-
15	-	+	+	+	-	-
16	+	+	+	+	+	+

Após a definição das condições experimentais, a realização dos experimentos foi aleatorizada, de acordo com a Tabela 6.

Tabela 6: Nomes das amostras após aleatorização.

Amostra	Experimento	Amostra	Experimento
HU-001	15	HU-009	4
HU-002	16	HU-010	7
HU-003	13	HU-011	12
HU-004	5	HU-012	11
HU-005	3	HU-013	14
HU-006	10	HU-014	8
HU-007	1	HU-015	6
HU-008	2	HU-016	9

Após realizadas as análises dos fatores que influenciaram nas propriedades ópticas do material, um segundo planejamento foi realizado, desta vez com 3

parâmetros: tempo de reação, concentração de cloreto de cobre e concentração de cloreto de alumínio.

Tabela 7: Parâmetros experimentais avaliados no segundo planejamento experimental da síntese de nanopartículas de sulfeto de cobre e alumínio.

Parâmetro	Símbolo	Nível (-)	Nível (+)
Tempo	t	2 min	5 min
Conc. De CuCl	C	0,3	0,75
Conc. De AlCl ₃	A	1	1,5

Nesta segunda etapa, foi realizado um planejamento fatorial 2^k completo, com número total de experimentos de 2³, ou seja, 8 experimentos. A lista completa das condições experimentais é apresentada na Tabela 8.

Tabela 8: Condições experimentais das sínteses do segundo planejamento experimental.

Experimento	t	C	A
1	-	-	-
2	+	-	-
3	-	+	-
4	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8	+	+	+

Após a definição das condições experimentais, a realização dos experimentos foi aleatorizada, de acordo com a Tabela 9.

Tabela 9: Nomes das amostras do segundo planejamento experimental após aleatorização.

Amostra	Experimento
HU-101	5
HU-102	8
HU-103	1
HU-104	6
HU-105	3
HU-106	7
HU-107	4
HU-108	2

A magnitude dos efeitos dos parâmetros é calculada de acordo com a equação 21:

$$E_i = \bar{y}_+ - \bar{y}_- \quad \text{Eq. 21.}$$

Onde E_i representa o efeito, \bar{y}_+ é a média dos resultados apresentados com o parâmetro avaliado no nível positivo (+) e \bar{y}_- é a média dos resultados apresentados com o parâmetro avaliado no nível negativo (-). A magnitude dos efeitos combinados é conhecida pela equação 22:

$$E_i = \bar{y}_{++} + \bar{y}_{--} - \bar{y}_{+-} - \bar{y}_{-+} \quad \text{Eq. 22.}$$

Onde \bar{y}_{++} , \bar{y}_{--} , \bar{y}_{+-} e \bar{y}_{-+} são as diferentes combinações de níveis de parâmetros experimentais.

4.1.1. Materiais

Cloreto de cobre (I) (Sigma-Aldrich, 97%), cloreto de alumínio (III) (Sigma-Aldrich, 98%), oleilamina (Sigma-Aldrich, 70%), octadeceno (Sigma-Aldrich, 90%), hexano (Fmaia, P.A.) e acetona (Fmaia, P. A.).

4.1.2. Síntese das nanopartículas via heat-up

Em um balão de três bocas, foram adicionados CuCl, AlCl₃, oleilamina e octadeceno. Em outro frasco, foi preparada uma solução de enxofre, oleilamina e octadeceno totalizando 5 mL de solução. Essa solução de enxofre foi mantida sob agitação e aquecimento a 60 °C até a dissolução do enxofre, observado pela formação de um líquido avermelhado translúcido. A solução de enxofre foi adicionada ao balão de três bocas e a solução resultante foi desgaseificada sob vácuo em três etapas, a 50, 70 e 90 °C, até que a formação de bolhas não fosse mais observada nas três temperaturas. Após, a temperatura foi elevada até a temperatura de síntese (200 ou 250 °C), e a reação foi mantida pelo tempo de reação determinado (2 ou 15 minutos). Após a reação, os nanocristais foram precipitados pela adição de acetona, e centrifugados por 10 minutos a 3000 rpm. O material sobrenadante foi descartado e as nanopartículas precipitadas foram redissolvidas em aproximadamente 5 mL de hexano. O processo de precipitação e centrifugação foi repetido duas vezes.

4.1.3. Preparação de amostras para microscopia confocal de fluorescência

Duas a três gotas de solução das nanopartículas em hexano foram depositadas sobre uma lâmina histológica de vidro e, em seguida, recoberta com uma lamínula de vidro, fixada com um agente de cola. As amostras foram então visualizadas em um Microscópio Confocal Espectral FV 1000 e excitada com lasers de 405, 473 e 559 nm.

4.1.4. Caracterização óptica

As propriedades ópticas das nanopartículas foram verificadas por medidas de absorção em um espectrofotômetro UV-Vis-NIR Cary 5000, com varredura entre 400 e 2000 nm e por medidas de emissão de fluorescência em um espectrofotômetro de fluorescência Shimadzu RF5301. Foram usadas suspensões das nanopartículas dispersas em hexano, em cubetas de quartzo tamanho padrão 10 mm.

4.1.5. Caracterização estrutural e morfológica

O tamanho das nanoestruturas e sua morfologia foram caracterizados por microscopia eletrônica de transmissão (MET). As imagens foram obtidas no CME/UFRGS, em um equipamento Jeol JEM 1200 EXII. Para obtenção das imagens as dispersões de nanopartículas em hexano foram depositadas sobre *grids* de cobre cobertos com carbono (300 mesh).

As medidas de difração de raios-x foram realizadas em um equipamento Rigaku Ultima IV, com radiação $K\alpha$ ($\lambda = 1.54 \text{ \AA}$), com geometria Bragg-Brentano. As varreduras foram realizadas entre 10° e 80° , com passo de $0,05^\circ$ e velocidade de varredura $3 (1^\circ/\text{min})$. Foram usados filmes das nanopartículas sobre placas histológicas foscas de vidro. As amostras foram preparadas depositando-se uma dispersão do material em hexano sobre as placas de vidro. Com a evaporação do solvente, um filme se formou sobre a placa e esta foi usada para a difratometria de raios-x.

Para identificação dos picos e das fases correspondentes nos difratogramas de raios-x, foi usado o *software* livre QualX, desenvolvida pelo Institute of Crystallography (IC)-CNR - Bari, Itália.^[102] A base de dados usada para comparação das cartas cristalográficas foi a POW_COD, desenvolvida pelos autores do QualX a partir da informação contida na Crystallography Open Database (COD).^[103]

A caracterização por espectroscopia Raman foi realizada em um espectrômetro Bruker, modelo SENTERRA, com microscópio, na Universidade Federal de Juiz de Fora. Os espectros foram obtidos através da excitação das amostras por um laser de 785 nm, com potência de 10 mW, aplicado durante 30 segundos, em uma lente de 50x. O espectro final é o resultado da média de dois espectros sobre o mesmo ponto. As amostras consistiam em filmes das nanopartículas, depositadas em uma placa histológica de vidro. As amostras foram preparadas depositando-se uma dispersão do material em hexano sobre as placas de vidro. Para mitigar o efeito da fluorescência, que impede a obtenção de informação relevante a partir dos espectros, foram depositadas nanopartículas de prata sobre os filmes a serem analisados.

4.2. RESULTADOS E DISCUSSÃO

A Figura 62 mostra as imagens obtidas por MET para as nanopartículas sintetizadas pelo método *heat up*, sob 16 condições de síntese diferentes. É possível observar que as amostras HU-001, HU-004, HU-007, HU-008 e HU-014 resultaram em nanopartículas com distribuição mais uniforme de tamanho e formato. Esses formatos foram encontrados em outros trabalhos de síntese em nanopartículas de CuAIS_2 .^[30-32] Nas amostras HU-001, HU-002, HU-003 e HU-005, surge a formação de nanoplacas de formato hexagonal, que é bastante comum em sulfetos de cobre.^[75,104] As demais condições resultaram em nanopartículas de múltiplos formatos e tamanhos na mesma síntese.

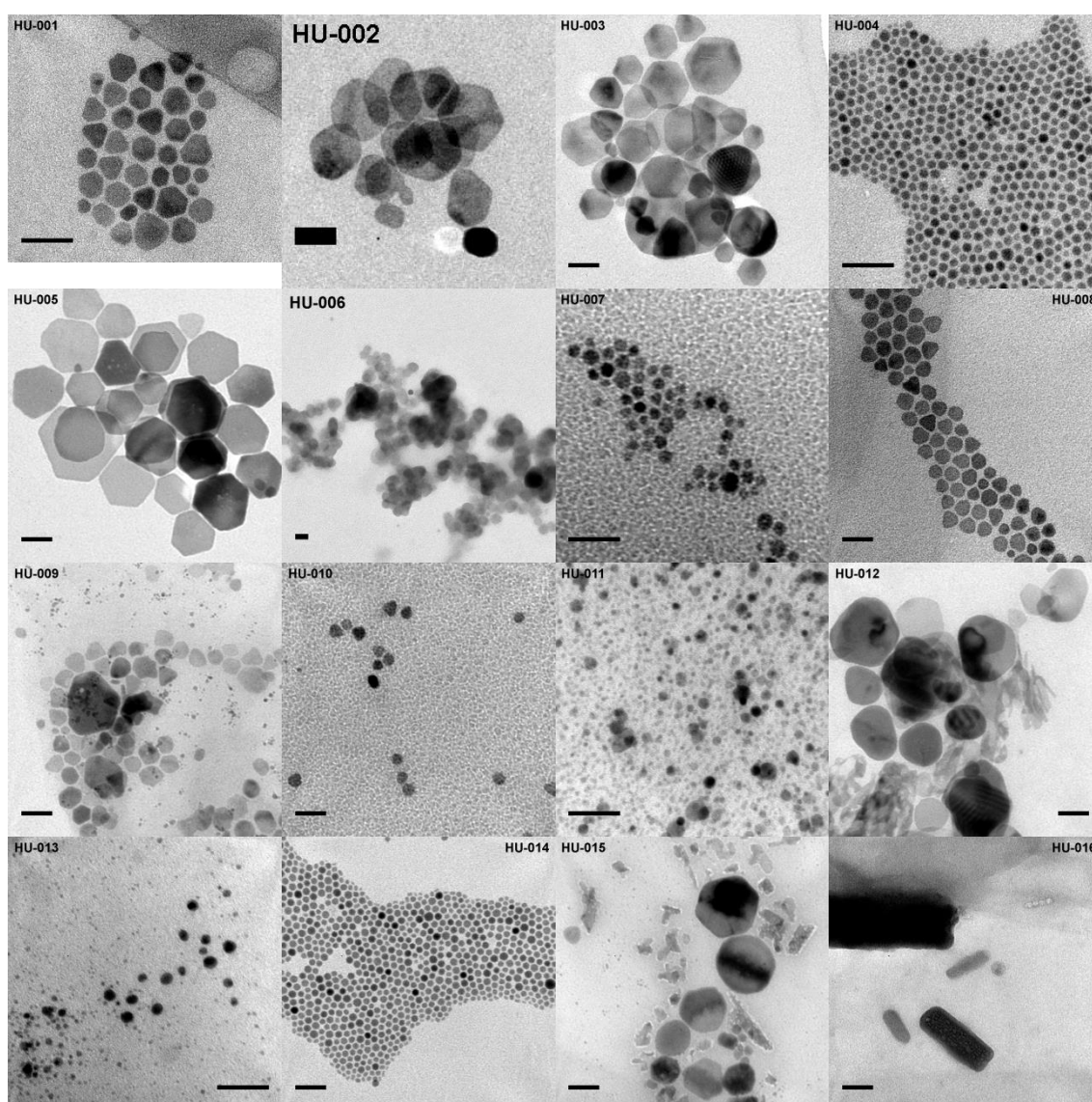


Figura 62: Imagens de MET das nanopartículas sintetizadas pelo método *heat up* sob 16 diferentes condições de síntese, de acordo com o planejamento experimental. Barra de escala: 50 nm.

As Figuras 63 a 66 mostram os histogramas dos lados dos hexágonos das amostras que apresentaram este tipo de formação (HU-001, HU-002, HU-003, HU-005). Como é possível observar nas imagens de MET (Figura 62), existe um certo grau de assimetria entre os lados dos hexágonos, o que fica evidenciado pela diferença de tamanho médio de cada lado. Os histogramas também mostram a possível presença de duas populações de placas hexagonais com tamanhos diferentes nas amostras HU-003 e HU-005.

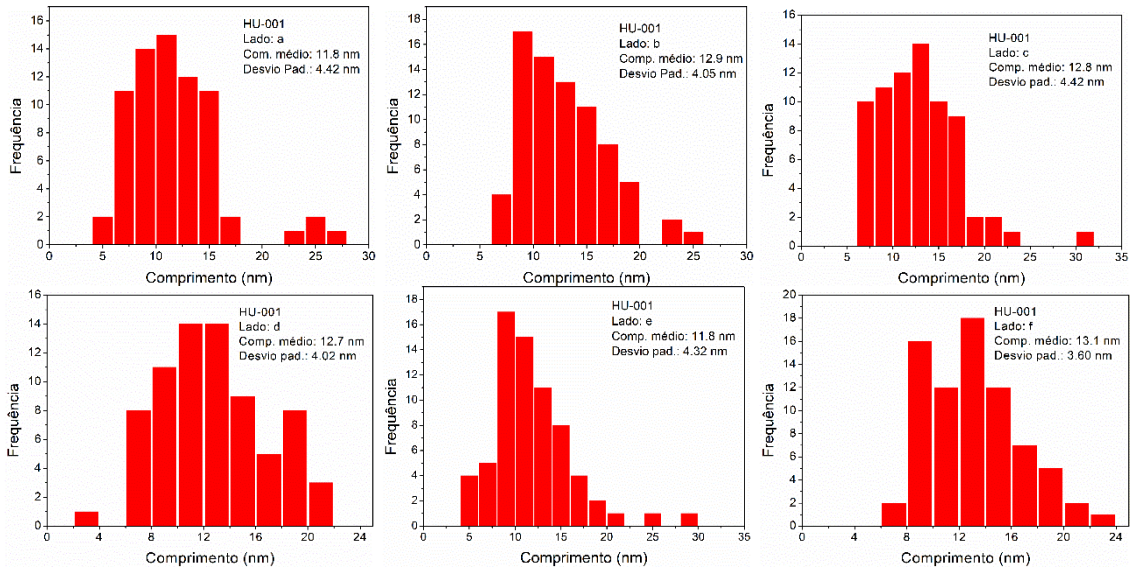


Figura 63: Histograma dos lados das placas hexagonais da amostra HU-001.

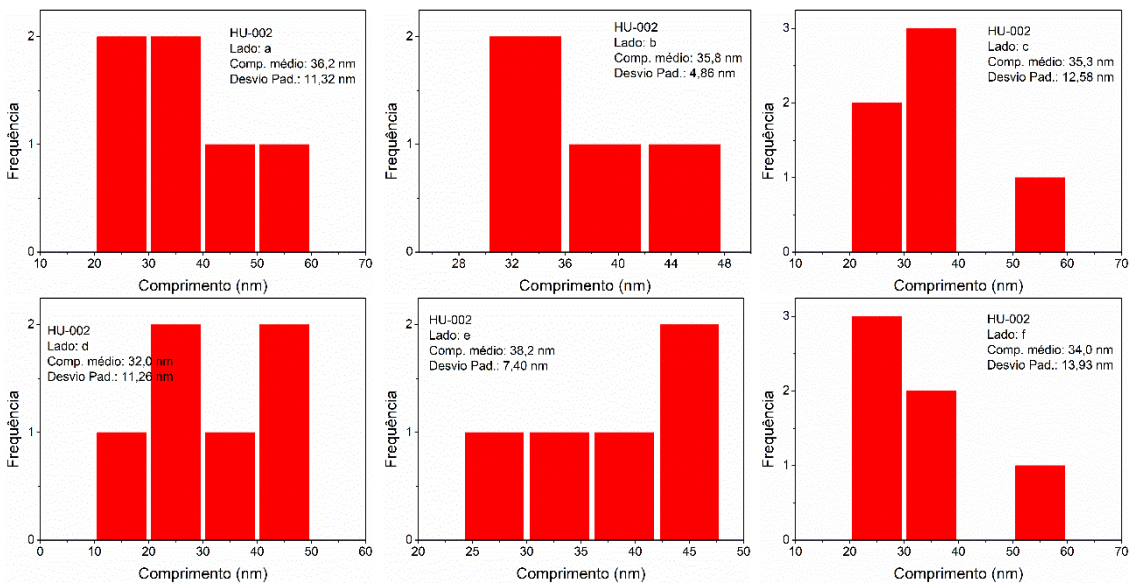


Figura 64: Histograma dos lados das placas hexagonais da amostra HU-002.

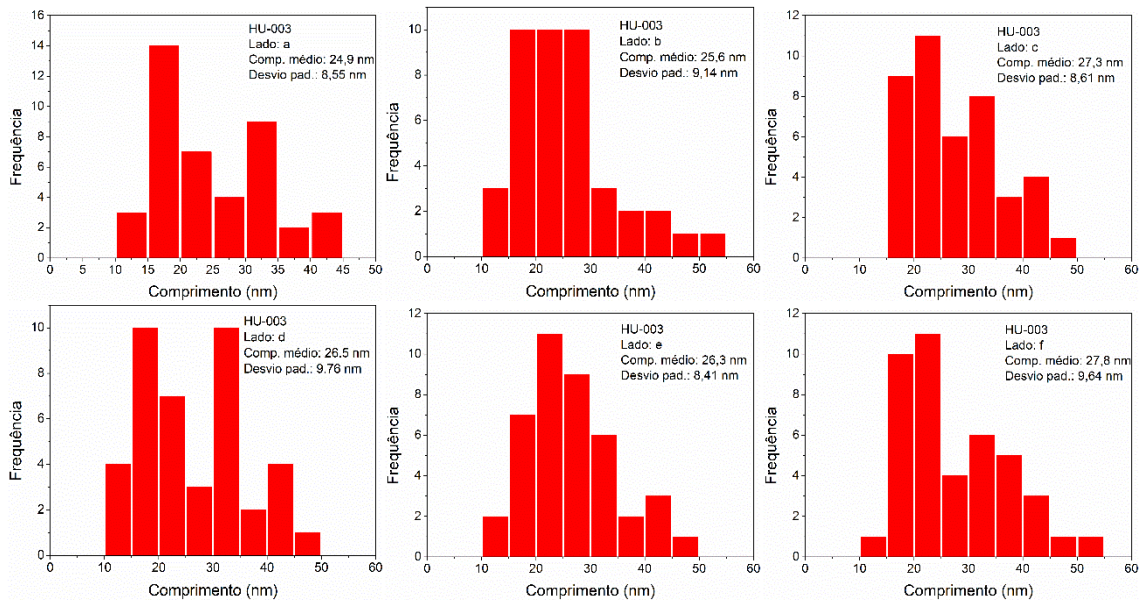


Figura 65: Histograma dos lados das placas hexagonais da amostra HU-003.

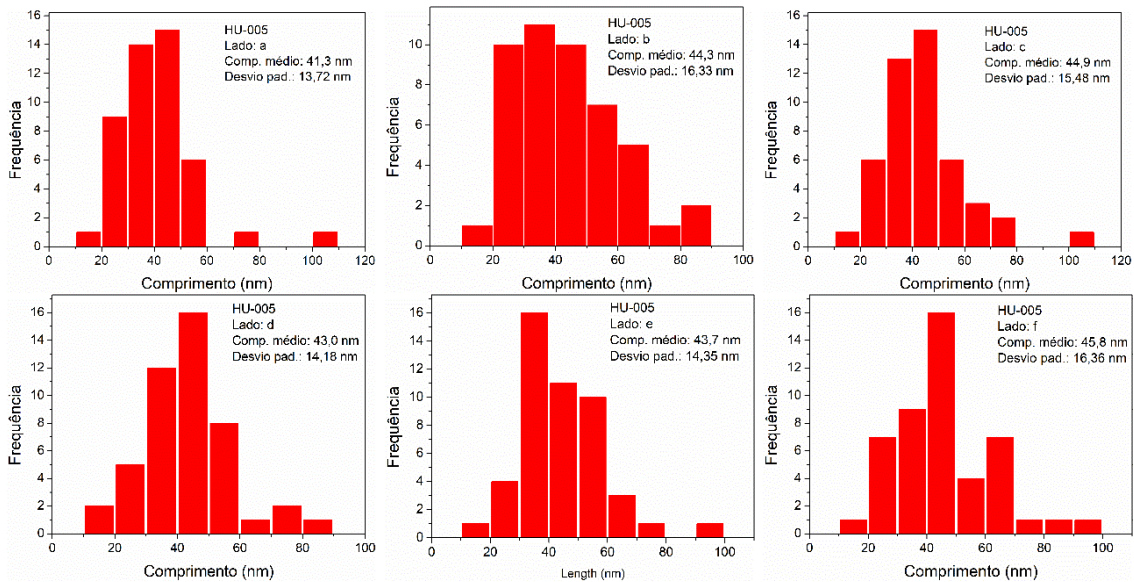


Figura 66: Histograma dos lados das placas hexagonais da amostra HU-005.

Os histogramas das amostras formadas por nanopartículas de formato esférico ou pseudo-esférico são mostrados na Figura 67. O tamanho médio das nanopartículas encontra-se entre $4,2 \pm 0,73$ e $68,6 \pm 24,63$ nm. As amostras HU-009, HU-010, HU-013 e HU-015 demonstram possuir mais de uma população com tamanhos diferentes na mesma amostra.

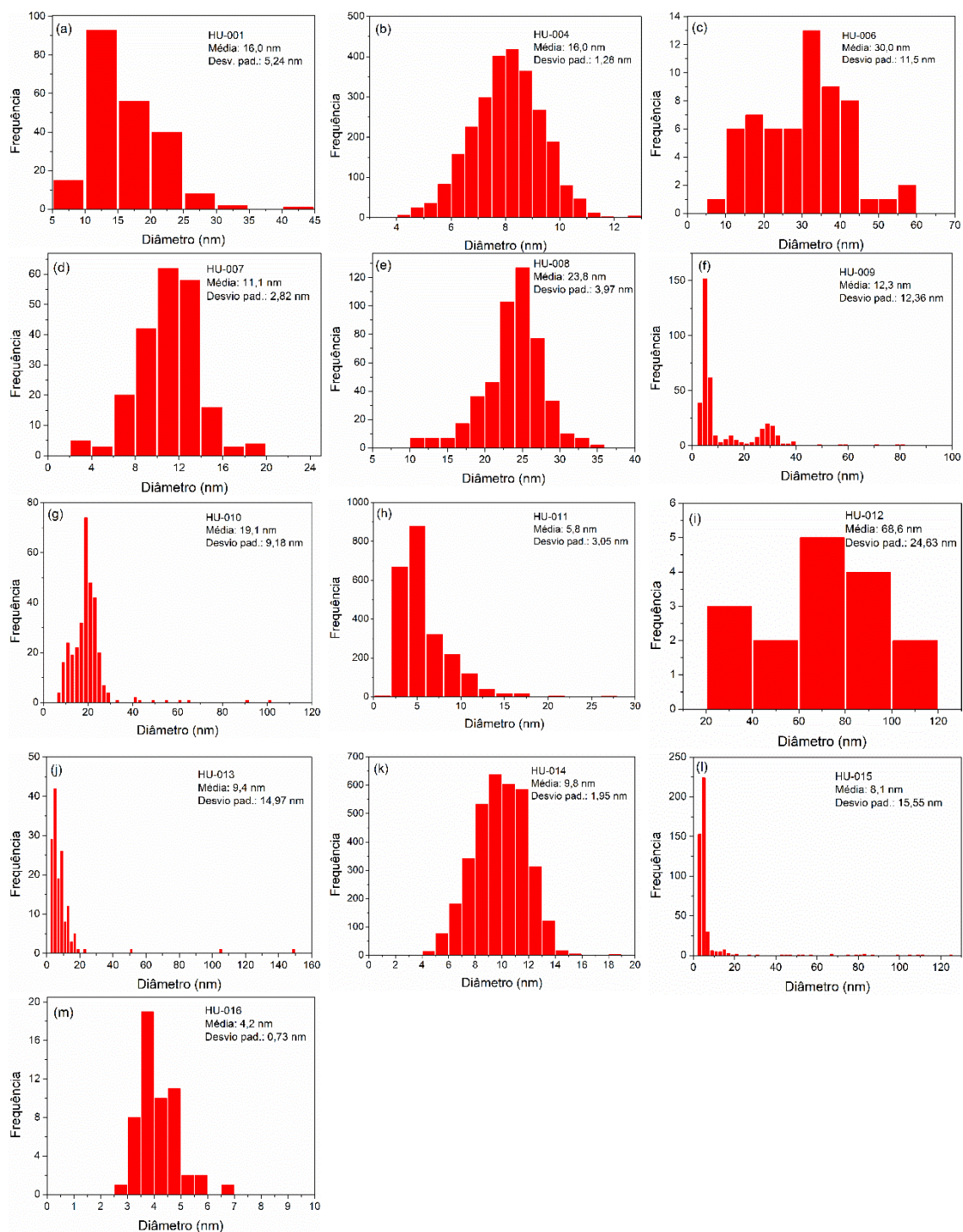


Figura 67: Histograma dos diâmetros das amostras HU-001 (a), HU-004 (b), HU-006 (c), HU-007 (d), HU-008 (e), HU-009 (f), HU-010 (g), HU-011 (h), HU-012 (i), HU-013 (j), HU-014 (k), HU-015 (l) e HU-016 (m).

A Figura 68 mostra os difratogramas de raios-x das amostras sintetizadas pelo método heat up. Nas amostras HU-001, HU-003, HU-005, HU-007, HU-010, HU-012, os picos em $29,2^\circ$, $48,4^\circ$ e $57,6^\circ$ estão relacionados com a presença de CuAlS_2 (COD 900-9930 e COD 154-2201). Além da formação de CuAlS_2 , as amostras HU-001, HU-003, HU-004, HU-005, HU-007, HU-010 e HU-

012 apresentaram picos em 16,15°, 17,62°, 31,53°, 32,18° e 39,76°, indicando a presença de subprodutos como Atacamite (COD 900-7718 e COD 900-7426); a amostra HU-011 apresentou pico em 16,25° e 39,90°, indicando a presença de Paratacamite (COD 900-7604); as amostras HU-001, HU-003, HU-005, HU-007, HU-010 apresentaram picos em 27,67°, 29,27°, 31,78° e 47,92°, indicando a presença de Covellite (COD 900-0523 e COD 900-0723); as amostras HU-011, HU-012 e HU-016 apresentaram picos em 11,53°, 27,2° e 43,3°, indicando a presença de AIOCl (oxicloreto de alumínio, COD 900-9194). A amostra HU-002 apresentou apenas picos relativos a CuS_2 (COD 210-1636). As tabelas com os difratogramas e as associações com os picos nas cartas cristalográficas encontram-se no Anexo 2.

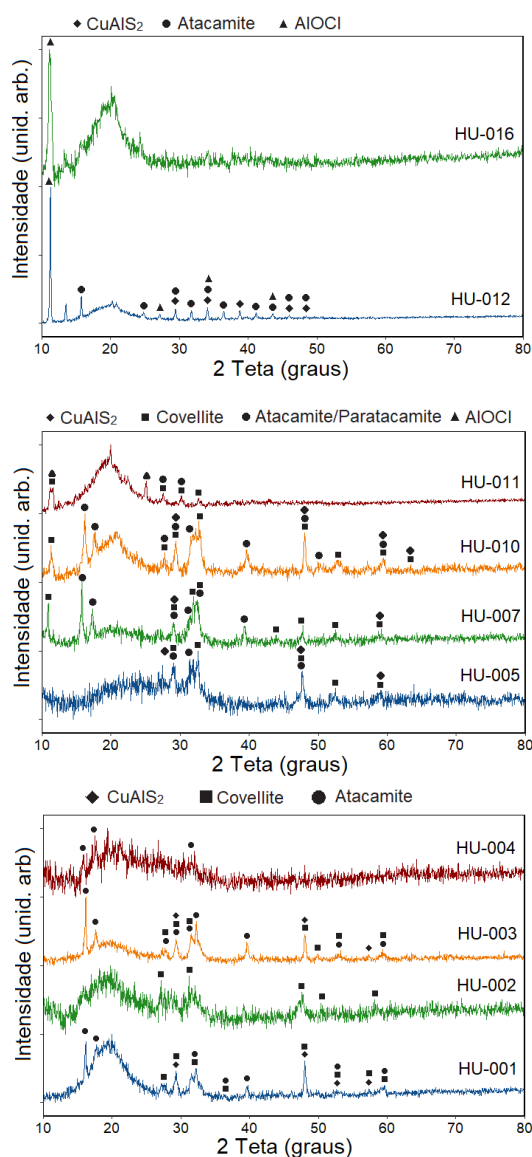


Figura 68: Difratogramas de raios-x das amostras sintetizadas pelo método *heat up*.

A Figura 69 mostra os espectros Raman das amostras sintetizadas pelo método *heat up*. Os pontos indicados pela linha tracejada em 137, 262, 272, 325 e 457 cm^{-1} são os picos relacionados à presença de CuAIS_2 .^[29,84,86,105-110] O pico em 470 cm^{-1} é tipicamente relacionado à estrutura Covellite (CuS).^[111-113] É possível observar também indicativos da presença de Atacamite, Paratacamite e Clinoatacamite. As tabelas completas com os picos e as associações com os picos contidos na literatura encontram-se no Anexo 3.

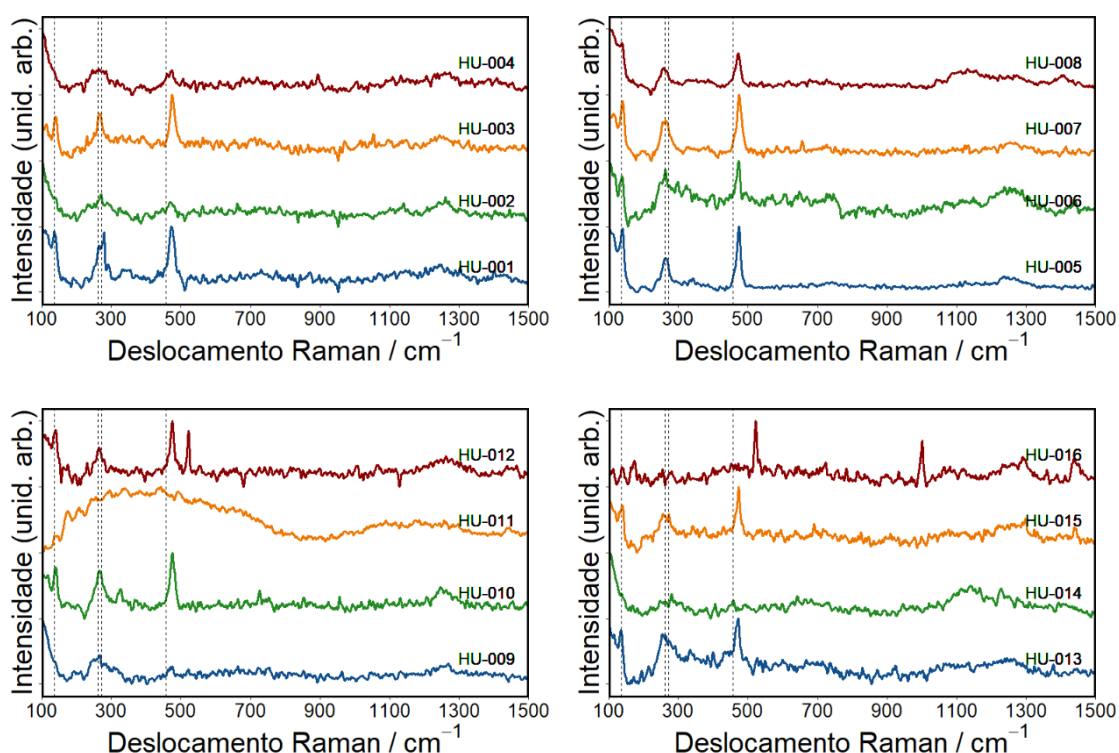


Figura 69: Espectros Raman das amostras sintetizadas pelo método *heat up*.

A Figura 70 ilustra os espectros de absorção no UV-Vis-NIR das amostras sintetizadas pelo método *heat up*, sob 16 condições de síntese distintas. É possível observar que algumas amostras apresentam uma larga banda de extinção no infravermelho próximo, típica de materiais que apresentam efeito de ressonância de plásmon de superfície. A presença de efeito plasmônico em nanoestruturas de calcogenetos de cobre (Cu_xS_y , Cu_xSe_y) já é amplamente aceita pela comunidade científica. A origem desse efeito tem sido atribuída à presença de portadores de cargas livres, gerados por deficiências de cobre na estrutura.^[104-116] Além de ter sido observado nos calcogenetos binários de cobre, o efeito plasmônico já foi relatado em nanoestruturas de $\text{Cu}_x\text{In}_y\text{S}_z$ e Cu-Zn-Sn

S.^[115,117] Entre as possíveis aplicações do efeito plasmônico estão as células solares plasmônicas^[117] e aplicações em terapias fototérmicas.^[75,115]

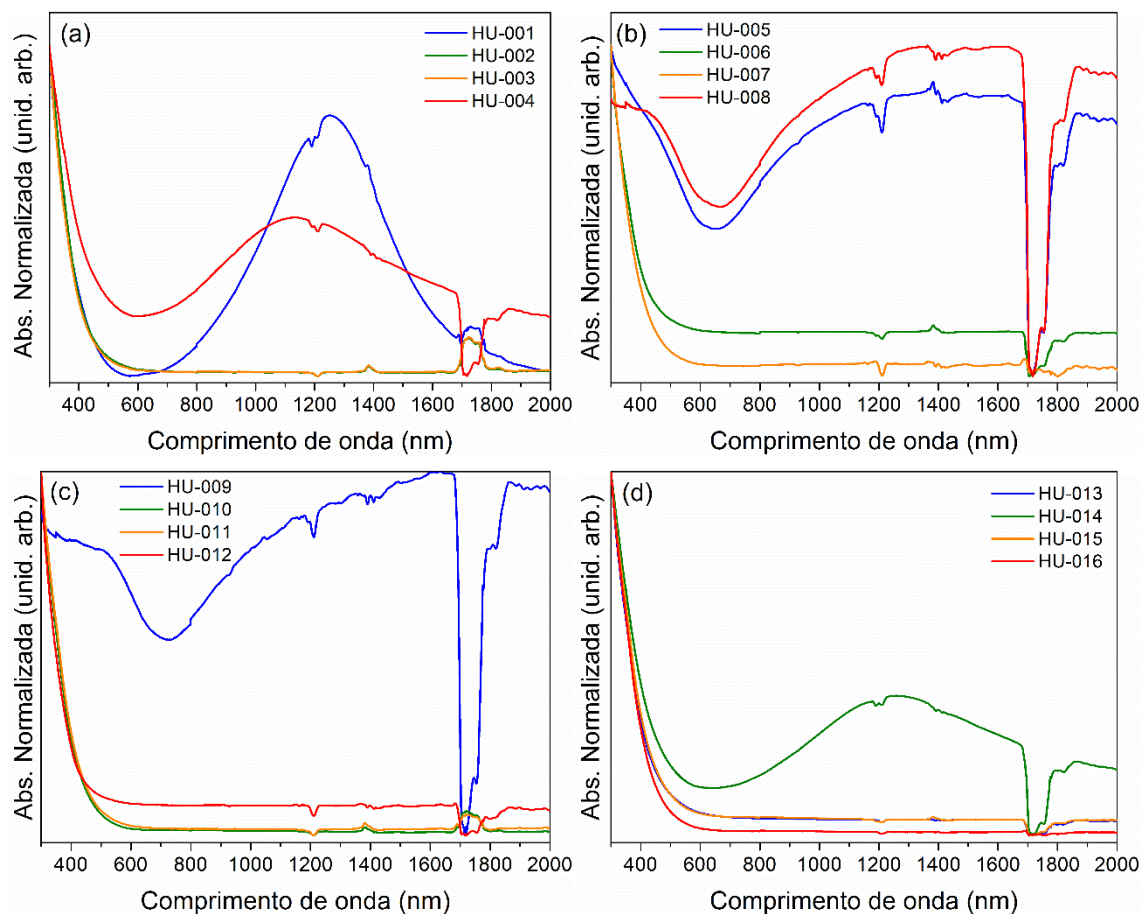


Figura 70: Espectros de absorção UV-Vis-NIR das nanopartículas de sulfeto de cobre e alumínio sintetizadas pelo método *heat up*.

Os valores de *bandgap* óptico de cada amostra, obtidos a partir do gráfico de Tauc (apresentados no Anexo 4), são mostrados na Tabela 7. É possível observar que o *bandgap* direto encontrado está entre 2,99 e 3,55 eV, valores compatíveis com os encontrados na literatura para CuAlS_2 .^[80,85,118] Por outro lado, as amostras também apresentaram transição indireta, com energias entre 1,52 e 2,24 eV. Foi realizado um teste t para determinar se os resultados se tratam de variação em um mesmo grupo ou se tratam de amostras de diferentes grupos. Na Tabela 7, amostras com valor p inferior a 0,05 são consideradas não pertencentes ao mesmo grupo. Ou seja, é possível dizer que há diferença significativa entre os valores de *bandgap* das amostras e que estes não fazem parte de um único grupo homogêneo. A média dos valores de *bandgap* direto foi de 3,27 eV, com desvio padrão de 0,19 eV. Para *bandgap* indireto, o valor médio encontrado foi de 1,94 eV, com desvio padrão de 0,2 eV.

A literatura não contém registros de CuAlS_2 possuindo *bandgap* indireto, sendo assim possível assumir que a transição indireta é proveniente de uma impureza ou co-produto.

Tabela 10: Valores de *bandgap* direto e indireto calculado a partir do gráfico de Tauc, das amostras sintetizadas pelo método *heat up*.

Amostra	<i>Bandgap</i> direto (eV)	Valor p	<i>Bandgap</i> indireto (eV)	Valor p
HU001	3,6	0,50	2,2	0,22
HU002	3,5	0,12	2,3	0,33
HU003	3,6	0,33	2,4	0,01
HU004	3,4	0,50	2,0	0,00
HU005	3,3	0,03	2,4	0,01
HU006	3,6	0,25	2,4	0,01
HU007	3,6	0,03	2,3	0,33
HU008	3,3	0,98	2,6	0,00
HU009	3,6	0,02	2,4	0,01
HU010	3,5	0,01	2,2	0,21
HU011	3,5	0,00	2,1	0,00
HU012	3,5	0,03	2,3	0,33
HU013	3,3	0,00	2,1	0,00
HU014	3,3	0,00	1,9	0,00
HU015	3,2	0,66	2,2	0,22
HU016	3,6	0,10	2,3	0,33

A Figura 71-a mostra um histograma dos valores, em elétron-volt, dos *bandgaps* direto e indireto para as amostras sintetizadas pelo método *heat-up*. A Figura 71-b, por sua vez, traz um diagrama de caixas que compara os resultados obtidos neste trabalho com os valores de *bandgap* existentes na literatura, tanto para estudos experimentais de CuAlS_2 , bem como para estudos teóricos. Observando esta comparação, é possível concluir que os valores de *bandgap* direto encontrados neste trabalho são compatíveis com a literatura. Considerando que, dos principais compostos encontrados através das análises de difratometria de raios-x e espectroscopia Raman, a Paratacamite possui um *bandgap* de natureza indireta de valor 2,67 eV,^[119] e a Covellite possui uma faixa de valores de *bandgap* entre 1,80 e 2,25 eV,^[112,120-122] é possível concluir que o material que gerou as transições diretas observadas nos espectros e resumidas na Tabela 10, trata-se de CuAlS_2 .

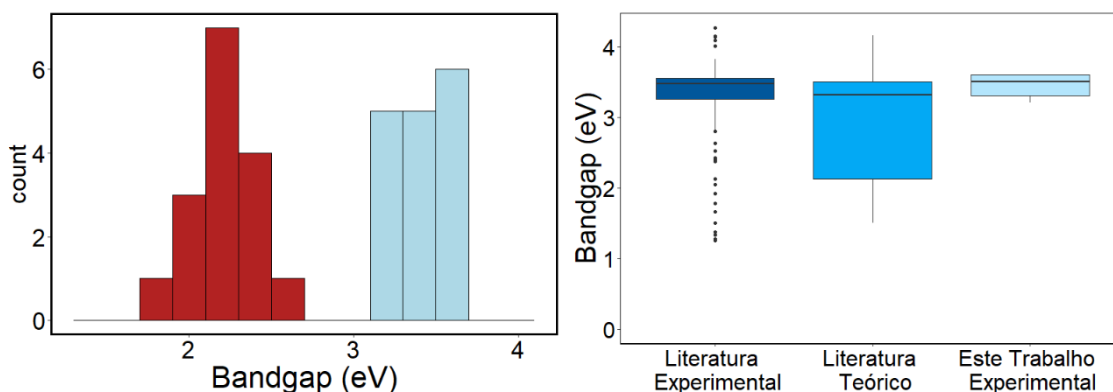


Figura 71: Histograma dos valores de bandgap indireto (vermelho) e direto (azul) das amostras sintetizadas pelo método *heat up* (a) e diagrama de caixas comparando os valores de *bandgap* direto da literatura com aqueles encontrados neste trabalho (b).

A Figura 72 traz os efeitos dos parâmetros de síntese no *bandgap* das amostras. Tanto para o valor de bandgap direto como para bandgap indireto, os parâmetros de síntese apresentaram uma influência muito pequena, tendo magnitude dos efeitos inferior a 0,1 eV.

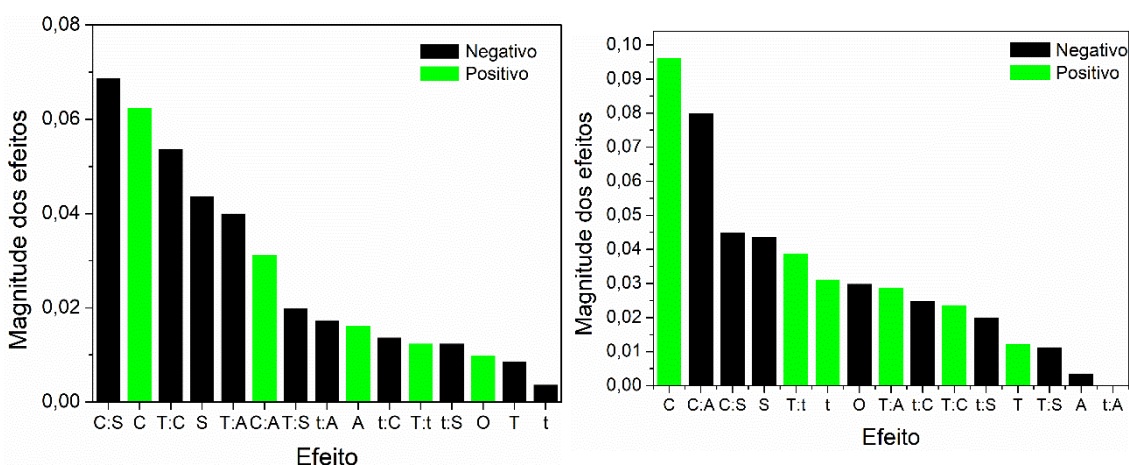


Figura 72: Gráfico de Pareto dos efeitos dos parâmetros de síntese no *bandgap* direto (esquerda) e indireto (direita) das amostras sintetizadas pelo método *heat up*.

As Figuras 73 a 76 mostram os espectros de emissão de fluorescência das amostras sintetizadas nas 16 diferentes condições de síntese. Em uma varredura inicial, os comprimentos de onda de excitação usados nos ensaios foram 350, 400, 450, 500 e 550 nm. Para a padronização das leituras, foram usadas dispersões das nanopartículas em hexano, cuja absorção em 300 nm estivesse entre 0,27 e 0,3. As maiores intensidades de emissão foram observadas para a excitação a 350 nm. Para este comprimento de onda, o pico de emissão ocorre entre 405 e 436 nm, sendo que a fotoluminescência

reportada por Poulouse e colaboradores para nanopartículas CuAIS₂ continha um pico centrado em cerca de 440 nm.^[30] Algumas amostras apresentam um pico estreito em 390 nm, que pode ser atribuído à oleilamina em excesso.

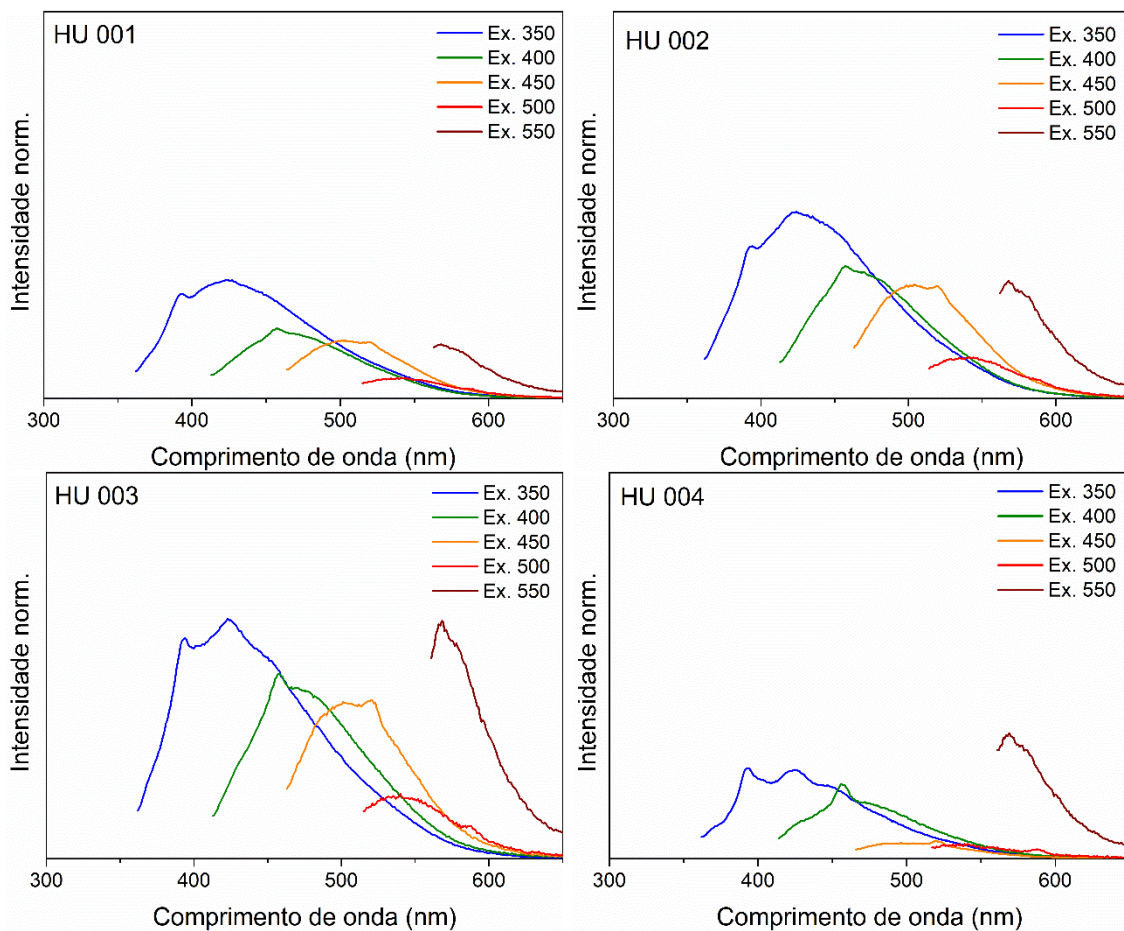


Figura 73: Espectros de fotoluminescência das nanopartículas sintetizadas pelo método *heat up* sob as condições HU-001, HU-002, HU-003 e HU-004.

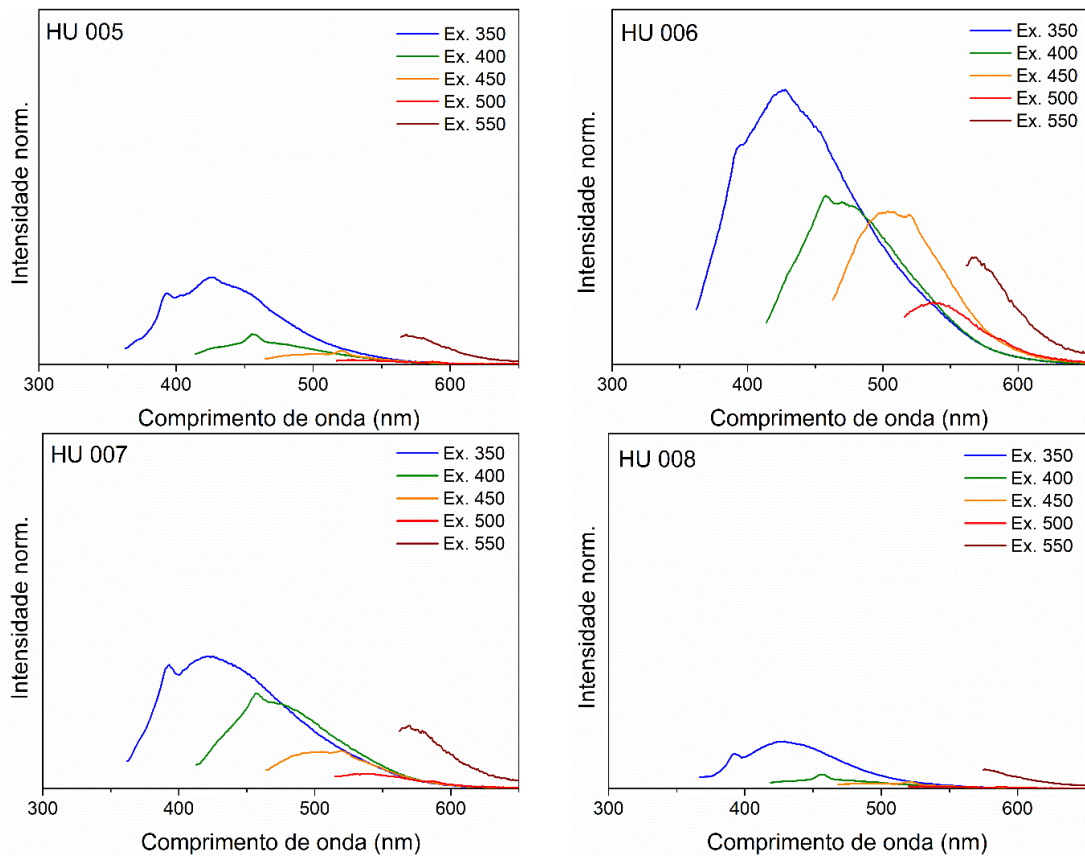


Figura 74: Espectros de fotoluminescência sintetizadas pelo método *heat up* sob as condições HU-005, HU-006, HU-007 e HU-008.

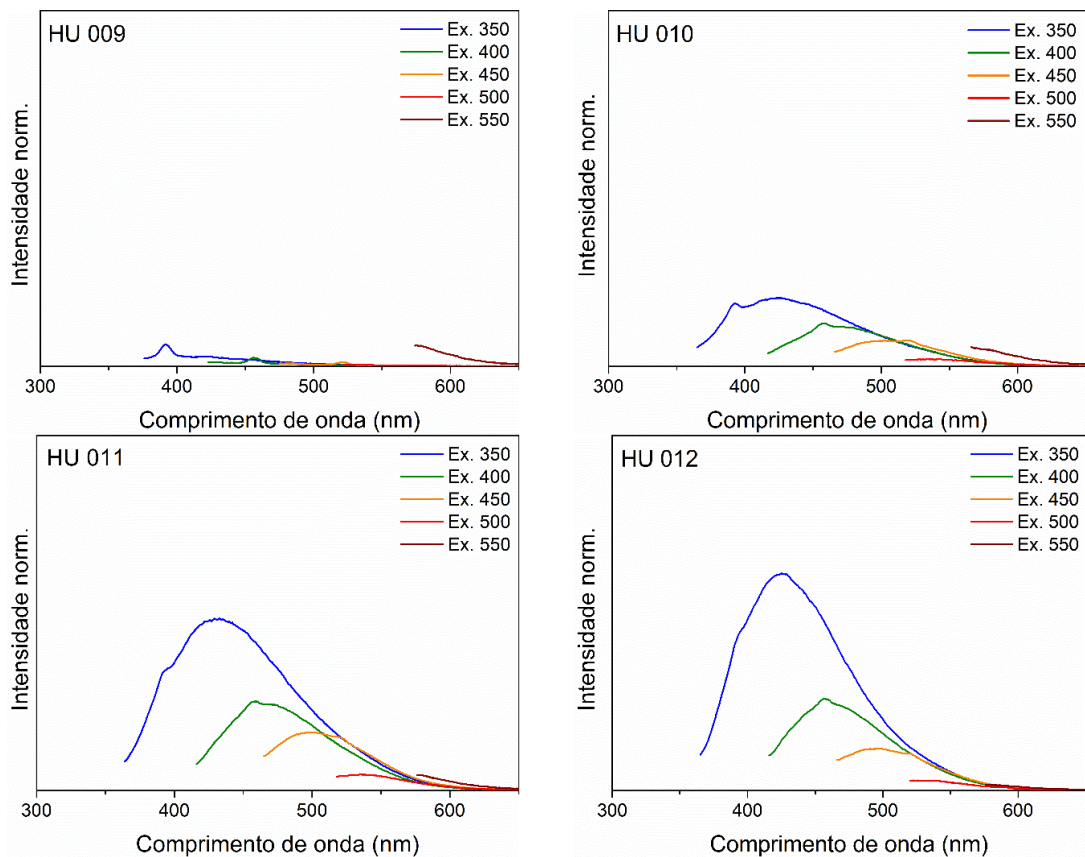


Figura 75: Espectros de fotoluminescência das nanopartículas sintetizadas pelo método *heat up* sob as condições HU-009, HU-010, HU-011 e HU-012.

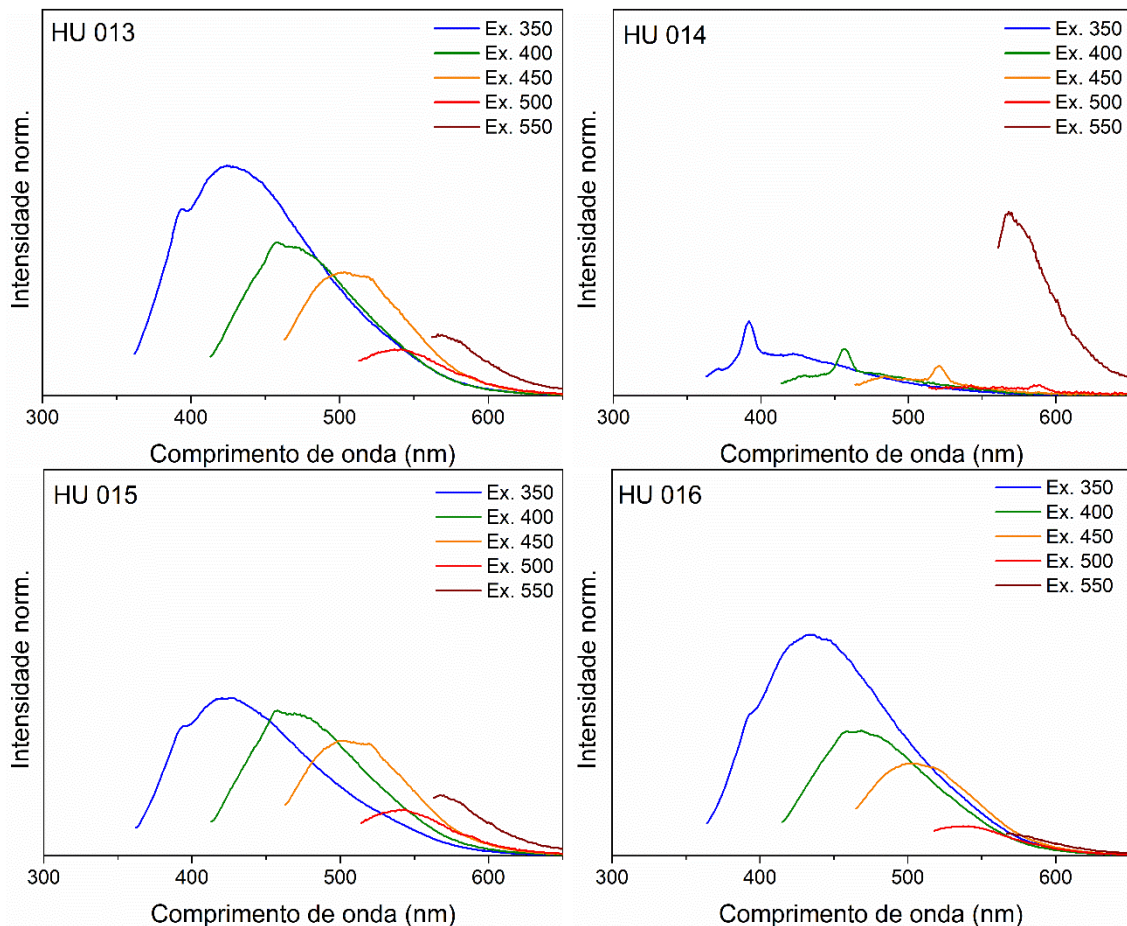


Figura 76: Espectros de emissão de fluorescência das nanopartículas sintetizadas pelo método *heat up* sob as condições HU-013, HU-014, HU-015 e HU-016.

Os valores da intensidade de emissão de cada amostra, obtidos a partir dos espectros de fotoluminescência, são mostrados na Tabela 11. É possível observar que a intensidade de emissão está entre 232,3 e 11,2 para amostras excitadas por um feixe de 350 nm, entre 130,5 e 10,2 para feixe de 400 nm, entre 96,8 e 3,7 para feixe de 450 nm e entre 31,3 e 0 para amostras excitadas por feixe de 500 nm. Também é possível observar que a amostra que apresenta maior intensidade de fotoluminescência quando excitada por um feixe de 350 nm (HU-012) não é a mesma que apresenta a intensidade maior nos demais comprimentos de onda (HU-016).

Tabela 11: Intensidades de fluorescência das nanopartículas.

Amostra	Intensidade			
	Ex. 350	Ex. 400	Ex. 450	Ex. 500
HU-001	124,4	63,9	48,7	-
HU-002	141,3	73,7	48,0	13,4
HU-003	87,6	45,0	31,0	9,0
HU-004	43,5	19,3	11,9	-
HU-005	61,2	17,8	8,3	-
HU-006	158,9	73,0	58,3	17,8
HU-007	78,3	44,3	27,4	7,5
HU-008	53,9	16,1	6,1	-
HU-009	11,2	10,2	8,3	-
HU-010	78,1	45,7	29,6	8,3
HU-011	196,6	99,0	66,5	18,3
HU-012	248,4	99,0	48,1	11,8
HU-013	159,2	88,2	62,7	20,5
HU-014	17,2	12,3	3,7	-
HU-015	166,2	82,2	60,1	19,4
HU-016	232,3	130,5	96,8	31,3

As posições dos picos de emissão, em elétron-volt, de cada amostra, obtidas a partir dos espectros de fotoluminescência, são mostradas na Tabela 9. É possível observar que a energia de emissão está entre 2,84 e 3,06 eV para amostras excitadas por um feixe de 350 nm, entre 2,65 e 2,71 eV para feixe de 400 nm, entre 2,38 e 2,56 eV para feixe de 450 nm e entre 2,29 e 2,48 eV para amostras excitadas por feixe de 500 nm.

Tabela 12: Energia de emissão das amostras.

Amostra	Energia (eV)			
	Ex. 350	Ex. 400	Ex. 450	Ex. 500
HU-001	2,92	2,68	2,47	-
HU-002	2,91	2,67	2,46	2,30
HU-003	2,93	2,66	2,47	2,48
HU-004	2,92	2,66	2,49	-
HU-005	2,92	2,66	2,48	-
HU-006	2,90	2,65	2,47	2,30
HU-007	2,94	2,66	2,48	2,30
HU-008	2,90	2,71	2,56	-
HU-009	3,05	2,71	2,38	-
HU-010	2,92	2,65	2,47	2,30
HU-011	2,87	2,66	2,48	2,32
HU-012	2,91	2,68	2,51	2,34
HU-013	2,92	2,66	2,47	2,30
HU-014	3,06	2,71	2,56	-
HU-015	2,93	2,65	2,47	2,29
HU-016	2,84	2,66	2,47	2,30

A Figura 77-a mostra um histograma dos valores, em elétron-volt, da energia de emissão das amostras sintetizadas pelo método *heat up*. Na Figura 77-b, por sua vez, traz um diagrama de caixas que compara os resultados obtidos neste trabalho com os valores de emissão existentes na literatura, tanto para estudos experimentais de CuAlS₂, bem como para estudos teóricos. Através desta comparação, é possível concluir que os valores encontrados neste trabalho são compatíveis com a literatura.

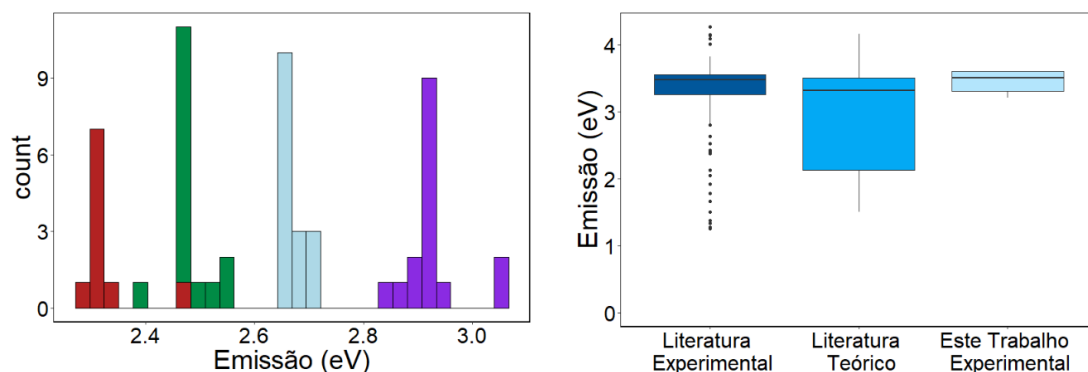


Figura 77: Histograma dos valores de energia de fotoluminescência das amostras (a) e diagrama de caixas comparando os valores da literatura com aqueles encontrados neste trabalho (b).

A Tabela 13 traz uma comparação dos valores de energia de emissão de fotoluminescência das amostras com valores encontrados na literatura em estudos teóricos e medições experimentais. É possível observar valores compatíveis com os encontrados neste trabalho, com uma diferença máxima de 0,09 eV. Este resultado corrobora com dados de XRD e Raman mostrando a formação de CuAlS₂ e que sua fotoluminescência é compatível com os valores reportados em literatura.

Tabela 13: Comparação da energia de emissão das amostras sintetizadas pelo método *heat up* com os valores encontrados na literatura.

Amostra	Ex. 350	Literatura				
		Teórico		Experimental		
HU-001	2,92	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-002	2,91	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	
HU-003	2,93	2,87 ^[123]	3,02 ^[123]	2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-004	2,92	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-005	2,92	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-006	2,90	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	
HU-007	2,94	2,87 ^[123]	3,02 ^[123]	2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-008	2,90	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	
HU-009	3,05		3,02 ^[123]			3,01 ^[127]
HU-010	2,92	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-011	2,87	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	

HU-012	2,91	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	
HU-013	2,92	2,87 ^[123]		2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-014	3,06		3,02 ^[123]			3,01 ^[127]
HU-015	2,93	2,87 ^[123]	3,02 ^[123]	2,95 ^[124,125]	2,90 ^[31,126]	3,01 ^[127]
HU-016	2,84	2,87 ^[123]			2,90 ^[31,126]	

A Figura 78 traz o gráfico de Pareto para os efeitos de cada parâmetro da síntese na intensidade de fotoluminescência. Quando as amostras são excitadas por diferentes comprimentos de onda, ocorre emissão de fotoluminescência por fenômenos diferentes, que podem ser tanto provenientes de diferentes defeitos que geram emissão, bem como diferentes populações de nanopartículas. Ao serem excitadas por um feixe com comprimento de onda de 350 e 400 nm, as amostras apresentam como principal parâmetro que influencia positivamente na intensidade de fotoluminescência, a concentração do precursor de alumínio. A interação da concentração do precursor de alumínio com o tempo, também positiva, mostra que pode haver um período de fixação deste precursor ao longo do tempo de reação, resultando em maior intensidade de fluorescência. Esse resultado é muito interessante, uma vez que, em estudos teóricos, foi mostrado que há uma emissão de fotoluminescência gerada por defeitos, ocasionada por uma substituição de cobre por alumínio na rede cristalina de calcopirita de sulfeto de cobre e alumínio.^[128]

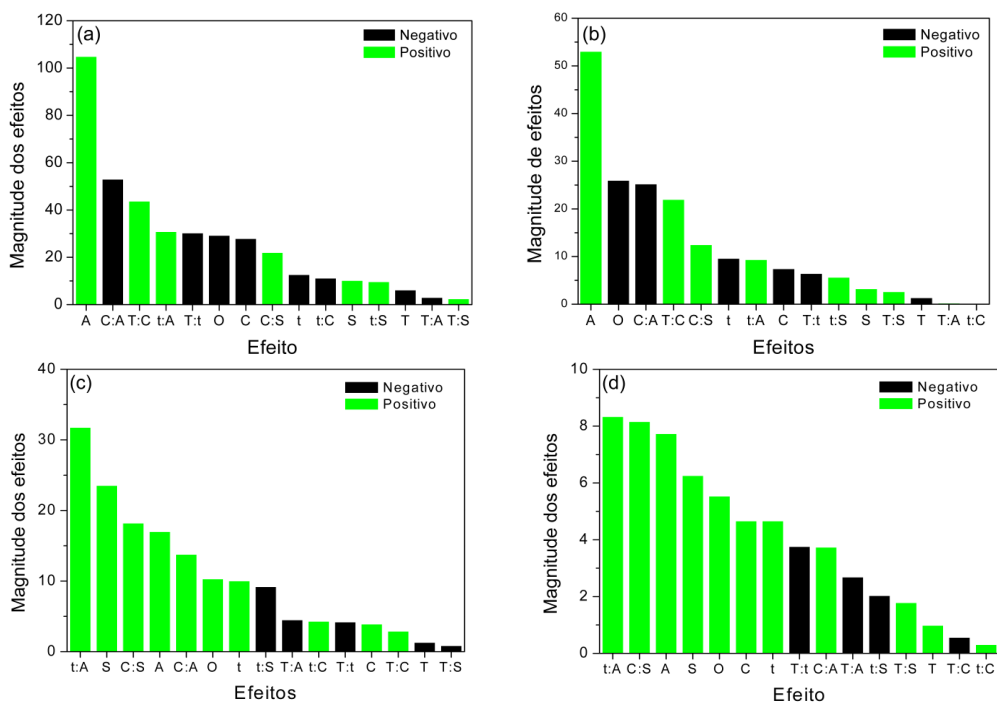


Figura 78: Gráfico de Pareto dos efeitos dos parâmetros da síntese *heat up* na intensidade da fluorescência das nanopartículas quando excitadas por um feixe com comprimento de onda de (a) 350, (b) 400, (c) 450 e (d) 500 nanômetros.

4.2.1. Segundo Planejamento Experimental

Após a concentração do precursor de alumínio ter sido identificado como um dos fatores que aumentam a intensidade de fluorescência e a concentração do precursor de cobre apresentar efeito no sentido oposto, um novo planejamento foi realizado com aumento da quantidade de cloreto de alumínio (III) e redução de cloreto de cobre (I) na reação (ver Tabela 7). A Figura 79 mostra os espectros de absorção no UV-Vis das amostras sintetizadas sob essas novas condições.

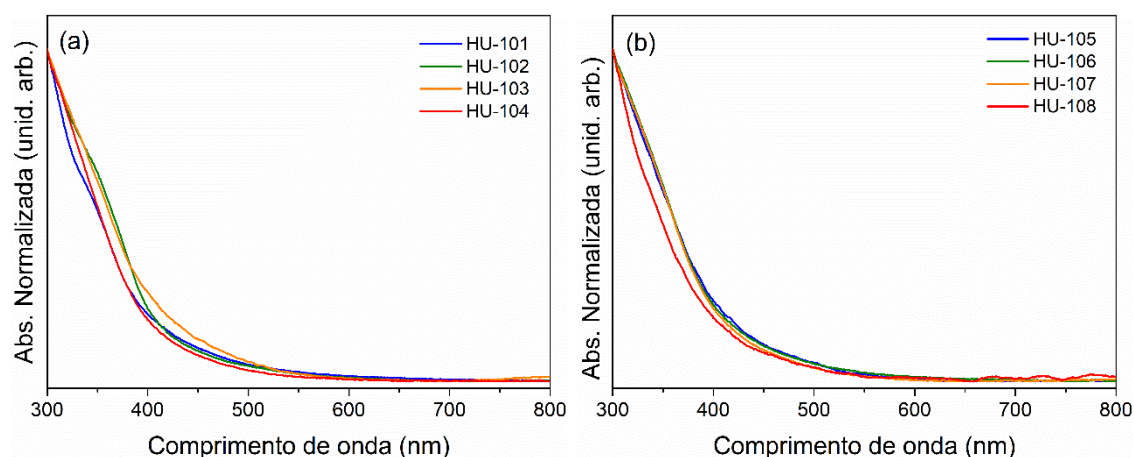


Figura 79: Espectros de absorção na região do UV-Vis das nanopartículas de sulfeto de cobre e alumínio sintetizadas nas condições descritas na Tabela 7.

As Figuras 80 e 81 mostram os espectros de emissão de fotoluminescência das amostras sintetizadas sob as condições constantes nas Tabelas 7 e 8. Assim como nos espectros anteriores, os comprimentos de onda de excitação usados foram 350, 400, 450, 500 e 550 nm e os picos referentes ao espalhamento de luz nestes comprimentos aparecem nos espectros. Para a padronização das leituras, foram usadas dispersões das nanopartículas em hexano, cuja absorção em 300 nm estivesse entre 0,27 e 0,3. As maiores intensidades de emissão foram observadas para a excitação a 350 nm.

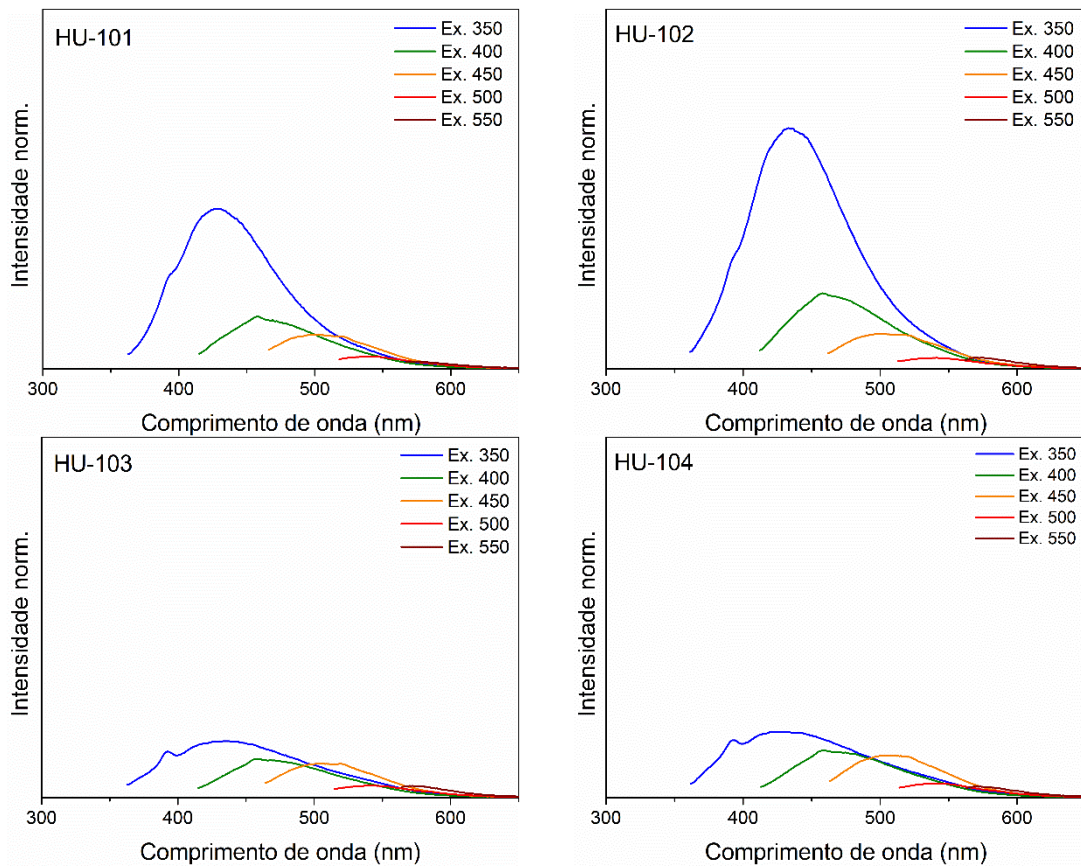


Figura 80: Espectros de fotoluminescência das nanopartículas obtida por método *heat up*, amostras HU-101, HU-102, HU-103 e HU-104.

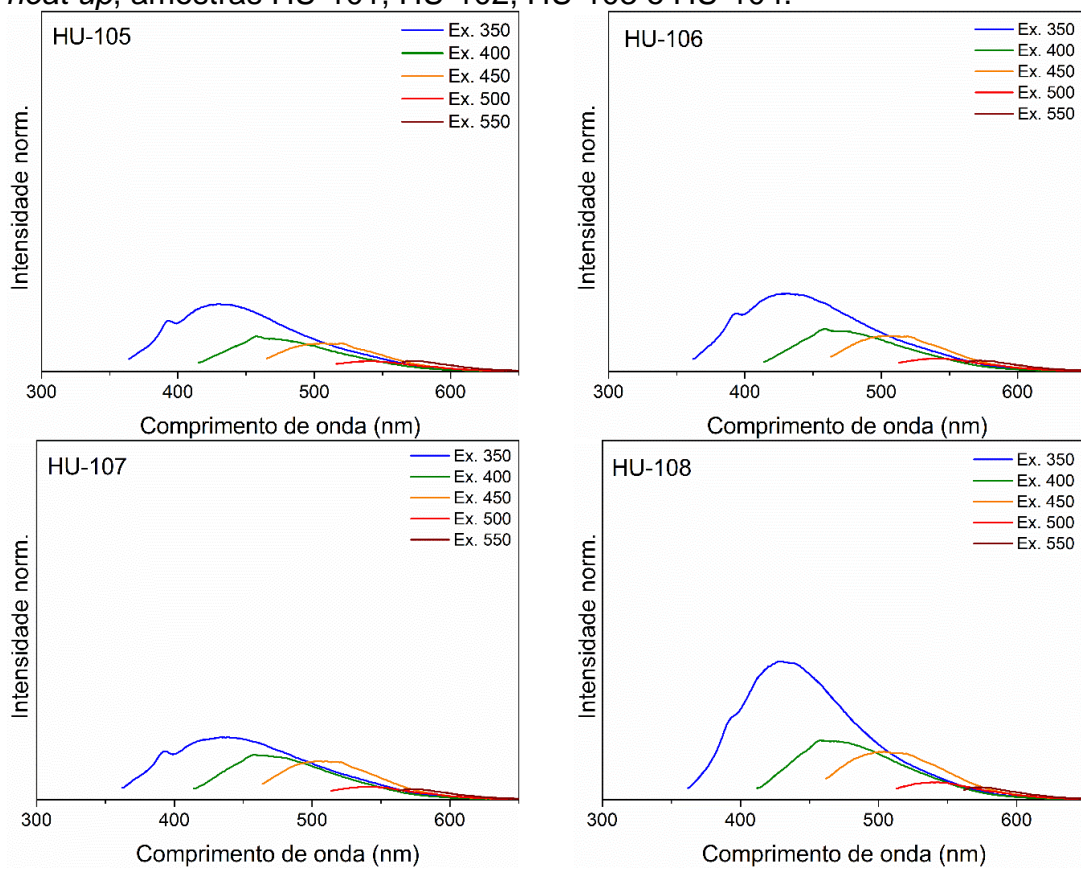


Figura 81: Espectros de fotoluminescência das nanopartículas obtida por método *heat up*, amostras HU-105, HU-106, HU-107 e HU-108.

A Tabela 14 resume as intensidades dos picos de emissão de cada amostra, sob excitação por um feixe com comprimento de onda de 350, 400, 450 e 500 nanômetros.

Tabela 14: Intensidades de fluorescência das nanopartículas sintetizadas pelo método *heat up* sob as condições de síntese daa Tabelas 7 e 8.

Amostra	Intensidade			
	Ex. 350	Ex. 400	Ex. 450	Ex. 500
HU-101	226,2	74,2	48,4	0
HU-102	340,4	107,0	49,7	16,1
HU-103	80,2	55,2	48,3	0
HU-104	93,2	66,7	59,7	20,5
HU-105	96,0	50,0	40,1	0
HU-106	110,2	60,1	50,4	19,1
HU-107	88,9	63,6	54,7	18,8
HU-108	195,7	84,2	68,0	25,2

A Tabela 15 resume energia de emissão de cada amostra, sob excitação por um feixe com comprimento de onda de 350, 400, 450 e 500 nanômetros.

Tabela 15: Energia de emissão das amostras sintetizadas pelo método *heat up* sob as condições de síntese das Tabelas 7 e 8.

Amostra	Energia (eV)			
	Ex. 350	Ex. 400	Ex. 450	Ex. 500
HU-101	2,88	2,71	2,49	-
HU-102	2,86	2,71	2,47	2,29
HU-103	2,84	2,71	2,45	-
HU-104	2,91	2,70	2,47	2,28
HU-105	2,88	2,71	2,46	-
HU-106	2,88	2,70	2,45	2,29
HU-107	2,86	2,71	2,43	2,28
HU-108	2,90	2,71	2,46	2,29

A Figura 82 traz o gráfico de Pareto para os efeitos de cada parâmetro da síntese do segundo planejamento experimental via *heat up* na intensidade de fluorescência. A intensidade da fotoluminescência emitida quando as amostras foram excitadas por feixes com comprimentos de onda de 350 e 400 nm é influenciada mais fortemente pela concentração do precursor de alumínio. O tempo de reação aparece como principal influência positiva na intensidade de fotoluminescência das amostras excitadas por feixe de comprimento de onda de 400, 450 e 500 nanômetros.

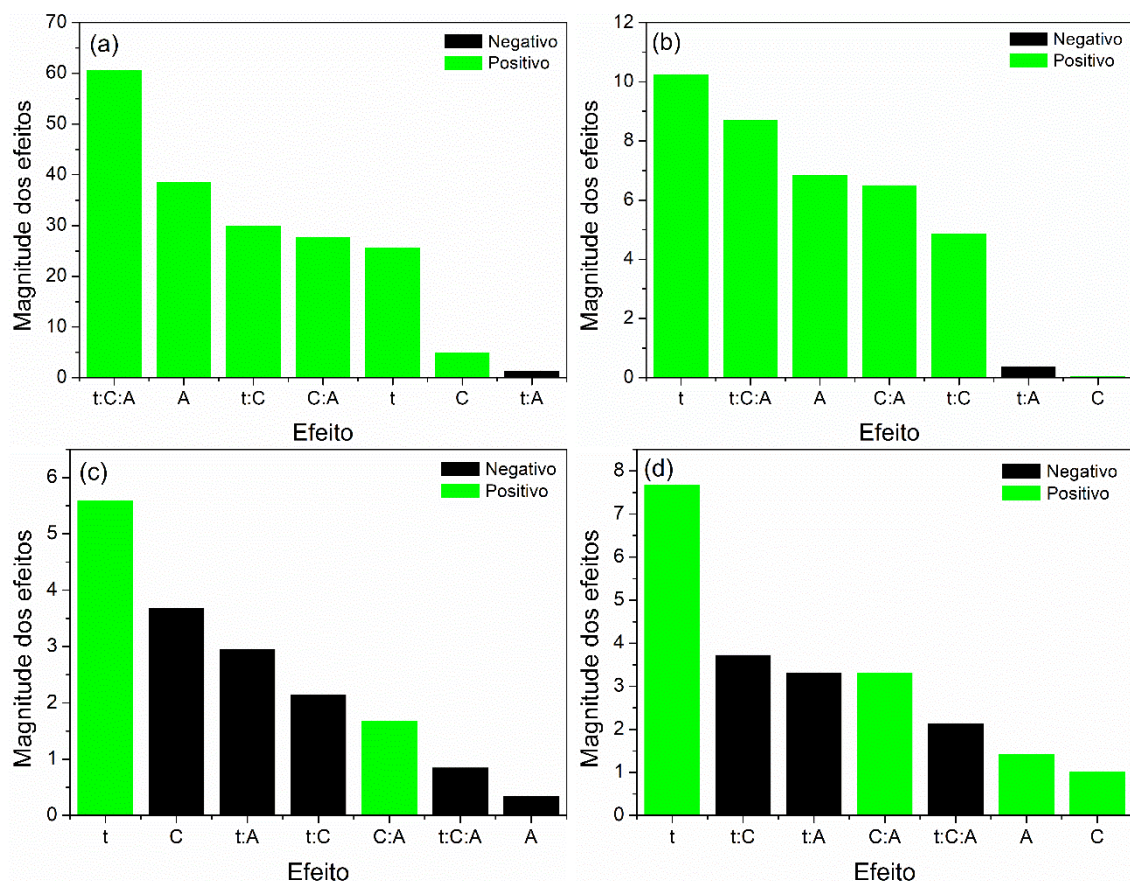


Figura 82: Gráfico de Pareto dos efeitos dos parâmetros da síntese *heat up* na intensidade da fluorescência das nanopartículas quando excitadas por um feixe com comprimento de onda de (a) 350, (b) 400, (c) 450 e (d) 500 nanômetros, sob as condições constantes nas Tabelas 7 e 8.

4.2.2. Aplicação em Microscopia Confocal de Fluorescência

A Figura 79 mostra as imagens obtidas por microscópio confocal de fluorescência, das amostras de nanopartículas de sulfeto de cobre e alumínio, sintetizadas sob as 16 condições diferentes de síntese, pelo método *heat up*. As imagens mostram um potencial de aplicação do material como marcadores fluorescentes, aderidos em material biológico.

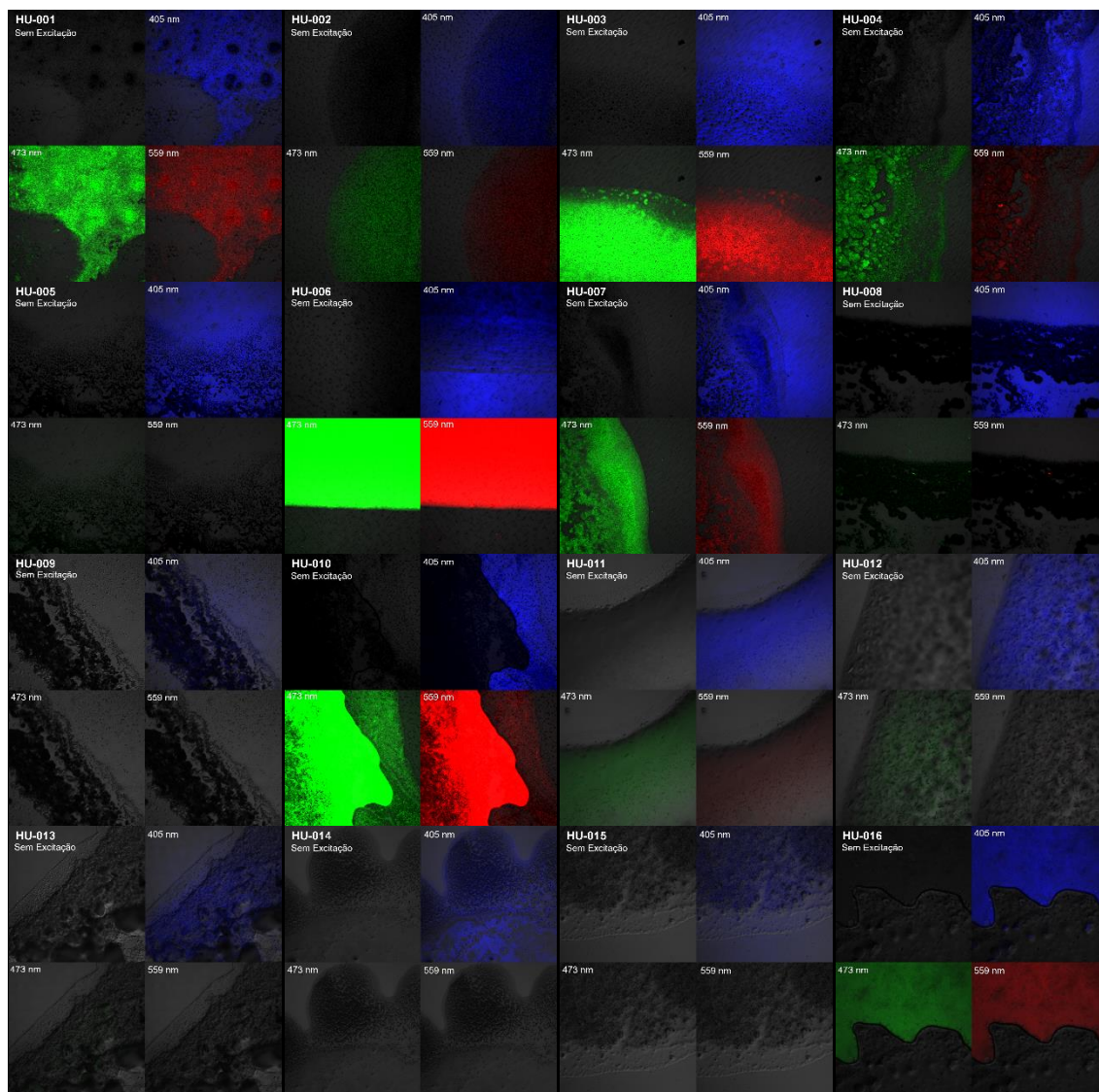


Figura 83: Imagens de microscopia confocal de fluorescência das nanopartículas sintetizadas pelo método *heat up*, depositadas sobre lâminas de vidro, sem excitação (superior esquerda), excitada por laser de 405 nm (nos quadros, na posição superior direita), 473 nm (inferior esquerda) e 559 nm (inferior direita).

5. CONCLUSÕES PARCIAIS

Nanopartículas de sulfeto de cobre e alumínio foram sintetizadas sob diferentes condições de síntese, pelo método *heat up* e injeção à quente. As imagens de microscopia eletrônica de transmissão revelaram uma miscelânea de tamanhos e formatos, de acordo com cada diferente condição de síntese. Análises de espectros Raman e padrões de difração de raios-x mostraram a formação de sulfeto de cobre e alumínio juntamente com subprodutos como covellite, atacamite, paratacamite e oxiclreto de alumínio. Análise dos espectros de absorção no UV-Vis-NIR e do gráfico de Tauc mostraram que as amostras apresentam tanto transição direta como transição indireta de *bandgap*. Os valores de *bandgap* direto são compatíveis com os dados contidos na literatura para sulfeto de cobre e alumínio, enquanto que os valores referentes ao *bandgap* indireto são compatíveis com dados para covellite. A análise dos parâmetros que influenciam na intensidade de emissão de fluorescência permite concluir que o principal parâmetro de síntese é a concentração do precursor de alumínio. Essa conclusão vai ao encontro de estudos teóricos anteriores, que demonstraram que existe um modo de fluorescência gerado por um defeito de substituição de alumínio em um sítio de cobre. As amostras foram aplicadas em microscopia confocal de fluorescência, demonstrando potencial para aplicações como marcadores biológicos.

Como perspectivas futuras, ficam as sugestões do uso de técnicas de precipitação seletiva, entre outras técnicas de separação, a fim de se obter nanopartículas de sulfeto de cobre e alumínio em alto grau de pureza. Adicionalmente, é recomendado que se faça o uso de uma estrutura do tipo core/shell como possível recurso para aumento na intensidade de fotoluminescência.

CAPÍTULO 5: CONCLUSÃO

Bases de dados de síntese *hot injection* de pontos quânticos de CdS, CdSe, PbS, PbSe e ZnSe foram criadas do zero, a partir de busca na literatura. Os algoritmos de *machine learning* foram aplicados a essas bases de dados para gerar modelos capazes de prever o tamanho médio final dos pontos quânticos e generalizar para demais condições de síntese. Através dos modelos, foi possível identificar que o tempo de reação, a temperatura de crescimento e a concentração do precursor metálico foram os fatores que mais influenciaram no crescimento das partículas. Em seguida, os modelos de *random forest*, *gradient boosting machine* e *deep learning* foram selecionados para produzir sugestões de sínteses para obtenção de pontos quânticos de CdSe com tamanhos entre 1,8 e 2,2 nm, 4,8 e 5,2 nm, e 7,8 e 8,2 nm. O método Monte Carlo foi usado para simular 5000 condições de reação com restrições baseadas na disponibilidade dos reagentes em nosso laboratório. Os modelos selecionados foram aqueles que demonstraram maior robustez a pequenas variações nas condições de síntese, como temperatura de injeção e tempo de reação. Finalmente, mapas de calor foram gerados, variando-se temperatura de injeção e tempo de reação, onde pode-se procurar a região com menor variação de diâmetro para realizar a síntese e reduzir a possibilidade de variação de tamanhos entre sínteses.

Na segunda parte do trabalho, foi realizado um planejamento experimental a fim de identificar os parâmetros que influenciam na intensidade de emissão de fotoluminescência na síntese de nanopartículas de sulfeto de cobre e alumínio. As amostras foram sintetizadas sob diferentes condições de síntese, pelo método *heat up*. A natureza do *bandgap* de menor energia para as amostras sintetizadas é de transição indireta, possivelmente relacionada com a fase Covellite, enquanto a transição direta, de maior energia, está relacionada com CuAlS₂. Relacionando as intensidades encontradas nas medidas de fotoluminescência, foi possível concluir que o principal parâmetro de síntese é a concentração do precursor de alumínio. Essa conclusão vai ao encontro de estudos teóricos anteriores, que demonstraram que existe um modo de fluorescência gerado por um defeito de substituição de alumínio em um sítio de cobre. As amostras foram aplicadas em microscopia confocal de fluorescência, demonstrando potencial para aplicações como marcadores biológicos.

REFERÊNCIAS

- 1 Kim W. D., Lee S., Pak C., Woo J. Y., Lee K., Baum F., Won J., Lee D. C. Metal tips on pyramid-shaped PbSe/CdSe/CdS heterostructure nanocrystal photocatalysts: study of Ostwald ripening and core/shell formation. *Chem Commun* 50(14):1719–1721, 2014.
- 2 Carey, G. H., Abdelhady, A. L., Ning, Z., Thon, S. M., Bakr, O. M., Sargent, E. H. Colloidal Quantum Dot Solar Cells, *Chem. Rev.*, 115, 12732-12763, 2015.
- 3 Bera, D., Qian, L., Tseng, T-K., Holloway, P. H. Quantum Dots and Their Multimodal Applications: A Review, *Materials*, 3, 2260-2345, 2010.
- 4 Dai, X.; Zhang, Z.; Jin, Y.; Niu, Y.; Cao, H.; Liang, X.; Chen, L.; Wang, J.; Peng, X. Solution-processed, high-performance light-emitting diodes based on quantum dots. *Nature*, v. 515, n. 1, p. 96 – 99, 2014.
- 5 Luo, Z., Xu D., Wu, S-T. Emerging Quantum-Dots-Enhanced LCDs. *Journal of Display Technology*, 10, 526–539, 2014.
- 6 National Nanotechnology Initiative, disponível em <https://www.nano.gov/timeline>, (acesso em 30/10/2018).
- 7 Bawendi, M. G., Murray, C. B., Norris, D. J. Synthesis and characterization of nearly monodisperse CdE (E = sulfur, selenium, tellurium) semiconductor nanocrystallites, *J. Am. Chem. Soc.*, 115 (19), 8706-8715, 1993.
- 8 <https://www.marketsandmarkets.com/Market-Reports/quantum-dots-qd-market-694.html>, accessed in 05/12/2020.
- 9 <https://www.futuremarketsinc.com/the-global-market-for-quantum-dots-2/>, accessed in 05/12/2020.
- 10 Braham, E. J., Cho, J., Forlano, K. M., Watson, D. F., Arròyave, R., Banerjee, S. Machine Learning-Directed Navigation of Synthetic Design Space: A Statistical Learning Approach to Controlling the Synthesis of Perovskite Halide Nanoplatelets in the Quantum-Confined Regime. *Chem. Mater.*, 31, 9, 3281–3292, 2019.
- 11 Dimiduk, D. M., Holm, E. A., Niezgoda, S. R. Perspectives on the Impact of Machine Learning, Deep Learning, and Artificial Intelligence on Materials, Processes, and Structures Engineering. *Integrating Materials and Manufacturing Innovation* vol. 7, 157–172, 2018.
- 12 Yang, W., Fidelis, T. T., Sun, W-H. Machine Learning in Catalysis, From Proposal to Practicing. *ACS Omega*, 5, 1, 83–88, 2020.
- 13 Mater, A. C. and Coote, M. L. Deep Learning in Chemistry. *J. Chem. Inf. Model.*, 59, 6, 2545–2559, 2019.

-
- 14 Cova, T. F. G. G. and Pais, A. A. C. C. Deep Learning for Deep Chemistry: Optimizing the Prediction of Chemical Patterns. *Front. Chem.*, 7:809, 2019.
- 15 Zhou, Z., Li, X., Zare R. N. Optimizing Chemical Reactions with Deep Reinforcement Learning. *ACS Cent. Sci.*, 3, 1337–1344, 2017.
- 16 Kim, E., Huang, K., Saunders, A., McCallum, A., Ceder, G., Olivetti, E. Materials Synthesis Insights from Scientific Literature via Text Extraction and Machine Learning. *Chem. Mater.*, 29, 21, 9436–9444, 2017
- 17 Jensen, Z., Kim, E., Kwon, S., Gani, T. Z. H., Román-Leshkov, Y., Moliner, M., Corma, A., Olivetti, E. A Machine Learning Approach to Zeolite Synthesis Enabled by Automatic Literature Data Extraction. *ACS Cent. Sci.*, 5, 5, 892–899, 2019.
- 18 Coley, C. W., Thomas III, D. A., Lummiss, J. A. M., Jaworski, J. N., Breen, C. P., Schultz, V., Hart, T., Fishman, J. S., Rogers, L., Gao, H., Hicklin, R. W., Plehiers, P. P., Byington, J., Piotti, J. S., Green, W. H., Hart, A. J., Jamison, T. F., Jensen, K. F. A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science*, Vol. 365, Issue 6453, eaax1566, 2019.
- 19 Voznyy, O., Levina, L., Fan, J. Z., Askerka, M., Jain, A., Choi, M-J., Ouellette, O., Todorović, P., Sagar, L. K., Sargent, E. H. Machine Learning Accelerates Discovery of Optimal Colloidal Quantum Dot Synthesis. *ACS Nano*, 13, 11122–11128, 2019.
- 20 Pang, H.; Meng, X.; Song, H.; Zhou, W.; Yang, G.; Zhang, H.; Izumi, Y.; Takei, T., Jewaswan, W.; Fukata, N.; Ye, J. Probing the role of nickel dopant in aqueous colloidal ZnS nanocrystals for efficient solar-driven CO₂ reduction. *Applied Catalysis B: Environmental*, v. 244, p. 1013–1020, 2019.
- 21 Wolf, M.; Fischer, N.; Claeys, M. Surfactant-free synthesis of monodisperse cobalt oxide nanoparticles of tunable size and oxidation state developed by factorial design. *Materials Chemistry and Physics*. v. 213, p. 305 – 312, 2018.
- 22 Fernandes, R. S.; Raimundo Jr, I. M.; Pimentel, M. F. Revising the synthesis of Stöber silica nanoparticles: A multivariate assessment study on the effects of reaction parameters on the particle size. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*. v. 577, p. 1 – 7, 2019.
- 23 Elazazy, M. S. ; Issa, A. A.; Al-Mashreky, M.; Al-Sulaiti, M.; Al-Saad, K. Application of fractional factorial design for green synthesis of cyano-modified silica nanoparticles: Chemometrics and multifarious response optimization. *Advanced Powder Technology*. v. 29, n. 5, p. 1204 – 1215, 2018.
- 24 Shah, M.; Pathak, K. Development and Statistical Optimization of Solid Lipid Nanoparticles of Simvastatin by Using 23 Full-Factorial Design. *AAPS PharmSciTech*, v. 11, n. 2, p. 489 – 496, 2010.

25 Núñez, R. N.; Veglia, A. V.; Pacioni, N. L. Improving reproducibility between batches of silver nanoparticles using an experimental design approach. *Microchemical Journal*, v. 141, p. 110-117, 2018.

26 Lima, M. M.; Macuvele, D. L. P.; Nones, J.; Silva, L. L.; Riella, H. G.; Fiori, M. A.; Soares, C. Synthesis of Fe₃O₄-Fe₂O₃@C Core-Shell Nanoparticles: Effect of Reactional Parameters on Structural and Magnetics Properties. *Journal of Inorganic and Organometallic Polymers and Materials*, 29, 1848-1861, 2019.

27 Castro, S. L., Bailey, S. G., Raffaele, R.P., Banger, K.K., Hepp, A.F. Nanocrystalline Chalcopyrite Materials (CuInS₂ and CuInSe₂) via Low-Temperature Pyrolysis of Molecular Single-Source Precursors. *Chem. Mater.*, 15, 3142-3147, 2003.

28 Baum, F. Dinâmica de crescimento de nanopartículas de sulfeto de cobre e antimônio, Dissertação de Mestrado, 96 pp., 2016.

29 Parlak, C. e Eryğit, R. Ab initio pressure-dependent vibrational and dielectric properties of chalcopyrite CuAlS₂. *Phys. Rev. B* 70, 075210, 2004.

30 Poulouse, A. C., Veerananarayanan, S., Aravind, A., Nagaoka, Y., Yoshida, Y., Maekawa, T., Kumar, D. S. Synthesis of CuAlS₂ Nanocrystals and Their Application in Bio-Imaging. *Mater. Express*, Vol. 2, No. 2, 94-104, 2012.

31 Bhattacharyya, B., Pandit, T., Rajasekar, G. P., Pandey, A. Optical Transparency Enabled by Anomalous Stokes Shift in Visible Light-Emitting CuAlS₂-Based Quantum Dots. *J. Phys. Chem. Lett.*, 9, 4451-4456, 2018.

32 Bhattacharyya, B., Simlandy, A. K., Chakraborty, A., Rajasekar, G. P., Aetukuri, N. B., Mukherjee, S., Pandey, P. Efficient Photosynthesis of Organics from Aqueous Bicarbonate Ions by Quantum Dots Using Visible Light. *ACS Energy Lett.*, 3, 1508-1514, 2018.

33 Strehlow, W. H. e Cook, E. L. Compilation of Energy Band Gaps in Elemental and Binary Compound Semiconductors and Insulators, *J. Phys. Chem.*, 2, 1, 163, 1973.

34 Callister, W. D. J. Ciência e Engenharia de Materiais: Uma Introdução, 5ª ed., Rio de Janeiro, ed. LTC, 2002.

35 Kittel, C. Introdução à Física do Estado Sólido, 5ª ed., Rio de Janeiro, ed. Guanabara Dois S. A., 1978.

36 International Organization for Standardization, Nanotechnologies -- Vocabulary -- Part 2: Nano-objects, ISO/TS 80004-2:2015.

37 Schiavon, M. A.; Silva, F. O.; Viol, L. C. S.; Ferreira, D. L.; Alves, J. L. A. O Estado da Arte da Síntese de Semicondutores Nanocristalinos Coloidais, *Quim. Nova*, Vol. 33, 9, 1933-1939, 2010.

38 Medintz, I. L., Uyeda H. T., Goldman, E. R., Mattoussi, H. Quantum dot bioconjugates for imaging, labelling and sensing. *Nature Materials*, 4, 435-446, 2005.

39 Pak, C., Woo, J. Y., Lee, K., Kim, W. D., Yoo, Y., Lee, D. C. Extending the Limit of Low-Energy Photocatalysis: Dye Reduction with PbSe/CdSe/CdS Core/Shell/Shell Nanocrystals of Varying Morphologies under Infrared Irradiation, *J. Phys. Chem. C*, 116, 25407-25414, 2012.

40 Wu, P., Hou, X., Xu, J-J., Chen, H-Y. Ratiometric fluorescence, electrochemiluminescence, and photoelectrochemical chemo/ biosensing based on semiconductor quantum dots. *Nanoscale*, 8, 8427-8442, 2016.

41 Oh, N., Kim, B. H., Cho, S-Y., Nam, S., Rogers, S. P., Jiang, Y., Flanagan, J. C., Zhai, Y., Kim, J-H., Lee, J., Yu, Y., Cho, Y. K., Hur, G., Zhang, J., Trefonas, P., Rogers, J. A., Shim, M. Double-heterojunction nanorod light-responsive LEDs for display applications. *Science*, 355, 616–619, 2017.

42 Carey, G. H., Abdelhady, A. L., Ning, Z., Thon, S. M., Bakr, O. M., Sargent, E. H. Colloidal Quantum Dot Solar Cells. *Chem. Rev.*, 115, 23, 12732–12763, 2015.

43 IEEE Spectrum, Your Guide to Television's Quantum Dot Future, disponível em <https://spectrum.ieee.org/consumer-electronics/audiovideo/your-guide-to-televisions-quantumdot-future>, acessado em 20/10/2018.

44 Techradar, Quantum of Promise: Acer packs Quantum Dot screen into new Android tablet, <https://www.techradar.com/news/quantum-of-promise-acer-packs-quantum-dot-screen-into-new-android-tablet>, acessado em 20/10/2018.

45 Supran, G. J., Shirasaki, Y., Song, K. W., Caruge, J.-M., Kazlas, P. T., Coe-Sullivan, S., Andrew, T. L., Bawendi, M. G., Bulović, V. QLEDs for displays and solid-state lighting. *MRS Bulletin*, 38(09), 703–711, 2013.

46 Ugarte, I., Castelló, I., Palomares, E., Pacios, R. Quantum Dots as Light Indicators for Emitting Diodes and Biological Coding, in *Quantum Dots – A Variety of New Applications*, Dr. Ameenah Al-Ahmadi (Ed.), ISBN 978-953-51-0483-4, 2012.

47 Wood, V. E Bukovic, V. Colloidal quantum dot light-emitting devices, *Nano Reviews*, 1, 5202-5208, 2010.

48 Jun, H. K., Careem, M. A., Arof, A. K. Quantum dot-sensitized solar cells – perspective and recent developments: A review of Cd chalcogenide quantum dots as sensitizers, *Renewable and Sustainable Energy Reviews*, 22, 148-167, 2013.

49 Kamat, P. V. Quantum Dot Solar Cells. Semiconductor Nanocrystals as Light Harvesters, *J. Phys. Chem.*, 112, 18737-18753, 2008.

50 Kamat, P V. Quantum Dot Solar Cells. The Next Big Thing in Photovoltaics, *J. Phys. Chem. Lett.*, 4, 903-918, 2013.

51 Bera, D., Qian, L., Tseng, T-K., Holloway, P. H. Quantum Dots and Their Multimodal Applications: A Review, *Materials*, 3, 2260-2345, 2010.

52 Ishikawa, M., Vasudevanpillai, B., Makita, Y., Sonoda, A., Yokoyama, H., Baba, Y. Temperature-Sensitive Photoluminescence of CdSe Quantum Dot Clusters, *J. Phys. Chem. B*, 109, 13899-13905, 2005.

53 Ijima, Y., Sakaue, H., Aikawa, A. Anodized-aluminum as quantum dot support for global temperature sensing from 100 to 500 K, *Sensors and Actuators B*, 150, 569-573, 2010.

54 Yang, J. M., Yang, H., Lin, L. Quantum Dot Nano Thermometers Reveal Heterogeneous Local Thermogenesis in Living Cells, *ACS Nano*, v. 5, No. 6, 5067-5071, 2011.

55 Parak, W. J., Aberasturi, D. J., Montenegro, J. M., Larramendi, I. R., Rojo, T., Klar, T. A., Alvarez-Puebla, R., Liz-Marzán, L. M. Optical Sensing of Small Ions with Colloidal Nanoparticles, *Chem. Mater.*, 24, 738-745, 2012.

56 Chen, Y. E. Rosenzweig, Z. Luminescent CdS Quantum Dots as Selective Ion Probes, *Anal. Chem.*, 74, 5132-5138, 2002.

57 Xu, C. E. Bakker, E. Multicolor Quantum Dot Encoding for Polymeric Particle-Based Optical Ion Sensors, *Anal. Chem.*, 79, 3716-3723, 2007.

58 Zhang, S. Ge, S., Zhang, C., Zhu, Y., Yu, J. BSA activated CdTe quantum dot nanosensor for antimony ion detection, *Analyst*, 135, 111-115, 2010.

59 Snee, P. T., Page, L. E., Zhang, X., Jawaid, A. M. Detection of toxic mercury ions using a ratiometric CdSe/ZnS nanocrystal sensor, *Chem. Commun.*, 47, 7773-7775, 2011.

60 Zhang, Z., Tu, R., Liu, B., Wang, Z., Gao, D., Wang, F., Fang, Q. Amine-Capped ZnS-Mn₂ Nanocrystals for Fluorescence Detection of Trace TNT Explosive, *Anal. Chem.*, 80, 3458-3465, 2008.

61 Fu, D., Wang, Y., Zheng, J., Zhang, Z., Yuan, C. CdTe nanocrystals as luminescent probes for detecting ATP, folic acid and L-cysteine in aqueous solution, *Colloids and Surfaces A: Physicochem. Eng. Aspects*, 342, 102-106, 2009.

62 Jiang, B., Chen, Z., Chen, J., Liang, Q., Wu, D., Zeng, Y. ZnSe quantum dots based fluorescence quenching method for determination of paeoniflorin, *Journal of Luminescence*, 145, 569-574, 2014.

63 Rogach, A. L., Susha, A. S., Javier, A. M., Parak, W. J. Luminescent CdTe nanocrystals as ion probes and pH sensors in aqueous solutions, *Colloids and Surfaces A: Physicochem. Eng. Aspects*, 281, 40-43, 2006.

64 Nocera, D. G., Snee, P. T., Somers, R. C., Nair, G., Zimmer, J. P., Bawendi, M. G. A Ratiometric CdSe/ZnS Nanocrystal pH Sensor, *J. Am. Chem. Soc.*, 128, 13320-13321, 2006.

65 Cao, S., Yu, D., Wang, Z., Liu, Y., Jin, L., Cheng, Y., Zhou, J. Quantum dot-based pH probe for quick study of enzyme reaction kinetics, *Enzyme and Microbial Technology*, 41, 127-132, 2007.

66 Chen, Z e Wu, D. Colloidal ZnSe quantum dots as pH probes for study of enzyme reaction kinetics by fluorescence spectroscopic technique, *Colloids and Surfaces A: Physicochem. Eng. Aspects*, 414, 174-179, 2012.

67 Azzazy, H. M. E., Mansour, M. M. H., Kazmierczak, S. C. From diagnostics to therapy: Prospects of quantum dots, *Clinical Biochemistry*, 40, 917-927, 2007.

68 Alivisatos, A. P., Gu, W., Larabell, C. Quantum Dots as Cellular Probes, *Annu. Rev. Biomed. Eng.*, 7, 55-76, 2005.

69 Drbohlavova, J., Adam, V, Kizek, R., Hubalek, J. Quantum Dots – Characterization, Preparation and Usage in Biological Systems, *Int. J. Mol. Sci.*, 10, 656-673, 2009.

70 Jamieson, T., Bakhshi, R., Petrova, D., Pocock, R., Imani, M., Seifalian, A. M. Biological applications of quantum dots, *Biomaterials*, 28, 4717-4732, 2007.

71 Hoffmann, M. R., Martin, S. T., Choi, W., Bahnemann, D. W. Environmental Applications of Semiconductor Photocatalysis, *Chemical Reviews*, 95, 69-96, 1995.

72 Kubacka, A., Fernández-García, M., Colón, G. Advanced Nanoarchitectures for Solar Photocatalytic Applications, *Chemical Reviews*, 112, 1555-1614, 2012.

73 Wilker, M. B., Schnitzenbaumer, K. J., Dukovic, G. Recent Progress in Photocatalysis Mediated by Colloidal II-VI Nanocrystals, *Isr. J. Chem.* 52, 1002-1015, 2012.

74 Zhao, J., Holmes, M. A., Osterloh, F. E. Quantum Confinement Controls Photocatalysis: A Free Energy Analysis for Photocatalytic Proton Reduction at CdSe Nanocrystals, *ACS Nano*, 7, No. 5, 4316-4325, 2013.

75 Coughlan, C., Ibáñez, M., Dobrozhan, O., Singh, A., Cabot, A., Ryan, K. M. Compound Copper Chalcogenide Nanocrystals, *Chem. Rev.*, 117, 9, 5865–6109, 2017.

76 Yoshino, K., Nomoto, K., Kinoshita, A., Ikari, T., Akaki, Y., Yoshitake, T. Dependence of Cu/In Ratio of Structural and Electrical Characterization of CuInS₂ Crystal. *J. Mater. Sci.: Mater. Electron.*, 19, 301, 2008.

77 Chen, B.; Chang, S.; Li, D.; Chen, L.; Wang, Y.; Chen, T.; Zou, B.; Zhong, H.; Rogach, A. L. Template Synthesis of CuInS₂ Nanocrystals from In₂S₃ Nanoplates and Their Application as Counter Electrodes in Dye-Sensitized Solar Cells. *Chem. Mater.* 2015, 27, 5949.

78 Chen, B.; Zhong, H.; Zhang, W.; Tan, Z. a.; Li, Y.; Yu, C.; Zhai, T.; Bando, Y.; Yang, S.; Zou, B. Highly Emissive and Color-Tunable CuInS₂-Based Colloidal

Semiconductor Nanocrystals: Off-Stoichiometry Effects and Improved Electroluminescence Performance. *Adv. Funct. Mater.*, 22, 2081, 2012.

79 Binsma, J. J. M.; Giling, L. J.; Bloem, J. Luminescence of CuInS₂: I. The Broad Band Emission and Its Dependence on the Defect Chemistry. *J. Lumin.* 1982, 27, 35.

80 Dai, X., Lei, H., Chen, C., Guo, Y., Fang, G. A simple synthesis of transparent and highly conducting p-type Cu_xAl_{1-x}S_y nanocomposite thin films as the hole transporting layer for organic solar cells. *RSC Adv.*, 8, 16887-16896, 2018.

81 Chaki, S. H., Mahato, K. S., Deshpande, M. P. Catalytic action of CuAlS₂ microparticles and nanoparticles in cellulose pyrolysis. *Phys. Scr.* 90, 045701, 2015.

82 Chichibu, S., Shirakata, S., Isomura, S., Nakanishi, H. Visible and Ultraviolet Photoluminescence from Cu-III-VI₂ Chalcopyrite Semiconductors Grown by Metalorganic Vapor Phase Epitaxy. *Jpn. J. Appl. Phys.* Vol. 36, 1703-1714, 1997.

83 Moreh, A. U., Momoh, M., Hamza, B. Influence Of Substrate Temperature on Optical Properties of Nanostructured CuAlS₂ Thin Films Grown By Two Stage Vacuum Thermal Evaporation Technique. *International Journal of Engineering Science Invention*, 2, 48-52, 2013.

84 Brini, R., Schmerber, G., Kanzari, M., Werckmann, J., Rezig, B. Study of the growth of CuAlS₂ thin films on oriented silicon (111). *Thin Solid Films* 517, 2191–2194, 2009.

85 Yue, G. H., Wang, X., Wang, L. S., Wang, W., Peng, D. L. Synthesis of single crystal CuAlS₂ nanowires via a low temperature direct polyol route. *Physics Letters A*, 372, 5995–5998, 2008.

86 Sukan, S., Baskar, K., Dhanasekaran, R. Structural, morphological and optical studies on CuAlS₂ and CuAlSe₂ nanorods prepared by hydrothermal method. *Journal of Alloys and Compounds*, 645, 85-89, 2015.

87 Domingos, P. A Few Useful Things to Know about Machine Learning. *Communications of the ACM*, Vol. 55, No. 10, 78-87, 2012.

88 James, G., Witten, D., Hastie, T., Tibshirani, R.: Unsupervised learning. In: *An Introduction to Statistical Learning*, Springer, New York (2013).

89 Géron, A. *Hands-On Machine Learning with Scikit-Learn and TensorFlow*, O'Reilly Media, Inc., Sebastopol (2017).

90 Chollet, F., Allaire, J. J. *Deep Learning with R*, Manning Publications, New York (2018).

91 Minaee, S. 20 Popular Machine Learning Metrics. Part 1: Classification & Regression Evaluation Metrics. 2019. Disponível em <https://towardsdatascience.com/20-popular-machine-learning-metrics-part-1-classification-regression-evaluation-metrics-1ca3e282a2ce>

92 Hale, J. Which Evaluation Metric Should You Use in Machine Learning Regression Problems? 2020. Disponível em <https://towardsdatascience.com/which-evaluation-metric-should-you-use-in-machine-learning-regression-problems-20cdaef258e>

93 Draper, N. R. e Smith, H. Applied Regression Analysis, 3ª Ed. John Wiley & Sons, Inc, 1998.

94 The R Foundation. R Project, disponível em <https://www.r-project.org/about.html>

95 Baum, F., Silva, M. F., Linden, G., Feijo, D., Rieder, E. S., Santos, M. J. L. Growth dynamics of zinc selenide quantum dots: the role of oleic acid concentration and synthesis temperature on driving optical properties. *J. Nanopart. Res.* 21, 42, 2019.

96 Thanh, N. T. K., Maclean, N., Mahiddine, S. Mechanisms of nucleation and growth of nanoparticles in solution. *Chemical Reviews*, 114, 7610–7630, 2014.

97 Newton, J. C., Ramasamy, K., Mandal, M., Joshi, G. K., Kumbhar, A., Sardar, R. Low-Temperature Synthesis of Magic-Sized CdSe Nanoclusters: Influence of Ligands on Nanocluster Growth and Photophysical Properties *J. Phys. Chem. C*, 116, 7, 4380–4389, 2012.

98 Yu, Q. e Liu, C-Y. Study of Magic-Size-Cluster Mediated Formation of CdS Nanocrystals: Properties of the Magic-Size Clusters and Mechanism Implication. *J. Phys. Chem. C*, 113, 12766-12771, 2009.

99 Wise, F. W. Lead Salt Quantum Dots: the Limit of Strong Quantum Confinement, *Acc. Chem. Res.*, 33, 773-780, 2000.

100 <https://www.h2o.ai/pt/> acessado em 01/11/2020.

101 Baum, F., Pretto, T., Köche, A., Santos, M. J. L. Machine Learning Tools to Predict Hot Injection Syntheses Outcomes for II–VI and IV–VI Quantum Dots. <https://doi.org/10.1021/acs.jpcc.0c05993>

102 Altomare, A., Corriero, N., Cuocci, C., Falcicchio, A., Moliterni, A., Rizzi, R. QUALX2.0: a qualitative phase analysis software using the freely available database POW_COD, *J. Appl. Cryst.* 48, 598-603, 2015.

¹⁰³ Grazulis, S., Chateigner, D., Downs, R.T., Yokochi, A.F.T., Quirós, M., Lutterotti, L., Manakova, E., Butkus, J., Moeck, P., Le Bail, A. Crystallography Open Database – an open-access collection of crystal structures. *J. Appl. Cryst.* 42, 726-729, 2009.

104 van der Stam, W., Gradmann, S., Altantzis, T., Ke, X., Baldus, M., Bals, S., Donega, C.M. Shape Control of Colloidal Cu_{2-x}S Polyhedral Nanocrystals by Tuning the Nucleation Rates. *Chem. Mater.*, 28, 6705–6715, 2016.

105 Kuroki Y., Osada M., Kato A., Okamoto T., Takata M. Exciton-Phonon Interaction in CuAlS₂ Powders. AMR, 11–12:175–178, 2006.

106 Brini, R., Schmerber, G., Kanzari, M., Rezig, B., Werckmann, J. Growth and optimization by post-annealing of chalcopyrite CuAlS₂ compound. Eur. Phys. J. Appl. Phys. 36, 11-15, 2006.

107 Bairamov, B. H. e Aydinli, A. High power gain for stimulated Raman amplification in CuAlS₂. Journal of Applied Physics 80, 5564, 1996.

108 Roa, L., González, J., Chervin, J. C., Chevy, A. High Pressure Raman Scattering Study of CuAlS₂. phys. stat. sol. (b) 211, 429, 1999.

109 Kushwaha, A. K., Khenata, R., Bouhemadou, A., Bin-Omran, S., Haddadi, K. Lattice Dynamical Properties and Elastic Constants of the Ternary Chalcopyrite Compounds CuAlS₂, CuGaS₂, CuInS₂, and AgGaS₂. Journal of Electronic Materials, vol. 46, 4109–4118, 2017.

110 Koschel, W.H., Hohler, V., Räuber, A., Baars, J. Optical phonons in CuAlS₂. Solid State Communications Vol. 13, 7, 1011-1016, 1973.

111 <https://rruff.info/covellite/R060129> acessado em 29/07/2020.

112 van der Stam, W., Gudjonsdottir, S., Evers, W. H., Houtepen, A. J. Switching between Plasmonic and Fluorescent Copper Sulfide Nanocrystals. J. Am. Chem. Soc., 139, 37, 13208–13217, 2017.

113 Xie, Y., Bertoni, G., Riedinger, A., Sathya, A., Prato, M., Marras, S., Tu, R., Pellegrino, T., Manna, L. Nanoscale Transformations in Covellite (CuS) Nanocrystals in the Presence of Divalent Metal Cations in a Mild Reducing Environment. Chem. Mater., 27, 21, 7531–7537 2015.

114 Luther, J. M., Jain, P. K., Ewers, T., Alivisatos, A. P. Localized surface plasmon resonances arising from free carriers in doped quantum dots. Nature Materials 2011, 10, 361-366.

115 Zhao, Y. and Burd, C. Development of plasmonic semiconductor nanomaterials with copper chalcogenides for a future with sustainable energy materials. Energy Environ. Sci., 2012, 5, 5564–5576.

116 Baum, F., Preto, T., Brolo, A.G., Santos, M.J.L. Uncovering the Mechanism for the Formation of Copper Thioantimonate (SbV) Nanoparticles and Its Transition to Thioantimonide (SbIII), Cryst. Growth Des., 18,11, 6521-6527, 2018.

117 Niezgoda, J. S., Yap, E., Keene, J. D., McBride, J. R., Rosenthal S. J. Plasmonic Cu_xIn_yS₂ Quantum Dots Make Better Photovoltaics Than Their Nonplasmonic Counterparts. Nano Lett., 14, 3262–3269, 2014.

118 Harichandran, G., Lalla, N.P. Facile synthesis of CuAlS₂ nanorods. Materials Letters, 62, 1267–1269, 2008.

119 Liu, X., Xua, L., Huang, Y., Cheng, H., Seo, H. J. Paratacamite phase stability and improved optical properties of $\text{Cu}_2(\text{OH})_3\text{Cl}$ crystal via Ni-doping. *Materials and Design*, 121, 194–201, 2017.

120 Achimovičová, M., Dutková, E., Tóthová, E., Bujňáková, Z., Briančin, J., Kitazono, S. Structural and optical properties of nanostructured copper sulfide semiconductor synthesized in an industrial mill. *Front. Chem. Sci. Eng.*, 13(1): 164–170, 2019.

121 Kalanura, S. S. e Seo, H. Tuning plasmonic properties of CuS thin films via valence band filling, *RSC Adv.*, 7, 11118, 2017.

122 Chaki, S. H., Deshpande, M. P., Tailor, J. P. Characterization of CuS nanocrystalline thin films synthesized by chemical bath deposition and dip coating techniques *Thin Solid Films*, 550, 291–297, 2014.

123 Liborio, L. M., Bailey, C. L., Mallia, G., Tomić, S., Harrison, N. M. Chemistry of defect induced photoluminescence in chalcopyrites: The case of CuAlS_2 . *Journal of Applied Physics*, 109, 023519, 2011.

124 Miyake, H., Yamada, M., Sugiyama, K. Vapor phase epitaxy of CuAlS_2 on CuGaS_2 substrates by the iodine transport method. *Journal of Crystal Growth*, 153, 180-183, 1995.

125 Iida, S. e Honda, H. Photoluminescence of as-grown CuAlS_2 prepared by iodine transport method. *Il Nuovo Cimento D* vol. 2, 1939–1943, 1983.

126 Shirakata, S., Aksenov, I., Sato, K., Isomura, S. Photoluminescence Studies in CuAlS_2 Crystals. *Jpn. J. Appl. Phys.* Vol. 31, 1071-1074, 1992.

127 Syrbu, N. N., Korzun, B. V., Fadzeyeva, A. A., Mianzelen, R. R., Ursaki, V. V., Galbi, I. Exciton spectra and energy band structure of CuAlS_2 crystals. *Physica B*, 405, 3243–3247, 2010.

128 Kuroki, Y., Okamoto, T., Takata, M. Impacts of intrinsic defects on luminescence properties of CuAlS_2 . *Applied Physics Letters* 89, 221117, 2006.

ANEXO 1

Sumário das bases de dados usadas nos algoritmos de *machine learning*

Variável	CdS	CdSe	PbS	PbSe	ZnSe
Artigos	21	37	25	18	28
Sínteses	161	357	114	84	121
T _{inj} (°C)	150 a 300	20 a 330	30 a 210	30 a 220	110 a 330
T _{grw} (°C)	150 a 260	45 a 330	5 a 220	30 a 220	110 a 320
t (min)	0,17 a 360	0,003 a 4180	0,08 a 60	0,1 a 200	0,17 a 900
M	Estearato de cádmio Nitrato de cádmio tetraidratado Cloreto de cádmio Óxido de cádmio	Acetato de cádmio Acetato de cádmio diidratado Óxido de cádmio Estearato de cádmio Dimetil-cádmio	Acetato de chumbo triidratado Cloreto de chumbo Estearato de chumbo Óxido de chumbo	Acetato de chumbo triidratado Cloreto de chumbo Óxido de chumbo Estearato de chumbo	Óxido de zinco Nonanoato de zinco Acetato de zinco heptaidratado Estearato de zinco Carbonato de zinco Dietil-zinco Cloreto de zinco Acetato de zinco
M _g	0,013 a 1,54	0,01 a 1,9	0,088 a 18	0,077 a 2,5	0,01 a 2,53
M _{mmol}	0,1 a 11,99	0,086 a 13,4	0,36 a 80,7	0,1 a 11,2	0,016 a 7,03
[M]	0,004 a 1,15	0,007 a 1,72	0,014 a 1,57	0,016 a 0,31	0,002 a 0,48
C	Enxofre Tioacetamida Tiouréia	Selênio	Sulfeto de bis(trimetilsilil) Difeniltiouréia Enxofre Tioacetamida	Seleneto de bis(trimetilsilil) Selênio	Ciclohexeno-1,2,3-selenodiazol Complexo Se-octilamina-oleilamina Selênio Dióxido de selênio
C _g	0,002 a 0,25	0,004 a 0,79	0,006 a 7,1	0,008 a 1,5	0,0016 a 1,0
C _{mmol}	0,05 a 7,813	0,05 a 9,99	0,056 a 39,8	0,1 a 19	0,02 a 12,66
[C]	0,004 a 0,48	0,004 a 12,4	0,008 a 0,58	0,016 a 0,67	0,002 a 1,04
M / C	0,2 a 20	0,1 a 6,0	0,98 a 22,12	0,125 a 7,91	0,77 a 32,3

Variável	CdS	CdSe	PbS	PbSe	ZnSe
CA	Sem ligante Ácido láurico Ácido oleico Ácido esteárico	Sem ligante Ácido oleico Ácido esteárico Ácido benzóico Ácido mirístico	Sem ligante Ácido beénico Ácido erúcico Ácido oleico	Sem ligante Ácido ascórbico Ácido oleico	Sem ligante Ácido decanóico Ácido nonanóico Ácido octadecilfosfónico Ácido oleico
CA _g	0 a 18,869	0 a 16,11	0 a 62	0 a 7,5	0 a 15,1
CA _{mmol}	0 a 24	0 a 57,03	0 a 219,5	0 a 26,55	0,69 a 53,5
M / CA	0 a 0,43	0 a 3,33	0,01 a 3,95	0 a 6,32	0,086 a 0,50
A	Sem amina Dodecilamina Oleilamina Trietanolamina	Sem amina Hexadecilamina Dodecilamina Dioctilamina Oleilamina Octadecilamina	Sem amina Oleilamina	Sem amina Oleilamina	Sem amina Decilamina Hexadecilamina Octadecilamina Oleilamina
A _g	0 a 27,4	0 a 10	0 a 22,8	0 a 12,51	0 a 10,16
P	Sem fosfina Trioctilfosfina	Sem fosfina Difenilfosfina Tributilfosfina Trioctilfosfina Trifenilfosfina	Sem fosfina Trioctilfosfina	Sem fosfina Tributilfosfina Trioctilfosfina	Sem fosfina Trioctilfosfina
P _g	0 a 0,831	0 a 20,78	0 a 1,5	0 a 16,62	0 a 9,0
P _{mmol}	0 a 2,242	0 a 56,1	0 a 4,05	0 a 44,85	2,42 a 24,3
C / P	0,04 a 6,97	0 a 4,264	0,2215	0 a 8,95	0,007 a 0,34
S ¹	Sem solvente 1 Diesel Etilenoglicol N-oleoylmorpholine Octadeceno Parafina líquida Ácido succínico	Sem solvente 1 Parafina líquida Octadeceno Difenil-éter Óxido de trioctilfosfina	Sem solvente 1 Octadeceno	Sem solvente 1 Difenil-éter Hexadecano Octadeceno Óxido de trioctilfosfina	Sem solvente 1 Difenil-éter Octadeceno Parafina líquida
S ¹ _g	0 a 66,6	0 a 101,4	0 a 891	0 a 23,7	0 a 79,0
S ²	Sem solvente 2 1-tetradecanol N-oleoylmorpholine Polivinilpirrolidona	Sem solvente 2 Phosphinic acid Óxido de tributilfosfina Óxido de trioctilfosfina	Sem solvente 2	Sem solvente 2	Sem solvente 2 Parafina líquida
S ² _g	0 a 5,36	0 a 28,31	0	0	0 a 23,75
T(g)	4 a 682,84	2,48 a 108,0	1,93 a 69,1	5,2 a 46,6	2,65 a 104,0
D (nm)	1,23 a 12	0,85 a 11,5	1,57 a 80	1,1 a 116,2	2,0 a 9,6

ANEXO 2

Tabelas de associação de picos de difração de raios-x

HU-001		Atacamite (900-7718)		Covellite (900-0523)		CuAlS ₂ (900-9930)	
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
10,85	204,6			10,82	31,5		
16,15	958,3	16,15	790,3				
17,75	450,6	17,62	564,1				
27,45	346,7			27,67	207,9		
29,30	703,9			29,27	497,3	29,18	868,5
31,55	650,0	31,53	470,2	31,78	827,3		
32,25	741,6	32,18	369,2				
35,50	256,5	35,53	4,6	35,02	16,4		
39,70	329,5	39,76	393,7				
48,05	1000,0	48,18	1,5	47,92	898,9	48,22	163,8
52,60	207,3	52,36	70,2	52,72	76,7	52,14	16,7
57,40	206,2	57,31	8,2	57,15	57,0	57,36	191,5
59,30	293,3	59,34	57,3	59,32	161,6		
68,05	203,6	67,32	55,2	67,31	1,8		

HU-002		CuS ₂ (210-1636)					
2θ	Int. relativa	2θ	Int. relativa				
11,50	478,9						
23,20	671,4						
27,10	922,7	27,04	195,7				
31,15	1000,0	31,31	407,6				
35,10	467,4	35,12	59,2				
39,15	407,7						
45,85	429,4						
47,20	485,2						
47,60	727,6	47,75	0,3				
50,35	438,8	50,51	1,5				
58,20	483,2	58,22	0,8				
70,85	441,8						

HU-003		Atacamite (900-7718)		Covellite (900-0523)		CuAlS ₂ (154-2201)	
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
13.40	63.6						
16.15	1000.0	16.15	714.6				
17.55	374.2	17.62	510.0				
27.30	220.5			27.12	78.4		
27.65	213.8	27.73	26.4	27.67	119.3		
29.30	347.4	29.60	14.1	29.27	285.4	29.29	444.3
31.50	444.2	31.53	425.2	31.78	474.8		
32.25	662.2	32.18	333.8				
35.70	107.2	35.67	142.7				
39.60	370.4	39.49	252.5				
41.60	167.7						
46.70	137.7	46.87	0.1			46.79	9.2
48.05	488.7	48.18	1.3	47.92	515.9	48.43	83.1
49.80	156.1	49.83	44.4				
52.90	128.8			52.72	44.0		
53.25	188.1	53.34	102.0	53.19	14.7		
57.35	121.1	57.31	7.4	57.15	32.7	57.60	97.2
59.40	211.2	59.34	51.8	59.32	92.8	58.90	5.5
62.85	136.8	63.02	0.0	63.31	1.7		
64.10	121.5	64.15	32.4			64.00	0.0
65.55	101.2	65.51	32.0				
67.65	163.0	67.84	0.2	67.31	1.0		
77.00	127.1	77.03	24.5	76.95	0.0		
79.30	118.5	79.46	0.0	79.02	51.2	79.28	30.5

HU-004		Atacamite (900-7718)					
2θ	Int. relativa	2θ	Int. relativa				
16.15	-	16.15	714.6				
17.55	-	17.62	510.0				
31.50	-	31.53	425.2				
32.25	-	32.18	333.8				

HU-005		Atacamite (900-7718)		Covellite (900-0523)		CuAlS ₂ (154-2201)	
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
29.30	-	29.60	14.1	29.27	285.4	29.29	444.3
31.50	-	31.53	425.2	31.78	474.8		
32.25	-	32.18	333.8				
48.05	-	48.18	1.3	47.92	515.9	48.43	83.1
52.90	-			52.72	44.0		

HU-007		Atacamite (900-7718)		Covellite (900-9723)		CuAIS ₂ (154-2201)	
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
10.90	501.1			10.79	48.3		
14.45	169.2						
15.75	1000.0	16.15	714.6				
17.25	516.4	17.62	510.0				
28.95	319.1	29.60	14.1	29.25	342.6	29.29	444.3
31.40	377.4	31.53	425.2				
31.90	751.9			31.75	755.1		
32.35	675.5	32.18	333.8	32.77	338.5		
39.35	352.5	39.49	252.5				
43.90	193.8			44.19	33.6		
47.80	355.1			47.89	501.3		
52.55	262.8			52.61	198.1		
58.75	224.8			58.60	101.7	58.90	5.5
59.15	255.9	59.34	51.8	59.24	271.9		

HU-010		Covellite (900-0523)		Atacamite (900-7426)		CuAIS ₂ (154-2201)	
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
11.30	519.8	10.82	31.2				
16.10	886.3			16.18	1006.7		
17.55	493.1			17.65	673.9		
23.75	246.2						
27.80	415.7	27.67	205.8	27.77	33.7		
29.35	639.7	29.27	492.3	29.71	6.4	29.287	1000.00
31.80	733.2	31.78	819.0	31.63	609.5		
32.25	842.9			32.17	433.5		
32.70	969.6	32.86	176.8	32.69	221.5		
39.60	420.7			39.45	300.0		
48.00	745.2	47.92	889.8	47.87	0.4	48.427	187.00
49.95	266.5			49.94	58.6		
52.75	262.8	52.72	76.0	52.36	0.2		
57.10	184.6	57.15	56.4	56.91	3.5	57.604	218.80
59.45	309.2	59.32	160.0	59.35	69.8		
63.40	159.8	63.31	2.9	63.35	12.4	63.744	5.64

HU-011		AlOCl (900-9194)		CuCl ₂ (H ₂ O) ₂ (152-7348)		Paratacamite (900-7604)	
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
11.250	902.57						
11.500	885.37	11.53	939.0				
13.600	250.87						
14.800	373.10						
16.250	296.84			16.17	433.3	16.25	366.9
18.700	409.30					18.95	56.1
19.950	964.58					19.61	11.5
22.400	331.92			21.92	205.4		
25.050	878.15						
25.650	369.42						
26.500	365.27			26.56	61.6	26.08	20.1
27.550	607.28	27.28	59.7				
30.150	419.89						
32.650	327.47			32.67	13.1	32.85	74.2
34.200	220.81	34.08	341.9	33.90	279.8		
35.350	300.77	35.08	54.9	35.32	55.8	35.36	1.8
37.900	281.45			37.94	31.4		
39.900	183.24					39.83	209.3
40.450	232.59			40.73	103.6	40.86	16.1
43.050	186.00	43.29	1.2	42.96	48.8		
75.350	236.07	75.32	13.2	75.65	5.0	75.01	0.2

HU-012		CuAIS ₂ (900-9930)		Atacamite (900-7426)		AIOCI (900-9194)	
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
11.25	1000.0					11.53	762.6
13.50	130.4						
15.75	160.7			16.18	93.5		
17.45	23.7			17.65	62.6		
20.25	54.1						
20.80	40.6						
24.75	50.8			24.47	0.2		
27.15	52.0					27.22	301.3
29.40	87.6	29.18	87.5	29.71	0.6		
31.70	75.7			31.63	56.6		
34.05	93.9	34.32	1.3	34.01	9.7	34.03	9.4
36.45	68.5			36.03	1.2		
38.80	71.9	38.69	6.8				
39.85	23.4			39.90	43.9		
41.15	49.8			41.16	8.5		
43.60	40.2			43.56	0.3	43.33	167.0
46.00	32.6	46.09	2.6	46.36	9.0		
48.40	29.0	48.22	16.5	48.34	0.4		
49.60	23.3			49.65	0.0		
57.80	22.1	58.22	0.3	57.51	0.6	58.23	3.9

HU-016		AIOCI (900-9194)					
2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa	2θ	Int. relativa
11.10	1000.0	11.53	762.6				
13.30	177.8						
15.65	150.7						
20.10	234.7						
24.35	167.6						
28.10	135.2						
34.15	129.5	34.03	9.4				
36.05	118.9						
37.70	164.4						
40.30	148.8						
48.00	103.5						
49.95	116.1						
76.40	82.5						

ANEXO 3

Tabelas de associação de picos de espectroscopia Raman.

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-001	470 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-001	135 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-001	257 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-001	459 cm ⁻¹	453 ^a ,457 ^d , 456 ^f ,		
HU-001	136 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-001	262 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-001	278 cm ⁻¹	271 ^{b,g} ,272 ^e , 270 ^f ,286 ^g ,278 ^h		277 ^{b,d}
HU-001	473 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-001	474 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-001	134 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-001	265 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-001	385 cm ⁻¹			386 ^c
HU-001	469 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-002	268 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-002	469 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-003	139 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-003	267 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-003	474 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-004	473 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-005	139 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-005	264 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-005	476 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-005	138 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-005	264 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-005	473 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-005	134 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-005	259 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-005	470 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-005	470 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-005	136 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-005	470 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-005	472 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-005	134 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-005	472 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-005	138 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-005	261 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-005	473 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-006	136 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-006	263 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-006	473 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-006	138 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-006	261 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-006	473 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-007	139 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-007	260 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-007	473 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-008	138 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-008	256 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-008	472 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-009	265 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-009	474 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-009	302 cm ⁻¹	304 ^a		
HU-009	138 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-009	268 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-009	474 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-009	138 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-009	474 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAIS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-010	139 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-010	265 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-010	326 cm ⁻¹	325 ^{a,j}		
HU-010	474 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAlS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-011				

Amostra	Posição do pico	CuAlS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-012	139 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-012	264 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		264 ^c
HU-012	474 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAlS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-013				

Amostra	Posição do pico	CuAlS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-014	279 cm ⁻¹	271 ^{b,g} ,272 ^e , 270 ^f ,286 ^g ,278 ^h		277 ^{b,d}
HU-014	135 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-014	253 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		
HU-014	470 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d

Amostra	Posição do pico	CuAlS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-015	139 cm ⁻¹	146 ^a ,137 ^{b,c,e}		138 ^c
HU-015	253 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		
HU-015	272 cm ⁻¹	271 ^{b,g} ,272 ^e , 270 ^f ,286 ^g ,278 ^h		264 ^c , 277 ^{b,d}
HU-015	472 cm ⁻¹		470 ^d ,472 ^{a,b,e} ,475 ^c	474 ^d
HU-015	174 cm ⁻¹			
HU-015	252 cm ⁻¹	263 ^{a,b,e} ,262 ^c , 250 ^d ,266 ^f ,262 ^h , 246 ⁱ		

Amostra	Posição do pico	CuAlS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
HU-016	1000 cm ⁻¹			

Referências do Anexo 3

Referência	CuAlS ₂	Covellite(CuS)	Atacamite, Paratacamite Clinoatacamite
a	Phys. Rev. B 70, 075210, 2004	J. Am. Chem. Soc., 139, 13208–13217, 2017	Appl. Phys. Lett. 98, 092508, 2011
b	Ternary Compounds, Organic Semiconductors Volume 41 E 10.1007/10717201_42	https://ruff.info/covellite/R060129	N. Jb. Miner. Abh., 178, 2, 197–215, 2003
c	AMR, 11–12:175–178, 2006	https://ruff.info/covellite/R060129	N. Jb. Miner. Abh., 178, 2, 197–215, 2003
d	Eur. Phys. J. Appl. Phys. 36, 11–15, 2006	https://ruff.info/covellite/R060143	N. Jb. Miner. Abh., 178, 2, 197–215, 2003
e	Journal of Applied Physics 80, 5564, 1996	Chem. Mater., 27, 7531–7537, 2015	
f	phys. stat. sol. (b), 211, 429, 1999		
g	Journal of Electronic Materials, vol. 46, 4109–4118, 2017		
h	Solid State Communications Vol. 13, 7, 1011-1016, 1973		
i	Journal of Alloys and Compounds, 645, 85- 89, 2015		
j	Thin Solid Films, 517, 2191–2194, 2009		

ANEXO 4

Gráficos de Tauc dos espectros de absorção UV-Vis-NIR mostrados na Figura 70.

