

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL**

JANAINA SILVA HASTENREITER KÜSTER

**ESTUDOS DE CORRELAÇÕES E COMPARAÇÕES ENTRE LIMITES
DE LIQUIDEZ DE SOLOS OBTIDOS PELOS MÉTODOS DE
CASAGRANDE E CONE**

VITÓRIA

2022

JANAINA SILVA HASTENREITER KÜSTER

**ESTUDOS DE CORRELAÇÕES E COMPARAÇÕES ENTRE LIMITES
DE LIQUIDEZ DE SOLOS OBTIDOS PELOS MÉTODOS DE
CASAGRANDE E CONE**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Civil, do Centro Tecnológico da Universidade Federal do Espírito Santo como requisito parcial para obtenção do título de Mestre em Engenharia Civil, na área de concentração Estruturas/Geotecnia e Fundações.

Orientadora: Prof.^a Dr.^a Kátia Vanessa Bicalho - UFES

Coorientador: Prof. Dr. Lázaro Valentin Zuquette – USP-EESC

VITÓRIA

2022

Ficha catalográfica disponibilizada pelo Sistema Integrado de
Bibliotecas - SIBI/UFES e elaborada pelo autor

K97e Küster, Janaina Silva Hastenreiter, 1992-
Estudos de correlações e comparações entre limites de liquidez
de solos obtidos pelos métodos de Casagrande e cone / Janaina
Silva Hastenreiter Küster. - 2022.
269 f. : il.

Orientadora: Kátia Vanessa Bicalho.
Coorientador: Lázaro Valentin Zuquette.
Dissertação (Mestrado em Engenharia Civil) - Universidade
Federal do Espírito Santo, Centro Tecnológico.

1. Solos - Umidade. 2. Limite de liquidez (LL). 3. Cone de
penetração. 4. Concha de Casagrande. 5. Correlação (Estatística). 6.
Redes neurais (computação). I. Bicalho, Kátia Vanessa. II.
Zuquette, Lázaro Valentin. III. Universidade Federal do Espírito
Santo. Centro Tecnológico. IV. Título.

CDU: 624

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

ESTUDO DE CORRELAÇÕES E COMPARAÇÕES ENTRE LIMITES DE LIQUIDEZ DE SOLOS OBTIDOS PELOS MÉTODOS DE CASAGRANDE E CONE

Janaina Silva Hastenreiter Kuster

Dissertação apresentada ao Curso de Mestrado em Engenharia Civil do Programa de Pós-Graduação em Engenharia Civil da Universidade Federal do Espírito, como requisito parcial para obtenção do título de Mestre em Engenharia Civil, área de Estruturas.

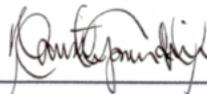
Aprovada no dia **07 de outubro de 2022** por:



Prof. Dra. Kátia Vanessa Bicalho
Doutora em Engenharia Civil
Orientadora – UFES



Prof. Dr. Lázaro Valentin Zuquette
Doutor em Geotecnia
Coorientador – USP



Prof. Dra. Maristela Gomes da Silva
Doutora em Engenharia Civil
Examinadora Interna – UFES

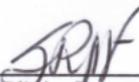


Prof. Dr. Adelmo Inacio Bertolde
Doutor em Estatística
Examinador Externo – UFES

gov.br

Documento assinado digitalmente
WILIAN HIROSHI HISATUGU
Data: 25/10/2022 06:47:58-0300
Verifique em <https://verificador.itl.br>

Prof. Dr. Wilian Hiroshi Hisatugu
Doutor em Engenharia Elétrica
Examinador Externo – UFES



Prof. Dr. Sílvio Romero de Melo Ferreira
Doutor em Engenharia Civil
Examinador Externo - UFPE

Vitória – ES, outubro de 2022

AGRADECIMENTOS

A Deus por me permitir viver essa experiência, por ser a minha força e por ter trilhado todo esse caminho ao meu lado.

Ao Programa de Pós-Graduação em Engenharia Civil - PPGEC da Universidade Federal do Espírito Santo pela oportunidade e por toda a aprendizagem proporcionada. Aos professores, funcionários e colegas de curso pela participação e contribuição ao longo dessa jornada.

Aos professores orientadores Kátia Vanessa Bicalho e Lázaro Valentin Zuquette por me guiarem na elaboração desta pesquisa, por toda a sua dedicação e por serem tão compreensivos.

Ao professor Wilian Hiroshi Hisatugu que tão pacientemente me auxiliou no estudo das redes neurais artificiais.

Aos professores Adelmo Inácio Bertolde, Silvio Romero de Melo Ferreira e Maristela Gomes da Silva por aceitarem o convite de fazer parte da Banca Examinadora e por assim contribuírem para o aprimoramento deste estudo.

Ao meu esposo Lucas por todo o incentivo, carinho e por compreender minha ausência em tantos momentos. Ao meu filho Benício, que ainda não chegou a este mundo, mas que tem sido a minha companhia nos dias e noites de estudo.

Aos meus pais por não medirem esforços para me proporcionarem o acesso aos estudos, por acreditarem tanto em mim e por toda a ajuda nesse período. Ao meu irmão Higne que sempre vibra com as minhas vitórias.

Aos amigos e familiares que estiveram ao meu lado e me incentivaram a prosseguir.

A Lugare Engenharia por consentir com a minha participação no programa e aos colegas de trabalho pelo apoio.

A todos que de alguma forma contribuíram para a realização de mais essa etapa da minha vida e celebraram essa vitória ao meu lado, meus sinceros agradecimentos.

RESUMO

Esta pesquisa analisa e compara valores de limites de liquidez (LL) obtidos pelo método de percussão de Casagrande, LL_c (aparelhos de base dura e macia), e pelos cones britânico e sueco, LL_p , para diferentes solos finos e uma ampla variação de valores de LL. Para as estimativas de LL_p foram utilizados regressões (lineares e não lineares) entre LL_c e LL_p e redes neurais artificiais (RNAs) do tipo *feed-forward* (FNN) treinadas com o algoritmo de retropropagação perceptron multicamadas (MLP), com uma ou duas camadas ocultas. Valores experimentais de LL_c e LL_p previamente compilados da literatura (507 amostras) foram selecionados para as análises estatísticas e RNAs. Os resultados experimentais foram separados em grupos de acordo com a dureza da base do aparelho de Casagrande e do tipo de cone utilizado, e divididos em subgrupos para a avaliação da influência do intervalo de LL. RNAs foram treinadas com LL_c , LL_p , IP e classificação SUCS dos solos como dados de entrada e comparadas com as redes de dados de entrada LL_c e LL_p (24 redes para cada conjunto de dados e obtido como saída LL_p). Análises estatísticas possibilitaram a seleção e o tratamento dos dados experimentais, com a eliminação dos *outliers*. Os dados de cada grupo e seus subgrupos foram submetidos a análises de regressão para o estabelecimento de correlações lineares e não lineares, e obtenção do coeficiente de determinação (R^2). As correlações lineares foram submetidas às análises de resíduos e aos testes de hipóteses para verificação da normalidade do modelo e independência das variáveis. A normalidade dos modelos foi verificada pela análise gráfica dos histogramas de frequência dos resíduos e do gráfico de Probabilidade Normal e pelos testes *Kolmogorov-Smirnov* (KS), *Shapiro-Wilk* (SW) e *Durbin-Watson* (DW). As correlações lineares e não lineares e as RNAs foram comparadas através de técnicas estatísticas que incluem os resultados obtidos para a raiz quadrada do erro médio (RMSE), erro médio absoluto (MAE), coeficiente de determinação (R^2), valores mínimos e máximos, média e desvio padrão (σ) das estimativas de LL_p . As análises dos dados indicam que os modelos propostos resultam em valores muito próximos para a previsão de LL_p . Os testes estatísticos mostraram que as correlações lineares obtidas nesta pesquisa, apesar dos elevados coeficientes de correlação ($R^2 > 0,74$), não foram significativas. Os resultados das RNAs mostram que além da variabilidade das propriedades geotécnicas dos resultados experimentais que compõem os conjuntos de dados, o número de amostras utilizadas na previsão de LL_p também influencia nos resultados. As RNAs treinadas têm potencial de aplicação para as estimativas de LL_p e representam uma ferramenta adicional para os métodos empíricos convencionais de regressão.

Palavras-chave: Limite de Liquidez, Plasticidade, Cone, Casagrande, Correlações de Regressão, Redes Neurais Artificiais.

ABSTRACT

This research analyzes and compares values of liquidity limits (LL) obtained by the Casagrande percussion method, LL_c (hard and soft base apparatus), and by the British and Swedish cones, LL_p , for different fine soils and a wide range of values of LL. For the LL_p estimates, regressions (linear and non-linear) were used between LL_c and LL_p and feed-forward artificial neural networks (ANNs) (FNN) trained using the multilayer perceptron backpropagation (MLP) algorithm, with one or two hidden layers. Experimental values of LL_c and LL_p previously compiled from the literature (507 samples) were selected for statistical analysis and ANNs. The experimental results were divided into groups according to the base hardness of the Casagrande appliance and the type of cone used, and divided into subgroups to assess the influence of the LL interval. ANNs were trained with LL_c , LL_p , IP and SUCS classification of soils as input parameters and compared with networks of input parameters LL_c and LL_p (24 networks for each dataset and obtained as LL_p output). Through the statistical analysis was possible to make the selection and treatment of data and eliminate outliers. Data from each group and its subgroups were submitted to regression analysis to establish linear and non-linear correlations, and to obtain the coefficient of determination (R^2). Linear correlations were submitted to residual analysis and hypothesis tests to verify the normality of the model and the independence of the variables. The normality of the models was verified by the graphic analysis of the residual frequency histograms and the Normal Probability plot and by the Kolmogorov-Smirnov (KS), Shapiro-Wilk (SW) and Durbin-Watson (DW) tests. Linear and nonlinear correlations and ANNs were compared using statistical techniques that include the results obtained for the root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), minimum and maximum values, mean and standard deviation (σ) of LL_p estimates. Data analysis indicates that the proposed models result in very close values for the LL_p prediction. Statistical tests showed that the linear correlations obtained in this research, despite the high correlation coefficients ($R^2 > 0,74$), were not significant. The ANN results show that in addition to the variability of the geotechnical properties of the experimental results that make up the data sets, the number of samples used in the LL_p prediction also influences the results. The trained ANNs have potential application for LL_p estimates and represent an additional tool for conventional empirical regression methods.

Keywords: Liquidity Limit, Plasticity, Cone, Casagrande, Regression Correlations, Artificial Neural Networks.

LISTA DE FIGURAS

Figura 1 - Novo gráfico de plasticidade do solo baseado no índice de fluxo e no limite de liquidez do cone de queda padrão britânico.	30
Figura 2 – Concha de Casagrande, riscador, cápsulas e espátulas.	34
Figura 3 – Riscador desenvolvido por Casagrande.....	38
Figura 4 – Riscador desenvolvido pela ASTM.....	39
Figura 5 – Cone de penetração para determinação do limite de liquidez.	41
Figura 6 – Avaliação de outliers: boxplot.....	65
Figura 7 – Avaliação de outliers: limites superiores e inferiores no gráfico de dispersão.	65
Figura 8 – Interpretação de Testes de Hipóteses.....	67
Figura 9 – Fluxograma da pesquisa.	80
Figura 10 - Correlações publicadas para o grupo B-01.....	85
Figura 11 - Correlações publicadas para o grupo B-02.....	87
Figura 12 - Correlações publicadas para o grupo B-03.....	88
Figura 13 – Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do grupo B-01.	92
Figura 14 – Histograma dos dados do grupo B-01: (a) LL_c ; (b) LL_p ; (c) IP; (d) porcentagem de argila; (e) porcentagem de finos; (f) classificação SUCS.	93
Figura 15 - Resultados experimentais dos dados selecionados para o grupo B-01.	94
Figura 16 – Ábaco de Casagrande para os dados de LL_c do grupo B-01.	95
Figura 17 - Ábaco de Casagrande para os dados de LL_p do grupo B-01.....	95
Figura 18 – Mudança de posição no ábaco de Casagrande para os dados do grupo B-01.	97
Figura 19 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01i.	98
Figura 20 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01ii.....	99
Figura 21 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01iii.....	100

Figura 22 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do grupo B-02.....	101
Figura 23 – Histograma dos dados do grupo B-02: (a) LL_c ; (b) LL_p ; (c) IP; (d) porcentagem de argila; (e) porcentagem de finos; (f) classificação SUCS.	102
Figura 24 - Resultados experimentais dos dados selecionados para o grupo B-02.	103
Figura 25 - Ábaco de Casagrande para os dados de LL_c do grupo B-02.....	104
Figura 26 - Ábaco de Casagrande para os dados de LL_p do grupo B-02.....	104
Figura 27 – Mudança de posição no ábaco de Casagrande para os dados do grupo B-02.	106
Figura 28 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02i.	107
Figura 29 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02ii.	108
Figura 30- Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02iii.	109
Figura 31 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do grupo B-03.	110
Figura 32 – Histograma dos dados do grupo B-03: (a) LL_c ; (b) LL_p ; (c) IP; (d) porcentagem de argila; (e) porcentagem de finos; (f) classificação SUCS.	111
Figura 33 - Resultados experimentais dos dados selecionados para o grupo B-03.	112
Figura 34 - Ábaco de Casagrande para os dados de LL_c do grupo B-03.....	113
Figura 35 - Ábaco de Casagrande para os dados de LL_p do grupo B-03.....	113
Figura 36 – Mudança de posição no ábaco de Casagrande para os dados do grupo B-03.	115
Figura 37 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03i.	116
Figura 38 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03ii.....	117
Figura 39 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do grupo B-03iii.....	118

Figura 40 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do grupo B-01 e as correlações linear e não linear definidas neste trabalho.	125
Figura 41 – Relatório de diagnóstico de resíduos - Grupo B-01.	128
Figura 42 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Grupo B-01.	128
Figura 43 – Resultados estatísticos das redes neurais e regressões do Grupo B-01.....	131
Figura 44 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01i e as correlações linear e não linear definidas neste trabalho.	133
Figura 45 - Relatório de diagnóstico de resíduos – Subgrupo B-01i.	135
Figura 46 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-01i.....	136
Figura 47 - Resultados estatísticos das redes neurais e regressões do Subgrupo B-01i. ...	139
Figura 48 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01ii e as correlações linear e não linear definidas neste trabalho.	141
Figura 49 – Relatório de diagnóstico de resíduos – Subgrupo B-01ii.	144
Figura 50 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-01ii.....	144
Figura 51 - Resultados estatísticos das redes neurais e regressões do Subgrupo B-01ii. .	147
Figura 52 - Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01iii e as correlações linear e não linear definidas neste trabalho.	149
Figura 53 – Relatório de diagnóstico de resíduos – Subgrupo B-01iii.....	152
Figura 54 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-01iii.	152
Figura 55 - Resultados estatísticos das redes neurais e regressões do Subgrupo B-01iii. .	155
Figura 56 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do grupo B-02 e as correlações linear e não linear definidas neste trabalho.	162

Figura 57 – Relatório de diagnóstico de resíduos – Grupo B-02.....	165
Figura 58 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Grupo B-02.	165
Figura 59 – Resultados estatísticos das redes neurais e regressões do Grupo B-02.....	168
Figura 60 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02i e as correlações linear e não linear definidas neste trabalho.	170
Figura 61 - Relatório de diagnóstico de resíduos – Subgrupo B-02i.	172
Figura 62 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-02i.....	173
Figura 63 - Resultados estatísticos das redes neurais e regressões do subgrupo B-02i....	176
Figura 64 - Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02ii e as correlações linear e não linear definidas neste trabalho.	178
Figura 65 - Relatório de diagnóstico de resíduos – Subgrupo B-02ii.	181
Figura 66 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-02ii.	181
Figura 67 - Resultados estatísticos das redes neurais e regressões do subgrupo B-02ii...	184
Figura 68 - Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02iii e as correlações linear e não linear definidas neste trabalho.	185
Figura 69 - Relatório de diagnóstico de resíduos – Subgrupo B-02iii.	188
Figura 70 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-02iii.	189
Figura 71 - Resultados estatísticos das redes neurais e regressões do subgrupo B-02iii. .	192
Figura 72 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do grupo B-03 e as correlações linear e não linear definidas neste trabalho.	198
Figura 73 – Relatório de diagnóstico de resíduos – Grupo B-03.....	200
Figura 74 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Grupo B-03.	201

Figura 75 - Resultados estatísticos das redes neurais e regressões do Grupo B-03.	204
Figura 76 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03i e as correlações linear e não linear definidas neste trabalho.	206
Figura 77 - Relatório de diagnóstico de resíduos – Subgrupo B-03i.	208
Figura 78 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-03i.	208
Figura 79 - Resultados estatísticos das redes neurais e regressões do subgrupo B-03i....	211
Figura 80 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03ii e as correlações linear e não linear definidas neste trabalho.	213
Figura 81 - Relatório de diagnóstico de resíduos – Subgrupo B-03ii.	215
Figura 82 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-03ii.	216
Figura 83 - Resultados estatísticos das redes neurais e regressões do subgrupo B-03ii...	218
Figura 84 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03iii e as correlações linear e não linear definidas neste trabalho.	220
Figura 85 - Relatório de diagnóstico de resíduos – Subgrupo B-03iii.	222
Figura 86 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-03iii.	222
Figura 87 - Resultados estatísticos das redes neurais e regressões do subgrupo B-03iii. .	225

LISTA DE TABELAS

Tabela 1 – Modelos de cone de queda usados em diversos países.....	42
Tabela 2 – Intervalos de LL e LP para diferentes minerais de argila.	53
Tabela 3 – Valores críticos do Teste KS para 5% de significância.	69
Tabela 4 - Valores críticos do Teste SW para 5% de significância.	70
Tabela 5 - Valores críticos do Teste DW para 5% de significância.	71
Tabela 6 – Algumas correlações lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-01.	82
Tabela 7 – Algumas correlações não lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-01.	84
Tabela 8 – Algumas correlações lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-02.	85
Tabela 9 – Algumas correlações não lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-02.	86
Tabela 10 – Algumas correlações lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-03.	87
Tabela 11 - Dados experimentais compilados da literatura (grupo B-01).	91
Tabela 12 - Descrição estatística dos dados do grupo B-01 (N=601).	93
Tabela 13 – Resumo da classificação SUCS de LL_c e LL_p para os dados do grupo B-01. ...	97
Tabela 14 – Dados experimentais compilados da literatura (grupo B-02).....	101
Tabela 15 - Descrição estatística dos dados do grupo B-02 (N=99).	102
Tabela 16 - Resumo da classificação SUCS de LL_c e LL_p para os dados do grupo B-02...	105
Tabela 17 – Dados experimentais compilados da literatura (grupo B-03).....	109
Tabela 18 - Descrição estatística dos dados do grupo B-03 (N=137).	111
Tabela 19 - Resumo da classificação SUCS de LL_c e LL_p para os dados do grupo B-03...	114
Tabela 20 – Resumo das topologias das redes neurais treinadas.	120
Tabela 21 – Descrição estatística dos dados investigados no grupo B-01.	123
Tabela 22 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01.	126

Tabela 23 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o Grupo B-01.	126
Tabela 24 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o Grupo B-01.	127
Tabela 25 – Resumo dos resultados dos testes estatísticos para o Grupo B-01.....	129
Tabela 26 – Resumo estatístico das estimativas de LL_p obtidas com as RNAs para as amostras do grupo B-01.	130
Tabela 27 - Descrição estatística dos dados investigados no subgrupo B-01i.	132
Tabela 28 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01i.	134
Tabela 29 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-01i.	134
Tabela 30 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-01i.	135
Tabela 31 – Resumo dos resultados dos testes estatísticos para o subgrupo B-01i.	136
Tabela 32 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01i.	137
Tabela 33 - Descrição estatística dos dados investigados no subgrupo B-01ii.....	140
Tabela 34 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01ii.	142
Tabela 35 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-01ii.....	142
Tabela 36 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-01ii.....	143
Tabela 37 – Resumo dos resultados dos testes estatísticos para o subgrupo B-01ii.	145
Tabela 38 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01ii.	145
Tabela 39 - Descrição estatística dos dados investigados no subgrupo B-01iii.....	147
Tabela 40 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01iii.	150
Tabela 41 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-01iii.....	150
Tabela 42 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-01iii.....	151

Tabela 43 – Resumo dos resultados dos testes estatísticos para o subgrupo B-01iii.	153
Tabela 44 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01iii.	153
Tabela 45 - Descrição estatística dos dados investigados no grupo B-02.....	161
Tabela 46 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02.	163
Tabela 47 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o Grupo B-02.	163
Tabela 48 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o Grupo B-02.	164
Tabela 49 – Resumo dos resultados dos testes estatísticos para o Grupo B-02.....	166
Tabela 50 – Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02.	167
Tabela 51 – Descrição estatística dos dados investigados no subgrupo B-01ii.....	169
Tabela 52 – Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02i.	171
Tabela 53 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-02i.	171
Tabela 54 – Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-02i.	172
Tabela 55 – Resumo dos resultados dos testes estatísticos para o subgrupo B-02i.	173
Tabela 56 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-02i.	174
Tabela 57 - Descrição estatística dos dados investigados no subgrupo B-02ii.....	177
Tabela 58 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02ii.	179
Tabela 59 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-02ii.....	179
Tabela 60 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-02ii.....	180
Tabela 61 – Resumo dos resultados dos testes estatísticos para o subgrupo B-02ii.	182
Tabela 62 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-02ii.	183
Tabela 63 - Descrição estatística dos dados investigados no subgrupo B-02iii.....	184

Tabela 64 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02iii.	186
Tabela 65 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-02iii.....	186
Tabela 66 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-02iii.....	187
Tabela 67 – Resumo dos resultados dos testes estatísticos para o subgrupo B-02iii.	189
Tabela 68 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-02iii.	190
Tabela 69 - Descrição estatística dos dados investigados no subgrupo B-03.	197
Tabela 70 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03.	199
Tabela 71 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o Grupo B-03.	199
Tabela 72 – Resumo dos resultados dos testes estatísticos para o Grupo B-03.....	201
Tabela 73 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03.	202
Tabela 74 - Descrição estatística dos dados investigados no subgrupo B-03i.	204
Tabela 75 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03i.	207
Tabela 76 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-03i.	207
Tabela 77 – Resumo dos resultados dos testes estatísticos para o subgrupo B-03i.	209
Tabela 78 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-03i.	210
Tabela 79 - Descrição estatística dos dados investigados no subgrupo B-03ii.....	212
Tabela 80 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03ii.	214
Tabela 81 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-03ii.....	214
Tabela 82 – Resumo dos resultados dos testes estatísticos para o subgrupo B-03ii.	216
Tabela 83 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-03ii.	217
Tabela 84 - Descrição estatística dos dados investigados no subgrupo B-03iii.....	219

Tabela 85 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03iii.	221
Tabela 86 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-03iii.....	221
Tabela 87 – Resumo dos resultados dos testes estatísticos para o subgrupo B-03iii.	223
Tabela 88 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-03iii.	224

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
ASTM	American Society for Testing and Materials
BS	British Standards
C	Argila
Cc	Índice de Compressão
CH	Argila de Alta Compressibilidade
CL	Argial de Baixa Compressibilidade
Cr	Índice de Recompressão
Cv	Coeficiente de Variação
CRAS	Curva de retenção de água no solo
d	Profundidade de Penetração do Cone
DL	Deep Learning
DW	Teste de independência dos resíduos de Durbin-Watson
e	Índice de Vazios
e_0	Índice de Vazios Inicial
ELM	Extreme Learning Machine (Máquina de Aprendizado Extremo)
FNN	Rede Neural Feedforward
H	Solos de Alta Compressibilidade
IA	Inteligência Artificial
IP	Índice de Plasticidade
IL	Índice de Liquidez
K	Fator de Cone

KS	Teste de aderência de Kolmogorov-Smirnov
L	Solos de Baixa Compressibilidade
LL	Limite de Liquidez
LL _c	Limite de Liquidez obtido pelo método de Casagrande
LL _{c-hardbase}	Limite de Liquidez obtido pelo método de Casagrande de base dura
LL _{c-softbase}	Limite de Liquidez obtido pelo método de Casagrande de base macia
LM	Levenberg-Marquardt
LL _p	Limite de Liquidez pelo método do Cone
LP	Limite de Plasticidade
M	Silte
MH	Silte de Alta Compressibilidade
ML	Silte de Baixa Compressibilidade
MLP	Perceptron Multi-camadas
MAE	Erro médio Absoluto
PL _c	Limite de plasticidade obtido pelo método do cone
RMSE	Raiz Quadrada do Erro Médio
R ²	Coeficiente de determinação
RNA	Rede Neural Artificial
S _u	Resistência Não Drenada
SUCS	Sistema Unificado de Classificação de Solos
SW	Teste de normalidade de Shapiro-Wilk
NBR	Norma da ABNT
w	Teor de Umidade
w _{nat}	Umidade natural inicial

W	Peso do cone de queda
x	Variável Dependente
y	Variável Independente
α	Nível de Significância
β	Índice de Penetração do Cone
μ	Média
ρ_{sat}	Peso específico saturado
σ	Desvio Padrão
σ^2	Variância

SUMÁRIO

1. INTRODUÇÃO	22
1.1 Justificativa.....	23
1.2 Objetivos	23
1.2.1 Objetivos Gerais.....	23
1.2.2 Objetivos Específicos.....	23
1.3 Estrutura da Dissertação.....	24
2. REVISÃO BIBLIOGRÁFICA.....	26
2.1 Fundamentos Gerais.....	26
2.1.1 Argila (Minerais de Argila).....	26
2.1.2 Plasticidade.....	27
2.1.3 Classificação dos Solos Finos	28
2.1.4 Limite de Plasticidade	31
2.1.5 Limite de Liquidez	33
2.2 Métodos para Obtenção do Limite de Liquidez.....	33
2.2.1 Método de Casagrande.....	34
2.2.2 Método do Cone.....	40
2.3 Comparações e Correlações entre os métodos de cone e de percussão	45
2.3.1 Influência do intervalo de LL	50
2.3.2 Influência do tipo de Argilomineral.....	52
2.3.3 Influência da Fração de Argila	55
2.3.4 Influência do Método de Preparação das Amostras	58
2.4 Outras Correlações Geotécnicas	59
2.4.1 Uso de LL para estimar resistência.....	60
2.4.2 Uso de LL para estimar parâmetros de compressibilidade.....	61
2.5 Métodos Estatísticos	63
2.5.1 Etapas de determinação de uma correlação	63

2.5.2	Análise de Regressão	66
2.5.3	Avaliação dos resíduos	67
2.5.4	Testes de Hipótese	68
2.6	Redes Neurais.....	72
2.6.1	Arquitetura de redes neurais artificiais.....	73
2.6.2	Aplicação de redes neurais artificiais.....	75
2.6.3	Aplicação de RNAs na engenharia geotécnica.....	76
2.6.4	Uso de RNA para estimar limites de consistência de solos.....	78
3.	MATERIAIS E MÉTODOS.....	80
3.1	Descrição do Banco de Dados	80
3.1.1	Correlações Publicadas entre LL_c e LL_p	81
3.1.2	Comparações entre LL_c e LL_p	89
3.2	Análises Estatísticas.....	118
3.3	Análises de Redes Neurais artificiais.....	120
3.4	Metodologia de avaliação dos resultados.....	122
4.	RESULTADOS E DISCUSSÕES	123
4.1	Resultados para os dados do grupo B-01	123
4.1.1	Avaliação estatística de amostras do grupo B-01	123
4.1.2	Aplicação das RNAs nas amostras do grupo B-01	129
4.1.3	Avaliação estatística dos dados do subgrupo B-01i	131
4.1.4	Aplicação das RNAs nas amostras do grupo B-01i	137
4.1.5	Avaliação estatística dos dados do subgrupo B-01ii.....	139
4.1.6	Aplicação das RNAs nas amostras do grupo B-01ii	145
4.1.7	Avaliação estatística dos dados do subgrupo B-01iii.....	147
4.1.8	Aplicação das RNAs nas amostras do grupo B-01iii.....	153
4.1.9	Conclusões parciais para o grupo de dados B-01	156
4.2	Resultados para os dados do grupo B-02	160
4.2.1	Avaliação estatística dos dados do subgrupo B-02	160

4.2.2	Aplicação das RNAs nas amostras do grupo B-02	167
4.2.3	Avaliação estatística dos dados do subgrupo B-02i	168
4.2.4	Aplicação das RNAs nas amostras do grupo subgrupo B-02i	174
4.2.5	Avaliação estatística dos dados do subgrupo B-02ii.....	176
4.2.6	Aplicação das RNAs nas amostras do grupo subgrupo B-02ii	182
4.2.7	Avaliação estatística dos dados do subgrupo B-02iii.....	184
4.2.8	Aplicação das RNAs nas amostras do grupo subgrupo B-02iii.....	190
4.2.9	Conclusões parciais para o grupo de dados B-02	192
4.3	Resultados para os dados do grupo B-03	197
4.3.1	Avaliação estatística dos dados do subgrupo B-03	197
4.3.2	Aplicação das RNAs nas amostras do grupo B-03	202
4.3.3	Avaliação estatística dos dados do subgrupo B-03i	204
4.3.4	Aplicação das RNAs nas amostras do grupo subgrupo B-03i	209
4.3.5	Avaliação estatística dos dados do subgrupo B-03ii.....	211
4.3.6	Aplicação das RNAs nas amostras do grupo subgrupo B-03ii	217
4.3.7	Avaliação estatística dos dados do subgrupo B-03iii.....	219
4.3.8	Aplicação das RNAs nas amostras do grupo subgrupo B-03iii.....	223
4.3.9	Conclusões parciais para o grupo de dados B-03	225
5.	CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS.....	230
5.1	Conclusões	230
5.2	Sugestões para trabalhos futuros	234
	REFERÊNCIAS	236
	APÊNDICE A.....	246

1. INTRODUÇÃO

A plasticidade é a propriedade que faz com que o sistema argila-água se deforme continuamente mediante uma força, e quando essa força é removida ou reduzida, a forma é mantida (Andrade, 2011; Bicalho et al., 2017). É quantificada pelo índice de plasticidade (IP) definido como a diferença no teor de umidade entre o limite de liquidez (LL) e o limite de plasticidade (LP). E, tem grande importância para a engenharia geotécnica, aplicações agrônomicas, na indústria cerâmica e no processo de fabricação de tijolos (O’Kelly, 2021).

Geralmente, dois métodos, o método de Casagrande e do cone são adotados para determinar o limite de liquidez de um solo (Sridharan e Prakash, 2000; Ozer, 2009; Mishra et al. , 2011; Quintela et al., 2014). Tradicionalmente, o procedimento mais utilizado para a determinação do limite de liquidez é o método de percussão padronizado por Casagrande (1932) que utiliza o aparelho desenvolvido pelo mesmo. O limite de liquidez determinado pelo método de Casagrande (LL_c) consiste então no teor de umidade do solo com o qual uma ranhura feita em uma amostra disposta em uma concha se fecha depois de 25 golpes. Devido a inúmeras variáveis intrínsecas à técnica do operador, teoria e aparelho de teste, muitos países adotaram um meio alternativo de determinação do limite de liquidez (Kestler, 1982). O uso do método do Cone de penetração para determinação do limite de liquidez de solos argilosos é uma alternativa atrativa por reduzir a interferência do operador nos resultados dos ensaios (Quintela et al., 2014; Bicalho et al., 2014). O limite de liquidez determinado pelo método de cone (LL_p), consiste na medição da penetração de um cone verticalmente em uma amostra de solo.

Vários estudos científicos concluíram que o ensaio com o cone apresenta resultados mais confiáveis relativamente ao método mais tradicional de Casagrande. Portanto, conforme Bicalho et al. (2014), o extenso banco de dados disponível com correlações entre LL_c e diferentes propriedades de engenharia, motivam o estudo de comparações entre os valores de LL_c e LL_p para diferentes solos finos. Além disso, outra importância da determinação do limite de liquidez está na sua aplicação na classificação dos solos finos (Sridharan et al., 2000) através do gráfico (carta ou ábaco) de plasticidade de Casagrande.

Nesta dissertação são analisados e comparados valores experimentais de LL_c (com aparelho de Casagrande base dura e base macia) e LL_p (com os cones britânico e sueco) e são utilizadas redes neurais artificiais *feed-forward* (FNN) treinadas com o algoritmo de retropropagação perceptron multicamadas (MLP) e análises de regressão para avaliar correlações entre LL_c e LL_p para diferentes solos argilosos e intervalos de LL.

1.1 Justificativa

A existência de diversos métodos para a obtenção dos limites de liquidez (LL) associados a variações observadas em padronizações internacionais para os referidos métodos e as variações de resultados obtidos tem motivado pesquisas sobre o tema.

O estudo da variabilidade de equações publicadas na literatura que correlacionam os resultados de LL obtidos por diferentes metodologias e solos é necessário, uma vez que os valores de LL são adotados em classificações e avaliações preliminares de solos finos para uso em obras geotécnicas. A avaliação das incertezas associadas a determinação do LL de solos finos é recomendada, uma vez que o uso de LL errado pode resultar na rejeição de materiais satisfatórios ou na aceitação de materiais inadequados. Os ensaios de determinação de LL possuem metodologias diferentes e fornecem valores distintos, e essas dispersões são maiores ou menores de acordo com o intervalo de valores de LL.

Assim, o desenvolvimento desta pesquisa é motivado pela variabilidade dos resultados de LL obtidos por diversos métodos e solos e as incertezas associadas a aplicações de LL em classificações de solos finos e correlações empíricas para estimativa de propriedades geotécnicas de solos argilosos de baixa a elevada plasticidade. Do ponto de vista científico, a pesquisa se justifica pela realização de uma análise crítica, com base em dados prévios, e a obtenção de correlações de regressão e de redes neurais artificiais. A justificativa técnica está associada a análise das correlações obtidas neste estudo, bem como das correlações da literatura, e os problemas enfrentados quanto a classificação dos solos finos.

1.2 Objetivos

1.2.1 Objetivos Gerais

O objetivo desta pesquisa foi desenvolver uma análise crítica dos resultados de limites de liquidez (LL_c e LL_p) com base no levantamento de dados prévios da literatura obtidos pelo método de percussão de Casagrande, LL_c (aparelhos de base dura e macia), e pelos cones britânico e sueco, LL_p , a elaboração de um banco de dados e a obtenção de correlações de regressão e de redes neurais artificiais.

1.2.2 Objetivos Específicos

São objetivos específicos deste trabalho:

- i. Avaliar os principais fatores associados as características dos solos e metodologias de ensaios que influenciam a variação dos valores de LL;

- ii. Examinar valores de LL obtidos por diferentes procedimentos e discutir a influência dessas diferenças em sistemas de classificação de solos;
- iii. Analisar estatisticamente correlações obtidas entre valores de LL determinados pelos métodos de percussão proposto por Casagrande, LL_c , e do cone de penetração, LL_p , considerando várias metodologias de ensaios, diferentes equipamentos e diversos solos argilosos (solos naturais, argilas puras e misturas de argila e areia fina);
- iv. Analisar por meio da técnica denominada Rede Neural Artificial (RNA) correlações entre LL_c e LL_p para diversos solos argilosos (solos naturais, argilas puras e misturas de argila e areia fina);
- v. Comparar os resultados de LL_p previstos pelas correlações de regressão e RNAs sugeridas nesta pesquisa.

1.3 Estrutura da Dissertação

Esta dissertação é estruturada em capítulos, conforme descrito a seguir.

O Capítulo 1 intitulado **Introdução** apresenta a justificativa para o desenvolvimento do trabalho, os objetivos gerais e específicos e a estrutura de organização da dissertação.

O Capítulo 2 intitulado **Revisão Bibliográfica** apresenta uma revisão da literatura acerca dos temas relevantes para o desenvolvimento da dissertação, com o embasamento teórico necessário para o entendimento das diferentes metodologias para determinação de LL_c e LL_p e dos fatores que influenciam esses resultados. Contém ainda os princípios necessários a compreensão das redes neurais artificiais e correlações de regressão e sua aplicabilidade na definição de correlações entre LL_c e LL_p .

O Capítulo 3 intitulado **Materiais e Métodos** contem a metodologia utilizada para a elaboração da dissertação. Nesse capítulo é feita a caracterização do banco de dados e apresentada a metodologia para o treinamento das redes neurais artificiais, para as análises estatísticas e a metodologia de comparação e avaliação dos resultados.

No Capítulo 4 intitulado **Resultados e Discussões** são apresentados os resultados obtidos com a utilização das redes neurais artificiais, das análises estatísticas, bem como a análise desses resultados. Nesse capítulo são apresentadas comparações de desempenho entre diferentes arquiteturas de RNAs com diferentes entradas de dados e correlações de regressão.

No Capítulo 5 intitulado **Conclusões e Sugestões para Trabalhos Futuros** são apresentadas e consolidadas as conclusões obtidas com a pesquisa, as limitações

encontradas, e apresentadas eventuais sugestões para pesquisas futuras. Por fim, são apresentados as **Referências Bibliográficas** utilizados e os **Apêndices**.

2. REVISÃO BIBLIOGRÁFICA

2.1 Fundamentos Gerais

2.1.1 Argila (Minerais de Argila)

A mineralogia é o principal fator que controla o tamanho, forma e propriedades das partículas do solo (Mitchell e Soga, 2005). O tipo e a quantidade de argilominerais são os principais responsáveis pela plasticidade dos materiais (Quintela et al., 2014). Portanto, a identificação dos minerais argilosos e a análise de sua estrutura é essencial, para compreensão do comportamento dos solos (Sowers, 1972).

As argilas, segundo Mitchell e Soga (2005) podem ser definidas por seu tamanho, se referindo a todos os constituintes do solo menor que 0,002mm e também pela classe de minerais, já que minerais de argila são principalmente silicatos de alumínio hidratado. Os minerais de argila, em sua maioria, possuem forma lamelar, o que segundo Cunha (2012) explica a existência de uma elevada superfície específica e, conseqüentemente, elevada sensibilidade à água. Baseado nesta estrutura e na composição química, os grupos de maior interesse são montmorilonita, illita e caulinita (Queiroz de Carvalho, 1975; Widjaja e Kurniawan, 2020).

As montmorilonitas, chamadas algumas vezes de esmectitas, são compostas por uma lâmina de alumínio entre duas de silício (Sowers, 1972). Segundo Mitchell e Soga (2005), a lâmina de alumínio no octaédrico pode ser substituída por magnésio, ferro, zinco, níquel, lítio ou outros cátions, resultando em uma deficiência de carga. E, na teoria, cada substituição produz um mineral distinto (Sowers, 1972), o que explica a grande variedade das montmorilonitas. Isso por exemplo permite a distinção entre a montmorilonita de cálcio natural e a montmorilonita de sódio (bentonita) (Bain, 1971). A estrutura das montmorilonitas caracteriza-se por apresentar moléculas de água entre as unidades estruturais, e, assim, a mesma apresenta elevada expansibilidade e plasticidade (Cunha, 2012). Ressalta-se que as bentonitas de ocorrência natural são caracterizadas por índices de plasticidade extremamente altos (Bain, 1971).

A illita é o mineral de argila mais comumente encontrado em solos na prática da engenharia (Mitchell e Soga, 2005). Segundo Sowers (1972), assim como as montmorilonitas, são formadas por uma lâmina de alumínio entre o silício, porém nas illitas, os átomos de silício são substituídos parcialmente por alumínio. Bain (1970) destaca que são difíceis de obter em estado suficientemente puro para testes e não são ativamente pesquisadas como um tipo específico de argila industrial, que, as variações nos valores de limite de liquidez e índice de

plasticidade para essas argilas são esperadas muito por conta da composição variável do argilomineral.

As caulinitas são formadas por lâminas de alumínio e sílica (Sowers, 1972). Segundo Cunha (2012), apresentam pequena expansão, difícil dispersão em água e baixa plasticidade, e, segundo Queiroz de Carvalho (1975), têm a capacidade de troca de cátions inferior a dos outros grupos. Os principais argilominerais do grupo da caulinita são: caulinita, nacrita, diquita e haloisita (Queiroz de Carvalho, 1975). Segundo Bain (1970), embora as caulinitas sejam virtualmente não plásticas e forneçam índices de plasticidade inferiores, a maioria apresenta algum comportamento plástico, com índices de plasticidade variando de acordo com o tamanho dos grãos.

Segundo Sowers (1972), o tipo de mineral argílico influencia muito a drenagem do solo, sua resistência, compressibilidade e particularmente sua reação às mudanças de umidade. As argilas têm propriedades de absorver água em virtude da insaturação negativa que apresentam em suas superfícies oriunda principalmente das ligações partidas de íons da superfície do cristal (Queiroz de Carvalho, 1975).

2.1.2 Plasticidade

A plasticidade pode ser definida como a propriedade de uma argila de modificar repetidamente sua forma sem sofrer ruptura. E, é influenciada por diversos fatores tais como umidade, composição mineralógica, distribuição do tamanho da partícula, tipo de cátion trocável, presença de sais ou matéria orgânica, e ao processo de moldagem da argila (Andrade et al., 2011).

Um solo argiloso pode apresentar-se nos estados líquido, plástico, semissólido ou sólido, a depender de sua umidade. O conceito de plasticidade, segundo Andrade et al. (2011), foi usado para explicar e caracterizar o comportamento dos materiais nos estados sólido ou líquido. Portanto, os limites inferiores e superiores do teor de água para cada estado do solo são denominados como limites de consistência ou limites de Atterberg.

Os limites de consistência foram introduzidos por Atterberg (1911) para caracterizar a relação entre a consistência da argila e o teor de água. Portanto, os limites de Atterberg são usados para identificação, descrição e classificação de solos finos e para avaliação preliminar de suas propriedades mecânicas (Mitchell e Soga, 2005). Esses parâmetros também são usados para prever a suscetibilidade de materiais de xisto argiloso ao intemperismo (Quintela et al., 2014). São, portanto, os parâmetros mais comumente especificados na engenharia geotécnica (Claveau-Mallet et al., 2012; O'Kelly, 2021; Crevelin, 2018)

À medida que o teor de umidade da argila aumenta, a plasticidade aumenta até um máximo, dependendo da natureza da argila (Andrade et al., 2011). Portanto, segundo Sousa (2011), o limite de plasticidade (LP) define a fronteira entre os comportamentos moldável e friável do solo enquanto o limite de liquidez (LL), entre o comportamento fluido e moldável, permitindo classificar os solos finos quanto à plasticidade.

2.1.3 Classificação dos Solos Finos

A classificação dos solos possibilita seu agrupamento de acordo com algumas de suas propriedades como o tamanho dos grãos, forma e composição mineralógica por exemplo. Casagrande (1948), argumentou que a plasticidade era a característica mais importante dos solos de granulação fina e que esta, ao invés do tamanho dos grãos, deveria ser, portanto, a base de um sistema de classificação de solos, e assim propôs um gráfico de plasticidade para interpretação dos limites de Atterberg em termos de propriedades de engenharia dos solos. Segundo Niazi et al. (2019), o conhecimento do teor de umidade em um solo e dos seus limites de consistência, além de facilitar a sua classificação, fornecem uma boa quantidade de informações sobre a resposta esperada desse solo sob o ponto de vista da engenharia e sobre as mudanças de tensão.

O uso do gráfico de plasticidade de Casagrande para classificar solos de granulação fina usando os limites de Atterberg é muito usado na engenharia geotécnica (Mitchell e Soga, 2005; Vardanega et al., 2021). O gráfico de plasticidade de Casagrande foi desenvolvido com base nos valores de LL e LP obtidos por meio dos métodos padrões: Método de Casagrande para obtenção de LL e método de Atterberg para a obtenção de LP. Assim, O'Kelly et al. (2018) discutem que do ponto de vista purista, apenas os limites de plasticidade e liquidez obtidos pelos métodos tradicionais devem ser adotados para a classificação dos solos finos, já que o gráfico foi desenvolvido com base nos valores dos métodos padrões ASTM descritos.

Segundo Manafi (2018), como a determinação da plasticidade dos solos tem um papel fundamental na sua classificação, a variabilidade na sua determinação pode até mesmo classificar um determinado solo em quatro categorias diferentes, o que justifica a importância da obtenção de outros métodos alternativos para a determinação da plasticidade do solo, a fim de se obter resultados mais precisos. Além disso, essa variabilidade afeta também o estabelecimento de correlações empíricas às propriedades de engenharia do solo.

Assim, Manafi (2018) discute que os valores de LL e LP obtidos pelos métodos padrão atuais não são valores definitivos, mas sim indicadores de quando o solo atinge os estados limites. Em seu estudo, o pesquisador faz uma análise dos efeitos dos valores de LL e LP na classificação dos solos por meio do método de regressão linear e observa que, conforme

esses valores variam, a posição desses solos no gráfico de plasticidade também varia, o que pode levar a classificação do solo em categorias diferentes. Vardanega et al. (2022) observou que, os solos que mudam de classificação são principalmente aqueles que originalmente ficavam muito próximos da linha A do gráfico de Casagrande.

Assim, conforme destacado por O' Kelly (2021), mudanças nos valores de LL, embora não signifiquem uma mudança fundamental no comportamento do material, podem ser suficientes para alterar a classificação de um solo de adequado a inadequado devido a valores de LL ou IP abaixo dos permitidos. Os valores dos limites de liquidez e plasticidade são afetados em vários graus dependendo das propriedades físicas e químicas do solo testado e pelos procedimentos laboratoriais usados para defini-los (Koester, 1992). No entanto, de acordo com O'Kelly et al. (2018) são frequentemente relatados na literatura sem referência aos métodos e (ou) padrões usados para sua determinação, o que introduz incerteza adicional no uso desses dados corretamente para a classificação dos solos ou em correlações.

Portanto, a introdução de valores de LL_p nos sistemas de classificação dos solos atuais pode afetar o uso de materiais em engenharia geotécnica, em alguns casos aumentando o custo para disposição e/ou melhoria do solo (Di Matteo 2012). Tendo isso em vista, pesquisadores (Shimobe e Spagnoli, 2020; Vardanega et al., 2022) tem discutido propostas alternativas para um sistema de classificação dos solos que se utilize dos resultados desses ensaios. O'Kelly (2021) por sua vez descreve avanços no sistema de classificação de solos inorgânicos de granulação fina (gráficos) que não dependam do LP obtido pelo método tradicional e discute propostas alternativas para determinações de limites de consistência e classificação de solo de granulação fina.

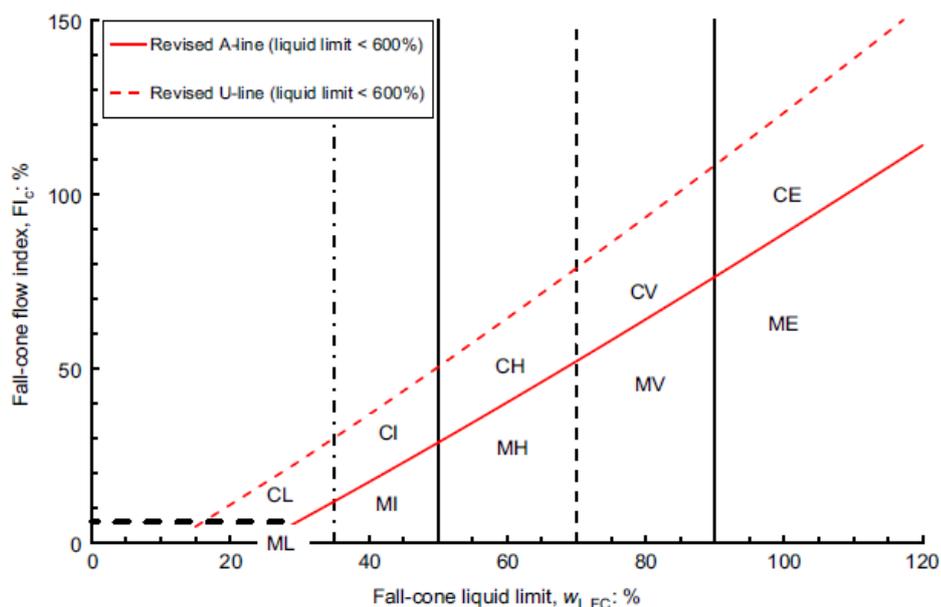
Gronbech et al. (2011) estudaram a relação entre os limites de liquidez de cone e de Casagrande para 33 amostras de argilas da Dinamarca, para profundidades de até 69 metros abaixo da superfície, com LL de até 350% e IP de até 300%, o que está bem fora da faixa normal do Gráfico de Plasticidade de Casagrande. O ábaco de plasticidade de Casagrande contempla um limite de liquidez de até 100% e um índice de plasticidade de até 60%, o que o torna inútil para classificação da argila dinamarquesa altamente plástica estudada. Assim, os pesquisadores propuseram duas novas categorias ao gráfico de plasticidade, argila de plasticidade super alta (CS) e argila de plasticidade extremamente alta (CE), cobrindo a faixa de limites de liquidez de até 200% e 350% respectivamente. Vale ressaltar que a argila estudada possui uma plasticidade raramente vista antes.

Shimobe e Spagnoli (2020) fizeram uso do método do cone de queda para avaliar as relações entre o teor de umidade (w), o índice de penetração do cone (β) e a tangente do ângulo de plasticidade (α). Os pesquisadores avaliaram mais de 500 dados obtidos por meio da literatura

e de ensaios realizados e discutiram sobre a potencial utilização de um diagrama que correlaciona o índice de penetração do cone (β) com a tangente do ângulo de plasticidade (α) para classificação dos solos finos, obtendo duas linhas características $\tan \alpha = 0,73$ e $\beta = 0,3$ para separação de siltes e argilas. E, concluíram que as correlações obtidas carecem de mais investigações.

Vardanega et al. (2022) utilizaram um banco de dados contendo 100 amostras de solos de granulação fina para redesenhar o gráfico de plasticidade com base na utilização do índice de fluxo e do limite de liquidez obtidos pelo método do cone. Para cada amostra, LL_p foi obtido utilizando o cone britânico, com uma penetração de 20mm, sendo relatado o valor da umidade (w). Adotando esse novo gráfico proposto pelos pesquisadores, o limite de plasticidade não é mais necessário para fins de classificação de solo, mudança essa que elimina a dependência desta classificação a um teste com alta variabilidade e dependência do operador como é o método tradicional para obtenção de LP. Cabe destacar ainda que as linhas A e U do gráfico de plasticidade de Casagrande foram determinadas através da utilização do método de Casagrande, assim, Vardanega et al. (2022) propuseram também uma atualização dessas linhas tendo como base nos resultados dos ensaios de cone. O gráfico proposto é apresentado na Figura 1.

Figura 1 - Novo gráfico de plasticidade do solo baseado no índice de fluxo e no limite de liquidez do cone de queda padrão britânico.



Fonte: Vardanega et al., 2022.

2.1.4 Limite de Plasticidade

O limite de plasticidade (LP) é o menor limite de teor de umidade dentro do qual o solo exibe comportamento plástico (Mitchell e Soga, 2005). O método mais utilizado para obtenção de LP é através do enrolar de uma amostra de solo sobre uma placa de vidro, sendo este o método tradicional descrito pela ABNT NBR 7180:2016 – Solo – Determinação do limite de plasticidade para determinação do limite de plasticidade de Atterberg. Esse método também é conhecido como o método de Atterberg para determinação do limite de plasticidade.

O procedimento de ensaio descrito pela ABNT NBR 7180 consiste na adição de água destilada a uma amostra de solo que deve ser homogeneizada por um tempo ideal compreendido entre 15 e 30 min, até a obtenção de uma pasta homogênea. Em sequência, toma-se uma amostra de cerca de 10g de solo previamente seca para formar uma bola que deve ser rolada sobre uma placa de vidro. Quando o cilindro formado nesse processo se fragmentar com diâmetro de 3mm e aproximadamente 100mm de comprimento, partes da amostra devem ser transferidas para um recipiente para determinação de sua umidade. Caso a amostra atinja o diâmetro sem se fragmentar o material deve ser amassado para repetição do procedimento. Caso se fragmente antes disso, a amostra recebe mais água destilada e os procedimentos seguintes são repetidos. Os procedimentos são repetidos de maneira a se obter ao menos 3 valores de umidade. A média dos resultados, expressa em porcentagem, indica o limite de plasticidade da amostra.

Manafi (2018) destaca que embora o procedimento de ensaio do método descrito pela ABNT NBR 7180:2016 pareça simples, o solo durante o ensaio se encontra sob tensões e sistemas de deformações complexos. Porém, embora o método possa parecer não científico, é atualmente a técnica mais adequada (O’Kelly et al., 2018).

Pesquisadores (O’Kelly et al., 2018; Manafi, 2018; O’Kelly, 2020) argumentam que os valores LP obtidos pelo método tradicional são dependentes do bom desempenho e julgamento do operador.

Manafi (2018) avaliou os resultados de ensaios de determinação do limite de plasticidade e concluiu que a variabilidade na determinação de LP aumenta conforme o aumento desse limite, mostrando que o julgamento do operador tem mais influência quanto maiores forem os valores de LP. O pesquisador discute ainda que a reprodutibilidade desse método diminui com o aumento dos valores de LP, já que diferentes técnicas podem ser usadas no procedimento de teste por vários operadores, especialmente para solos com altos valores de limite de plasticidade. Vardanega (2022) destaca que esta parcela de erro atribuído ao operador poderia ser então reduzida pela repetição dos ensaios e por um controle aprimorado do processo do teste.

Para investigar este ponto, O’Kelly et al. (2018) avaliou limites de plasticidade determinados independentemente por 4 laboratórios para 11 amostras de solos inorgânicos de granulação fina, com plasticidade intermediária a muito alta. A diferença máxima entre os menores e maiores valores de LP medidos no conjunto apresentado pelos pesquisadores para um determinado tipo de solo foi de 8%.

Como uma alternativa ao método tradicional, recentemente o teste do cone de queda não é usado apenas para a determinação do limite de liquidez, mas também estendido para a determinação simultânea de LL e LP usando um único cone (Shimobe e Spagnoli, 2019). Este método é mencionado na literatura como método de cone estendido.

Pesquisadores (Manafi, 2018; Shimobe e Spagnoli, 2019) citam que o método do cone de queda determina a força não drenada que é principalmente devido ao atrito não drenado, não podendo ser representativa da propriedade de plasticidade do solo. Segundo O’Kelly et al. (2018), é a deformação que indica se o solo está quebradiço (com formação de fissura) ou no seu estado plástico, em vez da magnitude da força aplicada ou da dureza de endentação como nos métodos de penetração do cone em que uma profundidade especificada para uma carga particular (ou seja, a resistência do solo) é tomada como a medição de força do limite de plasticidade. Assim, o método não pode determinar de forma consistente e confiável o limite de plasticidade no ponto de transição entre o estado plástico e o quebradiço (O’ Kelly, 2021).

Silva (2013), em contrapartida, discute que o método tradicional, além de ser bastante complicado de concretizar, também é bastante moroso, e que a determinação de LP por meio do método do cone apresenta as mesmas vantagens observadas, concluindo, portanto, que o método do cone para a obtenção do LP também é mais vantajoso. Embora as discussões, Shimobe e Spagnoli (2019) citam que há espaço para um estudo mais aprofundado do cálculo de LP com base no método do cone de queda estendido.

Outros métodos para a obtenção de LP também são identificados na literatura. Um dos métodos citados por O’Kelly et al. (2018), o aparelho de Barnes, é um método alternativo ao método tradicional para obtenção do LP, e que pode medir a tensão e os valores de tenacidade para um fio de solo, com controle da taxa de deformação. Embora os resultados desse método tenham sido valiosos, particularmente na obtenção de dados quantitativos de tenacidade do solo, tem desvantagens em termos de complexidade, lentidão e critérios empregados (O’Kelly, 2021). Shimobe e Spagnoli (2019) citam o ensaio de flexão de fio de solo para determinação do LP, proposto por Moreno-Maroto e Alonso-Azcárate (2015), que, segundo os pesquisadores, é simples e apresenta boa correspondência com o LP de Casagrande.

Como abordagens alternativas, baseadas em força, O’Kelly et al. (2018) citam os ensaios de cone de queda, penetração monotônica constante, carregamento estático e extrusão como meios de alcançar graus mais elevados de repetibilidade e reprodutibilidade dos resultados em relação ao método tradicional.

No entanto, embora os métodos alternativos citados, segundo O’Kelly et al. (2018), não foi desenvolvido nenhum método de medição do início da fragilidade melhor que o método tradicional.

Shimobe e Spagnoli (2019) ensaiaram 111 amostras de solos finos e obtiveram a seguinte correlação entre o limite de plasticidade obtido pelo método de Cone (PL_C) e o método tradicional PL:

$$PL_C = 0,86 PL + 3,14 \text{ com } R=0,876 \quad (1)$$

Assim, PL_C, de acordo com Shimobe e Spagnoli (2019) possui um erro relativo de cerca de 20% em comparação a PL, sendo a correspondência pior do que a de o limite de liquidez.

2.1.5 Limite de Liquidez

O limite de liquidez (LL) de um solo, descrito a primeira vez por Atterberg (1911), é a umidade em que ocorre a transição do comportamento líquido para o plástico (O’Kelly et al., 2018; Crevelin, 2018; Niazi et al., 2019), sendo, portanto, o valor máximo de umidade de uma argila em seu estado plástico. Segundo Bicalho et al. (2017), os valores de LL de argilas são frequentemente usados nas avaliações preliminares para seu uso em fundações e aterros para a construção de estradas e para obras de armazenamento e retenção de água ou contaminantes.

O’Kelly et al. (2018) citam que o valor de LL é fortemente dependente da classificação do solo, composição e propriedades mineralógicas, particularmente da fração de argila, da quantidade de água e da técnica usada para medi-lo, e seu uso requer cuidado devido à falta de uma padronização internacional das técnicas e equipamentos.

2.2 Métodos para Obtenção do Limite de Liquidez

Para a determinação do limite de liquidez de um solo argiloso, existem atualmente metodologias distintas, cujos resultados apresentam variações e são motivo de discussões e comparações. A concha de Casagrande e o cone de penetração são adotados como abordagens de medição (Mitchell e Soga, 2005; Mishra et al., 2011; Haigh, 2012; Quintela et al., 2014; O’Kelly et al., 2018; Crevelin, 2018; Shimobe e Spagnoli, 2019).

O método do cone foi aceito como o método padrão em muitos países como Inglaterra, Índia e Canadá, enquanto o método de Casagrande ainda é adotado oficialmente em muitos países, tais como Brasil e Portugal. Tendo em vista que as abordagens dos dois métodos são estabelecidas em vários padrões, sendo uma favorecida em relação à outra em diferentes partes do mundo, O'Kelly (2021) cita ser inevitável que ambos continuem a coexistir em prática até o momento em que um acordo internacional para a escolha definitiva de uma abordagem seja alcançado, juntamente com a padronização internacional de códigos para a alternativa escolhida.

2.2.1 Método de Casagrande

O método de Casagrande para a obtenção do limite de liquidez é adotado pela norma brasileira ABNT NBR 6459 – Solo – Determinação do limite de liquidez, que descreve os procedimentos de ensaio, bem como as características do aparelho. O aparelho consiste numa concha semiesférica que é deixada cair repetidamente de uma altura de 10 mm sobre uma base de borracha, que pode ser macia ou dura, através do uso de um mecanismo de manivela e engrenagens. O aparelho possui também um contador de golpes, e um cinzel com o qual se procede à abertura da ranhura.

Figura 2 – Concha de Casagrande, riscador, cápsulas e espátulas.



Fonte: Sousa, 2011.

O procedimento de ensaio descrito pela ABNT NBR 6459 consiste na adição de água destilada a uma amostra de solo até a obtenção de uma pasta homogênea, que transferida para a concha deve ser moldada para uma espessura da ordem de 10mm. Em sequência, o

deslocamento perpendicular do cinzel divide a amostra em duas partes, abrindo uma ranhura em sua parte central. A concha é então golpeada contra a base do aparelho, em queda livre, através do giro da manivela à razão de duas voltas por segundo. Por fim, anota-se a quantidade de golpes necessários para que as bordas inferiores da ranhura se unam por aproximadamente 13mm de comprimento. A umidade da amostra de solo é determinada e os procedimentos repetidos pelo menos 5 vezes, cobrindo o intervalo de 35 a 15 golpes, e adicionando água destilada na amostra a cada repetição. Com os resultados obtidos procede-se a construção do gráfico no qual as ordenadas sejam os números de golpes e as abcissas os teores de umidade correspondentes e ajustados a uma reta pelos pontos obtidos. O limite de liquidez do solo é então obtido na reta pelo teor de umidade correspondentes a 25 golpes.

2.2.1.1 Variáveis associadas ao ensaio de Casagrande

Casagrande (1958) criticou o teste para obtenção do limite de liquidez, apontando que existiam várias fontes de erro no teste incluindo equipamentos fora do padrão e erros oriundos da dureza da base sobre a qual o equipamento é utilizado. A existência de diferentes procedimentos de ensaio, bem como erros ocorridos durante sua execução, pode muitas vezes resultar em um valor equivocado de LL, este, quando utilizado, pode induzir a interpretações inadequadas de características de resistência e compressibilidade dos solos, por exemplo.

Embora os dispositivos de Casagrande sejam frequentemente distinguidos como dispositivos de base macia e dura, existe ainda uma considerável variabilidade mesmo dentro de cada uma dessas categorias. (O'Kelly et al., 2018). Segundo Manafi (2018), outros problemas relacionados a esse método são as diferentes técnicas de enchimento da concha com a amostra de solo, de formação da ranhura, da velocidade de queda em aparelhos manuais e o julgamento quanto ao fechamento do sulco. Silva (2013) cita outras variáveis como a dificuldade em fazer a ranhura nos solos arenosos, a influência do operador na profundidade da ranhura e ainda a tendência que os solos de baixa plasticidade tem de se liquidificar ao invés de fluir por plasticidade.

Tendo em vista as inúmeras variáveis que podem influenciar nos resultados dos ensaios pelo método de Casagrande, é fundamental descrever de forma mais detalhada os seguintes fatores:

- **Tipo de Solo**

Uma das desvantagens associadas à concha de Casagrande é a inadequação deste método para lidar com solos arenosos, já que a presença de areia pode dificultar a abertura da ranhura no solo (Kestler, 1982; Sousa, 2011). Além disso, alguns pesquisadores (Kestler, 1982; Silva,

2013) destacam que, nesse tipo de solo a resistência ao corte durante a execução da ranhura é maior do que a aderência e a fricção entre a superfície da concha e o solo, o que faz com que o solo seja arrastado ao longo da cavidade. Além disso, durante o ensaio, devido ao impacto com a base, a massa de solo de cada lado do sulco flui em direção ao centro para fechá-lo, em um escoamento viscoso, que segundo Kestler (1982) é extremamente difícil de discernir. Sridharan e Prakash (2000) portanto associam a impossibilidade da realização do ensaio em areias devido à completa ausência de cisalhamento viscoso.

Outra desvantagem é tendência dos solos de baixa plasticidade a escorregar ao invés de fluir por plasticidade (Kestler, 1982; Sridharan e Prakash, 2000; Mishra et al., 2011; Niazi et al., 2019). A baixa plasticidade dos solos tende ainda a segregar a água e alguns de seus constituintes (Kestler, 1982; Sousa, 2011). Pode-se, portanto, observar que o fechamento do sulco e, portanto, o limite de liquidez é apenas uma função da relação entre a resistência e a densidade do solo (Haigh, 2012)

- **Técnica do operador**

Outra inconveniência do aparelho de Casagrande e da metodologia do ensaio em questão é a sua sensibilidade à técnica do operador. Ozer (2009) cita que a inconsistência em alguns resultados pode ser explicada por fatores como a variação na massa e espessura da amostra de solo colocada na concha, a possível presença de bolhas de ar remanescentes no solo próximo ao sulco, uma diferença na disposição das partículas durante a colocação da amostra de solo na concha, o corte da ranhura, a velocidade de girar a manivela e a dificuldade de determinar o fechamento da ranhura.

A técnica do operador é, portanto, uma das variáveis de maior importância e que maior variação pode provocar na determinação de LL_c . Essa variação torna-se maior quando os operadores são de laboratórios diferentes (Kestler, 1982; Sousa, 2011), já que a metodologia adotada bem como o dispositivo utilizado podem ser diferentes. No entanto, a falta de experiência ou de precisão também são destacadas por Silva (2013) como causas da variabilidade dos resultados.

Liu e Thornburn (1964) realizou análises estatísticas para investigar quão bem um operador pode reproduzir os ensaios para obtenção dos limites de Atterberg e os efeitos da experiência do operador nos resultados. Dois operadores, um mais experiente e outro inexperiente, seguindo as recomendações propostas pelas normas ASTM, ensaiaram 20 amostras de solos silto arenosos, siltosos e argilosos. Os pesquisadores concluíram que para solos com baixa plasticidade, a variação de LL para o operador inexperiente é maior do que para o operador experiente e que, no entanto, essa quantidade de experiência parece ter pouco efeito sobre

a variação de LL nos solos de alta plasticidade. Em síntese, Liu e Thornburn (1964) observaram que para o ensaio de limite de liquidez pelo método de Casagrande pode-se esperar que um operador experiente reproduza os resultados do teste melhor do que um operador inexperiente.

A quantidade de solo colocado na concha também afeta diretamente os valores de LL_c . Solo em excesso irá provocar um sulco demasiado profundo, exigindo que o solo possua uma maior resistência ao cisalhamento para suportar 25 golpes (Kestler, 1982). Kestler (1982) explica que, como a resistência ao cisalhamento varia inversamente com o teor de umidade, para solo em excesso, o limite de liquidez tenderá a ser menor do que para a quantidade de solo apropriada. Assim, o inverso se aplica para os casos em que houver déficit de solo na concha, onde os valores de limite de liquidez tendem a ser maiores.

A frequência dos golpes aplicados na execução do ensaio encontra-se atualmente especificada pela ASTM e pela ABNT NBR 6459, na razão de dois golpes por segundo, e é também citada por pesquisadores como uma variável que influencia diretamente no valor de LL_c . Sousa (2011) destaca que essa influência é mais significativa nos casos em que a cadência fica abaixo de dois golpes por segundo e que quando estão acima, os resultados tendem a alinhar-se com a reta original correspondendo por tanto a dois golpes por segundo.

O ajuste da altura de queda pode ser outra importante fonte de erro para o ensaio de Casagrande (Claveau-Mallet et al., 2012). Claveau-Mallet et al. (2012) discutem a diferença entre as alturas de queda especificadas pelas normas britânica e a ASTM. Assim, enquanto a BS 1377 define a altura de queda de 10mm como a distância vertical máxima entre o ponto mais baixo da concha e da base, a ASTM D4318, define esta como a altura entre a base e o ponto da concha que atinge a base, que não corresponde ao ponto mais baixo quando a concha está totalmente levantada.

Visando minimizar a margem de erro humano, muitos pesquisadores (Ozer, 2009; Di Matteo, 2012; Spagnoli, 2012; Niazi et al., 2019) optaram pela realização dos ensaios em um mesmo laboratório e executados por um mesmo operador.

- **Aspectos físicos e mecânicos dos elementos que constituem o aparelho**

Casagrande (1958) propôs um aparelho no qual uma concha de latão é levantada e, em seguida, lançada a uma distância de 10 mm sobre uma base de borracha dura. O teste ainda realizado mundialmente difere pouco do proposto por Casagrande, mas é importante notar que existem variações nos equipamentos utilizados (Haigh, 2012).

Assim, destaca-se que o método de Casagrande também é sensível quanto a essas diferenças dos elementos que constituem o aparelho. Segundo Silva (2013) essas diferenças

estão relacionadas principalmente ao tipo de riscador para execução da ranhura bem como seu desgaste ao longo do tempo, a dureza da base e até a degradação do equipamento devido às sucessivas ranhuras realizadas.

A ferramenta de ranhura usada possui duas variantes: a ferramenta ASTM que corta a superfície do solo para impor um sulco profundo de 8 mm e a ferramenta curva da AASHTO sem cortador de superfície, usada com a especificação de 10 mm (Haigh, 2012).

Introduzida por Casagrande, a primeira ferramenta para execução da ranhura, na forma de uma espátula possui uma ponta aguçada e dois bordos que, segundo Silva (2013), são utilizados para nivelar e calibrar a camada de solo quando da execução da ranhura.

Figura 3 – Riscador desenvolvido por Casagrande.



Fonte: Silva (2013).

Algumas desvantagens desse riscador, segundo alguns pesquisadores (Kestler, 1982; Sousa, 2011) é a não conformidade das dimensões do sulco, de acordo com as especificações ASTM, e a dificuldade de manter a perpendicularidade com a concha para obtenção da direção e dimensões pretendidas.

O segundo utensílio, designado também por riscador ou cinzel, é longo, curvado e em forma de cunha. Segundo Silva (2013), este mantém as dimensões apropriadas na execução do sulco, porém contribui para a diminuição da plasticidade no momento da separação da camada de solo em duas partes, resultando no deslize das partes pela concha de Casagrande.

Figura 4 – Riscador desenvolvido pela ASTM.



Fonte: Silva (2013)

Esta ferramenta curva foi reconhecida por Casagrande (1958) como insatisfatória, por não controlar a altura do sulco, porém seu uso persistiu. Segundo Haigh (2012), o uso dessa ferramenta pode ocasionar um aumento na altura do declive, uma vez que não corta a superfície na altura necessária, que resultará, portanto, em resistências específicas maiores no limite de liquidez do que as provocadas pela ferramenta plana de Casagrande, o que corresponde a uma diminuição do limite de liquidez.

Kestler (1982) comparou os limites de liquidez obtidos pelas ferramentas de Casagrande e do padrão ASTM para dois tipos de solos argilosos, e observou, que para os tipos de solo examinados, a ferramenta de Casagrande produz um limite de líquido ligeiramente maior do que a ferramenta de ranhura padrão ASTM (que tende a fazer um sulco mais profundo).

O desgaste da ferramenta de ranhura tem um efeito duplo no comportamento observado no teste (Haigh, 2012). Assim, a diminuição da altura da ranhura devido ao desgaste, provoca o aumento da largura da folga que precisa fechar durante o ensaio, esses efeitos, segundo Haigh (2012), fazem com que o teste preveja uma resistência específica mais baixa no limite de liquidez. Como forma de eliminar os arranhões na superfície da concha, que são facilmente produzidos por ferramentas metálicas, pesquisadores (Mendoza e Orozco, 2001; Ozer, 2009) relatam a utilização em suas investigações de uma ferramenta de ranhura plástica.

A concha de Casagrande pode ainda apresentar bases com diferentes durezas (dura ou macia), o que irá influenciar nos resultados obtidos para LL_c. As especificações para esses dois tipos de dispositivos não estão bem definidas na maioria das normas existentes (Crevelin e Bicalho, 2019). Diversos padrões internacionais sugeriram o uso de diferentes tipos de bases para a realização do teste de Casagrande, de maneira que o padrão americano (ASTM D4318) prefere um aparelho com base dura, enquanto o padrão britânico (BS: 1377-parte 2 1990) uma base macia (Sridharan e Prakash, 2000; Mishra et al., 2012).

O limite de liquidez obtido com um aparelho de Casagrande de base dura é menor do que o obtido com o aparelho de base macia (O'Kelly et al., 2018; Shimobe e Spagnoli, 2019), e o

grau de desvio aumenta à medida que LL aumenta. Ozer (2009) explica que no método de percussão parte da energia produzida pela queda da concha é absorvida e parte é refletiva para a amostra de solo, e que, quando a base macia é usada, mais energia será absorvida pela base, e, portanto, será necessário aplicar um número maior de golpes para fechar a ranhura. O que, segundo O'Kelly et al. (2018), justifica os maiores valores de LL.

Ozer (2009) comparou os valores de limite de liquidez determinados pelo aparelho de Casagrande de base dura e base macia e pelo cone, com base em 42 amostras de solo natural de vários locais na Turquia, com LL variando entre 30 e 105%. Os resultados indicaram que os limites de liquidez obtidos pelo aparelho de Casagrande de base macia são até 5% superiores aos do aparelho de base dura. E, quando comparado com o cone, o aparelho de Casagrande de base dura geralmente fornece um limite de liquidez mais baixo, cerca de 70%, para solos de baixo LL, enquanto o aparelho de base macia fornece um limite de liquidez maior, cerca de 40%, para solos com um LL superior.

Crevelin e Bicalho (2019) no entanto, em sua avaliação da influência da dureza da base no valor de LL, verificaram que para as argilas montmorilonitas os resultados obtidos com o aparelho de base dura foram maiores do que os obtidos pelo aparelho de base macia. Com isso, os pesquisadores apontam para a necessidade de um estudo mais detalhado de influência da dureza da base em relação a mineralogia ou por faixas de valores de limite de liquidez.

Segundo Freitas Júnior (1974), variando-se também o tipo do material da base do aparelho, utilizando-se bases de diferentes coeficientes de restituição, pode-se mostrar que os valores de LL não sofrem variação sensível em seus resultados, desde que esse coeficiente não seja menor que 80%. Em sua pesquisa, Freitas Júnior (1974) observou que para três amostras ensaiadas, os valores de LL diminuíram em cerca de 10% em média utilizando-se vários tipos de bases com coeficientes de restituição variando de 53 a 90%.

Outros fatores associados como a variação na altura de queda da concha, os encaixes mecânicos entre as peças do aparelho, e o tipo de suporte (mesa de concreto, metal ou madeira) também são citados por Mendoza e Orozco (2001) como capazes de influenciar no valor do limite de liquidez. Freitas Júnior (1974) observou por meio da realização de ensaios com a variação das dimensões e peso da concha, que os valores de LL não são afetados sensivelmente por esses fatores.

2.2.2 Método do Cone

O teste do cone de queda foi desenvolvido originalmente na Suécia para medir a resistência ao cisalhamento não drenada dos solos de granulação fina (Shimobe e Spagnoli, 2019;

Karakan et al., 2020). Segundo Clemente (2018), atualmente é utilizado para a determinação dos limites de liquidez e plasticidade, sensibilidade de argilas e resistência não drenada para amostras indeformadas e amolgadas.

O ensaio consiste em medir a penetração vertical de um cone padronizado sob condições especificadas de peso, ângulo e tempo de queda, numa amostra de solo previamente preparada (Silva, 2013). Devem ser realizados no mínimo 4 ensaios repetidos, com diferentes teores de umidade, para a construção de um gráfico linear de penetração (mm) *versus* teor de umidade (%). O limite de liquidez será, portanto, o valor de umidade para uma determinada profundidade de penetração (10, 17, 20 ou 25mm), conforme o tipo de aparelho adotado.

Figura 5 – Cone de penetração para determinação do limite de liquidez.



Fonte: Crevelin, 2018.

O cone penetra até uma certa profundidade dependendo em parte da resistência ao cisalhamento não drenado do solo (Lee e Freeman, 2007). Segundo Sousa (2011), o ensaio é considerado não drenado porque a rapidez na penetração do cone não permite que as poropressões se dissipem. O limite de liquidez obtido pelo método do Cone (LL_p) corresponde a um valor fixo de resistência não drenada (S_u), que independe da densidade do solo (O'Kelly et al., 2018). A expressão seguinte é adotada para a obtenção de S_u do solo, com d sendo a profundidade de penetração do cone, W o peso e K um fator de cone.

$$Su = \frac{KW}{d^2} \quad (2)$$

Fonte: O'Kelly et al., 2018

Ao contrário do aparelho de Casagrande, não existe um único aparelho para o método de cone. Desde a descoberta deste novo mecanismo, diversos países desenvolveram estudos para testar as combinações de massa, ângulo do cone e profundidade de penetração mais favoráveis na determinação do limite de liquidez (Sousa, 2011). Portanto, o dispositivo, denominado *fall cone* ou cone de penetração, possui variantes em diferentes países. A Tabela 1 apresenta diversos tipos de cone identificados na literatura.

Tabela 1 – Modelos de cone de queda usados em diversos países.

País	Ângulo do cone (°)	Peso do conjunto (g)	Penetração no LL (mm)
Suécia	60	60	10
Índia	31	148	25,4
Estados Unidos	30	75	10
França	30	80	17
Inglaterra (BS 1377: Parte 2, 1990)	30	80	20
Canadá (BNQ 2501-092)	60	60	10
China	30	76	17

Fonte: Adaptado de Crevelin (2018).

Vale ressaltar que ainda não existe um padrão brasileiro disponível para a determinação do limite de liquidez usando o método de cone, mas alguns estudos nesse intuito já foram apresentados por Silveira (2001) e Queiroz de Carvalho (1986).

2.2.2.1 Variáveis associadas ao método de Cone

O cone de queda também possui suas próprias deficiências; entretanto, as variáveis associadas ao aparelho são em número mais limitado do que as associadas ao aparelho de Casagrande (Kestler, 1982). Além disso, muitos podem ser compensados. No entanto, Prakash (2012) destaca que há uma tendência crescente em usar o método do cone para determinar o limite de liquidez sem entender suas limitações.

O tempo de penetração, a nitidez da ponta, a rigidez do solo, a textura da superfície do cone e a técnica do operador estão entre as variáveis citadas por Kestler (1982) que podem ter algum efeito na repetibilidade do ensaio realizado pelo método de cone. Entretanto, ao

contrário do que se verifica no método de Casagrande, para o ensaio de cone, segundo Silva (2013), a técnica do operador é das variáveis menos relevantes.

Portanto, variáveis como o tempo de penetração, a rigidez do solo e as características do cone são descritas de forma mais detalhada a seguir:

- **Tempo de penetração**

O tempo decorrido entre a penetração e a medição da penetração é citado por Mendoza e Orozco (2001) com um fator que pode alterar o valor de limite de liquidez obtido pelo método do cone, já que a profundidade de penetração afeta os resultados (Claveau-Mallet et al., 2012). Segundo Sousa (2011), após o cone ser largado e penetrado na mistura de solo e água destilada, seu movimento deve ser livre por alguns segundos. A norma britânica (BS 1377: Parte 2, 1990) por sua vez, indica que o tempo necessário para atingir a profundidade de penetração adequada aos solos argilosos é de aproximadamente 5 ± 1 segundos, com o bloqueio em seguida do mecanismo de queda e do registro da leitura final. Silva (2013), no entanto, discute que para a maioria dos solos, com exceção dos solos lodosos, esse fator tem pouca influência e, verificou que a profundidade de penetração se mantém constante mesmo após o tempo determinado pela norma britânica.

Kestler (1982) investigou o tempo de penetração para o limite de liquidez determinado pelo cone, variando o tempo decorrido entre a penetração e a leitura. O pesquisador observou que na maioria dos casos, a profundidade de penetração permaneceu constante, ou quase, com o tempo. A exceção ocorreu para uma amostra de solo siltoso com um limite de liquidez muito baixo, o que com toda a probabilidade, devido a fluidez da água no interior dos poros provocou uma profundidade de penetração bem maior.

O uso de dispositivo temporizador permite que a penetração do cone na amostra seja desligada automaticamente após 5s (Di Matteo, 2012). Essa pode ser então uma maneira de mitigar os efeitos do tempo de penetração no valor do limite de liquidez e tem sido adotada por alguns pesquisadores (Ozer, 2009; Di Matteo, 2012). Vale ressaltar, conforme citado por Mendoza e Orozco (2001) que a precisão na medida da penetração também é um fator a ser observado, e que pode acarretar mudanças nos resultados obtidos com este ensaio.

- **Aspectos relacionados ao solo**

O tipo de solo e os aspectos de preparo da amostra também podem influenciar o valor de LL obtido com o cone de queda semelhante às influências para o método do copo de Casagrande (Lee e Freeman, 2007).

Em argilas de média a elevada rigidez, a penetração do cone é relativamente baixa, amplificando o erro associado à manobra (Sousa, 2011). Segundo Silva (2013), tal

inconveniente pode ser ponderado com a adição de pesos ao cone, desde que em conformidade com a norma seguida. No entanto, Crevelin (2018) destaca que essa variável carece de maiores investigações para determinação do incremento de peso adequado para que não se provoquem erros nos resultados obtidos.

Além disso, Lee e Freeman (2007) discutem que o processo de colocação do solo no copo especificado em BS1377 é influenciado pelo julgamento do operador e pode fornecer resultados inconsistentes se o posicionamento do solo não for uniforme.

Theron et al. (2019), cita que uma grande desvantagem do método de cone é que é necessária uma amostra de solo maior do que para o método de Casagrande e que associado ao cuidado necessário no enchimento do copo de teste para não inclusão de bolsas de ar, podem levar a falta de aceitabilidade do método. Bolsões de ar presos na argila ao redor do ponto de impacto também podem influenciar a profundidade de queda medida (Gronbech et al., 2011). Assim, Theron et al. (2019), estudou a substituição do copo por um anel de amostra de aço inoxidável do mesmo material e altura, com diâmetros de 54mm, 35mm, 29mm e 22mm, para 44 amostras de solo com grande variabilidade e com LL entre 25 e 80. O anel de 35mm foi escolhido por fornecer os mesmos resultados que o de tamanho padrão, indicando que o padrão de distorção de cisalhamento produzido pelo cone não é afetado significativamente.

- **Características do cone**

Dentre os diferentes aparelhos de cone sugeridos por diversos padrões internacionais, nota-se que grande parte dos ensaios registrados na literatura fazem referência aos cones do tipo sueco e britânico. No entanto, estudos indicam que ambos os tipos de cone fornecem essencialmente o mesmo valor de LL (Farrel et al., 1997). Ozer (2009) comparou os valores de limite de liquidez determinados pelo aparelho de Casagrande e pelos cones sueco e britânico e, concluiu que os valores de LL obtidos diferem ligeiramente.

Portanto, no que diz respeito a obtenção de um resultado independentemente de onde e quando o teste é realizado, O'Kelly et al. (2018) discutem que um dispositivo de cone padronizado é o meio mais apropriado para medir o limite de liquidez e que o ideal seria uma padronização do tipo de cone adotado internacionalmente. E, já que a maioria das localidades utilizam o padrão britânico (30° - 80g) com penetração de 20mm, esse, segundo os pesquisadores, tende a ser o padrão recomendado internacionalmente.

Mendoza e Orozco (2001) citam a variação do ângulo e massa do cone, a variação nas dimensões padrão do recipiente de solo ou alguma restrição mecânica na queda livre do cone como outros fatores atrelados às características do cone que podem influenciar no valor do limite de liquidez obtido por esta metodologia.

Outro fator importante que afeta a resistência no limite de liquidez, é a rugosidade do cone (Karakan et al., 2020). Segundo O'Kelly et al. (2018), o cone sueco é menos sensível a variações na rugosidade, podendo indiscutivelmente produzir maior repetibilidade entre os resultados se comparado ao cone britânico. A norma britânica BS 1377 menciona que a rugosidade da superfície tem mais influência na penetração do cone do que a variação na ponta do cone (Claveau-Mallet et al, 2012). No entanto, Claveau-Mallet et al. (2012) avaliou os resultados dos testes de limite de liquidez obtidos com o cone sueco e com o aparelho de Casagrande e observou que a profundidade de penetração do cone afeta os resultados, enquanto a massa e o grau de rugosidade do cone não.

Segundo Clemente (2018), a obtenção de resultados mais precisos através do método do cone depende ainda do bom estado do equipamento. Com o tempo, o uso repetido do cone resultará naturalmente em desgaste; entretanto, esse desgaste pode ser acelerado pelo uso intenso em solos arenosos (Kestler, 1982). Silva (2013) destaca que a verificação e identificação de possíveis desgastes no vértice do cone afasta a possibilidade da realização de leituras erradas e até mesmo da invalidação de ensaios. E, que tais desgastes podem ser suprimidos evitando a utilização do equipamento em amostras de solo com partículas arenosas.

Kestler (1982) cita que o cone possui vantagens quando comparado com o aparelho de Casagrande e a sua dependência ao desenho do aparelho, já que, essencialmente, a única parte do cone de queda que apresenta sinais de desgaste é o vértice e que, se necessário, isso pode ser compensado alinhando o vértice geométrico com a superfície do solo.

2.3 Comparações e Correlações entre os métodos de cone e de percussão

Os métodos do cone de queda e Casagrande medem diferentes propriedades mecânicas do solo, o teste de Casagrande é uma medida de força específica, isto é, força dividida pela densidade, enquanto o cone de queda é uma medida direta de força (Haigh, 2012). O dispositivo de Casagrande impõe uma carga de choque a amostra de solo com impactos repetidos contra a base do aparelho, iniciando uma falha de declive para fechar a abertura da ranhura feita na amostra (O'Kelly et al., 2018) enquanto o cone de queda penetra até uma certa profundidade, dependendo em parte da resistência ao cisalhamento não drenada do solo (Crevelin, 2018). Portanto, é esperado que os métodos forneçam diferentes valores de limite de liquidez (Haigh, 2012).

Atualmente o método de Casagrande está a decrescer a sua popularidade junto da comunidade técnica e científica devido à menor precisão dos resultados em comparação com

os obtidos pelo cone (Sousa, 2011). Casagrande (1958) reconhece que o método de percussão para determinar LL_c não fornece resultados uniformes de comparação para os solos de granulação fina, que apresentam reações diferentes quando submetidos a um ensaio dinâmico.

Segundo O'Kelly et al. (2018), embora o método de Casagrande seja utilizado há bastante tempo e exista uma quantidade enorme de dados usados inclusive em correlações, a falta de consistência entre os tipos de aparelhos torna o método não ideal para um teste tão amplamente adotado. O método, no entanto, apresenta como vantagens o baixo custo e a sensibilidade (Andrade et al., 2011) e continua a ser o método mais usado para a medição do limite de liquidez em grande parte do mundo (Haigh, 2012; Di Matteo, 2012).

O teste do Cone de queda por sua vez, segundo O'Kelly et al. (2018), sofre menos variabilidade em equipamento e execução do que o teste de Casagrande. Segundo Silva (2013), embora os métodos tendam a produzir resultados semelhantes, o método do cone pode ser dito mais vantajoso que o método de Casagrande pela maior facilidade e menor tempo de execução e por apresentar influência prática do operador praticamente nula. Clemente et al. (2020) discutem que o método do cone pode evitar alguns inconvenientes do ensaio de Casagrande como a dispersão dos resultados, as diferenças existentes entre os aparelhos, a diferença de material distribuída na concha, a observação do fechamento da abertura e o tempo de homogeneização incorreto. Entretanto, segundo Theron et al. (2019), o tamanho relativamente grande da amostra necessária para a realização do cone, juntamente com os cuidados atrelados ao enchimento do copo, tornou o método pouco atraente na África do Sul. Essa falta de aceitabilidade pode ter sido agravada pela alta variabilidade espacial das propriedades das argilas sul-africanas, levando a uma pobre correlação entre os testes em alguns casos (Theron et al., 2019).

Segundo Mishra et al. (2011), devido ao erro do operador, o método do cone se mostra mais preciso na determinação de LL. Assim, o método de Cone fornece uma melhor previsão de valores de limite de liquidez se comparado com o método de Casagrande (Silva, 2013; Niazi et al. 2019), principalmente por ser um método mais consistente e menos dependente do operador (Li et al., 2019).

A repetibilidade e a reprodutibilidade de uma medição são características importantes que podem ser quantificadas para permitir que os usuários compreendam a variabilidade dos resultados dos testes (Li et al., 2019). A reprodutibilidade dos ensaios através dos métodos de cone e de Casagrande refere-se à variação nas medidas de LL feitas em uma amostra sob diferentes condições (Rehman et al., 2020). Em quase todos os casos, qualquer literatura

aplicável afirma que o cone possui um grau muito maior de repetibilidade do que o teste convencional de Casagrande com diferentes operadores (Kestler, 1982).

Di Matteo (2012) avaliou a repetibilidade dos ensaios de limite de liquidez para o método de Cone e para o aparelho de Casagrande de base dura. Os ensaios foram repetidos três vezes, em um curto intervalo de tempo, com os mesmos solos, operador e condições de laboratório e a diferença observada para LL_p foi de 0,2%, enquanto para LL_c essa diferença chegou a 0,4%.

Ozer (2009) ensaiou amostras de argila de alta plasticidade para avaliar a repetibilidade dos ensaios de limite de liquidez, e, observou que a diferença entre os valores mais alto e mais baixo obtidos dos cinco testes replicados, executados pelo mesmo operador, usando os aparelhos de cone e de Casagrande foi de 0,8%.

Rehman et al. (2020) obtiveram o limite de liquidez de 30 amostras de solo de 10 países diferentes, utilizando os métodos de Casagrande de base dura e o cone britânico para avaliação da repetibilidade dos métodos. Os ensaios foram repetidos três vezes e, os resultados indicaram que o método de cone foi 16% mais repetível do que o método de Casagrande. Essa diferença, porém, segundo Rehman et al (2020) não foi estatisticamente significativa.

Li et al. (2019) ensaiou cinco amostras de misturas de solo contendo bentonita, também utilizando o método de cone britânico e o aparelho de Casagrande de base dura, e observaram que para o método de cone a fração do erro atribuído ao operador é de 50%, enquanto no método de Casagrande essa fração é de 89%. Os pesquisadores também correlacionaram os valores de LL_c e LL_p e observaram que ambos os métodos produzem variações progressivamente maiores com o aumento dos valores de LL, porém, as variações dos resultados do método de cone são muito menores do que as do método de Casagrande. Vale destacar que as amostras foram preparadas ao mesmo tempo, para minimização das possíveis variações causadas pelo preparo da amostra, e que, três operadores diferentes foram responsáveis pela execução do ensaio pelo método de Casagrande, embora tenham recebido um mesmo treinamento.

Manafi (2018), no entanto, cita que os resultados de seu estudo não confirmaram que o método do cone para determinação de LL é mais reproduzível em relação ao método de Casagrande, embora a norma britânica BS EN 1377-2 (1990) e o código europeu CEN EN 1997-2 (2007) se utilizem desse argumento para preferir o método do cone. Além disso, Manafi (2018) destaca que a ASTM D4318-17 (2017) ainda não reconhece o método do cone como o método padrão para determinação do limite de liquidez.

Rehman et al. (2020) destaca que os resultados de repetibilidade discutidos valem para um mesmo laboratório, já que a realização dos mesmos ensaios, porém com diferentes operadores, pode produzir resultados diversos. Portanto, quando são realizados estudos interlaboratoriais, o método do cone é relatado como tendo melhor repetibilidade (Claveau-Mallet et al, 2012).

Portanto, o cone de queda tem uma vantagem definitiva sobre o dispositivo de Casagrande, pois quase elimina fatores de julgamento (Kestler, 1982). As variáveis relacionadas ao método de Casagrande citadas no item 2.2.1.1 podem explicar a baixa repetibilidade do método de Casagrande.

Theron et al. (2019) investigou a adequação do método de cone para substituição do método de Casagrande e sua pesquisa mostra que a comparação de diferentes métodos de ensaio para amostras com grande variabilidade de solos pode conduzir a uma fraca correlação nos resultados, se a variabilidade não for tida em conta. Para a comparação entre LL_c e LL_p , amostras de três solos diferentes foram testadas para produzir funções de densidade de probabilidade. Os pesquisadores concluíram que, como as amostras estudadas são todas de variabilidade média a baixa, parece que cerca de 20 a 30 testes podem ser suficientes para obter uma correlação razoável entre os dois métodos, já que uma maior variabilidade pode exigir consideravelmente mais testes.

Além do mais, ao longo dos anos vários pesquisadores (Spagnoli, 2012; O’Kelly et al. (2018); Shimobe e Spagnoli (2019); Crevelin e Bicalho (2019); Niazi et al., 2019) têm se preocupado em estabelecer correlações entre LL_p e LL_c . Essas correlações, segundo O’Kelly et al. (2018) tem a finalidade de proporcionar as correções desejadas aos procedimentos de classificação dos solos utilizando as duas metodologias.

Spagnoli (2012) utilizou o resultado de 50 amostras de argila comercial pura, classificadas quanto a sua mineralogia como caulinitas e ilitas, e obteve a seguinte correlação para um intervalo de LL entre 20 e 61%:

$$LL_p = 0,99LL_c + 1,05 \quad (R^2 = 0,99) \quad (3)$$

Em sua pesquisa, Spagnoli (2012) não comparou a correlação obtida com as demais identificadas na literatura, devido às diferenças nas características geológicas das demais formações argilosas, já que estas utilizaram amostras de solos naturais enquanto o pesquisador utilizou amostras de argila pura. Assim, Spagnoli (2012) concluiu que a correlação obtida pode ser aplicada apenas para as argilas puras utilizadas na pesquisa.

O’Kelly et al. (2018) utilizaram um grande banco de dados obtidos por meio da literatura para estabelecer quatro correlações empíricas não lineares entre os métodos de cone e de Casagrande para diferentes padrões e apresentam as correlações seguintes, com LL_{FC} sendo o limite de liquidez obtido pelo método do Fall Cone, LL_{BScup} o limite de liquidez obtido pelo método de Casagrande usando o aparelho de base macia e $LL_{ASTMcup}$ o valor obtido pelo método de Casagrande de base dura recomendado pela ASTM.

Para uma gama completa de limite de liquidez O’Kelly et al. (2018) obteve:

$$LL_{FC} = 1,86 LL_{BScup}^{0,84} \quad (R^2 = 0,98, n = 216) \quad (4)$$

$$LL_{FC} = 1,90 LL_{ASTMcup}^{0,85} \quad (R^2 = 0,97, n = 199) \quad (5)$$

Para valores de limite de liquidez obtidos pelo método de Casagrande menor que 120%:

$$LL_{FC} = 1,62 LL_{BScup}^{0,88} \quad (R^2 = 0,96, n = 199) \quad (6)$$

$$LL_{FC} = 1,45 LL_{ASTMcup}^{0,92} \quad (R^2 = 0,97, n = 188) \quad (7)$$

Shimobe e Spagnoli (2019) apresentam uma tabela com a organização de 43 correlações obtidas por numerosas referências, dispostas em ordem cronológica, com os diferentes tipos de cone, de dureza da base do aparelho de Casagrande e tipos de solo. Além disso, analisaram 134 amostras de solo com limite de liquidez menor que 150%, utilizando o método de cone nos padrões britânico e sueco, e propuseram a seguinte correlação:

$$LL_C = 0,95 LL_p^{0,85} \quad (R^2 = 0,962) \quad (8)$$

Crevelin e Bicalho (2019) utilizaram resultados de 12 publicações, obtendo uma quantidade de dados igual a 184. Os dados analisados contêm os valores de limite de liquidez obtidos pelo cone britânico LL_p e pelos aparelhos de Casagrande de base dura $LL_{c-hard base}$ (117 amostras) e base macia $LL_{c-soft base}$ (67 amostras) e foram separados entre caulinitas e ilitas e montmorilonitas. Assim, as correlações propostas por Crevelin e Bicalho (2019) são as seguintes:

Para o aparelho de Casagrande de base dura em caulinitas e ilitas:

$$LL_{c-hardbase} = 1,0148 LL_p - 1,366 \quad (R^2 = 0,967, n = 71; 20 < LL_c < 100\%) \quad (9)$$

Para o aparelho de Casagrande de base dura em montmorilonitas:

$$LL_{c-hardbase} = 1,352 LL_p - 19,633 \quad (R^2 = 0,984, n = 39; 50 < LL_c < 460\%) \quad (10)$$

Para o aparelho de Casagrande de base macia em caulinitas e ilitas:

$$LL_{C\text{-softbase}} = 1,103 LL_p - 7,601 \quad (R^2 = 0,767, n = 28; 30 < LL_c < 70\%) \quad (11)$$

Para o aparelho de Casagrande de base macia em montmorilonitas:

$$LL_{C\text{-softbase}} = 1,216 LL_p - 8,834 \quad (R^2 = 0,998, n = 39; 50 < LL_c < 400\%) \quad (12)$$

Niazi et al. (2019) também apresentam uma revisão das correlações entre os valores de limite de liquidez para os métodos de Casagrande com o aparelho de base dura e do cone britânico. Os pesquisadores destacam que foram encontradas pelo menos 55 correlações entre esses dispositivos em 43 estudos diferentes, selecionando as correlações de 24 publicações em uma tabela. Assim, analisando 65 amostras de solos finos, obteve a seguinte correlação:

$$LL_p = 0,89 LL + 4,204 \quad (R^2 = 0,985) \quad (13)$$

Diferentes padrões adotam diferentes dispositivos e, portanto, correlações baseadas no limite de liquidez devem ser utilizadas com alguma cautela (Mitchell e Soga, 2005). Devido a variabilidade dos métodos utilizados em cada país, O'Kelly et al. (2018) destacam a importância da investigação de correlações entre os valores de LL obtidos pelos métodos de cone e de Casagrande, levando em conta as discrepâncias observadas. Cabe ressaltar que essas diferenças observadas entre os valores de LL terá enorme impacto em seu uso na classificação do solo e correlações com outras propriedades de engenharia do solo.

2.3.1 Influência do intervalo de LL

Os valores de limite de liquidez, plasticidade e o índice de plasticidade, tendem a crescer com a quantidade de argila presente no solo, de maneira que quando esses valores são elevados, os solos designam-se muito plásticos (Silva, 2013). Segundo Andrade et al. (2011), embora os limites de Atterberg sejam o método mais utilizado para avaliar a plasticidade, o grande número de variáveis envolvidas em sua determinação impede uma correlação detalhada dos parâmetros como o comportamento da argila.

Com o aumento da plasticidade do solo, a diferença entre os resultados dos métodos de cone e Casagrande torna-se sensivelmente maior (Sridharan e Prakash, 2000; Mishra et al., 2011). Para solos de menor plasticidade, o método do cone dá maior valor em relação ao método de Casagrande, enquanto para solos de maior plasticidade, o método de Casagrande apresenta valores maiores. (Mishra et al., 2011). Gronbech et al. (2011) observou que para argilas de plasticidade extremamente alta o método de cone subestimou o valor de LL em até 43 pontos percentuais.

O'Kelly et al. (2018) utilizando resultados de LL_p obtidos com o aparelho britânico e LL_c pelos aparelhos de Casagrande de base macia e dura, concluiu que os valores de LL_p são ligeiramente maiores para solos de LL baixo a intermediário e que para solos com valores de LL maiores ocorre uma forte divergência de valores entre os métodos.

Shimobe e Spagnoli (2019) analisaram os resultados de LL obtidos pelos cones britânico e sueco para 134 amostras de solos do Japão e concluíram que na faixa de LL menor que 120%, a correspondência entre os métodos de cone e de Casagrande é boa, enquanto para valores maiores, LL_p torna-se menor conforme LL aumenta. Assim, não há diferença entre os valores para uma faixa de LL menor que 100% e as diferenças aumentam conforme o limite de liquidez aumenta.

Manafi (2018) por sua vez obteve em seus resultados que para os valores de LL entre 25% e 50% os resultados obtidos pelo método do cone são mais altos que os obtidos pelo método de Casagrande e que o inverso ocorre a partir de valores entre 50% e 75%, sendo que esses intervalos aumentam conforme aumentam os valores de limite de liquidez.

Comparando os resultados obtidos para o limite de liquidez de 31 amostras de solo de diferentes locais do Brasil e diferentes propriedades geotécnicas, utilizando o cone sueco e o aparelho de Casagrande de base dura, Clemente et al. (2020), citam que para a maioria das amostras, percebe-se que os maiores valores de LL foram obtidos pelo método de Cone. O pesquisador observou ainda que quando os resultados são divididos em dois grupos, com LL inferior ou superior a 80%, para valores maiores que 80% existe uma maior dispersão dos resultados. A maioria dos solos testados apresentavam baixa atividade, sendo, portanto, necessário investigar mais a fundo a influência de diferentes minerais de argila nos resultados e correlações entre os dois métodos (Clemente et al., 2020).

Niazi et al. (2019) analisaram 65 amostras de solos finos coletadas de 32 condados da Indiana nos Estados Unidos, utilizando o cone britânico e o aparelho de Casagrande padrão ASTM, ou seja, de base dura, e objetivando minimizar a influência do operador nos ensaios, utilizaram um mesmo operador para a realização de todos os ensaios. Com base nos resultados dos ensaios os pesquisadores fizeram uma avaliação da classificação dos solos conforme o gráfico de plasticidade de Casagrande e concluíram que o limite de liquidez é subestimado para solos de alta plasticidade quando utilizado o método de Cone, e que para um valor de LL até cerca de 100%, os resultados entre os dois métodos se mostram bem parecidos.

Portanto, com base nas publicações analisadas, no geral, pode-se afirmar que para valores de limite de liquidez mais baixos, o método de Casagrande fornece resultados ligeiramente menores que os obtidos pelo método de Cone (Sridharan e Prakash, 2000; O'Kelly, 2018; Manafi, 2018; Karakan et al., 2020) e para valores mais altos, o método de Casagrande

fornece valores ligeiramente maiores que o de Cone (Niazi et al. 2019) para solos de média a alta plasticidade. Além disso, solos com limites de liquidez abaixo de 100% apresentam boa concordância entre os dois métodos (Wasti e Bezirci, 1986; Haigh, 2012), já que, segundo Wasti e Bezirci (1986), este é o valor aproximado esperado na maioria dos solos naturais.

O teste de limite e liquidez é uma forma de teste do cisalhamento dinâmico (Mitchell e Soga, 2005). Sridharan e Prakash (2000) hipotetizam que o principal componente da resistência ao cisalhamento que entra em ação no ensaio de percussão é a resistência ao cisalhamento viscoso, já que com o impacto com a base, a massa de solo de cada lado do sulco flui em direção ao centro para fechá-lo. E, que de maneira semelhante, o movimento do cone descendente mobiliza principalmente a resistência ao cisalhamento friccional entre as partículas do solo. Assim, segundo Shimobe e Spagnoli (2019) para as argilas de alta plasticidade, o método de Casagrande fornece valores mais confiáveis, já que esses solos possuem maior resistência ao cisalhamento viscoso principalmente devido à natureza viscosa da água de dupla camada. Para argilas de baixa plasticidade, porque são governados por resistências ao cisalhamento de atrito, o método do cone fornece dados mais confiáveis, pois não há cisalhamento viscoso mobilizado em valores baixos de umidade. Isso, segundo Sridharan e Prakash (2000) explica porque em solos de alta plasticidade os resultados obtidos pelo cone são inferiores ao de Casagrande e porque o inverso ocorre para os solos de baixa plasticidade.

Mishra et al. (2011), no entanto, discute que como a dissipação não drenada é mais fácil em solos com baixo limite de liquidez, a profundidade de penetração no caso do Cone e o abatimento do solo no ensaio de Casagrande, ocorrem com um teor de umidade comparável e resultam em um valor de LL quase igual.

Bicalho et al. (2017) citam que o intervalo de LL não é o único fator que define a diferença nos resultados de LL_c e LL_p ; o tipo e a quantidade do argilomineral e o tipo de cátion trocável também influenciam nas diferenças entre os valores obtidos pelos métodos de Cone e Casagrande.

2.3.2 Influência do tipo de Argilomineral

A mineralogia da argila apresenta um efeito notável na plasticidade. Sendo assim, é lamentável que seja ignorada na maioria das pesquisas que tratam de solos de granulação fina, embora a maioria dos problemas de campo associados a esses solos possam ser atribuídos à sua composição mineralógica argilosa (Prakash, 2012). Karakan et al. (2020) discutem que no passado vários estudos se preocuparam em determinar uma relação entre os índices de consistência e a resistência não drenada dos solos, no entanto, aponta para a

necessidade de estudos que relacionem as propriedades mecânicas e de consistência entre misturas de argilas e areias considerando os aspectos mineralógicos das argilas.

Segundo Sousa (2011), dado o apertado intervalo de variação na generalidade dos solos finos, a plasticidade não é tão representativa das características dos solos como o limite de liquidez. A variação nos valores de limite de liquidez entre diferentes grupos de minerais de argila é muito maior do que a variação no limite de plasticidade (Mitchell e Soga, 2005). Portanto, o comportamento do limite de liquidez é uma função dominante da mineralogia argilosa do solo (Prakash, 2012).

Mitchell e Soga (2005), apresentam os intervalos de LL e LP para os minerais de argila.

Tabela 2 – Intervalos de LL e LP para diferentes minerais de argila.

Mineral	LL (%)	LP (%)
Montmorilonita	100 - 900	50 - 100
Ilita	60 - 120	35 - 60
Caulinita	30 - 110	25 - 40

Fonte: adaptado de Mitchell e Soga, 2005.

Segundo Karakan et al. (2020), a mineralogia da argila apresenta um efeito notável na plasticidade, já que em seu estudo, os pesquisadores observaram que os índices de plasticidade de argilas contendo caulinita foram significativamente menores dos que os que continham montmorilonita.

Para as montmorilonitas propriamente ditas, alguma variação dos índices de plasticidade é esperada com substituições na rede cristalina, o que é geralmente ofuscado pelas diferenças no teor dos minerais argilosos (Bain, 1971). Segundo Bain (1970), o processamento comercial da argila bruta geralmente aumenta os valores de índice de plasticidade, o teor de argila, ou produzem um produto mais finamente dividido.

As argilas caulinita e montmorilonita possuem mecanismos diferentes que controlam o valor do limite líquido dessas argilas (Sridharan e Prakash, 2000): o LL de uma argila montmorilonita é controlado principalmente pela presença da camada de água adesiva presente na argila mineral, enquanto nos solos cauliniticos é controlado principalmente pelas forças entre as partículas do mineral argiloso (tração e repulsão). O limite de liquidez obtido pelo método de Casagrande por sua vez é uma medida predominante de resistência ao cisalhamento viscoso; enquanto para o método do cone, é uma medida predominante da resistência ao cisalhamento por fricção (Sridharan e Prakash, 2000; Mishra et al, 2011). Assim, Sridharan e Prakash (2000) destacam que o método do cone se adapta melhor aos solos cauliniticos e o método de Casagrande se mostra mais adequado para os solos montmoriloníticos.

Mishra et al. (2011) estudou a influência das propriedades mineralógicas e físicas de bentonita nos valores de LL para 12 misturas solo-bentonita preparadas na proporção 100:20. Os pesquisadores observaram que independentemente do tipo de bentonita os resultados obtidos pelos dois métodos são bastante idênticos para $LL < 116\%$, e que, para $LL > 116\%$, LL_c foi maior do que LL_p . Além disso, a diferença entre os dois métodos diminuiu com a diminuição de LL, fração de argila, superfície específica e capacidade de inchamento da bentonita presente na mistura. Mishra et al. (2011) discutem ainda que como a espessura da dupla camada difusa das bentonitas é influenciada por fatores como mineralogia e solução eletrolítica, os mesmos fatores podem influenciar o valor de limite de liquidez nas misturas.

Quintela et al. (2014) observou boas correlações entre o limite de liquidez, a quantidade de filossilicato e o tipo de argilominerais. Os pesquisadores analisaram 18 amostras de solos argilosos agrupados conforme a composição do tipo principal de argilomineral: caulinitas, montmorilonita, illita e um modelo contendo caulinita-montmorilonita. Os resultados indicaram que o LL aumenta com o incremento do teor de filossilicatos, quantidade de montmorilonita e área de superfície específica e, em contrapartida, o incremento no tamanho da partícula, a contribuição de minerais detríticos e o teor de caulinita se relacionam com um menor valor de LL.

Bicalho et al. (2017) avaliaram a influência da plasticidade e da mineralogia nas correlações empíricas entre LL_c e LL_p para as argilas naturais da Grande Vitória – ES e ensaios publicados de solos coesivos de baixa plasticidade de diferentes regiões do Brasil. Assim, observaram uma maior dispersão dos resultados entre LL_c e LL_p para misturas de uma argila bentonita com uma areia fina quartzosa, localizados próximo a linha U do gráfico de Casagrande, tal que LL_c é aproximadamente igual a $1,2LL_p$. As argilas naturais montmorilonitas com $LL_c < 90\%$, $LL_p < 85\%$ e $IP < 50\%$, localizadas próximo a linha A do gráfico de Casagrande, apresentaram menores diferenças entre LL_c e LL_p . Segundo Bicalho et al. (2017), para as misturas de bentonita e areia fina analisadas, o teor de argila, C, não influencia a variação entre LL_c e LL_p quando a matriz argilosa domina o comportamento da mistura, ou seja, não existe contato entre os grãos de areia ($C \geq 60\%$). E, neste caso predomina a influência da mineralogia da argila na relação entre LL_c e LL_p .

Crevelin (2018) por sua vez buscou avaliar a influência da mineralogia nas correlações entre LL_c e LL_p . A pesquisa mostrou que para solos essencialmente formados por caulinitas e illitas LL_p é maior do que LL_c em até 5%, e que essas diferenças são menores do que para solos com outras formações mineralógicas. Além disso, para solos argilosos formados essencialmente por montmorilonitas LL_c é maior do que LL_p em até 28% para LL_c entre 115 e 460%.

Crevelin e Bicalho (2019) analisando os resultados de solos argilosos de diferentes origens geológicas, para o método de cone utilizando o cone britânico e o de Casagrande utilizando os aparelhos de base dura e macia, concluíram que em solos essencialmente formados por caulinitas e ilitas, para baixos valores de LL, as diferenças entre os resultados dos dois métodos é menor do que para argilas motmorilonitas ou de alto LL, sendo o limite de liquidez obtido pelo método de Cone maior que o de Casagrande.

Vale ressaltar que, Mitchell e Soga (2005) citam que fatores como o grau de cristalinidade, tipo de cátions adsorvidos, ph, presença de matéria orgânica, tipo e quantidade de eletrólitos livres na água dos poros influenciam na variação das propriedades de engenharia das argilas.

2.3.3 Influência da Fração de Argila

O comportamento dos solos, é fortemente influenciado pela proporção de partículas grossas e finas, em peso (Karakan et al., 2020). Segundo Karakan et al. (2020), a classificação simplificada do comportamento dos solos, como coesivos e não coesivos, como na mecânica dos solos tradicional, resulta na simplificação da natureza complexa do comportamento do solo, o que pode ocasionar a negligência de aspectos mecânicos reais desses materiais.

Kayabali (2011) ressalta que enquanto a classificação dos solos de granulação grossa, como cascalho e areia, é geralmente realizada por uma análise de peneira, o agrupamento de solos de granulação fina como silte e argila é quase sempre realizado pelos limites de Atterberg. (Kayabali, 2011). Assim, o teor de argila e o teor de água tem uma relação (Widjaja e Kurniawan, 2020). Enquanto a resistência ao cisalhamento de solos de granulação grossa não é sensível à existência de água sob condições estáticas, para os solos finos, especialmente os argilosos, são fortemente influenciados pela presença de água. (Kayabali, 2011). Portanto, sendo os solos naturais formados de misturas de 2 ou mais minerais argilosos, constituintes em sua fração de argila, torna-se importante determinar o efeito dessas diferentes misturas nos resultados dos ensaios de limites de Atterberg (Freitas Júnior, 1974).

Entretanto, Karakan et al., 2020, afirma que antes de um certo teor de argila, os limites de Atterberg (LL e LP) e a fração de argila têm uma influência relativamente insignificante no comportamento geral dos solos. A composição do solo, o tipo de minerais de argila e as condições ambientais também causam resultados de correlação diferentes entre todos os solos (Widjaja e Kurniawan, 2020).

O limite de liquidez do cone pode ser maior que o limite de liquidez do método de Casagrande mesmo em altos teores de argila. Assim, parece que o teor de argila por si só não é responsável pela diferença observada (Sridharan e Prakash, 2000). Segundo Karakan et al.

(2020) vários mecanismos controlam os limites de liquidez e plasticidade das misturas argila-areia que são influenciados pela graduação da areia (bem ou mal graduada), teor de argila e origem da argila. Além disso, a capacidade para resistir à ruptura por cisalhamento nos dois métodos é desigual (Nini, 2014).

Sridharan e Prakash (2000) realizou uma série de ensaios em bentonitas, caulinitas e solos naturais afim de verificar sua teoria sobre a mobilização dos cisalhamentos viscoso (método de cone) e de fricção (método de Casagrande), bem como resultados da literatura, e observou que solos com menor teor de argila pode apresentar LL_c maior do que LL_p , se a proporção do argilomineral montmorilonita for apreciável e por outro lado, uma quantidade apreciável de caulinita pode apresentar LL_p maior que LL_c . Essas observações indicam que o tipo de mineral argiloso e sua proporção no teor de argila são responsáveis por um método dar um valor de limite de liquidez maior que o outro método, mas não apenas o teor de argila (Sridharan e Prakash, 2000). Segundo Widjaja e Kurniawan (2020), embora o teor de argila (C) também desempenhe um papel no aumento do limite de plasticidade e no índice de plasticidade, esse aumento não é tão significativo quanto o aumento de LL.

Nini (2014), discute que embora a água esteja mais associada à fração de argila dos solos, as frações de silte junto às de argila são as principais causas da consistência do solo, o que significa que, a contribuição das frações de silte e argila para essa consistência não é a mesma. Bain (1970) discute que o efeito da presença de silte e areia na plasticidade das argilas é de diluição, assumindo que os limites são reduzidos proporcionalmente à quantidade de material não-argiloso presente na amostra, e afirma ser portanto, conveniente a determinação dos limites de Atterberg apenas na fração de argila da amostra. Segundo Nini (2014), as frações de silte e argila são, portanto, os principais parâmetros que devem ser utilizados na comparação dos valores de limite de liquidez medidos pelos métodos de cone e de Casagrande.

A classificação de um solo como silte em vez de argila pode resultar em graves consequências para projetos geotécnicos críticos (Kayabali, 2011). Kayabali (2011) discute a utilização de solos passantes na peneira de nº 40, conforme aceito pela ASTM e norma britânica (BS 1377), para determinação dos limites de consistência de solos finos, quando a distinção entre os solos finos e grossos é feita pela peneira de nº 200. O pesquisador avaliou os efeitos da granulometria sobre os limites de consistência e comparou os resultados de limite de consistência obtidos com os materiais que passam na peneira de nº 40 e de nº 200. Os limites de liquidez obtidos pelo método de Casagrande para os solos que passam na peneira de nº 200 são de 10 a 20% maiores que os valores correspondentes para os solos que passam na peneira de nº 40 em pelo menos mais da metade das 60 amostras de solo investigadas. E

concluiu que, o aumento dos limites de liquidez e plasticidade para os materiais passantes na peneira de nº 200 causa mudanças notáveis tanto nas classes de solo quanto nos níveis de plasticidade. No geral, um terço das amostras ensaiadas mostraram uma mudança no tipo de solo, nível de plasticidade, ou ambos, quando a classe USCS do material de nº 200 é comparada com o material passante na peneira de nº 40. Essa mudança, segundo o pesquisador, torna-se bastante significativa quando a classe de um solo de grão fino está próxima das fronteiras. Assim, o pesquisador alerta para a possibilidade de uma quantidade significativa de material grosseiro inclusa nas amostras para realização dos ensaios de determinação dos limites de Atterberg, e cita que uma escolha mais racional seria apenas o uso do material de nº 200 para tal finalidade.

Rehman et al. (2020) avaliaram o efeito do tamanho das partículas em 28 amostras de solo da Venezuela, que passaram pelas peneiras de 2mm e 425 μ m. Os pesquisadores concluíram que, para amostras com partículas de tamanhos menores que 2mm, não foram observadas diferenças significativas para $LL < 35\%$, no entanto, houve maiores discrepâncias entre os dois tamanhos de partícula para amostras com $LL > 36\%$. As análises de regressão mostraram que as estimativas de LL para os dois tamanhos de partícula estavam fortemente relacionados, embora as discrepâncias observadas para valores de $LL > 36\%$.

Karakan et al. (2020) estudaram as relações entre os limites de Atterberg e a fração de argila para 44 misturas de argilas caulinitas e montmorilonitas com areias bem e mal graduadas em teores variando de 0 a 100%. Os pesquisadores observaram que com a adição de 10% de areia, as misturas de caulinitas têm seu limite de liquidez reduzido de 68,8% para 51,6% no caso das areias bem graduadas e para 55,3% para as mal graduadas. Para as argilas montmorilonitas essa diferença é de 142,6% para 125,6% para as areias bem graduadas e de 116,6% para as mal graduadas. Assim, destacam que a inclusão de uma pequena quantidade de areia altera significativamente a plasticidade das misturas areia-argila.

Widjaja e Kurniawan (2020) estudaram 4 amostras de solo com testes diretos de laboratório e 17 amostras de solo de uma pesquisa anterior em West Java, Indonésia. O LL ficou na faixa de 41,46% a 115,80% e a fração de argila variou de 12,24% a 66,18%. Os resultados mostraram que o aumento do teor de argila é acompanhado pelo aumento do limite de liquidez em uma relação linear. Os pesquisadores, no entanto, ressaltam que os locais de extração das amostras são únicos, porque os solos são influenciados pela atividade vulcânica, e que por esta razão, os solos podem ter um comportamento diferente de outros locais.

2.3.4 Influência do Método de Preparação das Amostras

Os métodos de preparação e tratamento da amostra, para a realização dos ensaios, pode conduzir os resultados a uma larga faixa de dispersão, conduzindo a interpretações errôneas, quanto ao comportamento dos solos (Freitas Júnior, 1974).

Um dos problemas principais nas preparações das amostras é a secagem prévia do material, principalmente para solos argilosos, com umidades naturais acima do limite de plasticidade (Freitas Júnior, 1974). A umidade do ar circundante também afeta os resultados do limite de líquido tanto do cone de queda quanto do dispositivo de limite de líquido Casagrande (Kestler, 1982). Os dias secos, em particular, fazem com que a amostra de solo seque com o tempo, o que, segundo Kestler (1982) resulta em um número sucessivamente crescente de golpes por execução ou leituras decrescentes do penetrômetro. Em certos solos, segundo Freitas Júnior (1974), os ensaios realizados a partir da amostra na umidade natural, podem alcançar valores significativamente maiores que os determinados a partir da amostra seca previamente. Secar o solo ao ar e reidratar antes do teste pode tender a diminuir o LL (Lee e Freeman, 2007).

O processo de secagem tradicional é realizado em forno de convecção (CO), e, a amostra de solo deve ficar no forno por tempo suficiente para atingir massa constante, geralmente levando várias horas, e de forma prática de um dia para o outro. Tendo em vista a conveniência para diversas atividades geotécnicas de se conhecer os limites de plasticidade e o teor de umidade dos solos de grão fino, o mais rápido possível, Mendonza e Orozco (2001) usaram a secagem das amostras no forno micro-ondas (MWO) como complemento ao ensaio de cone. Segundo os pesquisadores, com isso a produtividade e a eficiência dos laboratórios podem ser aumentadas, reduzindo consideravelmente o tempo de ensaio e com uma classificação geotécnica confiável que pode estar disponível em 1h.

Assim, Mendonza e Orozco (2001) concluíram que estatisticamente a correlação entre LL_p obtida com o cone sueco e LL_c com a secagem MWO é, para fins práticos, igual ao estabelecido entre as mesmas técnicas com secagem a CO. Por outro lado, é alta similaridade entre os teores de umidade obtidos com as duas técnicas de secagem. Portanto, os pesquisadores afirmam que o uso do MWO em testes de rotina é viável e confiável, mesmo com massas de solo relativamente pequenas. E, como os resultados são iguais em um curto tempo, conseqüentemente o uso da secagem MWO reduz custos.

Sabe-se que a eficácia da homogeneização e a tixotropia influenciam em certos solos, de maneira intensa, as interações entre os resultados de LL e a resistência ao cisalhamento da amostra (Freitas Júnior, 1974). Tanto para o método de cone como para o de Casagrande, a homogeneização é uma parte muito importante para encontrar o limite de liquidez. Gronbech et al. (2011) destaca que não homogeneizar a argila ou misturar ar na argila pode ter uma

grande influência no resultado do limite de liquidez. Claramente, uma amostra de solo mal misturada tem maior probabilidade de produzir resultados incorretos, independentemente do aparelho usado; portanto, o cone de queda não tem vantagem sobre o dispositivo de Casagrande e vice-versa neste caso (Kestler, 1982). Contudo, o processo de homogeneização é difícil e demorado, tornando-se um processo caro.

Gronbech et al. (2011) avaliou a influência da homogeneização e a relação entre os limites de liquidez de cone e de Casagrande para 33 amostras de argilas da Dinamarca de alta plasticidade. Com base na alta plasticidade da argila estudada, foi utilizado um processo especial de homogeneização continuada até que não fossem deixadas imperfeições visuais e o material tivesse uma superfície brilhante. Através dos testes, ficou evidente que o processo de homogeneização tem grande influência nos resultados, tornando-se muito importante ser metuculoso.

A água a ser acrescentada ao solo deve apresentar teor iônico semelhante ao da contida nos vazios da amostra, de forma a não alterar o comportamento químico-coloidal característico do material ensaiado (Freitas Júnior, 1974). Segundo Freitas Júnior (1974), em vista da dificuldade que tal processo introduzirá nas práticas rotineiras de laboratório, aconselha-se padronizar o uso de água destilada, que sendo inerte do ponto de vista iônico, apresentará, pelo menos, a vantagem de introduzir a mesma influência, o que não acontece com águas comuns cuja composição varia de lugar para lugar, formando soluções com distintos valores de pH.

2.4 Outras Correlações Geotécnicas

Uma série de propriedades de engenharia dos solos, como resistência ao cisalhamento não drenado, compressibilidade, permeabilidade, comportamento de expansão, área de superfície específica, capacidade de troca catiônica e liquefação têm correlações diretas ou indiretas com o valor do limite de liquidez (Ozer, 2009).

Os valores de limite de liquidez e plasticidade são, portanto, usados tanto para a classificação do solo quanto para determinação de parâmetros do solo por meio de correlações, conforme citado por alguns pesquisadores (Ozer, 2009; Shimobe e Spagnoli, 2019; Manafi, 2018).

A seguir são apresentados a relação entre o limite de liquidez e a estimativa dos parâmetros de resistência e compressibilidade obtidos a partir da literatura.

2.4.1 Uso de LL para estimar resistência

O limite líquido de um solo é definido como o teor de água correspondente a um valor de resistência ao cisalhamento não drenado de cerca de 2,5 kPa (Casagrande, 1932). Essa resistência diminui com o aumento do limite de liquidez (Crevelin e Bicalho, 2019; O' Kelly, 2021). Segundo O'Kelly (2021), o LL_c pode ser estimado a partir da umidade (w) tomando como valor médio a razão entre a resistência ao cisalhamento não drenada e o peso específico saturado (ρ_{sat}). Assim, segundo o pesquisador, o LL_c é dependente da estabilidade do declive dinâmico da pasta de solo ranhurada no dispositivo de Casagrande e, portanto, da força específica do solo (ou seja, S_u / ρ_{sat}), de modo que para solos de alta plasticidade as resistências são mais baixas. O limite de liquidez de Casagrande é uma medida da resistência específica (isto é, resistência não drenada dividida pela densidade) da argila em uma alta taxa de deformação (Haigh, 2012).

O teste de limite de liquidez tipo cone é uma medição direta da resistência (Haigh, 2012). Segundo O'Kelly (2021), o método de cone, está associado a uma resistência ao cisalhamento não drenada, dependente da massa do cone, rugosidade da superfície e características do ângulo de vértice e do valor de LL_p obtido. O'Kelly et al. (2018) citam que S_u varia ligeiramente entre os diferentes solos, dependendo de sua dependência da taxa de deformação, bem como das características físicas do cone e da profundidade de penetração. Isso porque, segundo o pesquisador, para o método do cone, a taxa de deformação muda continuamente conforme o cone acelera sob a gravidade, penetrando no corpo de prova e, em seguida, desacelera antes de descansar, com a taxa de deformação também dependente das características do cone. E, recomendam o uso mais amplo do método de cone para argilas moles a médio-duras no estabelecimento da variabilidade da resistência com a mudança do teor de umidade e outros parâmetros.

Clemente (2018) avaliou a resistência ao cisalhamento não drenada de solos de origem marinha obtidas por meio do método de cone. Segundo o pesquisador, a metodologia de ensaio utilizando o cone para a determinação de S_u é muito simples e fatores como o tempo de execução do ensaio e o cálculo são menores do que comparado a outros métodos. Clemente (2018) concluiu ainda que a massa do cone afeta os resultados, já que quanto maior a massa do cone, maior é o intervalo de penetração atingido durante o ensaio; entretanto, não há influência considerável da caracterização (percentual de finos, índice de plasticidade e sensibilidade) nos resultados de S_u .

Haigh (2012) cita que como muitas correlações foram obtidas com base no método do limite de liquidez de Casagrande, um método mais repetível de medir a resistência específica seria mais atraente. Segundo Claveau-Mallet et al. (2012), as principais vantagens do método de

cone (sueco) ao aparelho de Casagrande são a possibilidade de estudar muitos problemas relacionados com a resistência ao cisalhamento da argila intacta e remoldada, e a alegada melhor repetibilidade de seus resultados.

Karakan et al. (2020) citam que a estrutura mineralógica da argila e a distribuição do tamanho dos grãos tem um efeito notável na resistência ao cisalhamento não drenada (S_u), já que seus resultados apontaram para um valor de S_u maior em misturas de argila e areia bem graduadas se comparado com as misturas mal graduadas. Ainda de acordo com os pesquisadores, o valor de S_u , é independente das características geotécnicas das misturas e diminui linearmente com o aumento da profundidade de penetração do cone. E, as misturas contendo montmorilonita possuem maior S_u quando comparadas com as misturas que contém caulinita.

O índice de liquidez (IL) também é usado como um parâmetro de dimensionamento do teor de umidade para avaliação do comportamento do solo, bem como a variação da resistência ao cisalhamento (Niazi et al., 2019; O'Kelly et al., 2018). Karakan et al. (2020) em seus resultados, observaram que os valores de IL mais baixos, correspondem a maiores valores de S_u .

2.4.2 Uso de LL para estimar parâmetros de compressibilidade

As correlações empíricas que utilizam as propriedades índices, consideram o princípio que a compressibilidade do solo é resultante de fatores como a mineralogia e o estado de tensões (Filho, 2019). E, são de importância prática, especialmente para estudos de fundações, já que o índice de compressão (C_c) é usado para calcular o recalque de adensamento dos solos argilosos.

Segundo Bicalho et al. (2014), várias equações empíricas têm sido publicadas na literatura para estimar C_c para diferentes solos argilosos em função de propriedades índices como o limite de liquidez obtido pelo método de Casagrande (LL_c), índice de plasticidade (IP), índice de vazios inicial (e_0) e umidade natural inicial (w_{nat}). No entanto, as correlações comumente utilizadas baseiam-se no limite de liquidez pelo fato de que tal propriedade é facilmente determinada e a compressibilidade está relacionada com a superfície específica das partículas que as compõem e, portanto, com a sua mineralogia. (Giasi et al., 2003).

Cunha (2012) avaliou a relação entre os limites de Atterberg e o índice de compressão e seus resultados indicaram o índice de liquidez como um melhor parâmetro para estimar a compressibilidade dos solos, se comparado com o índice de plasticidade e o limite de plasticidade.

Bicalho et al. (2014) avaliou resultados de ensaios realizados em argilas naturais da Grande Vitória – ES e, por meio do método de regressão linear, obteve a seguinte correlação empírica simples para C_c em função de LL_c :

$$C_c = 0,011 (LL_c - 8,3) \quad (14)$$

Filho (2019) a partir de uma base de dados de propriedades índices avaliou a capacidade de previsão do índice de compressão (C_c) e da taxa de compressão (C_r) a partir de correlações empíricas publicadas na literatura e por correlações de ajustes em comparação com o uso de redes neurais artificiais. O autor obteve as seguintes correlações empíricas para estimar C_c e C_r a partir de LL_c .

Para solos de diversos países com $19 < LL_c < 177$:

$$C_c = 0,1265 LL_c - 0,2036 \quad (15)$$

Para solos brasileiros com $28 < LL_c < 140$:

$$C_c = 0,01875 LL_c - 0,3688 \quad (16)$$

Para solos de diversos países com $19 < LL_c < 112$:

$$C_r = 0,00319 LL_c - 0,01580 \quad (17)$$

Para solos brasileiros com $28 < LL_c < 135$:

$$C_r = 0,00423 LL_c - 0,0034 \quad (18)$$

Filho (2019) destaca que em alguns casos as correlações $LL_c \times C_c$ apresentam maior variabilidade nos resultados dentre as propriedades índices. Tal variabilidade, de acordo com o pesquisador, pode ser justificada pela complexidade da estrutura dos argilominerais e sua relação com o comportamento dos solos moles, já que LL_c é a variável mais dependente das características e mineralogias do solo se comparada com as demais. No entanto, para algumas amostras de solo LL_c se mostrou como a propriedade de melhor correlação com C_r .

Cabe ressaltar, que para as pesquisas citadas foram utilizados resultados dos ensaios obtidos pelo método de Casagrande, não sendo, portanto, avaliados os resultados de ensaios de limite de liquidez obtidos pelo método do cone. E, a existência de um vasto banco de dados com resultados de LL_p publicados na literatura, motivam o estudo de correlações entre C_c , C_r e LL_c .

2.5 Métodos Estatísticos

Segundo Santos (2017) a estatística na engenharia busca encontrar ou aperfeiçoar sistemas matemáticos capazes de descrever fenômenos da natureza, de modo que, após a identificação do problema a ser estudado, identificam-se os fatores determinantes que podem afetar os resultados, coletam-se dados por meio de experimentos e, por fim, refina-se e determina-se o melhor modelo que expresse o problema.

Assim, a estatística, segundo com Crevelin (2018), fornece uma estrutura para descrição da variabilidade e para aprender sobre quais fontes potenciais de variabilidade são mais relevantes.

As equações de correlação, por sua vez, segundo Crevelin (2018), são modelos empíricos que tendem a explicar a relação entre as variáveis aleatórias observadas em experimentos, levando-se em conta as fontes potenciais de variabilidade que as influenciam.

2.5.1 Etapas de determinação de uma correlação

Para definir uma correlação empírica entre parâmetros, o primeiro passo é definir a amostragem do banco de dados, que deve conter um número suficiente de observações (n) para garantir a representatividade do que está sendo estudado (Santos, 2017). Santos (2017), discute a relevância de um levantamento detalhado das características das amostras, considerações e limitações do método de amostragem, evitando assim obtenção de resultados limitados e tendenciosos.

Em seguida, o tratamento dos dados é realizado por meio de tabelas e gráficos. Nesta etapa são obtidas, portanto, as medidas estatísticas de dispersão dos dados x e y para analisar a sua variabilidade (Crevelin, 2018). Assim, as principais medidas estatísticas adotadas são:

- ✓ Variância (σ^2):

$$\sigma^2 = \frac{1}{n-1} \sum (i - \mu)^2 \quad (19)$$

- ✓ Desvio Padrão (σ)

$$\sigma = \sqrt{\sigma^2} \quad (20)$$

- ✓ Coeficiente de Variação (C_v)

$$C_v = \frac{\sigma}{\mu} \cdot 100 \quad (21)$$

O comportamento de uma variável dependente y como uma função de variáveis independentes x é usado para descrever os modelos de regressão (Chrystello, 2017). Assim,

os modelos de regressão, segundo Chrystello (2017), são usados para encontrar e descrever padrões de homogeneidade dentre a heterogeneidade em um conjunto de observações. Ainda segundo o pesquisador, a linearidade do modelo de regressão linear pode ser avaliada por meio do gráfico de dispersão entre as variáveis x e y , pelo exame do valor da estimativa do coeficiente de correlação de Pearson e a partir de um teste de hipótese.

De acordo com Santos (2017), a regressão linear é a ferramenta mais utilizada para se correlacionar duas ou mais variáveis, onde a e b são parâmetros constantes de ajustes, e e_i os resíduos, responsáveis por simular as aleatoriedades que provocam variações em y (variável dependente) independentes de x (variável independente ou regressor).

$$y = a + bx_i + e_i \quad (22)$$

A solução da equação é feita pelo Método dos Mínimos Quadrados, onde de acordo com Santos (2017), se minimiza a soma dos quadrados dos resíduos e_i e obtém-se os parâmetros por derivações, até chegar as expressões para as médias amostrais de x e y apresentadas a seguir.

$$\sum_{i=1}^n e_i = \sum_{i=1}^n (y_i - a - bx_i)^2 \quad (23)$$

$$a = \bar{x} - b\bar{y} \quad (24)$$

$$b = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (25)$$

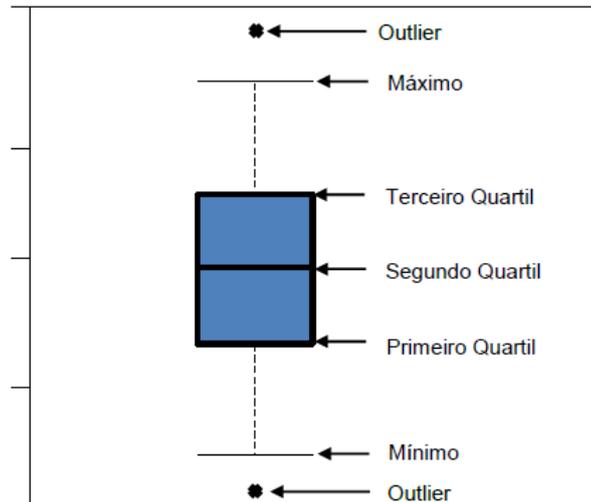
O método dos mínimos quadrados, de acordo com Chrystello (2017), é uma técnica de otimização matemática que busca encontrar o melhor ajuste para um conjunto de dados observados, tendo por objetivo definir o ajuste da função de regressão que minimize a soma das distâncias entre os pontos estimados e os pontos observados.

O diagrama de dispersão, por sua vez, é constituído por pontos onde cada ponto P_i representa um par de valores observados (x_i, y_i) , que posicionados em torno da reta de regressão tem a dupla função de analisar a existência de uma associação entre as variáveis e de possibilitar a identificação de qual o modelo matemático (equação) é mais apropriado para descrever essa associação. Portanto, quando o diagrama de dispersão indica uma tendência para uma relação linear, então os dados encontram-se bem ajustados pela reta de regressão (Crevelin, 2018).

A plotagem dos dados através do diagrama do tipo *boxplot*, permite identificar a presença de valores atípicos às observações, chamados de *outliers*. Os *outliers* podem ser ditos aqueles valores muito diferentes do restante do conjunto de dados (Reis e Reis, 2002). Em estatística,

a partir do retângulo para cima e para baixo, são delimitados os limites superior e inferior, respectivamente, tal que os valores fora desses limites são denominados de pontos exteriores e podem significar erro de amostragem.

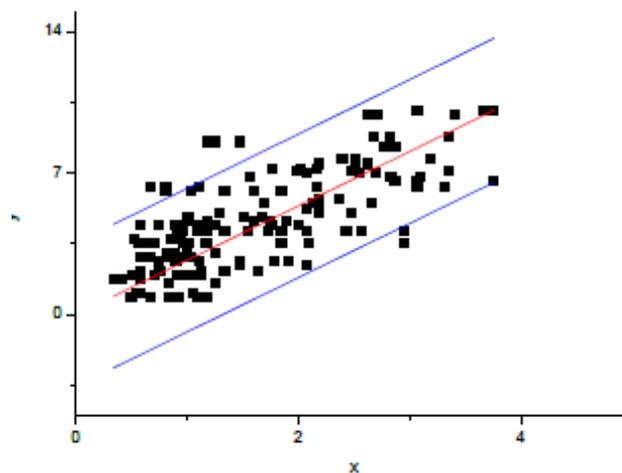
Figura 6 – Avaliação de outliers: boxplot



Fonte: Filho (2019)

No caso de x sendo variável contínua, os *outliers* são identificados pelo traçado de linhas inferiores e superiores no gráfico de dispersão, dadas pelo intervalo de confiança (Crevelin, 2018).

Figura 7 – Avaliação de outliers: limites superiores e inferiores no gráfico de dispersão.



Fonte: Adaptado de Santos (2017).

Portanto, os dados acima do máximo ou abaixo do mínimo de ambos os gráficos são os *outliers* e sua exclusão tem como objetivo eliminar as observações infieis, que podem ser ocasionadas por erros de amostragem, por exemplo (Crevelin, 2018; Santos, 2017).

2.5.2 Análise de Regressão

A análise de regressão consiste numa técnica para modelar e investigar a relação entre duas ou mais variáveis (Santos, 2017; Crevelin, 2018). Crevelin (2018), aponta que esta relação se faz necessária nos problemas no âmbito da engenharia. Para Santos (2017), consiste em etapa indispensável para validação de uma correlação, ou seja, quando se deseja verificar se o modelo é ou não adequado para expressar matematicamente os dados da amostra.

Esta análise, é feita rotineiramente pelo coeficiente de determinação (R^2), no entanto, o método mais correto é por inferência estatística com uso dos testes de hipótese (Santos, 2017). O coeficiente de determinação mede a fração da variabilidade das observações y pela aplicação da equação em novos valores de x, portanto, pode variar de 0 a 100% (ou de 1,0) (Santos, 2017; Crevelin, 2018). Assim, de acordo com Chrystello (2017) é uma das formas de avaliação da qualidade do ajuste do modelo de regressão linear. Dessa forma, quanto mais próximo de 1,0 estiver o coeficiente de determinação, melhor será o grau de explicação da variação de y em termos da variável x. Em outras palavras, o R^2 indica o quanto o modelo de regressão consegue explicar os valores observados, de modo que quanto maior o R^2 , mais explicativo é o modelo e melhor o modelo se ajusta à amostra (valores observados).

Os testes de hipótese fazem parte do ramo da estatística chamado de inferência, onde se procura afirmar certa condição sobre uma amostra (Santos, 2017). A inferência estatística de acordo com Crevelin (2018) consiste em utilizar dados amostrais para generalizar certa condição sobre os valores dos parâmetros populacionais, de modo que, em um processo de melhoria contínua, permite confirmar a eficácia das medidas de melhorias adotadas.

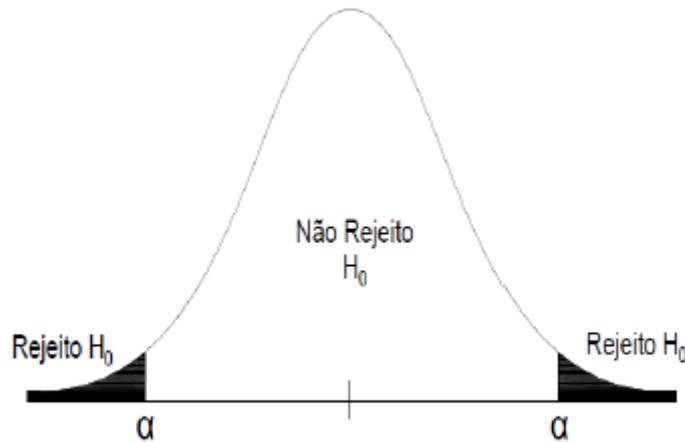
No teste de hipóteses a condição é escrita por duas hipóteses, nula e alternativa, H_0 e H_1 respectivamente, excludentes entre si. A conclusão, aceitar ou rejeitar H_0 , dependerá da comparação entre o resultado do teste e o valor crítico, formulado para cada teste de hipótese (Santos, 2017; Crevelin, 2018). No entanto, é prática comum em estatística a comparação do valor-p com o nível de significância (α) (Santos, 2017; Crevelin, 2018), conforme a equação 26.

$$\alpha = 1 - \text{Nível de Confiança} \quad (26)$$

Segundo Crevelin (2018), o p-valor é definido como a probabilidade de que a estatística do teste tenha valor extremo em relação ao valor observado quando a hipótese H_0 é verdadeira.

Assim, caso o valor-p seja superior ao nível de significância, aceita-se H_0 , caso contrário, rejeita-se H_0 e se aceita H_1 com o nível de confiança arbitrado (geralmente 95%) (Santos, 2017; Crevelin, 2018).

Figura 8 – Interpretação de Testes de Hipóteses.



Fonte: Adaptado de Santos (2017)

Segundo Santos (2017), uma vez que existe uma grande variedade de testes, a escolha deve ser feita de acordo com a aplicação e com o objetivo da análise, sendo o teste F, utilizado na análise da correlação como um todo e o teste t que avalia os parâmetros (a e b) separadamente.

Nestes casos, as hipóteses são:

- ✓ H_0 : o parâmetro não é significativo;
- ✓ H_1 : o parâmetro é significativo.

A estatística do teste t de Student, de acordo com Chrystello (2017), é igual à diferença entre a inclinação estimada a partir dos dados observados, e a inclinação dos dados observados, dividida pelo desvio padrão da inclinação.

Os resultados dos testes de hipóteses (Testes F e t), segundo Crevelin, 2018, podem ser avaliados pelo valor-p, assim, os parâmetros de regressão são significantes quando o valor-p obtido para os testes é menor do que o nível de significância de 5%.

2.5.3 Avaliação dos resíduos

A validação de uma correlação encontrada depende da análise dos resíduos, ou seja, avaliação dos erros de ajuste do modelo (Santos, 2017; Crevelin, 2018). Segundo Santos (2017), como se assume que os resíduos são independentes e possuem distribuição normal, ou seja, média zero e variância constante (homoscedasticidade), devem-se aplicar testes e/ou

análises gráficas para verificar se esses erros possuem essas características. Tais técnicas, de acordo com Crevelin (2018), podem ser gráficas ou através de testes, porém o ideal é que sejam combinadas já que as análises gráficas podem ser subjetivas e os testes indicados para tomada de decisão.

2.5.3.1 Análise Gráfica

Graficamente a normalidade pode ser testada por meio do Histograma, da Plotagem de Probabilidade Normal dos Resíduos e pela Plotagem dos Resíduos (Santos, 2017; Crevelin, 2018).

O histograma é um gráfico de barras justapostas em que no eixo horizontal está a variável de interesse dividida em classes e no eixo vertical a frequência da classe correspondente (Chrystello, 2017). Deve ser apresentado em formato de sino e simétrico e utilizado em casos com grande quantidade de amostras.

Já no gráfico de Probabilidade Normal, conforme destaca Crevelin (2018), os valores observados são plotados contra os valores esperados e caso o ajuste seja o modelo normal, se apresenta como uma reta se aproximando de uma diagonal.

No gráfico de Plotagem de Resíduos, por sua vez, é possível observar a variância de erros pela plotagem em x das variáveis independentes e em y os resíduos das variáveis dependentes correspondentes, sendo desejável que os dados se comportem em uma faixa horizontal de baixa amplitude. Já a observância dos dados nos demais formatos apresentados evidencia a heterocedasticidade (variância não constante) do modelo (Crevelin, 2018). Neste gráfico é possível avaliar a independência dos resíduos, já que quando os erros estão distribuídos conforme uma possível tendência (linear, parabólica, etc.), há um forte indicativo de que são dependentes entre si (Santos, 2018; Crevelin, 2018). Crevelin (2018) destaca que além de diagnosticar heteroscedasticidade, o gráfico também auxilia na detecção de *outliers*, de modo que se um *outlier* for influente, ele pode interferir a função de regressão.

2.5.4 Testes de Hipótese

Segundo Pereira et al. (2020), a análise de regressão também associa outros procedimentos de diagnóstico, como a análise de resíduos e o teste de independência de *Durbin-Watson*, por exemplo, que indica a consistência do modelo e proporciona a visibilidade de como melhorá-lo. Uma vez que análises gráficas podem ser subjetivas dependendo de quem as analisa, recomenda-se o uso de testes de hipótese para avaliar a normalidade dos resíduos (Santos, 2017). Dentre os testes de hipótese destacam-se o teste de aderência de

Kolmogorov-Smirnov, o teste de normalidade de *Shapiro-Wilk* e o teste de independência dos resíduos de *Durbin-Watson*.

O teste de aderência de *Kolmogorov-Smirnov*, segundo Crevelin (2018), é utilizado para observar a máxima diferença absoluta entre a função de distribuição acumulada, comparada com um valor crítico, com $1-\alpha$ de confiança, por meio das seguintes hipóteses:

H_0 : a distribuição é normal ($D_{KS} < D'_{KS}$);

H_1 : a distribuição não é normal ($D_{KS} > D'_{KS}$).

Onde DKS é o resultado do teste e D'_{KS} é o valor crítico. A estatística do teste, DKS, mede as distâncias verticais entre os gráficos das funções teórica e empírica $F(x)$ e $F_n(x)$, respectivamente (Crevelin, 2018).

$$D_{KS} = \max | F(x) - F_n(x) | \quad (27)$$

O valor de $F_n(x)$ é dado pela razão entre a posição do dado i e o valor total de dados, n . Já o valor de $F(x)$ é obtido na tabela da distribuição normal padrão em termos de $Z(i)$, onde s é o desvio padrão (Crevelin, 2018).

$$Z_i = \frac{x_i - \bar{x}}{s} \quad (28)$$

Tabela 3 – Valores críticos do Teste KS para 5% de significância.

Observações (n)	D'KS
5	0,56
10	0,41
15	0,34
20	0,29
25	0,27
30	0,24
35	0,23
40	0,21
45	0,20
50	0,19
Amostras Maiores	$1,36/\sqrt{n}$

Fonte: Adaptado de Crevelin, 2018.

O teste de normalidade de *Shapiro-Wilk* (SW), é dado por meio das seguintes hipóteses:

H_0 : a distribuição é normal ($D_{SW} > D'_{SW}$);

H_1 : a distribuição não é normal ($D_{SW} < D'_{SW}$).

A estatística do teste (D_{SW}) é dada pela equação 29, onde $a_{(i)}$ são constantes geradas a partir das médias, variâncias e covariâncias de n ordens, bem como os valores críticos do teste (D'_{SW}).

$$W = \frac{(\sum_{i=1}^n a_i x_i)^2}{(\sum_{i=1}^n x_i \bar{x})^2} \quad (29)$$

Tabela 4 - Valores críticos do Teste SW para 5% de significância.

n	D'_{sw}								
3	0,767	13	0,866	23	0,914	33	0,931	43	0,943
4	0,748	14	0,874	24	0,916	34	0,933	44	0,944
5	0,762	15	0,881	25	0,918	35	0,934	45	0,945
6	0,788	16	0,887	26	0,920	36	0,935	46	0,945
7	0,803	17	0,892	27	0,923	37	0,936	47	0,946
8	0,818	18	0,897	28	0,924	38	0,938	48	0,947
9	0,829	19	0,901	29	0,926	39	0,939	49	0,947
10	0,842	20	0,905	30	0,927	40	0,940	50	0,947
11	0,850	21	0,908	31	0,929	41	0,941		
12	0,859	22	0,911	32	0,930	42	0,942		

Fonte: Adaptado de Crevelin, 2018.

Vale ressaltar que o teste *Shapiro-Wilk* (SW) é utilizado em função do tamanho da amostra, assim, embora a tendência de melhores resultados para os testes de uma maneira geral para amostras grandes, neste estudo, o teste SW é limitado a amostras com tamanho entre 8 e 30 observações.

Para verificar a existência de autocorrelação nos resíduos da regressão utilizamos, frequentemente, o teste de *Durbin-Watson* (Hoffman, 2016). O teste de independência dos valores residuais de *Durbin-Watson*, segundo Crevelin (2018), é utilizado para avaliar a dependência sequencial entre os erros, ou seja, testa-se se a magnitude de um resíduo e_i influencia na magnitude do resíduo e_{i+1} , por meio das hipóteses:

H_0 : não existe autocorrelação dos resíduos

H_1 : existe autocorrelação dos resíduos

Portanto, a autocorrelação indica se os erros de uma mesma variável estão associados, apresentando uma determinada tendência, ou padrão de relacionamento.

A estatística do teste é dada pela equação seguinte:

$$d = \frac{(\sum_{i=2}^n e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (30)$$

O valor de d varia entre 0 e 4. Um valor próximo de 0 indica a existência de autocorrelação positiva nos erros e um valor próximo de 4 indica que os erros estão negativamente autocorrelacionados (Hoffman, 2016). Segundo Pereira et al. (2020), o resultado do teste de *Durbin-Watson* indica que não há multicolinearidade entre as variáveis independentes se o resultado se encontra entre 1,5 e 2,5. Ou seja, para um resultado próximo de 2, os erros não são autocorrelacionados.

A decisão pode ser tomada em função dos valores críticos, dL e dU , da Tabela de Durbin-Watson (Tabela 5).

Tabela 5 - Valores críticos do Teste DW para 5% de significância.

n	dL	dU	n	dL	dU	n	dL	dU	n	dL	dU
6	0,610	1,400	26	1,302	1,461	70	1,583	1,641	320	1,810	1,823
7	0,700	1,356	27	1,316	1,469	75	1,598	1,652	330	1,813	1,825
8	0,763	1,332	28	1,328	1,476	80	1,611	1,662	340	1,816	1,828
9	0,824	1,320	29	1,341	1,483	85	1,624	1,671	350	1,819	1,830
10	0,879	1,320	30	1,352	1,489	90	1,635	1,678	360	1,821	1,833
11	0,927	1,324	31	1,363	1,496	95	1,645	1,687	370	1,824	1,835
12	0,971	1,331	32	1,373	1,502	100	1,654	1,694	380	1,826	1,837
13	1,101	1,340	33	1,383	1,508	150	1,720	1,747	390	1,829	1,829
14	1,045	1,350	34	1,393	1,514	200	1,758	1,779	400	1,831	1,841
15	1,077	1,361	35	1,402	1,519	210	1,764	1,786	410	1,833	1,823
16	1,106	1,371	36	1,411	1,525	220	1,770	1,788	420	1,835	1,845
17	1,133	1,381	37	1,419	1,530	230	1,775	1,793	430	1,837	1,846
18	1,158	1,391	38	1,427	1,535	240	1,780	1,797	440	1,839	1,848
19	1,180	1,401	39	1,435	1,540	250	1,785	1,801	450	1,841	1,840
20	1,201	1,411	40	1,442	1,544	260	1,789	1,804	460	1,843	1,851
21	1,221	1,420	45	1,475	1,566	270	1,793	1,808	470	1,844	1,853
22	1,239	1,429	50	1,503	1,585	280	1,797	1,811	480	1,846	1,864
23	1,257	1,437	55	1,528	1,601	290	1,801	1,814	490	1,848	1,856
24	1,273	1,446	60	1,549	1,616	300	1,804	1,817	500	1,849	1,857
25	1,288	1,454	65	1,567	1,629	310	1,807	1,820	550	1,856	1,864

Fonte: Adaptado de Crevelin, 2018.

Nesse caso, o valor de d é comparado com dL e dU . Portanto, segundo Hoffman (2016), para erros positivamente autocorrelacionados, se $d < dL$, o resultado é significativo; se $d > dU$, o resultado não é significativo; e, se $dL < d < dU$, o resultado é inconclusivo. Já para a autocorrelação negativa, o resultado é significativo se $d > 4-dL$; não é significativo se $d > 4-dU$; e, será inconclusivo caso $4-dU < d < 4-dL$.

Segundo Crevelin (2018), em todos os testes citados, é mais usual a comparação do valor-p com o nível de significância. Assim, para o teste de *Durbin-Watson*, valor-p < 0,05, indica que os resíduos não são independentes. E, quando os testes de *Kolmogorov-Smirnov* e o teste

de *Shapiro-Wilk* apresentam valor- $p > 0,05$, segundo Pereira et al. (2020), sugere-se que deve ser aceita a hipótese nula e confirmada a normalidade da distribuição da variável dependente.

2.6 Redes Neurais

Redes neurais artificiais (RNAs) são uma forma de inteligência artificial que tenta imitar a função do cérebro humano e do sistema nervoso (Shahin et al., 2001, Moayedí et al., 2019, Totola, 2020). Esse processo ocorre por meio de um neurônio artificial que é responsável por reproduzir as experiências sinápticas de comunicação existentes entre os neurônios biológicos. Assim, os modelos de rede neural, segundo Moayedí et al. (2019), são sistemas de computação matemáticos inspirados na rede neural biológica, e, têm sido recentemente estendidas, apresentadas e aplicadas por muitos acadêmicos de pesquisa na área de engenharia geotécnica.

As técnicas de Inteligência Artificial (AI) são abordagens matemáticas desenvolvidas para buscar a solução mais adequada de acordo com determinados critérios dentro do tempo e recursos disponíveis (Edib, 2020). Segundo Filho (2019), as discussões no campo da inteligência artificial apontam para o potencial de aplicação da inteligência automatizada na solução de problemas complexos do cotidiano da sociedade. Cada técnica (AI) tem suas próprias características, vantagens e desvantagens e segundo Edib et al. (2021), essa variedade de técnicas de IA existentes (redes neurais artificiais, programação genética, máquinas de vetores de suporte, etc.) possibilita aos pesquisadores a resolução de um mesmo problema ao mesmo tempo em que descobrem a melhor técnica para sua resolução.

Nesse mesmo contexto, emerge a lógica do Aprendizado de Máquina (*Machine Learning*), um subconjunto da IA, que fornece aos sistemas a capacidade de aprender e melhorar automaticamente com a experiência, sem ser explicitamente programado (Shahin et al., 2001). Mais específico para a *Deep Learning* (DL), ou aprendizado neural profundo, este é um tipo particular de *Machine Learning* (ML) que, segundo Shahin et al. (2001) atinge grande poder e flexibilidade ao aprender a representar o mundo como uma hierarquia aninhada de conceitos, sem extrair recursos manualmente, um meio muito potencial para resolver problemas relacionados com muitos fatores incertos em geotecnia. Seguindo os avanços tecnológicos do estado da arte atual em IA, as técnicas de *Deep Learning* (DL) e *Extreme Learning Machines* (ELM) se destacam pela sua utilização em técnicas de aprendizagem, no entanto, seu emprego, conforme destacado por Filho (2019), é indicado quando se dispõe de uma base de dados ampla e um recurso computacional mais eficiente. Assim, para determinadas pesquisas, sugere-se o treinamento de redes neurais utilizando algoritmos que demandem menor esforço computacional e menor base de dados amostrais.

As Redes Neurais Artificiais são, portanto, uma das técnicas mais bem-sucedidas de Inteligência Artificial (IA) (Juwaied, 2018; Edib, 2020), e tem apresentado alta capacidade preditiva em comparação aos métodos tradicionais (Juwaied, 2018). Os modelos possuem sensibilidade para lidar com dados limitados, incompletos, incoerentes e até mesmo incorretos (Edib, 2020), e a eficiência da pesquisa depende tanto da técnica usada quanto das características do problema. Uma vantagem das RNA é a capacidade de aprender e generalizar, produzindo um modelo não-linear, com saídas adequadas para entradas inexistentes durante o processo de aprendizagem (Fleck et al., 2016). Segundo Moayed et al. (2019), as RNA oferecem ferramentas necessárias para engenheiros geotécnicos melhorar a qualidade, desempenho e reduzir riscos em projetos de engenharia.

2.6.1 Arquitetura de redes neurais artificiais

As redes neurais artificiais consistem em uma série de neurônios artificiais conhecidos como “elementos de processamento”, “nós” ou “unidades” (Shahin et al., 2008). Segundo Edib et al. (2021), a rede geralmente consiste em um número de neurônios para importar as entradas chamados de “camada de entrada” e um outro número de neurônios para exportar as saídas, chamados “camada de saída”, sendo que entre essas duas camadas pode haver um número de neurônios organizados em camadas ditas “camadas ocultas”, conectados por “links” que contém uma determinada força ou “peso”. São os pesos que, segundo Shahin et al. (2008), determinam a força das conexões entre os neurônios interconectados. Assim, cada neurônio atua como um elemento de processamento que soma os dados multiplicados pelos pesos dos links da camada anterior e aplica uma função não linear na soma para produzir os dados de saída para a próxima camada (Edib, 2020).

A determinação da arquitetura da rede, segundo Shahin et al. (2008), é uma das tarefas mais importantes e difíceis no desenvolvimento do modelo de RNA, já que requer a seleção do número ótimo de camadas e do número de nós de cada uma delas. Filho (2019), destaca a falta de consenso entre os pesquisadores quando a uma arquitetura ótima de uma RNA, já que devido a diversidade dos problemas que são apresentados às redes, com variáveis de entrada e saída distintos, as respostas particulares são obtidas em função da arquitetura e das funções de ativação utilizadas. Shahin et al. (2008) cita que não existe uma maneira direta e precisa para determinação do melhor número de nós de uma camada oculta, entretanto, o procedimento de tentativa e erro, geralmente usado na engenharia geotécnica, pode ser usado.

Fleck et al. (2016) sintetiza as redes neurais em três classes de arquitetura fundamentalmente diferentes: redes alimentadas adiante com camada única, redes alimentadas diretamente com múltiplas camadas e redes recorrentes, onde a sua definição é fundamental para a eficiência

do processo ao qual se destina. Já o processo em RNAs é geralmente dividido em supervisionado e não supervisionado. A principal diferença entre os modelos é a presença ou não de um supervisor externo. Segundo Shahin et al. (2008), na aprendizagem supervisionada, a rede é apresentada a um conjunto histórico de entradas do modelo e as saídas desejadas, sendo comparadas as saídas real e desejada para cálculo do erro que será utilizado para ajuste dos pesos e conseqüente redução deste erro. Na aprendizagem não supervisionada, a rede é apresentada apenas com os estímulos de entrada e não há saídas desejadas, já que, segundo Shahin et al. (2008), a própria rede é que ajusta os pesos das conexões de acordo com os valores de entrada, que neste caso são agrupados em classes de recursos semelhantes.

A rede neural *feedforward* (FNN) é o modelo básico e mais simples de rede neural artificial (Zhang et al., 2021). O processamento dos sinais ocorre de forma unidimensional de uma camada à outra até atingir a camada de saída (Totola, 2020). Segundo Totola (2020), as redes *feed-forward* são redes estáticas, que produzem apenas um conjunto de valores de saída para um determinado conjunto de entrada e a resposta independe do estado da rede anterior. O algoritmo de treinamento amplamente utilizado em FNN é o algoritmo de propagação reversa (*backpropagation*) (Zhang et al., 2021)

Na rede neural *backpropagation feed-forward*, segundo Sitton et al. (2017), o treinamento ocorre pelo ajuste dos pesos sinápticos de modo que os pesos são inicializados como números aleatórios para uma primeira passagem do conjunto de dados de treinamento. Assim, caso a saída não corresponda a meta aceitável predefinida, os pesos e tendências dentro da rede são ajustados minimizando uma função de erro. O treinamento é completo quando o erro entre as saídas simuladas e o alvo é suficientemente pequeno, em um sentido médio, ao longo de todo o treinamento e dados de validação. O principal tipo é chamado de *perceptron* que pode possuir uma (*perceptron simples*) ou múltiplas camadas (*perceptron multi-camadas*, MLP) (Totola, 2020).

As redes do tipo múltiplas camadas (*perceptron multi-camadas*, MLP), são constituídas por um conjunto de neurônios que formam a camada de entrada, uma ou mais camadas ocultas de neurônios e uma camada de saída (Shahin et al., 2008; Totola, 2010) e, segundo Shahin et al. (2008), são as RNA's mais comumente usadas em engenharia geotécnica. Com esse tipo de arquitetura, a camada oculta passa a utilizar a função de ativação não linear, com a função sigmoideal, conferindo à rede um modelo genuinamente não linear (Filho, 2019). Ainda segundo Filho (2019), diferentemente da *perceptron simples*, o erro não é obtido apenas da diferença entre a saída desejada e a saída calculada pela rede, já que existem camadas intermediárias, sendo o algoritmo de retropropagação (*backpropagation*) utilizado para

estimar o erro das camadas intermediárias a partir do efeito que estas provocam no erro da camada de saída.

Totola (2020) cita um outro tipo de redes neurais, as redes recorrentes (*recurrent networks*), que são ditas dinâmicas, já que nelas ocorre a modificação do estado da rede conforme novos padrões de entrada são apresentados. Se distinguem pela ocorrência de realimentação dos pesos, ou seja, os valores de saída de uma camada podem servir como dados de entrada da mesma camada ou de uma camada intermediária (Totola, 2020).

Outro algoritmo de destaque é o de Levenberg-Marquardt (LM) (Filho, 2019). Segundo Filho (2019), é uma técnica matemática para otimização do algoritmo *backpropagation* (incorporação de algoritmo para mínimos quadrados não lineares).

2.6.2 Aplicação de redes neurais artificiais

A aplicação de uma rede neural obedece à necessidade do uso de uma técnica de otimização de processamento de dados diversos, para a resolução de um problema complexo (Filho, 2019). Segundo Shahin et al. (2008), os modelos de RNA precisam ser desenvolvidos de maneira sistemática, abordando os principais fatores, como a determinação de entradas, divisão de dados e pré-processamento, escolha da arquitetura da rede, e a seleção cuidadosa de alguns parâmetros que controlam o método de otimização, critérios de parada e validação do modelo.

Por essa razão, os dados de entrada devem ser escolhidos com atenção, de forma a serem suficientemente representativos do domínio de interesse do problema a ser estudado, já que, segundo Shahin et al. (2008), um bom subconjunto de variáveis de entrada pode melhorar substancialmente o desempenho do modelo.

As fases de treinamento e teste da rede consistem em atividade de definição dos pesos adequados para melhorar o desempenho da RNA (Filho, 2019). Segundo Shahin et al. (2008), o conjunto de treinamento é usado para ajustar os pesos de conexão, enquanto o conjunto de teste é usado para verificar o desempenho do modelo em vários estágios de treinamento e para determinar quando interromper o treinamento, evitando assim, o ajuste excessivo.

Ao contrário dos modelos estatísticos convencionais, os modelos de RNA geralmente têm um grande número de parâmetros de modelo (pesos de conexão) e podem, portanto, ajustar os dados de treinamento (Shahin et al., 2008). Portanto, o treinamento, segundo Filho (2019), consiste na apresentação à RNA do conjunto de dados selecionado, de modo que a rede processe os dados de entrada e os multiplique pelos pesos sinápticos. Assim, é importante que os dados de treinamento sejam representativos de todo o domínio do problema processado (Filho, 2019). Assim, de acordo com Totola (2020), o pré-processamento dos

dados é feito para garantir que todas as variáveis de entrada recebam a mesma atenção durante o treinamento, sendo uma etapa fundamental da utilização das RNAs na Engenharia geotécnica. As saídas, por sua vez, segundo Filho (2019), são as respostas aos estímulos, que são comparados com os resultados previstos para a extração dos erros da fase de treinamento. A partir desse mecanismo de comparação, busca-se então ajustar os valores dos pesos sinápticos de ponderação das entradas, a partir de algoritmos matemáticos (Filho, 2019).

Portanto, como síntese do treinamento e teste, o propósito da validação, segundo Filho (2019), é assegurar que a rede possua a habilidade de generalizar o aprendizado obtido por meio da etapa de treinamento. O conjunto de validação é usado para estimar o desempenho da rede treinada no ambiente implantado (Shahin et al., 2008). Segundo Totola (2020), o conjunto de validação é utilizado para verificação do desempenho do modelo durante o treinamento, funcionando também como um critério de parada, já que quando o erro aumenta para o conjunto de validação, o treinamento é encerrado (*early stopping technique*). Ainda segundo Totola (2020), a importância da divisão dos dados se dá pelo problema do *overfitting*, que pode ocorrer em função da perda da capacidade de generalização da rede neural. Segundo Shahin et al. (2008), o coeficiente de determinação (R^2), a raiz do erro quadrático médio e o erro médio absoluto, são os principais critérios frequentemente usados para avaliar o desempenho de predição dos modelos de RNA

2.6.3 Aplicação de RNAs na engenharia geotécnica

A engenharia geotécnica envolve vários materiais (Shahin et al., 2001; Juwaied, 2018) com comportamentos distintos devido aos processos de formação desses materiais. Associada a sua essência variada, problemas de engenharia geotécnica também são dificultados pelos processos de amostragem e ensaio desses materiais, como forma de caracterizar e obter parâmetros para as análises de engenharia.

O custo e a dificuldade de realizar testes de campo para todos os tipos de projetos levam os projetistas a prever propriedades de engenharia geotécnica com o uso de testes de fácil aplicação e baixo custo (Arama et al., 2021). Modelos matemáticos para prever, simular e analisar o comportamento dos diferentes tipos de solos e rochas foram propostos ao longo dos anos, e continuam sendo estudados e melhorados, entretanto, a imprecisão e incerteza dos dados de entrada diminuem consideravelmente a precisão dos resultados e aumentam as incertezas envolvidas nos projetos.

Nos últimos anos, alguns estudiosos têm tentado gerar, estender e apresentar novas ferramentas e abordagens, métodos e técnicas de RNA no campo da engenharia civil

(Moayed et al., 2019). Alguns pesquisadores (Shahin et al., 2001; Juwaied, 2018) observaram que a partir dos anos 90 há cada vez mais a IA empregada como uma ferramenta eficaz na engenharia geotécnica. Segundo Edib (2020), nos últimos 35 anos o uso de técnicas de IA na engenharia geotécnica aumentou de uma pesquisa para cerca de 100 pesquisas por ano. Moayed et al. (2019) relatam que desde o início de 1900 até a data de publicação de sua pesquisa havia 152 artigos indexados na *web of science* que se referiam ao uso de RNA na área de engenharia geotécnica. E, o número de publicações, segundo os pesquisadores, aumentou de uma publicação no ano de 2000 para 22 publicações em 2017.

As redes neurais artificiais têm sido utilizadas na modelagem da deflexão de paredes de contenção, escavação, comportamento de solo, estruturas de contenção de terra, previsão da capacidade de suporte de estacas, avaliação da estabilidade de taludes, mapeamento de suscetibilidade a escorregamentos, na caracterização local, liquefação e na classificação de solos.

Shahin et al. (2001), apresenta um exame do estado da arte das RNAs em engenharia geotécnica e discute as direções da pesquisa atual que precisam de mais atenção no futuro. Os pesquisadores concluem que apesar das RNAs se mostrarem eficazes na engenharia geotécnica e em outras disciplinas, a técnica sofre algumas deficiências como a falta de transparência do modelo, limitação quanto à extrapolação dos dados e a falta de avaliação das incertezas.

Sitton et al. (2017) apresentou uma estrutura de rede neural artificial que processa dados de ensaios de campo qualitativos e quantitativos, para atribuir classificações dos solos com rapidez e precisão. Assim, segundo os pesquisadores, a precisão do sistema neural aumenta à medida que mais amostras de solo são analisadas e mais dados são fornecidos para treinar a rede. E, concluíram que as redes neurais são capazes de fazer uso de dados de análise qualitativa e quantitativa do solo obtidos em campo e são promissoras como ferramentas para pesquisas futuras.

Filho (2019), a partir de uma base de dados contendo 2.022 amostras de solos argilosos moles de 25 sítios geológicos ao redor do mundo, compostas pelas propriedades índices limite de liquidez (LL), umidade (w), índice de vazios (e) e índice de plasticidade (IP) avaliou a capacidade de previsão do índice de compressão (C_c) e da taxa de compressão (C_r) a partir de correlações empíricas publicadas na literatura e por correlações de ajustes em comparação com o uso de redes neurais artificiais. Os resultados obtidos indicaram que as redes neurais artificiais têm potencial de aplicação como alternativa às correlações empíricas na previsão de C_c e C_r para solos moles do litoral brasileiro e também para outros países.

Totola (2020), avaliou o desempenho do uso de RNAs para estimativa da curva de retenção de água no solo (CRAS) limite superior de secagem para a faixa de sucção de 0 a 1500 kPa utilizando 565 amostras do banco de dados hidrofísicos de solos no Brasil denominado *HYBRAS 1.0*. Os resultados indicaram as RNAs como uma ferramenta potencial para estimativa da CRAS limite superior de secagem ao longo de um amplo intervalo de sucção, sendo útil para estudos e projetos preliminares em solos não saturados.

Yousefour et al. (2021), propôs uma nova abordagem para a previsão da rigidez e resistência a compressão não confinada de solos orgânicos estabilizados por meio do método de mistura de solo úmido. Os pesquisadores analisaram e compararam dois algoritmos de redes neurais e o método de regressão linear, além disso, adotaram um método de análise de sensibilidade para investigar o impacto de cada variável de controle e identificar a faixa ideal para maximização da resistência e rigidez dos solos orgânicos. Como resultado, o método de regressão linear demonstrou um valor significativamente mais baixo de precisão. Além disso, a resposta gerada mostrou como as redes neurais poderiam explicar as relações não lineares entre as propriedades de solos orgânicos estabilizados e as variáveis de controle.

2.6.4 Uso de RNA para estimar limites de consistência de solos

Existem muitas publicações sobre o uso de técnicas estatísticas convencionais para desenvolver equações empíricas para prever os limites de Atterberg a partir de propriedades do solo (Zolfaghari et al., 2015). No entanto, nos últimos anos, novas técnicas de computação, como as redes neurais artificiais (RNA), têm sido empregadas como poderosos modelos preditivos para este fim.

Zolfaghari et al. (2015) avaliaram o uso de dados de solos do Irã para previsão de valores de LL usando modelos de RNAs. Os valores de avaliação de desempenho sugerem que a técnica de RNA pode ter um desempenho aceitável para a previsão dos limites de consistência dos solos investigados, principalmente, quando associados a outras propriedades do solo, índices topográficos e de sensoriamento remoto, utilizados como dados de entrada. Os pesquisadores observaram melhor precisão das previsões para os índices LL, LP e IP respectivamente, uma vez, que sendo dependente de LL e LP, o valor de IP mostra-se menos previsível. E, que a significância do teor de argila para as predições é uma relação não linear. Por fim, os pesquisadores concluíram que os modelos de RNAs fornecem estimativas aceitáveis dos limites de consistência em escala regional, e que além do tamanho das partículas, teor de argila, e do tipo de argila, o teor de matéria orgânica influencia muito os valores de limites de consistência.

Arama et al. (2021) investigaram uma base de dados de 1253 diferentes amostras de solos argilosos de alta plasticidade localizados na Turquia. Análises de regressão e técnicas de inteligência artificial foram utilizadas para obtenção dos valores de LP e IP a partir dos resultados de LL, objetivando uma relação direta entre LL e IP, e permitindo a eliminação da determinação de LP na classificação dos solos. Os valores de LL e IP foram considerados como entrada e saída de dados, respectivamente, realizando a pré-seleção entre diferentes parâmetros de solo (compactação, profundidade, conteúdo de finos, etc.). Foi utilizada uma estrutura de rede multicamadas e feed-forwarded, e o treinamento do modelo também foi realizado por meio do algoritmo de retropropagação de Levenberg-Marquardt. Os pesquisadores concluíram que a aplicação das RNAs dá resultados mais precisos que as expressões empíricas obtidas por regressão para a determinação do valor do índice de plasticidade dos solos argilosos investigados.

Edib et al. (2021) por sua vez, utilizaram as técnicas de programação genética (PG) e RNA para prever os limites de consistência, LL e LP de um solo não saturado tratado com um ligante conhecido como cimento híbrido. Os pesquisadores avaliaram 121 amostras de solo que foram divididas entre dados de treinamento (81 amostras) e dados de validação (40 amostras). Foi utilizado uma RNA simples de retropropagação, sem camadas ocultas e função de ativação linear, para comparar a precisão das técnicas. Embora a simplicidade da configuração da RNA, segundo os pesquisadores, o modelo foi capaz de determinar as relações entre as entradas e as saídas com precisão. Os resultados de predição mostraram que técnica de RNA superou a técnica de programação genética em termos de avaliação de desempenho conduzida por métodos estatísticos que utilizaram a soma do erro quadrático (RMSE) e o coeficiente de determinação (R^2).

Diaz et al. (2021) utilizaram diferentes algoritmos de aprendizado de máquina (ML) para correlacionar LL_p e LL_c levando em consideração outros parâmetros previamente determinados, como LP e granulometria. Os dados investigados incluem solos de diversas regiões e valores de plasticidade. A pesquisa incluiu a avaliação indireta do efeito da areia fina e das porções de finos nos valores de LL_c e LL_p . Para a obtenção de LL_p foi utilizado o cone britânico enquanto para o LL_c um aparelho de Casagrande de base dura. Os resultados demonstraram o potencial das técnicas de aprendizado de máquina para relacionar os valores de LL obtidos pelo método de Casagrande e do cone em solos de finos de baixa a média plasticidade, principalmente para valores de $LL > 30\%$ para os quais abordagens clássicas de regressão linear fornecem piores ajustes de desempenho.

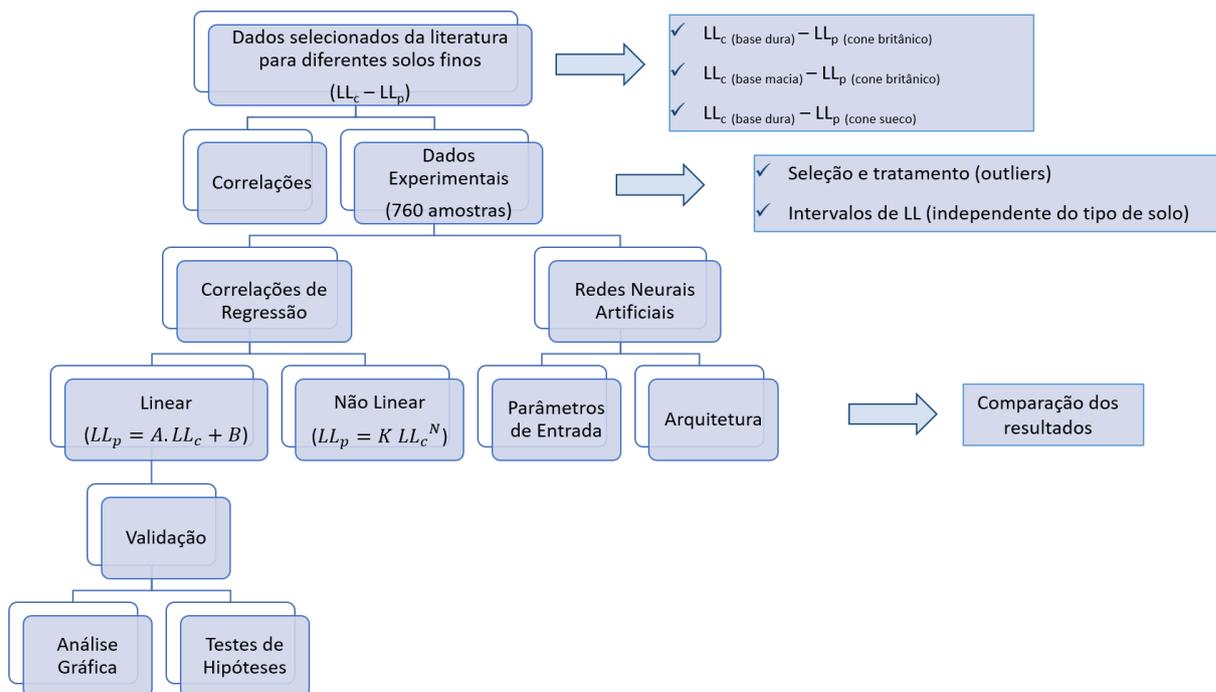
3. MATERIAIS E MÉTODOS

Este capítulo apresenta a base de dados organizada para a pesquisa e utilizada para as análises estatísticas e de redes neurais artificiais e em seguida expõe a metodologia empregada na execução de cada uma destas etapas.

É descrita, portanto, a metodologia para o estabelecimento de correlações de regressão, considerando as ferramentas computacionais utilizadas, os métodos de análises gráficas e testes de hipóteses a que os dados foram submetidos. Também é apresentada a metodologia utilizada no desenvolvimento e treinamento das RNAs e da descrição das características dos modelos adotados. E, descrita a metodologia de avaliação e comparação dos resultados.

A Figura 9 apresenta um fluxograma da metodologia empregada na pesquisa.

Figura 9 – Fluxograma da pesquisa.



Fonte: Autor (2022)

Os itens seguintes contêm a descrição de cada uma das etapas elencadas no fluxograma da Figura 9.

3.1 DESCRIÇÃO DO BANCO DE DADOS

Para atingir o principal objetivo desta pesquisa, a elaboração do banco de dados contempla a compilação de resultados experimentais de ensaios de laboratório de amostras de solos argilosos de diferentes locais (Brasil, Egito, Equador, Estados Unidos, Grécia, Itália, Índia, Portugal e Turquia).

A base de dados foi compilada a partir de pesquisas publicadas previamente na literatura brasileira e internacional. Portanto, para a composição do banco de dados utilizados nesta pesquisa foram compilados as correlações e os resultados de ensaios de LL_c e LL_p obtidos das publicações consultadas. No entanto, foram utilizados nas análises estatísticas e de redes neurais artificiais apenas os resultados experimentais de publicações que especificaram a dureza da base do aparelho de Casagrande e o tipo de cone. As pesquisas que não continham tais informações foram descartadas pois, esta padronização, tende a garantir uma maior confiabilidade dos resultados. Neste estudo foram avaliados os métodos do cone de penetração sugeridos pelas normas britânica e sueca e os aparelhos de Casagrande de base macia e dura. Outras informações como o local de origem da amostra de solo, intervalo de LL , o tipo de solo ensaiado, o tipo de argilomineral predominante, e a granulometria, quando informados, também foram registradas neste trabalho.

3.1.1 Correlações Publicadas entre LL_c e LL_p

Ao longo dos anos vários pesquisadores (Wasti, 1987; Di Matteo, 2012; Mishra et al., 2011, Spagnoli, 2012, Bicalho et al., 2017; Sousa, 2011; Crevelin e Bicalho, 2019) têm sugerido correlações entre LL_p e LL_c . Estas funções, quando lineares, podem ser escritas como (Bicalho et al., 2014):

$$LL_p = A.LL_c + B \quad (31)$$

$$LL_c = A^*. LL_c - B^* \quad (32)$$

Onde A, B, A* e B* são constantes de ajuste obtidas através da correlação linear.

Já as correlações não lineares podem ser escritas como:

$$LL_p = K LL_c^N \quad (33)$$

Onde K e N são constantes de ajuste, que variam sistematicamente conforme o material e o tipo de dureza da base do aparelho de Casagrande.

As correlações compiladas da literatura foram inicialmente separadas de acordo com a dureza da base do aparelho de Casagrande e do tipo de cone utilizado, formando quatro grupos:

- Grupo B-01: Base dura do aparelho de Casagrande e Cone Britânico;
- Grupo B-02: Base macia do aparelho de Casagrande e Cone Britânico;
- Grupo B-03: Base dura do aparelho de Casagrande e Cone Sueco;
- Grupo B-04: Base macia do aparelho de Casagrande e Cone Sueco.

Nesta pesquisa, as correlações foram apresentadas com os valores de LL_p em função de LL_c . E, organizadas conforme os intervalos de LL, a dureza da base do aparelho de Casagrande, o tipo de cone, o tipo de solo e sua origem, quando informados. Ao todo 64 correlações foram compiladas de 53 publicações e classificadas conforme os grupos a que pertencem. Cabe ressaltar que não foram observadas correlações para o grupo B-04.

3.1.1.1 Correlações do grupo B-01

Na Tabela 6 são apresentadas 37 correlações lineares catalogadas de 31 publicações da literatura pertencentes ao grupo B-01.

Tabela 6 – Algumas correlações lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-01.

Número	Referências	A	B	Varição de LL_c (%)	Tipo de Solo	Nº Ensaio s	R ²	Origem Solo
1	Sowers et al. (1960)	0,790	8,390	10-130	Solos naturais micáceos	21	-	Georgia, USA
2	Wires (1984)	0,940	0,970	38-55	Argilas	4	-	-
3	Belviso et al. (1985)	0,990	4,290	34-134	Argilas naturais	16	0,984	Sul da Itália
4	Sampson and Netterberg (1985)	1,010	4,200	20-120	Solos naturais	43	0,986	África do Sul
5	Wasti e Bezirci (1986)	1,010	4,920	27-110	Solo natural e misturas solo-bentonita	25	0,928	Turquia
6	Wasti (1987)	0,587	32,320	145-526	Misturas Bentonitas e solos naturais	10	-	-
7	Koester (1992)	0,955	4,410	≤600	Solos naturais de grãos finos e misturas comerciais de areia fina	27	-	Misissipi, Estados Unidos
8	Campos et al. (1999)	0,880	8,130	24-72	Argilas caulinitas naturais	15	0,960	Paraíba, Brasil
9	Feng (2001)	0,940	2,600	25-76	Solos naturais	66	-	Mundiais
10	Mendoza e Orozco (2001)	0,832	13,280	30-350	-	9	0,989	-
11	Mendoza e Orozco (2001)	1,066	1,840	30-100	-	6	0,971	-
12	Li (2004)	1,010	0,680	30-455	Solos naturais	19	-	Taiwan, China
13	Orhan et al. (2006)	1,020	2,130	27-98	Solos naturais	16	-	Turquia
14	Dragoni et al. (2008)	1,020	2,870	28-74	Solos naturais argilosos	41	0,980	Centro da Itália
15	Fojtová et al. (2009)	1,000	2,440	20-51	Solos naturais finos	52	0,978	Bacia de Ostrava, República Tcheca
16	Ozer (2009)	0,900	6,040	29-104	Solos naturais	32	0,996	Turquia

Número	Referências	A	B	Varição de LL _c (%)	Tipo de Solo	Nº Ensaio S	R ²	Origem Solo
17	Cevallos Luna (2012)	1,118	-6,308	19-85	Solos naturais	36	0,985	Quito, Equador
18	Spagnoli (2012)	0,990	1,050	20-61	Caulinita e Illita	50	0,990	Alemanha
19	Mishra et al. (2011)	0,842	11,686	70-150	Misturas de Bentonitas	12	0,992	Japão
20	Di Matteo (2012)	1,000	2,200	24-40	Misturas de areia fina, silte e argila	6	0,980	Sul da Itália
21	Silva (2013)	1,050	0,610	38-45	Illita e Caulinita	10	0,479	Região de Chaves em Portugal
22	Kayabali et al. (2016)	1,000	2,000	28-166	Solos naturais e misturas com bentonita	275	0,990	Vários locais da Turquia
23	Kollaros (2016)	1,030	-4,930	53-75	Solos naturais e misturas solo-cal	6	0,990	Grécia
24	Couto et al. (2016)	1,020	8,870	28-64	Siltes e Argilas	7	0,811	Minas Gerais
25	Matusiewicz et al. (2016)	0,910	3,200	≤150				
26	Bicalho et al. (2017)	1,000	2,700	14-100	Caulinita	5	0,980	Argilas naturais da Grande Vitória
27	Bicalho et al. (2017)	0,900	-9,000	>100	Misturas Bentonita e areia fina	4	0,980	Argilas naturais da Grande Vitória
28	El-Shinawi (2017)	0,910	5,640	28-70	Argilas naturais	40	0,974	Distrito de Cairo-Suez do Norte, Egito
29	Crevelin e Bicalho (2019)	0,985	1,346	20-100	Caulinitas e ilitas	76	0,967	Mundiais - obtidos de diversas publicações
30	Crevelin e Bicalho (2019)	0,739	14,521	50-460	Montmorilonitas	39	0,984	
31	Rehman et al. (2020)	0,924	3,460	20-80	Argila	41	0,990	Mundiais Amostras de Indiana nos Estados Unidos.
32	Niazi et. al (2019)	0,890	4,200	11-55	Solos naturais	65	0,985	
33	Karakan et.al (2020)	0,927	5,522	10-150	Mistura Caulinita e areia bem graduada	9	-	Turquia
34	Karakan et.al (2020)	0,885	2,575	10-150	Mistura Caulinita e areia mal graduada	9	-	Turquia
35	Karakan et.al (2020)	0,774	11,656	10-150	Mistura Cálcio-Montmorilonita e areia bem graduada	9	-	Turquia
36	Karakan et.al (2020)	0,807	7,091	10-150	Mistura Cálcio-Montmorilonita e areia mal graduada	9	-	Turquia

Número	Referências	A	B	Varição de LL _c (%)	Tipo de Solo	Nº Ensaio S	R ²	Origem Solo
37	Diaz et al. (2021)	0,841	2,881	17-63	Solos de granulação fina com plasticidade baixa a média	113	0,990	Espanha

Fonte: Autor (2022)

A correlação proposta por Di Matteo (2012), além dos dados experimentais citados, inclui os resultados de LL de 94 solos naturais publicados na literatura por Campbell (1975), Belviso et al. (1985), Wasti & Bezirci (1986), Dragoni et al. (2008), Ozer (2009), Fojtová et al. (2009) e Zentar et al. (2009).

As correlações propostas por Bicalho et al. (2017), além dos dados das amostras das Argilas da Grande Vitória e das Misturas de bentonita e areia fina, incluem os resultados experimentais publicados por Pinto & Castro (1971) e Silveira (2001) para solos tropicais brasileiros de diferentes locais.

Crevelin e Bicalho (2019) para a correlação apresentada para os solos que possuem a caulinita e a illita como minerais predominantes utilizou os resultados de 76 amostras de solos publicadas na literatura por Di Matteo (2012), Spagnoli (2012), Verástegui-Flores e Di Emidio (2014), Quintela et al. (2014), e Bicalho et al. (2017). Já a correlação apresentada para o grupo que possui as montmorilonitas como mineral principal foi obtida a partir do resultado de 41 amostras de solos publicadas na literatura por Wasti & Bezirci (1986), Deka et al. (2009), Mishra et al. (2011), Verástegui-Flores e Di Emidio (2014), e Bicalho et al. (2017).

O'Kelly et al. (2018) utilizaram 368 resultados experimentais obtidos da literatura para estabelecer quatro correlações empíricas não lineares entre os métodos de cone e de Casagrande para diferentes padrões. Na Tabela 7 são apresentadas 2 correlações não lineares publicadas por O'Kelly et al. (2018) pertencentes ao grupo B-01.

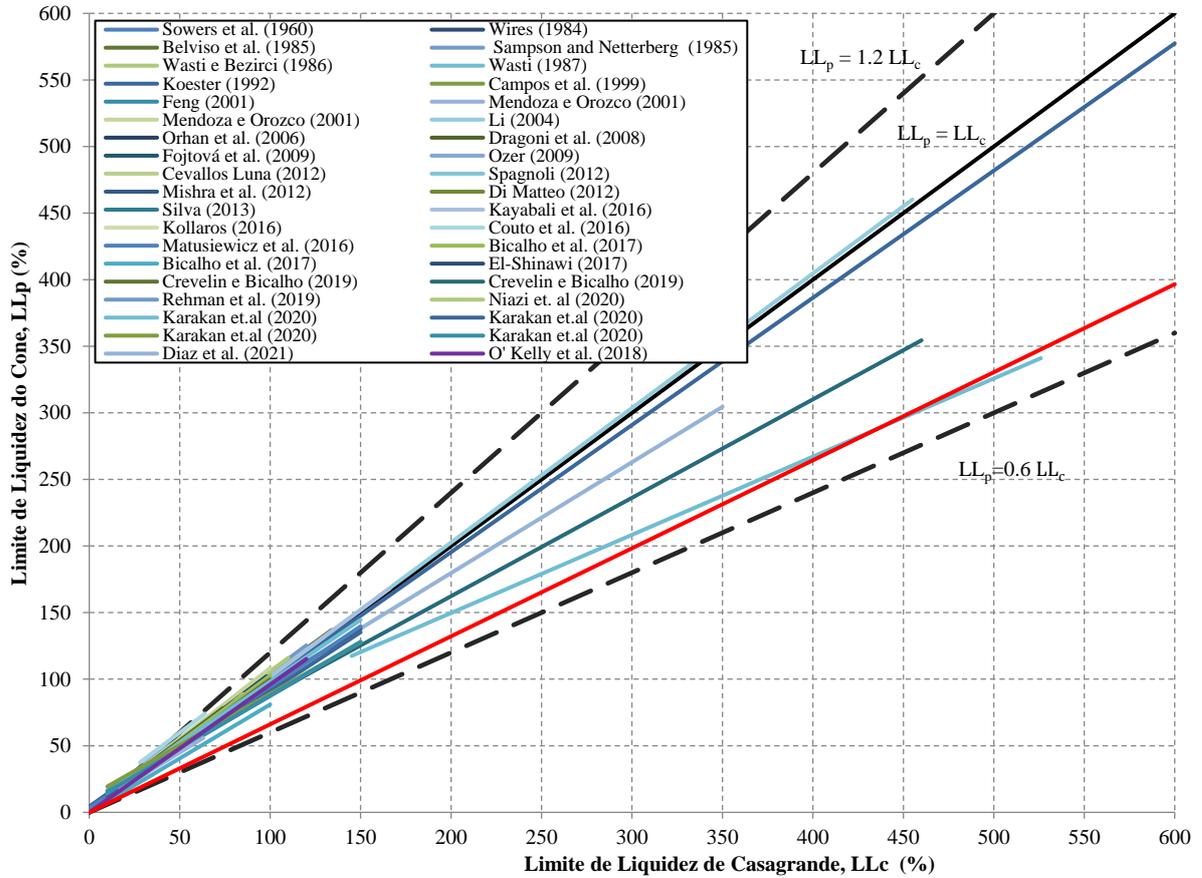
Tabela 7 – Algumas correlações não lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-01.

Número	Referências	A	B	Varição de LL _c (%)	Tipo de Solo	Nº Ensaio	R ²	Origem Solo
1	O' Kelly et al. (2018)	1,450	0,920	<120	Solos naturais e	188	0,970	Mundiais
2	O' Kelly et al. (2018)	1,900	0,850	<600	manufaturados	199	0,970	

Fonte: Autor (2022)

As correlações lineares e não lineares pertencentes ao grupo B-01, são apresentadas ainda na Figura 10. As equações encontradas variam no intervalo de $LL_p=0,6LL_c$ e $LL_p=1,2LL_c$.

Figura 10 - Correlações publicadas para o grupo B-01.



Fonte: Autor (2022).

3.1.1.2 Correlações do grupo B-02

Na Tabela 8 são apresentadas 14 correlações lineares catalogadas de 12 publicações da literatura pertencentes ao grupo B-02.

Tabela 8 – Algumas correlações lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-02.

Número	Referências	A	B	Variação de LL _c (%)	Tipo de Solo	Nº Ensaios	R ²	Origem Solo
1	Sherwood and Ryley (1968)	0,909	2,730	25-90	-	-	-	-
2	Sherwood and Ryley (1970)	0,950	0,950	30-72	Argilas arenosas, siltosas e orgânicas; caulinitas	25	-	Inglaterra e África
3	Littleton and Farmilo (1977)	0,970	1,600	20-450	Argilas arenosas e argilas	19	-	Mundiais
4	Budhu (1985)	0,935	4,020	20-100	-	-	-	-
5	Moon and White (1985)	1,163	-5,095	≤150	-	-	-	-

Número	Referências	A	B	Varição de LL _c (%)	Tipo de Solo	Nº Ensaios	R ²	Origem Solo
6	Queiroz de Carvalho (1986)	1,070	2,334	13-48	-	-	-	-
7	Sridharan e Prakash (1998)	0,827	8,933	29-92	-	-	-	-
8	Sridharan e Prakash (2000)	0,824	8,196	33-92	Caulinitas	5	-	-
9	Ozer (2009)	0,900	6,040	29-104	Solos naturais	32	0,996	Turquia
10	Sousa (2011)	1,414	-10,077	38-40	Caulinitas e Ialitas	16	-	Península de Setúbal, Portugal
11	Sousa (2011)	1,104	-1,235	42-45	Caulinitas e Ialitas	16	-	Mundiais - obtidos de diversas publicações
12	Crevelin e Bicalho (2019)	0,907	6,891	30-70	Caulinitas e Ialitas	28	0,767	Mundiais - obtidos de diversas publicações
13	Crevelin e Bicalho (2019)	0,822	7,265	50-400	Montmorilonitas	39	0,998	Mundiais - obtidos de diversas publicações
14	Shimobe e Spagnoli (2019)	0,950	-0,850	< 150	Argilas, siltes, solos vulcânicos e orgânicos	156	0,962	Amostras do Japão

Fonte: Autor (2022)

A correlação proposta por Crevelin e Bicalho (2019) para os solos que possuem a caulinita e a ilita como minerais predominantes utilizou os resultados de 28 amostras de solos publicadas na literatura por Sridharan e Prakash (2000), Grabowska-Olszewska (2003), Sousa (2011) e Nagaraj et al. (2012). Já a correlação apresentada para o grupo que possui as montmorilonitas como mineral principal foi obtida a partir do resultado de 39 amostras de solos publicadas na literatura por Sridharan e Prakash (2000), Grabowska-Olszewska (2003) e Nagaraj et al. (2012).

Na Tabela 9 são apresentadas 2 correlações não lineares publicadas por O'Kelly et al. (2018) pertencentes ao grupo B-02.

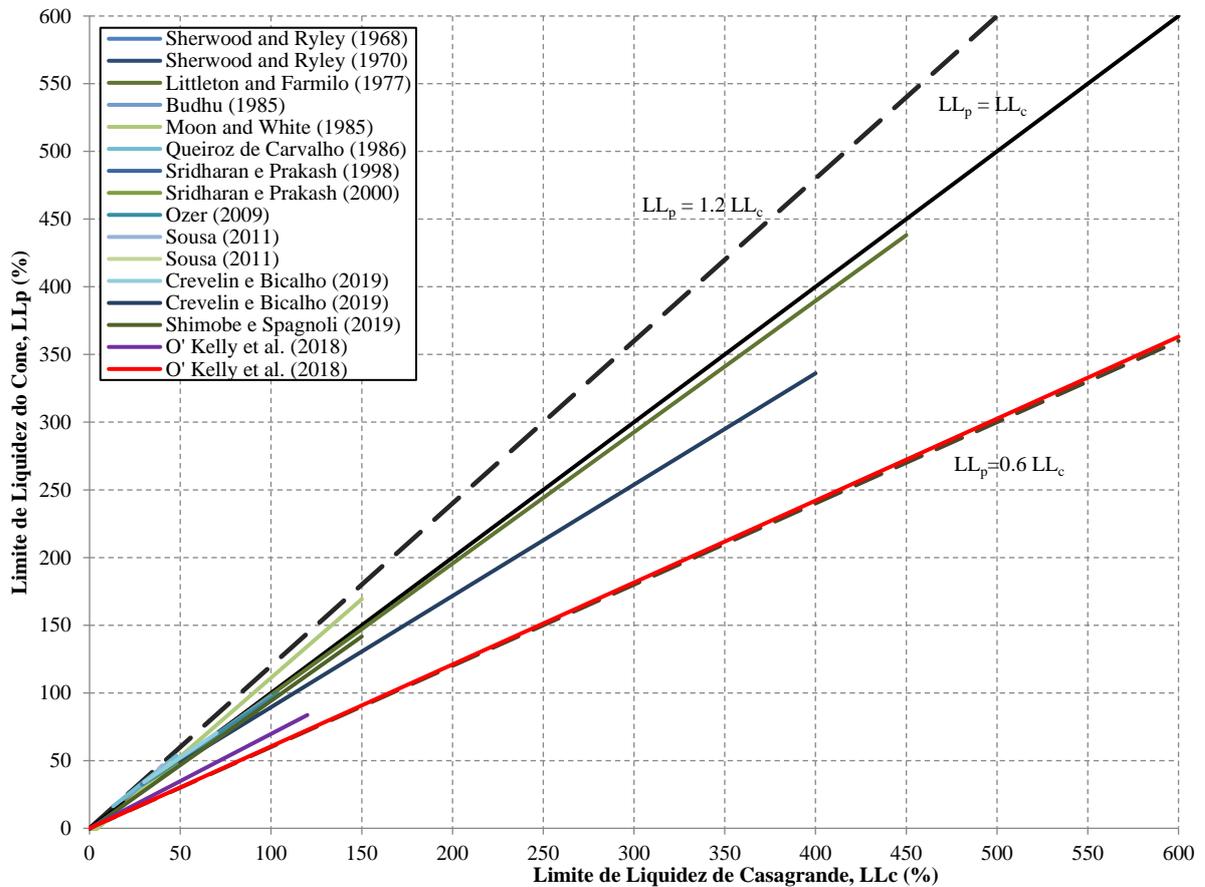
Tabela 9 – Algumas correlações não lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-02.

Número	Referências	A	B	Varição de LL _c (%)	Tipo de Solo	Nº Ensaios	R ²	Origem Solo
1	O' Kelly et al. (2018)	1,620	0,880	<120	Solos naturais e	199	0,960	Mundiais
2	O' Kelly et al. (2018)	1,860	0,840	<600	manufaturados	216	0,980	Mundiais

Fonte: Autor (2022)

As correlações lineares e não lineares pertencentes ao grupo B-02, são apresentadas ainda na Figura 11. As equações encontradas variam no intervalo de $LL_p=0,6LL_c$ e $LL_p=1,2LL_c$.

Figura 11 - Correlações publicadas para o grupo B-02.



Fonte: Autor (2022).

3.1.1.3 Correlações do grupo B-03

Na Tabela 10 são apresentadas 9 correlações lineares catalogadas de 8 publicações da literatura pertencentes ao grupo B-03.

Tabela 10 – Algumas correlações lineares previamente publicadas obtidas por diferentes pesquisadores e formulações apresentadas no grupo B-03.

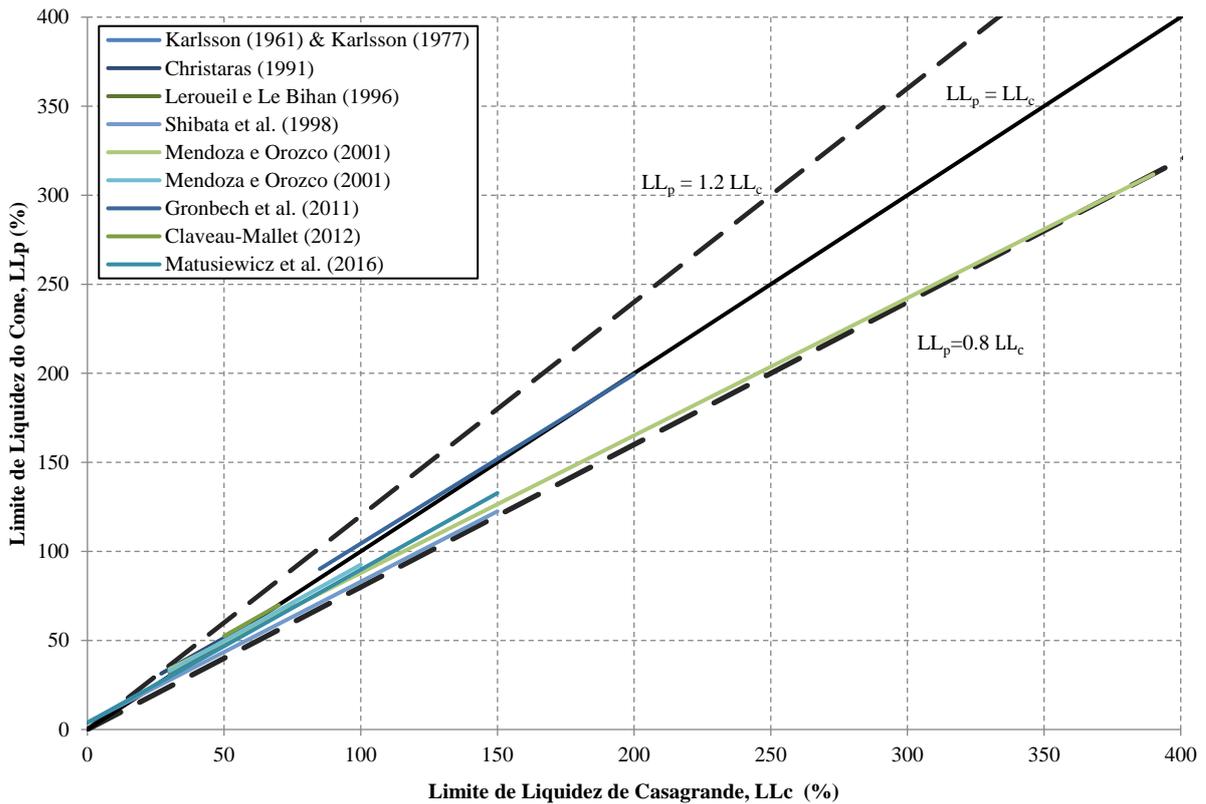
Número	Referências	A	B	Varição de LL _c (%)	Mineralogia Solo	Nº Ensaio	R ²	Origem Solo
1	Karlsson (1961) & Karlsson (1977)	0,850	5,020	30-76	Argilas remoldadas	91	-	Suécia
2	Christaras (1991)	0,870	7,930	27-59	Argila de plasticidade intermediária a alta	24	0,992	Creta, Grécia
3	Leroueil e Le Bihan (1996)	0,860	6,340	30-74	Siltos argilosos, argilas marinhas siltosas e argilas lacustres	43	-	Canadá

Número	Referências	A	B	Varição de LL _c (%)	Mineralogia Solo	Nº Ensaí os	R ²	Origem Solo
4	Shibata et al. (1998)	0,790	4,000	≤150	-	-	-	-
5	Mendoza e Orozco (2001)	0,772	10,710	30-390	Caulinitas, Bentonitas, Misturas	33	0,9934	México
6	Mendoza e Orozco (2001)	0,855	7,100	30-100	-	22	0,9874	-
7	Gronbech et al. (2011)	0,950	9,400	85-200	Argilas de alta plasticidade	33	0,952	Dinamarca
8	Claveau-Mallet (2012)	0,870	8,980	50-70	Argila marinha	8	-	Quebec
9	Matusiewicz et al. (2016)	0,860	3,750	≤150	-	-	-	-

Fonte: Autor (2022).

Para o grupo B-03 não foram catalogadas correlações não lineares. As correlações lineares pertencentes a este grupo, são apresentadas ainda na Figura 12. As equações encontradas variam no intervalo de $LL_p=0,8LL_c$ e $LL_p=1,2LL_c$.

Figura 12 - Correlações publicadas para o grupo B-03.



Fonte: Autor (2022).

De uma maneira geral, as correlações lineares, assim como nos grupos B-01 e B-02, estão localizadas entre as linhas de $LL_p = 1,2 LL_c$ e $LL_p = 0,8 LL_c$. No entanto, observa-se uma maior tendência em se aproximarem da linha e $LL_p = 0,8 LL_c$, exceto a correlação publicada por Gronbech et al. (2011) que se aproxima da linha $LL_c = LL_p$.

3.1.2 Comparações entre LL_c e LL_p

Para as análises e comparações entre os valores de LL_p e LL_c e para a determinação de correlações entre esses resultados, foram compilados também da literatura os resultados de ensaios realizados em amostras de solos argilosos.

Embora o extenso número de correlações publicadas e de pesquisas sobre o tema, os resultados dos ensaios não são informados por todas as publicações. Portanto, os dados utilizados para as comparações entre os valores de LL_p e LL_c não são coincidentes com as publicações utilizadas para o levantamento das correlações compiladas da literatura no item 3.1.1. Tratam-se, portanto, de bancos de dados distintos.

Os resultados experimentais dos ensaios foram classificados de acordo com a publicação, tipo de aparelho de Casagrande, tipo de cone, tipo e mineralogia do solo, quantidade de ensaios e o local de origem das amostras e são apresentados nos itens seguintes. Estes resultados foram separados de acordo com a dureza da base do aparelho de Casagrande e do tipo de cone utilizado, definindo quatro grupos:

- Grupo B-01: Base dura do aparelho de Casagrande e Cone Britânico;
- Grupo B-02: Base macia do aparelho de Casagrande e Cone Britânico;
- Grupo B-03: Base dura do aparelho de Casagrande e Cone Sueco;
- Grupo B-04: Base macia do aparelho de Casagrande e Cone Sueco.

De maneira análoga ao item 3.1.1, não foram encontrados na literatura pesquisas que apresentassem os resultados de ensaios de LL_c e LL_p para os equipamentos descritos no grupo B-04.

Dos resultados experimentais previamente compilados da literatura foram selecionados para as análises estatísticas e redes neurais artificiais, RNAs, os que apresentam informação complementar dos valores de IP. Os valores de IP das amostras de solo dos grupos B-01, B-02 e B-03 possibilitaram a inclusão destes resultados experimentais no Ábaco de Casagrande.

O Ábaco de Casagrande é frequentemente utilizado para a classificação dos solos de granulação fina. O gráfico relaciona os valores de índice de plasticidade (IP) no eixo das ordenadas e os valores do limite de liquidez (LL) no eixo das abcissas. O gráfico de Casagrande foi desenvolvido com base nos valores de LL obtidos pelo método de Casagrande. Portanto, conforme abordado no item 2.1.3, a utilização de valores de LL_p pode comprometer a classificação dos solos. Entretanto, para a avaliação da mudança na classificação SUCS dos solos quando utilizado os resultados do limite de liquidez obtido pelo

método do cone, os resultados de LL_c e LL_p para os grupos B-01, B-02 e B-03 desta pesquisa foram plotados no ábaco de Casagrande.

Os valores de IP, assim como a classificação dos solos pelo SUCS, foram utilizados como parâmetro de entrada para as algumas análises das RNAs.

Posteriormente, para a avaliação da influência do intervalo de LL, estes grupos foram divididos conforme os subgrupos:

- Subgrupo i: $LL < 50$;
- Subgrupo ii: $50 < LL \leq 120$;
- Subgrupo iii: $LL > 120$.

3.1.2.1 Descrição dos dados do Grupo B-01

O grupo de dados experimentais B-01 contempla os resultados experimentais compilados da literatura que utilizaram o cone no padrão britânico e o aparelho de Casagrande de base dura para a realização dos ensaios.

Foram investigados no grupo B-01 os resultados de ensaios realizados em 601 amostras de solo, compilados de 22 publicações: Wasti e Bezirci (1986); Diniz (1994); Silveira (1998); Campos et al. (1999); Silveira (2001); Deka et al. (2009); Fojtová et al. (2009); Ozer (2009); Kayabali (2011); Di Matteo (2012); Mishra et al. (2011); Spagnoli (2012); Cevallos Luna (2012); Pinto et al. (2012); Flores e Emidio (2014); Quintela et al. (2014); Nini (2014); Couto et al. (2016); Bicalho et al. (2017); El-Shinawi (2017); Theron et al. (2019); Niazi et al. (2019).

A Tabela 11 apresenta um resumo dos resultados experimentais de LL_p e LL_c investigados no grupo B-01, e a Figura 13 os resultados dos dados avaliados.

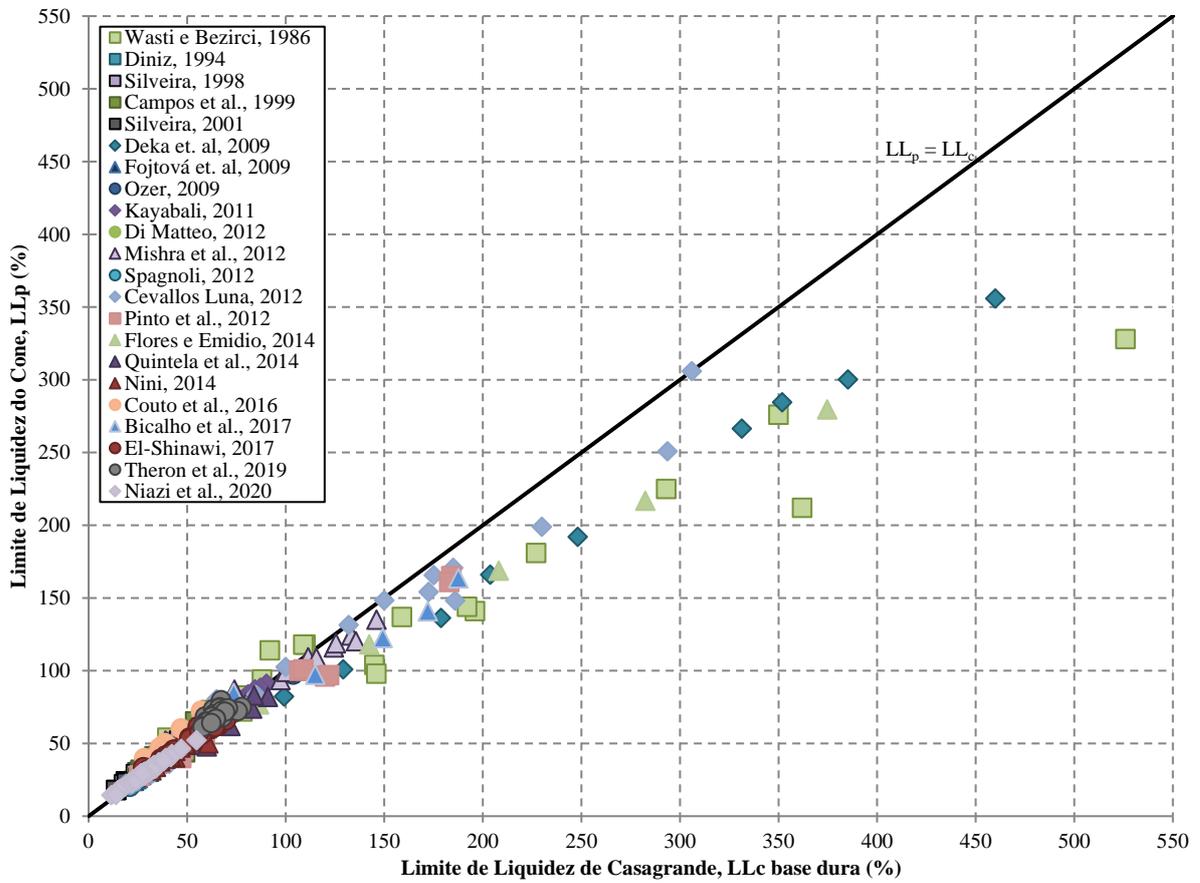
Tabela 11 - Dados experimentais compilados da literatura (grupo B-01).

Referências	Variação de LL _c (%)	Tipo de Solo	Nº Amostras	Granulometria			Origem Solo
				%média Argila	% média Silte	%média Areia	
Wasti e Bezirci (1986)	27-526	Solo natural e misturas solo-bentonita	25	53,92	-	-	Turquia
Diniz (1994)	22-44	Montmorilonita, ilita, quartzo, magnetita	11	49,18		43,00	Paraíba, Brasil
Silveira (1998)	14-46	Solos Tropicais	16	-		-	Brasil
Campos et al. (1999)	24-72	Argilas caulinitas naturais	15	-	-	-	Paraíba, Brasil
Silveira (2001)	14-40	Solos Tropicais	5	-	-	-	Brasil
Deka et al. (2009)	45-460	Montmorilonita, ilita, quartzo, magnetita	11	50,50	27,50	22,00	-
Fojtová et al. (2009)	20-51	Solos naturais finos	52	-	-	-	Bacia de Ostrava, República Tcheca
Ozer (2009)	29-104	Solos naturais	32	20,55	53,16	23,18	Turquia
Kayabali (2011)	42-91	Argilas e siltes	119	-	-	-	-
Cevallos Luna (2012)	19-306	Solos naturais	36	69,01		30,42	Quito, Equador
Spagnoli (2012)	19-61	Caulinita e Ilita	50	100,00	-	-	Alemanha
Mishra et al. (2011)	77-147	Misturas de Bentonitas	12	-	-	-	Japão
Di Matteo (2012)	25-38	Misturas de areia fina, silte e argila	6	23,83	54,00	22,18	Sul da Itália
Pinto et al. (2012)	26-184	Misturas de caulinita e ilita	16	88,75	-	11,25	-
Flores e Emidio (2014)	53-375	Mistura de caulinita e bentonita	6	100,00	-	-	-
Quintela et al. (2014)	39-91	Caulinitas	14	55,36	38,00	8,55	Portugal
Couto et al. (2016)	28-58	Siltes e Argilas	6	12,17	52,32	32,85	Minas Gerais
Bicalho et al. (2017)	14-100	Caulinita	5	68,59		28,23	Argilas naturais da Grande Vitória
Bicalho et al. (2017)	>100	Misturas Bentonita e areia fina	4	-	-	-	Argilas naturais da Grande Vitória
El-Shinawi (2017)	287-70	Argilas naturais	40	36,78	46,41	16,90	Distrito de Cairo-Suez do Norte, Egito

Referências	Variação de LL _c (%)	Tipo de Solo	Nº Amostras	Granulometria			Origem Solo
				%média Argila	% média Silte	%média Areia	
Theron et al. (2019)	57-78	Solos finos naturais	26	-	-	-	África do Sul
Niazi et al. (2019)	11-55	Argilas e siltes	65	24,07	48,99	26,94	Estados Unidos

Fonte: Autor (2022)

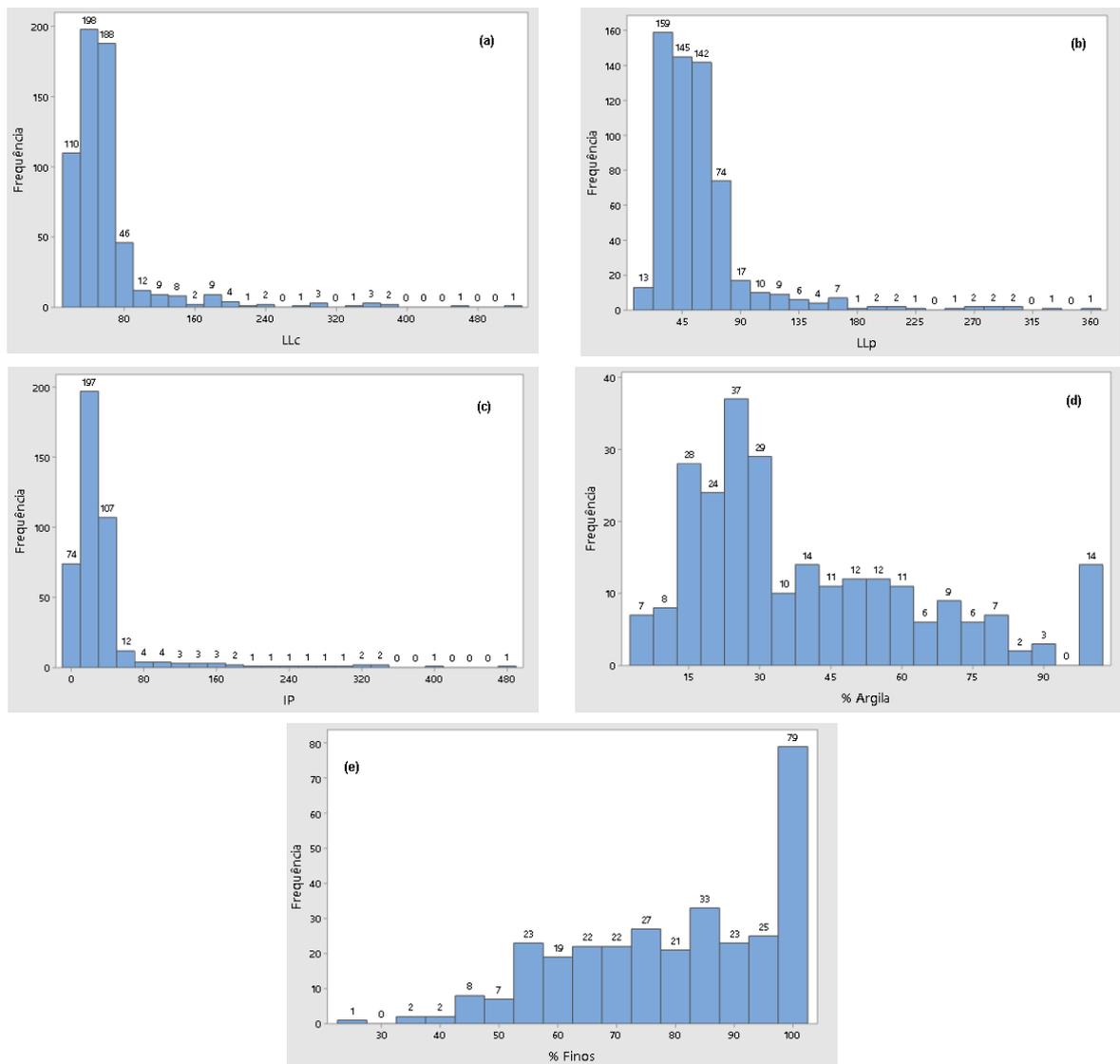
Figura 13 – Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do grupo B-01.



Fonte: Autor (2022)

A Figura 14 apresenta os histogramas com a distribuição dos dados do grupo B-01.

Figura 14 – Histograma dos dados do grupo B-01: (a) LL_c; (b) LL_p; (c) IP; (d) porcentagem de argila; (e) porcentagem de finos; (f) classificação SUCS.



Fonte: Autor (2022)

A descrição estatística dos dados do grupo B-01 é apresentada na Tabela 12. Por meio da tabela observa-se que do total de 601 amostras, apenas 421 possuem a informação de dados de IP, 250 a porcentagem de argila e 314 a porcentagem de finos.

Tabela 12 - Descrição estatística dos dados do grupo B-01 (N=601).

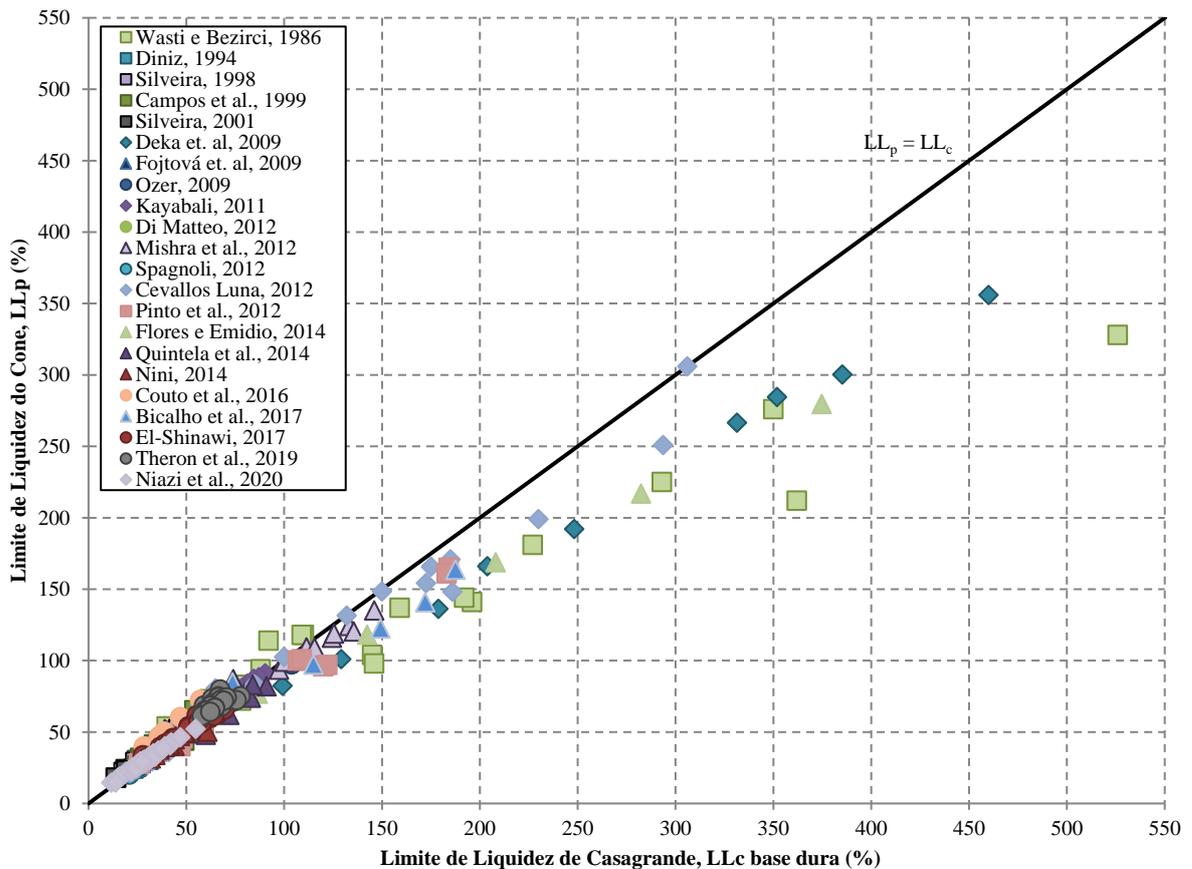
Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Máximo
LL _c	601	0	60,44	54,93	3017,63	90,89	11,60	526,00
LL _p	601	0	59,17	41,95	1759,48	70,89	14,50	355,91
IP	421	180	35,25	54,56	2976,28	154,76	0,00	488,00
% Argila	250	351	40,05	24,86	617,92	62,07	3,40	100,00
% Finos	314	287	79,911	17,570	308,709	21,99	27,300	100,000

Legenda: N = número de amostras; N* = número de dados faltantes; DesvPad = desvio padrão; CoefVar = Coeficiente de variação.

Fonte: Autor (2022)

Assim, foram selecionados dentre o grupo B-01, conforme os critérios citados no item 3.1.2, os resultados de ensaios realizados em 421 amostras de solo, compilados de 16 publicações: Wasti e Bezirci (1986); Diniz (1994); Campos et al. (1999); Silveira (2001); Deka et al. (2009); Ozer (2009); Kayabali (2011); Di Matteo (2012); Cevallos Luna (2012); Flores e Emidio (2014); Quintela et al. (2014); Nini (2014); Couto et al. (2016); Bicalho et al. (2017); El-Shinawi (2017); Niazi et al. (2019).

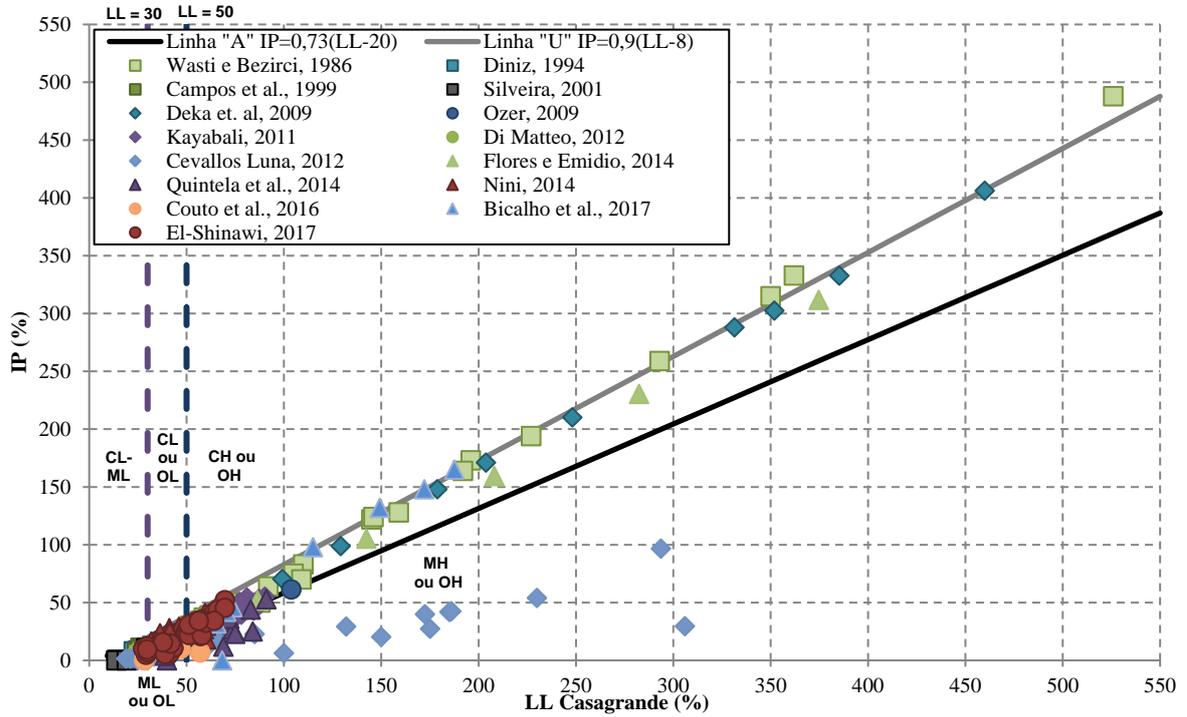
Figura 15 - Resultados experimentais dos dados selecionados para o grupo B-01.



Fonte: Autor (2022)

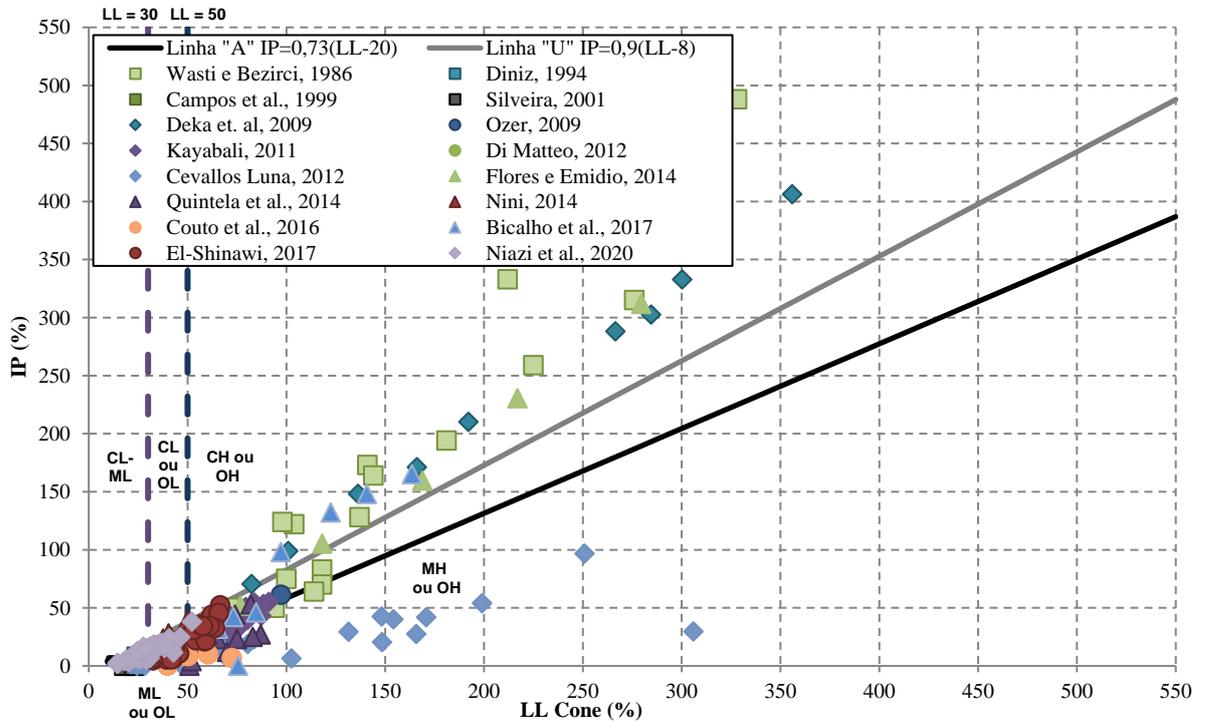
Os valores de IP, LL_c e LL_p foram utilizados para a classificação dos solos através do ábaco de Casagrande, conforme apresentado nas Figura 16 e Figura 17.

Figura 16 – Ábaco de Casagrande para os dados de LL_c do grupo B-01.



Fonte: Autor (2022)

Figura 17 - Ábaco de Casagrande para os dados de LL_p do grupo B-01.



Fonte: Autor (2022)

Os dados de Wasti e Bezirci (1986) com valores de LL_c maiores que 120, classificados como misturas de bentonita e com porcentagem de argila superior a 50%, tenderam a linha “U” do ábaco. Tais amostras mostraram uma maior dispersão no ábaco de LL_p , com as amostras posicionadas acima da linha “U” do ábaco. Nestes casos, o valor de LL_p se mostrou menor que o de LL_c . Isso também foi observado para os resultados das amostras publicados por Deka et al. (2009), classificadas como misturas de argila expansiva contendo montmorilonita, illita e quartzo e argila não expansiva composta por quartzo e magnetita, e com as amostras de mistura de bentonita e areia fina ($C > 60\%$) publicadas por Bicalho et al. (2017).

Flores e Emidio (2014) avaliaram misturas de bentonita e caulinita, e as amostras se posicionaram entre as linhas “A” e “U” do ábaco de Casagrande. No entanto, quando plotados os dados de LL_p no ábaco, as amostras com maior porcentagem de bentonita, e maior LL , se moveram para cima da linha “U”. Nesse caso, LL_p também se mostrou menor que LL_c .

As amostras de caulinita publicadas por Bicalho et al. (2017) tenderam a linha “A” do ábaco de Casagrande, e apresentaram LL_p superior ao LL_c . Entretanto, as amostras investigadas por Campos et al. (1999) embora também classificadas como caulinitas, em alguns casos se mostraram abaixo da linha “A” do ábaco de Casagrande. Quando analisado o ábaco LL_p , mais resultados são vistos abaixo da linha “A”, já que se observa um valor de LL_p maior que o de LL_c . O mesmo ocorreu para os dados publicados por Diniz (1994), Silveira (2001), Ozer (2009), Kayabali (2011), Di Matteo (2012), Quintela et al. (2014), Nini (2014) e Niazi et al. (2019), já que a classificação de amostras como siltes, argilas siltosas ou areias argilosas justifica seu posicionamento no ábaco. De uma maneira geral, esses dados tendem a linha “A” do ábaco de Casagrande. Embora os resultados de Couto et al. (2016) também tenham apresentado valores baixos de LL , as amostras de solo são compostas em sua maioria por silte e areia, o que justifica seu posicionamento abaixo da linha “A” do ábaco.

Os resultados das amostras de solo publicados por Cevallos Luna (2012), em sua maioria são classificados como silte e em geral se encontram abaixo da linha “A” do ábaco de Casagrande. Já os dados publicados por El-Shinawi (2017), descritos como argila, também apresentaram a tendência citada, já que embora a classificação das amostras, a granulometria dos solos indica a presença de silte e areia. Entretanto, os dados de LL maiores que 50, em sua maioria, se posicionaram entre as linhas “A” e “U” do ábaco.

A Tabela 13 sintetiza a distribuição por classificação SUCS para os resultados de LL_c e LL_p , bem como a porcentagem das amostras que sofreu mudança de classificação.

Tabela 13 – Resumo da classificação SUCS de LL_c e LL_p para os dados do grupo B-01.

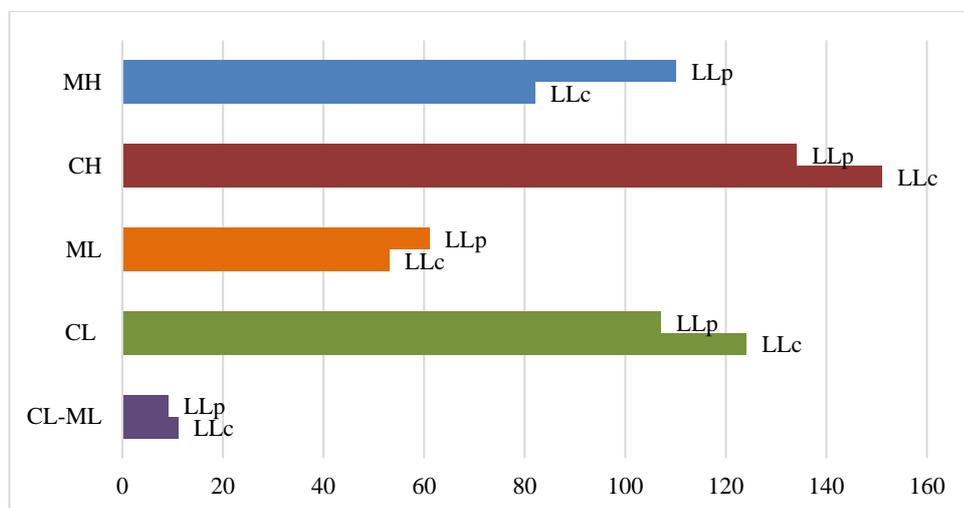
Referência	Amostras	Classificação SUCS LL _c					Classificação SUCS LL _p					Mudança de posição no Ábaco (%)
		CL-ML (%)	CL (%)	ML (%)	CH (%)	MH (%)	CL-ML (%)	CL (%)	ML (%)	CH (%)	MH (%)	
Wasti e Bezirci, 1986	25	0,00	8,00	0,00	92,00	0,00	0,00	4,00	0,00	80,00	16,00	16,00
Diniz, 1994	11	9,09	81,82	9,09	0,00	0,00	9,09	54,55	36,36	0,00	0,00	27,27
Campos et al., 1999	15	6,67	46,67	20,00	13,33	13,33	6,67	33,33	26,67	13,33	20,00	33,33
Silveira, 2001	5	0,00	60,00	0,00	0,00	0,00	0,00	20,00	40,00	0,00	0,00	40,00
Deka et. al, 2009	11	0,00	0,00	9,09	90,91	0,00	0,00	0,00	9,09	90,91	0,00	0,00
Ozer, 2009	32	0,00	40,63	12,50	40,63	6,25	0,00	34,38	15,63	43,75	6,25	18,75
Kayabali, 2011	119	0,00	2,52	3,36	54,62	39,50	0,00	0,00	0,84	42,86	56,30	19,33
Di Matteo, 2012	6	0,00	100,00	0,00	0,00	0,00	0,00	83,33	16,67	0,00	0,00	16,67
Cevallos Luna, 2012	36	5,56	0,00	47,22	0,00	38,89	2,78	2,78	47,22	0,00	38,89	5,56
Flores e Emidio, 2014	6	0,00	0,00	0,00	83,33	16,67	0,00	0,00	0,00	83,33	16,67	0,00
Quintela et al., 2014	14	0,00	0,00	7,14	35,71	50,00	0,00	14,29	0,00	28,57	50,00	4,00
Nini, 2014	29	0,00	58,62	6,90	20,69	13,79	0,00	55,17	6,90	24,14	13,79	3,45
Couto et al., 2016	6	0,00	0,00	66,67	0,00	16,67	0,00	0,00	33,33	0,00	50,00	50,00
Bicalho et al., 2017	9	0,00	0,00	0,00	77,78	11,11	0,00	0,00	0,00	66,67	22,22	11,11
El-Shinawi, 2017	40	0,00	30,00	27,50	35,00	7,50	0,00	20,00	37,50	35,00	7,50	10,00
Niazi et al., 2019	65	10,77	80,00	7,69	1,54	0,00	9,23	78,46	10,77	1,54	0,00	4,62

Fonte: Autor (2022)

Nota-se da Tabela 13 que os dados que apresentaram maior mudança de posição no ábaco de Casagrande quando analisados os resultados de LL_c e LL_p, possuíam LL_c<80: Couto et al. (2016), Silveira (2001), Campos et al. (1999) e Diniz (1994).

A Figura 18 sintetiza a variação dessa classificação dos dados do grupo B-01.

Figura 18 – Mudança de posição no ábaco de Casagrande para os dados do grupo B-01.



Fonte: Autor (2022)

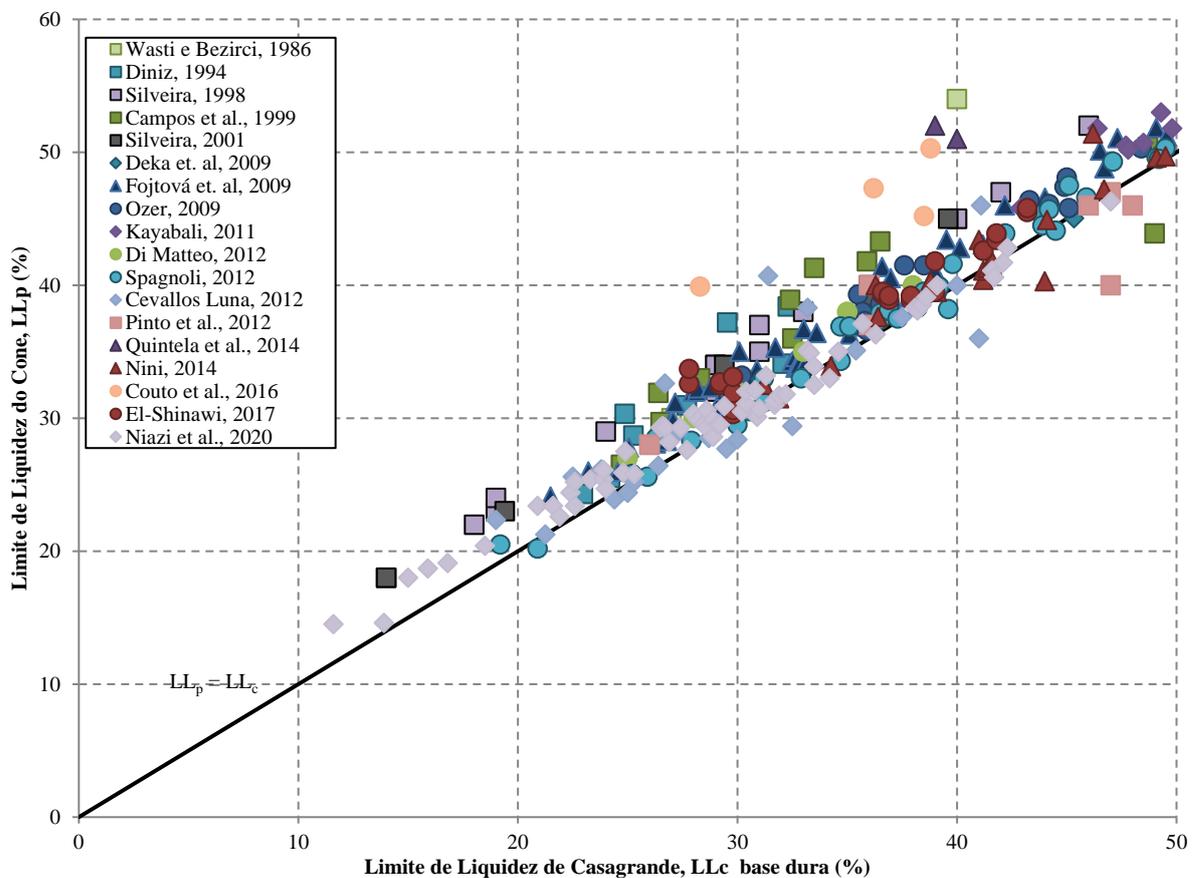
A caracterização dos dados segundo as informações e resultados dispostos na literatura permite sua padronização e utilização com garantia de confiabilidade. Assim, objetivando avaliar a influência do intervalo de LL nas correlações entre LL_c e LL_p , conforme citado no item 3.1.2, os dados do grupo B-01 foram divididos nos seguintes subgrupos:

- B-01i: $LL_c < 50$;
- B-01ii: $50 < LL_c \leq 120$;
- B-01iii: $LL_c > 120$.

Descrição dos dados do subgrupo B-01i

O subgrupo B-01i contempla os resultados de ensaios do grupo B-01 que apresentaram $LL_c < 50$. Assim, dentro deste subgrupo, foram investigados os resultados de 188 ensaios, oriundos de 13 publicações: Wasti e Bezirci (1986); Diniz (1994); Campos et al. (1999); Silveira (2001); Ozer (2009); Kayabali (2011); Di Matteo (2012); Cevallos Luna (2012); Quintela et al. (2014); Nini (2014); Couto et al. (2016); El-Shinawi (2017); Niazi et al. (2019).

Figura 19 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01i.

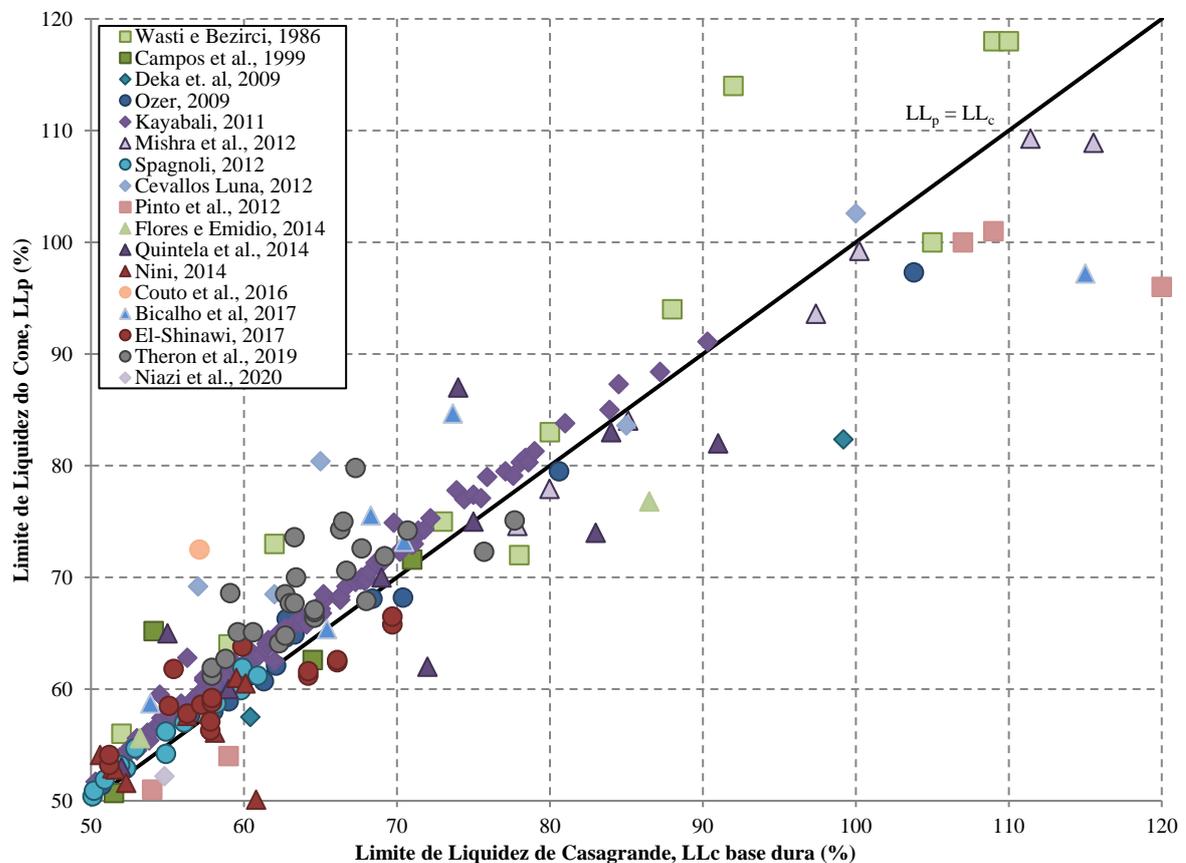


Fonte: Autor (2022)

Descrição dos dados do subgrupo B-01ii

O subgrupo B-01ii contempla os resultados de ensaios do grupo B-01 que apresentaram $50 < LL_c \leq 120$. Assim, dentro deste subgrupo, foram investigados os resultados de 199 ensaios, oriundos de 13 publicações: Wasti e Bezirci (1986); Campos et al. (1999); Deka et al. (2009); Ozer (2009); Kayabali (2011); Cevallos Luna (2012); Flores e Emidio (2014); Quintela et al. (2014); Nini (2014); Couto et al. (2016); Bicalho et al. (2017); El-Shinawi (2017); Niazi et al. (2019).

Figura 20 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01ii.

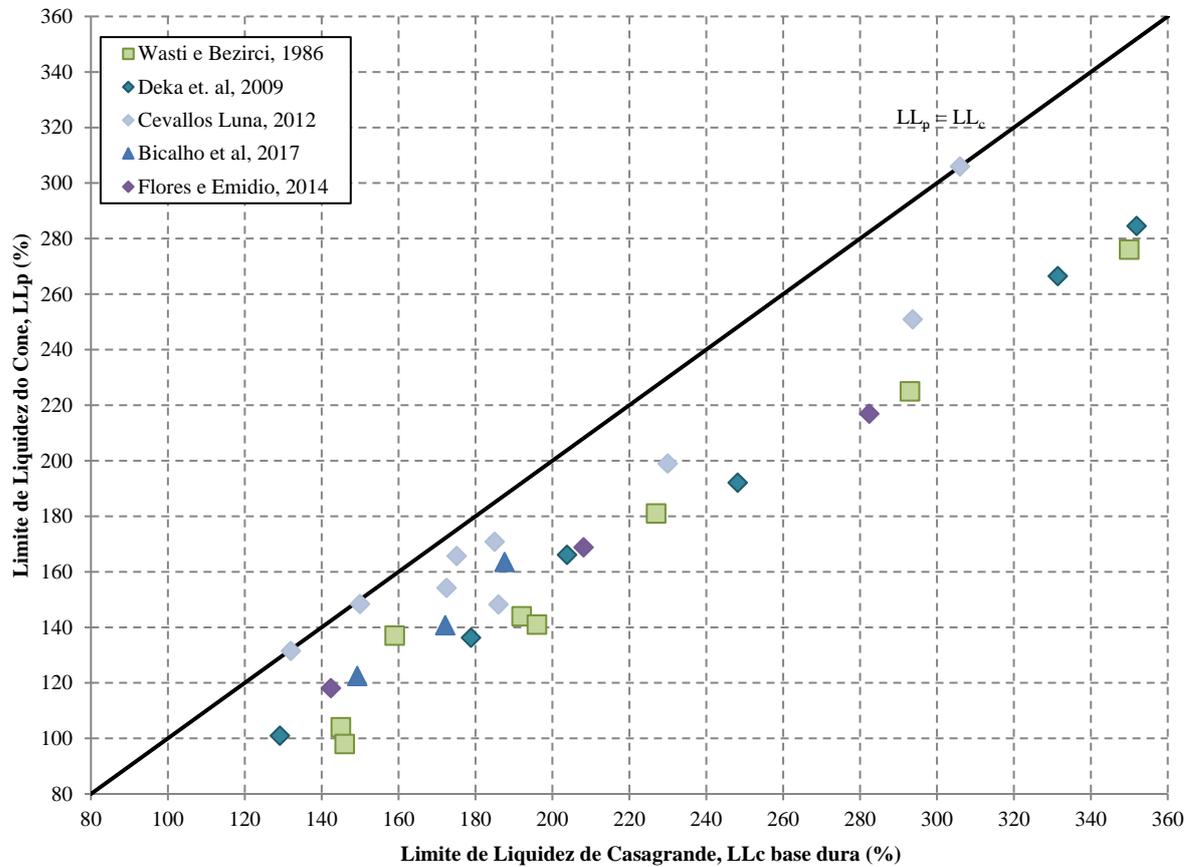


Fonte: Autor (2022)

Descrição dos dados do subgrupo B-01iii

O subgrupo B-01iii contempla os resultados de ensaios do grupo B-01 que apresentaram $LL_c > 120$. Assim, dentro deste subgrupo, foram investigados os resultados de 34 ensaios, oriundos de 5 publicações: Wasti e Bezirci (1986); Deka et al. (2009); Cevallos Luna (2012); Flores e Emidio (2014); Bicalho et al. (2017).

Figura 21 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01iii.



Fonte: Autor (2022).

3.1.2.2 Descrição dos dados do Grupo B-02

O grupo de dados experimentais B-02 abrange os resultados experimentais compilados da literatura que utilizaram o cone no padrão britânico e o aparelho de Casagrande de base macia para a realização dos ensaios.

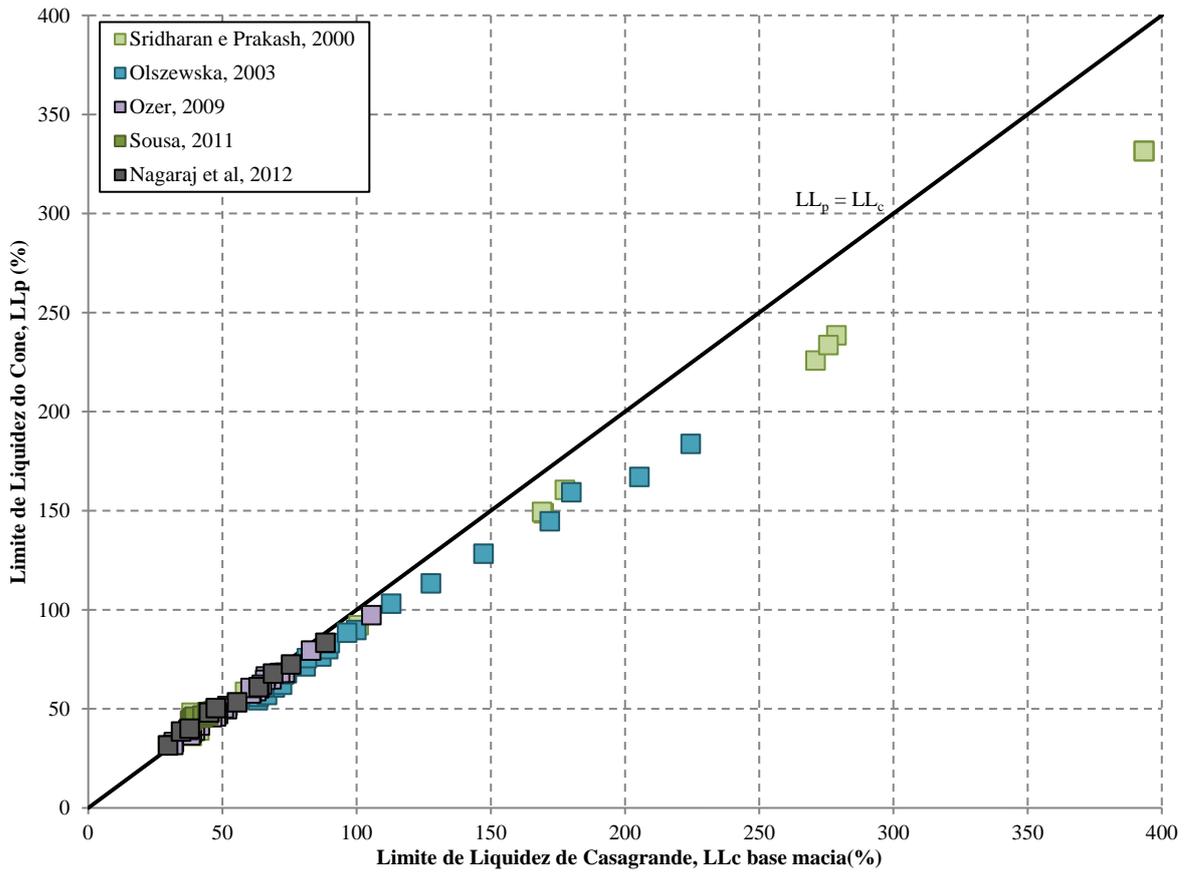
Foram investigados no grupo B-02 os resultados de ensaios de 99 amostras de solo, compilados de 5 publicações: Sridharan e Prakash (2000); Olszewska (2003); Ozer (2009); Sousa (2011); Nagaraj et al. (2012).

Tabela 14 – Dados experimentais compilados da literatura (grupo B-02).

Referências	Variação de LL _c (%)	Tipo de Solo	Nº Amostras	Granulometria			Origem Solo
				%média Argila	% média Silte	%média Areia	
Sridharan e Prakash (2000)	29-394	Caulinitas, montmorilonitas e bentonitas	19	75,89	-	-	-
Olszewska (2003)	63-225	Caulinitas e bentonitas	22	100,00	-	-	-
Ozer (2009)	31-106	Solos naturais	32	20,55	53,16	23,18	Turquia
Sousa (2011)	38-46	Caulinitas e ilitas	16	21,65	77,10	1,25	Portugal
Nagaraj et al. (2012)	29-89	Caulinitas e montmorilonitas	10	29,60	45,00	25,40	Índia

Fonte: Autor (2022)

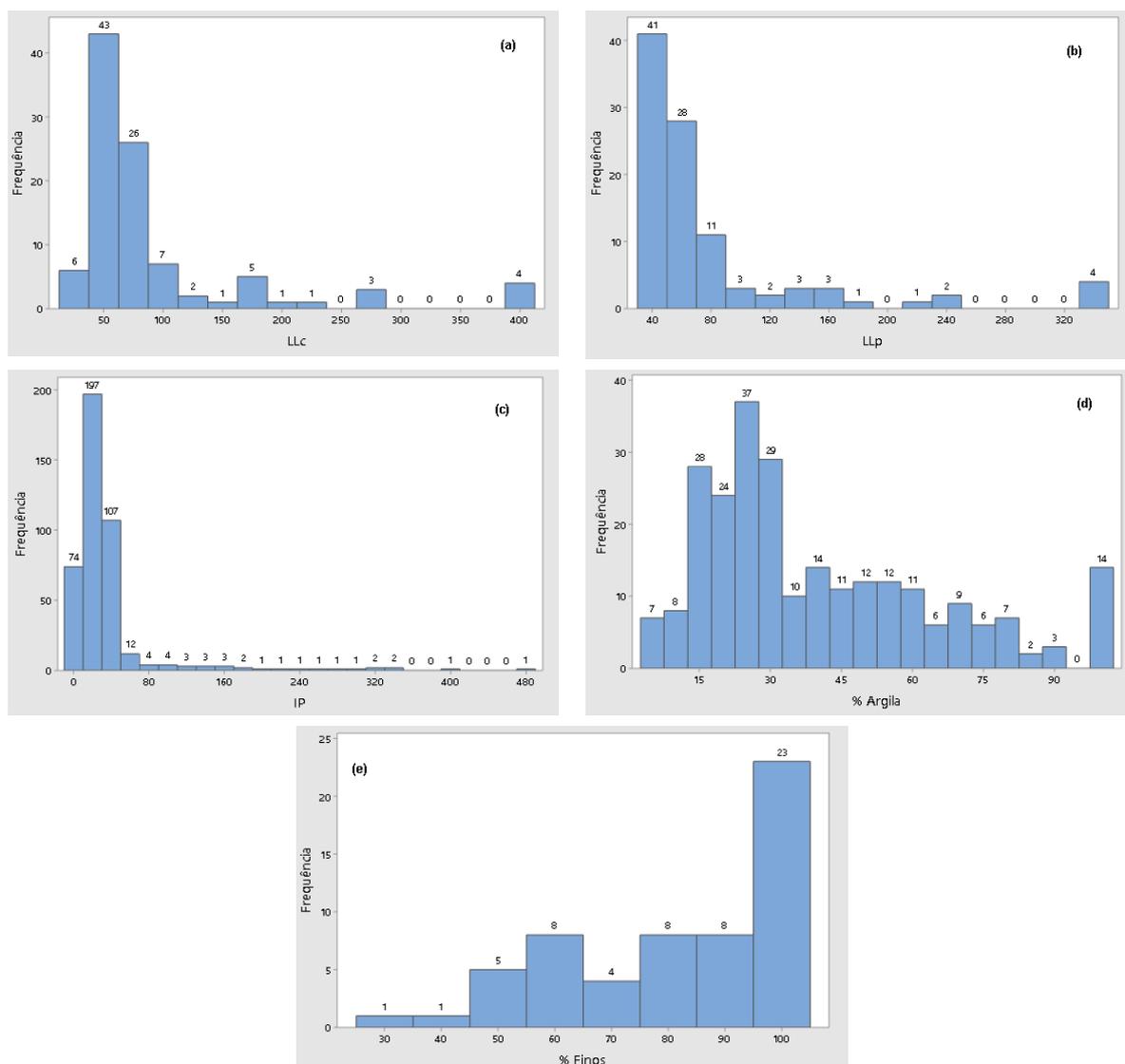
Figura 22 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do grupo B-02.



Fonte: Autor (2022)

A Figura 23 apresenta os histogramas com a distribuição dos dados do grupo B-02.

Figura 23 – Histograma dos dados do grupo B-02: (a) LL_c; (b) LL_p; (c) IP; (d) porcentagem de argila; (e) porcentagem de finos; (f) classificação SUCS.



Fonte: Autor (2022)

A descrição estatística dos dados do grupo B-02 é apresentada na Tabela 15. Por meio da tabela observa-se que do total de 199 amostras, apenas 80 possuem a informação de dados de IP, 99 a porcentagem de argila e 58 a porcentagem de finos.

Tabela 15 - Descrição estatística dos dados do grupo B-02 (N=99).

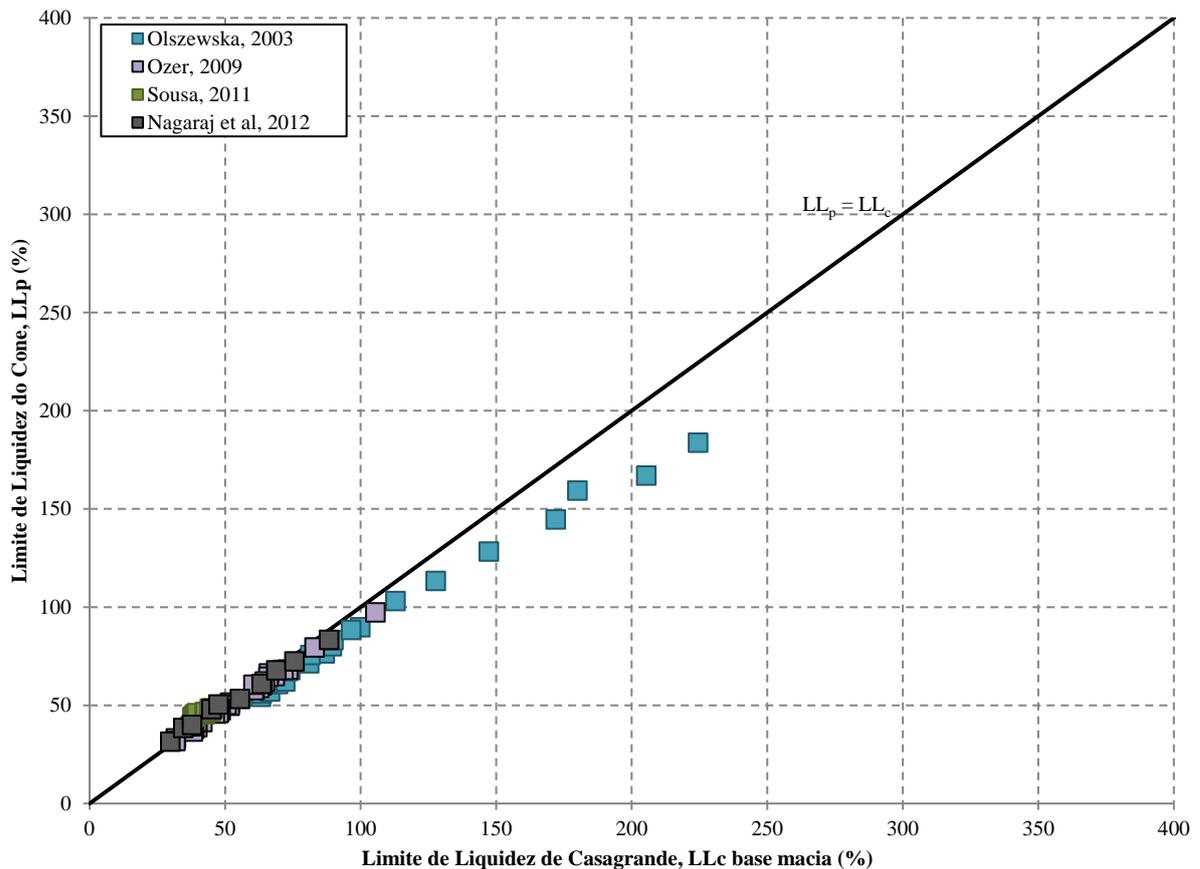
Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Máximo
LL _c	99	0	87,88	82,34	6779,49	93,69	29,80	393,40
LL _p	99	0	80,71	67,14	4508,19	83,19	31,50	331,50
IP	80	19	33,32	27,26	743,23	81,83	10,70	152,00
% Argila	99	0	49,92	37,17	1381,97	74,47	3,40	100,00
% Finos	58	41	80,77	19,27	371,19	23,85	27,30	100,00

Legenda: N = número de amostras; N* = número de dados faltantes; DesvPad = desvio padrão; CoefVar = Coeficiente de variação.

Fonte: Autor (2022)

Assim, foram selecionados dentre o grupo B-02 os resultados de ensaios realizados em 80 amostras de solo, compilados de 4 publicações: Olszewska (2003); Ozer (2009); Sousa (2011); Nagaraj et al. (2012). Ou seja, apenas os dados publicados por Sridharan e Prakash (2000) não continham a informação do IP.

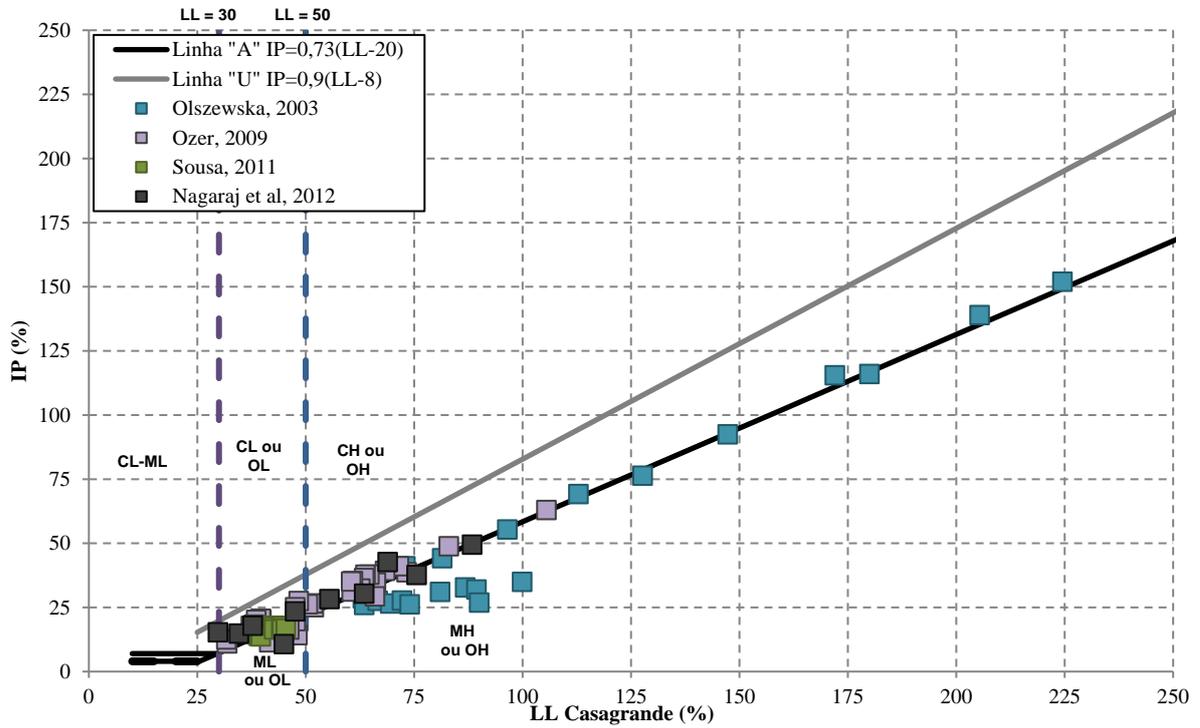
Figura 24 - Resultados experimentais dos dados selecionados para o grupo B-02.



Fonte: Autor (2022)

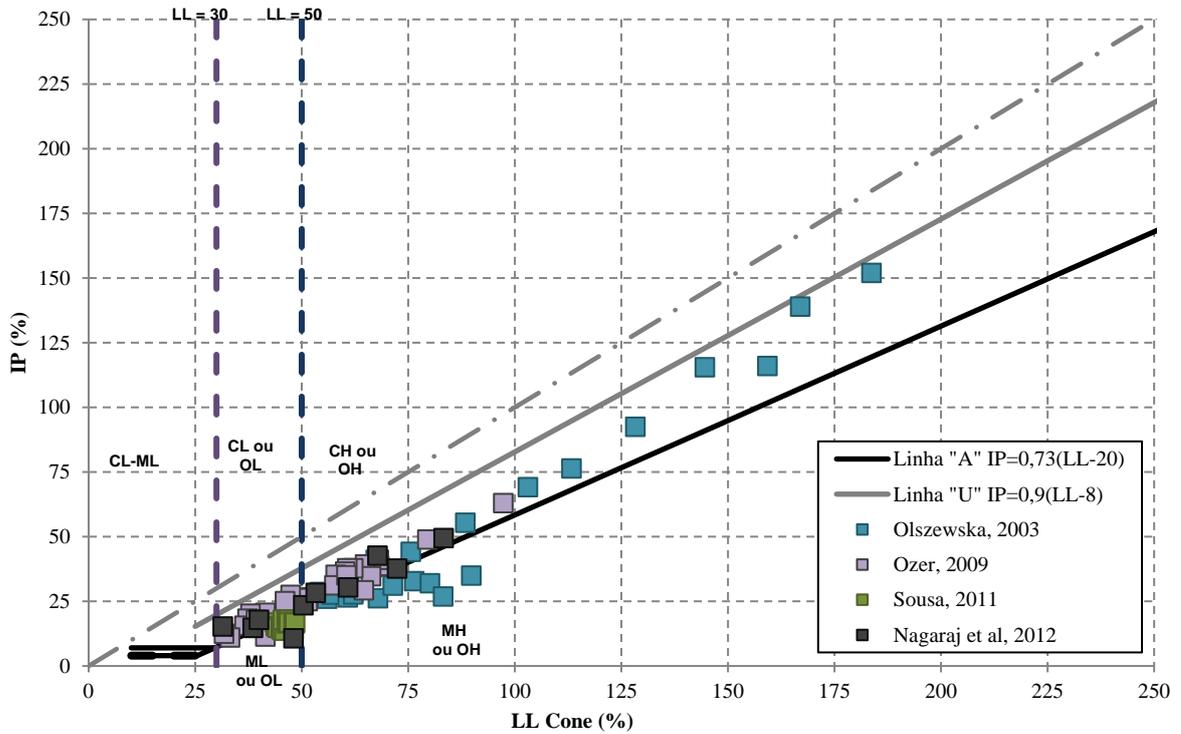
Os valores de IP, LL_c e LL_p foram utilizados para a classificação dos solos através do ábaco de Casagrande, conforme apresentado nas Figura 25 e Figura 26.

Figura 25 - Ábaco de Casagrande para os dados de LL_c do grupo B-02.



Fonte: Autor (2022)

Figura 26 - Ábaco de Casagrande para os dados de LL_p do grupo B-02.



Fonte: Autor (2022)

Os resultados dos ensaios publicados por Olszewska (2003), embora tratem de misturas de caulinita e bentonita em diferentes proporções e diferentes valores de LL, possui dados localizados abaixo da linha “A” e outros que tendem a linha “A” do ábaco de Casagrande. Nota-se que quando os resultados de LL_p são plotados no ábaco (Figura 26), aquelas amostras que antes apresentavam uma tendência à linha “A”, se aproximam da linha “U”. Nestes caso os valores de LL_p são menores que os de LL_c .

Os dados publicados por Ozer (2009) de uma maneira geral, encontram-se entre as linhas “A” e “U” do ábaco, exceto aqueles classificados como silte, que se posicionaram abaixo da linha “A”. Neste caso também foi observado que os valores de LL_p são menores que os de LL_c .

Já os dados de Nagaraj et al. (2012), embora a mesma tendência de posicionamento no ábaco de Casagrande, apresentaram tendências distintas quanto aos valores de LL_c e LL_p . Assim, tais resultados indicam valores maiores de LL_c para as amostras classificadas como montmorilonitas, enquanto para as amostras de caulinitas foram observados maiores valores de LL_p .

Os resultados publicados por Sousa (2011) por sua vez, embora classificados como argilas, devido ao alto percentual de silte em sua composição granulométrica, apresentou amostras abaixo da linha “A” do ábaco de Casagrande. Quando plotados no ábaco de LL_p , os resultados dos ensaios se posicionaram abaixo da linha “A”. Os valores de LL_p se mostraram maiores que os de LL_c .

A Tabela 16 sintetiza a distribuição por classificação SUCS para os resultados de LL_c e LL_p , bem como a porcentagem das amostras que sofreu mudança de classificação.

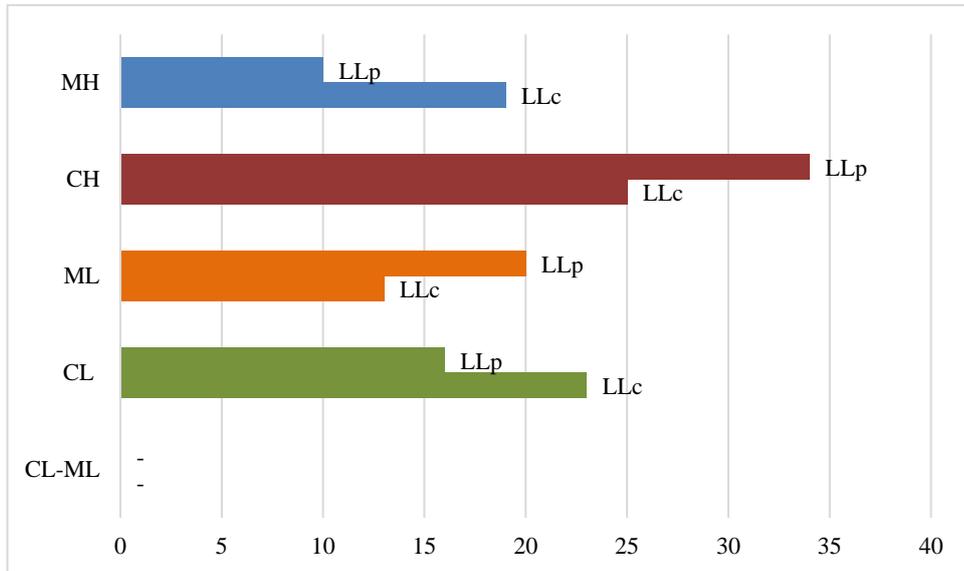
Tabela 16 - Resumo da classificação SUCS de LL_c e LL_p para os dados do grupo B-02.

Referência	Amostras	Classificação SUCS LL_c					Classificação SUCS LL_p					Mudança de posição no Ábaco (%)
		CL-ML (%)	CL (%)	ML (%)	CH (%)	MH (%)	CL-ML (%)	CL (%)	ML (%)	CH (%)	MH (%)	
Olszewska, 2003	22	0,00	0,00	0,00	36,36	63,64	0,00	0,00	0,00	63,64	36,36	27,27
Ozer, 2009	32	0,00	34,38	12,50	46,88	6,25	0,00	40,63	9,38	46,88	3,13	9,38
Sousa, 2011	16	0,00	50,00	50,00	0,00	0,00	0,00	0,00	100,00	0,00	0,00	50,00
Nagaraj et al, 2012	10	0,00	40,00	10,00	20,00	30,00	0,00	30,00	10,00	50,00	10,00	30,00

Fonte: Autor (2022)

Da Tabela 16 infere-se um percentual de mudança de posicionamento no ábaco de Casagrande, e consequente classificação SUCS das amostras de solo do grupo B-02, superior a 30% para os resultados de ensaios publicados por Sousa (2011) e Nagaraj et al. (2012). A Figura 27 sintetiza essa variação da classificação dos dados do grupo B-02.

Figura 27 – Mudança de posição no ábaco de Casagrande para os dados do grupo B-02.



Fonte: Autor (2022)

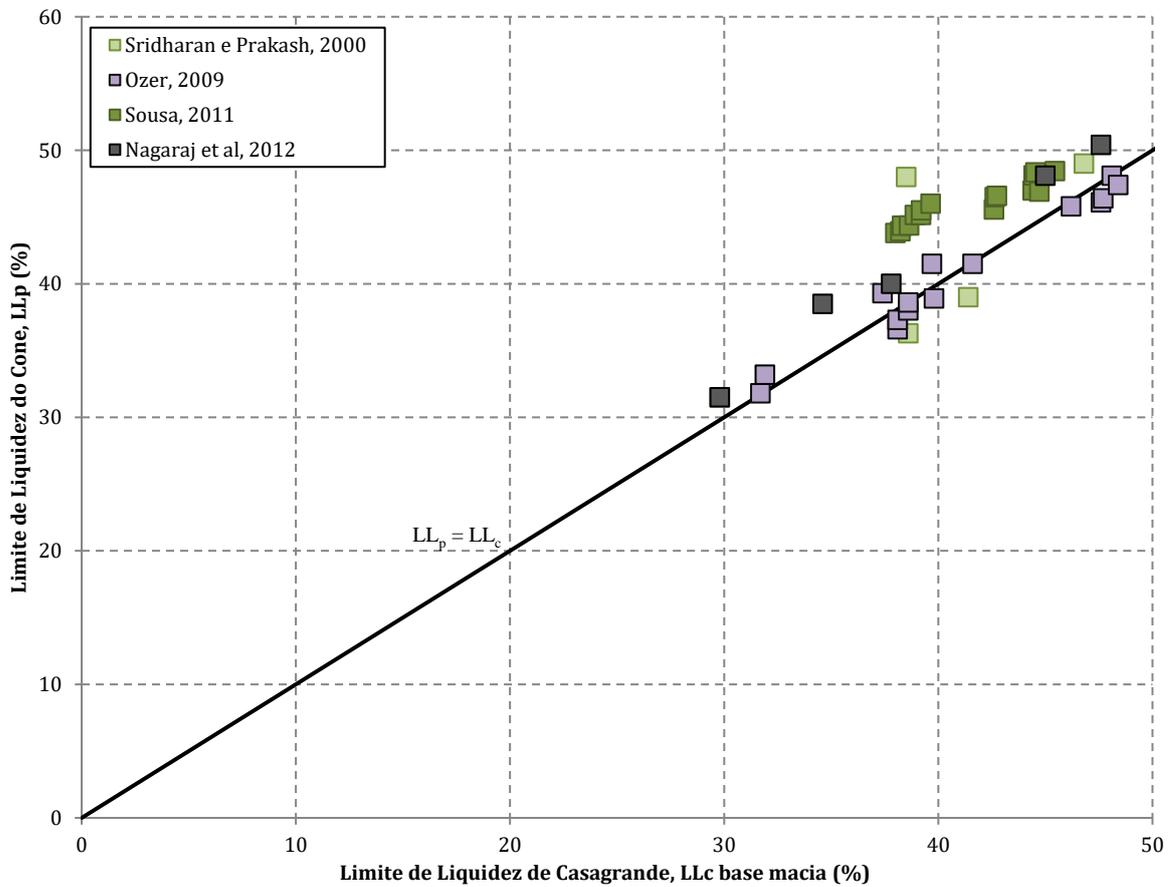
A partir dos dados compilados da literatura, conforme citado no item 3.1.2, para avaliação da influência do intervalo de LL nas correlações entre LL_c e LL_p os dados do grupo B-02 foram divididos nos subgrupos:

- B-02i: $LL_c < 50$;
- B-02ii: $50 < LL_c \leq 120$;
- B-02iii: $LL_c > 120$.

Descrição dos dados do subgrupo B-02i

O subgrupo B-02i contempla os resultados de ensaios do grupo B-02 que apresentaram $LL_c < 50$. Assim, dentro deste subgrupo, foram investigados os resultados de 36 ensaios, oriundos de 3 publicações: Ozer (2009); Sousa (2011); Nagaraj et al. (2012).

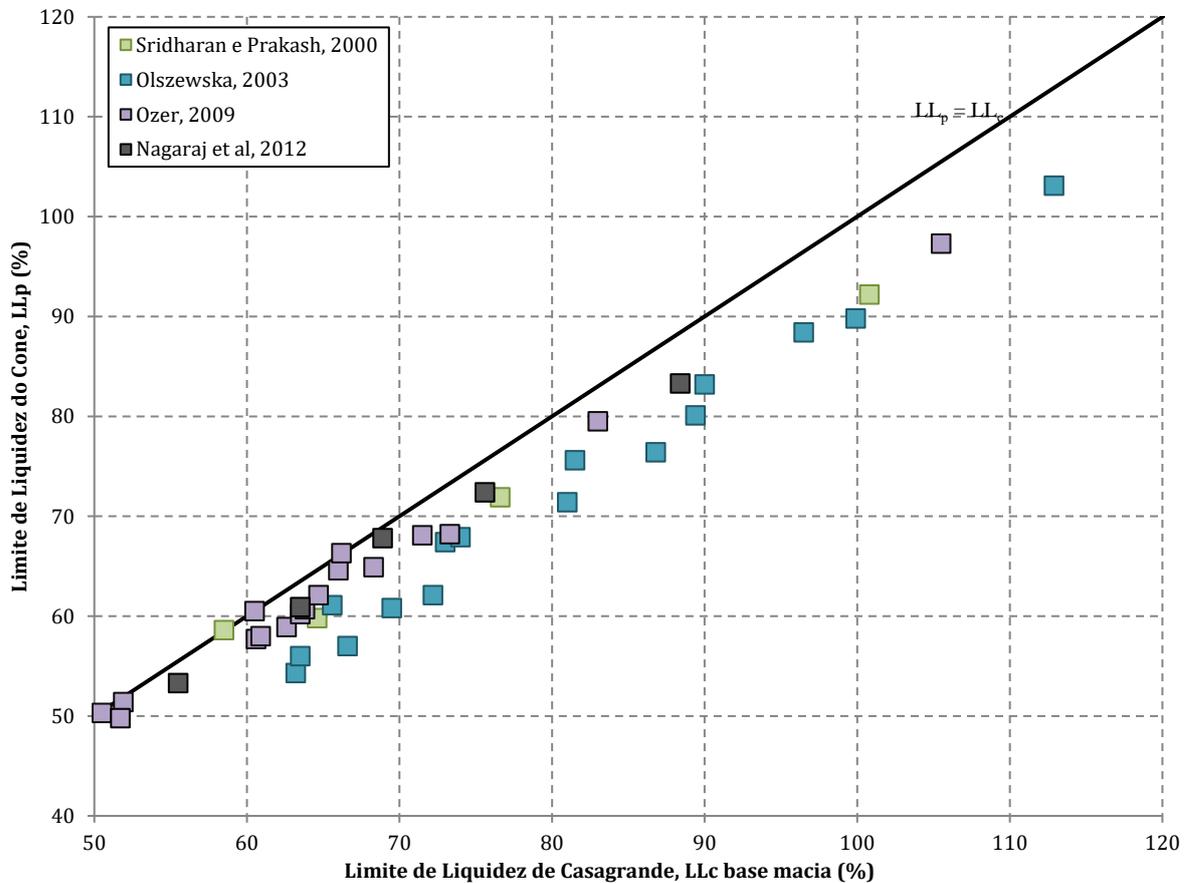
Figura 28 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02i.



Descrição dos dados do subgrupo B-02ii

O subgrupo B-02ii contempla os resultados de ensaios do grupo B-02 que apresentaram $50 < LL_c \leq 120$. Assim, dentro deste subgrupo, foram investigados os resultados de 38 ensaios, oriundos de 3 publicações: Olszewska (2003); Ozer (2009); Nagaraj et al. (2012).

Figura 29 - Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02ii.

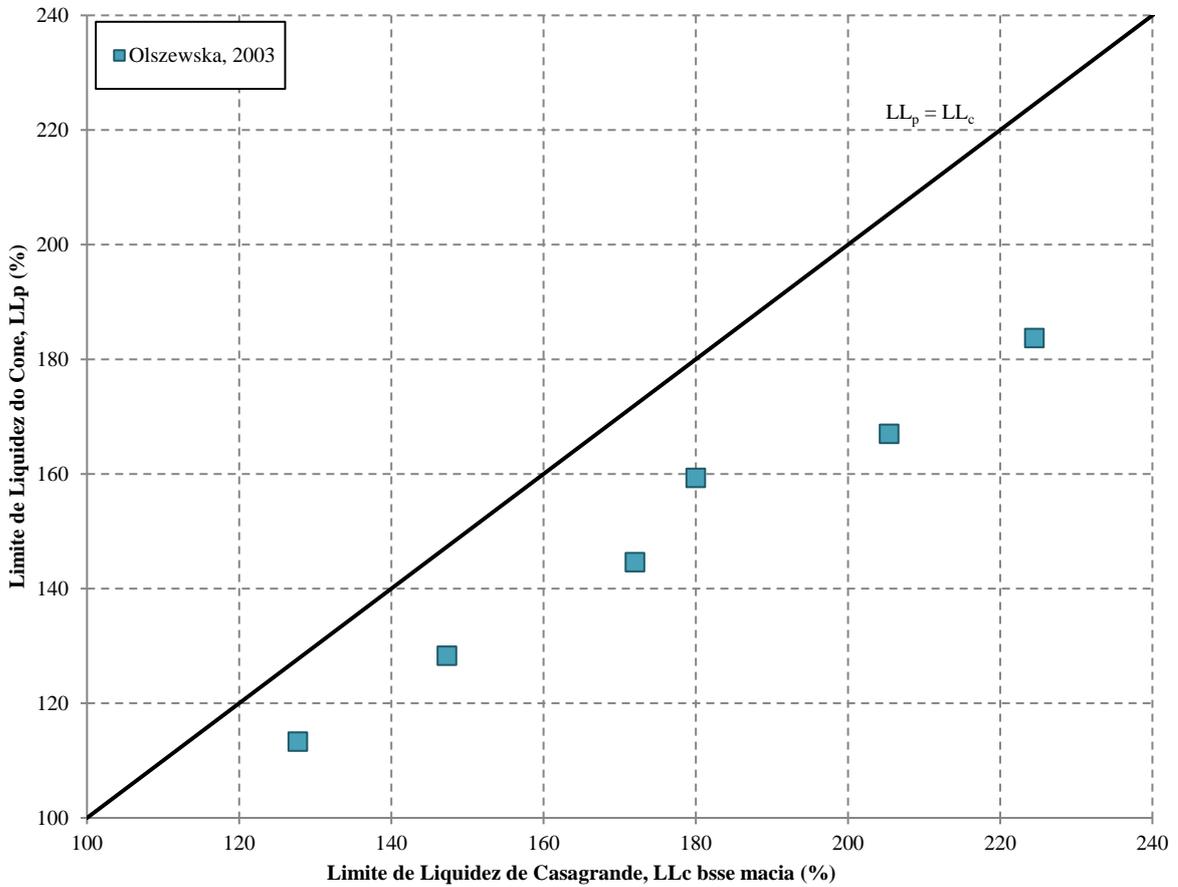


Fonte: Autor (2022)

Descrição dos dados do subgrupo B-02iii

O subgrupo B-02iii contempla os resultados de ensaios do grupo B-02 que apresentaram $LL_c > 120$. Assim, dentro deste subgrupo, foram investigados os resultados de 6 ensaios, publicados por Olszewska (2003).

Figura 30- Valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02iii.



Fonte: Autor (2022)

3.1.2.3 Descrição dos dados do Grupo B-03

O grupo de dados B-03 apresentam os resultados experimentais compilados da literatura que utilizaram o cone no padrão sueco e o aparelho de Casagrande de base dura para a realização dos ensaios.

Foram investigados no grupo B-03 os resultados de ensaios realizados em 137 amostras de solo, compilados de 5 publicações: Christaras (1991); Mendoza e Orozco (2011); Gronbech et al. (2011); Clemente (2018); Clemente et al. (2020).

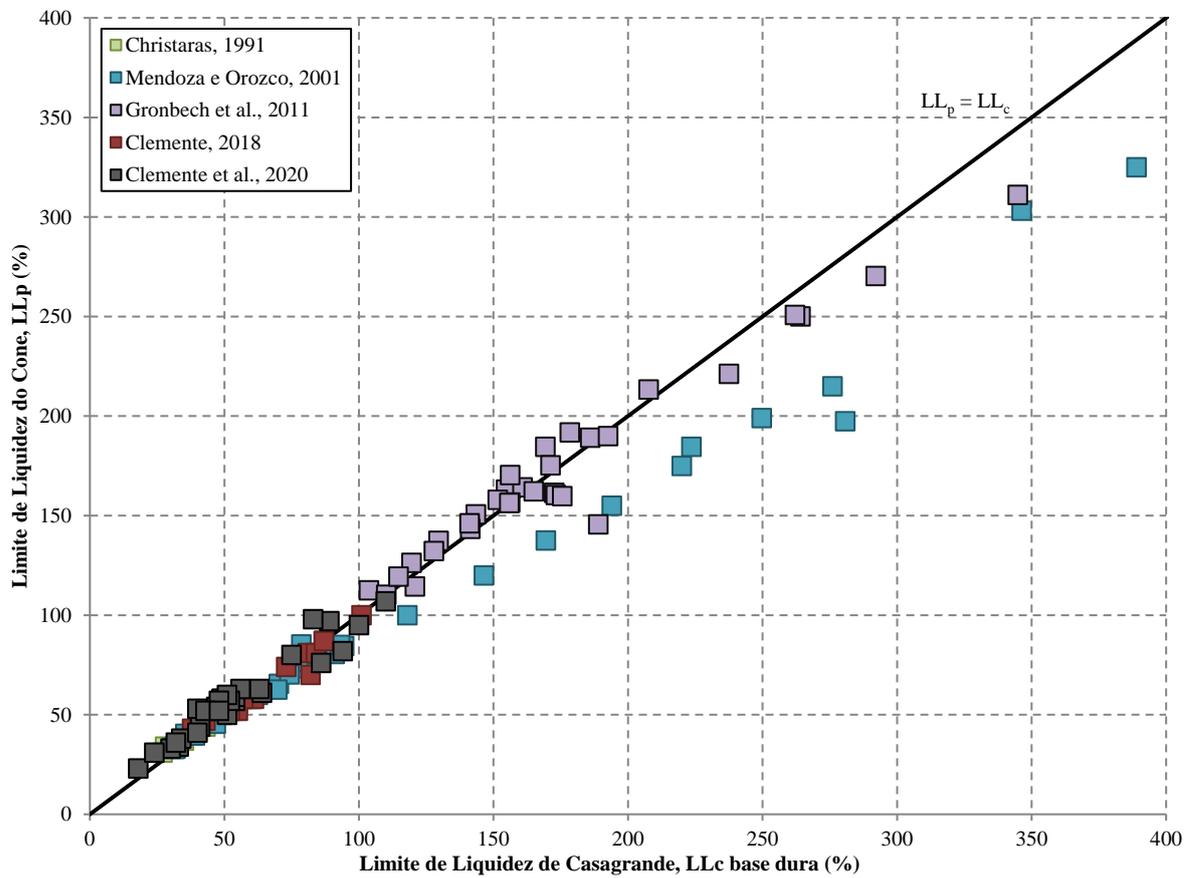
Tabela 17 – Dados experimentais compilados da literatura (grupo B-03).

Referências	Variação de LL _c (%)	Tipo de Solo	Nº Amostras	Granulometria			Origem Solo
				%média Argila	% média Silte	%média Areia	
Christaras (1991)	27-59	Argila de plasticidade intermediária a alta	10	30,00	-	-	Grécia

Referências	Variação de LL _c (%)	Tipo de Solo	Nº Amostras	Granulometria			Origem Solo
				%média Argila	% média Silte	%média Areia	
Mendoza e Orozco (2001)	30-389	Caulinitas e bentonitas	33	-	-	-	México
Gronbech et al. (2011)	87-345	Argilas de alta plasticidade	33	-	-	-	Dinamarca
Clemente (2018)	38-101	Argilas e siltes	16	40,68	35,30	24,04	Rio de Janeiro, Brasil
Clemente et al. (2020)	18-110	Argila, silte, e solos de origem marinha	31	73,74		26,26	Brasil

Fonte: Autor (2022)

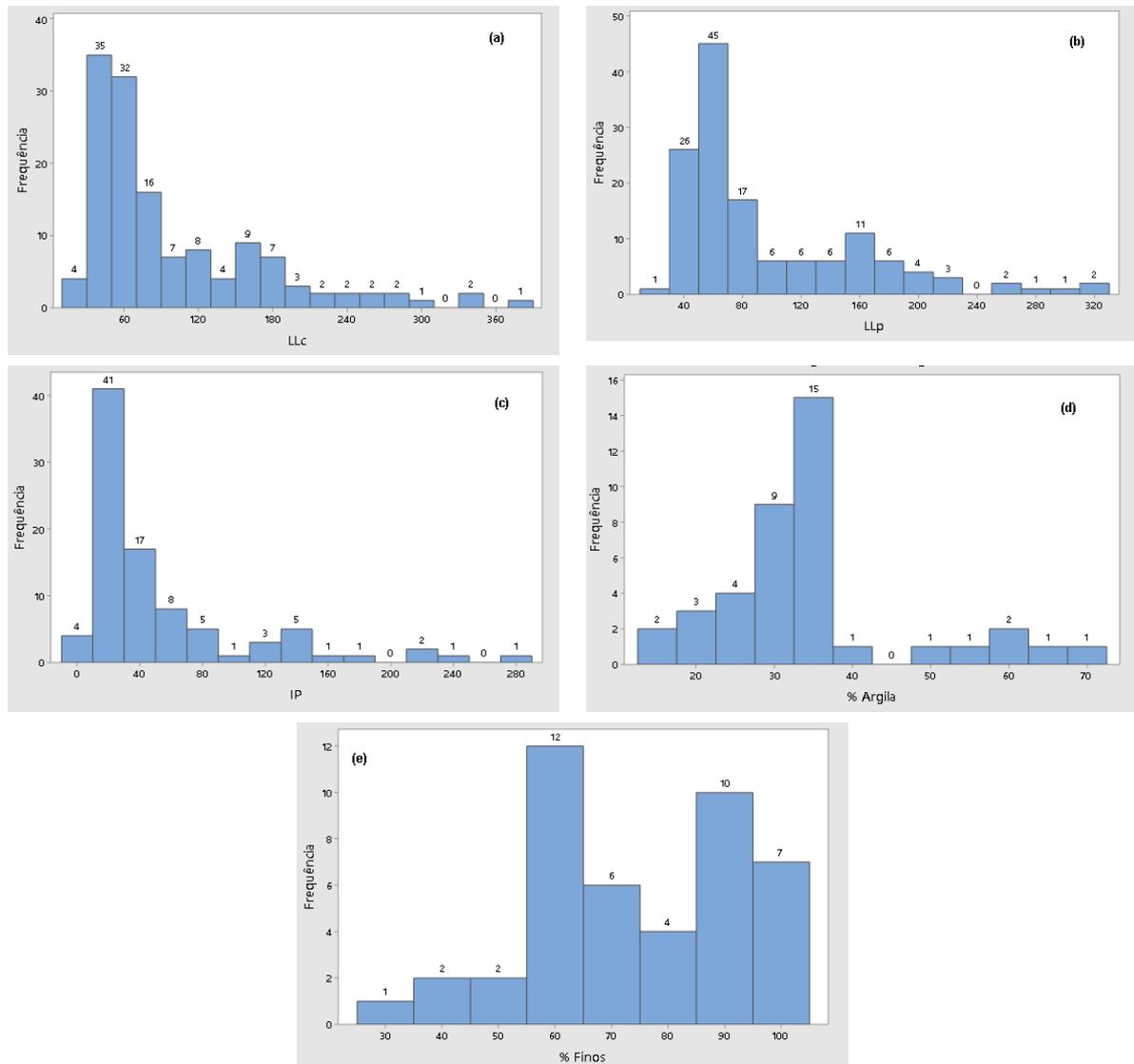
Figura 31 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do grupo B-03.



Fonte: Autor (2022)

A Figura 32 apresenta os histogramas com a distribuição dos dados do grupo B-02.

Figura 32 – Histograma dos dados do grupo B-03: (a) LL_c; (b) LL_p; (c) IP; (d) porcentagem de argila; (e) porcentagem de finos; (f) classificação SUCS.



Fonte: Autor (2022)

A descrição estatística dos dados do grupo B-03 é apresentada na Tabela 18. Por meio da tabela observa-se que do total de 137 amostras, apenas 90 possuem a informação de dados de IP, 40 a porcentagem de argila e 44 a porcentagem de finos.

Tabela 18 - Descrição estatística dos dados do grupo B-03 (N=137).

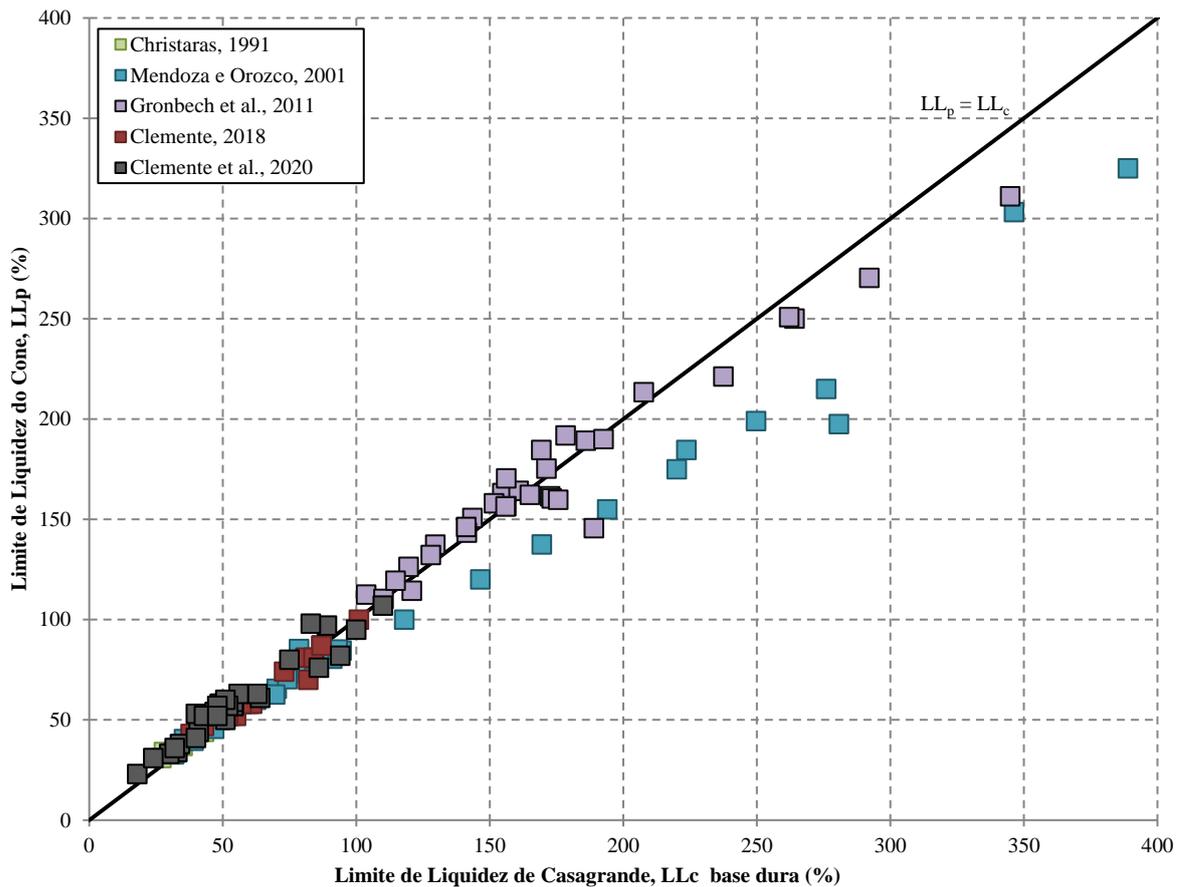
Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Máximo
LL _c	137	0	98,24	74,84	5600,74	76,18	18,00	389,00
LL _p	137	0	94,94	64,91	4212,77	68,36	23,00	325,00
IP	90	47	53,50	56,75	3220,37	106,07	7,00	284,70
% Argila	40	97	34,27	12,51	156,62	36,52	14,29	67,90
% Finos	44	93	74,55	18,07	326,39	24,23	31,50	100,00

Legenda: N = número de amostras; N* = número de dados faltantes; DesvPad = desvio padrão; CoefVar = Coeficiente de variação.

Fonte: Autor (2022)

Assim, foram selecionados dentre o grupo B-03 os resultados de ensaios realizados em 90 amostras de solo, compilados de 4 publicações: Christaras (1991); Gronbech et al. (2011); Clemente (2018); Clemente et al. (2020). Assim, apenas os dados coletados de Mendoza e Orozco (2011) não possuíam a informação do IP.

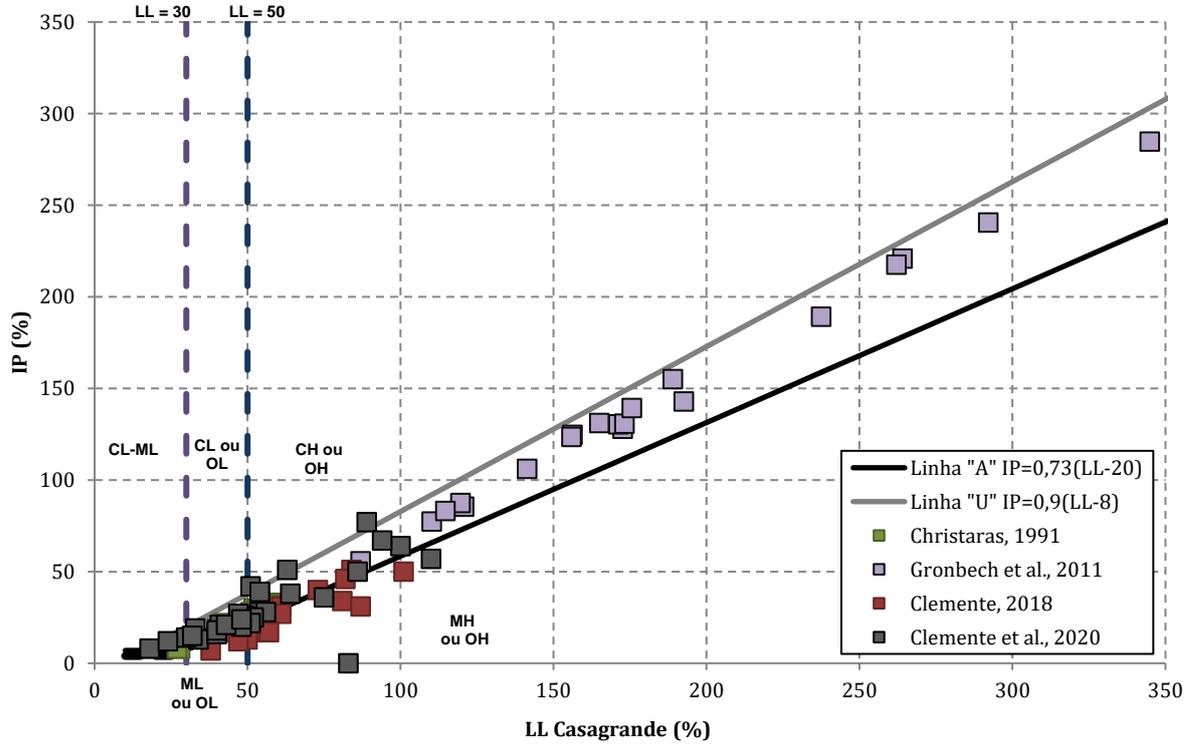
Figura 33 - Resultados experimentais dos dados selecionados para o grupo B-03.



Fonte: Autor (2022)

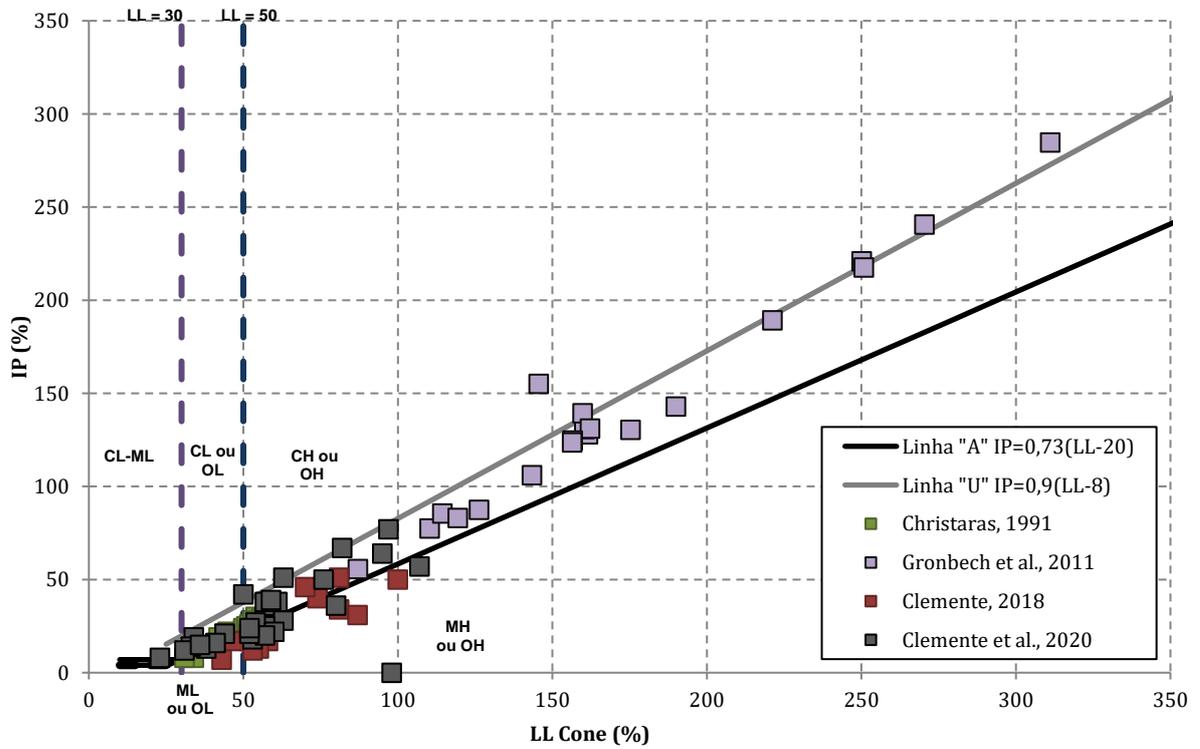
Os valores de IP, LL_c e LL_p foram utilizados para a classificação dos solos através do ábaco de Casagrande, conforme apresentado nas Figura 34 e Figura 35.

Figura 34 - Ábaco de Casagrande para os dados de LL_c do grupo B-03.



Fonte: Autor (2022)

Figura 35 - Ábaco de Casagrande para os dados de LL_p do grupo B-03.



Fonte: Autor (2022)

Os dados publicados por Gronbech et al. (2011), compostos por argilas de alta plasticidade, se posicionaram entre as linhas “A” e “U” do ábaco de Casagrande. Entretanto, quando analisados os resultados de LL_p , as amostras apresentaram uma tendência a linha “U” e em alguns casos se posicionam acima desta linha.

Os resultados publicados por Christaras (1991) se mostraram próximos a linha “A” do ábaco, e apesar de serem classificados como argilas, quando analisados os resultados de LL_p , a mudança de classificação levou a classificação de algumas amostras como silte. Cabe ressaltar que a granulometria destas amostras indica um percentual de argila inferior a 50%.

Já os resultados de Clemente (2018) e Clemente et al. (2020) se mostraram bem dispersos no ábaco. As amostras possuem uma composição variada de argilas e siltes, com amostras de origem marinha e solo residual.

A Tabela 19 sintetiza a distribuição por classificação SUCS para os resultados de LL_c e LL_p , bem como a porcentagem das amostras que sofreu mudança de classificação.

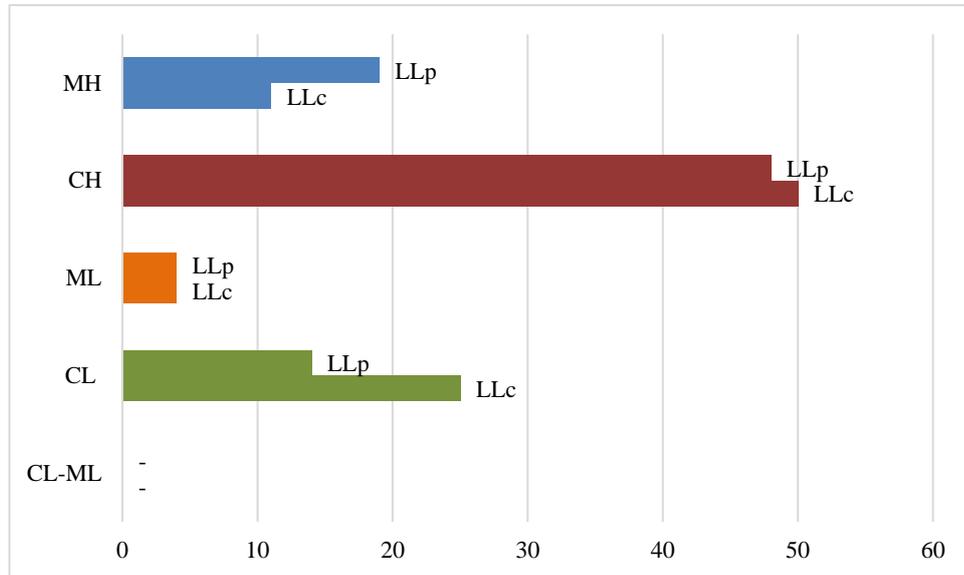
Tabela 19 - Resumo da classificação SUCS de LL_c e LL_p para os dados do grupo B-03.

Referência	Amostras	Classificação SUCS LL_c					Classificação SUCS LL_p					Mudança de posição no Ábaco (%)
		CL-ML (%)	CL (%)	ML (%)	CH (%)	MH (%)	CL-ML (%)	CL (%)	ML (%)	CH (%)	MH (%)	
Christaras, 1991	24	0,00	45,83	0,00	54,17	0,00	0,00	25,00	8,33	66,67	0,00	20,83
Gronbech et al., 2011	33	0,00	0,00	0,00	60,61	0,00	0,00	0,00	0,00	45,45	0,00	15,15
Clemente, 2018	16	0,00	0,00	18,75	31,25	50,00	0,00	0,00	12,50	31,25	56,25	6,25
Clemente et al., 2020	31	0,00	45,16	3,23	38,71	9,68	0,00	25,81	0,00	38,71	32,26	29,03

Fonte: Autor (2022)

De acordo com a Tabela 19, os dados de Clemente et al. (2020) foram os que apresentaram maior percentual de mudança de classificação. Cerca de 29% das amostras de solo de apresentou classificação distinta no ábaco e Casagrande quando utilizado o valor de LL_p . A Figura 36 sintetiza a variação da classificação dos dados do grupo B-03.

Figura 36 – Mudança de posição no ábaco de Casagrande para os dados do grupo B-03.



Fonte: Autor (2022)

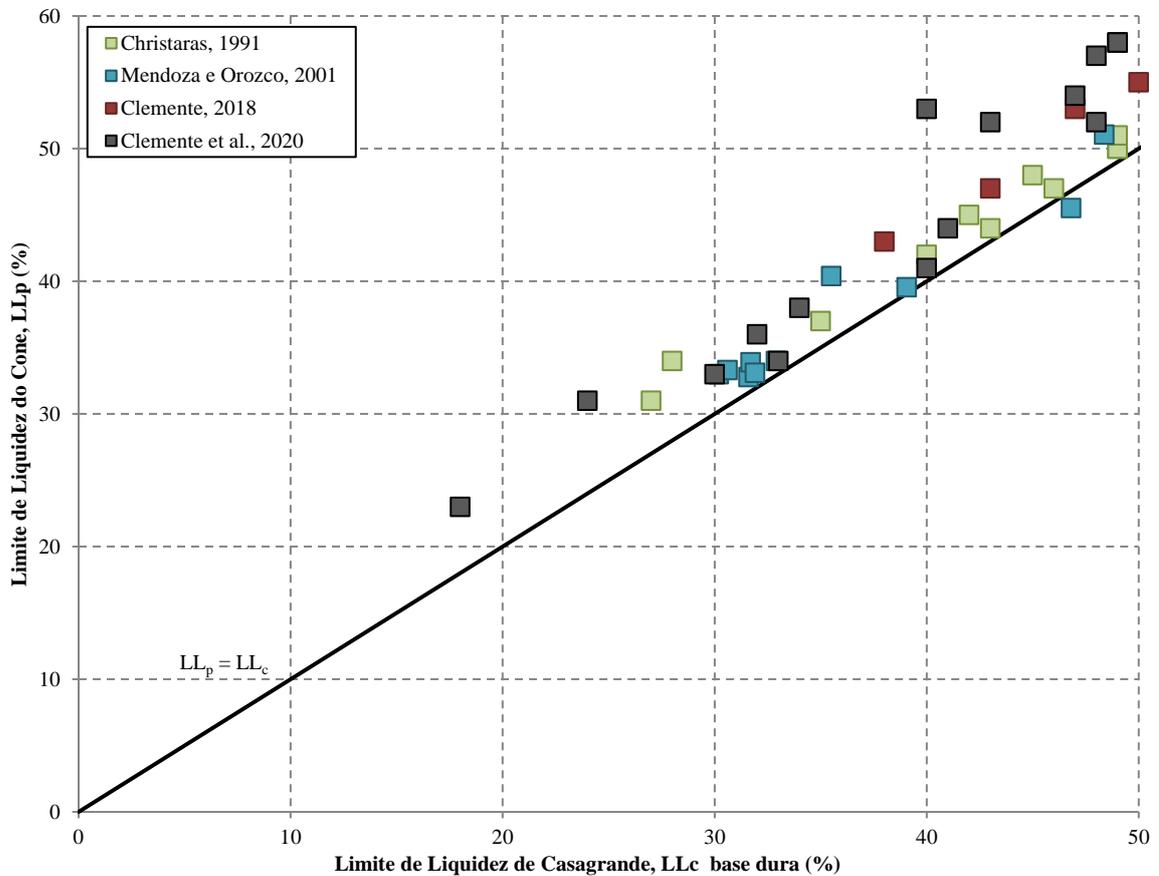
A partir dos dados compilados da literatura, conforme citado no item 3.1.2, para avaliação da influência do intervalo de LL nas correlações entre LL_c e LL_p os dados do grupo B-03 foram divididos nos subgrupos:

- B-03i: $LL_c < 50$;
- B-03ii: $50 < LL_c \leq 120$;
- B-03iii: $LL_c > 120$.

Descrição dos dados do subgrupo B-03i

O grupo B-03i contempla os resultados de ensaios do grupo B-03 que apresentaram $LL_c < 50$. Assim, dentro deste subgrupo, foram investigados os resultados de 30 ensaios, oriundos de 3 publicações: Christaras (1991); Clemente (2018); Clemente et al. (2020).

Figura 37 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03i.

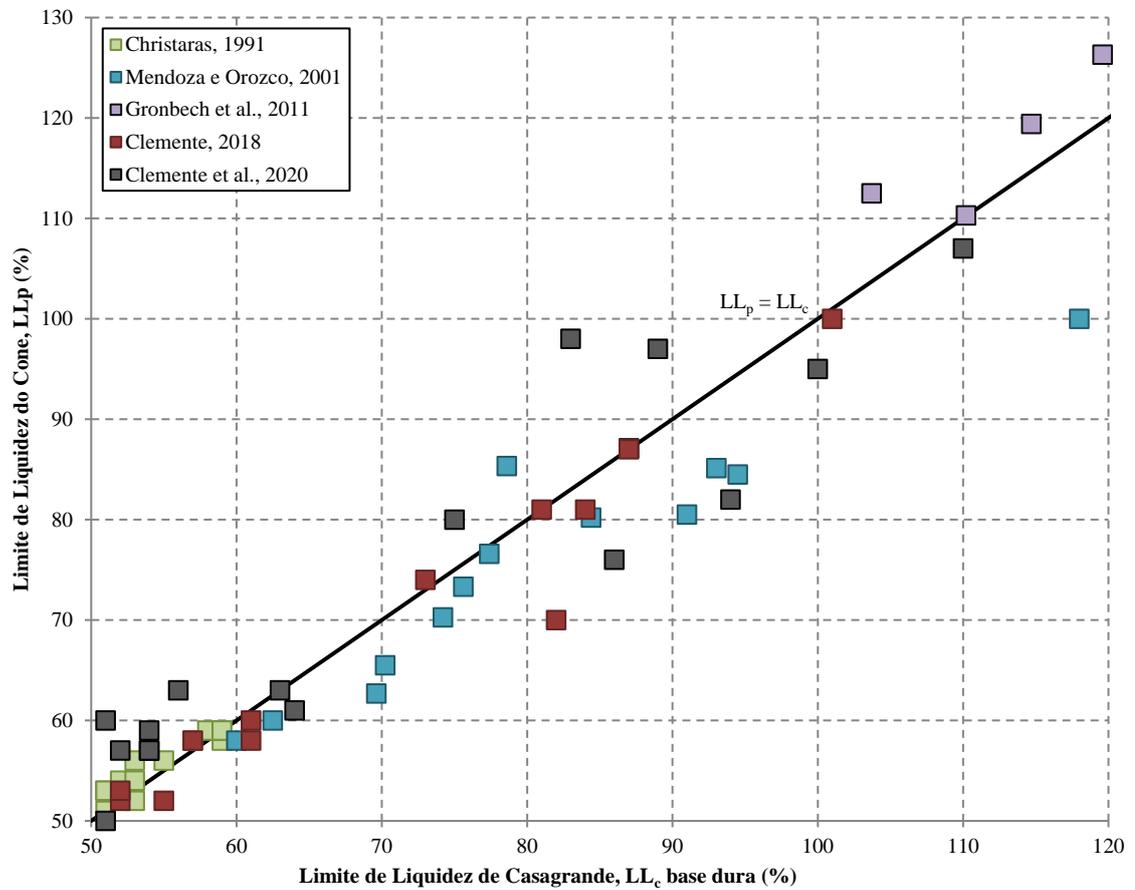


Fonte: Autor (2022)

Descrição dos dados do subgrupo B-03ii

O grupo B-03ii contempla os resultados de ensaios do grupo B-03 que apresentaram $50 < LL_c \leq 120$. Assim, dentro deste subgrupo, foram investigados os resultados de 44 ensaios, oriundos de 4 publicações: Christaras (1991); Gronbech et al. (2011); Clemente (2018); Clemente et al. (2020).

Figura 38 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03ii.

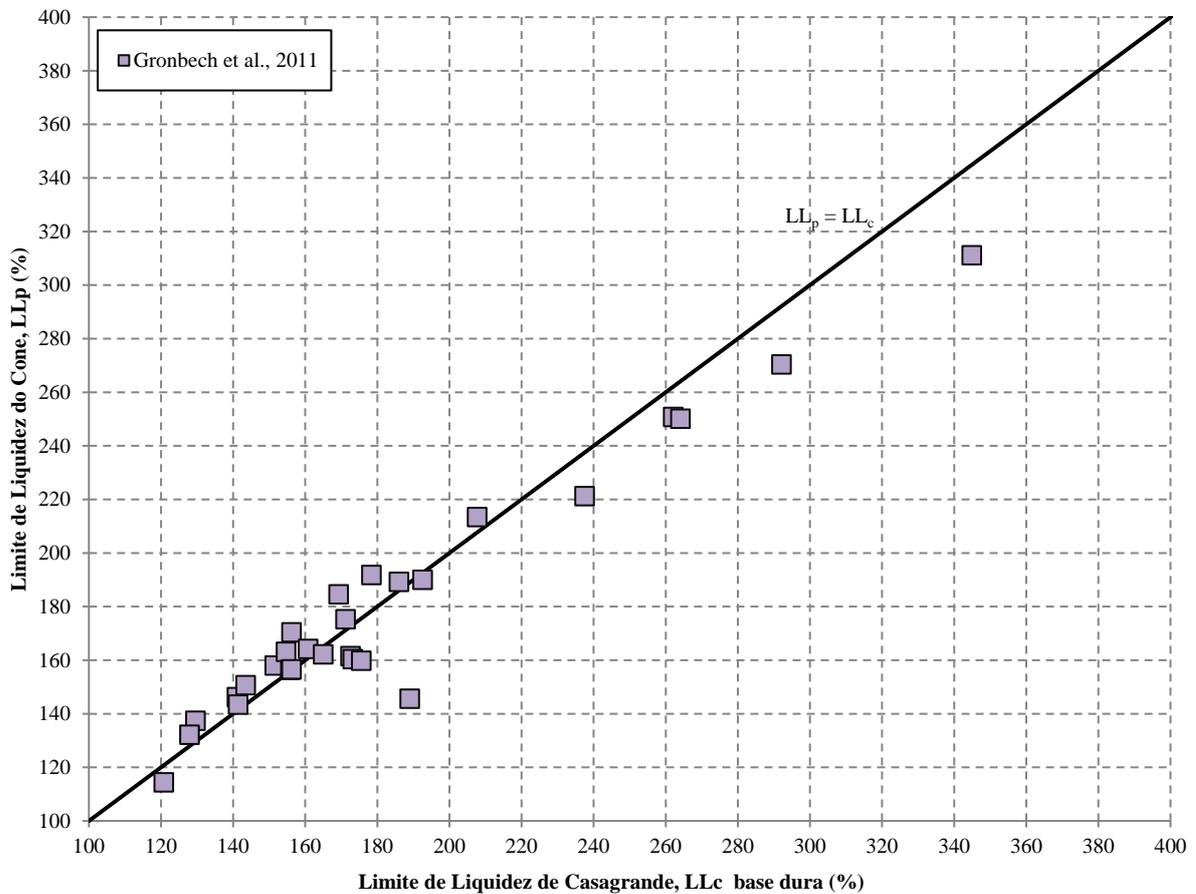


Fonte: Autor (2022)

Descrição dos dados do subgrupo B-03iii

O grupo B-03iii contempla os resultados de ensaios do grupo B-03 que apresentaram $LL_c > 120$. Assim, dentro deste subgrupo, foram investigados os resultados de 16 ensaios, publicados por Gronbech et al. (2011).

Figura 39 - Valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do grupo B-03iii.



Fonte: Autor (2022)

3.2 ANÁLISES ESTATÍSTICAS

Os métodos estatísticos empregados nesta pesquisa foram aplicados através de ferramentas computacionais como os programas *Microsoft Excel*, *Minitab* e do Programa-R.

O *Microsoft Excel* é um popular editor de planilhas que possui ferramentas de cálculo e de construção de gráficos, além do armazenamento, organização e análise de diferentes tipos de dados. O *Minitab* é um programa estatístico utilizado para a análise de dados e estatística. R é uma linguagem de programação de computadores, utilizada para a manipulação de dados estatísticos e gráficos, amplamente difundida para o desenvolvimento de *software* estatístico e análise de dados.

Com o intuito de identificar a presença de valores atípicos às amostras de solo selecionadas nesta pesquisa, os resultados dos ensaios descritos no item 3.1.2 foram submetidos a avaliação de *outliers* e ao recurso gráfico *boxplot* do programa *Minitab*. Na identificação dos *outliers* e na análise dos gráficos *boxplot* foram avaliados os resultados de LL_c, LL_p e IP das

amostras selecionadas para um nível de significância de 5%. Cabe ressaltar que o processo de identificação dos *outliers* por meio dos recursos mencionados ocorre de forma interativa no programa Minitab. A cada eliminação dos dados identificados como *outliers*, o processo de análise é repetido para a identificação de outros possíveis resultados. Assim, a cada nova avaliação, como o conjunto de dados sobre alteração, são identificados novos *outliers*, até que enfim todos sejam removidos e não restem *outliers* para aquele novo grupo de dados.

Após a eliminação dos *outliers* os dados de cada grupo foram submetidos a análises de regressão para o estabelecimento de correlações de regressão lineares e não lineares entre os dados, e obtenção do coeficiente de determinação (R^2). As correlações lineares e não lineares foram comparadas através dos resultados obtidos para a raiz quadrada do erro médio (RMSE), erro médio absoluto (MAE), coeficiente de determinação (R^2), valores mínimos e máximos, média e desvio padrão (σ) das estimativas de LL_p .

Para a comparação das correlações de regressão com dados da literatura foram selecionadas correlações lineares e não lineares entre LL_c e LL_p obtidas para diferentes intervalos de LL , assim como nesta pesquisa. As correlações lineares obtidas para os grupos B-01 e B-02 foram comparadas, observados os referidos intervalos de LL , com as correlações lineares publicadas por Di Matteo (2012) e Crevelin e Bicalho (2019) e com as correlações não lineares de O'Kelly et al. (2018). Essas publicações foram selecionadas por apresentarem correlações obtidas a partir de resultados de ensaios de diversos países, compiladas também da literatura, assim como proposto nesta pesquisa. Para o grupo B-03, os resultados foram comparados com as correlações lineares de Christaras (1991) e Gronbech et al. (2011). Não foram encontradas na literatura correlações não lineares para o grupo B-03.

Por último, para a validação dos modelos de correlações propostos, as correlações lineares foram submetidas às análises de resíduos e aos testes de hipóteses para verificação da normalidade do modelo e independência das variáveis. O gráfico de resíduos foi utilizado para a avaliação da homoscedasticidade, auxiliando na identificação de pontos de dados atípicos ao modelo proposto. A normalidade do modelo foi verificada pela análise gráfica dos histogramas de frequência dos resíduos e do gráfico de Probabilidade Normal e pelos testes de aderência de *Kolmogorov-Smirnov*, teste de normalidade de *Shapiro-Wilk* e teste de independência dos resíduos de *Durbin-Watson*. Os testes citados foram avaliados pela comparação do valor-p com o nível de significância do modelo. Assim, para o teste de *Durbin-Watson*, valor-p < 0,05, indica que os resíduos não são independentes. E, para os testes de *Kolmogorov-Smirnov* e *Shapiro-Wilk*, valor-p > 0,05 confirma a normalidade da distribuição da variável independente.

3.3 ANÁLISES DE REDES NEURAIS ARTIFICIAIS

O treinamento das RNAs nesta pesquisa fez uso do software *RapidMiner*. O *RapidMiner* trata-se de um programa de ciência de dados utilizado na preparação de dados, aprendizado de máquina e análise preditiva. O software realiza o treinamento de redes *feed-forward* (FNN) a partir do uso de um conjunto de entradas e de saídas previamente organizados pelo usuário e permite a escolha das porcentagens de divisão nos conjuntos de treinamento, validação e teste e do número de neurônios na camada escondida. A rede neural *feed-forward* (FNN) é treinada por um algoritmo de retropropagação perceptron multicamadas (MLP) e a função de ativação padrão é a função sigmoide.

Nesta pesquisa, várias redes foram treinadas e posteriormente analisadas seguindo critérios estatísticos para avaliação de sua capacidade de previsão dos resultados. Para as etapas de treinamento e validação cruzada, os dados foram separados na proporção de 70% para a fase de treinamento e 30% para a validação cruzada. A seleção das amostras para esses dois conjuntos foi realizada pelo software de maneira aleatória.

Adotou-se como dados de entrada LL_c , LL_p , IP e classificação SUCS dos solos, e como saída LL_p . Para a utilização da classificação SUCS das amostras de solo como dado de entrada, foram atribuídos valores numéricos as classificações existentes, para a compreensão desses dados por meio das redes. A fim de avaliar o desempenho das redes quanto aos dados de entrada, também foram investigados os resultados obtidos utilizando apenas LL_c e LL_p como dados de entrada.

Optou-se nesta pesquisa por investigar a previsão de LL_p a partir de RNAs estruturadas em uma e duas camadas ocultas. Assim, foram treinadas diversas redes para cada grupo e subgrupo descritos neste capítulo. O número de neurônios adotados na primeira camada oculta foi de 3, 6, 9, 12, 15 e 18, enquanto na segunda camada foi adotado a metade do número de neurônios da primeira camada: 1, 3, 4, 6, 7 e 9 (com os números arredondados). Portanto, para cada grupo de amostras foram treinadas 24 redes, sendo as primeiras apenas com uma camada oculta e as demais com duas camadas ocultas.

A Tabela 20 seguinte apresenta a topologia utilizada para a identificação das redes neurais treinadas.

Tabela 20 – Resumo das topologias das redes neurais treinadas.

Rede	Neurônios 1ª camada	Neurônios 2ª camada	Dados de Entrada
RN1.1	3	-	LL_p , LL_c , IP e Classificação SUCS
RN1.2	6	-	LL_p , LL_c , IP e Classificação SUCS
RN1.3	9	-	LL_p , LL_c , IP e Classificação SUCS

Rede	Neurônios 1ª camada	Neurônios 2ª camada	Dados de Entrada
RN1.4	12	-	LL _p , LL _c , IP e Classificação SUCS
RN1.5	15	-	LL _p , LL _c , IP e Classificação SUCS
RN1.6	18	-	LL _p , LL _c , IP e Classificação SUCS
RN1.7	3	1	LL _p , LL _c , IP e Classificação SUCS
RN1.8	6	3	LL _p , LL _c , IP e Classificação SUCS
RN1.9	9	4	LL _p , LL _c , IP e Classificação SUCS
RN1.10	12	6	LL _p , LL _c , IP e Classificação SUCS
RN1.11	15	7	LL _p , LL _c , IP e Classificação SUCS
RN1.12	18	9	LL _p , LL _c , IP e Classificação SUCS
RN2.1	3	-	LL _p e LL _c
RN2.2	6	-	LL _p e LL _c
RN2.3	9	-	LL _p e LL _c
RN2.4	12	-	LL _p e LL _c
RN2.5	15	-	LL _p e LL _c
RN2.6	18	-	LL _p e LL _c
RN2.7	3	1	LL _p e LL _c
RN2.8	6	3	LL _p e LL _c
RN2.9	9	4	LL _p e LL _c
RN2.10	12	6	LL _p e LL _c
RN2.11	15	7	LL _p e LL _c
RN2.12	18	9	LL _p e LL _c

Fonte: Autor (2022)

As combinações de arquiteturas e dados de entrada propostos para as RNAs foram comparadas através dos resultados obtidos para a raiz quadrada do erro médio (RMSE), erro médio absoluto (MAE), coeficiente de determinação (R^2), valores mínimos e máximos, média e desvio padrão (σ) das estimativas de LL_p para a indicação de um modelo mais adequado a cada grupo de dados analisado.

Na retropropagação, os valores de saída são comparados com a resposta correta para calcular o valor de alguma função de erro predefinida. Esse erro é realimentado através da rede. Usando essas informações, o algoritmo ajusta os pesos de cada conexão para reduzir o valor da função de erro em uma pequena quantidade. Esse processo é repetido por um número de vezes especificada pelo usuário. Nesta pesquisa, foram definidos mil ciclos de treinamento para o treinamento da RNA. A otimização deste processo é interrompida se o erro de treinamento ficar abaixo do valor especificado pelo usuário. Nesta pesquisa, o valor de erro utilizado para essa interrupção foi de 0,0001.

3.4 METODOLOGIA DE AVALIAÇÃO DOS RESULTADOS

Nesta pesquisa foram confrontados os resultados obtidos por meio das redes neurais artificiais e das correlações de regressão, a fim de indicar o método que melhor se adequa a cada um dos modelos propostos. A rede neural artificial representa uma ferramenta adicional para métodos empíricos convencionais de regressão e que permite o processamento de uma grande quantidade de dados, com aplicabilidade à problemas complexos e com alto grau de linearidade (Bicalho et al., 2022).

Foram, portanto, comparados os resultados obtidos para a raiz quadrada do erro médio (RMSE), erro médio absoluto (MAE), coeficiente de determinação (R^2), valores mínimos e máximos, média e desvio padrão (σ) das estimativas de LL_p obtidas das correlações de regressão e das RNAs.

A raiz quadrada do erro médio (RMSE) é a raiz do somatório do quadrado da diferença entre os valores estimados e medidos, dividido pelo número de observações, o que consequentemente atribui maior peso aos maiores erros. Valores mais próximos de zero indicam um melhor desempenho do modelo.

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n (LL_{p\text{estimado}} - LL_{p\text{medido}})^2} \quad (34)$$

O erro médio absoluto (MAE) é calculado a partir da média dos erros absolutos.

$$MAE = \sqrt{\frac{1}{n} \sum_1^n |LL_{p\text{estimado}} - LL_{p\text{medido}}|} \quad (35)$$

O coeficiente de determinação (R^2) mede a relação existente entre duas variáveis distintas. Conforme citado no item 2.5.2, o valor de R^2 mais próximo de 1 indica um melhor ajuste da correlação.

Os valores mínimos e máximos se referem aos menores e maiores valores do conjunto de dados. A média é obtida através da soma de todos os valores do conjunto de dados dividida pelo número de elementos do conjunto.

O desvio padrão (σ) mede a dispersão dos resultados e indica o quanto o conjunto de dados é uniforme. É calculado conforme a equação 20.

4. RESULTADOS E DISCUSSÕES

Neste capítulo são apresentados os resultados das análises estatísticas e das redes neurais artificiais obtidos para os dados experimentais investigados (pares de LL_c e LL_p compilados e descritos no item 3.1.2) e as correspondentes discussões desses resultados.

4.1 RESULTADOS PARA OS DADOS DO GRUPO B-01

4.1.1 Avaliação estatística de amostras do grupo B-01

Os dados do grupo B-01 foram submetidos a testes estatísticos para a identificação e tratamento dos *outliers* conforme a metodologia citada no item 3.2. Foram identificados 41 *outliers* do total de amostras de solo selecionadas, resultando o grupo, portanto, em 380 amostras aprovadas para serem utilizadas no estabelecimento das correlações empíricas e no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c, LL_p e IP são apresentados no Anexo I.

Para o grupo B-01 nota-se que a maioria dos dados identificados como *outliers* são de resultados de ensaios de altos valores de LL. A maior parcela desses dados parte das publicações de Wasti e Bezirci (1986), Deka et al. (2009), Cevallos Luna (2012) e Bicalho et al. (2017). Destaca-se que dentre os resultados publicados por Bicalho et al. (2017), todos aqueles referentes às misturas de bentonita e areia fina (C>60%) foram identificados como *outliers*, sendo, portanto, excluídos dos dados utilizados para o estabelecimento das correlações e das análises das RNAs para este grupo. A Tabela 21 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 21 – Descrição estatística dos dados investigados no grupo B-01.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	380	41	48,759	17,422	303,539	35,73	11,600	33,050	50,650	61,675	100,000
LL _p	380	41	50,686	17,422	303,511	34,37	14,500	35,000	51,800	63,950	102,600
IP	380	41	22,414	12,381	153,280	55,24	1,180	12,222	21,600	31,675	55,100

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

Fonte: Autor (2022)

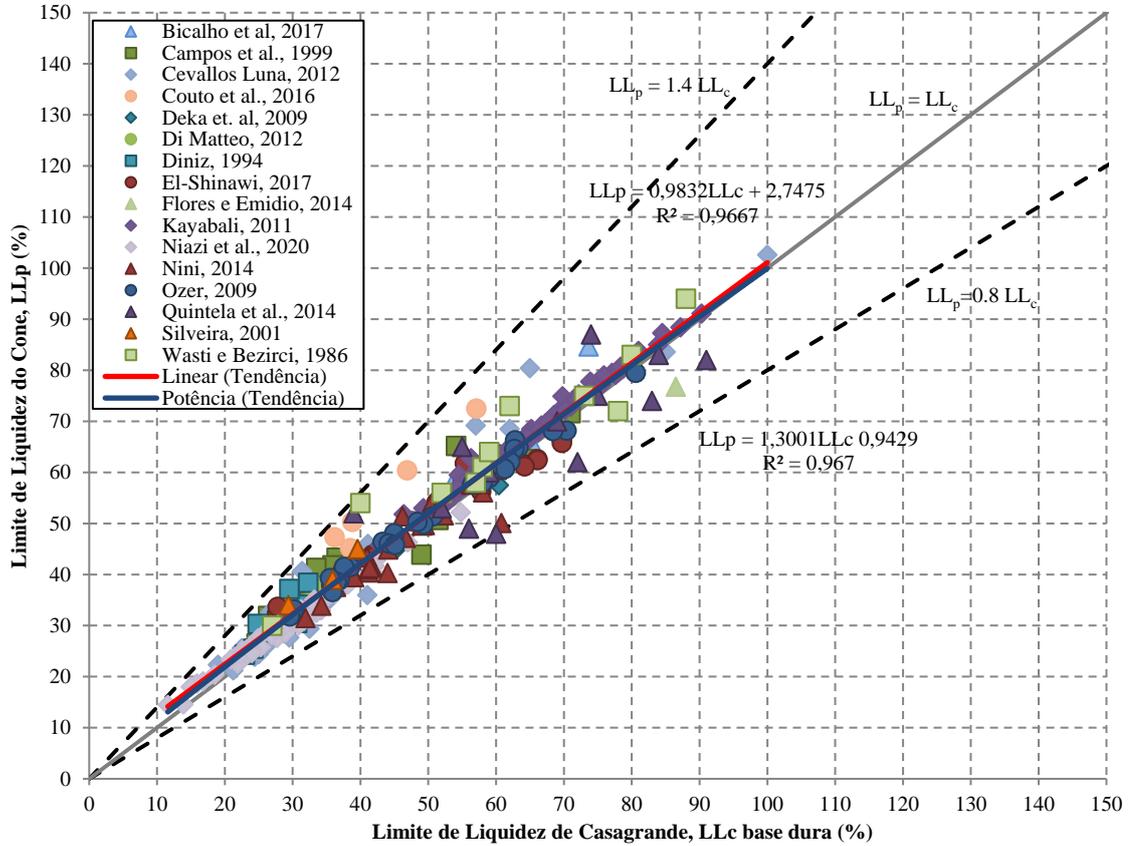
Os resultados apresentados na Tabela 21 foram comparados com a descrição estatística das 601 amostras designadas inicialmente para o grupo B-01 (Tabela 12). Foi observado, após a seleção e tratamento dos dados, uma redução de cerca de 70% nos valores de desvio padrão, 80% da variância e 50% do coeficiente de variação para as variáveis LL_c, LL_p e IP. Esses resultados indicam que o tratamento dos dados possibilitou a formação de um grupo com dados menos dispersos, mais próximos da média. Entretanto, os coeficientes de variação ainda se mostraram superiores a 30% para as variáveis LL_c e LL_p e superior a 50% para a

variável IP. Portanto, embora os resultados positivos obtidos com o tratamento dos dados, o grupo B-01 ainda é composto de dados heterogêneos.

A variabilidade dos dados do grupo B-01 é descrita no item 3.1.2.1. Na Tabela 13, o resumo das classificações SUCS das amostras indica que, embora os dados concentrem-se em sua maioria (mais de 60%) entre argilas de alta e baixa compressibilidade (CL e CH), cerca de 19% das amostras são classificadas como siltes de baixa compressibilidade (MH). O posicionamento das amostras no ábaco de Casagrande e consequentemente sua classificação SUCS, corroboram as observações feitas quanto a granulometria, já que o grupo contempla amostras de solos com frações de silte e areia. A mineralogia dos solos estudados não é informada em todas as publicações, no entanto, Campos et al. (1999), Di Matteo (2012), Flores e Emidio (2014), Quintela et al. (2014) e Bicalho et al. (2017) relatam amostras compostas por caulinitas. Apenas uma amostra publicada por Flores e Emídio (2014) trata da mistura de caulinita (80%) e bentonita (20%).

Conforme citado, os dados de valores mais altos de LL foram de uma maneira geral identificados como *outliers*, assim, o grupo B-01 foi limitado ao intervalo de $LL_c < 100$. A Figura 40 apresenta os pares de dados e as correlações obtidas.

Figura 40 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do grupo B-01 e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-01:

$$LL_p = 0,9832LL_c + 2,7475 \quad R^2 = 0,9667 \quad (36)$$

$$LL_p = 1,3001LL_c^{0,9429} \quad R^2 = 0,967 \quad (37)$$

Nota-se da Figura 40, que as amostras de solo do grupo B-01 encontram-se entre as linhas de $LL_p = 1,4 LL_c$ e $LL_p = 0,8 LL_c$ e cerca de 82% das amostras apresenta resultados experimentais de LL_p maiores do que LL_c . As correlações de regressão se situam entre as linhas $LL_p=LL_c$ e $LL_p = 1,4 LL_c$. Observa-se que as correlações preveem resultados muito próximos, principalmente para a faixa de valores de LL_p entre 34 e 68. A Tabela 22 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-01.

Tabela 22 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01.

Regressão	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
Linear - RL	3,174	1,795	0,967	14,153	101,068	50,687	17,130
Não linear - RNL	3,162	1,944	0,967	13,112	99,949	50,595	17,127

Fonte: Autor (2022)

O resumo estatístico apresentado na Tabela 22 confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Ambos, apresentam inclusive um mesmo coeficiente de determinação (R^2). O modelo não linear possui erro médio absoluto (MAE), raiz quadrada do erro médio (RMSE) e desvio padrão ligeiramente menores.

As correlações obtidas para o Grupo B-01 foram ainda comparadas com correlações obtidas da literatura, conforme a metodologia descrita no item 3.2. Os resultados destas comparações constam das Tabela 23 e Tabela 24.

Tabela 23 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o Grupo B-01.

LL _c	LL _p			
	Grupo B-01	Di Matteo (2012)	Crevelin e Bicalho (2019)	Crevelin e Bicalho (2019)
	LL _c <120	20 < LL _c <50	20 < LL _c <100	50 < LL _c <460
20	22,41	22,20	21,05	29,30
30	32,24	32,20	30,90	36,69
40	42,08	42,20	40,75	44,08
50	51,91	52,20	50,60	51,47
60	61,74	-	60,45	58,86
70	71,57	-	70,30	66,25
80	81,40	-	80,15	73,64
90	91,24	-	90,00	81,03
100	101,07	-	99,85	88,42

Fonte: Autor (2022)

A correlação linear obtida para o grupo B-01 apresenta resultados mais próximos aos obtidos por Di Matteo (2012) para valores de LL entre 20 e 50 para misturas de areia fina e silte argiloso e Crevelin e Bicalho (2019) para amostras de caulinitas e ilitas com valores de LL entre 20 e 100. A semelhança observada entre as correlações pode ser explicada pela quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa. Assim, em relação a correlação de Di Matteo (2012), a correlação obtida no presente estudo para o grupo B-01 inclui os dados publicados pelo próprio pesquisador e os dados de Wasti & Bezirci (1986) e Ozer (2009) também utilizados por Di Matteo (2012) na obtenção de sua correlação. Quando comparada com Crevelin e Bicalho (2019), foram utilizados em comum

os dados de Di Matteo (2012), Flores e Emidio (2014), Quintela et al. (2014) e Bicalho et al. (2017).

Tabela 24 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o Grupo B-01.

LL _c	LL _p		
	Grupo B-01	O' Kelly et al. (2018)	O' Kelly et al. (2018)
	LL _c <120	LL _c <120	LL _c <600
20	21,91	22,82	24,25
30	32,12	33,14	34,22
40	42,13	43,18	43,70
50	51,99	53,02	52,83
60	61,74	62,70	61,69
70	71,40	72,25	70,32
80	80,98	81,70	78,77
90	90,50	91,05	87,07
100	99,95	100,32	95,23

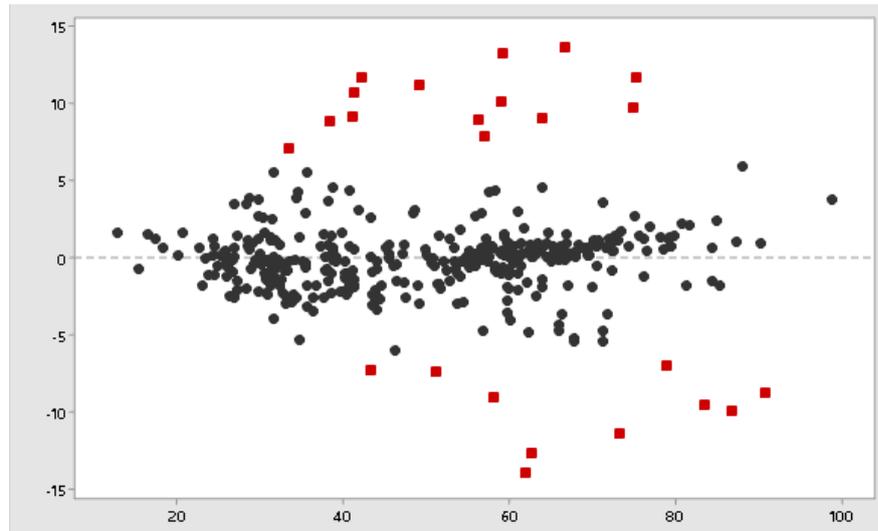
Fonte: Autor (2022)

A correlação não linear obtida para o grupo B-01 mostra-se muito próxima da publicada por O'Kelly et al. (2018) para LL_c de base dura < 120%. Isso pode ser justificado pelo número coincidente de dados analisados em ambas as pesquisas, já que os dados analisados por O'Kelly et al. (2018) incluem, entre outros, os dados publicados por Wasti & Bezirci (1986), Ozer (2009) e Di Matteo (2012).

Além das semelhanças observadas quanto aos dados utilizados para a obtenção das correlações, destaca-se o intervalo de LL adotado como outro fator de influência para a proximidade entre os resultados obtidos na presente pesquisa e os resultados das publicações analisadas. Conforme supracitado, após o tratamento dos dados com a eliminação dos *outliers*, o grupo B-01 foi limitado ao intervalo de $LL_c < 100$.

A correlação linear obtida para o grupo B-01 foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Os testes, conforme metodologia descrita no item 3.2 foram utilizados como critério para a validação do modelo. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos apresentado na Figura 41.

Figura 41 – Relatório de diagnóstico de resíduos - Grupo B-01.

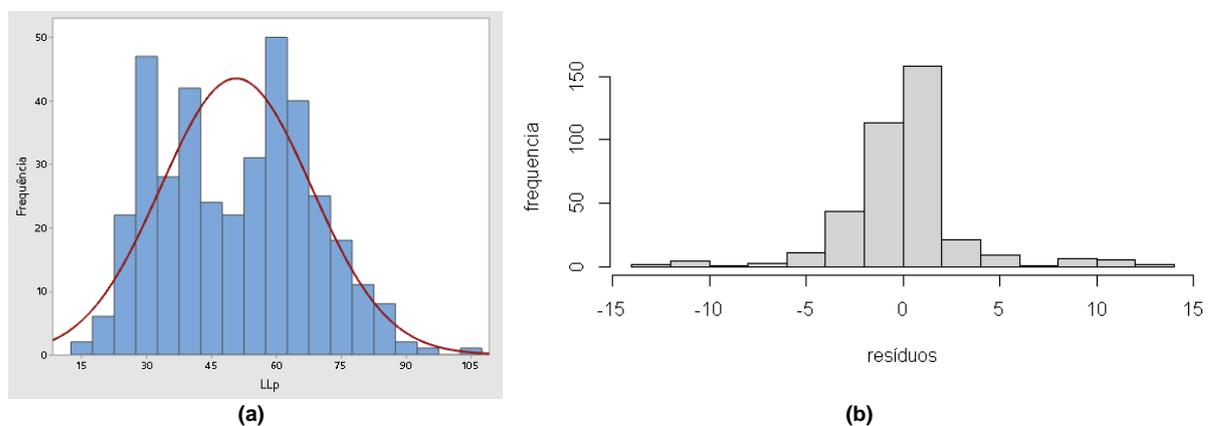


Fonte: Autor (2022)

Na Figura 41 é possível identificar 24 pontos de dados com grandes resíduos. Estes pontos, representados pelos pontos vermelhos da figura, não estão bem ajustados pela equação proposta. Desse total, 16 possuem pontos atípicos de LL_c , o que pode fazer com que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 42.

Figura 42 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Grupo B-01.



Fonte: Autor (2022)

Através da Figura 42 infere-se que a variável dependente apresenta um gráfico de distribuição com assimetria moderada à direita. Logo, observa-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 25 apresenta o resumo dos resultados dos testes estatísticos realizados para o grupo B-01.

Tabela 25 – Resumo dos resultados dos testes estatísticos para o Grupo B-01.

Teste	valor-p	Conclusão
Teste t (intercepto)	3,07E-08	Parâmetro significativo
Teste t (x)	2,00E-16	Parâmetro significativo
Teste F	2,20E-16	Parâmetro significativo
Teste KS	8,76E-11	Distribuição dos resíduos não é normal
Teste DW	2,15E-06	Resíduos não são independentes

Fonte: Autor (2022)

Os resultados dos testes de hipóteses (Testes F e t) indicam que os parâmetros de regressão são significantes em relação aos dados observados, uma vez que o valor-p obtido é maior do que a significância de 5%.

O resultado do teste de aderência de *Kolmogorov Smirnov* (teste KS) aponta para a distribuição não normal dos resíduos, com valor-p inferior a significância de 5%. Este resultado vai de encontro à assimetria observada na análise dos resíduos por meio dos histogramas.

O resultado do teste de *Durbin Watson* (DW), que avalia a independência entre os resíduos, indica que os resíduos do grupo B-01 não são independentes, já que o valor-p é menor que o nível de significância de 5%. O resultado do teste DW é de 1,5337, ou seja, por estar próximo de 2, indica que não há autocorrelação entre as variáveis independentes.

Ressalta-se que, neste caso, em função do tamanho da amostra, o teste *Shapiro-Wilk* (SW) não foi utilizado.

Portanto, apesar do alto valor de R^2 ($> 0,96$) para a regressão linear, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios dos testes estatísticos indicaram que a correlação linear apresentada não é significativa.

4.1.2 Aplicação das RNAs nas amostras do grupo B-01

Utilizando os dados filtrados nas análises estatísticas, as 380 amostras aprovadas no grupo B-01 foram submetidas as análises das redes neurais artificiais. Um resumo estatístico das RNAs treinadas para este grupo é apresentado na Tabela 26.

Tabela 26 – Resumo estatístico das estimativas de LL_p obtidas com as RNAs para as amostras do grupo B-01.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	3,109	2,520 \pm 1,821	0,961	14,414	74,564	40,154	13,286
RN1.2	3,265	2,795 \pm 1,688	0,963	16,342	75,034	40,601	13,370
RN1.3	2,902	2,122 \pm 1,979	0,961	13,861	75,138	39,730	13,775
RN1.4	2,751	1,879 \pm 2,009	0,963	15,023	73,791	39,249	13,398
RN1.5	2,940	1,803 \pm 2,323	0,962	12,905	72,809	37,751	13,719
RN1.6	2,767	1,933 \pm 1,981	0,963	15,005	73,871	39,379	13,409
RN1.7	2,881	2,106 \pm 1,966	0,962	18,135	74,471	39,513	13,197
RN1.8	2,803	1,964 \pm 2,000	0,963	17,452	73,813	39,306	13,148
RN1.9	2,756	1,907 \pm 1,990	0,963	16,153	73,930	39,265	13,331
RN1.10	2,776	1,683 \pm 2,207	0,962	13,791	73,599	38,359	13,549
RN1.11	3,490	3,067 \pm 1,665	0,962	16,490	75,758	40,946	13,387
RN1.12	2,920	2,267 \pm 1,841	0,963	14,456	74,723	39,903	13,526
RN2.1	3,017	1,926 \pm 2,322	0,954	15,326	75,046	39,186	13,715
RN2.2	3,036	1,799 \pm 2,446	0,955	14,065	73,711	38,324	13,549
RN2.3	3,225	1,952 \pm 2,567	0,955	13,160	73,051	37,622	13,608
RN2.4	2,981	1,848 \pm 2,339	0,955	13,859	74,752	38,869	13,817
RN2.5	3,079	2,104 \pm 2,248	0,955	14,917	75,487	39,595	13,752
RN2.6	3,033	1,991 \pm 2,289	0,955	15,143	74,255	39,289	13,433
RN2.7	3,688	3,027 \pm 2,106	0,952	19,161	76,836	40,855	13,539
RN2.8	3,078	1,922 \pm 2,404	0,952	16,977	74,854	38,861	13,518
RN2.9	3,600	2,949 \pm 2,064	0,953	18,434	76,315	40,739	13,430
RN2.10	3,160	2,217 \pm 2,252	0,953	17,428	75,123	39,607	13,381
RN2.11	3,218	2,373 \pm 2,174	0,954	16,866	75,480	39,945	13,525
RN2.12	3,110	1,830 \pm 2,515	0,953	15,503	73,677	38,180	13,437

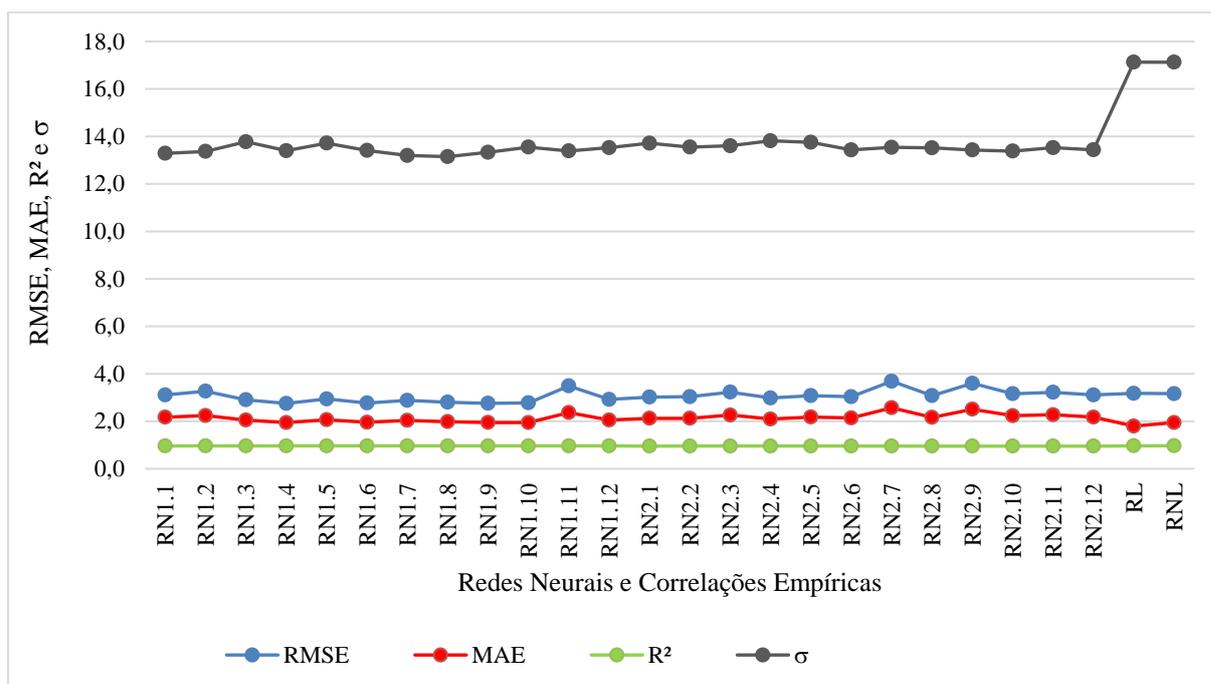
Fonte: Autor (2022)

Observa-se com as análises das RNAs que as arquiteturas utilizadas para cada conjunto de dados de entrada possuem resultados muito próximos. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) proporcionou um aumento da média dos erros MAE de 2,06 para 2,24, RMSE de 2,95 para 3,19, e um leve aumento do valor médio de R^2 de 0,95 para 0,96 em comparação com os obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Analisando os erros MAE e RMSE das redes RN1.1 a RN1.12, as arquiteturas com 12 neurônios em uma única camada (RN1.4) e com 9 neurônios na primeira camada e 4 na segunda (RN1.9) apresentaram resultados menores quando comparados com as demais arquiteturas.

Em comparação com os resultados obtidos pelas análises de regressão, se compararmos apenas o critério de coeficiente de determinação, R^2 , observa-se um resultado relativamente maior para as equações de regressão ($R^2 = 0,97$). No entanto, quando analisado desvio

padrão (σ), observa-se que as RNAs apresentam resultados menores do que as das correlações de regressão. O desvio padrão médio observado para as RNAs é de 13,50 enquanto o das correlações de regressão é de 17,10. A Figura 43 apresenta graficamente a comparação das RNAs e correlações de regressão.

Figura 43 – Resultados estatísticos das redes neurais e regressões do Grupo B-01.



Fonte: Autor (2022)

Da Figura 43 nota-se a semelhança entre os modelos de correlação e RNAs, com destaque para as pequenas diferenças entre os resultados. Uma exceção, no entanto, são os maiores valores de σ das correlações de regressão. Portanto, os resultados indicam que as RNAs se adequam melhor a previsão de valores de LL_p para o grupo de dados B-01.

4.1.3 Avaliação estatística dos dados do subgrupo B-01i

Para os dados do subgrupo B-01i, aplicada a metodologia citada no item 3.2, foram identificados 3 *outliers*. Portanto, resultaram neste subgrupo 185 amostras aprovadas para serem utilizadas no estabelecimento das correlações empíricas, bem como no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c , LL_p e IP são apresentados no Anexo I.

Foram identificados como *outliers* os dois maiores resultados de LL das amostras de Nini (2014) e o maior dentre os dados publicados por Couto et al. (2016). A Tabela 27 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 27 - Descrição estatística dos dados investigados no subgrupo B-01i.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	185	3	33,379	8,030	64,481	24,06	11,600	27,800	32,500	39,000	49,800
LL _p	185	3	35,373	8,307	68,999	23,48	14,500	29,550	34,100	41,300	54,000
IP	185	3	12,493	5,976	35,715	47,84	1,180	8,433	12,367	16,350	28,000

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

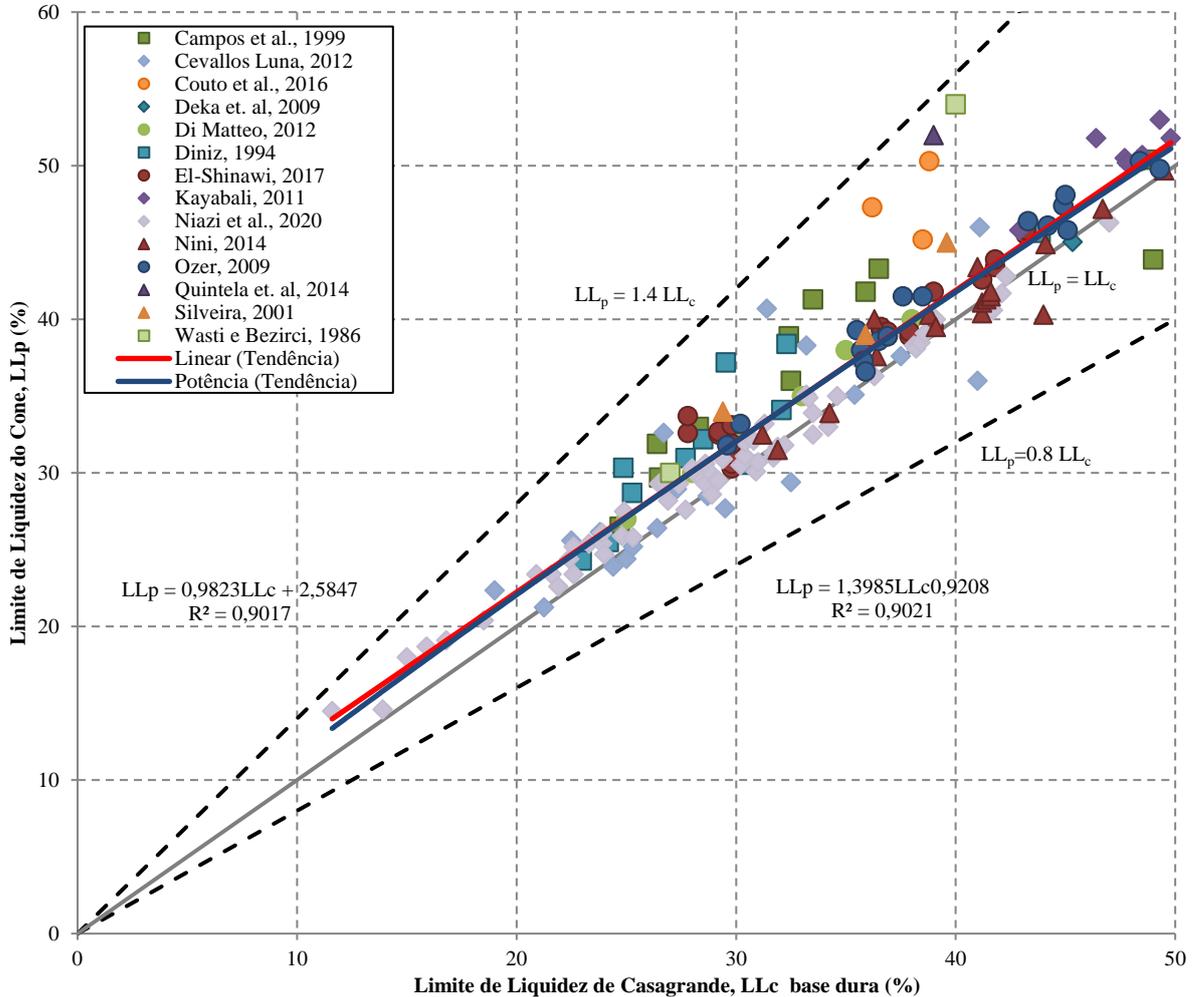
Fonte: Autor (2022)

Os resultados apresentados na Tabela 27, quando comparados com a descrição estatística do grupo B-01, indicam que este subgrupo apresenta dados menos dispersos. Foi observado para o subgrupo B-01i, uma redução de cerca de 52% nos valores de desvio padrão, 77% da variância e 24% do coeficiente de variação para as variáveis LL_c, LL_p e IP quando comparado com o grupo B-01. Esses resultados indicam que a subdivisão dos dados do subgrupo B-01i resultou em um conjunto de dados menos dispersos face aos obtidos para o grupo B-01. No entanto, a variável IP possui um coeficiente de variação próximo de 50%. Portanto, embora o tratamento dos dados, nota-se a variabilidade das amostras de solo para este conjunto de dados.

Os dados do subgrupo B-01i, em função do intervalo de LL em que se estão inseridos, são classificados como argilas e siltes de baixa compressibilidade (CL-ML, CL e ML). Embora mais de 60% dos dados sejam descritos como amostras de argila (CL), a granulometria das amostras selecionadas para este subgrupo, de maneira análoga ao grupo B-01, é caracterizada por parcelas de siltes e areias, justificando, portanto, a variação observada quanto a classificação das amostras. A mineralogia dos solos descritas para as amostras obtidas de Campos et al. (1999), Di Matteo (2012) e Quintela et al. (2014) indicam a presença de argilas caulinitas. No entanto, a mineralogia não é informada para os resultados experimentais das demais publicações que compõem o subgrupo de dados.

A Figura 44 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 44 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01i e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-01i:

$$LL_p = 0,9823LL_c + 2,5847 \quad R^2 = 0,9017 \quad (38)$$

$$LL_p = 1,3985LL_c^{0,9208} \quad R^2 = 0,9021 \quad (39)$$

Nota-se da Figura 44, que as amostras de solo analisadas se encontram entre as linhas $LL_p=0,8 LL_c$ e $LL_p = 1,4 LL_c$ e cerca de 81% dos resultados de LL_p observados são maiores que LL_c . As correlações de regressão obtidas se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 1,4 LL_c$, seguindo a tendência observada para o grupo B-01. Observa-se que as correlações preveem resultados muito próximos, principalmente para valores de LL_p entre 23 e 42. A

Tabela 28 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-01i.

Tabela 28 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01i.

Regressão	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
Linear - RL	2,597	1,750	0,902	13,979	51,503	35,373	7,888
Não linear - RNL	2,594	1,755	0,902	13,360	51,106	35,281	7,851

Fonte: Autor (2022)

O resumo estatístico confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Ambos, apresentam inclusive um mesmo coeficiente de determinação (R^2) e valores praticamente coincidentes para os erros RMSE e MAE. O modelo não linear, no entanto, possui σ ligeiramente menor. A análise de regressão para o subgrupo B-01i segue, portanto, a tendência observada para o grupo B-01, com resultados semelhantes entre as correlações de regressão.

As correlações obtidas para o subgrupo B-01i foram ainda comparadas com correlações da literatura, conforme a metodologia citada no item 3.2. Os resultados destas comparações constam das Tabela 29 e Tabela 30.

Tabela 29 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-01i.

LL_c	LL_p		
	Subgrupo B-01i	Di Matteo (2012)	Crevelin e Bicalho (2019)
	$LL_c < 50$	$20 < LL_c < 50$	$20 < LL_c < 100$
20	22,23	22,20	21,05
30	32,05	32,20	30,90
40	41,88	42,20	40,75
50	51,70	52,20	50,60

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-01i, seguindo a tendência observada no Grupo B-01, apresenta resultados mais próximos aos obtidos por Di Matteo (2012) para valores de LL entre 20 e 50 para misturas de areia fina e silte argiloso e Crevelin e Bicalho (2019) para amostras de caulinitas e ilitas com valores de LL entre 20 e 100. Assim, em relação a correlação de Di Matteo (2012), a correlação obtida no presente estudo para o subgrupo B-01i inclui os dados publicados pelo próprio pesquisador, bem como os dados de Wasti & Bezirci (1986) e Ozer (2009) também utilizados por Di Matteo (2012) na obtenção de sua correlação. Quando comparada com Crevelin e Bicalho (2019), foram utilizados em comum apenas os dados de Di Matteo (2012 e Quintela et al. (2014).

Tabela 30 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-01i.

LL_c	LL_p		
	Subgrupo B-01i	O' Kelly et al. (2018)	O' Kelly et al. (2018)
	$LL_c < 50$	$LL_c < 120$	$LL_c < 600$
20	22,06	22,82	24,25
30	32,05	33,14	34,22
40	41,77	43,18	43,70
50	51,29	53,02	52,83

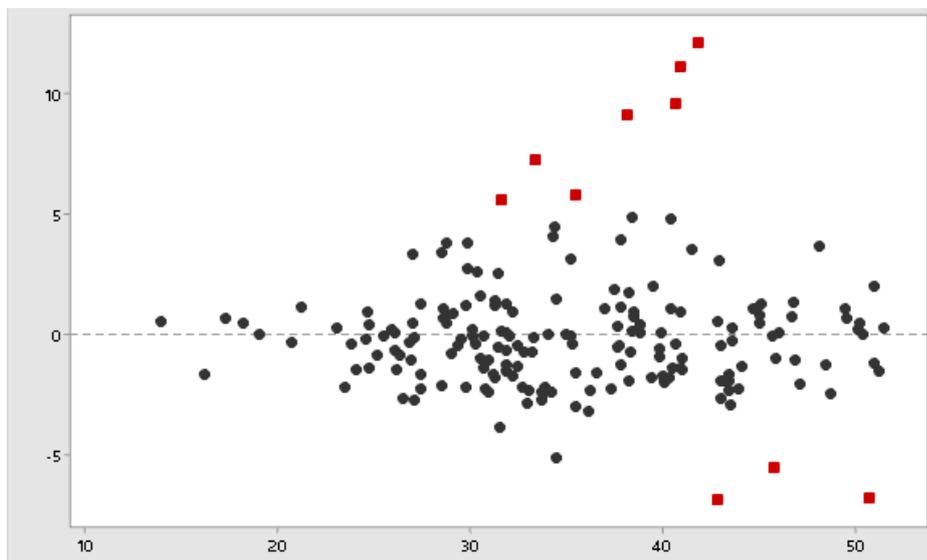
Fonte: Autor (2022)

A correlação não linear obtida para o subgrupo B-01i, assim como observado para o Grupo B-01, mostra-se muito próxima da publicada por O'Kelly et al. (2018) para LL_c de base dura < 120%. Os dados analisados por O'Kelly et al. (2018) incluem, entre outros, os dados publicados por Wasti & Bezirci (1986), Ozer (2009) e Di Matteo (2012), ambos utilizados na obtenção da correlação não linear do subgrupo B-01i.

A semelhança observada entre as correlações desta pesquisa e as da literatura, pode ser explicada pela quantidade de dados utilizados em comum para a obtenção das correlações, bem como ao intervalo de LL analisado, assim como observado para o Grupo B-01.

A correlação linear obtida para o subgrupo B-01i foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, o gráfico de resíduos da Figura 45 foi utilizado para a verificação da homoscedasticidade do modelo.

Figura 45 - Relatório de diagnóstico de resíduos – Subgrupo B-01i.

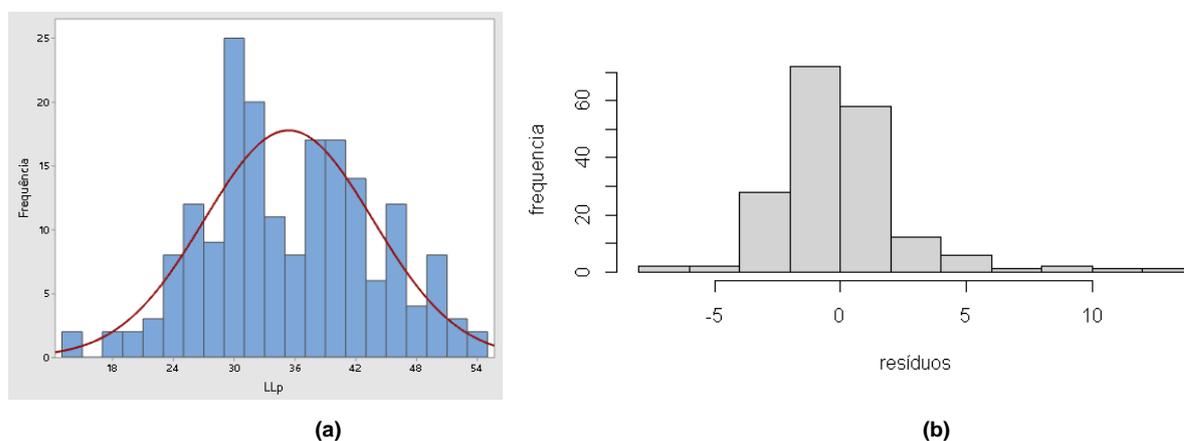


Fonte: Autor (2022)

Na Figura 45 são identificados 10 pontos de dados com grandes resíduos. Estes pontos, representados pelos pontos vermelhos da figura, não estão bem ajustados pela equação proposta. Desse total, 3 possuem pontos atípicos de LL_c , o que pode fazer com que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 46.

Figura 46 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-01i.



Fonte: Autor (2022)

Através da Figura 46 infere-se que a variável dependente apresenta um gráfico de distribuição com assimetria moderada à esquerda. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 31 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-01i.

Tabela 31 – Resumo dos resultados dos testes estatísticos para o subgrupo B-01i.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,00196	Parâmetro significativo
Teste t (x)	2,00E-16	Parâmetro significativo
Teste F	2,20E-16	Parâmetro significativo
Teste KS	1,76E-08	Distribuição dos resíduos não é normal
Teste DW	2,52E-05	Resíduos não são independentes

Fonte: Autor (2022)

Analisando os resultados da Tabela 31, verifica-se que os resultados obtidos para o subgrupo B-01i são similares aos obtidos para o Grupo B-01.

Assim, os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados. O resultado do teste de aderência de Kolmogorov Smirnov (KS) aponta para a distribuição não normal dos resíduos, em consonância com a assimetria observada na análise dos resíduos por meio dos histogramas. O resultado do teste de Durbin Watson (DW), indica que os resíduos não são independentes, e, o resultado do teste DW é de 1,4158, ou seja, por se aproximar de 0, indica que há autocorrelação positiva entre as variáveis independentes.

Portanto, apesar de um R^2 alto ($> 0,90$), e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios dos testes estatísticos dos resíduos indicaram que a correlação linear apresentada não é significativa.

4.1.4 Aplicação das RNAs nas amostras do grupo B-01i

Utilizando os dados filtrados nas análises estatísticas, as 185 amostras aprovadas no grupo B-01i foram submetidas as análises das redes neurais artificiais. Um resumo do desempenho das RNAs treinadas para este subgrupo é apresentado na Tabela 32.

Tabela 32 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01i.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	2,806	1,935 \pm 2,031	0,753	29,215	50,415	41,362	4,893
RN1.2	2,691	1,721 \pm 2,069	0,757	28,721	50,237	40,814	4,982
RN1.3	2,849	2,065 \pm 1,962	0,759	28,929	51,226	41,564	5,088
RN1.4	2,675	1,711 \pm 2,057	0,760	28,241	50,336	40,670	5,107
RN1.5	2,709	1,857 \pm 1,972	0,767	29,313	50,438	41,301	4,889
RN1.6	2,982	2,274 \pm 1,928	0,763	29,139	52,063	41,919	5,232
RN1.7	3,377	2,778 \pm 1,920	0,757	30,204	50,382	42,666	4,709
RN1.8	2,770	1,932 \pm 1,986	0,754	28,802	49,692	41,263	4,868
RN1.9	2,702	1,822 \pm 1,996	0,761	28,824	49,735	41,111	4,861
RN1.10	2,724	1,879 \pm 1,972	0,763	28,955	49,783	41,273	4,897
RN1.11	2,918	2,180 \pm 1,941	0,760	29,704	50,439	41,814	4,827
RN1.12	2,658	1,679 \pm 2,060	0,760	29,359	49,383	40,523	4,871
RN2.1	3,020	2,233 \pm 2,033	0,733	28,938	50,762	41,683	4,993
RN2.2	2,919	2,071 \pm 2,057	0,734	29,044	50,948	41,366	5,022
RN2.3	2,918	1,791 \pm 2,303	0,734	27,238	49,469	39,893	5,105
RN2.4	2,823	1,881 \pm 2,105	0,733	27,992	50,028	40,743	5,001
RN2.5	3,387	2,717 \pm 2,023	0,733	29,656	51,559	42,473	5,049
RN2.6	2,807	1,837 \pm 2,122	0,733	28,482	49,920	40,693	4,855
RN2.7	3,177	2,435 \pm 2,041	0,730	29,496	50,214	42,068	4,768

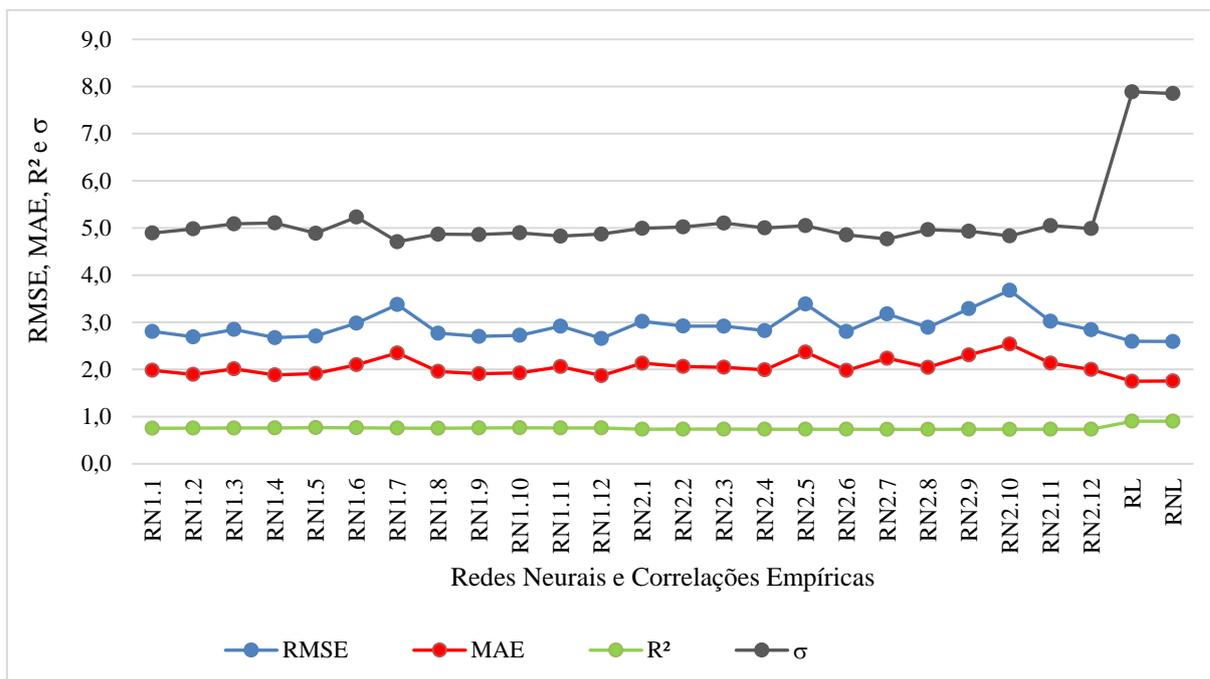
Rede	RMSE	MAE	R²	Mínimo	Máximo	Média	σ
RN2.8	2,893	2,060 \pm 2,031	0,730	28,185	49,600	41,189	4,964
RN2.9	3,289	2,582 \pm 2,037	0,731	29,435	50,939	42,293	4,933
RN2.10	3,678	3,109 \pm 1,965	0,732	30,515	51,536	42,975	4,832
RN2.11	3,022	2,233 \pm 2,036	0,733	28,924	50,984	41,680	5,050
RN2.12	2,842	1,793 \pm 2,206	0,733	27,631	49,416	40,228	4,988

Fonte: Autor (2022)

Observa-se, conforme resultados do Grupo B-01, que as arquiteturas utilizadas, para cada conjunto de dados de entrada, possuem resultados muito próximos. Assim, a utilização apenas dos dados de entrada LL_c e LL_p proporcionou um aumento da média dos erros MAE de 1,99 para 2,15, RMSE de 2,82 para 3,06, e uma diminuição do valor médio de R² de 0,76 para 0,73 em comparação com os obtidos com a utilização de LL_c, LL_p, IP e Classificação SUCS (RN1.1 a RN1.12). Analisando os erros MAE e RMSE, as arquiteturas com 12 neurônios em uma única camada (RN1.4), 15 neurônios em uma única camada (RN1.5) e 18 neurônios na primeira camada e 9 na segunda (RN1.12) apresentaram resultados ligeiramente menores.

Em comparação com os resultados obtidos pelas análises de regressão, se compararmos apenas o critério de coeficiente de determinação, R², observa-se um resultado relativamente maior para as equações de regressão (R² = 0,90). No entanto, quando analisado o desvio padrão (σ), menores valores são obtidos com as RNAs. O desvio padrão médio observado para as RNAs é de 4,95 enquanto o das correlações de regressão é de 7,87. A Figura 47 apresenta graficamente a comparação das RNAs e das correlações de regressão.

Figura 47 - Resultados estatísticos das redes neurais e regressões do Subgrupo B-01i.



Fonte: Autor (2022)

Da Figura 47 nota-se a semelhança entre os modelos de correlação e RNAs, com destaque para as pequenas diferenças entre os resultados. Uma exceção, no entanto, são os maiores valores de σ das correlações de regressão. Portanto, os resultados indicam que as RNAs se adequam melhor a previsão de valores de LL_p para o subgrupo de dados B-01i.

4.1.5 Avaliação estatística dos dados do subgrupo B-01ii

Para os dados do subgrupo B-01ii, aplicada a metodologia citada no item 3.2, foram identificados 16 *outliers*. Resultaram, portanto, 183 amostras aprovadas para a obtenção das correlações empíricas e no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c , LL_p e IP são apresentados no Anexo I.

Foram identificados como *outliers* os maiores valores de LL dentre os dados publicados por Wasti e Bezirci (1986), Bicalho et al. (2017), Deka et al. (2009), Ozer (2009), Flores e Emídio (2014) e Quintela et al. (2014). Os dados publicados por Nini (2014) e Kayabali (2011), tiveram como *outliers* os maiores valores de IP. O inverso ocorreu para os dados de Cevallos Luna (2012), onde as análises estatísticas apontaram para a identificação dos menores valores de IP como *outliers*.

A Tabela 33 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 33 - Descrição estatística dos dados investigados no subgrupo B-01ii.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	183	14	62,694	8,160	66,586	13,02	50,300	56,300	61,400	68,000	85,000
LL _p	183	14	64,453	8,442	71,275	13,10	48,000	58,000	62,800	69,800	87,300
IP	183	14	31,688	7,785	60,611	24,57	12,000	26,000	31,400	37,000	52,100

Legenda: N = número de amostras; N* = número de outliers; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

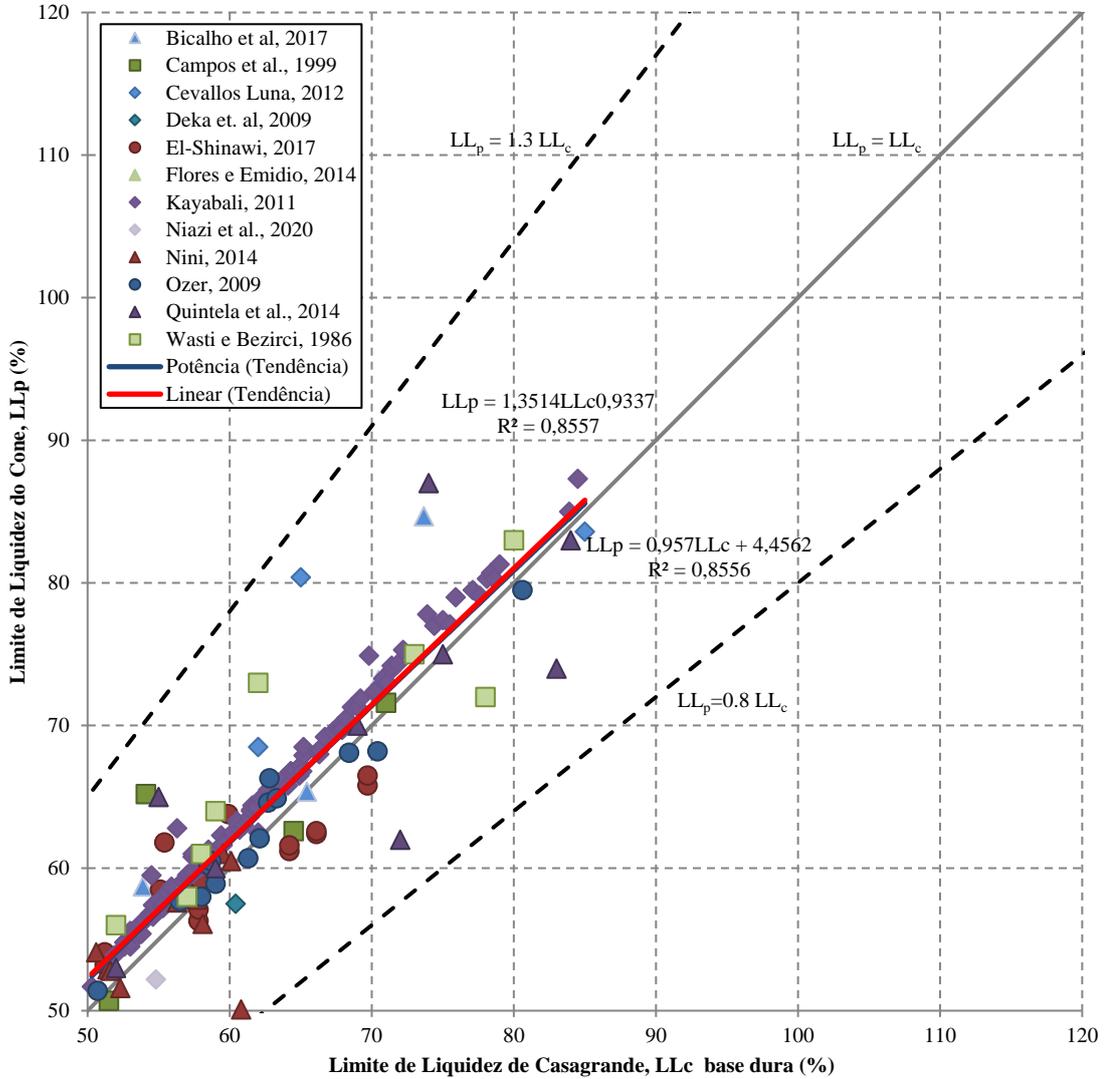
Fonte: Autor (2022)

Os resultados apresentados na Tabela 33, quando comparados com a descrição estatística do grupo B-01, indicam que este subgrupo apresenta dados menos dispersos. Foi observado uma redução de cerca de 48% nos valores de desvio padrão, 74% da variância e 60% do coeficiente de variação para as variáveis LL_c, LL_p e IP. Esses resultados indicam que o tratamento dos dados possibilitou a formação de um grupo com dados menos dispersos, mais próximos da média. O coeficiente de variação observado para as variáveis LL_c, LL_p e IP ficou abaixo de 25%, indicando que este subgrupo é composto por mais homogêneos se comparado ao grupo B-01.

Os dados do subgrupo B-01ii são classificados, segundo o sistema de classificação SUCS, como argilas e siltes de alta compressibilidade (CH e MH). Embora cerca de 62% das amostras sejam descritas como amostras de argila, a granulometria destas amostras, de maneira análoga ao grupo B-01, é caracterizada por parcelas de siltes e areias. Assim, aproximadamente 38% das amostras são classificadas como silte de alta compressibilidade (MH). A mineralogia dos solos descritas para as amostras obtidas de Bicalho et al. (2017), Campos et al. (1999), Flores e Emidio (2014) e Quintela et al. (2014) indicam a presença de argilas caulinitas. Apenas uma amostra de solo, obtida de Deka et al. (2009), se trata da mistura de argilas não expansivas e expansivas com minerais de montmorilonita, illita e quartzo. A mineralogia não é informada para os resultados experimentais das demais publicações que compõem o subgrupo de dados.

Conforme observado na Tabela 33, apesar do subgrupo B-01ii ser composto inicialmente por resultados de LL_c entre 50 e 120, após a filtragem dos dados, com a eliminação dos outliers, o valor de LL_c para o subgrupo é inferior a 100. A Figura 48 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 48 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01ii e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-01ii:

$$LL_p = 0,957LL_c + 4,4562 \quad R^2 = 0,8556 \quad (40)$$

$$LL_p = 1,3514LL_c^{0,9337} \quad R^2 = 0,8557 \quad (41)$$

Nota-se da Figura 48 que as amostras de solo do subgrupo B-01ii encontram-se entre as linhas $LL_p=LL_c$ e $LL_p = 1,3 LL_c$ e cerca de 83% das amostras apresenta resultados experimentais de LL_p maiores do que LL_c . As correlações de regressão se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 1,3 LL_c$. Observa-se que as correlações descritas para o subgrupo B-01ii preveem resultados muito próximos, principalmente para valores de LL_p entre 58 e 78.

Esses resultados podem ser ditos semelhantes aos obtidos para o grupo B-01, já que, ambas as correlações se mostram acima da linha de igualdade $LL_p=LL_c$.

A Tabela 34 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o subgrupo B-01ii.

Tabela 34 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01ii.

Regressão	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
Linear - RL	3,200	1,824	0,856	52,593	85,801	64,454	7,809
Não linear - RNL	3,204	1,847	0,856	52,425	85,562	64,362	7,806

Fonte: Autor (2022)

Confirma-se por meio do resumo estatístico, a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Ambos apresentam inclusive um mesmo coeficiente de determinação (R^2) e, os erros RMSE e σ praticamente coincidentes. O modelo não linear possui erro MAE ligeiramente menor. A análise de regressão para o subgrupo B-01ii segue, portanto, a tendência observada para o grupo B-01, com resultados semelhantes entre as correlações lineares e não lineares.

As correlações obtidas para o subgrupo B-01ii foram ainda comparadas com correlações obtidas da literatura, conforme metodologia citada no item 3.2. Os resultados destas comparações constam das Tabela 35 e Tabela 36.

Tabela 35 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-01ii.

LL_c	LL_p		
	Subgrupo B-01ii	Crevelin e Bicalho (2019)	Crevelin e Bicalho (2019)
	$50 < LL_c < 120$	$20 < LL_c < 100$	$50 < LL_c < 460$
50	51,70	50,60	51,47
60	61,52	60,45	58,86
70	71,35	70,30	66,25
80	81,17	80,15	73,64
90	90,99	90,00	81,03
100	100,81	99,85	88,42
110	110,64	0,00	95,81
120	120,46	0,00	103,20

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-01ii, apresenta resultados mais próximos aos obtidos por Crevelin e Bicalho (2019) para amostras de caulinitas e ilitas com valores de LL entre 20 e 100. A semelhança observada entre as correlações pode ser explicada pela quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa.

Assim, em relação a correlação de Crevelin e Bicalho (2019), foram utilizados em comum os dados de Flores e Emidio (2014), Quintela et al. (2014) e Bicalho et al. (2017).

Tabela 36 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-01ii.

LL_c	LL_p		
	Subgrupo B-01ii	O' Kelly et al. (2018)	O' Kelly et al. (2018)
	$50 < LL_c < 120$	$LL_c < 120$	$LL_c < 600$
50	52,13	53,02	52,83
60	61,81	62,70	61,69
70	71,38	72,25	70,32
80	80,85	81,70	78,77
90	90,25	91,05	87,07
100	99,58	100,32	95,23
110	108,85	109,51	103,26
120	118,06	118,64	111,19

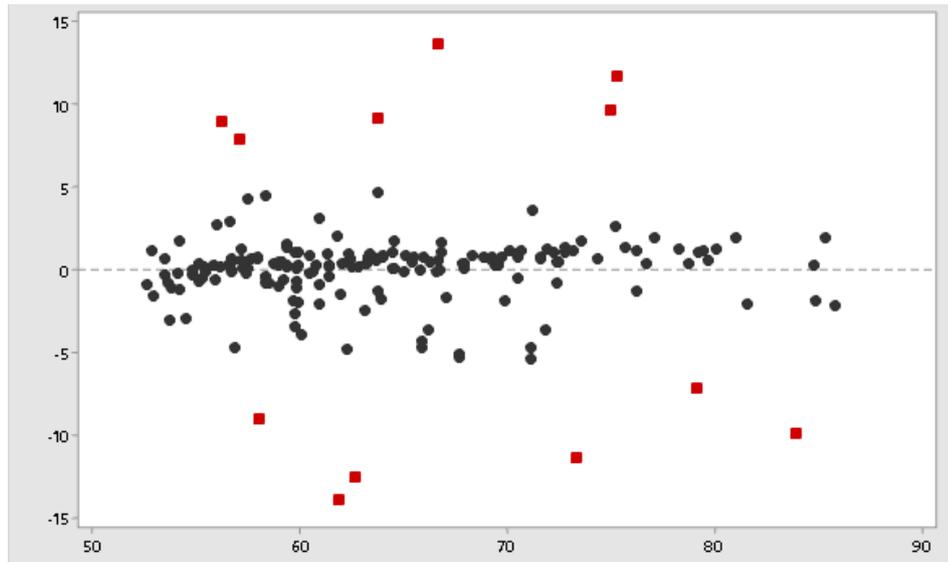
Fonte: Autor (2022)

A correlação não linear obtida para o subgrupo B-01ii, assim como observado para o grupo B-01, mostra-se muito próxima da publicada por O'Kelly et al. (2018) para LL_c de base dura < 120%. Em relação a correlação de O'Kelly et al. (2018) para LL_c de base dura < 600%, nota-se a semelhança entre os resultados para valores de LL_c entre 50 e 70. Os dados analisados por O'Kelly et al. (2018) incluem, entre outros, os dados publicados por Wasti & Bezirci (1986) e Ozer (2009), ambos utilizados para a obtenção da correlação não linear do subgrupo B-01ii.

A semelhança observada entre as correlações lineares e não lineares pode ser explicada pela quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa, bem como ao intervalo de LL analisado, assim como observado para o grupo B-01.

A correlação linear obtida para o subgrupo B-01ii foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos da Figura 49.

Figura 49 – Relatório de diagnóstico de resíduos – Subgrupo B-01ii.

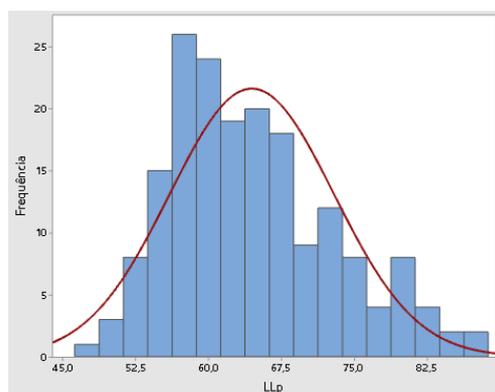


Fonte: Autor (2022)

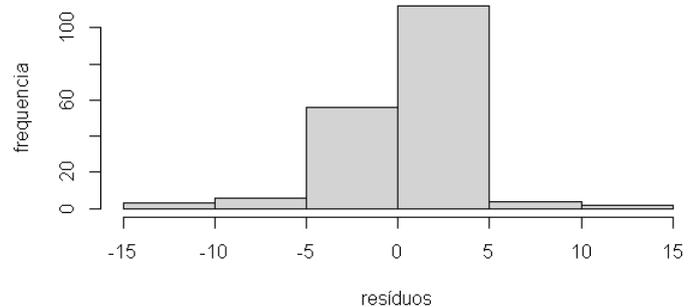
No gráfico da Figura 49 é possível identificar que 12 pontos de dados possuem grandes resíduos. Estes pontos, representados pelos pontos vermelhos da figura, não estão bem ajustados pela equação proposta. Desse total, 5 possuem pontos atípicos de LL_c , o que pode fazer com que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 50.

Figura 50 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-01ii.



(a)



(b)

Fonte: Autor (2022)

Através da Figura 50 observa-se que a variável dependente apresenta um gráfico de distribuição com assimetria moderada à direita. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 37 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-01ii.

Tabela 37 – Resumo dos resultados dos testes estatísticos para o subgrupo B-01ii.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,0169	Parâmetro significativa
Teste t (x)	2,00E-16	Parâmetro significativa
Teste F	2,20E-16	Parâmetro significativa
Teste KS	9,75E-05	Distribuição dos resíduos não é normal
Teste DW	7,64E-02	Resíduos independentes

Fonte: Autor (2022)

Os resultados dos testes de hipóteses (Testes F e t) indicam que os parâmetros de regressão são significantes em relação aos dados observados. O resultado do teste de aderência de Kolmogorov Smirnov (KS) aponta para a distribuição não normal dos resíduos, em consonância com a assimetria observada na análise dos resíduos por meio dos histogramas. O resultado do teste de Durbin Watson (DW) indica que os resíduos do subgrupo B-01ii são independentes, e, o resultado do teste DW é de 1,7983, próximo de 2, indica que não há autocorrelação entre as variáveis independentes.

Portanto, apesar de um R^2 significativo ($>0,85$) para a regressão linear, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios obtidos com a maioria dos testes estatísticos dos resíduos indicaram que a correlação linear apresentada não é significativa.

4.1.6 Aplicação das RNAs nas amostras do grupo B-01ii

Utilizando os dados filtrados nas análises estatísticas, as 183 amostras aprovadas no grupo B-01ii foram submetidas as análises das redes neurais artificiais. Um resumo do desempenho das RNAs treinadas é apresentado na Tabela 38.

Tabela 38 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01ii.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	5,012	3,512 \pm 3,576	0,822	53,457	87,884	67,773	10,842
RN1.2	5,381	3,878 \pm 3,730	0,814	53,756	90,313	68,373	11,194
RN1.3	5,263	3,748 \pm 3,694	0,821	54,272	89,503	68,376	11,004
RN1.4	5,271	3,744 \pm 3,710	0,828	54,938	89,458	68,566	11,099
RN1.5	5,585	4,156 \pm 3,732	0,818	54,925	90,476	68,943	11,206

Rede	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
RN1.6	5,330	3,790 ± 3,748	0,803	53,355	91,291	67,310	11,718
RN1.7	5,233	3,832 ± 3,563	0,823	54,654	87,159	68,426	10,675
RN1.8	4,698	3,230 ± 3,412	0,841	54,524	84,230	67,610	10,188
RN1.9	4,670	3,210 ± 3,392	0,841	54,317	84,132	67,550	10,222
RN1.10	4,714	3,274 ± 3,392	0,841	54,426	84,712	67,666	10,209
RN1.11	4,806	3,376 ± 3,421	0,845	54,978	84,442	68,086	10,118
RN1.12	4,524	3,190 ± 3,207	0,840	52,758	83,623	66,449	10,004
RN2.1	4,732	3,123 ± 3,555	0,847	54,104	84,475	67,919	9,929
RN2.2	4,992	3,428 ± 3,629	0,848	54,861	85,954	68,587	10,089
RN2.3	4,552	2,818 ± 3,575	0,848	53,400	85,254	67,447	10,358
RN2.4	4,541	2,804 ± 3,572	0,847	52,998	85,151	67,379	10,378
RN2.5	4,842	3,202 ± 3,632	0,847	54,278	85,898	68,250	10,175
RN2.6	4,421	2,847 ± 3,383	0,847	52,228	83,435	66,254	10,023
RN2.7	4,391	2,863 ± 3,329	0,848	53,045	82,229	66,244	10,107
RN2.8	4,493	2,881 ± 3,447	0,848	53,679	85,528	67,045	9,926
RN2.9	4,417	2,864 ± 3,362	0,848	52,976	82,387	66,441	9,921
RN2.10	4,644	3,127 ± 3,433	0,846	53,671	82,598	67,397	9,585
RN2.11	4,478	2,969 ± 3,352	0,847	53,092	81,833	66,506	9,622
RN2.12	4,719	3,104 ± 3,555	0,847	54,005	84,471	67,901	9,987

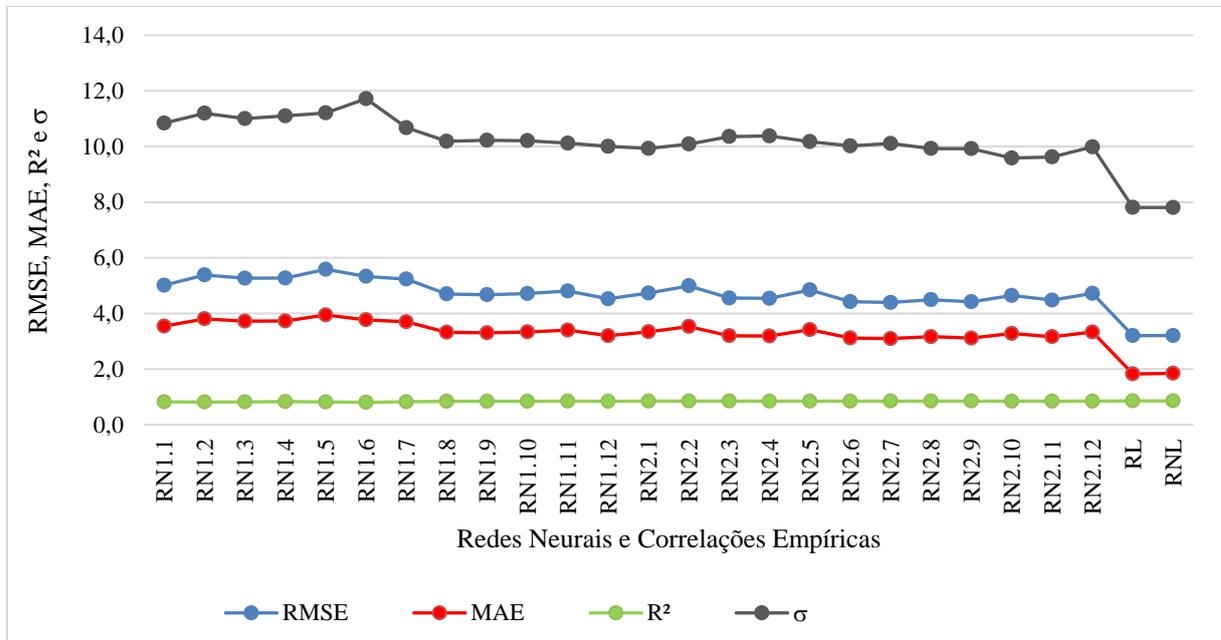
Fonte: Autor (2022)

Observa-se, em conformidade com os resultados do grupo B-01, que as arquiteturas utilizadas, para cada conjunto de dados de entrada, possuem resultados muito próximos. Entretanto, diferentemente do grupo B-01, a utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) apresentou menores erros e σ e um maior R², face aos obtidos com a utilização de LL_c, LL_p, IP e Classificação SUCS. A utilização destes dados proporcionou uma diminuição da média dos erros MAE de 3,56 para 3,24, RMSE de 5,04 para 4,60, e um leve aumento do valor médio de R² de 0,83 para 0,85 em comparação com os obtidos com a utilização de LL_c, LL_p, IP e Classificação SUCS (RN1.1 a RN1.12). Os valores de R² observados para as RN2.1 a RN2.12 são praticamente coincidentes. Analisando, no entanto, os erros MAE e RMSE e o σ , as arquiteturas com 18 neurônios em uma única camada (RN2.6) e com 3 neurônios na primeira camada e 1 na segunda (RN2.7), apresentaram resultados ligeiramente menores quando comparados com as demais arquiteturas.

Em comparação com os resultados obtidos pelas análises de regressão, se compararmos apenas o critério de coeficiente de determinação, R², observa-se um mesmo resultado para as RNAs RN2.1 a 2.12 e as equações de regressão (R² = 0,85). No entanto, quando analisado os erros RMSE e MAE e o desvio padrão (σ), observa-se que as RNAs apresentam resultados maiores do que as das correlações de regressão. O desvio padrão médio observado para as RNAs é de 10,40 enquanto o das correlações de regressão é de 7,80. Os erros médios RMSE e MAE são de 4,82 e 3,40 para as RNAs e de 3,20 e 1,84, respectivamente, para as

correlações de regressão. A Figura 51 apresenta graficamente a comparação das RNAs e correlações de regressão.

Figura 51 - Resultados estatísticos das redes neurais e regressões do Subgrupo B-01ii.



Fonte: Autor (2022).

Da Figura 51 nota-se as diferenças destacadas entre os modelos de correlação linear e não linear e entre as diferentes arquiteturas de RNAs. Observa-se, portanto, um resultado melhor para as equações de regressão, diferentemente da tendência observada para o grupo B-01 de uma maneira geral.

4.1.7 Avaliação estatística dos dados do subgrupo B-01iii

Para os dados do subgrupo B-01iii os gráficos de *outliers* e o diagramas de *boxplot* para os dados de LL_c , LL_p e IP (Anexo I) não evidenciaram a presença de *outliers*. Portanto, 34 amostras foram utilizadas para a obtenção das correlações empíricas e no treinamento das redes neurais artificiais.

A Tabela 39 apresenta um resumo estatístico das amostras de solo investigadas no subgrupo B-01iii.

Tabela 39 - Descrição estatística dos dados investigados no subgrupo B-01iii.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL_c	34	0	242,1	100,9	10179,0	41,67	129,2	168,8	199,9	312,4	526,0
LL_p	34	0	192,2	70,8	5017,1	36,86	98,0	139,8	167,4	254,8	355,9
IP	34	0	173,3	118,7	14080,5	68,48	20,4	86,1	153,8	266,3	488,0

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

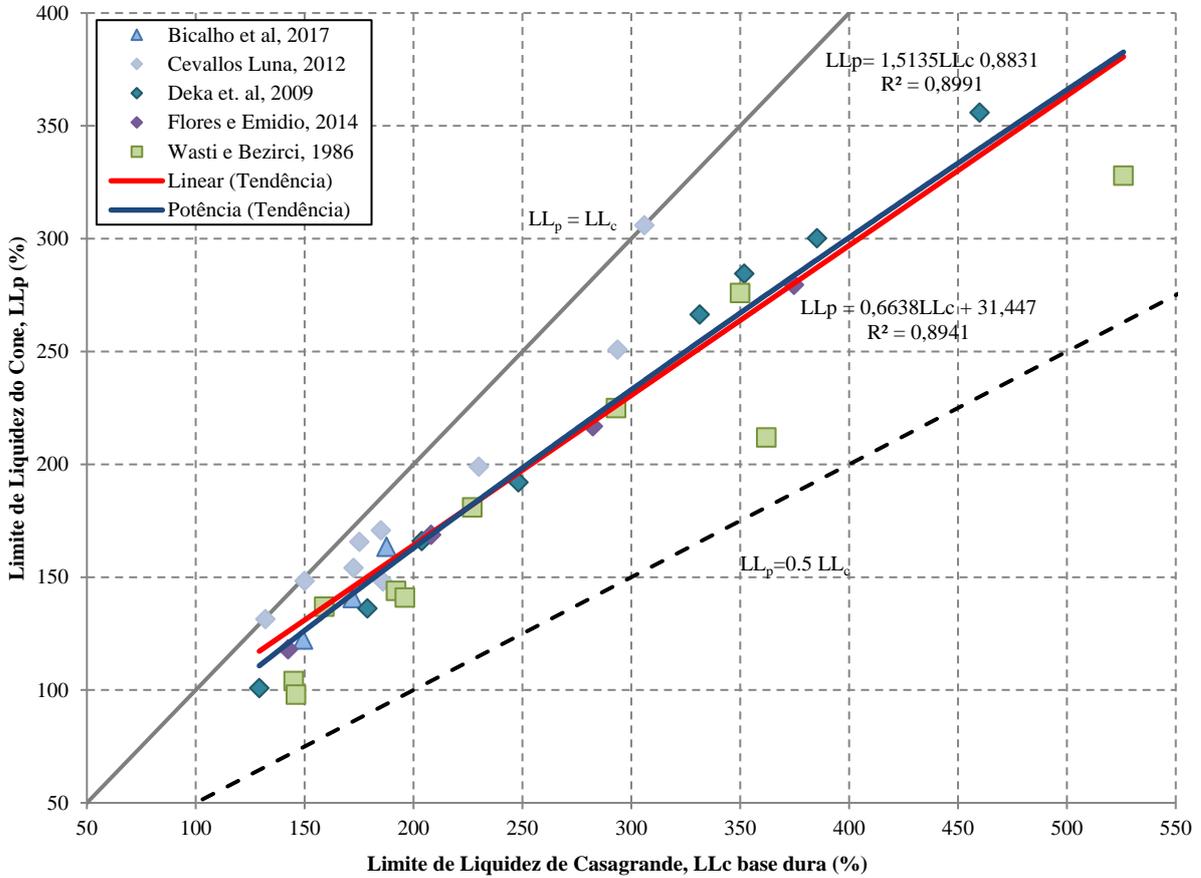
Fonte: Autor (2022)

Os resultados apresentados na Tabela 39 mostram altos valores de desvio padrão, variância e coeficiente de variação para este subgrupo de dados, quando comparados com os valores obtidos para o grupo B-01. Foi observado um aumento de cerca de 84% nos valores de desvio padrão, 97% da variância e 15% do coeficiente de variação para as variáveis LL_c , LL_p e IP. Esses resultados indicam que este subgrupo apresenta dados mais dispersos quando comparado com o grupo B-01 e com os subgrupos B-01i e B-01ii. Os coeficientes de variação observados para as variáveis LL_c , LL_p e IP, acima de 38%, corroboram que se trata de um conjunto de dados heterogêneos.

Os dados do subgrupo B-01iii são classificados, segundo o sistema de classificação SUCS, como argilas e siltes de alta compressibilidade (CH e MH). Embora mais de 73% das amostras sejam descritas como amostras de argila, a granulometria destas amostras, de maneira análoga ao grupo B-01, também é caracterizada por parcelas de siltes e areias. Um exemplo disto são as misturas de bentonita e areia fina ensaiadas por Bicalho et al. (2017) e que compõem este subgrupo de dados. Assim, cerca de 27% das amostras são classificadas como silte de alta compressibilidade (MH). Quanto a mineralogia, o subgrupo é composto de misturas de bentonita (Wasti e Bezirci, 1986), bentonita e areia (Bicalho et al., 2017), bentonita e caulinita (Flores e Emídio, 2014) e misturas de argila argilas expansivas (montmorilonita, illita e quartzo) e não expansivas (Deka et al., 2009).

Além das características das amostras de solo citadas, as diferenças entre os valores de LL_c e LL_p para este subgrupo de dados são maiores do que as observadas para os demais subgrupos. Essa dispersão dos resultados pode ser observada na Figura 52 que apresenta os pares de dados e as correlações avaliadas para os dados do subgrupo B-01iii.

Figura 52 - Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-01iii e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-01iii:

$$LL_p = 0,6638LL_c + 31,447 \quad R^2 = 0,8941 \quad (42)$$

$$LL_p = 1,5135LL_c^{0,8831} \quad R^2 = 0,8991 \quad (43)$$

Nota-se da Figura 52, que as correlações obtidas se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 0,5 LL_c$, bem como as amostras de dados analisadas, divergindo, portanto, dos resultados obtidos para o grupo B-01. Diferentemente dos demais resultados, observa-se que mais de 97% das amostras de solo analisadas para o subgrupo B-01iii apresenta resultados experimentais de LL_p menores do que LL_c . Além disso, os resultados experimentais apresentam diferenças de resultados maiores do que as observadas para o grupo B-01. Isso ocorre, conforme citado, devido a maior dispersão dos dados observadas para este subgrupo.

A Tabela 40 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-01iii.

Tabela 40 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01iii.

Regressão	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
Linear – RL	22,714	15,876	0,894	117,190	380,606	192,153	66,971
Não linear - RNL	22,339	15,267	0,899	110,751	382,723	191,340	69,869

Fonte: Autor (2022)

Observa-se por meio do resumo estatístico a proximidade dos resultados obtidos para os métodos de regressão linear e não linear, assim como observado na Figura 52. A correlação de regressão não linear apresentou MAE, RMSE e σ ligeiramente menores, e um coeficiente de determinação praticamente coincidente. A análise de regressão para o subgrupo B-01iii segue, portanto, a tendência observada para o grupo B-01, com resultados semelhantes entre as correlações lineares e não lineares. Vale ressaltar que, embora apresente um valor de R² alto (>0,89), os modelos de regressão apresentados, possuem altos valores de erro e desvio padrão para a previsão dos resultados de LL_p . Isso ocorre, conforme discutido, devido a variabilidade dos dados utilizados para a obtenção destas correlações.

As correlações obtidas para o subgrupo B-01iii foram ainda comparadas com correlações obtidas da literatura, conforme metodologia citada no item 3.2. Os resultados destas comparações constam das Tabela 41 e Tabela 42.

Tabela 41 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-01iii.

LL _c	LL _p	
	Subgrupo B-01iii	Crevelin e Bicalho (2019)
	LL _c >120	50<LL _c <460
120	111,10	103,20
150	131,02	125,37
200	164,21	162,32
250	197,40	199,27
300	230,59	236,22
350	263,78	273,17
400	296,97	310,12
450	330,16	347,07
500	363,35	384,02
550	396,54	420,97
600	429,73	457,92

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-01iii, apresenta resultados próximos aos obtidos por Crevelin e Bicalho (2019) para amostras de argilas montmorilonitas para valores de $LL_c < 600$. Nota-se que a semelhança entre as correlações é maior para valores de $LL_c < 400$.

Esta semelhança está relacionada com a concentração dos dados do subgrupo B-01iii nessa mesma faixa de valores e com a quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa. Assim, em relação a correlação de Crevelin e Bicalho (2019), foram utilizados em comum os dados de Wasti & Bezirci (1986), Deka et al. (2009), Flores e Emidio (2014) e Bicalho et al. (2017). Os dados de Mishra et al. (2011) utilizados também por Crevelin e Bicalho (2019), não foram utilizados neste estudo, tendo em vista que não apresentam a informação do IP do solo.

Tabela 42 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-01iii.

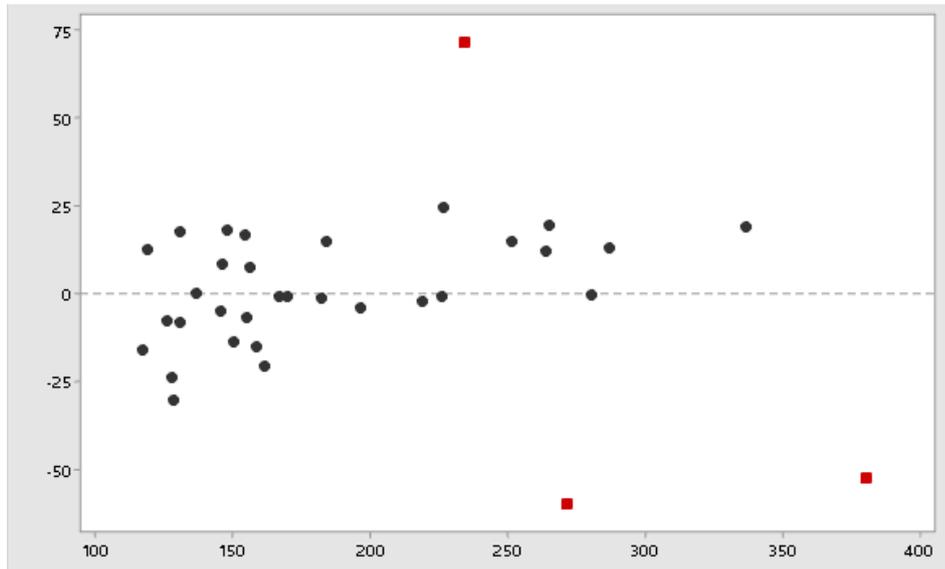
LL_c	LL_p	
	Subgrupo B-01iii	O' Kelly et al. (2018)
	$LL_c > 120$	$LL_c < 600$
120	103,78	111,19
150	126,38	134,41
200	162,94	171,64
250	198,43	207,49
300	233,09	242,27
350	267,08	276,19
400	300,51	309,39
450	333,45	341,97
500	365,97	374,01
550	398,10	405,57
600	429,90	436,70

Fonte: Autor (2022)

A correlação não linear obtida para o subgrupo B-01iii apresenta resultados próximos aos da correlação publicada por O'Kelly et al. (2018) para LL_c de base dura < 600%. Neste caso, os dados analisados por O'Kelly et al. (2018) incluem, entre outros, apenas os dados publicados por Wasti & Bezirci (1986), dados estes inseridos no grupo de dados para obtenção da correlação para o subgrupo B-01iii. A semelhança entre os resultados pode ser justificada pela faixa de LL adotada em ambas as pesquisas, bem como pela utilização de amostras de solo com alta variabilidade.

A correlação linear obtida para o subgrupo B-01iii foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos da Figura 53.

Figura 53 – Relatório de diagnóstico de resíduos – Subgrupo B-01iii.

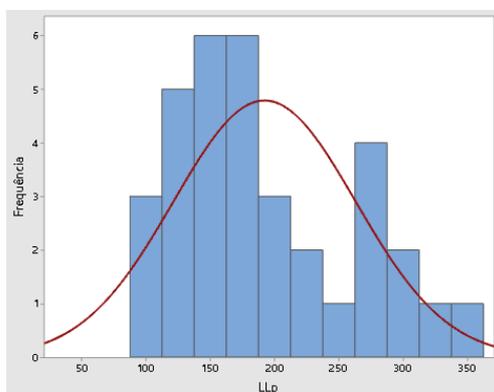


Fonte: Autor (2022)

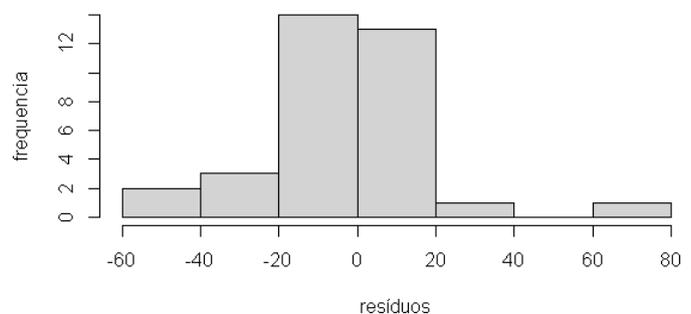
No gráfico da Figura 53 é possível identificar que 3 pontos de dados possuem grandes resíduos. Estes pontos, representados pelos pontos vermelhos da figura, não estão bem ajustados pela equação proposta. Desse total, apenas 1 possui pontos atípicos de LL_c . Devido ao comportamento dos dados no gráfico de resíduos, distribuídos em uma faixa horizontal de alta amplitude, observa-se também a tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 54.

Figura 54 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-01iii.



(a)



(b)

Fonte: Autor (2022)

Através da Figura 54 observa-se que a variável dependente apresenta um gráfico de distribuição com uma assimetria moderada à esquerda. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 43 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-01iii.

Tabela 43 – Resumo dos resultados dos testes estatísticos para o subgrupo B-01iii.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,00554	Parâmetro significativa
Teste t (x)	2,00E-16	Parâmetro significativa
Teste F	2,20E-16	Parâmetro significativa
Teste KS	9,93E-06	Distribuição dos resíduos não é normal
Teste DW	3,42E-03	Resíduos não são independentes

Fonte: Autor (2022)

Analisando os resultados da Tabela 43, verifica-se que os resultados obtidos para o subgrupo B-01iii são similares aos obtidos para o grupo B-01. Os resultados dos testes de hipóteses (Testes F e t) indicam que os parâmetros de regressão são significantes em relação aos dados observados. O resultado do teste de aderência de Kolmogorov Smirnov (teste KS) aponta para a distribuição não normal dos resíduos, em consonância com a assimetria observada na análise dos resíduos por meio dos histogramas. O resultado do teste de Durbin Watson (DW), indica que os resíduos não são independentes, e, o resultado do teste DW é de 1,1608, ou seja, por se aproximar de 0, indica que há autocorrelação positiva entre as variáveis independentes.

Portanto, apesar de um R^2 significativo ($> 0,89$) para a correlação de regressão linear, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios obtidos com a maioria dos testes estatísticos dos resíduos indicaram que a correlação linear apresentada não é significativa.

4.1.8 Aplicação das RNAs nas amostras do grupo B-01iii

Utilizando os dados filtrados nas análises estatísticas, as 34 amostras aprovadas no grupo B-01iii foram submetidas as análises das redes neurais artificiais. Um resumo do desempenho das RNAs é apresentado na Tabela 44.

Tabela 44 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-01iii.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	28,615	19,257 \pm 21,167	0,924	120,024	381,179	200,331	86,721
RN1.2	29,362	19,703 \pm 21,770	0,926	118,976	385,182	201,306	87,712

Rede	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
RN1.3	32,425	22,127 ± 23,702	0,924	124,782	390,807	206,140	87,381
RN1.4	30,033	19,930 ± 22,467	0,920	115,620	393,968	198,659	90,449
RN1.5	28,756	19,285 ± 21,331	0,921	113,335	390,439	196,300	90,182
RN1.6	29,565	19,636 ± 22,103	0,931	119,959	380,356	203,147	86,879
RN1.7	27,613	18,904 ± 20,128	0,920	124,336	370,226	200,364	83,223
RN1.8	27,837	18,908 ± 20,430	0,921	121,119	373,760	199,617	85,067
RN1.9	28,369	19,084 ± 20,990	0,922	119,234	377,311	199,623	86,492
RN1.10	29,638	20,048 ± 21,829	0,925	122,115	378,451	202,871	85,901
RN1.11	28,603	19,075 ± 21,313	0,924	117,497	379,719	199,730	87,479
RN1.12	30,659	20,542 ± 22,759	0,928	120,133	381,825	203,852	87,519
RN2.1	35,235	28,441 ± 20,800	0,931	129,827	366,446	213,041	81,984
RN2.2	33,643	25,664 ± 21,754	0,931	125,176	373,813	210,037	84,921
RN2.3	32,852	24,766 ± 21,584	0,932	124,570	374,792	209,027	84,869
RN2.4	35,133	27,821 ± 21,455	0,933	128,214	376,851	212,421	84,103
RN2.5	37,106	29,365 ± 22,682	0,931	129,356	385,176	213,965	86,030
RN2.6	39,755	32,663 ± 22,663	0,932	131,427	389,502	217,263	86,408
RN2.7	31,854	23,547 ± 21,453	0,922	127,608	353,138	208,000	80,928
RN2.8	32,737	24,919 ± 21,232	0,924	128,218	354,875	209,519	80,774
RN2.9	34,208	26,851 ± 21,194	0,928	129,274	364,668	211,451	82,087
RN2.10	31,553	23,637 ± 20,901	0,929	125,415	362,650	207,927	82,431
RN2.11	32,657	24,931 ± 21,093	0,930	125,820	364,412	209,403	82,776
RN2.12	32,201	24,682 ± 20,681	0,931	124,864	363,183	209,071	82,487

Fonte: Autor (2022)

Observa-se com as análises das redes, que as arquiteturas utilizadas, para cada conjunto de dados de entrada, possuem resultados muito próximos. Assim, a utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) apresentou maior R² e menor erro MAE, embora o maior erro RMSE, face aos obtidos com a utilização de LL_c, LL_p, IP e Classificação SUCS. A utilização destes dados proporcionou uma leve diminuição dos erros MAE de 20,69 para 20,53, e um leve aumento do valor médio de R² de 0,92 para 0,93 em comparação com os obtidos com a utilização de LL_c, LL_p, IP e Classificação SUCS (RN1.1 a RN1.12).

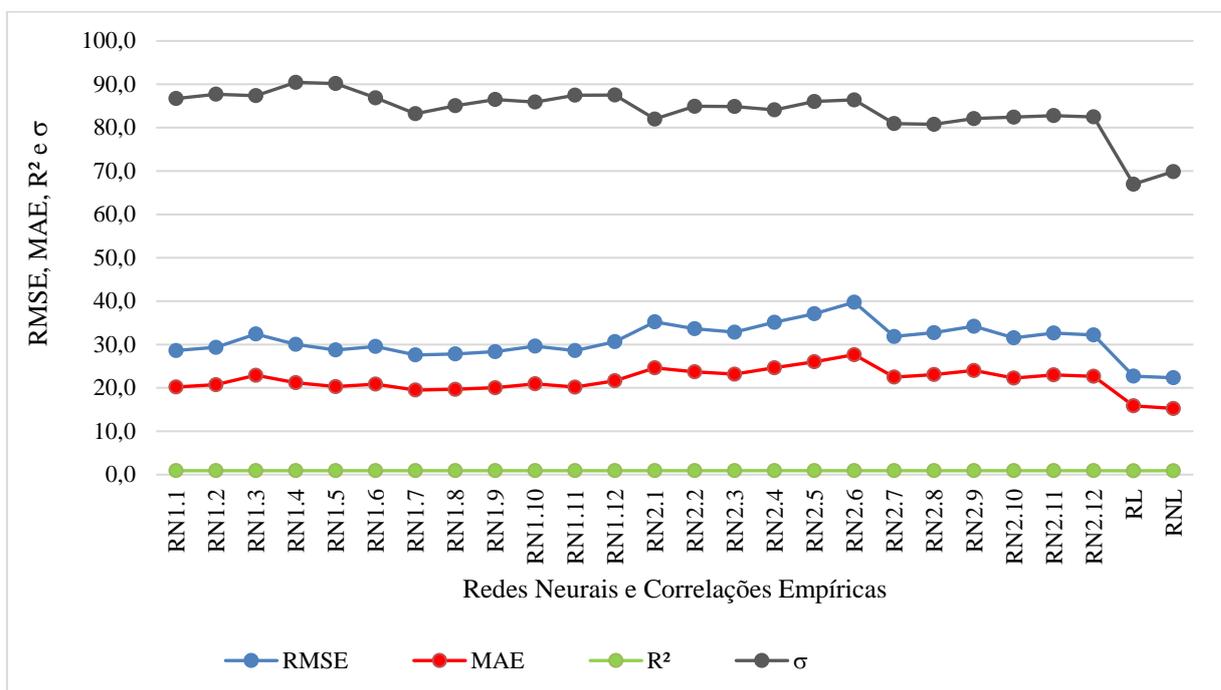
Analisando os erros MAE e RMSE e o σ das redes RN1.1 a RN1.12, as arquiteturas com 6 neurônios na primeira camada e 3 na segunda (RN2.8) e com 12 neurônios na primeira camada e 6 na segunda (RN2.10), foram as que apresentaram menores valores. Entretanto, um valor ligeiramente menor de R² é notado na arquitetura com 12 neurônios na primeira camada e 6 na segunda (RN2.10).

Em comparação com os resultados obtidos pelas análises de regressão, se compararmos apenas o critério de coeficiente de determinação, R², observa-se um resultado maior para as RNAs RN2.1 a 2.12 (>0,92) se comparado com as equações de regressão (R² = 0,89). No

entanto, quando analisado os erros RMSE e MAE e o desvio padrão (σ), observa-se que as RNAs apresentam resultados maiores do que as das correlações de regressão. O desvio padrão médio observado para as RNAs é de 85,20 enquanto o das correlações de regressão é de 68,42. Os erros médios RMSE e MAE são de 31,68 e 20,61 para as RNAs e de 22,53 e 15,57, respectivamente, para as correlações de regressão.

A Figura 55 apresenta graficamente o desempenho das redes neurais em comparação com as regressões.

Figura 55 - Resultados estatísticos das redes neurais e regressões do Subgrupo B-01iii.



Fonte: Autor (2022)

Da Figura 55 nota-se as diferenças destacadas entre os modelos de correlação linear e não linear e entre as diferentes arquiteturas de RNAs. Observa-se, portanto, um resultado melhor para as equações de regressão, diferentemente da tendência observada para o grupo B-01 de uma maneira geral.

4.1.9 Conclusões parciais para o grupo de dados B-01

Inicialmente apresentam-se os resultados das análises estatísticas e das redes neurais artificiais dos dados experimentais correspondentes ao grupo B-01 e seus subgrupos (B-01i, B-01ii e B-01iii) e suas discussões. O grupo B-01 e seus subgrupos correspondem aos resultados experimentais de LL_c obtidos pelo aparelho de Casagrande de base dura e LL_p obtido pelo cone britânico.

Dos resultados experimentais previamente compilados da literatura foram selecionados para as análises estatísticas e redes neurais artificiais os que apresentam informação complementar de IP. Os valores de IP também foram utilizados como parâmetro de entrada para algumas das análises das RNAs, conforme abordado no item 3.1.2. Para a análise da influência do intervalo de LL nos valores de LL_c e LL_p , os grupos foram separados em três subgrupos: B-01i ($LL < 50$), B-01ii ($50 < LL \leq 120$) e B-01iii ($LL > 120$).

Devido a variabilidade observada nas amostras do banco de dados do grupo B-01 e seus subgrupos que possuem resultados experimentais de diversos locais (Brasil, Egito, Equador, Estados Unidos, Itália, Portugal e Turquia), inicialmente foram identificados *outliers* para as variáveis LL_c , LL_p e IP. Os *outliers* foram definidos por meio dos gráficos de *outliers* e diagramas *boxplot*, conforme metodologia descrita no item 3.2. De uma maneira geral, a eliminação dos *outliers* resultou em uma redução no número de amostras analisadas nesta dissertação, além de reduzirem o intervalo inicial de LL para cada grupo ou subgrupo. A maioria dos *outliers* corresponde aos maiores valores de LL_c , LL_p ou IP das amostras de cada conjunto de dados. Como resultado, o grupo B-01 e o subgrupo B-01ii, foram limitados a valores de $LL_c < 100$, obtidos pelos aparelhos de Casagrande de base dura.

A descrição estatística mostrou que os resultados obtidos com a seleção dos dados e eliminação dos *outliers* proporcionaram uma redução de cerca de 70% nos valores de desvio padrão, 80% da variância e 50% do coeficiente de variação para as variáveis LL_c , LL_p e IP quando comparados com os dados brutos compilados da literatura. Da divisão do grupo B-01 em subgrupos de dados, notou-se uma redução superior a 48% no desvio padrão e 74% na variância para os subgrupos B-01i e B-01ii. Essa redução foi menor para o coeficiente de variação que diminuiu em cerca de 24% para o subgrupo B-01i e 60% para o subgrupo B-01ii. O subgrupo B-01iii apresentou comportamento distinto dos demais subgrupos, com um aumento de cerca de 84% nos valores de desvio padrão, 97% da variância e 15% do coeficiente de variação. Quando analisado o coeficiente de variação, apenas o subgrupo B-01ii apresentou valores inferiores a 25% para as variáveis LL_c , LL_p e IP. Isto indica que o tratamento dos dados possibilitou a formação de conjuntos de dados menos dispersos, mais

próximos da média. Assim, conclui-se que o subgrupo B-01ii possui dados mais homogêneos se comparados com os demais subgrupos de dados.

A variabilidade das amostras também foi observada nas diferentes classificações de solo, granulometria e tipos de argilominerais que caracterizam os resultados experimentais que compõem o banco de dados. Em todos os conjuntos de dados (grupo B-01, subgrupo B-01i, subgrupo B-01ii e subgrupo B-01iii) mais de 60% das amostras foram classificadas como argilas de alta e baixa compressibilidade (CL e CH). As demais amostras foram classificadas como siltes de baixa e alta compressibilidade (ML e MH). Da granulometria dos solos informada nas publicações de Diniz (1994), Deka et al. (2009), Ozer (2009), Di Matteo (2012), Quintela et al. (2014), Nini (2014), Couto et al. (2016), Bicalho et al. (2017), El-Shinawi (2017) e Niazi et al. (2019) foi observado que os dados contemplam amostras de solos com frações de silte e areia. O posicionamento dos valores de LL_c das amostras no ábaco de Casagrande apresentado no item 3.1.2.1 e conseqüentemente sua classificação SUCS, corroboram as informações da granulometria citada nas referidas publicações. No grupo B-01 foi notada uma variedade maior de classificações, contendo amostras de argilas e siltes de baixa e alta compressibilidade, ou seja, CL-ML, CL, ML, CH e MH. Os demais subgrupos tiveram as classificações atribuídas ao intervalo de LL em que estão inseridos. Assim, o subgrupo B-01i, $LL < 50$, contém amostras de solo de baixa compressibilidade, CL-ML, CL e ML. Os subgrupos B-01ii e B-01iii, $LL > 50$, solos de alta compressibilidade, CH e MH.

A mineralogia dos solos estudados não é informada em todas as publicações, no entanto, alguns autores relatam amostras compostas por caulinitas, ilitas, bentonitas e motmorilonitas, e em alguns casos a mistura desses argilominerais. A presença desses argilominerais também é variável conforme os intervalos de LL analisados. Assim, para o grupo B-01, subgrupo B-01i e B-01ii, com $LL < 120$, são relatados, em sua maioria, amostras compostas por caulinitas. Para o subgrupo B-01iii, $LL > 120$, são relatados em sua maioria misturas de bentonitas.

Dentre os resultados experimentais de LL_c e LL_p para o intervalo de $LL_c < 120$ (grupo B-01, subgrupo B-01i e subgrupo B-01iii), acima de 80% das amostras apresenta $LL_c < LL_p$. Esses resultados divergem para valores de $LL_c > 120$ (subgrupo B-01iii) com cerca de 97% das amostras apresentando $LL_c > LL_p$, com fortes divergências entre os resultados. Destaca-se, portanto, que as diferenças entre LL_c e LL_p não refletem a totalidade dos resultados experimentais. Para o intervalo de $LL_c < 120$, por exemplo, ao menos 20% das amostras apresentaram $LL_c > LL_p$, divergindo da maior parcela de dados analisados.

Após a eliminação dos *outliers*, os dados de cada grupo foram submetidos a análises de regressão para definição das correlações de regressão lineares e não lineares. As correlações

de regressão foram comparadas através dos resultados obtidos para a raiz quadrada do erro médio (RMSE), erro médio absoluto (MAE), coeficiente de determinação (R^2), valores mínimos e máximos, média e desvio padrão (σ) dos valores previstos de LL_p , conforme citado no item 3.4. A comparação entre as correlações lineares e não lineares dos conjuntos de dados analisados indicam que os modelos propostos resultam em valores muito próximos para a previsão de LL_p . As correlações não lineares apresentaram valores de desvio padrão ligeiramente menores que as correlações lineares. E, os erros RMSE e MAE foram praticamente coincidentes para os subgrupos B-01i e B-01ii. Portanto, conclui-se pela semelhança entre as correlações de regressão na explicação dos modelos propostos.

Os resultados experimentais de LL_c e LL_p que compõem os conjuntos de dados, juntamente com as correlações de regressão obtidas, foram plotados no gráfico de $LL_c \times LL_p$ para a análise gráfica dos resultados nos itens 4.1.1, 4.1.3, 4.1.5 e 4.1.7. As correlações de regressão obtidas para o intervalo de $LL_c < 120$ (grupo B-01 e subgrupos B-01i e B-01ii) se mostraram acima da linha de igualdade $LL_p = LL_c$. Isso indica que para este intervalo, aplicando as correlações propostas, os valores previstos de LL_p são superiores aos resultados experimentais de LL_c . O subgrupo B-01iii, no entanto, apresentou resultados diferentes dos observados para os demais. Portanto, para o intervalo de $LL_c > 120$, notou-se que as correlações se mostraram abaixo da linha de igualdade $LL_p = LL_c$, já que o conjunto de dados amostrais se concentrou entre o intervalo de $LL_p = LL_c$ e $LL_p = 0,5 LL_c$. Portanto, para $LL_c > 120$, aplicando as correlações propostas, os valores previstos de LL_p são menores que os resultados experimentais de LL_c . Destaca-se, conforme supracitado, que o subgrupo B-01iii, $LL_c > 120$, foi o que apresentou dados com maior variabilidade, ou seja, trata-se um grupo de dados heterogêneos. As diferenças entre os valores de LL_c e LL_p observadas para este subgrupo são maiores do que as observadas para os demais conjuntos de dados. Como resultado, as correlações obtidas apresentam diferenças de resultados maiores, inclusive com valores de desvio padrão superiores a 60%, conforme apresentado na Tabela 40 .

As correlações lineares propostas foram submetidas às análises de resíduos e aos testes de hipóteses para verificação da normalidade do modelo e independência das variáveis, conforme metodologia citada no item 3.2. Para as correlações lineares propostas, em todos os conjuntos de dados analisados foram observados a heterocedasticidade e que a distribuição dos resíduos para os modelos não é normal. Apenas os resíduos verificados para o subgrupo B-01ii foram classificados como independentes, os demais foram todos ditos dependentes. Esses resultados refletem a variabilidade dos dados analisados, conforme discutido anteriormente. Portanto, os testes estatísticos mostraram que as correlações

lineares propostas para os conjuntos de dados do grupo B-01 e subgrupos B-01i, B-01ii e B-01iii não são significativas, apesar dos elevados valores do coeficiente de determinação R^2 .

A eliminação dos *outliers*, conforme citado, levou em consideração os resultados obtidos dos gráficos de *outliers* e diagramas *boxplot*. Entretanto, as barras isoladas nas extremidades dos histogramas (Figura 42, Figura 46, Figura 50 e Figura 54) e sua assimetria, e os pontos de dados identificados como resíduos (Figura 41, Figura 45, Figura 49 e Figura 53) indicam a presença de dados atípicos nos conjuntos de dados analisados. A presença desses dados que se diferenciam muito de todos os outros, ou pontos fora da curva, justificam a heterocedasticidade e a distribuição não normal dos resíduos, resultados determinantes para a conclusão de que as correlações lineares não são significativas. Assim, verifica-se que, para o tratamento e eliminação dos *outliers*, devem ser examinados além dos gráficos de *outliers* e diagramas *boxplot* os gráficos de dispersão dos resíduos e histogramas.

Os conjuntos de dados também foram analisados pelas redes neurais artificiais para previsão de LL_p , conforme metodologia citada no item 3.3. Diferentes arquiteturas e dados de entrada foram testados para a avaliação da capacidade de previsão das redes e sua consequente validação. Foram considerados arquiteturas com uma e duas camadas ocultas e número de neurônios variáveis. Foram comparadas as arquiteturas considerando os dados de entrada LL_c e LL_p e os dados de entrada LL_c , LL_p , IP e classificação SUCS. As combinações de arquiteturas e dados de entrada propostos para as RNAs também foram comparadas através dos resultados obtidos de RMSE, MAE, R^2 , valores mínimos e máximos, média σ das estimativas de LL_p .

Foram observadas pequenas diferenças nos resultados obtidos para as arquiteturas testadas e dados de entrada de dados. Quanto ao número de neurônios e camadas ocultas, as arquiteturas propostas para os conjuntos de dados do grupo B-01 e seus subgrupos apresentam resultados semelhantes. Os resultados, portanto, diferem quanto aos dados de entrada. Para o subgrupo B-01iii as diferenças observadas foram de 4,8 para o erro RMSE, 3,3 para o erro MAE e 3,8 para o desvio padrão quando comparados a utilização de LL_c e LL_p e LL_c , LL_p , IP e classificação SUCS como dados de entrada. Para os demais subgrupos as maiores diferenças observadas foram menores que 0,5 para o erro RMSE, 0,3 para o erro MAE e 0,7 para o desvio padrão. Nota-se que o subgrupo B-01iii foi o que apresentou maiores diferenças entre os resultados, entretanto, essas diferenças são inferiores a 15% para os erros RMSE, MAE e σ . Conclui-se, portanto, pela semelhança entre os modelos de RNAs apresentados. De uma maneira geral, conclui-se pela semelhança entre as RNAs testadas.

Os resultados obtidos com as correlações de regressão foram confrontados com os resultados obtidos pelas redes neurais artificiais, a fim de indicar o método que melhor se adequa a cada

um dos modelos propostos. Foram comparados os resultados obtidos de RMSE, MAE, R^2 , valores mínimos e máximos, média e σ das estimativas de LL_p das correlações de regressão e das RNAs. Observou-se resultados semelhantes entre as correlações lineares e não lineares e as RNAs propostas para os conjuntos de dados analisados no grupo B-01 e subgrupos B-01i, B-01ii e B-01iii. Os resultados obtidos com a comparação dos erros RMSE e MAE e R^2 apresentaram diferenças pequenas para o grupo B-01 e para os subgrupos B-01i. Para esses conjuntos de dados, as diferenças foram menores que 15% para os erros RMSE e MAE. Para os subgrupos B-01ii e B-01iii, $LL_c > 50$, essas diferenças foram superiores a 28% para o erro RMSE e 30% para o erro MAE. Quando analisado o σ foram observadas diferenças superiores a 20% para todos os conjuntos de dados. Tendo como premissa os menores valores de σ , as correlações de regressão indicaram uma menor variabilidade para a previsão de LL_p para os subgrupos B-01ii e B-01iii. Entretanto, ressalta-se que apesar destes resultados, as correlações lineares não foram significativas.

É importante destacar também, que para a maioria dos casos citados, os coeficientes de determinação R^2 das correlações de regressão e RNAs foram muito semelhantes. Uma exceção foram os valores obtidos com as RNAs do subgrupo B-01i ($LL_c < 50$). Assim, os demais conjuntos de dados apresentaram valores próximos ou superiores a 0,9. Ou seja, se levado em consideração apenas o R^2 , conclui-se que as correlações de regressão e RNAs são explicativas dos conjuntos de dados analisados. Entretanto, um R^2 próximo de 1 não é suficiente para a tomada de decisão quanto a aceitação de um determinado modelo, conforme destacado por Crevelin (2019). Nesta pesquisa observou-se que embora os altos valores de R^2 não foi possível a validação dos modelos de correlações lineares propostos. Portanto, nota-se por meio dos resultados obtidos que, embora o R^2 forneça uma estimativa da força de correlação entre o modelo e a variável de resposta, não indica se um modelo de regressão é adequado, já que é possível ter um valor baixo de R^2 para um bom modelo ou um valor alto de R^2 para um modelo que não se encaixa nos dados.

4.2 RESULTADOS PARA OS DADOS DO GRUPO B-02

4.2.1 Avaliação estatística dos dados do subgrupo B-02

Os dados do grupo B-02 foram submetidos a testes estatísticos para a identificação e tratamento dos *outliers* conforme a metodologia citada no item 3.2. Assim, foram identificados 8 *outliers* do total de amostras de solo selecionadas, resultando o grupo, portanto, em 72 amostras aprovadas para serem utilizadas no estabelecimento das correlações empíricas e no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c , LL_p e IP são apresentados no Anexo I.

Para o grupo B-02, nota-se que os maiores valores de LL das publicações de Olszewska (2003) e o maior valor de LL publicado por Ozer (2009) foram apontados pelas análises estatísticas como *outliers*. A Tabela 45 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 45 - Descrição estatística dos dados investigados no grupo B-02.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	72	8	55,74	17,66	311,95	31,68	29,80	39,66	49,45	67,88	99,90
LL _p	72	8	54,52	14,06	197,56	25,78	31,50	45,15	50,05	63,97	89,80
IP	72	8	25,58	10,64	113,14	41,58	10,70	16,41	25,55	33,78	55,40

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

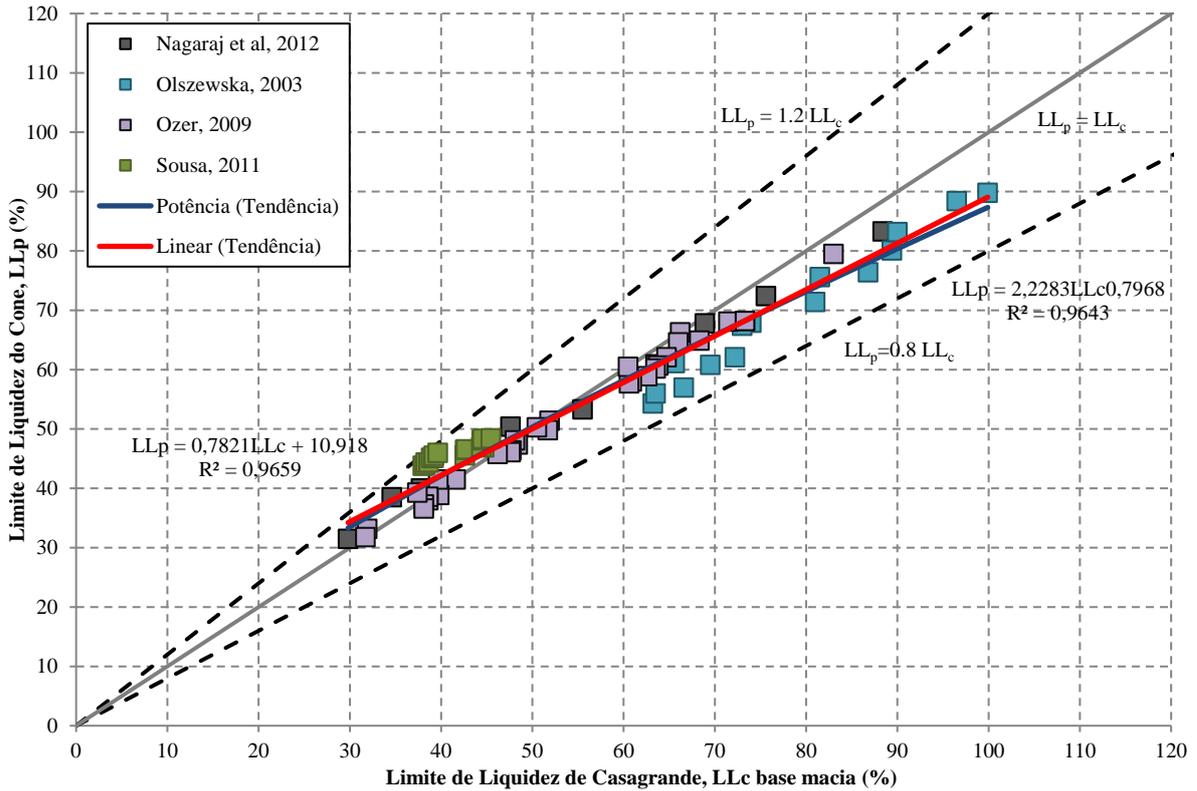
Fonte: Autor (2022)

Os resultados apresentados na Tabela 45 foram comparados com a descrição estatística das 90 amostras designadas inicialmente para o grupo B-02 (Tabela 15). Foi observado, após a seleção e tratamento dos dados, uma redução de cerca de 70% nos valores de desvio padrão, 92% da variância e 54% do coeficiente de variação para as variáveis LL_c, LL_p e IP. Esses resultados indicam que o tratamento dos dados possibilitou a formação de um grupo com dados menos dispersos, mais próximos da média. Entretanto, os coeficientes de variação para as variáveis LL_c e IP ainda se mostraram superiores a 30%. Portanto, embora os resultados positivos obtidos do tratamento dos dados, o grupo B-02 ainda é composto de dados heterogêneos.

A variabilidade dos dados do grupo B-02 é descrita no item 3.1.2.2. Na Tabela 16, o resumo das classificações SUCS das amostras indica que, embora os dados concentrem-se em sua maioria entre argilas de alta e baixa compressibilidade (CL e CH), cerca de 40% das amostras são classificadas como siltes de baixa e alta compressibilidade (ML e MH). O posicionamento das amostras no ábaco de Casagrande e conseqüentemente sua classificação SUCS, corroboram as observações feitas quanto a granulometria, já que o grupo contempla amostras de solos com frações significativas de silte e areia. A maioria das amostras publicadas por Ozer (2009), Sousa (2011) e Nagaraj et al. (2012) apresentaram frações de argila inferiores a 30%. O grupo B-02 é composto ainda por amostras de solo de diferentes mineralogias. Foram observados dentre este grupo a presença de ilitas (Sousa, 2011), caulinitas e montmorilonitas (Nagaraj et al., 2012), além das misturas de caulinita e bentonita (Olszewska, 2003), e de ilitas e caulinitas (Sousa, 2011). Apenas os dados publicados por Ozer (2009) não continham a informação do tipo de argilomineral.

Conforme citado, os dados de valores mais altos de LL foram de uma maneira geral identificados como *outliers*, assim, o grupo B-02 foi limitado ao intervalo de $LL_c < 100$. A Figura 56 apresenta os pares de dados e as correlações obtidas.

Figura 56 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do grupo B-02 e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-02:

$$LL_p = 0,7821LL_c + 10,918 \quad R^2 = 0,9659 \quad (44)$$

$$LL_p = 2,2283LL_c^{0,7958} \quad R^2 = 0,9643 \quad (45)$$

Nota-se da Figura 56, que o grupo de dados amostrados está concentrado entre as linhas $LL_p = 1,2 LL_c$ e $LL_p = 0,8 LL_c$ e cerca de 60% das amostras apresenta resultados experimentais de LL_p menores do que LL_c . As correlações obtidas se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 1,2 LL_c$, para valores de LL_c menores que 50 e entre as linhas $LL_p=LL_c$ e $LL_p = 0,8 LL_c$ para valores de LL_c superiores a 50. Observa-se que as correlações descritas para o grupo B-02 preveem resultados muito próximos, principalmente para valores de LL_p entre 45 e 67.

A Tabela 46 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-02.

Tabela 46 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02.

Regressão	RMSE	MAE	R²	Mínimo	Máximo	Média	σ
Linear - RL	2,579	2,134	0,966	34,225	89,050	54,516	13,813
Não linear - RNL	2,661	2,204	0,964	33,201	86,942	54,227	13,623

Fonte: Autor (2022)

Observa-se também por meio do resumo estatístico, que o modelo linear possui menor erro médio absoluto (MAE), raiz quadrada do erro médio (RMSE) e coeficiente de determinação (R^2), enquanto o modelo não linear apresenta apenas um menor desvio padrão.

As correlações obtidas para o Grupo B-02 foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados destas comparações constam das Tabela 47 e Tabela 48.

Tabela 47 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o Grupo B-02.

LL_c	LL_p		
	Grupo B-02	Crevelin e Bicalho (2019)	Crevelin e Bicalho (2019)
	LL_c<100	30<LL_c<70	50<LL_c<400
30	34,38	34,10	36,69
40	42,20	43,17	44,08
50	50,02	52,24	51,47
60	57,84	61,31	58,86
70	65,67	70,38	66,25
80	73,49	-	73,64
90	81,31	-	81,03
100	89,13	-	88,42

Fonte: Autor (2022)

A correlação linear obtida para o grupo B-02 apresenta resultados mais próximos aos obtidos por Crevelin e Bicalho (2019) para amostras de caulinitas e ilitas com valores de LL entre 30 e 70. A semelhança observada entre as correlações pode ser explicada pela quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa. Assim, em relação a correlação Crevelin e Bicalho (2019), foram utilizados em comum os dados de Olszewska (2003), Özer (2009), Sousa (2011) e Nagaraj et al. (2012).

Tabela 48 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o Grupo B-02.

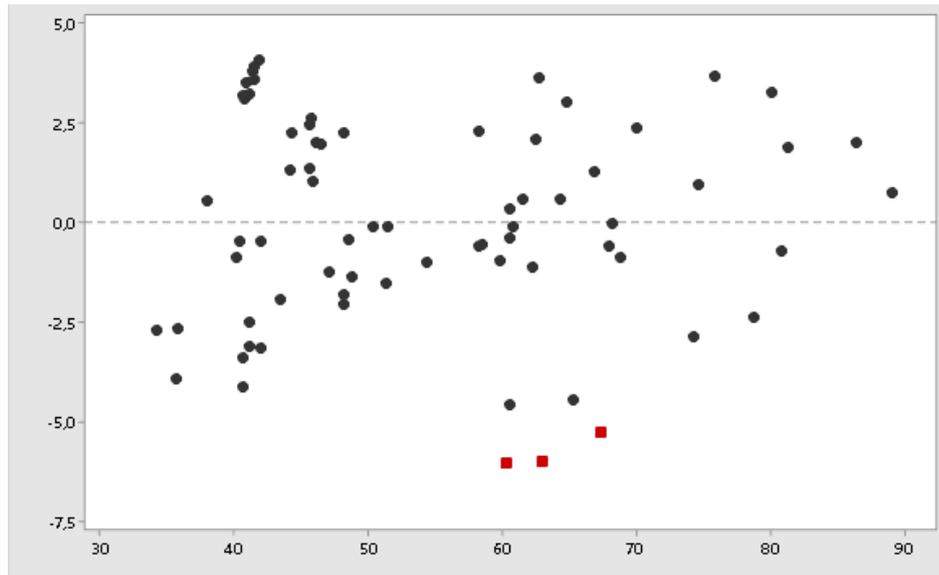
LL_c	LL_p		
	Grupo B-02	O' Kelly et al. (2018)	O' Kelly et al. (2018)
	$LL_c < 100$	$LL_c < 120$	$LL_c < 600$
30	33,38	32,31	32,38
40	41,97	41,62	41,23
50	50,12	50,65	49,73
60	57,95	59,47	57,96
70	65,51	68,11	65,98
80	72,85	76,60	73,81
90	80,01	84,97	81,48
100	87,01	93,22	89,03

Fonte: Autor (2022)

A correlação não linear obtida para o grupo B-02 mostra-se muito próxima da publicada por O'Kelly et al. (2018) para LL_c de base macia < 120%, principalmente para $LL_c < 70$. Dentre os dados analisados por O'Kelly et al. (2018) apenas os dados publicados por Ozer (2009) são comuns aos dados utilizados para a obtenção da correlação não linear do subgrupo B-02. Destaca-se, portanto, o intervalo de LL adotado como um fator de influência para a proximidade entre os resultados obtidos na presente pesquisa e os resultados de O'Kelly et al. (2018). Isso ocorre já que, embora o Grupo B-02 contemple os valores de LL_c e LL_p obtidos pelos métodos de Casagrande de base macia e cone britânico das publicações em geral, a eliminação dos *outliers* limitou o grupo a valores de $LL_c < 100$.

A correlação linear obtida para o grupo B-02 foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Os testes, conforme metodologia descrita no item 3.2, foram utilizados como critério para a validação do modelo. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos (Figura 57)

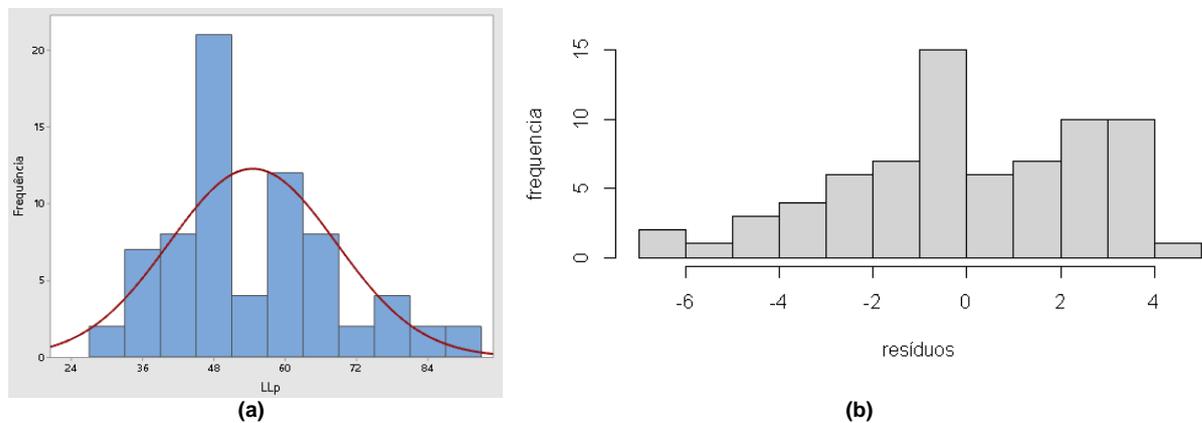
Figura 57 – Relatório de diagnóstico de resíduos – Grupo B-02.



Fonte: Autor (2022)

No gráfico da Figura 57 é possível identificar que 3 pontos de dados possuem grandes resíduos, representados pelos pontos vermelhos da figura, não estando bem ajustados pelas equações propostas. Desse total, 2 possuem pontos atípicos de LL_c , o que pode fazer com que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 58.

Figura 58 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Grupo B-02.

Fonte: Autor (2022)

Através da Figura 58 observa-se que a variável dependente apresenta um gráfico de distribuição com uma assimetria moderada à direita. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 49 apresenta o resumo dos resultados dos testes estatísticos realizados para o grupo B-02.

Tabela 49 – Resumo dos resultados dos testes estatísticos para o Grupo B-02.

Teste	valor-p	Conclusão
Teste t (intercepto)	2,94E-16	Parâmetro significativo
Teste t (x)	2,00E-16	Parâmetro significativo
Teste F	2,20E-16	Parâmetro significativo
Teste KS	2,09E-06	Distribuição dos resíduos não é normal
Teste DW	3,08E-14	Resíduos não são independentes

Fonte: Autor (2022)

Os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados, uma vez que o valor-p obtido para os testes é maior do que a significância de 5%.

O resultado do teste de aderência de *Kolmogorov Smirnov* (teste KS) aponta para a distribuição não normal dos resíduos, uma vez que o valor-p é inferior a significância de 5%. Esse resultado vai de encontro à assimetria observada na análise dos resíduos por meio dos histogramas.

O resultado do teste de *Durbin Watson* (DW), que avalia a independência entre os resíduos, indica que os resíduos do grupo B-012 não são independentes, já que o valor-p é menor que o nível de significância de 5%. Além disso, o resultado do teste DW é de 0,54736, ou seja, por estar próximo de 0, indica que há autocorrelação positiva entre as variáveis independentes. Portanto, os resultados do teste DW não foram satisfatórios, apontando para a inconsistência do modelo.

Ressalta-se que, neste caso, em função do tamanho da amostra, o teste *Shapiro-Wilk* (SW) não foi utilizado.

Portanto, apesar de um R^2 relativamente alto ($> 0,96$) para as regressões lineares e não lineares, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios dos testes estatísticos dos resíduos indicaram que a correlação linear apresentada não é significativa.

4.2.2 Aplicação das RNAs nas amostras do grupo B-02

Utilizando os dados filtrados nas análises estatísticas, as 72 amostras aprovadas no grupo B-02 foram submetidas as redes neurais artificiais. Um resumo estatístico das RNAs treinadas é apresentado na Tabela 50.

Tabela 50 – Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02.

Rede	RMSE	MAE		R^2	Mínimo	Máximo	Média	σ
RN1.1	2,357	2,094	\pm 1,082	0,981	34,229	81,763	48,468	11,269
RN1.2	3,710	3,399	\pm 1,488	0,972	33,835	81,477	47,062	11,508
RN1.3	3,895	3,565	\pm 1,569	0,973	33,131	81,079	46,832	11,571
RN1.4	3,024	2,658	\pm 1,440	0,971	34,323	81,668	47,902	11,439
RN1.5	3,027	2,653	\pm 1,457	0,969	34,488	81,355	47,958	11,207
RN1.6	2,897	2,510	\pm 1,446	0,972	34,326	82,364	48,071	11,602
RN1.7	5,706	5,483	\pm 1,580	0,963	33,809	78,273	44,975	11,162
RN1.8	3,543	3,195	\pm 1,531	0,967	34,811	81,695	47,354	11,547
RN1.9	3,936	3,593	\pm 1,606	0,967	34,090	81,044	46,890	11,548
RN1.10	3,317	2,933	\pm 1,551	0,968	34,490	81,848	47,614	11,589
RN1.11	3,576	3,213	\pm 1,568	0,971	33,769	81,255	47,241	11,547
RN1.12	3,843	3,460	\pm 1,672	0,969	33,446	80,925	46,965	11,546
RN2.1	2,617	2,343	\pm 1,167	0,974	35,607	81,102	48,383	11,142
RN2.2	3,433	3,171	\pm 1,315	0,974	34,391	80,641	47,340	11,308
RN2.3	3,295	3,006	\pm 1,351	0,977	33,675	80,588	47,440	11,409
RN2.4	4,480	4,222	\pm 1,499	0,976	32,478	79,123	46,115	11,339
RN2.5	3,870	3,622	\pm 1,364	0,976	33,289	79,547	46,789	11,243
RN2.6	2,517	2,242	\pm 1,144	0,978	34,206	80,631	48,373	11,280
RN2.7	2,524	2,250	\pm 1,143	0,970	36,620	81,127	48,673	11,100
RN2.8	4,640	4,416	\pm 1,425	0,970	33,695	79,408	46,032	11,318
RN2.9	2,491	2,213	\pm 1,145	0,973	35,766	80,795	48,587	11,156
RN2.10	3,418	3,136	\pm 1,358	0,974	34,198	80,321	47,357	11,376
RN2.11	3,702	3,463	\pm 1,308	0,972	34,513	79,949	47,059	11,192
RN2.12	3,340	3,052	\pm 1,355	0,975	34,037	80,283	47,426	11,364

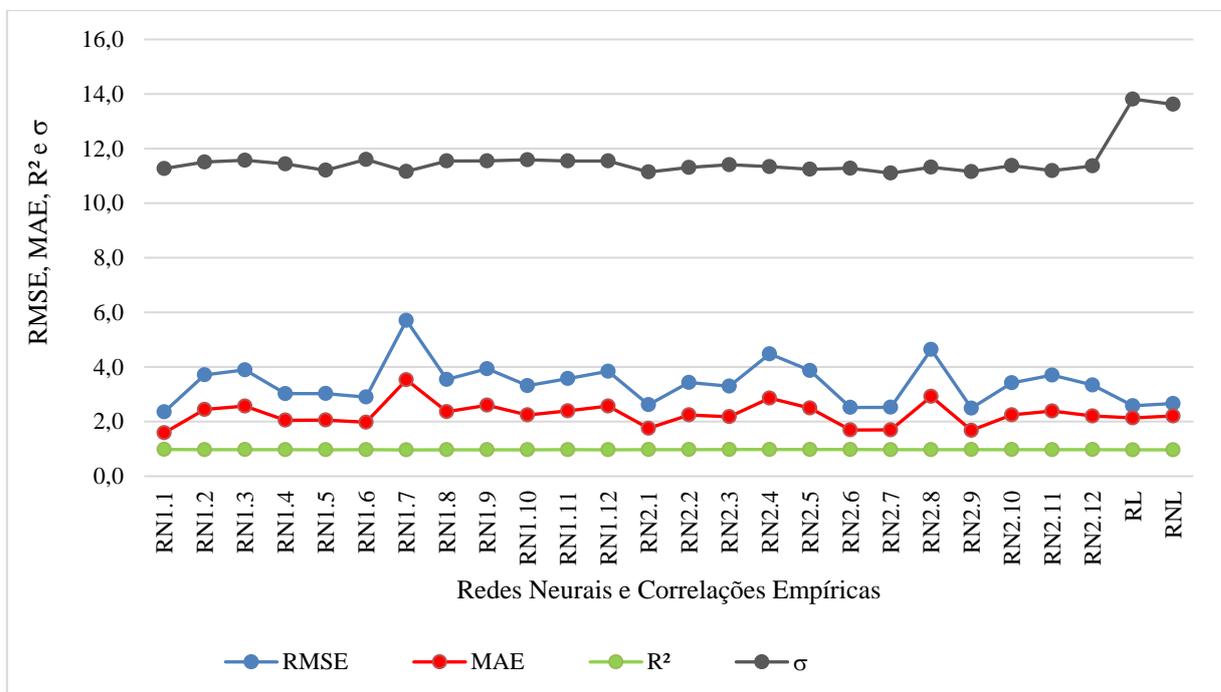
Fonte: Autor (2022)

Observa-se com as análises das RNAs que as arquiteturas utilizadas para cada conjunto de dados de entrada possuem resultados muito próximos. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) proporcionou uma redução da média dos erros MAE de 2,36 para 2,20, RMSE de 3,57 para 3,36, com um mesmo valor médio de R^2 em comparação com os obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Entretanto, como essas diferenças são pequenas, analisando os erros MAE, RMSE e coeficiente de determinação (R^2), a arquitetura com 3 neurônios em uma única camada

(RN1.1) foi a que apresentou menores resultados quando comparados com as demais arquiteturas.

Em comparação com os resultados obtidos pelas análises de regressão, observa-se um mesmo valor médio de R^2 que o obtido com as correlações de regressão. No entanto, quando analisado o desvio padrão (σ), observa-se que as RNAs apresentam resultados menores do que as das correlações de regressão. O desvio padrão médio observado para as RNAs é de 11,37 enquanto o das correlações de regressão é de 13,72. A Figura 59 apresenta graficamente a comparação das RNAs e correlações de regressão.

Figura 59 – Resultados estatísticos das redes neurais e regressões do Grupo B-02.



Fonte: Autor (2022)

Da Figura 59 nota-se a semelhança entre os modelos de correlação e RNAs, com destaque para as pequenas diferenças entre os resultados. Uma exceção, no entanto, são os maiores valores de σ das correlações de regressão. Portanto, os resultados indicam que as RNAs se adequam melhor a previsão de valores de LL_p para o grupo de dados B-02.

4.2.3 Avaliação estatística dos dados do subgrupo B-02i

Para os dados do subgrupo B-02i, aplicada a metodologia citada no item 3.2, foram identificados 3 *outliers*. Portanto, resultaram neste subgrupo 33 amostras aprovadas para serem utilizadas para o estabelecimento das correlações empíricas, bem como no

treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c, LL_p e IP são apresentados no Anexo I.

Foram identificados como *outliers* o maior valor de LL para os dados publicados por Nagaraj et al. (2012) e os maiores valores de IP dentre os resultados de ensaio publicados por Ozer (2009). A Tabela 51 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 51 – Descrição estatística dos dados investigados no subgrupo B-01ii.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	33	3	40,193	4,430	19,621	11,02	29,800	38,100	39,200	44,435	48,100
LL _p	33	3	42,752	4,944	24,445	11,56	31,500	38,750	44,370	46,545	48,440
IP	33	3	15,933	2,518	6,339	15,80	10,700	14,555	16,370	17,800	20,600

Legenda: N = número de amostras; N* = número de outliers; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

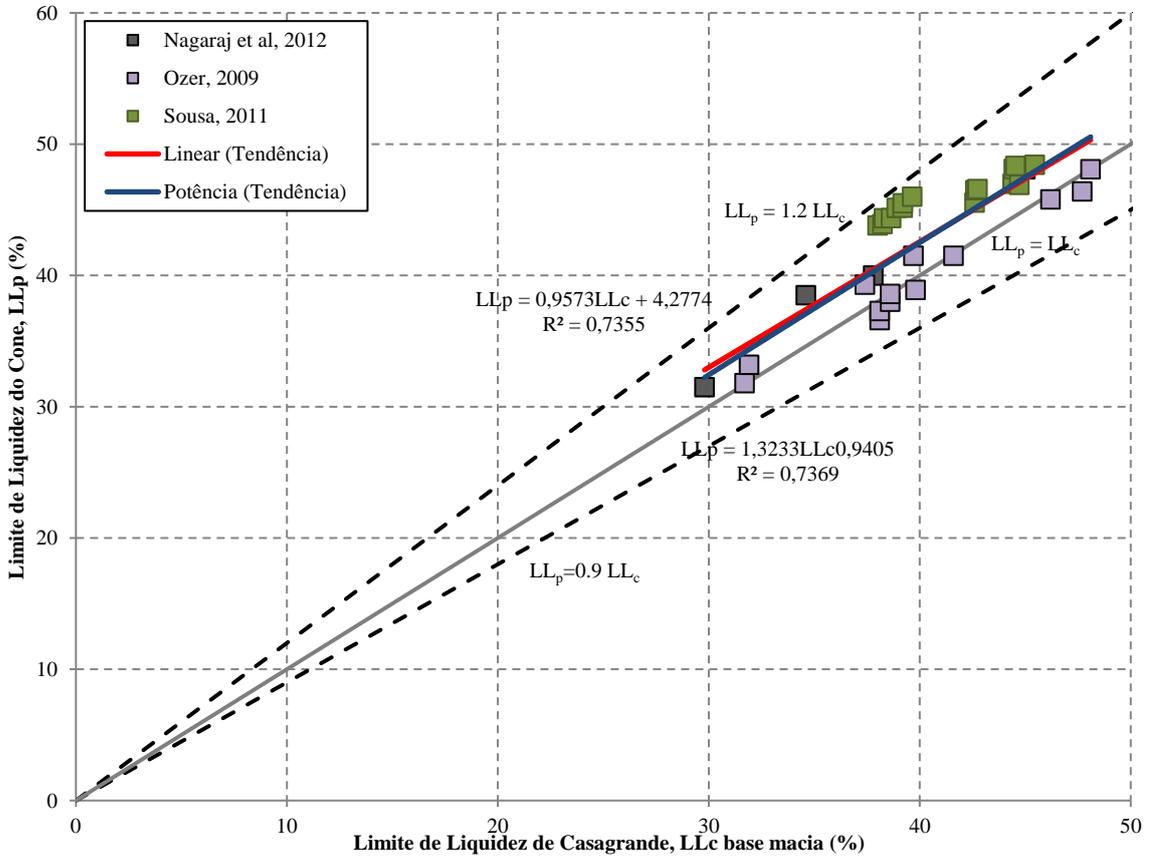
Fonte: Autor (2022)

Os resultados apresentados na Tabela 51, quando comparados com a descrição estatística do grupo B-02, indicam que este grupo apresenta dados menos dispersos. Foi observado para o subgrupo B-02i, uma redução de cerca de 72% nos valores de desvio padrão, 92% da variância e 61% do coeficiente de variação para as variáveis LL_c, LL_p e IP quando comparado com o grupo B-02. O coeficiente de variação observado para as variáveis LL_c, LL_p e IP ficou abaixo de 15%, indicando que este subgrupo é composto por dados mais homogêneos se comparado ao grupo B-02.

Os dados do subgrupo B-02i, em função do intervalo de LL em que se estão inseridos, são classificados como argilas e siltes de baixa compressibilidade (CL e ML). Embora mais de 60% dos dados sejam descritos como amostras de argila (CL), a granulometria das amostras selecionadas para este subgrupo, é caracterizada por parcelas de siltes e areias, justificando, portanto, a variação observada quanto a classificação das amostras. A mineralogia dos solos descritas para as amostras obtidas de Nagaraj et al. (2012) e Sousa (2011) indicam a presença de argilas caulinitas e ilitas. No entanto, a mineralogia não é informada para os resultados experimentais de Ozer (2009).

A Figura 60 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 60 – Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02i e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-02i:

$$LL_p = 0,9573LL_c + 4,2774 \qquad R^2 = 0,7355 \qquad (46)$$

$$LL_p = 1,3233LL_c^{0,9405} \qquad R^2 = 0,7369 \qquad (47)$$

Nota-se da Figura 60 que as amostra de solo analisadas se encontram entre as linhas $LL_p = 0,9 LL_c$ e $LL_p = 1,2 LL_c$ e cerca de 73% dos resultados de LL_p observados são maiores que LL_c . As correlações de regressão obtidas se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 1,2 LL_c$, seguindo, portanto, a tendência observada para o grupo B-02, para valores de LL_c menores que 50. Observa-se que estas correlações apresentam resultados muito próximos, principalmente para valores de LL_p superiores a 40.

A Tabela 52 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-02i.

Tabela 52 – Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02i.

Regressão	RMSE	MAE	R²	Mínimo	Máximo	Média	σ
Linear – RL	2,504	2,178	0,736	32,805	50,324	42,754	4,240
Não linear – RNL	2,505	2,139	0,737	32,222	50,549	42,680	4,431

Fonte: Autor (2022)

O resumo estatístico confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Entretanto, observa-se uma variação de 4,2 para o σ da correlação linear para 4,4 para a correlação não linear. Ou seja, as correlações não lineares possuem valores pouco menores de σ . Os resultados obtidos para os erros RMSE, MAE e R^2 são praticamente coincidentes.

As correlações obtidas para o subgrupo B-02i foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados destas comparações constam das Tabela 53 e Tabela 54.

Tabela 53 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-02i.

LL_c	LL_p		
	Subgrupo B-02i	Crevelin e Bicalho (2019)	Crevelin e Bicalho (2019)
	LL_c<50	30<LL_c<70	50<LL_c<400
30	33,00	34,10	36,69
40	42,57	43,17	44,08
50	52,14	52,24	51,47

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-02i apresenta resultados próximos aos obtidos por Crevelin e Bicalho (2019). Entretanto, os resultados podem ser ditos mais próximos da correlação obtida para amostras de caulinitas e ilitas com valores de LL entre 30 e 70, isso já que o intervalo se mostra mais próximo ao utilizado para a obtenção da correlação linear do subgrupo B-02i ($LL_c < 50$). A semelhança observada entre as correlações pode ser explicada ainda pela quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa. Assim, em relação a correlação Crevelin e Bicalho (2019), foram utilizados em comum os dados de Olszewska (2003), Özer (2009), Sousa (2011) e Nagaraj et al. (2012).

Tabela 54 – Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-02i.

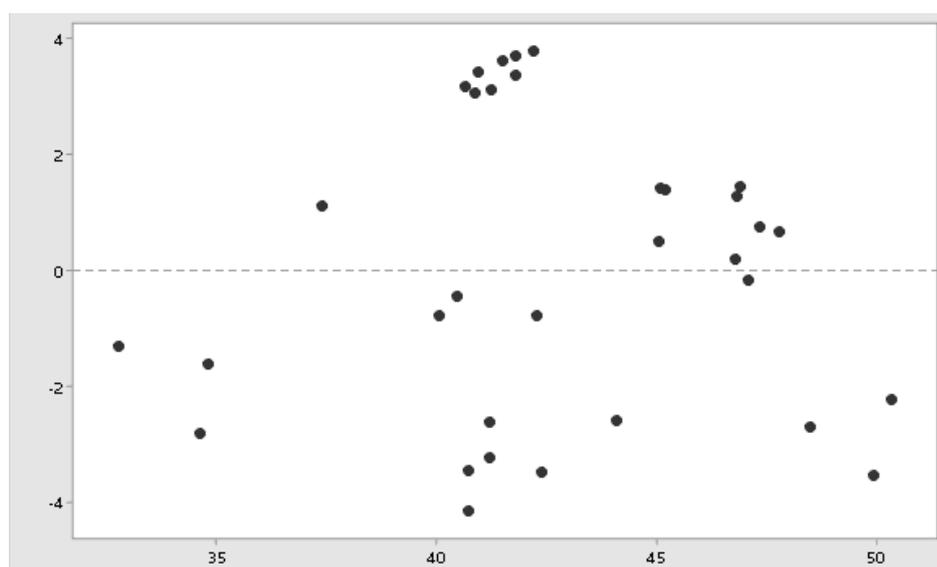
LL_c	LL_p		
	Subgrupo B-02i	O' Kelly et al. (2018)	O' Kelly et al. (2018)
	$LL_c < 50$	$LL_c < 120$	$LL_c < 600$
30	32,43	32,31	32,38
40	42,50	41,62	41,23
50	52,43	50,65	49,73

Fonte: Autor (2022)

A correlação não linear obtida para o subgrupo B-02i mostra-se muito próxima das correlações publicadas por O'Kelly et al. (2018) para LL_c de base macia < 120% e LL_c de base macia < 600%. Dentre os dados analisados por O'Kelly et al. (2018) apenas os dados publicados por Ozer (2009) são comuns aos dados utilizados para a obtenção da correlação não linear para o subgrupo B-02i. Destaca-se, portanto, o intervalo de LL adotado como um fator de influência para a proximidade entre os resultados obtidos na presente pesquisa e os resultados de O'Kelly et al. (2018), já que embora os bons resultados obtidos da comparação das duas correlações, a correlação de O'Kelly et al. (2018) para LL_c de base macia < 120% apresentou maior proximidade com a correlação do subgrupo B-02i.

A correlação linear obtida para o subgrupo B-02i foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos (Figura 61).

Figura 61 - Relatório de diagnóstico de resíduos – Subgrupo B-02i.

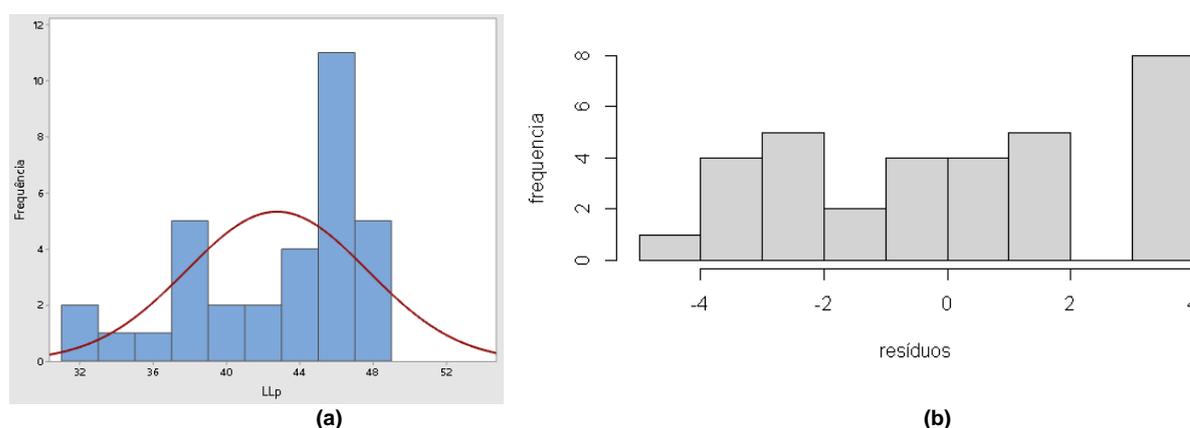


Fonte: Autor (2022)

No gráfico da Figura 61 não foram identificados pontos de dados com grandes resíduos. O comportamento dos dados no gráfico de resíduos, que se apresentam distribuídos em uma faixa horizontal de alta amplitude, apontam para uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 62.

Figura 62 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-02i.



Fonte: Autor (2022)

Através da Figura 62 observa-se que a variável dependente apresenta um gráfico de distribuição com uma assimetria moderada à esquerda. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 55 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-02i.

Tabela 55 – Resumo dos resultados dos testes estatísticos para o subgrupo B-02i.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,313	Parâmetro não significativo
Teste t (x)	1,83E-10	Parâmetro significativo
Teste F	1,83E-10	Parâmetro significativo
Teste KS	5,98E-03	Distribuição dos resíduos não é normal
Teste DW	1,16E-08	Resíduos não são independentes

Fonte: Autor (2022)

Analisando os resultados da Tabela 55, verifica-se que os resultados obtidos para o subgrupo B-02i são similares aos obtidos para o Grupo B-02.

Assim, os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados, embora o Teste

t (intercepto) aponte para a insignificância do parâmetro. O resultado do teste de aderência de Kolmogorov Smirnov (teste KS) aponta para a distribuição não normal dos resíduos, em consonância com a assimetria observada na análise dos resíduos por meio dos histogramas. O resultado do teste de Durbin Watson (DW), indica que os resíduos do subgrupo B-01i não são independentes, e, o resultado do teste DW é de 0,47286, ou seja, por se aproximar de 0, indica que há autocorrelação positiva entre as variáveis independentes. Portanto, os resultados do teste DW para o subgrupo B-02i também não foram satisfatórios, apontando para a inconsistência do modelo.

Portanto, apesar de um R^2 significativo ($> 0,73$) para as regressões lineares e não lineares, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios dos testes estatísticos dos resíduos indicaram que a correlação linear apresentada não é significativa.

4.2.4 Aplicação das RNAs nas amostras do grupo subgrupo B-02i

Utilizando os dados filtrados nas análises estatísticas, as 33 amostras aprovadas no subgrupo B-02i foram submetidas as redes neurais artificiais. Um resumo estatístico das RNAs treinadas é apresentado na Tabela 56.

Tabela 56 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-02i.

Rede	RMSE	MAE		R^2	Mínimo	Máximo	Média	σ
RN1.1	5,758	4,732	± 3,281	0,222	34,944	46,307	42,201	3,897
RN1.2	6,139	5,413	± 2,894	0,097	34,972	47,886	42,107	4,254
RN1.3	5,958	5,263	± 2,793	0,094	35,490	47,529	42,141	3,975
RN1.4	5,462	4,637	± 2,887	0,165	35,707	47,048	42,539	3,771
RN1.5	5,975	5,239	± 2,873	0,108	35,193	47,663	42,189	4,100
RN1.6	6,242	5,542	± 2,872	0,090	34,505	48,803	42,331	4,695
RN1.7	5,835	5,050	± 2,922	0,128	35,494	47,125	42,219	3,925
RN1.8	5,756	5,113	± 2,644	0,090	35,943	47,549	42,317	3,855
RN1.9	5,835	5,161	± 2,723	0,088	35,935	47,259	42,150	3,749
RN1.10	5,844	5,108	± 2,838	0,096	36,135	46,695	41,978	3,501
RN1.11	5,369	4,842	± 2,319	0,076	36,665	47,894	42,717	3,696
RN1.12	5,494	4,947	± 2,392	0,074	36,581	47,711	42,539	3,666
RN2.1	2,567	2,407	± 0,890	0,594	41,704	45,936	44,488	1,472
RN2.2	2,437	2,260	± 0,911	0,597	41,791	46,128	44,635	1,510
RN2.3	2,369	2,192	± 0,898	0,593	41,899	46,155	44,703	1,483
RN2.4	2,600	2,447	± 0,879	0,620	41,684	46,023	44,448	1,497
RN2.5	2,762	2,598	± 0,939	0,605	41,374	45,874	44,297	1,563
RN2.6	2,189	1,997	± 0,895	0,600	42,096	46,384	44,898	1,489
RN2.7	2,361	2,171	± 0,926	0,577	41,850	46,132	44,724	1,500
RN2.8	2,353	2,132	± 0,996	0,582	41,690	46,293	44,763	1,611
RN2.9	2,867	2,723	± 0,899	0,589	41,370	45,603	44,172	1,478

Rede	RMSE	MAE		R ²	Mínimo	Máximo	Média	σ
RN2.10	1,930	1,696	± 0,921	0,591	42,330	46,673	45,199	1,514
RN2.11	2,585	2,420	± 0,910	0,601	41,632	45,990	44,475	1,515
RN2.12	2,820	2,657	± 0,944	0,597	41,304	45,775	44,238	1,558

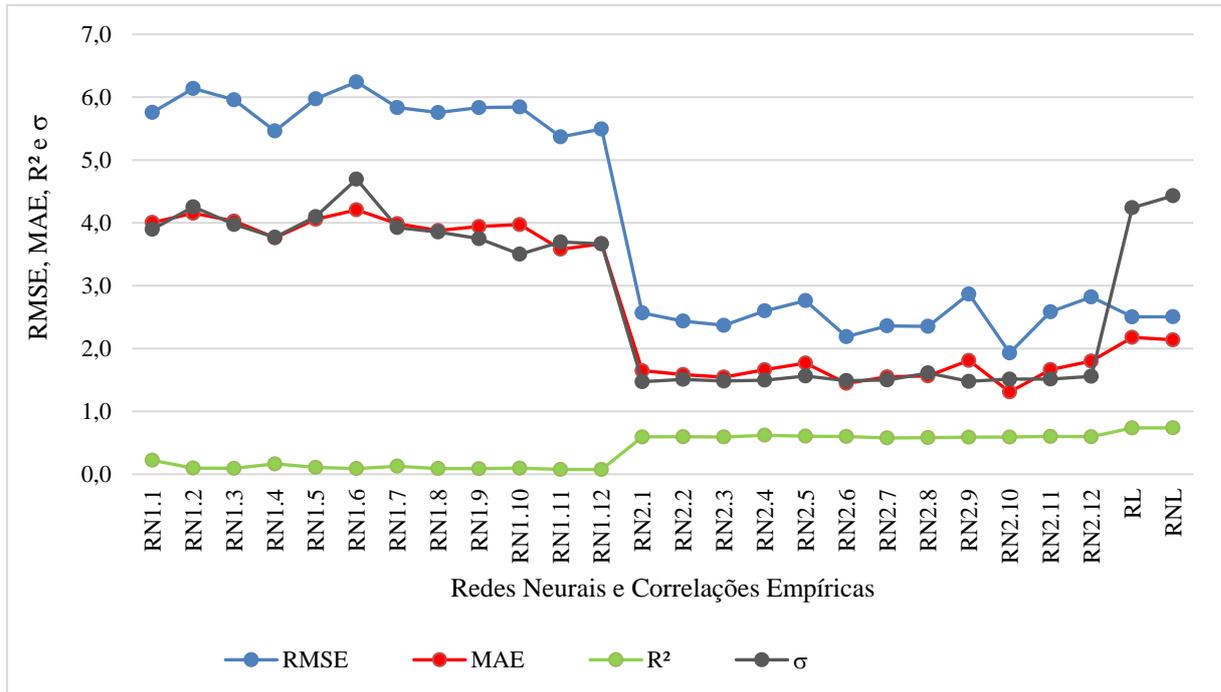
Fonte: Autor (2022)

Observa-se, conforme resultados do Grupo B-02, que as arquiteturas utilizadas, para cada conjunto de dados de entrada, possuem resultados muito próximos. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) proporcionou uma diminuição da média dos erros MAE de 3,94 para 1,61, RMSE de 5,81 para 2,49, e um aumento do valor médio de R² de 0,11 para 0,60 em comparação com os obtidos com a utilização de LL_c, LL_p, IP e Classificação SUCS (RN1.1 a RN1.12). Destaca-se que as redes RN1.1 a RN1.12, apresentaram resultados de R² inferiores a 0,16.

Portanto, analisando os erros MAE e RMSE das redes RN2.1 a RN2.12, a arquitetura com 12 neurônios na primeira camada e 6 na segunda (RN2.10) foi a que apresentou menores erros. Utilizando ainda, os critérios de desvio padrão (σ) e R², para diferenciar os resultados obtidos, a arquitetura com 12 neurônios em uma única camada (RN2.4), apresentou maiores resultados (R²=0,62), entretanto possui maiores erros quando comparada com a RN2.10.

Em comparação com os resultados obtidos pelas análises de regressão, se compararmos apenas o critério de coeficiente de determinação, R², observa-se um resultado maior para as equações de regressão (R² = 0,74). Entretanto, as correlações de regressão apresentaram valores médios de σ superiores. O σ médio foi de 4,34 para as correlações, enquanto para as RNAs esse valor foi de 2,72. A Figura 63 apresenta graficamente o desempenho das redes neurais em comparação com as regressões.

Figura 63 - Resultados estatísticos das redes neurais e regressões do subgrupo B-02i.



Fonte: Autor (2022)

Da Figura 63 nota-se a semelhança entre os modelos de correlação e RNAs obtidas com a utilização dos dados de entrada LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12) (RN1.1 a RN1.12), com destaque para as pequenas diferenças entre os resultados. Uma exceção, no entanto, são os maiores valores de σ das correlações de regressão. Portanto, os resultados indicam que as RNAs, quando adotados os dados de entrada LL_c , LL_p , IP e Classificação SUCS, se adequam melhor a previsão de valores de LL_p para o subgrupo de dados B-02i.

4.2.5 Avaliação estatística dos dados do subgrupo B-02ii

Para os dados do subgrupo B-02ii, aplicada a metodologia citada no item 3.2, foram identificados 4 *outliers*. Portanto, resultaram neste subgrupo 34 amostras aprovadas para serem utilizadas para o estabelecimento das correlações empíricas. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c , LL_p e IP são apresentados no Anexo I.

Foram identificados como *outliers* os maiores valores de LL publicados por Olszewska (2003) e Ozer (2009). A Tabela 57 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 57 - Descrição estatística dos dados investigados no subgrupo B-02ii.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	34	4	69,04	10,64	113,23	15,41	50,50	63,05	66,40	74,40	90,00
LL _p	34	4	64,48	9,17	84,03	14,22	49,80	57,92	62,10	69,00	83,30
IP	34	4	33,81	6,67	44,56	19,74	25,10	27,67	32,60	38,02	49,50

Legenda: N = número de amostras; N* = número de outliers; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

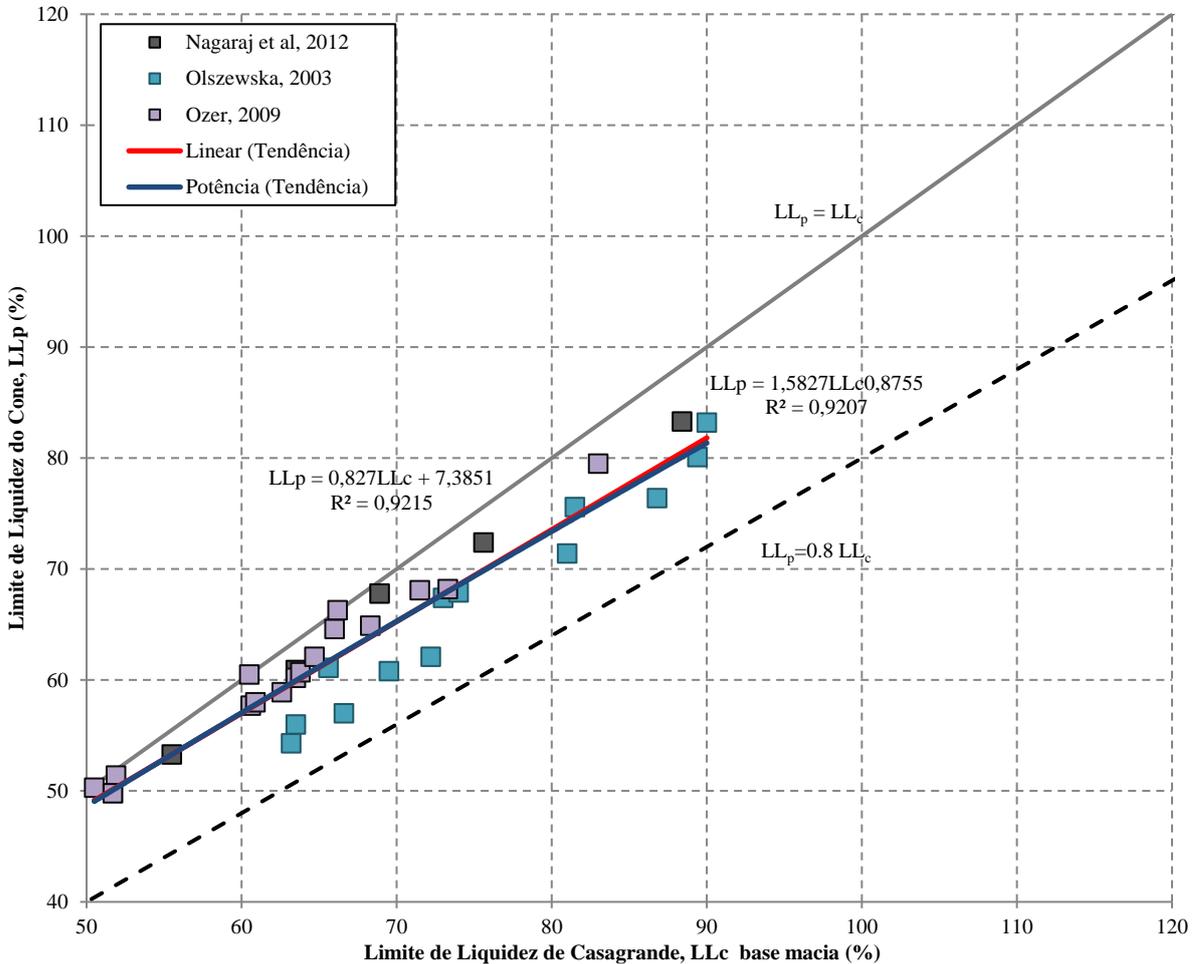
Fonte: Autor (2022)

Os resultados apresentados na Tabela 57, quando comparados com a descrição estatística do grupo B-02, indicam que este subgrupo apresenta dados menos dispersos. Foi observado para o subgrupo B-02ii, uma redução de cerca de 37% nos valores de desvio padrão, 61% da variância e 50% do coeficiente de variação para as variáveis LL_c, LL_p e IP quando comparado com o grupo B-02. O coeficiente de variação observado para as variáveis LL_c, LL_p e IP ficou abaixo de 20%, indicando que este subgrupo é composto por dados mais homogêneos se comparado ao grupo B-02.

Os dados do subgrupo B-02ii, em função do intervalo de LL em que se estão inseridos, são classificados como argilas e siltes de alta compressibilidade (CH e MH). Pouco mais de 50% dos dados são descritos como amostras de argila (CH), entretanto, a granulometria das amostras selecionadas para este subgrupo, de maneira análoga ao grupo B-02, é caracterizada por parcelas de siltes e areias, justificando, portanto, a variação observada quanto a classificação das amostras. A mineralogia dos solos descritas para as amostras de solo que compõem o subgrupo indicam a presença de argilas caulinitas, montmorilonitas e bentonitas. Nagaraj et al. (2012) apresenta amostras compostas por montmorilonitas, enquanto Olszewska (2003) amostras de caulinitas e misturas de caulinita e bentonita em diferentes proporções. No entanto, a mineralogia não é informada para os resultados experimentais compilados de Ozer (2009).

Conforme observado na Tabela 57, apesar do subgrupo B-02ii ser composto inicialmente por resultados de LL_c entre 50 e 120, após a filtragem dos dados, com a eliminação dos outliers, o valor de LL_c para o subgrupo é inferior a 100. A Figura 64 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 64 - Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02ii e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-02ii:

$$LL_p = 0,827LL_c + 7,3851 \qquad R^2 = 0,9215 \qquad (48)$$

$$LL_p = 1,5827LL_c^{0,8755} \qquad R^2 = 0,9207 \qquad (49)$$

Nota-se da Figura 64, que as amostras de solo analisadas se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 0,8 LL_c$ e cerca de 94% dos resultados de LL_p observados são menores que LL_c . As correlações de regressão obtidas também se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 0,8 LL_c$, seguindo, portanto, a tendência observada para o grupo B-02, para valores de LL_c maiores que 50. Observa-se que as correlações descritas para o grupo B-02ii apresentam resultados muito próximos, principalmente para valores de LL_p inferiores a 70, onde as

correlações se mostram praticamente coincidentes. A Tabela 58 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-02ii.

Tabela 58 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02ii.

Regressão	RMSE	MAE	R²	Mínimo	Máximo	Média	σ
Linear - RL	2,530	1,935	0,922	49,149	81,815	64,477	8,800
Não linear - RNL	2,547	1,943	0,921	49,049	81,346	64,412	8,683

Fonte: Autor (2022)

O resumo estatístico confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Ambos, apresentam inclusive valores praticamente coincidentes para os erros RMSE e MAE e para o coeficiente de determinação (R^2). O modelo não linear, no entanto, possui σ pouco menor. A análise de regressão para o subgrupo B-02ii segue, portanto, a tendência observada para o grupo B-02, com resultados semelhantes entre as correlações de regressão.

As correlações obtidas para o subgrupo B-02ii foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados destas comparações constam das Tabela 59 e Tabela 60.

Tabela 59 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-02ii.

LL_c	LL_p		
	Subgrupo B-02ii	Crevelin e Bicalho (2019)	Crevelin e Bicalho (2019)
	50<LL_c<120	30<LL_c<70	50<LL_c<400
50	48,74	52,24	51,47
60	57,01	61,31	58,86
70	65,28	70,38	66,25
80	73,55	-	73,64
90	81,82	-	81,03
100	90,09	-	88,42
110	98,36	-	95,81
120	106,63	-	103,20

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-02ii apresenta resultados próximos aos obtidos por Crevelin e Bicalho (2019) para amostras de montmorilonitas com valores de LL entre 50 e 400. A semelhança observada entre as correlações pode ser explicada ainda pela quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa. Assim, em relação a correlação Crevelin e Bicalho (2019), foram utilizados em comum os

dados de Olszewska (2003), Özer (2009), e Nagaraj et al. (2012). Destaca-se que as amostras utilizadas para o subgrupo B-02ii continha argilas e siltes de baixa e alta plasticidade contendo montmorilonita, obtidos de Nagaraj et al. (2012), argilas de baixa e alta plasticidade de Özer (2009) e misturas de caulinita e bentonita de Olszewska (2003).

Tabela 60 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-02ii.

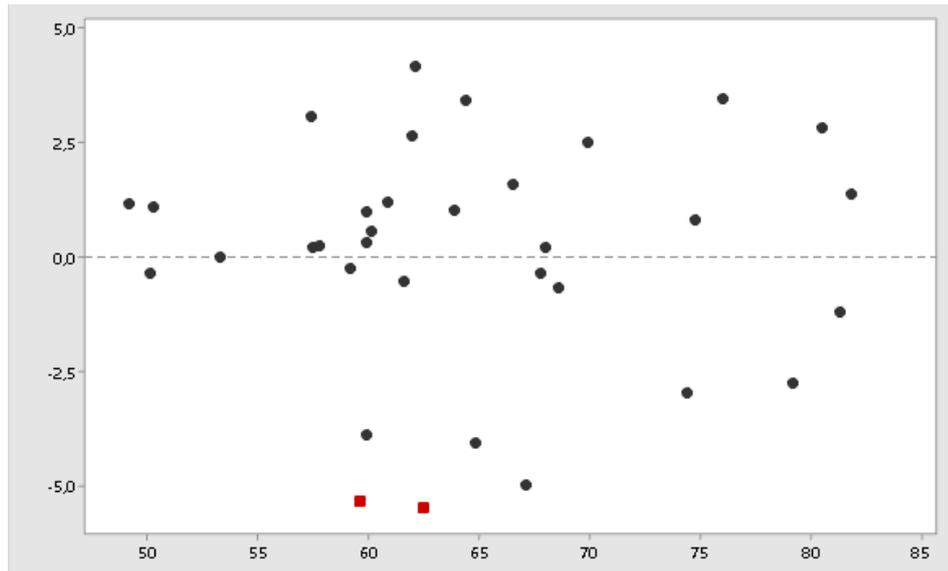
LL_c	LL_p		
	Subgrupo B-02ii	O' Kelly et al. (2018)	O' Kelly et al. (2018)
	$50 < LL_c < 120$	$LL_c < 120$	$LL_c < 600$
50	48,62	50,65	49,73
60	57,04	59,47	57,96
70	65,28	68,11	65,98
80	73,38	76,60	73,81
90	81,35	84,97	81,48
100	89,21	93,22	89,03
110	96,97	101,38	96,45
120	104,65	109,44	103,76

Fonte: Autor (2022)

A correlação não linear obtida para o subgrupo B-02ii mostra-se muito próxima da correlação publicada por O'Kelly et al. (2018) para LL_c de base macia < 600%, embora a correlação para LL_c de base macia < 120% seja mais compatível com o intervalo de LL analisado. Dentre os dados analisados por O'Kelly et al. (2018) apenas os dados publicados por Özer (2009) são comuns aos dados utilizados para a obtenção da correlação não linear para o subgrupo B-02ii. Os resultados foram, portanto, diferentes daqueles obtidos para o Grupo B-02, embora este abrangesse também dados de $LL_c < 120$. Ressalta-se que a concentração dos dados em um dado intervalo de LL para os grupos analisados é que podem ter influenciado na diferença dos resultados obtidos dentre eles, já que no Grupo B-02 os dados concentram-se no intervalo de LL_c entre 30 e 100 e no subgrupo B-02ii entre 50 e 90.

A correlação linear obtida para o subgrupo B-02ii foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos (Figura 65).

Figura 65 - Relatório de diagnóstico de resíduos – Subgrupo B-02ii.

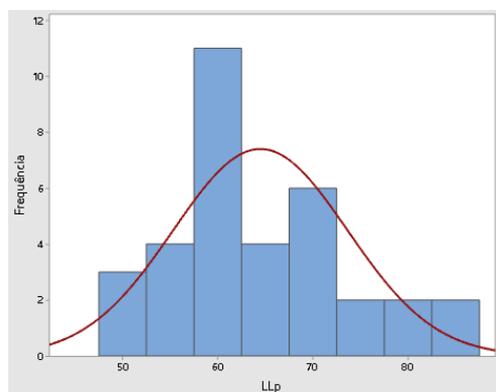


Fonte: Autor (2022)

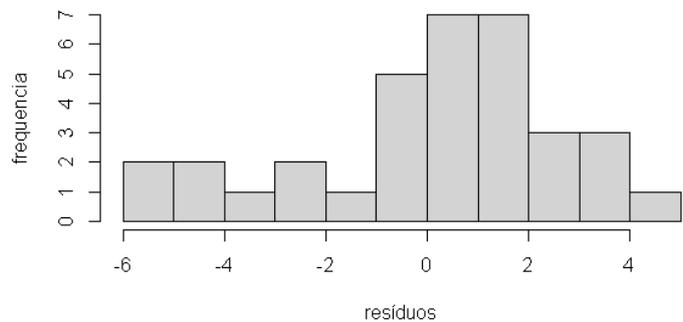
No gráfico da Figura 65 é possível identificar que 2 pontos de dados possuem grandes resíduos, representados pelos pontos vermelhos da figura, não estando bem ajustados pelas equações propostas. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 66.

Figura 66 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-02ii.



(a)



(b)

Fonte: Autor (2022)

Através da Figura 66 observa-se que a variável dependente apresenta um gráfico de distribuição com uma assimetria moderada à esquerda. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 61 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-02ii.

Tabela 61 – Resumo dos resultados dos testes estatísticos para o subgrupo B-02ii.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,0186	Parâmetro significativo
Teste t (x)	2,00E-16	Parâmetro significativo
Teste F	2,20E-16	Parâmetro significativo
Teste KS	2,07E-02	Distribuição dos resíduos não é normal
Teste DW	7,62E-04	Resíduos não são independentes

Fonte: Autor (2022)

Analisando os resultados da Tabela 61, verifica-se que os resultados obtidos para o subgrupo B-02ii são similares aos obtidos para o Grupo B-02.

Assim, os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados. O resultado do teste de aderência de Kolmogorov Smirnov (teste KS) aponta para a distribuição não normal dos resíduos, em consonância com a assimetria observada na análise dos resíduos por meio dos histogramas. O resultado do teste de Durbin Watson (DW), indica que os resíduos do subgrupo B-02ii não são independentes, e, o resultado do teste DW é de 1,0251, ou seja, por se aproximar de 0, indica que há autocorrelação positiva entre as variáveis independentes. Portanto, os resultados do teste DW para o subgrupo B-02ii também não foram satisfatórios, apontando para a inconsistência do modelo.

Portanto, apesar de um R^2 relativamente alto ($> 0,92$) para as regressões lineares e não lineares, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios dos testes estatísticos dos resíduos indicaram que a correlação linear apresentada não é significativa.

4.2.6 Aplicação das RNAs nas amostras do grupo subgrupo B-02ii

Utilizando os dados filtrados nas análises estatísticas, as 34 amostras aprovadas no subgrupo B-02ii foram submetidas as redes neurais artificiais. Um resumo do desempenho das RNAs é apresentado na Tabela 62.

Tabela 62 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-02ii.

Rede	RMSE	MAE		R^2	Mínimo	Máximo	Média	σ
RN1.1	2,249	1,305	± 1,831	0,891	58,729	80,362	64,439	6,538
RN1.2	2,449	1,389	± 2,017	0,880	58,340	78,879	63,970	6,136
RN1.3	2,251	1,250	± 1,872	0,883	58,784	80,534	64,559	6,521
RN1.4	2,223	1,188	± 1,879	0,891	58,977	80,395	64,481	6,507
RN1.5	2,163	1,203	± 1,798	0,888	59,062	79,862	64,492	6,263
RN1.6	2,213	1,862	± 1,196	0,854	59,934	79,650	65,725	5,839
RN1.7	2,491	1,444	± 2,029	0,860	58,367	77,395	64,041	5,759
RN1.8	2,225	1,177	± 1,888	0,881	58,889	79,598	64,509	6,306
RN1.9	2,530	1,444	± 2,078	0,875	58,249	78,873	63,919	6,218
RN1.10	2,327	1,188	± 2,001	0,884	58,661	79,328	64,221	6,286
RN1.11	2,081	1,210	± 1,693	0,883	59,336	78,967	64,658	5,977
RN1.12	2,032	1,296	± 1,565	0,886	59,499	79,923	64,926	6,188
RN2.1	2,846	2,511	± 1,339	0,946	58,110	76,056	62,839	5,487
RN2.2	3,455	3,191	± 1,326	0,947	57,343	75,354	62,159	5,509
RN2.3	2,362	1,951	± 1,330	0,946	58,676	76,727	63,399	5,509
RN2.4	2,656	2,282	± 1,359	0,947	58,460	76,001	63,068	5,348
RN2.5	2,733	2,378	± 1,346	0,946	58,351	76,301	62,972	5,473
RN2.6	2,108	1,661	± 1,297	0,947	59,115	76,840	63,732	5,403
RN2.7	3,036	2,726	± 1,338	0,944	57,785	76,244	62,624	5,661
RN2.8	2,897	2,572	± 1,334	0,945	57,999	76,212	62,778	5,574
RN2.9	3,324	3,047	± 1,328	0,946	57,531	75,738	62,303	5,565
RN2.10	2,896	2,570	± 1,337	0,946	58,041	76,106	62,780	5,524
RN2.11	2,862	2,531	± 1,337	0,946	58,092	76,113	62,819	5,507
RN2.12	2,576	2,193	± 1,352	0,946	58,498	76,281	63,157	5,435

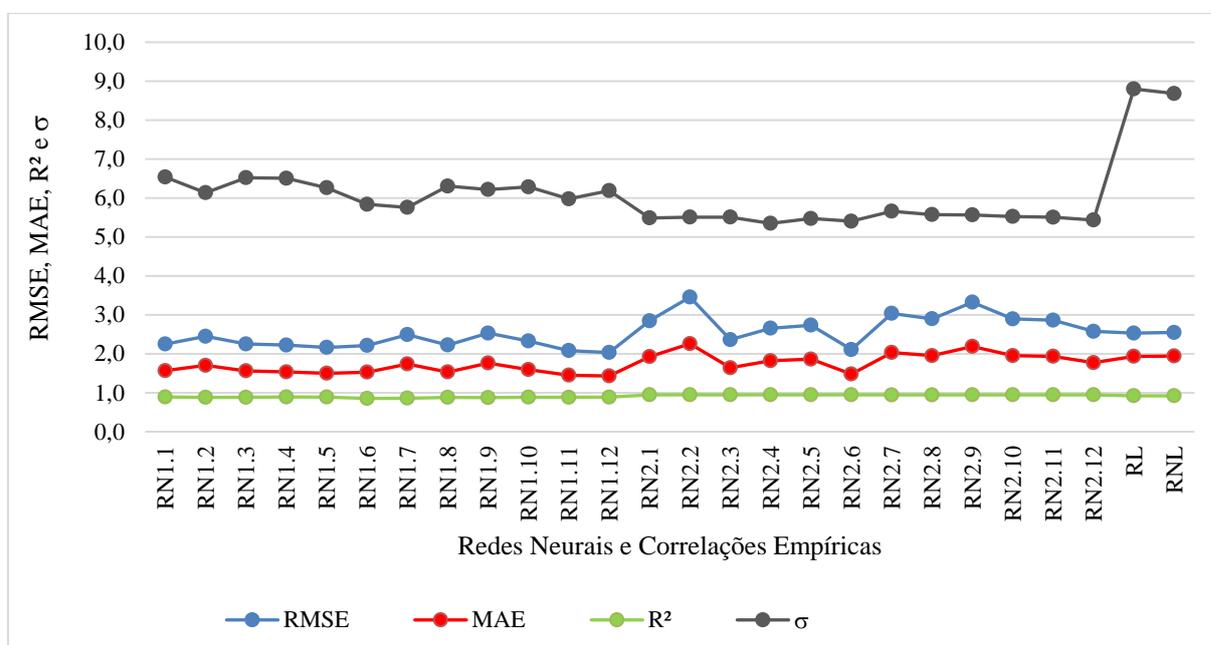
Fonte: Autor (2022)

Observa-se, conforme resultados do Grupo B-01, que as arquiteturas utilizadas, para cada conjunto de dados de entrada, possuem resultados muito próximos. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) apresentou menor σ e maior R^2 , face aos obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Os valores médios de σ foram reduzidos de 6,2 para 5,5 e os valores de médios de R^2 aumentados de 0,88 para 0,95. Apesar dos menores erros observados entre RN1.1 e RN1.12, ressalta-se que a diferença entre os parâmetros foi mais significativa entre o σ e R^2 . Portanto, analisando os parâmetros de erro MAE, RMSE, e R^2 das redes RN2.1 a RN2.12, a arquitetura com 18 neurônios em uma única camada (RN2.6) foi a que apresentou menores valores de erro (RMSE=2,10 e MAE=1,48) e maior R^2 (0,95).

Em comparação com os resultados obtidos pelas análises de regressão, observa-se resultados muito próximos para os erros RMSE, MAE e para o R^2 . Entretanto, quando analisado o σ nota-se um valor médio de 5,86 para as RNAs e 8,74 para as correlações de

regressão. A Figura 67 apresenta graficamente o desempenho das redes neurais em comparação com as regressões.

Figura 67 - Resultados estatísticos das redes neurais e regressões do subgrupo B-02ii.



Fonte: Autor (2022)

Da Figura 67 nota-se a semelhança entre os modelos de correlação e RNAs, com destaque para as pequenas diferenças entre os resultados. Uma exceção, no entanto, são os maiores valores de σ das correlações de regressão. Portanto, os resultados indicam que as RNAs se adequam melhor a previsão de valores de LL_p para o subgrupo de dados B-02ii.

4.2.7 Avaliação estatística dos dados do subgrupo B-02iii

Para os dados do subgrupo B-02iii não foram identificados *outliers*, portanto, 6 amostras foram utilizadas para o estabelecimento das correlações empíricas, bem como no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c , LL_p e IP são apresentados no Anexo I.

A Tabela 63 apresenta um resumo estatístico das amostras de solo investigadas.

Tabela 63 - Descrição estatística dos dados investigados no subgrupo B-02iii.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL_c	6	0	176,2	35,8	1281,0	20,32	127,7	142,4	176,0	210,2	224,5
LL_p	6	0	149,4	25,9	671,1	17,34	113,3	124,6	151,9	171,2	183,7
IP	6	0	115,2	28,1	788,4	24,37	76,4	88,5	115,8	142,3	152,0

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

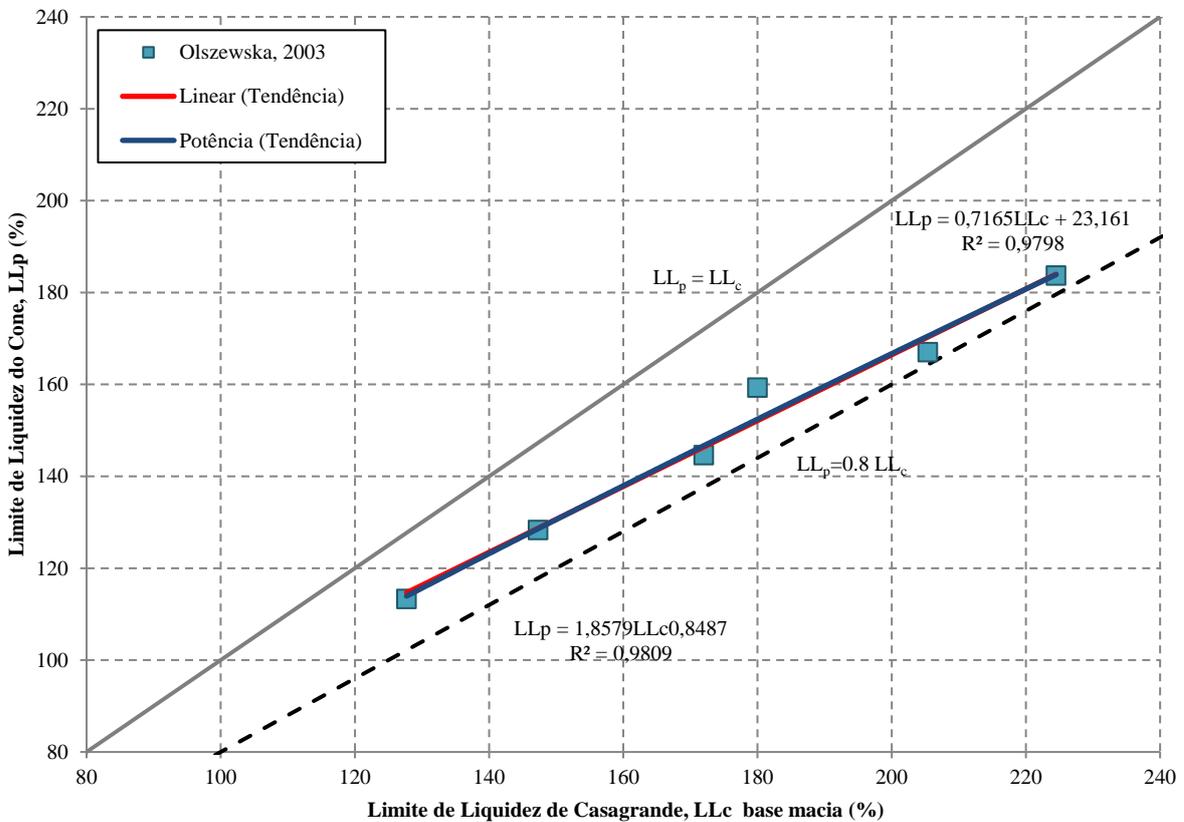
Fonte: Autor (2022)

Os resultados apresentados na Tabela 63, quando comparados com a descrição estatística do grupo B-02, indicam que este subgrupo apresenta dados menos dispersos. Foi observado para o subgrupo B-01i, um aumento de cerca de 53% nos valores de desvio padrão e 77% da variância e uma redução de aproximadamente 37% do coeficiente de variação para as variáveis LL_c , LL_p e IP quando comparado com o grupo B-02. Esses resultados indicam que a subdivisão dos dados do subgrupo B-02iii não resultou em um conjunto de dados menos dispersos face aos obtidos para o grupo B-02.

Esta homogeneidade dos dados também é observada quando analisados a classificação SUCS das amostras de solo e sua mineralogia. Os dados do subgrupo B-02iii, $LL_c > 120$, em função do intervalo de LL em que se estão inseridos, são classificados como argilas de alta compressibilidade (CH). E, a mineralogia dos solos descritas para as amostras indicam a mistura de caulinitas e bentonitas em diferentes proporções. Ressalta-se que este subgrupo é composto apenas pelos dados publicados por Olszewska (2003).

A Figura 68 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 68 - Comparações entre os valores de LL obtidos pelo método do cone britânico e pelo método de Casagrande de base macia para os dados compilados do subgrupo B-02iii e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-02iii:

$$LL_p = 0,7165LL_c + 23,161 \quad R^2 = 0,9787 \quad (50)$$

$$LL_p = 1,8579LL_c^{0,8487} \quad R^2 = 0,9809 \quad (51)$$

Nota-se da Figura 68, que as correlações obtidas possuem valores próximos, e se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 0,8 LL_c$, assim como o conjunto de dados analisados, seguindo, portanto, a tendência observada para o grupo B-02, para valores de LL_c maiores que 50. Além disso, todas as amostras de solo analisadas apresentaram valores de LL_p maiores que LL_c . A Tabela 64 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-02iii.

Tabela 64 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-02iii.

Regressão	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
Linear - RL	3,362	2,395	0,979	114,658	184,015	149,372	25,644
Não linear - RNL	3,270	2,247	0,981	113,906	183,865	149,324	25,857

Fonte: Autor (2022)

O resumo estatístico confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Ambos, apresentam inclusive um mesmo coeficiente de determinação (R^2) e valores muito próximos para os erros RMSE, MAE e σ . A análise de regressão para o subgrupo B-02iii segue, portanto, a tendência observada para o grupo B-02, com resultados semelhantes entre as correlações de regressão.

As correlações obtidas para o subgrupo B-02iii foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados obtidos nesta pesquisa não puderam ser comparados com os de Olszewska (2003), embora o conjunto de dados esteja limitado aos resultados de sua publicação, já que o pesquisador não apresentou uma correlação entre os valores de LL_c e LL_p . Os resultados destas comparações constam das Tabela 65 e Tabela 66.

Tabela 65 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-02iii.

LL _c	LL _p	
	Subgrupo B-02iii	Crevelin e Bicalho (2019)
	LL _c >120	50<LL _c <400
30	44,66	36,69
40	51,82	44,08

LL _c	LL _p	
	Subgrupo B-02iii	Crevelin e Bicalho (2019)
	LL _c >120	50<LL _c <400
50	58,99	51,47
60	66,15	58,86
70	73,32	66,25
80	80,48	73,64
90	87,65	81,03
100	94,81	88,42
110	101,98	95,81
120	109,14	103,20
150	130,64	125,37
200	166,46	162,32
250	202,29	199,27
300	238,11	236,22
350	273,94	273,17
400	309,76	310,12

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-02iii apresenta resultados próximos aos obtidos por Crevelin e Bicalho (2019) para amostras de montmorilonitas com valores de LL entre 50 e 400, para o intervalo de LL>200. Portanto, embora os dados analisados se concentrem em uma faixa de valores inferior a 200, nota-se que a extrapolação dos resultados, com a utilização da correlação para explicação do modelo de LL>200 se mostrou eficiente. A semelhança observada entre as correlações pode ser explicada ainda pela quantidade de dados utilizados em comum para a obtenção das correlações desta pesquisa, já que a correlação de Crevelin e Bicalho (2019) utilizou os dados de Olszewska (2003).

Tabela 66 - Comparação entre os resultados previstos de LL_p entre correlações não lineares da literatura e a obtida para o subgrupo B-02iii.

LL _c	LL _p	
	Subgrupo B-02iii	O' Kelly et al. (2018)
	>120	<600
Intervalo de LL		
30	33,32	32,38
40	42,53	41,23
50	51,40	49,73
60	60,00	57,96
70	68,38	65,98
80	76,59	73,81
90	84,64	81,48
100	92,56	89,03
110	100,36	96,45
120	108,05	103,76

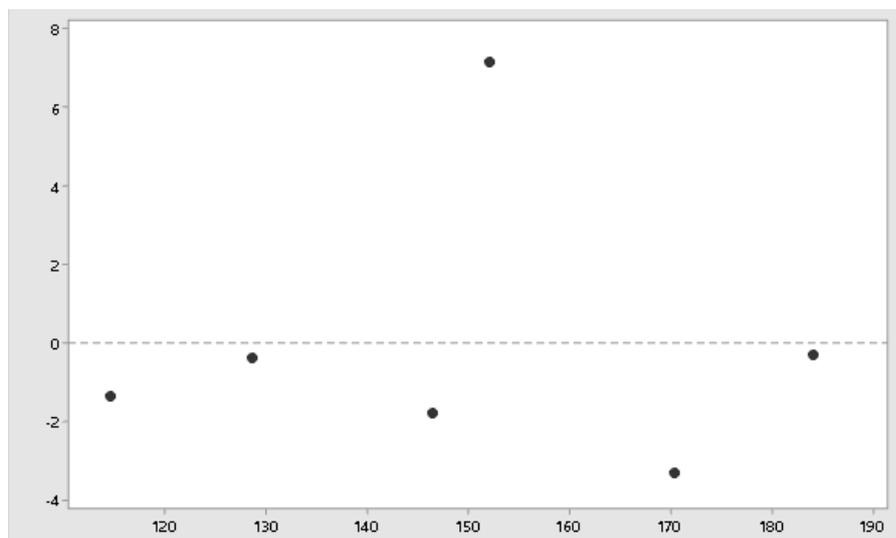
LL _c	LL _p	
	Subgrupo B-02iii	O' Kelly et al. (2018)
Intervalo de LL	>120	<600
150	130,58	125,15
200	166,69	159,36
250	201,44	192,21
300	235,16	224,02
350	268,02	254,99
400	300,19	285,26
450	331,74	314,93
500	362,78	344,07
550	393,34	372,75
600	423,49	401,01

Fonte: Autor (2022)

A correlação não linear obtida para o subgrupo B-02ii mostra-se muito próxima da correlação publicada por O'Kelly et al. (2018) para LL_c de base macia < 600%, para o intervalo de LL<200. Observa-se neste caso a influência da concentração dos dados na faixa de valores inferior a 200 para o subgrupo B-02iii. Vale ressaltar que os dados do subgrupo B-02iii não possuem dados em comuns aos dados de O'Kelly et al. (2018), não sendo, portanto, este um fator de influência para os resultados analisados.

A correlação linear obtida para o subgrupo B-02iii foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos (Figura 69).

Figura 69 - Relatório de diagnóstico de resíduos – Subgrupo B-02iii.

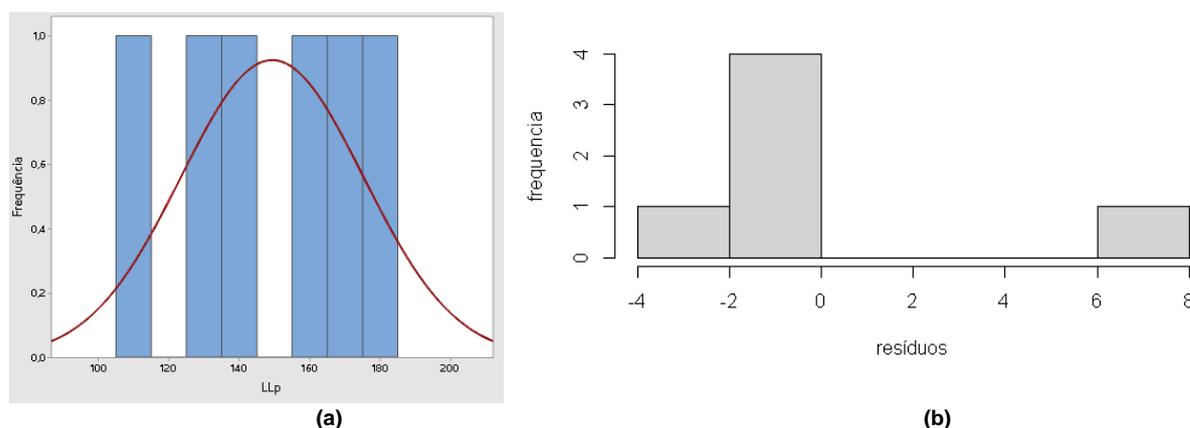


Fonte: Autor (2022)

No gráfico da Figura 69 não foram identificados pontos de dados com grandes resíduos, entretanto, o comportamento dos dados no gráfico de resíduos, e sua distribuição em uma faixa horizontal de alta amplitude, apontam para uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 70.

Figura 70 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-02iii.



Fonte: Autor (2022)

Através da Figura 70 observa-se que a variável dependente apresenta um gráfico de distribuição uniforme, em que todas as barras do gráfico apresentam a mesma altura, ou seja, todos os valores de LL_p são igualmente prováveis de ocorrer. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 67 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-02iii.

Tabela 67 – Resumo dos resultados dos testes estatísticos para o subgrupo B-02iii.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,06588	Parâmetro não significativo
Teste t (x)	1,54E-04	Parâmetro significativo
Teste F	1,54E-04	Parâmetro significativo
Teste KS	0,1189	Distribuição Normal dos resíduos
Teste DW	0,8115	Resíduos independentes

Fonte: Autor (2022)

Analisando os resultados da Tabela 67, verifica-se que os resultados obtidos para o subgrupo B-02iii são distintos aos obtidos para o Grupo B-02.

Assim, os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados, embora o Teste t (intercepto) aponte para a insignificância do parâmetro. O resultado do teste de aderência de Kolmogorov Smirnov (teste KS) aponta para a distribuição normal dos resíduos. O resultado do teste de Durbin Watson (DW), indica que os resíduos do subgrupo B-02iii são independentes, e, o resultado do teste DW é de 2,9872, ou seja, por se aproximar de 4, indica que há autocorrelação negativa entre as variáveis independentes, conforme observado no gráfico de diagnósticos de resíduos. Portanto, os resultados do teste DW para o subgrupo B-02iii não foram satisfatórios, apontando para a inconsistência do modelo.

Portanto, o modelo apresenta um R^2 alto ($> 0,97$) para as regressões lineares e não lineares, significância dos parâmetros de regressão observados com os testes de hipóteses, e distribuição normal dos resíduos. Entretanto, tendo em vista a heterocedasticidade e a autocorrelação negativa dos resíduos, de uma maneira geral, a correlação linear apresentada não é significativa.

4.2.8 Aplicação das RNAs nas amostras do grupo subgrupo B-02iii

Utilizando os dados filtrados nas análises estatísticas, as 34 amostras aprovadas no subgrupo B-02iii foram submetidas as redes neurais artificiais. Um resumo do desempenho das RNAs é apresentado na Tabela 68.

Tabela 68 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-02iii.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	4,990	3,765 ± 3,276	1,000	174,041	183,211	178,626	6,484
RN1.2	5,926	5,225 ± 2,795	1,000	175,020	186,130	180,575	7,856
RN1.3	6,151	5,513 ± 2,727	1,000	175,240	186,486	180,863	7,952
RN1.4	7,075	6,813 ± 1,907	1,000	175,720	188,605	182,163	9,111
RN1.5	6,829	6,503 ± 2,085	1,000	175,589	188,118	181,853	8,859
RN1.6	7,019	6,773 ± 1,840	1,000	175,613	188,633	182,123	9,207
RN1.7	25,211	23,861 ± 8,140	1,000	151,279	151,699	151,489	0,297
RN1.8	18,052	16,277 ± 7,805	1,000	158,528	159,618	159,073	0,771
RN1.9	16,301	14,360 ± 7,715	1,000	160,356	161,625	160,990	0,897
RN1.10	9,745	7,245 ± 6,517	1,000	166,272	169,938	168,105	2,592
RN1.11	5,877	5,470 ± 2,149	1,000	170,321	176,081	173,201	4,073
RN1.12	5,277	5,238 ± 0,644	1,000	171,594	177,818	174,706	4,401
RN2.1	5,095	4,826 ± 1,634	1,000	173,460	180,508	176,984	4,984
RN2.2	4,981	4,344 ± 2,437	1,000	173,782	181,793	177,787	5,665
RN2.3	5,074	3,641 ± 3,534	1,000	174,175	185,593	178,884	6,659
RN2.4	5,063	3,918 ± 3,208	1,000	174,125	184,410	179,268	7,272
RN2.5	5,497	4,515 ± 3,136	1,000	174,651	185,079	179,865	7,374
RN2.6	5,575	4,838 ± 2,770	1,000	174,608	185,769	180,188	7,892

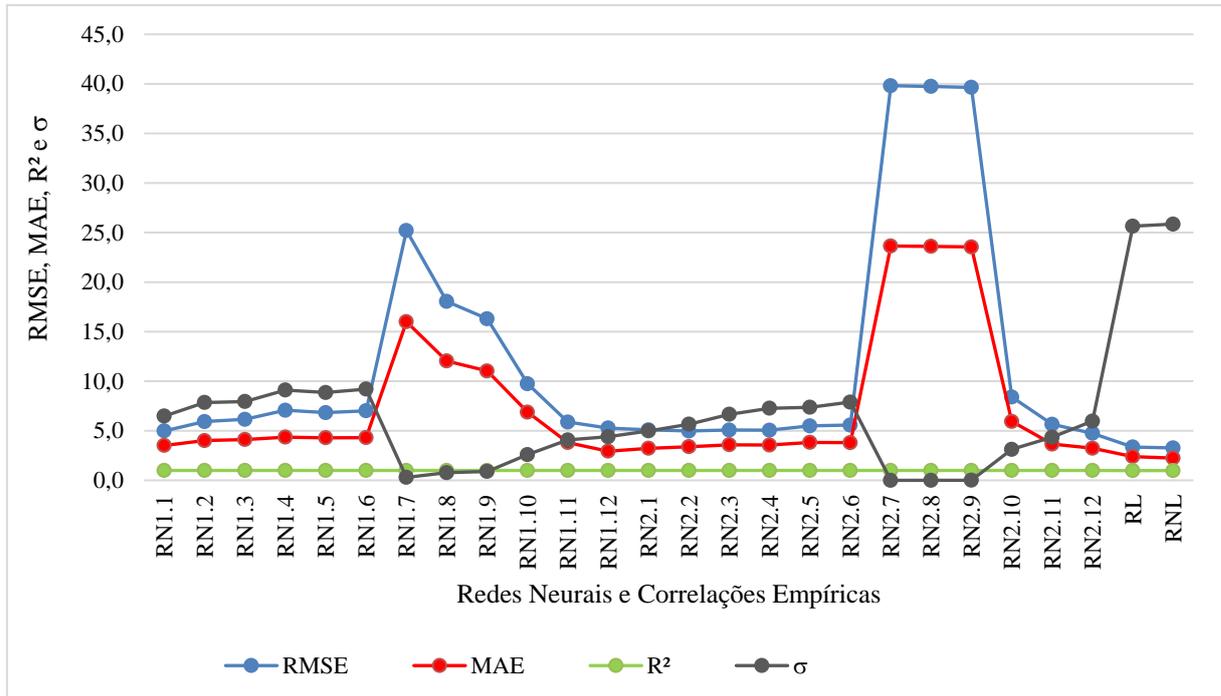
Rede	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
RN2.7	39,818	38,933 ± 8,350	1,000	136,417	136,417	136,417	0,000
RN2.8	39,744	38,858 ± 8,349	1,000	136,491	136,494	136,492	0,002
RN2.9	39,650	38,763 ± 8,342	1,000	136,579	136,595	136,587	0,011
RN2.10	8,412	6,138 ± 5,751	1,000	167,387	171,810	169,599	3,128
RN2.11	5,658	5,275 ± 2,048	1,000	170,227	176,378	173,302	4,349
RN2.12	4,745	4,121 ± 2,352	1,000	173,473	181,931	177,702	5,981

Fonte: Autor (2022)

Observa-se, conforme resultados do Grupo B-02, que as arquiteturas utilizadas, para cada conjunto de dados de entrada, possuem resultados muito próximos. Dentre as arquiteturas analisadas para as redes neurais, observa-se um comportamento atípico das redes que continham 3 neurônios na primeira camada e 1 na segunda, 6 neurônios na primeira camada e 3 na segunda e 9 neurônios na primeira camada e 4 na segunda, ou seja, RN1.7 a RN1.9 e RN2.7 a RN2.9. Tais arquiteturas apresentaram erros RMSE e MAE muito superiores aos verificados para as demais redes e regressões, embora os pequenos valores de σ registrados. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) apresentou desempenho semelhante aos obtidos com a utilização de LL_c, LL_p, IP e Classificação SUCS (RN1.1 a RN1.12), já que os dados analisados se mostraram bem variáveis. Entretanto, analisando o erro MAE e o σ , a arquitetura com 18 neurônios na primeira camada e 9 na segunda (RN1.12) foi a que apresentou menores valores.

Em comparação com os resultados obtidos pelas análises de regressão, observa-se um resultado semelhante para os valores de R². Entretanto, a média dos erros RMSE e MAE e do σ apresentam grandes diferenças de valores quando comparadas as correlações de regressão e as RNAs. Assim, os erros médios RMSE variam de 11,99 para 3,32 e os erros MAE de 7,60 para 2,32 das RNAs para as correlações de regressão. O σ médio observado foi de 4,83 para as RNAs e de 25,75 para as correlações de regressão. A Figura 71 apresenta graficamente o desempenho das redes neurais em comparação com as regressões.

Figura 71 - Resultados estatísticos das redes neurais e regressões do subgrupo B-02iii.



Fonte: Autor (2022)

Da Figura 71 nota-se maiores diferenças entre os modelos de correlação e RNAs do que os observados para os demais subgrupos. No entanto, observa-se maiores valores de σ das correlações de regressão. Portanto, se desconsiderados os resultados das RN1.7 a RN1.9 e RN2.7 a RN2.9, as demais arquiteturas de RNAs se adequam melhor a previsão de valores de LL_p para o subgrupo de dados B-02iii.

4.2.9 Conclusões parciais para o grupo de dados B-02

Apresentam-se os resultados das análises estatísticas e das redes neurais artificiais dos dados experimentais correspondentes ao grupo B-02 e seus subgrupos (B-02i, B-02ii e B-02iii) e suas discussões. O grupo B-02 e seus subgrupos correspondem aos resultados experimentais de LL_c obtidos pelo aparelho de Casagrande de base macia e LL_p obtido pelo cone britânico.

Conforme citado nas Conclusões parciais para o grupo de dados B-01, item 4.1.9, os resultados experimentais previamente compilados da literatura foram selecionados para as análises estatísticas e redes neurais artificiais conforme a metodologia abordada no item 3.1.2. Para a análise da influência do intervalo de LL nos valores de LL_c e LL_p , os grupos foram separados em três subgrupos: B-02i ($LL < 50$), B-02ii ($50 < LL \leq 120$) e B-02iii ($LL > 120$).

Os dados, devido a variabilidade observada nas amostras do banco de dados, também foram submetidos a definição e eliminação dos *outliers* por meio dos gráficos de *outliers* e diagramas

boxplot, conforme metodologia descrita no item 3.2. O grupo B-02 e seus subgrupos possuem resultados experimentais de diferentes países como Índia, Portugal e Turquia. De uma maneira geral, a eliminação dos *outliers* resultou em uma redução do número de amostras analisadas nesta dissertação, além de reduzirem o intervalo inicial de LL para cada grupo ou subgrupo. A maioria dos *outliers* corresponde aos maiores valores de LL_c , LL_p ou IP das amostras de cada conjunto de dados. Como resultado, o grupo B-02 e o subgrupo B-02ii, foram limitados a valores de $LL_c < 100$, obtidos pelos aparelhos de Casagrande de base macia. Para o subgrupo B-02iii não foram identificados *outliers*.

A descrição estatística mostrou que os resultados obtidos com a seleção dos dados e eliminação dos *outliers* proporcionaram uma redução de cerca de 70% nos valores de desvio padrão, 92% da variância e 54% do coeficiente de variação para as variáveis LL_c , LL_p e IP. Esses resultados foram muito próximos aos obtidos para o grupo B-01. Da divisão do grupo B-02 em subgrupos de dados, notou-se uma redução superior a 37% no desvio padrão, 61% na variância e 37% para o coeficiente de variação. Os coeficientes de variação foram inferiores a 25% para todos os subgrupos de dados. Assim, conclui-se que em relação ao grupo de dados B-02, a divisão dos dados em subgrupos possibilitou a formação de conjuntos de dados mais homogêneos.

A variabilidade das amostras também foi avaliada segundo as classificações dos solos, granulometria e tipos de argilominerais. Em todos os conjuntos de dados (grupo B-02, subgrupo B-02i, subgrupo B-02ii e subgrupo B-02iii) em torno de 50% das amostras foram classificadas como argilas de alta e baixa compressibilidade (CL e CH). As demais amostras foram classificadas como siltes de baixa e alta compressibilidade (ML e MH). Apenas o subgrupo B-02iii apresentou toda as amostras classificadas como argilas de alta compressibilidade. Da granulometria informada nas publicações de Ozer (2009), Sousa (2011) e Nagaraj et al (2012), foi observado que as amostras são compostas por frações de silte e areia. O posicionamento dos valores de LL_c das amostras no ábaco de Casagrande apresentado no item 3.1.2.2 e conseqüentemente sua classificação SUCS, corroboram as informações da granulometria citada nas referidas publicações. No grupo B-02 foi notada uma variedade maior de classificações, contendo amostras de argilas e siltes de baixa e alta compressibilidade, ou seja, CL, ML, CH e MH. Os demais subgrupos tiveram as classificações atribuídas ao intervalo de LL em que estão inseridos. Assim, o subgrupo B-02i, $LL < 50$, contém amostras de solo de baixa compressibilidade, CL e ML. Os subgrupos B-02ii e B-02iii, $LL > 50$, solos de alta compressibilidade, CH e MH. Estas observações são semelhantes as destacadas para o grupo B-01.

A mineralogia dos solos estudados não é informada por Ozer (2009), no entanto, Olszewska (2003), Sousa, (2011) e Nagaraj et al. (2012) relatam amostras compostas por caulinitas, ilitas, bentonitas e montmorilonitas, e em alguns casos a mistura desses argilominerais. A presença desses argilominerais também é variável conforme os intervalos de LL analisados. Assim, o subgrupo B-02i $LL_c < 50$, são relatados, em sua maioria, amostras compostas por caulinitas e ilitas. Para os subgrupos B-02ii e B-02iii, $LL_c > 50$, são relatados em sua maioria montmorilonitas e misturas de bentonitas. Os intervalos citados diferem daqueles observados para o grupo B-01.

Os resultados obtidos nesta pesquisa indicam que, para $LL_c < 50$, subgrupo B-02i, cerca de 72% das amostras apresentam $LL_c < LL_p$, enquanto para valores de $LL_c > 50$, subgrupos B-02ii e B-02iii, mais de 94% apresentam $LL_c > LL_p$. Destaca-se, portanto, conforme citado para o grupo B-01, que as diferenças entre LL_c e LL_p não refletem a totalidade dos resultados experimentais.

De maneira análoga a descrição apresentada para o grupo B-01, item 4.1.9, os dados de cada grupo e subgrupo foram submetidos a análises de regressão para definição das correlações de regressão lineares e não lineares e comparados através da metodologia abordada no item 3.4. A comparação entre as correlações lineares e não lineares dos conjuntos de dados analisados indicam que os modelos propostos resultam em valores muito próximos para a previsão de LL_p . Os valores de R^2 foram praticamente coincidentes para todos os conjuntos de dados do grupo B-02 e as diferenças entre os erros RMSE, MAE e σ foram pequenas. Portanto, conclui-se pela semelhança entre as correlações de regressão na explicação dos modelos propostos.

Os resultados experimentais de LL_c e LL_p que compõem os conjuntos de dados, juntamente com as correlações de regressão obtidas de maneira análoga ao grupo B-01, também foram plotados no gráfico $LL_c \times LL_p$ para a análise gráfica dos resultados (itens 4.2.1, 4.2.3, 4.2.5 e 4.2.7). Foi observado que as correlações obtidas para o grupo B-02 se encontram entre as linhas $LL_p = LL_c$ e $LL_p = 1,2 LL_c$, para valores de LL_c menores que 50 e entre as linhas $LL_p = LL_c$ e $LL_p = 0,8 LL_c$ para valores de LL_c superiores a 50, sendo essa tendência, portanto, confirmada pelas correlações obtidas para os subgrupos B-02i, B-02ii e B-02iii. Ou seja, as correlações se mostram acima da linha de igualdade $LL_c = LL_p$ para $LL_c < 50$ e abaixo da linha de igualdade para $LL_c > 50$. Portanto, para $LL_c > 50$, aplicando as correlações propostas, os valores previstos de LL_p são menores que os resultados experimentais de LL_c .

As correlações lineares propostas também foram submetidas às análises de resíduos e aos testes de hipóteses para verificação da normalidade do modelo e independência das variáveis, conforme metodologia citada no item 3.2. Para as correlações lineares propostas,

em todos os conjuntos de dados analisados foram observados a heterocedasticidade e que a distribuição dos resíduos para os modelos não é normal. Apenas os resíduos verificados para o subgrupo B-02iii foram classificados como independentes, os demais foram todos ditos dependentes. Portanto, os testes estatísticos mostraram que as correlações lineares propostas para os conjuntos de dados do grupo B-02 e subgrupos B02i, B-02ii e B-02iii não são significativas, apesar dos elevados valores de coeficiente de determinação R^2 ($R^2 > 0,74$). Esses resultados são semelhantes aos observados para os conjuntos de dados do grupo B-01, conforme descrito no item 4.1.9, embora tenha sido observada uma maior homogeneidade das amostras de solo para os subgrupos B-02i, B-02ii e B-02iii. Ainda em conformidade com as conclusões apresentadas para o grupo B-01, item 4.1.9, a presença de dados atípicos nos conjuntos de dados analisados verificados com a assimetria dos histogramas (Figura 58, Figura 62, Figura 66 e Figura 70) e com os pontos de dados identificados como resíduos (Figura 57, Figura 61, Figura 65 e Figura 69) justificam a heterocedasticidade e a distribuição não normal dos resíduos, resultados determinantes para a conclusão de que as correlações lineares não são significativas.

Os conjuntos de dados também foram analisados pelas redes neurais artificiais para previsão de LL_p , conforme metodologia citada no item 3.3. De maneira análoga ao grupo B-01, diferentes arquiteturas e dados de entrada foram testados para a avaliação da capacidade de previsão das redes e sua consequente validação. Quanto ao número de neurônios e camadas ocultas, as arquiteturas propostas para os conjuntos de dados do grupo B-02 e subgrupos B-02i e B-02ii apresentam resultados semelhantes. Entretanto, para o subgrupo B-02iii, foi observado um comportamento atípico das redes que continham 3 neurônios na primeira camada e 1 na segunda, 6 neurônios na primeira camada e 3 na segunda e 9 neurônios na primeira camada e 4 na segunda, ou seja, RN1.7 a RN1.9 e RN2.7 a RN2.9. Esses resultados indicaram diferenças superiores a 50% para as médias de erros RMSE e MAE quando comparado com as médias obtidas para todas as RNAs avaliadas. Com esses resultados, nota-se que a utilização de arquiteturas com duas camadas ocultas, para o subgrupo B-02iii, proporcionou o aumento dos erros RMSE e MAE em mais de 66% se comparado com as arquiteturas de apenas uma camada oculta. Entretanto, os valores de σ foram cerca de 70% menores para as arquiteturas com apenas uma camada oculta.

Os resultados, quanto a entrada e dados, indicaram uma melhor previsão dos valores de LL_p considerando apenas os dados de entrada LL_c e LL_p . Para o subgrupo B-02i, entretanto, as diferenças entre a utilização dos dados de entrada foram maiores do que as observadas para os demais subgrupos. Para esse subgrupo, a utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) proporcionou uma diminuição da média dos erros MAE de 3,94 para

1,61, RMSE de 5,81 para 2,49, e um aumento do valor médio de R^2 de 0,11 para 0,60 em comparação com os obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Além disso, as redes RN1.1 a RN1.12 do subgrupo B-02i, apresentaram resultados de R^2 inferiores a 0,16. Portanto, assim como concluído para o grupo B-01, item 4.1.9, os resultados indicam uma maior influência das variáveis LL_c e LL_p na definição dos modelos, tendo em vista que os valores de IP e da classificação SUCS dependem de LL.

Ainda de maneira análoga ao grupo B-01, os resultados obtidos com as correlações de regressão foram confrontados com os resultados obtidos pelas redes neurais artificiais. Foi observado que os valores de σ são maiores para as correlações de regressão de todos os conjuntos de dados analisados (grupo B-02, subgrupo B-02i, subgrupo B-02ii e subgrupo B-02iii). Maiores diferenças foram observadas para o subgrupo B-02iii, onde o σ médio das correlações de regressão foi cerca de 82% maior do que o observado para as RNAs. Para os demais conjuntos de dados, essa diferença média foi inferior a 40%. Portanto, observadas as a variação das RNAs quanto aos diferentes tipos de arquitetura e de dados de entrada para os subgrupos B-02i e B-02iii, as RNAs resultam em uma menor variabilidade na previsão dos valores de LL_p . Ou seja, essas conclusões são válidas para a seleção adequada das arquiteturas e dados de entrada, conforme citado nos itens 4.2.4 e 4.2.8.

Destaca-se ainda que para o subgrupo B-02iii a média dos erros RMSE e MAE e do σ apresentam grandes diferenças de valores quando comparadas as correlações de regressão e as RNAs. Assim, os erros médios RMSE variam de 11,99 para 3,32 e os erros MAE de 7,60 para 2,32 das RNAs para as correlações de regressão. O σ médio observado foi de 4,83 para as RNAs e de 25,75 para as correlações de regressão. Essas divergências dos resultados podem ser justificadas pelo número pequeno de dados utilizados para este subgrupo ($n=6$), tendo em vista a homogeneidade citada para este conjunto de amostras de solo.

É importante ressaltar que o tamanho das amostras dos subgrupos analisados ($n<34$) não é grande o suficiente para fornecer uma estimativa precisa da força de correlação, assim, estimativas da força de relação, como o R^2 podem variar muito. Embora o número de observações para os subgrupos B-02i e B-02ii seja adequado para representar a distribuição dos dados, uma melhor representação da distribuição de um histograma pode ser obtida com um número maior de observações. No caso do subgrupo B-02iii, o número pequeno de dados (6 amostras) requer cautela na interpretação dos resultados obtidos, uma vez que, com amostras pequenas, a exatidão do valor-p é sensível a erros de resíduos não-normais. Assim, é preciso destacar que este fator pode explicar a impossibilidade de validação das correlações embora os valores de R^2 obtidos e a homogeneidade citada para estes conjuntos de dados.

4.3 RESULTADOS PARA OS DADOS DO GRUPO B-03

4.3.1 Avaliação estatística dos dados do subgrupo B-03

Os dados do grupo B-03 foram submetidos a testes estatísticos para a identificação e posterior tratamento dos *outliers* conforme a metodologia citada no item 3.2. Foram identificados 35 *outliers* do total de amostras de solo selecionadas para o grupo B-03, resultando o grupo, portanto, em 55 amostras aprovadas para serem utilizadas para o estabelecimento das correlações empíricas e no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c, LL_p e IP são apresentados no Anexo I.

A análise estatística para a identificação dos *outliers* resultou na identificação dos dados publicados por Gronbech et al. (2011), que possuíam valores altos de IP, quando comparados aos demais dados publicados. Além disso, os maiores valores de LL_c e LL_p dos resultados de ensaio publicados por Clemente (2018) e Clemente et al. (2020), bem como os menores resultados de LL_c e LL_p publicados por Clemente et al (2020). A Tabela 69 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*. Cabe ressaltar, que o grupo B-03 apresentou um percentual de 70% dos dados identificados como *outliers*, um número muito superior aos resultados identificados para os demais grupos.

Tabela 69 - Descrição estatística dos dados investigados no subgrupo B-03.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	55	35	48,18	8,78	77,00	18,21	27,00	43,00	50,00	54,00	64,00
LL _p	55	35	51,02	8,09	65,43	15,85	31,00	47,00	53,00	58,00	63,00
IP	55	35	23,40	7,84	61,47	33,50	7,00	18,00	24,00	29,00	42,00

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

Fonte: Autor (2022)

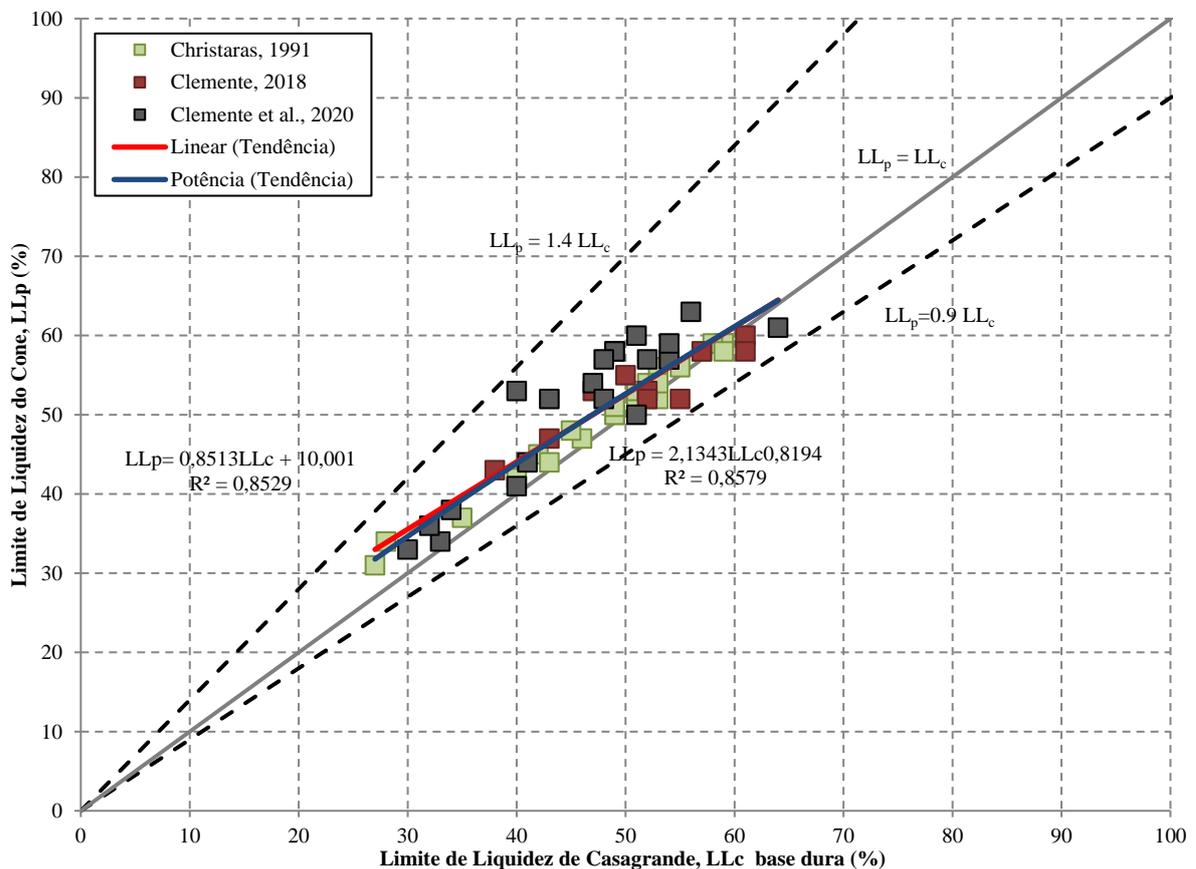
Os resultados apresentados na Tabela 69 foram comparados com a descrição estatística das 137 amostras designadas inicialmente para o grupo B-03 (Tabela 18). Foi observado, após a seleção e tratamento dos dados, uma redução de cerca de 82% nos valores de desvio padrão, 97% da variância e 64% do coeficiente de variação para as variáveis LL_c, LL_p e IP. Esses resultados indicam que o tratamento dos dados possibilitou a formação de um grupo com dados menos dispersos, mais próximos da média. O coeficiente de variação observado para as variáveis LL_c, LL_p e IP ficou abaixo de 35%, indicando que o grupo é composto de dados mais homogêneos.

A variabilidade dos dados do grupo B-03 é descrita no item 3.1.2.3. Na Tabela 19, o resumo das classificações SUCS das amostras indica que, embora os dados concentrem-se em sua maioria (cerca de 82%) entre argilas de baixa e alta compressibilidade (CL e CH), cerca de 18% das amostras são classificadas como siltes de baixa e alta compressibilidade (ML e MH).

O posicionamento das amostras no ábaco de Casagrande e conseqüentemente sua classificação SUCS, corroboram as observações feitas quanto a granulometria, já que o grupo contempla amostras de solos com frações de silte e areia. A mineralogia dos solos estudados não é informada nas publicações.

Conforme citado, os dados de valores mais altos de LL foram de uma maneira geral identificados como *outliers*, assim, o grupo B-03 passou a ser composto por valores de LL_c inferiores a 80. A Figura 72 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 72 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do grupo B-03 e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-03:

$$LL_p = 0,8513LL_c + 10,001 \quad R^2 = 0,8529 \quad (52)$$

$$LL_p = 2,1343LL_c^{0,8194} \quad R^2 = 0,8579 \quad (53)$$

Nota-se da Figura 72, que as amostras de solo do grupo B-03 encontram-se entre as linhas $LL_p = 1,4 LL_c$ e $LL_p = 0,9 LL_c$ e 80% das amostras apresenta resultados experimentais de LL_p maiores do que LL_c . As correlações de regressão se situam entre as linhas $LL_p=LL_c$ e $LL_p = 1,4 LL_c$. Observa-se que as correlações preveem resultados muito próximos, principalmente para a faixa de valores de LL_p superiores a 40. A Tabela 70 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-03.

Tabela 70 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03.

Regressão	RMSE	MAE	R ²	Mínimo	Máximo	Média	σ
Linear - RL	3,074	2,379	0,853	32,987	64,486	51,020	7,471
Não linear - RNL	3,032	2,327	0,858	31,777	64,452	50,943	7,735

Fonte: Autor (2022)

O resumo estatístico apresentado na Tabela 70 confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Observa-se também por meio do resumo estatístico, que o modelo não linear possui menor erro médio absoluto (MAE), raiz quadrada do erro médio (RMSE) e desvio padrão, e um maior coeficiente de determinação (R²), quando comparado com a regressão linear.

As correlações obtidas para o Grupo B-03 foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados destas comparações constam da Tabela 71.

Tabela 71 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o Grupo B-03.

LL _c	LL _p		
	Grupo B-03	Christaras (1991)	Gronbech et al. (2011)
	LL _c <80	27<LL _c <59	85<LL _c <200
30	35,54	34,03	37,90
40	44,05	42,73	47,40
50	52,57	51,43	56,90
60	61,08	60,13	66,40
70	69,59	-	75,90
80	78,11	-	85,40

Fonte: Autor (2022)

A correlação linear obtida para o grupo B-03 apresenta resultados mais próximos aos obtidos por Christaras (1991) para o intervalo de LL entre 27 e 59. Nota-se que quando comparado com os dados de Gronbech et al. (2011), as diferenças entre os resultados aumentam conforme o aumento de LL. Ressalta-se que, embora o Grupo B-03 com a eliminação dos

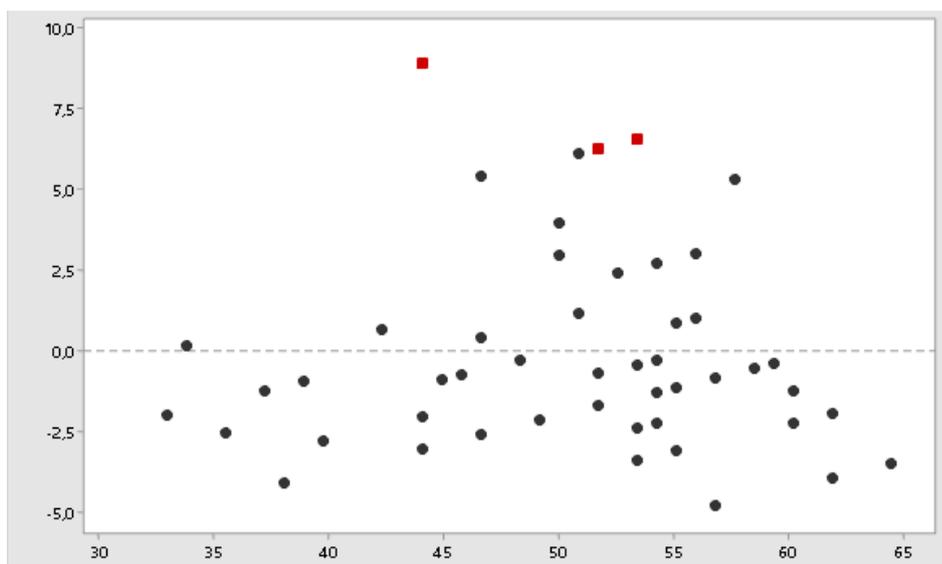
outliers tenha sido limitado ao grupo a valores de $LL_c < 80$, a comparação com os dados de Gronbech et al. (2011) mostra-se pertinente já que o objetivo do grupo de dados é correlacionar resultados de limites de liquidez obtidos pelos métodos de Casagrande de base dura e cone sueco de uma maneira geral.

Assim, embora as diferenças regionais observadas entre as amostras de dados utilizadas, nota-se que por incluírem os dados de Christaras (1991) e Gronbech et al. (2011), os resultados do grupo B-03 se mostram próximos aos publicados pelos pesquisadores. Os dados de Clemente (2018) e Clemente et al. (2020) não foram utilizados a título de comparação, já que as publicações não apresentaram correlações entre os valores de LL_c e LL_p analisados.

Vale ressaltar que não foram encontradas correlações não lineares entre os dados de limite de liquidez obtidos pelos métodos de Casagrande de base dura e de cone Sueco, motivo pelo qual a correlação obtida não pode ser comparada com resultados da literatura.

A correlação linear obtida para o grupo B-03 foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos apresentado na Figura 73.

Figura 73 – Relatório de diagnóstico de resíduos – Grupo B-03.



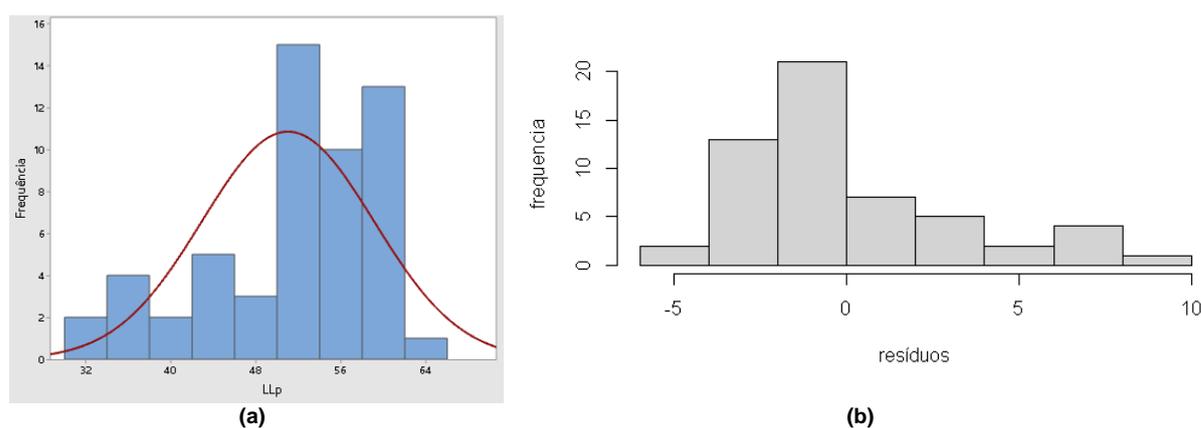
Fonte: Autor (2022)

No gráfico da Figura 73 é possível identificar que 4 pontos de dados possuem grandes resíduos, representados pelos pontos vermelhos da figura, não estando bem ajustados pelas equações propostas. Desse total, 2 possuem pontos atípicos de LL_c , o que pode fazer com

que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 74.

Figura 74 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Grupo B-03.



Fonte: Autor (2022)

Através da Figura 74 observa-se que a variável dependente apresenta um gráfico de distribuição com uma assimetria moderada à esquerda. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 72 apresenta o resumo dos resultados dos testes estatísticos realizados para o grupo B-03.

Tabela 72 – Resumo dos resultados dos testes estatísticos para o Grupo B-03.

Teste	valor-p	Conclusão
Teste t (intercepto)	1,00E-04	Parâmetro significativo
Teste t (x)	2,00E-16	Parâmetro significativo
Teste F	2,20E-16	Parâmetro significativo
Teste KS	5,73E-05	Distribuição dos resíduos não é normal
Teste DW	2,69E-03	Resíduos não são independentes

Fonte: Autor (2022)

Os resultados obtidos com os testes de hipóteses (Testes F e t) indicam que os parâmetros de regressão são significantes em relação aos dados observados, uma vez que o valor-p obtido para os testes é maior do que a significância de 5%.

O resultado do teste de aderência de *Kolmogorov Smirnov* (teste KS) aponta para a distribuição não normal dos resíduos, uma vez que o valor-p é inferior a significância de 5%.

Esse resultado vai de encontro à assimetria observada na análise dos resíduos por meio dos histogramas.

O resultado do teste de *Durbin Watson* (DW), que avalia a independência entre os resíduos, indica que os resíduos do grupo B-02 não são independentes, já que o valor-p é menor que o nível de significância de 5%. Além disso, o resultado do teste DW é de 1,2832, ou seja, por estar próximo de 0, indica que há autocorrelação positiva entre as variáveis independentes. Portanto, os resultados do teste DW não foram satisfatórios, apontando para a inconsistência do modelo.

Ressalta-se que, neste caso, em função do tamanho da amostra, o teste *Shapiro-Wilk* (SW) não foi utilizado.

Portanto, apesar de um R^2 significativo ($> 0,85$) para as regressões lineares e não lineares, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses, os resultados não satisfatórios dos testes estatísticos dos resíduos indicaram que a correlação linear apresentada não é significativa.

4.3.2 Aplicação das RNAs nas amostras do grupo B-03

Utilizando os dados filtrados nas análises estatísticas, as 55 amostras aprovadas no grupo B-03 foram submetidas as redes neurais adotadas artificiais. Um resumo do desempenho das RNAs treinadas para este grupo é apresentado na Tabela 73.

Tabela 73 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	5,018	4,041 \pm 2,975	0,914	33,300	56,827	45,759	7,800
RN1.2	5,770	4,694 \pm 3,355	0,911	32,460	55,461	44,811	7,806
RN1.3	5,097	4,106 \pm 3,020	0,913	33,351	56,794	45,738	7,648
RN1.4	4,142	3,326 \pm 2,469	0,909	34,715	58,873	47,242	7,711
RN1.5	4,599	3,610 \pm 2,849	0,913	33,382	57,398	46,307	7,920
RN1.6	3,808	3,323 \pm 1,861	0,916	35,179	58,539	48,040	7,525
RN1.7	4,798	3,829 \pm 2,891	0,911	33,762	57,114	46,107	7,779
RN1.8	4,888	3,927 \pm 2,910	0,913	33,595	57,170	45,980	7,747
RN1.9	4,131	3,308 \pm 2,475	0,917	34,642	57,725	47,179	7,646
RN1.10	4,499	3,566 \pm 2,743	0,916	33,919	57,435	46,559	7,667
RN1.11	4,771	3,809 \pm 2,873	0,912	33,783	57,297	46,163	7,741
RN1.12	4,924	3,961 \pm 2,926	0,915	33,505	56,883	45,967	7,629
RN2.1	4,829	3,893 \pm 2,857	0,915	34,266	56,558	46,307	7,319
RN2.2	4,782	3,877 \pm 2,800	0,915	34,422	56,595	46,430	7,245
RN2.3	5,292	4,313 \pm 3,066	0,915	33,741	55,836	45,705	7,193
RN2.4	5,053	4,095 \pm 2,960	0,916	33,886	56,068	46,003	7,251
RN2.5	5,186	4,217 \pm 3,019	0,916	33,805	55,935	45,847	7,200

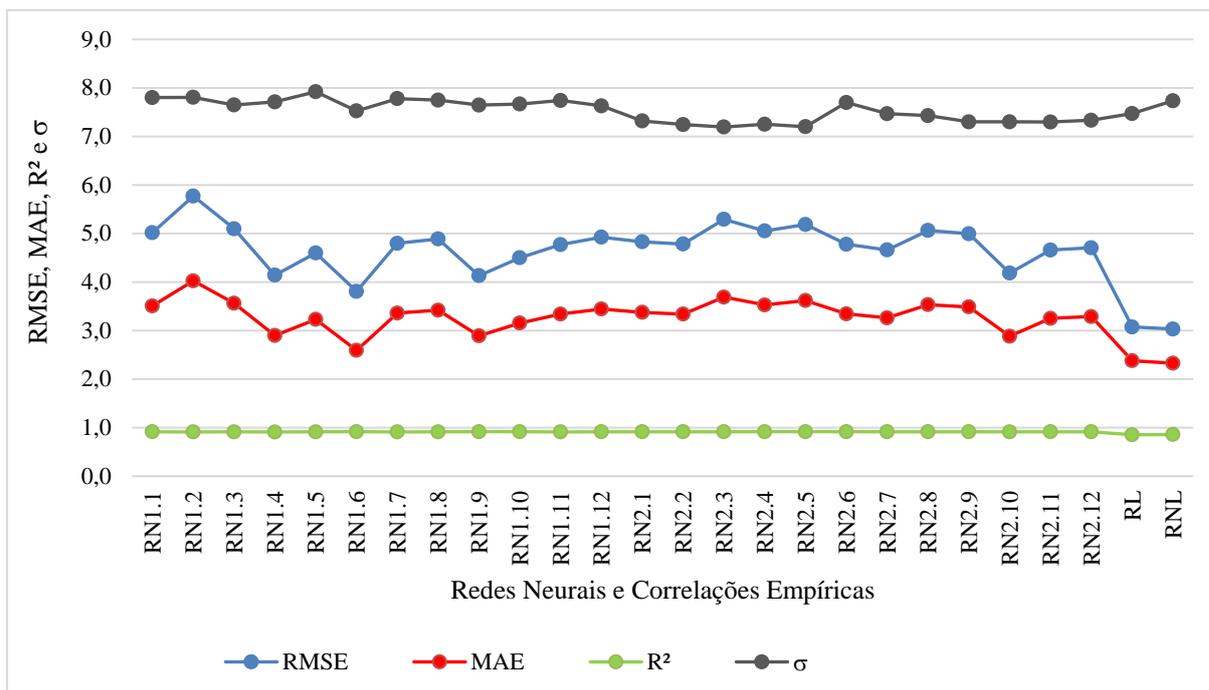
Rede	RMSE	MAE		R ²	Mínimo	Máximo	Média	σ
RN2.6	4,778	3,831	± 2,855	0,917	33,493	56,420	46,118	7,699
RN2.7	4,661	3,732	± 2,792	0,914	34,437	56,816	46,473	7,469
RN2.8	5,062	4,103	± 2,964	0,915	33,750	56,174	45,888	7,429
RN2.9	4,994	4,044	± 2,930	0,915	34,022	56,193	46,063	7,300
RN2.10	4,184	3,546	± 2,221	0,913	35,683	57,862	47,494	7,301
RN2.11	4,659	3,784	± 2,718	0,915	34,602	56,825	46,592	7,297
RN2.12	4,705	3,799	± 2,776	0,914	34,435	56,826	46,496	7,333

Fonte: Autor (2022)

Observa-se com as análises das RNAs que as arquiteturas utilizadas para cada conjunto de dados de entrada possuem resultados muito próximos. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) proporcionou um aumento da média dos erros MAE de 3,29 para 3,38, RMSE de 4,70 para 4,85 e R^2 de 0,91 para 0,92, e uma redução dos valores médios de σ de 7,72 para 7,32 em comparação com os obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Analisando, entretanto, os erros MAE e RMSE e o σ das redes RN1.1 a RN1.12, a arquitetura com 18 neurônios em uma única camada (RN1.6) e com 9 neurônios na primeira camada, foi a que apresentou menores valores (RMSE = 3,80, MAE = 2,59 e σ = 7,525).

Em comparação com os resultados obtidos pelas análises de regressão, se compararmos os critérios de coeficiente de determinação, R^2 , observa-se um resultado médio maior para as RNA ($R^2 = 0,91$). Os erros médios RMSE e MAE obtidos com as RNAs foram de 4,78 e 3,33 enquanto os obtidos com as correlações de regressão foram 3,05 e 2,35. No entanto, quando analisado desvio padrão (σ), observa-se que as RNAs apresentam resultados pouco menores do que as das correlações de regressão. O desvio padrão médio observado para as RNAs é de 7,53 enquanto o das correlações de regressão é de 7,60. A Figura 75 apresenta graficamente a comparação das RNAs e correlações de regressão.

Figura 75 - Resultados estatísticos das redes neurais e regressões do Grupo B-03.



Fonte: Autor (2022)

Da Figura 75 nota-se pequenas diferenças entre os modelos de correlação e RNAs, com destaque para os maiores valores de σ e menor R^2 das correlações de regressão. Portanto, os resultados indicam que as RNAs se adequam melhor a previsão de valores de LL_p para o grupo de dados B-03.

4.3.3 Avaliação estatística dos dados do subgrupo B-03i

Para os dados do subgrupo B-03i, aplicada a metodologia citada no item 3.2, não foram identificados *outliers*, portanto, 30 amostras foram utilizadas para o estabelecimento das correlações empíricas, bem como no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c , LL_p e IP são apresentados no Anexo I.

A Tabela 74 apresenta um resumo estatístico das amostras de solo investigadas.

Tabela 74 - Descrição estatística dos dados investigados no subgrupo B-03i.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL_c	30	0	40,23	8,62	74,25	21,42	18,00	33,75	42,50	48,00	50,00
LL_p	30	0	44,70	9,24	85,39	20,67	23,00	36,75	46,00	52,25	58,00
IP	30	0	17,63	5,71	32,59	32,37	7,00	13,00	19,00	22,25	27,00

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

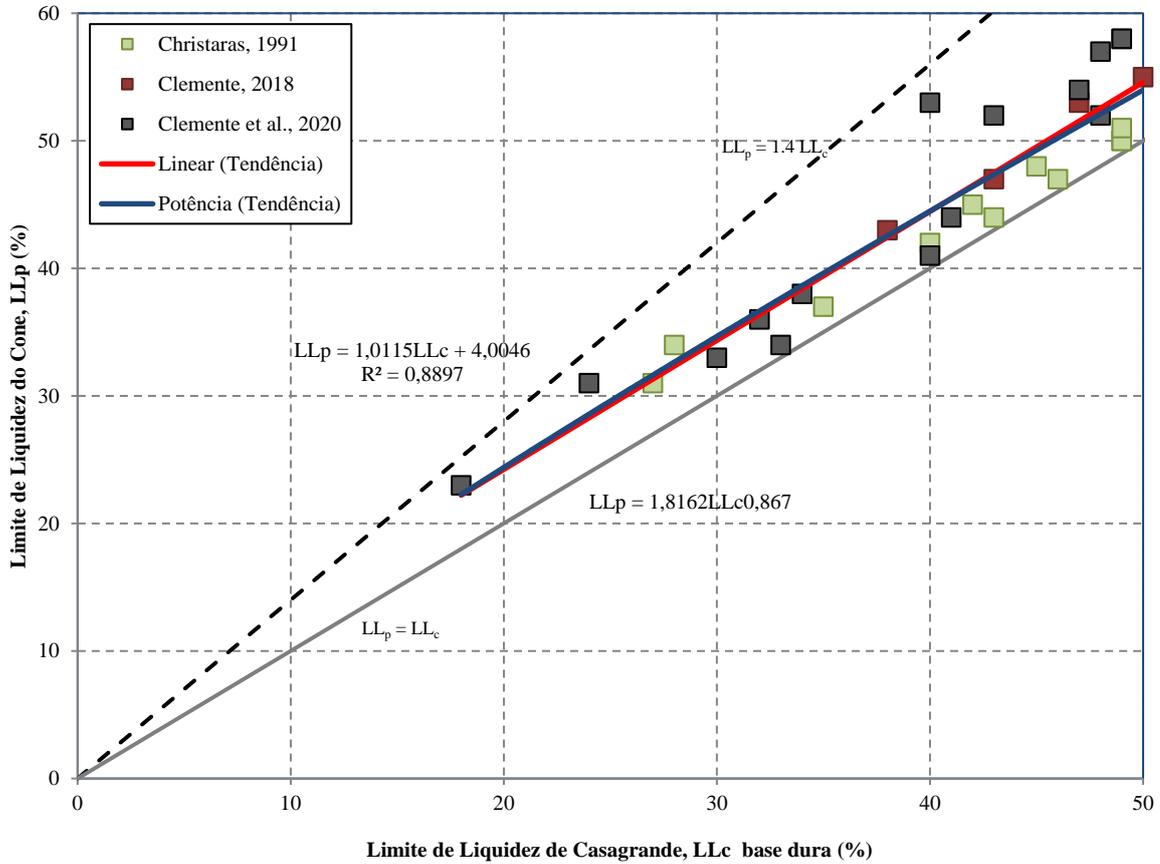
Fonte: Autor (2022)

Os resultados apresentados na Tabela 74, foram comparados com a descrição estatística do grupo B-03. Foi observado para o subgrupo B-03i, uma redução de cerca de 5% nos valores de desvio padrão, 6% da variância e um aumento de 9% do coeficiente de variação para as variáveis LL_c , LL_p e IP quando comparado com o grupo B-03. Essas diferenças foram menores se comparadas com os resultados observados para os grupos B-01 e B-02. Além disso, o aumento do coeficiente de variação observado para as variáveis LL_c , LL_p e IP, não indica que os dados deste subgrupo, após a eliminação dos *outliers* são mais homogêneos se comparado ao grupo B-03.

As classificações SUCS das amostras indica que, embora os dados concentrem-se em sua maioria (cerca de 84%) entre argilas de baixa compressibilidade (CL), cerca de 13% das amostras são classificadas como siltes de baixa compressibilidade (ML) e 3% de siltes de alta compressibilidade (MH). O posicionamento das amostras no ábaco de Casagrande e conseqüentemente sua classificação SUCS, corroboram as observações feitas quanto a granulometria, já que o grupo contempla amostras de solos com frações de silte e areia. A mineralogia dos solos estudados não é informada nas publicações.

A Figura 76 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 76 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03i e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-03i:

$$LL_p = 1,0115LL_c + 4,0046 \qquad R^2 = 0,8897 \qquad (54)$$

$$LL_p = 1,8162LL_c^{0,867} \qquad R^2 = 0,8876 \qquad (55)$$

Nota-se da Figura 76, que as correlações obtidas se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 1,4 LL_c$, bem como o conjunto de dados analisados, seguindo, portanto, a tendência observada para o grupo B-03. Observa-se também que para todas as amostras os resultados de LL_p são maiores que LL_c . As correlações de regressão descritas para o grupo B-03i apresentam resultados aproximados, principalmente para os valores de LL_p entre 40 e 45.

A Tabela 75 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-03i.

Tabela 75 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03i.

Regressão	RMSE	MAE	R²	Mínimo	Máximo	Média	σ
Linear - RL	3,018	2,428	0,890	22,212	54,580	44,701	8,716
Não linear - RNL	3,060	2,458	0,888	22,258	53,972	44,577	8,437

Fonte: Autor (2022)

O resumo estatístico confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Observa-se por meio do resumo estatístico, que o modelo linear possui valores de erros MAE, RMSE e coeficiente de determinação (R^2) praticamente coincidentes. Entretanto nota-se um menor σ obtido com as correlações não lineares.

As correlações obtidas para o subgrupo B-03i foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados destas comparações constam da Tabela 76.

Tabela 76 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-03i.

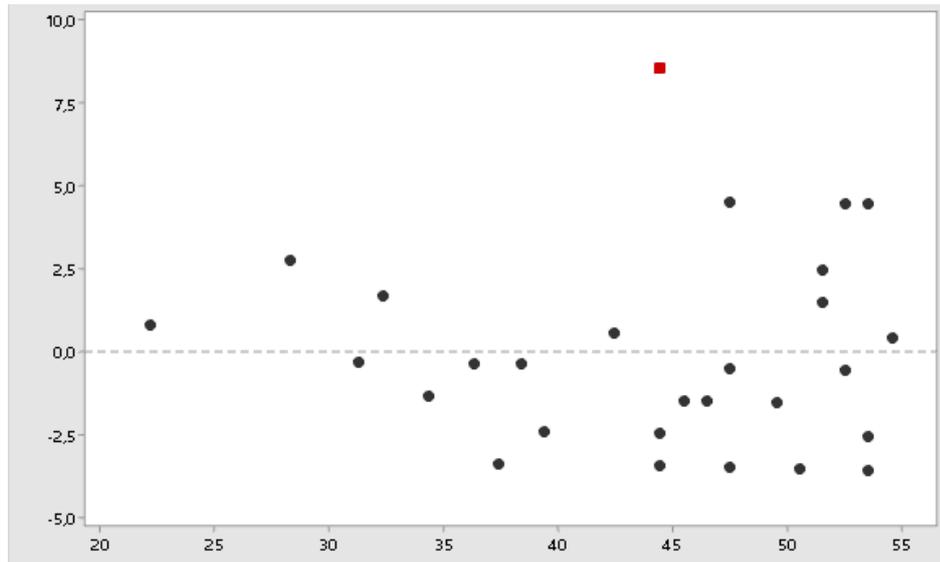
LL_c	LL_p	
	Subgrupo B-03i	Christaras (1991)
	LL_c<50	27<LL_c<59
30	34,35	34,03
40	44,46	42,73
50	54,58	51,43

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-03i apresenta resultados semelhantes aos observados para o grupo B-03. Assim, os resultados se mostraram próximos aos obtidos por Christaras (1991) para o intervalo de LL entre 27 e 59.

A correlação linear obtida para o subgrupo B-03i foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos da Figura 77.

Figura 77 - Relatório de diagnóstico de resíduos – Subgrupo B-03i.

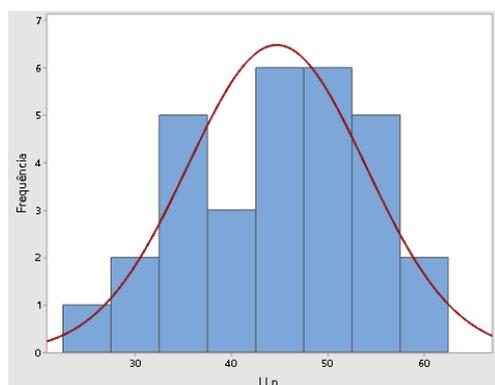


Fonte: Autor (2022)

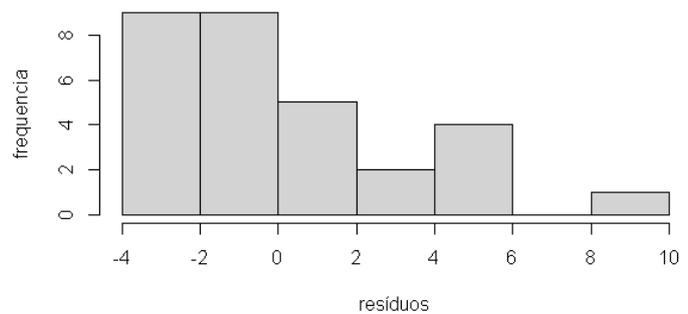
No gráfico da Figura 77 é possível identificar que um ponto de dados possui grandes resíduos, representado pelo ponto vermelho da figura, não estando, portanto, bem ajustado pelas equações propostas. Trata-se de um ponto atípico de LL_c , o que pode fazer com que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, que se apresentam distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 78.

Figura 78 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-03i.



(a)



(b)

Fonte: Autor (2022)

Através da Figura 78 observa-se que a variável dependente apresenta um gráfico de distribuição com uma assimetria moderada à esquerda. Logo, verifica-se que os resíduos também apresentam assimetria em sua distribuição.

A Tabela 77 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-03i.

Tabela 77 – Resumo dos resultados dos testes estatísticos para o subgrupo B-03i.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,159	Parâmetro não significante
Teste t (x)	6,24E-15	Parâmetro significante
Teste F	6,24E-15	Parâmetro significante
Teste KS	1,60E-03	Distribuição dos resíduos não é normal
Teste DW	9,81E-02	Resíduos independentes
Teste SW	0,01821	Distribuição dos resíduos não é normal

Fonte: Autor (2022)

Assim, os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados, embora o Teste t (intercepto) aponte para a insignificância do parâmetro. O resultado do teste de aderência de Kolmogorov Smirnov (teste KS) aponta para a distribuição não normal dos resíduos, em consonância com a assimetria observada na análise dos resíduos por meio dos histogramas e com o resultado do teste de Shapiro Wilk, aplicável para o número de observações entre 8 e 30, confirmando que a distribuição dos resíduos não é normal. O resultado do teste de Durbin Watson (DW), indica que os resíduos do subgrupo B-03i são independentes, e, o resultado do teste DW é de 1,5876, ou seja, por se aproximar de 2, indica que não há autocorrelação entre as variáveis independentes. Portanto, os resultados do teste DW para o subgrupo B-03i foram satisfatórios, apontando para a consistência do modelo.

Portanto, apesar de um R^2 significativo ($> 0,88$) para as regressões lineares e não lineares, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses e da independência dos resíduos observadas pelo teste DW, os resultados das análises de distribuição normal dos resíduos e a heterocedasticidade do modelo indicaram que a correlação linear apresentada não é significativa.

4.3.4 Aplicação das RNAs nas amostras do grupo subgrupo B-03i

Utilizando os dados filtrados nas análises estatísticas, as 30 amostras aprovadas no subgrupo B-03i foram submetidas as redes neurais artificiais. Um resumo do desempenho das RNAs é apresentado na Tabela 78.

Tabela 78 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-03i.

Rede	RMSE	MAE		R^2	Mínimo	Máximo	Média	σ
RN1.1	5,476	4,609	± 2,957	0,796	43,120	50,850	48,417	2,358
RN1.2	4,662	3,802	± 2,697	0,757	43,717	52,260	49,405	2,542
RN1.3	5,034	4,288	± 2,638	0,786	41,754	51,864	48,712	3,015
RN1.4	4,183	3,309	± 2,559	0,764	43,365	53,300	49,854	2,902
RN1.5	5,505	5,012	± 2,278	0,867	40,496	50,196	47,988	3,202
RN1.6	4,467	3,441	± 2,850	0,666	43,531	53,979	49,683	2,980
RN1.7	5,486	4,629	± 2,943	0,729	44,095	51,210	48,614	2,081
RN1.8	4,725	3,868	± 2,713	0,807	43,317	51,960	49,202	2,612
RN1.9	4,785	3,913	± 2,753	0,790	38,000	50,000	46,556	3,779
RN1.10	4,842	4,036	± 2,675	0,815	42,636	51,915	48,964	2,804
RN1.11	4,682	3,834	± 2,689	0,802	42,815	52,366	49,166	2,837
RN1.12	5,055	4,425	± 2,445	0,839	41,521	51,496	48,575	3,075
RN2.1	4,975	4,247	± 2,591	0,845	42,094	50,961	48,753	2,810
RN2.2	4,581	3,862	± 2,465	0,849	42,087	51,538	49,138	2,994
RN2.3	4,589	3,845	± 2,505	0,850	42,285	51,531	49,155	2,925
RN2.4	3,990	3,214	± 2,365	0,855	42,705	52,416	49,812	3,067
RN2.5	4,980	4,337	± 2,448	0,851	41,605	51,124	48,663	3,011
RN2.6	5,331	4,753	± 2,414	0,854	41,149	50,817	48,247	3,052
RN2.7	4,565	3,761	± 2,588	0,836	42,738	51,324	49,295	2,727
RN2.8	4,498	3,698	± 2,561	0,845	42,560	51,540	49,307	2,846
RN2.9	5,201	4,541	± 2,535	0,848	41,653	50,764	48,459	2,886
RN2.10	5,284	4,651	± 2,507	0,850	41,499	50,724	48,349	2,919
RN2.11	5,160	4,542	± 2,447	0,851	41,399	50,923	48,458	3,012
RN2.12	5,216	4,619	± 2,423	0,853	41,285	50,903	48,381	3,042

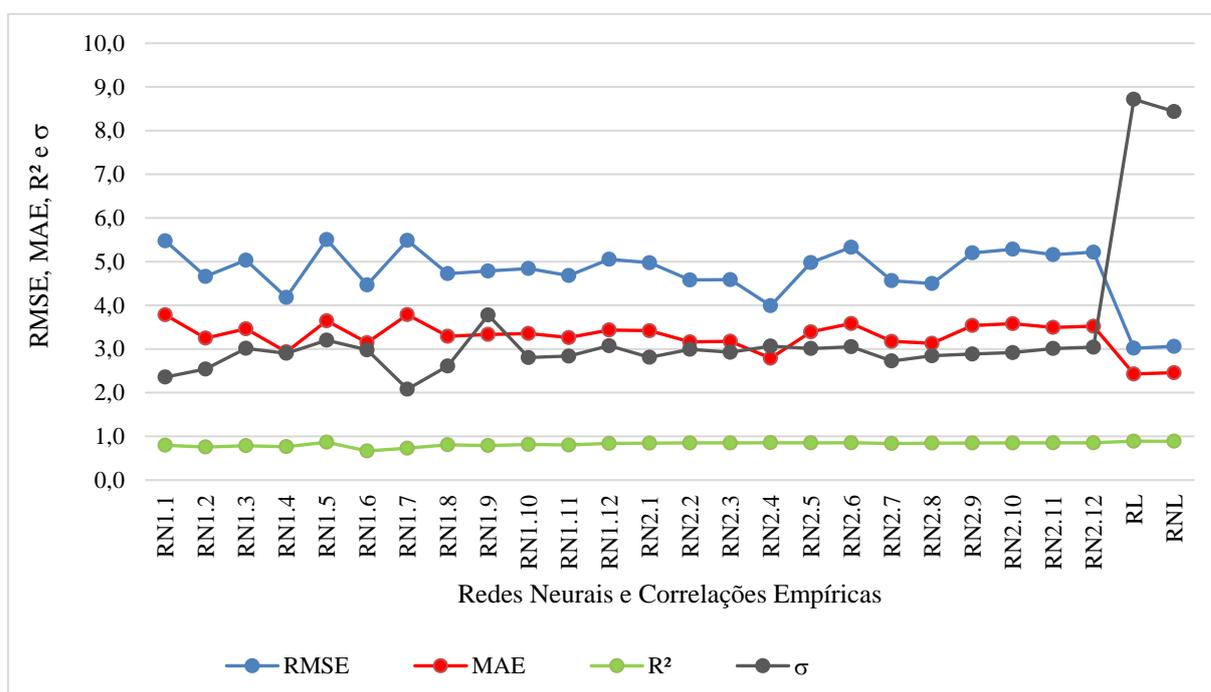
Fonte: Autor (2022)

Observa-se, conforme resultados do Grupo B-03, que as arquiteturas utilizadas, para cada conjunto de dados de entrada, possuem resultados muito próximos. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12), proporcionou uma diminuição da média dos erros MAE de 4,91 para 4,86 e RMSE de 3,39 para 3,33 e um aumento do valor médio de R^2 de 0,78 para 0,85, quando comparadas com os resultados obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Quanto as arquiteturas, os resultados obtidos com a utilização de apenas uma camada oculta apresentou uma maior diferença de valores, se comparado com as arquitetura com duas camadas ocultas. Entretanto, analisando os erros MAE e RMSE e o R^2 das redes RN2.1 a RN2.12, a arquitetura com 12 neurônios em uma única camada (RN2.4) apresentou menores valores para os erros RMSE e MAE (RMSE=3,99 e MAE = 2,79) um maior R^2 ($R^2 = 0,85$).

Em comparação com os resultados obtidos pelas análises de regressão, observa-se um resultado melhor para as equações de regressão quando analisados os critérios de erro e R^2 .

As correlações de regressão apresentaram erros médios RMSE e MAE de 3,04 e 2,44 enquanto para as RNAs esses valores foram de 4,89 e 3,36. O R^2 médio observado para as correlações de regressão foi de 0,89, enquanto o R^2 médio das RNAs foi de 0,82. Entretanto, as maiores diferenças foram observadas quando analisado o σ , com valores médios superiores das correlações de regressão. O σ médio foi de 8,58 para as correlações, enquanto para as RNAs esse valor foi de 2,89. A Figura 79 apresenta graficamente o desempenho das redes neurais em comparação com as regressões.

Figura 79 - Resultados estatísticos das redes neurais e regressões do subgrupo B-03i.



Fonte: Autor (2022)

Da Figura 79 nota-se a semelhança entre os modelos de correlação e RNAs obtidas com a utilização das arquiteturas com duas camadas ocultas (RN1.7 a RN1.12 e RN2.7 a RN2.12), com destaque para as pequenas diferenças entre os resultados. Uma exceção, no entanto, são os maiores valores de σ e os menores erros RMSE das correlações de regressão. Entretanto, os resultados indicam que as RNAs, quando adotadas as arquiteturas com duas camadas ocultas, se adequam melhor a previsão de valores de LL_p para o subgrupo de dados B-03i.

4.3.5 Avaliação estatística dos dados do subgrupo B-03ii

Para os dados do subgrupo B-03ii, aplicada a metodologia citada no item 3.2, foram identificados 17 *outliers* do total de amostras de solo selecionadas para o grupo B-03. Portanto, resultaram neste subgrupo 27 amostras aprovadas para serem utilizadas para o

estabelecimento das correlações empíricas, bem como no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c, LL_p e IP são apresentados no Anexo I.

Foram identificados como *outliers* os resultados dos ensaios publicados por Gronbech et al. (2011) e aqueles que apresentavam maior valor de LL entre os dados publicados por Clemente (2018). Além do mais, foram identificados dentre os resultados publicados por Clemente et al. (2020) aqueles que possuíam os maiores valores de LL e IP.

A Tabela 79 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 79 - Descrição estatística dos dados investigados no subgrupo B-03ii.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL _c	27	13	55,000	3,669	13,462	6,67	51,000	52,000	54,000	58,000	64,000
LL _p	27	13	56,259	3,415	11,661	6,07	50,000	53,000	57,000	59,000	63,000
IP	27	13	28,81	6,05	36,62	21,00	17,00	25,00	29,00	32,00	42,00

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

Fonte: Autor (2022)

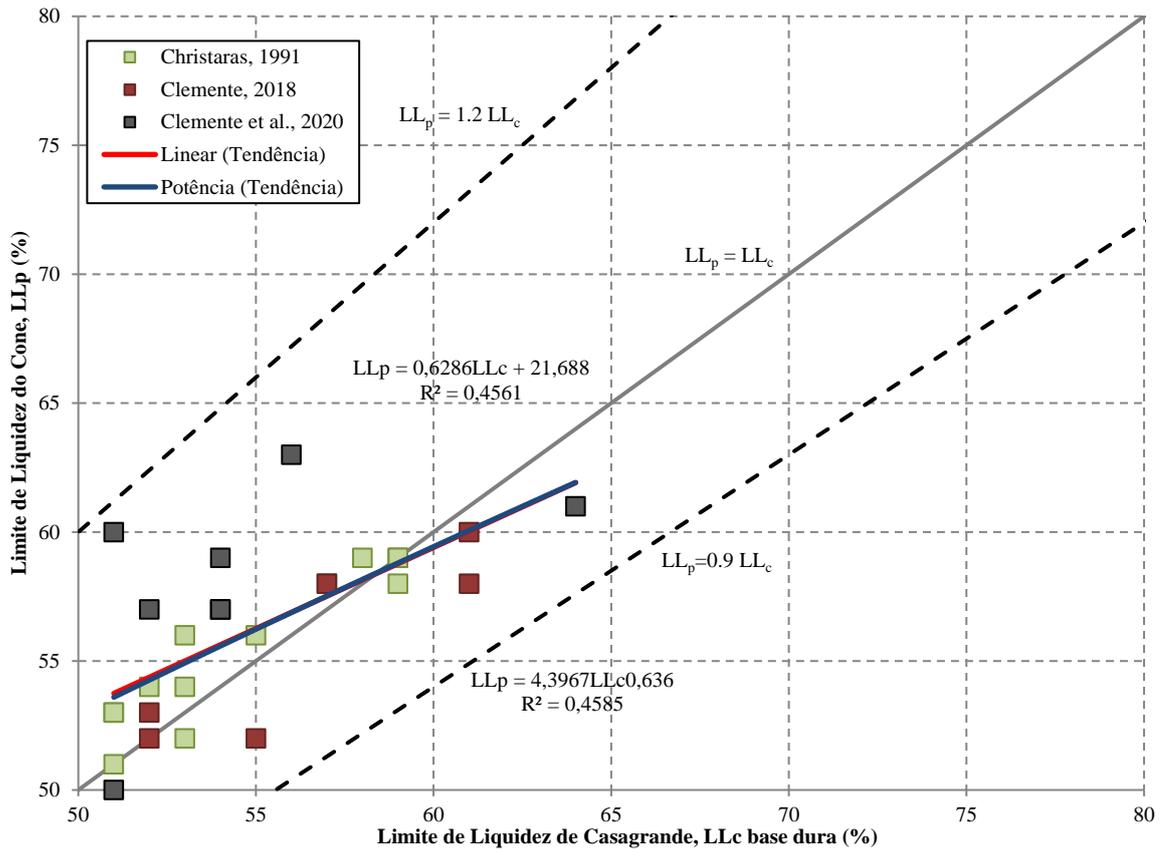
Os resultados apresentados na Tabela 79, quando comparados com a descrição estatística do grupo B-03, indicam que este grupo apresenta dados menos dispersos. Foi observado para o subgrupo B-03ii, uma redução de cerca de 47% nos valores de desvio padrão, 70% da variância e 50% do coeficiente de variação para as variáveis LL_c, LL_p e IP quando comparado com o grupo B-03. O coeficiente de variação observado para as variáveis LL_c, LL_p e IP ficou abaixo de 21%, indicando que este subgrupo é composto por dados mais homogêneos se comparado ao grupo B-03.

Os dados do subgrupo B-03i, em função do intervalo de LL em que se estão inseridos, são classificados como argilas e siltes de alta compressibilidade (CH e MH). Embora mais de 81% dos dados sejam descritos como amostras de argila (CH), a granulometria das amostras selecionadas para este subgrupo, é caracterizada por parcelas de siltes e areias, justificando, portanto, a variação observada quanto a classificação das amostras. A mineralogia das amostras de solo não é informada pelas publicações.

Conforme observado na Tabela 79, apesar do subgrupo B-03ii ser composto inicialmente por resultados de LL_c entre 50 e 120, após a filtragem dos dados, com a eliminação dos *outliers*,

o valor de LL_c para o subgrupo é inferior a 80. A Figura 80 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 80 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03ii e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-03ii:

$$LL_p = 0,6286LL_c + 21,688 \quad R^2 = 0,4561 \quad (56)$$

$$LL_p = 4,3867LL_c^{0,636} \quad R^2 = 0,4585 \quad (57)$$

Nota-se da Figura 80 que as amostras de solo analisadas se encontram entre as linhas $LL_p = 1,2 LL_c$ e $LL_p = 0,9 LL_c$ e cerca de 59% dos resultados de LL_p observados são maiores que LL_c . As correlações de correlação descritas para o grupo B-03ii se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 1,2 LL_c$, para valores de LL_c inferiores a 58, e entre as linhas $LL_p=LL_c$ e $LL_p = 0,9 LL_c$ para valores superiores a 58. Observa-se que estas correlações apresentam

resultados aproximados, e para valores de LL_p acima de 57 os resultados são praticamente coincidentes.

O R^2 obtido para as correlações lineares e não lineares se mostrou abaixo de 0,46. Ou seja, os modelos explicam apenas cerca de 46% da variação de LL_p . No gráfico da Figura 80 isso pode ser observado pelo afastamento dos pontos de dados das linhas de regressão.

A Tabela 80 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-03ii.

Tabela 80 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03ii.

Regressão	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
Linear - RL	2,471	1,806	0,456	53,747	61,918	56,261	2,306
Não linear - RNL	2,471	1,774	0,459	53,499	61,811	56,104	2,362

Fonte: Autor (2022)

O resumo estatístico confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Observa-se por meio do resumo estatístico, que os dois modelos de regressão apresentados possuem o mesmo erro RMSE e resultados muito próximos de erro MAE, R^2 e σ .

As correlações obtidas para o subgrupo B-03ii foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados destas comparações constam da Tabela 81.

Tabela 81 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-03ii.

LL_c	LL_p		
	Subgrupo B-03ii	Christaras (1991)	Gronbech et al. (2011)
	50< LL_c <120	27< LL_c <59	85< LL_c <200
50	53,12	51,43	56,90
60	59,40	60,13	66,40
70	65,69	-	75,90
80	71,98	-	85,40
90	78,26	-	94,90
100	84,55	-	104,40
110	90,83	-	113,90
120	97,12	-	123,40

Fonte: Autor (2022)

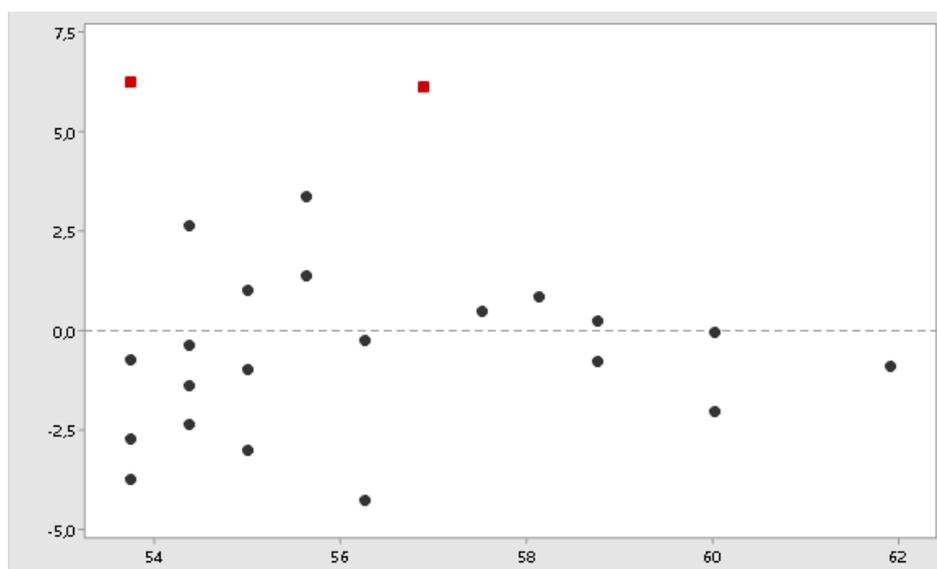
A correlação linear obtida para o subgrupo B-03ii segue a tendência observada para o grupo B-03 em geral, e apresenta resultados mais próximos aos obtidos por Christaras (1991) para o intervalo de LL entre 27 e 59. Nota-se que quando comparado com os dados de Gronbech

et al. (2011), as diferenças entre os resultados aumentam conforme o aumento de LL, além disso destaca-se que os dados do subgrupo B-03ii não inclui os dados publicados por Gronbech et al. (2011).

Vale ressaltar que não foram encontradas correlações não lineares entre os dados de limite de liquidez obtidos pelos métodos de Casagrande de base dura e de cone Sueco, motivo pelo qual a correlação obtida não pode ser comparada com resultados da literatura.

A correlação linear obtida para o subgrupo B-03ii foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos da Figura 81.

Figura 81 - Relatório de diagnóstico de resíduos – Subgrupo B-03ii.

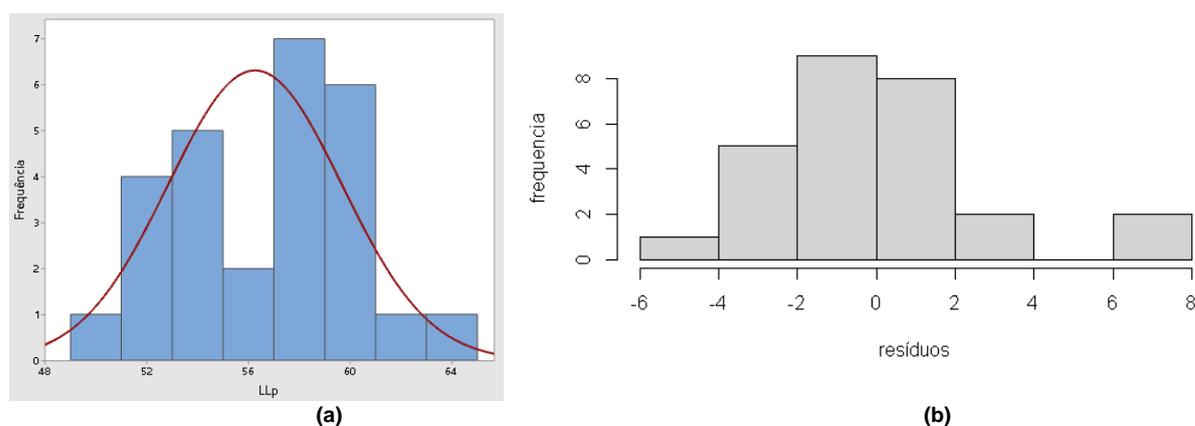


Fonte: Autor (2022)

No gráfico da Figura 81 é possível identificar que 2 pontos de dados possuem grandes resíduos, representados pelos pontos vermelhos da figura, não estando bem ajustados pelas equações propostas. Desse total, 1 trata-se de ponto atípico de LL_c , o que pode fazer com que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, se apresentam distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 82.

Figura 82 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-03ii.



Fonte: Autor (2022)

Através da Figura 82 observa-se que a variável dependente apresenta um gráfico de distribuição normal. Entretanto, verifica-se que os resíduos apresentam assimetria em sua distribuição.

A Tabela 82 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-03ii.

Tabela 82 – Resumo dos resultados dos testes estatísticos para o subgrupo B-03ii.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,00831	Parâmetro significativo
Teste t (x)	1,11E-04	Parâmetro significativo
Teste F	1,11E-04	Parâmetro significativo
Teste KS	1,60E-01	Distribuição Normal dos resíduos
Teste DW	1,09E-01	Resíduos independentes
Teste SW	0,0866	Distribuição Normal dos resíduos

Fonte: Autor (2022)

Os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados. O resultado do teste de aderência de Kolmogorov Smirnov (teste KS) aponta para a distribuição normal dos resíduos, em consonância com a simetria observada na análise dos resíduos por meio dos histogramas e com o resultado do teste de Shapiro Wilk, aplicável para o número de observações entre 8 e 30, confirmando que a distribuição dos resíduos é normal. O resultado do teste de Durbin Watson (DW), indica que os resíduos do subgrupo B-03ii são independentes, e, o resultado do teste DW é de 1,5808, ou seja, por se aproximar de 2, indica que não há autocorrelação entre as variáveis independentes. Portanto, os resultados do teste DW para o subgrupo B-03ii foram satisfatórios, apontando para a consistência do modelo.

Portanto, embora a distribuição normal dos resíduos, devido a heterocedasticidade do modelo, a correlação linear apresentada não é significativa.

4.3.6 Aplicação das RNAs nas amostras do grupo subgrupo B-03ii

Utilizando os dados filtrados nas análises estatísticas, as 27 amostras aprovadas no subgrupo B-03ii foram submetidas as redes neurais artificiais. Um resumo do desempenho das RNAs é apresentado na Tabela 83.

Tabela 83 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-03ii.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	6,719	5,541 ± 3,801	0,000	53,620	67,109	60,727	4,251
RN1.2	7,157	5,863 ± 4,105	0,000	53,494	68,304	61,039	4,641
RN1.3	8,780	7,091 ± 5,177	0,000	53,650	71,917	62,238	5,912
RN1.4	8,296	6,776 ± 4,788	0,000	53,509	70,943	61,944	5,557
RN1.5	8,246	6,610 ± 4,929	0,000	53,446	70,115	61,832	5,428
RN1.6	9,159	7,368 ± 5,441	0,000	53,812	72,599	62,631	6,107
RN1.7	4,164	3,220 ± 2,640	0,077	53,948	59,571	58,538	1,922
RN1.8	5,503	4,391 ± 3,316	0,000	53,359	63,493	59,405	3,180
RN1.9	6,086	4,871 ± 3,649	0,000	53,469	64,512	60,011	3,558
RN1.10	6,498	5,116 ± 4,005	0,000	53,493	65,505	60,584	3,944
RN1.11	7,002	5,620 ± 4,177	0,000	53,561	66,966	60,905	4,338
RN1.12	7,137	5,764 ± 4,208	0,000	53,823	67,252	61,198	4,343
RN2.1	2,884	2,262 ± 1,789	0,573	54,282	59,044	57,318	1,980
RN2.2	2,940	2,292 ± 1,842	0,565	54,158	58,455	57,073	1,856
RN2.3	2,923	2,354 ± 1,732	0,567	54,408	58,899	57,449	1,938
RN2.4	2,982	2,390 ± 1,784	0,548	54,255	58,514	57,120	1,820
RN2.5	3,090	2,423 ± 1,918	0,500	54,055	58,230	56,854	1,758
RN2.6	2,849	2,154 ± 1,866	0,562	54,351	59,933	57,423	2,160
RN2.7	2,888	2,279 ± 1,774	0,574	54,406	59,204	57,421	1,988
RN2.8	2,845	2,195 ± 1,810	0,573	54,338	59,568	57,431	2,116
RN2.9	2,811	2,110 ± 1,857	0,566	54,201	59,854	57,422	2,260
RN2.10	2,830	2,111 ± 1,884	0,555	54,306	60,133	57,445	2,279
RN2.11	2,869	2,057 ± 1,999	0,531	54,430	61,108	57,748	2,561
RN2.12	2,957	2,314 ± 1,840	0,529	54,700	61,538	58,093	2,625

Fonte: Autor (2022)

Observa-se que os resultados obtidos foram semelhantes entre as diferentes arquiteturas testadas, no entanto, variaram para cada conjunto de dados de entrada. A utilização apenas dos dados de entrada LL_c e LL_p (RN 2.1 a RN 2.12), proporcionou uma diminuição da média dos erros MAE de 4,94 para 2,04, RMSE de 7,06 para 2,91, e um aumento do valor médio de R^2 de 0,01 para 0,55 em comparação com os obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Os resultados obtidos com as redes RN1.1 a RN1.12

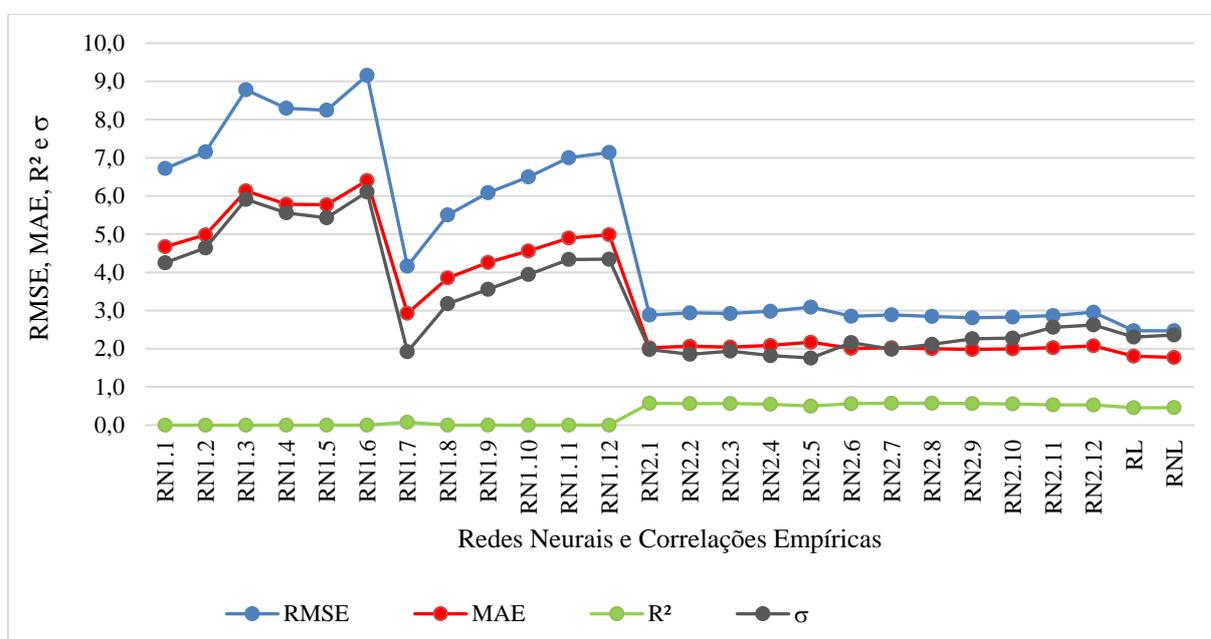
foram inconsistentes, com valores de R^2 iguais a zero para a maioria das arquiteturas adotadas.

Portanto, analisando os erros MAE e RMSE das redes RN2.1 a RN2.12, a arquitetura com 9 neurônios na primeira camada e 4 na segunda (RN2.9) mostraram menores resultados (RMSE=2,81 e MAE=1,98). Entretanto, a arquitetura com 3 neurônios na primeira camada e 1 (RN2.7) na segunda apresentou melhor R^2 ($R^2=0,57$). Entretanto, conforme supracitado, as diferenças observadas entre as arquiteturas foram pequenas.

A utilização apenas dos dados de entrada LL_c e LL_p (RN 2.1 a RN 2.12) apresentou resultados muito próximos aos obtidos com as análises de regressão. Os erros médios RMSE e MAE obtidos das RNAs foram de 2,91 e 2,04 enquanto o obtido com as correlações foram de 2,47 e 1,79 respectivamente. O R^2 médio verificado para as RNAs foi de 0,55 enquanto o das correlações de regressão foi de 0,46. Menores diferenças foram observadas quanto ao σ , já que para as RNAs o valor médio obtido foi de 2,11 e para as correlações de regressão foi de 2,33.

A Figura 83 apresenta graficamente o desempenho das redes neurais em comparação com as regressões.

Figura 83 - Resultados estatísticos das redes neurais e regressões do subgrupo B-03ii.



Fonte: Autor (2022)

Da Figura 83 nota-se a semelhança citada entre os modelos de correlação e RNAs obtidas com a utilização dos dados de entrada LL_c e LL_p , (RN2.1 a RN2.12). Uma exceção, no entanto, são os maiores valores de σ das correlações de regressão e o menor valor de R^2 . Portanto,

os resultados indicam que as RNAs, quando adotados os dados de entrada LL_c e LL_p , se adequam melhor a previsão de valores de LL_p para o subgrupo de dados B-03ii.

4.3.7 Avaliação estatística dos dados do subgrupo B-03iii

Para os dados do subgrupo B-03iii, aplicada a metodologia citada no item 3.2, não foram identificados *outliers*, portanto, 16 amostras foram utilizadas para o estabelecimento das correlações empíricas, bem como no treinamento das redes neurais artificiais. Os gráficos de *outliers* e o diagramas de *boxplot* obtidos para os dados de LL_c , LL_p e IP são apresentados no Anexo I.

A Tabela 84 apresenta um resumo estatístico das amostras de solo investigadas após a remoção dos *outliers*.

Tabela 84 - Descrição estatística dos dados investigados no subgrupo B-03iii.

Variável	N	N*	Média	DesvPad	Variância	CoefVar	Mínimo	Q1	Mediana	Q3	Máximo
LL_c	16	0	200,9	61,5	3780,1	30,61	120,8	158,4	174,4	256,0	344,9
LL_p	16	0	189,3	54,8	3006,4	28,96	114,4	156,5	161,8	242,9	311,1
IP	16	0	159,4	54,9	3018,6	34,47	85,5	125,5	135,2	210,5	284,7

Legenda: N = número de amostras; N* = número de *outliers*; DesvPad = desvio padrão; CoefVar = Coeficiente de variação; Q1 = primeiro quartil; Q3 = terceiro quartil.

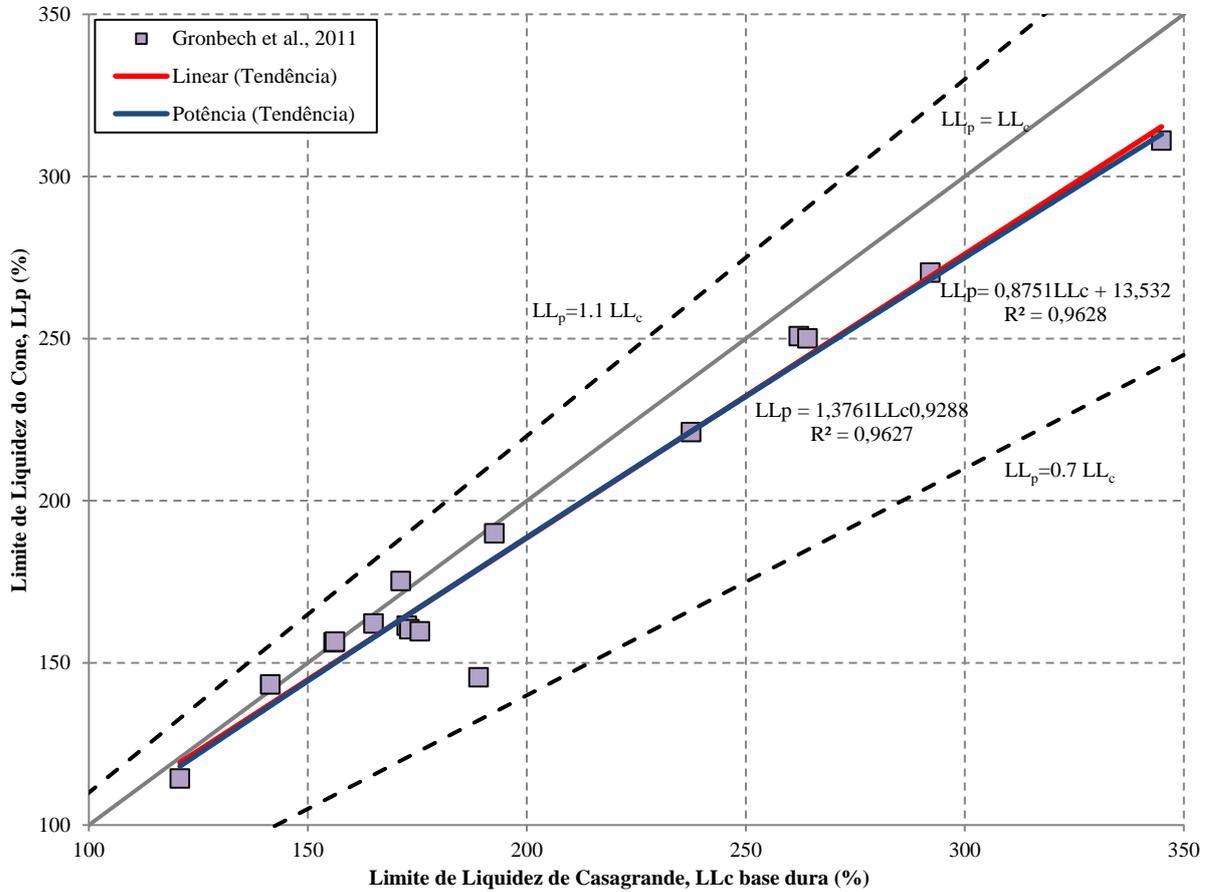
Fonte: Autor (2022)

Os resultados apresentados na Tabela 84, quando comparados com a descrição estatística do grupo B-03, indicam um aumento de cerca de 86% nos valores de desvio padrão, 98% da variância e 28% do coeficiente de variação para as variáveis LL_c , LL_p e IP. O coeficiente de variação observado para as variáveis LL_c , LL_p e IP ficou acima de 30%. Portanto, embora o tratamento dos dados, não houve redução do desvio padrão, variância e coeficiente de variação em relação ao grupo B-03. Nota-se, portanto, uma maior variabilidade das amostras de solo para este conjunto de dados.

Quanto a classificação dos solos, não é observada variabilidade entre as amostras. Os dados do subgrupo B-03iii, em função do intervalo de LL em que se estão inseridos, são classificados como argilas de alta compressibilidade (CH). A mineralogia não é informada, entretanto, as amostras selecionadas para este grupo são compostas apenas dos dados publicados por Gronbech et al. (2011).

A Figura 84 apresenta os pares de dados e as correlações avaliadas para os dados filtrados.

Figura 84 - Comparações entre os valores de LL obtidos pelo método do cone sueco e pelo método de Casagrande de base dura para os dados compilados do subgrupo B-03iii e as correlações linear e não linear definidas neste trabalho.



Fonte: Autor (2022)

Assim, por meio da análise de regressão foram obtidas duas funções, uma linear e outra não linear, para os dados do grupo B-03iii:

$$LL_p = 0,8751LL_c + 13,532 \qquad R^2 = 0,9628 \qquad (58)$$

$$LL_p = 1,3761LL_c^{0,9288} \qquad R^2 = 0,9627 \qquad (59)$$

Nota-se da Figura 84 que as amostras de solo analisadas se encontram entre as linhas $LL_p = 0,7 LL_c$ e $LL_p = 1,1 LL_c$ e 75% dos resultados de LL_p observados são menores que LL_c . As correlações de regressão obtidas se encontram entre as linhas $LL_p=LL_c$ e $LL_p = 0,7 LL_c$ e apresentam resultados aproximados. Para os valores de LL_p entre 155 e 175 os resultados são praticamente coincidentes. Nota-se, portanto, que os dados e correlações sugeridas também seguem a tendência observada para os dados do grupo B-03.

A Tabela 85 apresenta um resumo estatístico dos dados previstos a partir de cada uma das correlações obtidas para o grupo B-03iii.

Tabela 85 - Resumo estatístico das estimativas de LL_p para as amostras do grupo B-03iii.

Regressão	RMSE	MAE	R²	Mínimo	Máximo	Média	σ
Linear – RL	10,234	7,217	0,963	119,244	315,354	189,329	53,803
Não linear - RNL	10,259	7,165	0,963	118,161	313,083	188,995	53,473

Fonte: Autor (2022)

O resumo estatístico confirma a proximidade dos resultados obtidos para os métodos de regressão linear e não linear. Ambos possuem o coeficiente de determinação (R^2) e os resultados obtidos para os erros RMSE, MAE e R^2 apresentam pequenas diferenças.

As correlações obtidas para o subgrupo B-03iii foram ainda comparadas com correlações obtidas da literatura, conforme citado no item 3.2. Os resultados destas comparações constam da Tabela 86.

Tabela 86 – Comparação entre os resultados previstos de LL_p entre correlações lineares da literatura e a obtida para o subgrupo B-03iii.

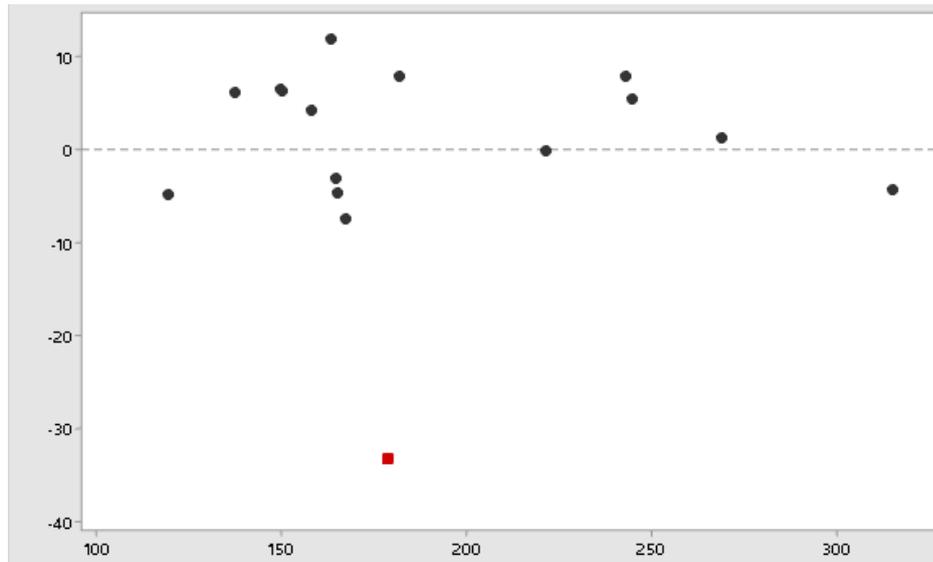
LL_c	LL_p	
	Grupo B-03	Gronbech et al. (2011)
	LL_c>120	85<LL_c<200
120	118,54	123,40
150	144,80	151,90
200	188,55	199,40

Fonte: Autor (2022)

A correlação linear obtida para o subgrupo B-03iii segue a tendência observada para o Grupo B-03 em geral, e quando comparada com os dados de Gronbech et al. (2011), as diferenças entre os resultados aumentam conforme o aumento de LL. Vale destacar que a correlação obtida para o subgrupo B-03iii contém apenas os dados de Gronbech et al. (2011) para o intervalo de dados analisados, ou seja, das 33 amostras analisadas pelos pesquisadores, apenas 16 foram utilizadas para a obtenção da presente correlação.

A correlação linear obtida para o subgrupo B-03iii foi submetida a avaliação dos resíduos e as verificações da normalidade de distribuição e da independência dos valores residuais. Assim, a homocedasticidade do modelo pode ser verificada através do gráfico de resíduos da Figura 85.

Figura 85 - Relatório de diagnóstico de resíduos – Subgrupo B-03iii.

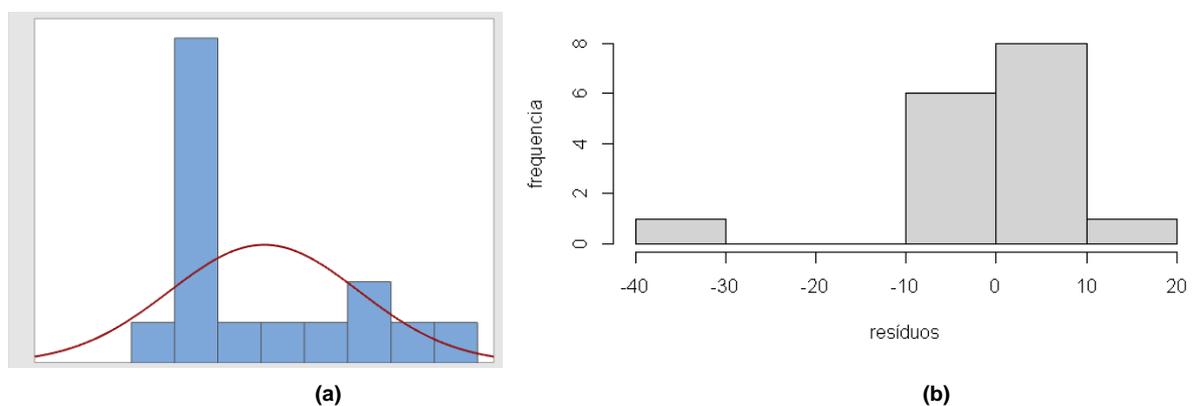


Fonte: Autor (2022)

No gráfico da Figura 85 é possível identificar que um ponto de dados possui grandes resíduos, representado pelo ponto vermelho da figura, não estando, portanto, bem ajustado pelas equações propostas. Trata-se de um ponto atípico de LL_c , o que pode fazer com que a linha ajustada seja puxada para mais perto dos valores atípicos e afastada dos outros pontos. Portanto, devido aos grandes resíduos identificados, e ao comportamento dos dados no gráfico de resíduos, que se apresentam distribuídos em uma faixa horizontal de alta amplitude, observa-se uma tendência de heterocedasticidade (variação não constante) do modelo.

Seguindo com a análise dos resíduos, o histograma de distribuição dos valores residuais é utilizado para identificação da distribuição normal, conforme apresentado na Figura 86.

Figura 86 – Análise dos resíduos: Histograma de LL_p (a); Histograma dos resíduos (b) – Subgrupo B-03iii.



Fonte: Autor (2022)

Através da Figura 86 observa-se que a variável dependente apresenta um gráfico de distribuição uniforme. Verifica-se ainda que os resíduos apresentam assimetria em sua distribuição.

A Tabela 87 apresenta o resumo dos resultados dos testes estatísticos realizados para o subgrupo B-03iii.

Tabela 87 – Resumo dos resultados dos testes estatísticos para o subgrupo B-03iii.

Teste	valor-p	Conclusão
Teste t (intercepto)	0,182	Parâmetro não significante
Teste t (x)	2,08E-11	Parâmetro significante
Teste F	2,09E-11	Parâmetro significante
Teste KS	0,00032	Distribuição dos resíduos não é normal
Teste DW	0,227	Resíduos independentes
Teste SW	0,00151	Distribuição dos resíduos não é normal

Fonte: Autor (2022)

Os resultados obtidos com os testes de hipóteses (Testes F e t), indicam que os parâmetros de regressão são significantes em relação aos dados observados, embora o Teste t (intercepto) aponte para a insignificância do parâmetro. O resultado do teste de aderência de Kolmogorov Smirnov (teste KS) aponta para a distribuição não normal dos resíduos, em consonância com a assimetria observada na análise dos resíduos por meio dos histogramas e com o resultado do teste de Shapiro Wilk, aplicável para o número de observações entre 8 e 30, confirmando que a distribuição dos resíduos não é normal. O resultado do teste de Durbin Watson (DW), indica que os resíduos do subgrupo B-03iii são independentes, e, o resultado do teste DW é de 1,7769, ou seja, por se aproximar de 2, indica que não há autocorrelação entre as variáveis independentes. Portanto, os resultados do teste DW para o subgrupo B-03iii foram satisfatórios, apontando para a consistência do modelo.

Portanto, apesar de um R^2 alto ($> 0,96$) para as regressões lineares e não lineares, e apesar da significância dos parâmetros de regressão observados com os testes de hipóteses e da independência dos resíduos observadas pelo teste DW, os resultados das análises de distribuição normal dos resíduos e a heterocedasticidade do modelo indicaram que a correlação linear apresentada não é significativa.

4.3.8 Aplicação das RNAs nas amostras do grupo subgrupo B-03iii

Utilizando os dados filtrados nas análises estatísticas, as 16 amostras aprovadas no subgrupo B-03iii foram submetidas as redes neurais artificiais. Um resumo do desempenho das RNAs é apresentado na Tabela 88.

Tabela 88 - Resumo estatístico das estimativas de LL_p para as amostras do subgrupo B-03iii.

Rede	RMSE	MAE	R^2	Mínimo	Máximo	Média	σ
RN1.1	123,956	118,533 ± 36,261	0,000	134,720	157,380	142,187	8,831
RN1.2	118,972	113,889 ± 34,403	0,000	141,594	159,589	146,831	7,330
RN1.3	118,253	113,170 ± 34,299	0,000	142,486	159,975	147,550	7,139
RN1.4	116,572	111,443 ± 34,196	0,000	144,100	161,615	149,277	7,096
RN1.5	131,721	125,498 ± 40,007	0,000	120,776	154,588	135,222	12,237
RN1.6	115,126	110,131 ± 33,541	0,000	145,217	162,050	150,589	6,630
RN1.7	108,017	103,860 ± 29,681	0,986	156,860	156,861	158,860	0,000
RN1.8	86,273	81,167 ± 29,238	0,793	178,638	180,083	179,553	0,557
RN1.9	106,570	102,447 ± 29,357	0,996	157,863	158,830	158,273	0,364
RN1.10	84,028	78,984 ± 28,676	0,884	179,884	183,081	181,736	1,199
RN1.11	88,606	83,505 ± 29,631	0,032	176,741	177,489	177,215	0,322
RN1.12	79,803	74,871 ± 27,618	0,946	182,518	188,842	185,849	2,377
RN2.1	92,443	87,749 ± 29,111	0,928	172,021	173,782	172,980	0,662
RN2.2	93,449	88,796 ± 29,120	0,990	172,127	172,821	171,924	0,631
RN2.3	89,199	84,677 ± 28,040	0,994	173,817	178,733	176,043	1,840
RN2.4	81,961	77,219 ± 27,475	0,992	180,431	187,067	183,501	2,477
RN2.5	87,799	83,281 ± 27,802	0,995	174,926	180,545	177,439	2,107
RN2.6	84,344	80,231 ± 26,017	0,994	175,851	186,721	180,489	4,111
RN2.7	107,626	103,457 ± 29,665	0,988	157,244	157,291	157,263	0,018
RN2.8	72,843	67,344 ± 27,765	0,890	189,914	195,943	193,376	2,282
RN2.9	75,256	69,923 ± 27,825	0,893	187,462	193,291	190,797	2,206
RN2.10	67,441	62,130 ± 26,234	0,937	192,878	203,528	198,590	3,999
RN2.11	63,569	58,387 ± 25,140	0,955	195,121	209,042	202,333	5,218
RN2.12	67,646	62,475 ± 25,940	0,961	192,375	203,826	198,245	4,279

Fonte: Autor (2022)

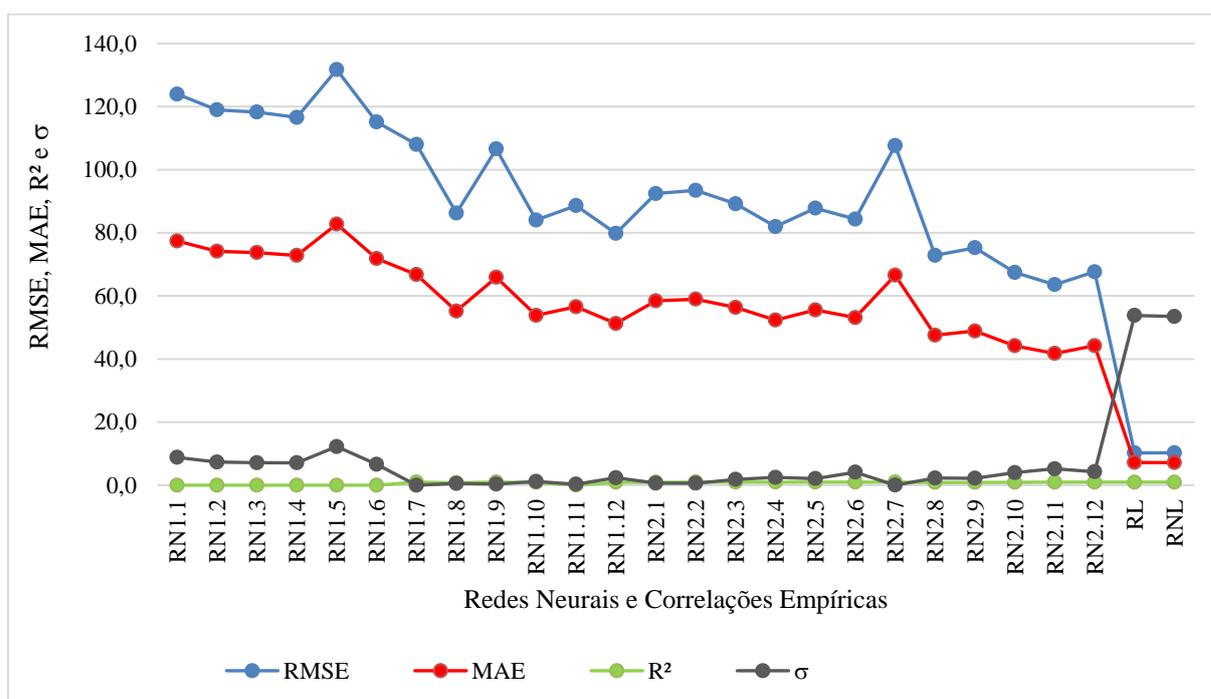
Observa-se, conforme resultados do grupo B-03, que as arquiteturas utilizadas para cada conjunto de dados de entrada, possuem resultados muito variados. A utilização apenas dos dados de entrada LL_c e LL_p (RN2.1 a RN2.12) proporcionou uma diminuição da média dos erros MAE de 66,85 para 52,33, RMSE de 106,49 para 81,96, e um aumento do valor médio de R^2 de 0,39 para 0,96 em comparação com os obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS (RN1.1 a RN1.12). Dentre as redes RN 1.1 e 1.12, nota-se ainda que, para as arquiteturas contendo apenas uma camada oculta, o valor de R^2 foi nulo.

Portanto, analisando os erros MAE e RMSE das redes RN2.1 a RN2.12, a arquitetura com 15 neurônios na primeira camada e 7 na segunda (RN2.11) foi a que mostrou menores valores (RMSE= 63,57 e MAE=41,76). Entretanto, o maior valor de R^2 , 1,0, foi observado para a arquitetura contendo 12 neurônios em uma única camada oculta (RN2.5).

Em comparação com os resultados obtidos pelas análises de regressão, observa-se que os erros obtidos com as redes neurais se mostraram muito maiores. Em comparação com os

valores médios, as RNAs apresentaram erros RMSE e MAE de 94,23 e 59,59 e as correlações de regressão 10,25 e 7,19 respectivamente. Os valores médios de σ foram de 3,50 para as RNAs e de 53,64 para as correlações de regressão. A Figura 87 apresenta graficamente o desempenho das redes neurais em comparação com as regressões.

Figura 87 - Resultados estatísticos das redes neurais e regressões do subgrupo B-03iii.



Fonte: Autor (2022)

Da Figura 87 nota-se as diferenças citadas entre as correlações de regressão e as RNAs. Assim, embora os altos valores de σ das correlações de regressão, nota-se que estas se adequam melhor a previsão de valores de LL_p para o subgrupo de dados B-03iii.

4.3.9 Conclusões parciais para o grupo de dados B-03

Apresentam-se os resultados das análises estatísticas e das redes neurais artificiais dos dados experimentais do grupo B-03 e seus subgrupos (B-03i, B-03ii e B-03iii) e suas discussões. O grupo B-03 e seus subgrupos correspondem aos resultados experimentais de LL_c obtidos pelo aparelho de Casagrande de base dura e LL_p obtido pelo cone sueco.

Conforme citado nas Conclusões parciais para os grupos de dados B-01 e B-02, itens 4.1.9 e 4.2.9, os resultados experimentais previamente compilados da literatura foram selecionados para as análises estatísticas e redes neurais artificiais conforme a metodologia abordada no

item 3.1.2. Para a análise da influência do intervalo de LL nos valores de LL_c e LL_p , os grupos foram separados em três subgrupos: B-03i ($LL < 50$), B-03ii ($50 < LL \leq 120$) e B-03iii ($LL > 120$).

Os dados, devido a variabilidade das amostras do banco de dados, também foram submetidos a definição e eliminação dos *outliers* por meio dos gráficos de *outliers* e diagramas *boxplot*, conforme metodologia descrita no item 3.2. O grupo B-03 e seus subgrupos possuem resultados experimentais do Brasil e da Grécia. De uma maneira geral, a eliminação dos *outliers* provocou uma redução do número de amostras analisadas nesta dissertação, além de diminuir o intervalo de LL para cada grupo ou subgrupo. A maioria dos *outliers* corresponde aos maiores valores de LL_c , LL_p ou IP das amostras de cada conjunto de dados. Como resultado, o grupo B-03 e o subgrupo B-03ii foram limitados a valores de $LL_c < 80$, obtidos pelos aparelhos de Casagrande de base dura. Para os subgrupos B-03i e B-03iii não foram identificados *outliers*.

A descrição estatística mostrou que os resultados obtidos com a seleção dos dados e eliminação dos *outliers* proporcionaram uma redução de cerca de 82% nos valores de desvio padrão, 97% da variância e 64% do coeficiente de variação para as variáveis LL_c , LL_p e IP quando comparados com os dados brutos compilados da literatura. Da divisão do grupo B-03 em subgrupos de dados, notou-se uma redução superior a 47% nos valores de desvio padrão e 70% da variância para o subgrupo B-03ii. Essa redução foi bem menor para o subgrupo B-03i, com uma diminuição de 5% nos valores médios de desvio padrão e 6% da variância. Quando analisados os coeficientes de variação, houve um aumento para 9% para o subgrupo B-03i e um aumento de 28% para o subgrupo B-03iii e uma diminuição de 50% do subgrupo B-03ii quando comparados com o grupo B-03. Os valores médios observados para os coeficientes de variação dos subgrupos B-03i e B-03ii foram inferiores a 25%, o que possibilitou concluir pela homogeneidade dos dados destes conjuntos. Entretanto, para o subgrupo B-03iii foi observado um aumento de cerca de 86% nos valores de desvio padrão, 98% da variância e 28% do coeficiente de variação para as variáveis LL_c , LL_p e IP. O coeficiente de variação observado para as variáveis LL_c , LL_p e IP ficou acima de 30%. Nota-se, portanto, uma maior variabilidade das amostras de solo para o conjunto de dados do subgrupo B-03iii.

A variabilidade das amostras também foi observada nas diferentes classificações de solo, granulometria e tipos de argilominerais que caracterizam os resultados experimentais que compõem o banco de dados. Em todos os conjuntos de dados analisados foi observado que as amostras são compostas por mais de 80% de argilas de alta e baixa compressibilidade (CL e CH). As demais amostras foram classificadas como siltes de baixa e alta compressibilidade (ML e MH). Da granulometria dos solos informada nas publicações de Clemente (2018) e

Clemente et al. (2020) foi observado que os dados contemplam amostras de solo com frações de silte e argila. Observa-se que para estes dados, cerca de 70% das amostras são compostas por solos finos, argilas e siltes. As amostras de solo publicadas por Christaras (1991) apresentaram um valor médio de 30% de argila. No grupo B-03 foi notada uma variedade maior de classificações, contendo amostras de argilas e siltes de baixa e alta compressibilidade, ou seja, CL, ML, CH e MH. Os demais subgrupos tiveram as classificações atribuídas ao intervalo de LL em que estão inseridos. Assim, o subgrupo B-03i, $LL_c < 50$, contém em sua maioria amostras de solo de baixa compressibilidade CL e ML. Os subgrupos B-03ii e B-03iii, $LL_c > 50$, solos de alta compressibilidade, CH e MH. A mineralogia das amostras de solo não é informada nas publicações.

Dentre os resultados experimentais de LL_c e LL_p para o intervalo de para $LL_c < 120$ (grupo B-03, subgrupo B-03i e subgrupo B-03ii), a maioria das amostras (80% para o subgrupo B-03, 100% para o subgrupo B-03i e 59% para o subgrupo B-03ii) apresentou $LL_p > LL_c$ e para $LL_c > 120$, subgrupo B-03iii, 75% das amostras apresentou valores de $LL_p < LL_c$. Destaca-se, portanto, que as diferenças entre LL_c e LL_p não refletem a totalidade dos resultados experimentais. Nota-se com a porcentagem de resultados de $LL_p > LL_c$ do subgrupo B-02ii (59%), que conforme LL_c aumenta, os dados tendem a $LL_p < LL_c$.

De maneira análoga a descrição apresentada para o grupo B-01, item 4.1.9, os dados de cada grupo e subgrupo foram submetidos a análises de regressão para definição das correlações de regressão lineares e não lineares e comparados através da metodologia abordada no item 3.4. A comparação entre as correlações lineares e não lineares dos conjuntos de dados analisados indicam que os modelos propostos resultam em valores muito próximos para a previsão de LL_p . Os valores de R^2 foram praticamente coincidentes para todos os conjuntos de dados do grupo B-03 e as diferenças entre os erros RMSE, MAE e σ foram pequenas. Portanto, conclui-se pela semelhança entre as correlações de regressão na explicação dos modelos propostos.

Os resultados experimentais de LL_c e LL_p que compõem os conjuntos de dados, juntamente com as correlações de regressão obtidas de maneira análoga aos grupos B-01 e B-02, também foram plotados no gráfico $LL_c \times LL_p$ para a análise gráfica dos resultados (itens 4.3.1, 4.3.3, 4.3.5 e 4.3.7). As correlações obtidas para o subgrupo B-03i, $LL_c < 50$, se mostraram acima da linha de igualdade $LL_c = LL_p$, de maneira semelhante ao observado para o grupo B-03 em geral ($LL_c < 80$). No entanto, nos resultados do subgrupo B-03ii as correlações se encontram entre as linhas $LL_p = LL_c$ e $LL_p = 1,2 LL_c$, para valores de LL_c inferiores a 58, e entre as linhas $LL_p = LL_c$ e $LL_p = 0,9 LL_c$ para valores superiores a 58, ou seja, para $LL_c > 58$, se encontram abaixo da linha de igualdade $LL_c = LL_p$. Esse resultado obtido para o subgrupo B-

03ii é confirmado pelos resultados do subgrupo B-03iii, com $LL_c > 120$, com as correlações também abaixo de $LL_c = LL_p$.

As correlações lineares propostas também foram submetidas às análises de resíduos e aos testes de hipóteses para verificação da normalidade do modelo e independência das variáveis, conforme metodologia citada no item 3.2. Para as correlações lineares propostas, em todos os conjuntos de dados analisados foram observados a heterocedasticidade. Apenas os resíduos verificados para o subgrupo B-03 foram classificados como dependentes, os demais foram todos ditos independentes. A distribuição dos resíduos foi normal apenas para o subgrupo B-03ii. Portanto, os testes estatísticos mostraram que as correlações lineares propostas para os conjuntos de dados do grupo B-03 e subgrupos B-03i, B-03ii e B-03iii não são significativas. Esses resultados são semelhantes aos observados para os conjuntos de dados dos grupos B-01 e B-02, conforme descrito nos itens 4.1.9 e 4.2.9, embora tenha sido observada uma maior homogeneidade das amostras de solo para o grupo B-03 e os subgrupos B-03i e B-03ii. A presença de dados atípicos nos conjuntos de dados analisados verificados com a assimetria dos histogramas de resíduos (Figura 74, Figura 78, Figura 82 e Figura 86) e com os pontos de dados identificados como resíduos (Figura 73, Figura 77, Figura 81 e Figura 85) justificam a heterocedasticidade dos modelos, resultados determinantes para a conclusão de que as correlações lineares não são significativas.

Os conjuntos de dados também foram analisados pelas redes neurais artificiais para previsão de LL_p , conforme metodologia citada no item 3.3. De maneira análoga aos grupos B-01 e B-02, diferentes arquiteturas e dados de entrada foram testados para a avaliação da capacidade de previsão das redes e sua consequente validação. Quanto ao número de neurônios e camadas ocultas, as arquiteturas propostas foram semelhantes para o grupo B-03 e subgrupo B-03i, $LL_c < 50$. Os demais subgrupos, B-03ii e B-03iii, apresentaram melhores resultados com a utilização de duas camadas ocultas. Para o subgrupo B-03ii foram observados valores nulos de R^2 quando utilizados apenas uma camada oculta. E, para o subgrupo B-03iii, nota-se que a utilização de arquiteturas com uma camada oculta, proporcionou o aumento dos erros RMSE e MAE em cerca de 20% e em mais de 62% para o σ se comparado com as arquiteturas de duas camadas ocultas. Quando comparados os dados de entrada, para os subgrupos B-03i, B-03ii e B-03iii a utilização apenas de LL_c e LL_p (RN2.1 a RN2.12) proporcionou a redução dos erros RMSE, MAE e no σ . Os valores médios de R^2 observados para as redes RN2.1 a RN2.12 também foram superiores aos obtidos com a utilização de LL_c , LL_p , IP e Classificação SUCS como dados de entrada. Para os subgrupos B-03ii e B-03iii, $LL_c > 50$, as diferenças observadas foram superiores a 60%.

Ainda de maneira análoga aos grupos B-01 e B-02, os resultados obtidos com as correlações de regressão foram confrontados com os resultados obtidos pelas redes neurais artificiais. Foi observado que os valores de σ são menores para as correlações de regressão apenas do subgrupo B-03ii. Maiores diferenças foram observadas para os subgrupos B-03iii, onde o σ médio das correlações de regressão foi cerca de 93% maior do que o observado para as RNAs. Portanto, observadas as a variação das RNAs quanto aos diferentes tipos de arquitetura e de entrada de dados, as RNAs indicaram uma menor variabilidade para a previsão dos valores de LL_p do grupo B-03 e subgrupos B-03i e B-03ii, $LL_c < 120$. Ou seja, essas conclusões são válidas para a seleção adequada das arquiteturas e entrada de dados. Nota-se que as diferenças observadas entre as correlações de regressão e as RNAs foram maiores para o subgrupo B-03iii, $LL_c > 120$, dada a maior variedade dos dados que o compõe. Portanto, as correlações de regressão, para $LL_c > 120$, indicaram uma menor variabilidade para a previsão dos valores de LL_p . Entretanto, ressalta-se que apesar destes resultados, as correlações lineares não foram significativas.

De maneira análoga as conclusões apresentadas para o grupo B-02, destaca-se que o tamanho das amostras dos subgrupos analisados não é grande o suficiente para fornecer uma estimativa precisa da força de correlação, assim, estimativas da força de relação, como o R^2 podem variar muito. O subgrupo B-03ii apresentou um baixo valor de R^2 ($< 0,46$) para as correlações lineares e não lineares se comparado ao grupo B-03 ($R^2 > 0,85$) e subgrupos B-03i ($R^2 = 0,89$) e B-03iii ($R^2 = 0,96$). Este subgrupo, apesar do R^2 , apresentou baixos coeficientes de variação para as variáveis LL_c , LL_p e IP ($< 21\%$), o que nos permitiu concluir que é composto de dados homogêneos. Foi observado ainda, por meio dos testes de hipótese, que, apesar da heterocedasticidade do modelo, a distribuição dos resíduos é normal. Portanto, conclui-se que para o subgrupo B-03ii, o R^2 não pode ser utilizado para avaliação da adequação dos modelos de regressão. Destaca-se ainda que o subgrupo de dados B-03iii ($LL_c > 120$) possui um número muito reduzido de amostras (16 amostras), com alta variabilidade, o que requer cautela na interpretação dos dados obtidos.

Nota-se ainda que o R^2 obtido com as redes neurais para os subgrupos B-03ii e B-03iii, utilizando como dados de entrada LL_c , LL_p , IP e Classificação SUCS, em alguns casos foi igual a 0, indicando que o modelo em nada explica a variabilidade dos dados de resposta ao redor de sua média. Nestes casos, conforme citado, recomenda-se a seleção de modelos de arquitetura com número de neurônios, camadas ocultas e dados de entrada adequados para a utilização das RNAs na previsão dos valores de LL_p .

5. CONCLUSÕES E SUGESTÕES PARA TRABALHOS FUTUROS

5.1 CONCLUSÕES

Nesta pesquisa procurou-se analisar e comparar valores limites de liquidez (LL) obtidos pelo método de percussão de Casagrande, LL_c (aparelhos de base dura e macia), e pelos cones britânico e sueco, LL_p , para diferentes solos finos e uma ampla variação de valores de LL. Foram investigados, portanto, correlações de regressão entre LL_c e LL_p e a utilização da técnica de redes neurais artificiais para previsão dos valores de LL_p analisando a influência do intervalo de LL nos modelos estudados. Os resultados e discussões e as conclusões parciais apresentados no item 4 possibilitaram a consideração de algumas importantes conclusões apresentadas neste capítulo.

Devido a variabilidade observada nas amostras dos bancos de dados dos grupos B-01, B-02, B-03 e seus subgrupos que possuem resultados experimentais de diversos locais (Brasil, Egito, Equador, Estados Unidos, Grécia, Itália, Índia, Portugal e Turquia), os procedimentos de categorização dos dados, a identificação dos *outliers* e as análises estatísticas foram eficientes na redução da heterogeneidade das amostras. A descrição estatística mostrou que os resultados obtidos com a seleção dos dados e eliminação dos *outliers* proporcionaram, em todos os casos, uma redução dos valores de desvio padrão, variância e coeficiente de variação para as variáveis LL_c , LL_p e IP. Entretanto, quando da divisão dos dados para avaliação da influência dos intervalos de LL, para $LL_c > 120$ (subgrupos B-01iii, B-02iii e B-03iii) não foram identificados *outliers*. Para este intervalo, $LL_c > 120$, notou-se dados mais heterogêneos se comparados com os obtidos para $LL_c < 120$. Quando analisados os coeficientes de variação, para verificação da homogeneidade dos dados, os resultados divergiram para cada um dos grupos de dados. Apenas os resultados obtidos para os subgrupos B-01i, B-01iii e B-03iii apresentaram coeficientes de variação acima de 25%. Cabe ressaltar que embora o coeficiente de variação seja muito utilizado para expressar a variabilidade dos dados estatísticos, não há um consenso entre pesquisadores quanto aos intervalos que definem as baixas, médias e altas dispersões. Nesta pesquisa, foi adotado o valor de 25% como indicativo de altas dispersões entre os resultados, caracterizando os dados como heterogêneos.

A variabilidade das amostras também foi observada nas diferentes classificações de solo, granulometria e tipos de argilominerais que caracterizam os resultados experimentais que compõem os bancos de dados. Em todos os conjuntos de dados foram observados mais de 53% das amostras compostas por argilas de baixa e alta compressibilidade (CL e CH).

A mineralogia dos solos estudados não é informada em todas as publicações, no entanto, alguns autores relatam amostras compostas por caulinitas, ilitas, bentonitas e montmorilonitas, e em alguns casos a mistura desses argilominerais. Os tipos de argilominerais identificados nos conjuntos de dados também é variável conforme o intervalo de LL em que estão inseridos e corroboram para a diversidade das amostras de solo citadas para cada conjunto de dados. No grupo B-01 foram registradas amostras compostas por caulinitas no intervalo de $LL_c < 120$ (subgrupos B-01i e B-01ii). No grupo B-02, as amostras de caulinitas e ilitas foram registradas para o intervalo de $LL_c < 50$ (subgrupo B-02i). Para os demais intervalos destes grupos foram registradas amostras compostas por montmorilonitas e bentonitas. Para o grupo B-03 não foi mencionada a mineralogia das amostras. Dada as divergências citadas quanto aos intervalos e a falta de informações das publicações, não foi possível definir nesta pesquisa um intervalo de LL relacionado ao tipo de argilomineral. Ressalta-se que a divisão dos limites de LL_c estabelecidos para a formação dos subgrupos nesta pesquisa ocorreu de maneira aleatória.

Dentre os resultados experimentais de LL_c e LL_p para $LL_c < 120$ nos grupos B-01 e B-03, acima de 80% das amostras apresentou $LL_c < LL_p$. Esses grupos possuem em comum a utilização do aparelho de Casagrande de base dura, porém divergem quanto ao tipo de cone. Para o grupo B-02, composto de amostras obtidas com o aparelho de Casagrande de base macia e cone britânico, cerca de 72% das amostras no intervalo de $LL_c < 50$ apresentou $LL_c < LL_p$. Portanto, não foi possível estabelecer um único intervalo de LL para o qual o método de cone apresente resultados menores do que o de Casagrande sem que sejam considerados os aparelhos utilizados. Além disso, destaca-se que as diferenças entre LL_c e LL_p não refletem a totalidade dos resultados experimentais. Ou seja, não é possível afirmar que para determinada combinação de aparelhos (B-01, B-02, B-03) e intervalo de LL os resultados de LL_c serão sempre menores que os de LL_p ou o inverso.

A avaliação da mudança na classificação SUCS dos solos quando utilizado os resultados do limite de liquidez obtido pelo método do cone indicaram que cerca de 14% das amostras de solo do grupo B-01, 25% do grupo B-02 e 22% do grupo B-03 mudam de classificação em comparação aos resultados obtidos pelo método de Casagrande. Embora as diferenças nas classificações tenham sido inferiores a 25% para todos os conjuntos de dados analisados, essas mudanças podem ser suficientes para alterar a classificação de um solo de adequado a inadequado devido a valores de LL ou IP abaixo dos permitidos. No grupo B-01 os resultados experimentais que apresentaram maior mudança de posição no ábaco de Casagrande quando analisados LL_c e LL_p , estavam no intervalo de $LL_c < 80$. Portanto, a utilização de valores de LL_p no Sistema Unificado de Classificação de Solos (SUCS) pode afetar o uso de materiais em engenharia geotécnica.

Das correlações de regressão obtidas, em todos os casos, a comparação entre as correlações lineares e não lineares dos conjuntos de dados analisados indicam que os modelos propostos resultam em valores muito próximos para a previsão de LL_p . Portanto, conclui-se pela semelhança entre as correlações de regressão lineares e não lineares na explicação dos modelos propostos.

As correlações de regressão obtidas nesta pesquisa se mostraram acima da linha de igualdade para o intervalo de $LL_c < 120$ no grupo B-01, $LL_c < 50$ no grupo B-02 e $LL_c < 58$ no grupo B-03. Conclui-se que para estes intervalos, aplicadas as correlações propostas, os valores previstos de LL_p são superiores aos resultados experimentais de LL_c . Para os demais intervalos, dentro que cada grupo de dados, é observado o inverso, ou seja, as correlações se encontram abaixo da linha de igualdade e preveem resultados de LL_p inferiores aos resultados experimentais de LL_c . Observa-se, portanto, que os resultados variam quanto aos tipos de equipamentos empregados na obtenção de LL_c e LL_p e com o intervalo de LL .

Os resultados dos testes estatísticos mostraram que as correlações lineares propostas para os conjuntos de dados não são significativas. Esses resultados refletem a variabilidade dos dados analisados. Portanto, embora a eliminação dos *outliers* tenha proporcionado a diminuição da heterogeneidade dos conjuntos de dados (B-01, B-02 e B-03), o tratamento desses dados levou em consideração apenas os resultados obtidos dos gráficos de *outliers* e diagramas *boxplot*. As assimetrias observadas nos histogramas de LL_p e dos resíduos, e os pontos de dados identificados como resíduos indicaram a presença de dados atípicos nos conjuntos de dados analisados, ou pontos fora da curva, que justificam a heterocedasticidade e a distribuição não normal dos resíduos, observada para a maioria dos conjuntos de dados. Assim, conclui-se que, para o tratamento e eliminação dos *outliers*, devem ser examinados além dos gráficos de *outliers* e diagramas *boxplot* os gráficos de dispersão dos resíduos e histogramas.

Vale ressaltar ainda que, de uma maneira geral, as correlações consideradas não significativas pelos testes de resíduos e de normalidade (*Kolmogorov-Smirnov*, *Shapiro-Wilk* e *Durbin-Watson*) apresentaram em sua maioria, elevados valores de coeficiente de determinação, R^2 . Entretanto, um R^2 próximo de 1 não é suficiente para a tomada de decisão quanto a aceitação de um determinado modelo. Observou-se que embora os altos valores de R^2 ($R^2 > 0,74$) para a maioria dos conjuntos de dados, os modelos de regressão linear não foram significativos. Destaca-se que o subgrupo B-03ii, embora a homogeneidade dos dados e distribuição normal dos resíduos, foi o único a apresentar um baixo valor de R^2 ($< 0,46$) para as correlações de regressão lineares e não lineares. Portanto, nota-se por meio dos resultados obtidos que o R^2 embora forneça uma estimativa da força de correlação entre o modelo e a

variável de resposta, não indica se um modelo de regressão é adequado, já que é possível ter um valor baixo de R^2 para um bom modelo ou um valor alto de R^2 para um modelo que não se encaixa nos dados. Em todos os casos, recomenda-se a identificação e tratamento dos pontos de dados atípicos identificados como resíduos nesta pesquisa como um esforço necessário à validação das correlações lineares propostas.

Os resultados obtidos com as redes neurais artificiais indicaram a semelhança entre as arquiteturas e entrada de dados testadas para o grupo B-01. Entretanto, quando analisados os grupos B-02 e B-03, os resultados variaram em maior parte quanto aos dados de entrada de dados. Para esses grupos, a utilização de LL_c e LL_p como dados de entrada proporcionou a redução dos erros RMSE, MAE e do σ quando comparados com a utilização de LL_c , LL_p , IP e classificação SUCS. Portanto, a dependência do valor de LL_c no valor correspondente de IP limita os resultados obtidos com as RNAs para definir a possível associação entre valores de LL_c e LL_p , o que nos permite concluir que as variáveis LL_c e LL_p exercem uma maior influência na definição dos modelos. Ressalta-se que os bancos de dados com resultados experimentais dos grupos B-02 (aparelho de Casagrande de base macia e cone britânico) e B-03 (aparelho de Casagrande de base dura e cone sueco), são menores que o banco de dados que correlaciona os valores de limites de liquidez obtidos pelos métodos de Casagrande de base dura e cone britânico, grupo B-01. Ou seja, além da variabilidade das propriedades geotécnicas das amostras dos conjuntos de dados, o número de amostras utilizadas na previsão de LL_p pelas RNAs também influencia nos resultados.

Quando avaliados o número de neurônios e camadas ocultas adotados para as RNAs, as diferenças observadas foram pontuais, ou seja, na maioria dos casos foi observada a semelhança entre as arquiteturas propostas. No subgrupo B-02iii a utilização de arquiteturas com duas camadas ocultas proporcionou o aumento dos erros RMSE e MAE em mais de 66% e uma redução de cerca de 70% nos valores de σ se comparado com as arquiteturas de apenas uma camada oculta. No subgrupo B-03ii a utilização de uma camada oculta resultou em valores nulos de R^2 . E, no subgrupo B-03iii, a utilização de arquiteturas com uma camada oculta, proporcionou o aumento dos erros RMSE e MAE em cerca de 20% e em mais de 62% para o σ se comparado com as arquiteturas de duas camadas ocultas. Portanto, não foi possível aderir por um número de neurônios e/ou camadas ocultas que indicasse um modelo de arquitetura mais adequada a previsão de LL_p .

Quando confrontados os resultados obtidos com as correlações de regressão e redes neurais artificiais, a fim de indicar o método que melhor se adequa a cada um dos modelos propostos, os resultados variaram conforme cada conjunto de dados. De uma maneira geral, para $LL_c > 120$ (subgrupos B-01iii, B-02iii e B-03iii), os resultados apontaram para maiores

diferenças entre as correlações de regressão e as RNAs. Esses resultados ocorrem em função da maior variabilidade das amostras de solo observadas neste intervalo. Tendo como premissa os valores de σ , as RNAs indicaram uma menor variabilidade para previsão de LL_p para o grupo B-01 e subgrupo B-01i. Os resultados foram semelhantes para o grupo B-02 e os subgrupos B-02i e B-02iii e para o grupo B-03 e subgrupos B-03i e B-03ii ($LL_c < 120$). Portanto, não foi possível estabelecer um intervalo de LL para o qual as RNAs se mostram mais adequadas a previsão de LL_p quando comparadas com as correlações de regressão para todos os grupos de dados investigados. Conforme destacado por Bicalho et al. (2022), as redes neurais representam uma ferramenta adicional para os métodos empíricos convencionais de regressão permitindo o processamento de uma grande quantidade de dados, com aplicabilidade à problemas complexos e com alto grau de linearidade. Desta forma, cabe ressaltar que não há uma única metodologia a ser utilizada. Nesta pesquisa foram realizadas as análises detalhadas dos dados e as análises de regressão. Como resultado, verificou-se que cada modelo (correlação de regressão ou RNAs) se ajusta melhor a determinada metodologia tendo em vista a distribuição dos dados nos intervalos de LL estudados, os métodos de ensaio empregados e suas combinações.

5.2 SUGESTÕES PARA TRABALHOS FUTUROS

Tendo em vista que esta pesquisa não esgota os estudos acerca da comparação entre os valores de limite de liquidez obtidos por diferentes metodologias, sugere-se para trabalhos futuros:

- i. A obtenção de um banco de dados mais detalhado acerca da caracterização dos solos e dos ensaios, com um maior número de amostras que contenham especificações do tipo de argilomineral predominante, granulometria, teor de argila, bem como do método de preparação das amostras para a utilização dessas informações na avaliação de sua influência nos valores de limite de liquidez obtidos pelos métodos de Casagrande e do cone;
- ii. A investigação de outras arquiteturas possíveis de redes neurais artificiais para previsão dos valores de LL_p , bem como a avaliação de outros dados de entrada para as redes que melhor se ajustem aos modelos analisados;
- iii. A investigação mais aprofundada dos intervalos de LL representativos das variações entre as diferenças de LL_c e LL_p , tendo em vista as divergências observadas da literatura;
- iv. A análise de correlações empíricas que utilizam LL para estimar propriedades geotécnicas de solos e avaliação do efeito da variabilidade dos resultados considerando diversas metodologias de ensaios e equipamentos utilizados para

determinar valores de LL nos resultados estimados das propriedades geotécnicas investigadas.

- v. Reunir amostras de solos finos de diferentes regiões do Brasil e desenvolver os ensaios de caracterização para avaliação dos valores de limite de liquidez obtidos pelos métodos de Casagrande e do cone em solos brasileiros.

REFERÊNCIAS

- ANDRADE, F. A., AL-QURESHI, H. A HOTZA, D. Measuring the plasticity of clays: A review. **Applied Clay Science**, [S.L], v. 51, n.1-2, p. 1-7, jan. 2011. Elsevier BV. <https://doi.org/10.1016/j.clay.2010.10.028>
- ARAMA, Z. A., MELDA, Y., AKIN, M. S., DALYAN, I. A comparative study on the application of artificial intelligence networks versus regression analysis for the prediction of clay plasticity. **Arabian Journal of Geosciences**, 14:534, p. 1-16, 2021. <https://doi.org/10.1007/s12517-021-06894-x>
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR-7180**: “Solo - Determinação do limite de plasticidade”. Rio de Janeiro, 2016.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR-6459**: “Solo - Determinação do limite de liquidez”. Rio de Janeiro, 2017.
- ATTERBERG, A. Die Plastizität der Tone. **Internationale Mitteilungen der Bodenkunde**, v. 1, p. 4-37, 1911.
- BAIN, J. A. Plasticity Chart as an Aid to the Identification and Assessment of Industrial Clays. **Clay Minerals**, [S.L.], v. 9, n. 1, p.1-17, jul. 1971. Mineralogical Society. <http://dx.doi.org/10.1180/claymin.1971.009.1.01>.
- BELVISO, R., CIAMPOLI, S., COTECCHIA, V. FREDERICO, A. Use of the cone penetrometer to determine consistency limits. **Ground Eng.**, v. 18, n. 5, p. 21-22, 1985.
- BICALHO, K. V., GRAMELICH, J. C., CUNHA, C. L. S. Comparação entre os valores do limite de liquidez obtidos pelos métodos de Casagrande e cone para solos argilosos brasileiros. **Comunicações Geológicas** (2014) 101, Especial III, 1097-1099, 2014.
- BICALHO, K. V.; BATISTITI, V. G.; XIMMENES, Roberta B. Compressibilidade das argilas marinhas, Grande Vitória, ES, Brasil. **Comunicações Geológicas** (2014). Especial III, CNG/2º CoGePLiP. Rio Grande do Sul, Porto Alegre, 2014a.
- BICALHO, K. V., GRAMELICH, J. C., CUNHA, C. L. S., JUNIOR, R. G. S. Estudo dos valores do limite de liquidez obtidos pelos métodos de Casagrande e Cone para diferentes argilas. **Geotecnia**, [S.L], n. 140, p. 63-72, jun. 2017. Coimbra University Press. <http://dx.doi.org/10.24849/j.geot.2017.140.04>.
- BICALHO, K. V., KÜSTER, J. S. H, TOTOLA, L. B., CRISTELLO, L. G. C. Discussion of. **Soils And Rocks**, [S.L.], v. 44, n. 2, p. 1-2, 18 jun. 2021. ABMS - Brazilian Association for Soil Mechanics and Geotechnical Engineering. <http://dx.doi.org/10.28927/sr.2021.062721>.

BICALHO, K. V., MORESCHI, M., KÜSTER, J. S. H., FILHO, A. G. O., HISATUGU, W. Estudo do uso da Técnica de Redes Neurais Artificiais em Engenharia Geotécnica. In: ANAIS DO XX COBRAMSEG, 2022, Campinas. **Anais eletrônicos...** Campinas, Galoá, 2022. Disponível em: <https://proceedings.science/cobramseg-2022/papers/estudo-do-uso-da-tecnica-de-redes-neurais-artificiais-em-engenharia-geotecnica>> Acesso em: 12 set. 2022.

CAMPBELL, D. J. Liquid Limit Determination of Arable Topsoils Using a Drop-Cone Penetrometer. **Journal of Soil Science**, [S.L.], v. 26, n. 23, p. 234-240, st. 1975. Wiley. <http://dx.doi.org/10.1111/j.1365-2389.1975.tb01946.x>.

CAMPOS, L. F. A., MACEDO, R. S. De, KIYOHARA, P. K., FERREIRA, H. C. Características de plasticidade de argilas para uso em cerâmica vermelha ou estrutural. **Cerâmica**, [S.L.], v. 45, n. 295, p. 140-145, maio 1999. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s0366-69131999000500006>.

CASAGRANDE, A. Research on the Atterberg limits of soils. *Public Road*, v. 13, n. 8, p. 121-136, 1932.

CASAGRANDE, A. Notes on the Design of the Liquid Limit Device. **Geotechnique**, [S.L.], v. 8, n. 2, p. 84-91, jun. 1958. Thomas Telford Ltda. <http://dx.doi.org/10.1680/geot.1958.8.2.84>

CASAGRANDE, A. Classification and identification of soils. *Transactions of the American Society of Civil Engineers*, vol. 113, p. 901-930, 1948.

CEVALLOS LUNA, A. S. **Determinación del limite líquido y plástico de los suelos mediante el uso del penetrómetro cónico**. 2012. 126 f. Dissertação - Pontifica Universidad Católica del Ecuador, Quito, 2012.

CHRISTARAS, B. A comparison of the Casagrande and fall cone penetrometer methods for liquid limit determination in marls from Crete, Greece. **Engineering Geology**, v. 31, p. 131-142. Elsevier Science Publishers B.V., 1991.

CHRYSTELLO, A. C. **Avaliação de curvas de calibração do método do papel filtro tipo Whatman nº42 utilizadas para estimar sucção matricial em solos não saturados**. 2017. 136 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Espírito Santo, Vitória, 2017.

CLAVEAU-MALLET, D., DUHAIME, F., CHAPUIS, R. P. Pratical Considerations when Using the Swedish Fall cone. **Geotechnical Testing Journal**, Vol. 35, n. 4, p.1-11, 2012.

CLEMENTE, C. W. **Determinação da resistência não drenada através do método do cone em solos marinhos – Análise, conceito e diretrizes**. 2018. 108 f. Dissertação (Mestrado em Engenharia de Construção Civil) – Universidade Federal do Paraná, Curitiba, 2018.

CLEMENTE, C. W., FARO, V. P., MONCADA, M. P. H. Determination of liquid limit by the Fall Cone method. **Soils and Rocks**, [S.L], v. 43, n. 4, p. 661-667, 31 dez. 2020. ABMS – Brazilian Association for Soil Mechanics and Geotechnical Engineering. <https://doi.org/10.28927/SR.434661>.

COUTO, B. O. C., GOMES, R. C., FERREIRA, L. D. Correlação entre os Valores do limite de liquidez obtidos pelos métodos de Casagrande e cone de queda livre para diferentes materiais. In: ANAIS DO XVIII COBRAMSEG, 2016, Belo Horizonte. **Anais eletrônicos...** Belo Horizonte, 2016.

CREVELIN, L. C. **Comparação entre os valores de limite de liquidez obtidos pelos métodos de Casagrande e de Cone para diferentes solos argilosos**. 2018. 162 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Espírito Santo, Vitória, 2018.

CREVELIN, L. C., BICALHO, K. V. Comparison of the Casagrande and Fall Cone Methods for Liquid Limit Determinations in Different Clay Soils. **Revista Brasileira de Ciência do Solo**, v. 43, 2019. <https://doi.org/10.1590/18069657rbc20180105>.

CUNHA, C. L. S. **Estudo das características de compressibilidade unidimensional e plasticidade de misturas de argila e areia**. 2012. 127 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Espírito Santo, Vitória, 2012.

DEKA, S., SREEDEEP, S., DASH, S. K. Re-evaluation of Laboratory Cone Penetration Method for High Liquid Limit Based on Free Swell Property of Soil. **Geotechnical Testing Journal**, v. 32, n. 6, p. 1-6, 2009.

DIAZ, E.; PASTOR, J. L; RABAT, Á; TOMÁS, R. Machine learning techniques for relating liquid limit obtained by Casagrande cup and fall cone test in low-medium plasticity fine grained soils. **Engineering Geology**, [S.L], v. 294, p. 1-10, dez 2021, Elsevier BV. <https://doi.org/10.1016/j.enggeo.2021.106381>

DI MATTEO, L. Liquid limit of low- to medium-plasticity soils: comparison between casagrande cup and cone penetrometer test. **Bulletin Of Engineering Geology And The Environment**, [S.L.], v. 71, n. 1, p. 79-85, 2012. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10064-011-0412-5>.

DRAGONI, W. PROSPERINI, N., VINTI, G. Some observation on the procedures for the determination of the liquid limit: an application on Plio-Pleistocenic clayey soils from Umbria region (Italy). **Ital J Eng Geol Environ Spec Issue**, v. 1, p. 185-197, 2008.

EDIB, A. M. 35 Years of (AI) in Geotechnical Engineering: State of the Art. **Geotechnical And Geological Engineering**, [S.L.], v. 39, n. 2, p. 637-690, 17 set. 2020. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10706-020-01536-7>.

EDIB, A. M., NWOBIA, L. I., ONYELOWE, K. C., ANEKE, F. I. Predicting Nanobinder-Improved Unsaturated Soil Consistency Limits Using Genetic Programming and Artificial Neural Networks. **Applied Computation Intelligence And Soft Computing**, [S.L.], v. 2021, p. 1-13, jul. 2021. Hindawi Limited. <http://dx.doi.org/10.1155/2021/5992628>.

EL-SHINAWI, A. A Comparison of Liquid Limit Values for Fine Soils: A Case Study at the North Cairo-Suez District, Egypt. **Journal Geological Society of India**, v. 89, p. 339-343, mar. 2017.

FARRELL, E., SCHUPPENER, B. AND WASSING, B. ETC. 5. Fall-cone study. **Ground Eng.**, v. 30, n. 1, p. 33-36, 1997.

FILHO, A. G. O. **Análise comparativa da estimativa do índice de compressão de argilas por redes neurais artificiais e correlações empíricas**. 2019. 203 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Espírito Santo, Vitória, 2019.

FOJTOVÁ, L. MARSCHALKO, M., FRANEKOVÁ, R. KOVAR, L. Study of compatibility of methods for liquid limit measurement according to Czech State Standart and newly adopted European Standart. **GeoScience Engineering LV**, v. 1, p. 233-241, 2009.

FREITAS JÚNIOR, M. S. Limites de Consistência: Correlações e causas principais de erros de interpretação, p. 163-176, 1974.

GIASI, C. I.; CHERUBINI, C.; PACCAPELO, F. Evaluation of compression index of remoulded clays by means of Atterberg limits. **Bulletin Of Engineering Geology And The Environment**, [S.L.], v. 62, n. 4, p. 333-340, nov. 2003. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10064-003-0196-3>.

GRONBECH, L. G., NIELSEN, B. N., IBSEN, L. B. Comparison of liquid limit f highly plastic clay by means casagrande and fall cone apparatus. In: **Symposium Proceedings: 64th Canadian Geotechnical Conference and 14th Pan-American Conference on Soil Mechanics and Engineering, 5th Pan-American Conference on Teaching and Learning of Geotechnical Engineering. Pan-AM CGS Geotechnical Conference**, 2011.

FLECK, L., TAVARES, M. H. F., EYNG, E. HELMANN, A. C., ANDRADE, M. A. M. Redes Neurais Artificiais: Princípios Básicos. **Revista Científica Eletrônica Inovação e Tecnologia**, v. 1, n. 13, p. 47-57, 2016.

GRABOWSKA-OLSZEWSKA, B. Modelling physical properties of mixtures of clays: example of a two-component mixture of kaolinite and montmorillonite. **Applied Clay Science**, [S.L.], v. 22, n. 5, p. 251-259, abr. 2003. Elsevier BV. [http://dx.doi.org/10.1016/s0169-1317\(03\)00078-4](http://dx.doi.org/10.1016/s0169-1317(03)00078-4).

HAIGH, S. K. Mechanics of the Casagrande liquid limit test. **Canadian Geotechnical Journal**, [S.L.], v. 49, n.9, p. 1015-1023, set. 2012. Canadian Science Publishing. <http://dx.doi.org/10.1139/t2012-066>.

HIND, K. J. The Casagrande plasticity chart – does it help or hinder the NZGS soil classification process? In: **Proc. 20th NZG Geotechnical Symposium**, p. 1-8, Eds. GJAlexander & CY Chin, Napier, 2017.

HOFFMAN, R. **Análise de regressão: uma introdução à econometria** [recurso eletrônico]. 5. Ed. Piracicaba: O Autor, 2016.

KARAKAN, E., SHIMOBÉ, S., SEZER, A. Effect of clay fraction and mineralogy on fall cone results of clay-sand mixtures. **Engineering Geology**, [S.L.], v. 279, p. 1-19, dez. 2020. Elsevier BV. <https://doi.org/10.1016/j.enggeo.2020.105887>.

KAYABALI, K. Determination of Consistency Limits: A Comparison between -#40 and -#200 Materials. *Electronic Journal of Geotechnical Engineering*, v. 16, p. 1547-1561, 2011.

KESTLER, M. A. Correlations and comparisons between the Casagrande liquid limit device and the fall Cone. M. S., Thesis, Massachusetts Institute of Technology, 1982.

JUWAIED, N. S. Applications of artificial intelligence in geotechnical engineering. **ARP Journal of Engineering and Applied Sciences**, v. 13, n. 8, p. 2764-2785, abr. 2018.

KOESTER, J. P. The Influence of Test Procedure on Correlation of Atterberg Limits with Liquefaction in Fine-grained Soils. **Geotechnical Testing Journal**, v. 15, n. 4, p. 352–361, 1992.

LEE, L. T, FREEMAN, R. B. An Alternative Test Method for Assessing Consistency Limits. **Geotechnical Testing Journal**, v. 30, n. 4, p. 1-8, 2007.

LI, C., ASHLOCK, J., WANG, X. Quantifying Repeatability Reproducibility Sources of Error and Capacity of a Measurement: demonstrated using laboratory soil plasticity tests. **Advances In Civil Engineering**, [S.L.], v. 2019, p. 1-11, jul. 2019. Hindawi Limited. <http://dx.doi.org/10.1155/2019/4539549>.

LIU, T. K.; THORNBURN, T. M. Study of the Reproducibility of Atterberg Limits, **Highway Research Record**, v. 63, p. 22-30, 1964.

MANAFI, M. S. G. Soil Plasticity Variability and Its Effect on Soil Classification. **Geotechnical Testing Journal**, [S.L], v. 42, n. 2, p. 457-470, ago. 2018. ASTM International. <http://dx.doi.org/10.1520/gtj20170273>.

MITCHELL, J. K., SOGA, K. **Fundamentals of Soil Behavior**. 3rd ed. John Wiley & Sons, 2005.

MISHRA, A. K, OHTSUBO, M., LI, L. Y., HIGASHI, T. **Influence of various factors on the difference in the liquid limit values determined by Casagrande's and fall cone method**. Environmental Earth Sciences, [S.L.], v. 65, n. 1, p. 21-27, abr. 2011. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s12665-011-1061-5>.

MENDOZA, M. J., OROZCO, M. Quick and Reliable Procedure for Liquid Limit Determination of Fine-Grained Soils. **Geotechnical Testing Journal**, v. 24, n. 1, p. 103-108, mar. 2001.

MORENO-MAROTO, J. M., ALONSO-AZCÁRATE, J. An accurate, quick and simple method to determine the plastic limit and consistency changes in all types of clay and soil: the thread bending test. **Applied Clay Science**, [S.L.], v. 114, p. 497-508, set. 2015. Elsevier BV. <http://dx.doi.org/10.1016/j.clay.2015.06.037>.

MOAYEDI, H., MOSALLANEZHAD, M., RASHID, A. S. A., JUSOH, W. A. W., MUAZU, M. A. A systematic review and meta-analysis of artificial neural network application in geotechnical engineering: theory and applications. **Neural Computing & Applications**, [S.L.], v. 32, n.2, p. 495–518, mar. 2019. <https://doi.org/10.1007/s00521-019-04109-9>

NAGARAJ, H. B., SRIDHARAN, A. MADHU, B. V. Comparative Study of Determination of Liquid Limit by Percussion Cup, Cone and K_0 -Stress Methods. In: **8th International Symposium on Lowland Technology**, 2012.

NIAZI, S. F., PINAN-LLAMAS, A., CHOLEWA, C., AMSTUTZ, C. Liquid limit determination of low to medium plasticity Indian soils by hard base Casagrande percussion cup vs. BS fall-cone methods. **Bulletin of Engineering Geology and The Environment**, [S.L.], v. 79, n. 4, p. 2141-2158, dez. 2019. Springer Science and Business Media LLC. <https://doi.org/10.1007/s10064-019-01668-y>.

NINI, R. Effect of the silt and clay fractions on the liquid limit measurements by Atterberg cup and fall cone penetrometer. **International Journal Of Geotechnical Engineering**, [S.L.], v. 8, n. 2, p. 239-241, fev. 2014. Informa UK Limited. <http://dx.doi.org/10.1179/1939787913y.0000000018>.

O'KELLY, B. C., VARDANEGA, P. J., HAIGH, S. K. Use of fall cones to determine Atterberg limits: a review. **Géotechnique**, [S.L], v. 68, n. 10, p. 843-856, out. 2018. Thomas Telford Ltda. <https://doi.org/10.1680/jgeot.17.R.039>.

O'KELLY, B. C., VARDANEGA, P. J., HAIGH, S. K., BICALHO, K. V., FLEUREAU, J. M. & CUI, Y. J. (2020). Use of fall cones to determine Atterberg limits: a review. **Géotechnique**, [S.L], v. 70, n. 7, p. 652-654, jul. 2020. Thomas Telford Ltda. <https://doi.org/10.1680/jgeot.18.D.001>.

O'KELLY, B. C. Review of Recent Developments and Understanding of Atterberg Limits Determinations. **Geotechnics**, [S.L], v. 1, n. 1, p. 59-75, maio 2021. MDPIAG. <https://doi.org/10.3390/geotechnics1010004>.

O'KELLY, B. C., SOLTANI, A. Machine learning techniques for relating liquid limit obtained by Casagrande cup and fall cone test in low-medium plasticity fine grained soils [Eng. Geol. (2021) 294, 106381]. **Engineering Geology**, [S.L], v. 306, p. 1-4, set. 2022. Elsevier BV. <http://dx.doi.org/10.1016/j.enggeo.2022.106746>.

OZER, M. Comparison of liquid limit values determined using the hard and soft base Casagrande apparatus and the cone penetrometer. **Bulletin Of Engineering Geology And The Environment**, [S.L], v. 68, n. 3, p. 289-296, fev. 2009. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10064-009-0191-4>.

PRAKASH, K. Discussion of "Mechanics of the Casagrande liquid limit test" [Apperars in the Canadian Geotechnical Journal, 49(9): 1015-1023]. **Canadian Geotechnical Journal**, [S.L], v. 49, n. 11, p. 1325-1326, nov. 2012. Canadian Science Publishing. <http://dx.doi.org/10.1139/t2012-089>.

PEREIRA, G. M. B., BASTOS, L. B., GONÇALVES, K., WITECK, G. R., PALADINI, E. P. APLICAÇÃO DA TÉCNICA REGRESSÃO LINEAR MÚLTIPLA: Um estudo de caso de uma fábrica de cimentos. **Revista de Ensino de Engenharia**, [S.L], v. 39, n.1, p. 50-60, 2020. Revista de Ensino em Engenharia. <http://dx.doi.org/10.37702/ree2236-0158.v39p50-60.2020>.

PINTO S., CASTRO, P. F. Determinação do Limite de Liquidez pelo método do cone de penetração. **IPR/DNER**, 24 p., 1971.

PINTO, P. B., ROSA, K. L., ALVEZ, A. M. L., BASTOS, C. A. B. **Comparação entre Resultados de Ensaios de Casagrande e de Cone de Queda Livre para Determinação do Limite de Liquidez**. 2012.

QUINTELA, A., COSTA, C., TERROSO, D., ROCHA, F. Liquid Limit determination of clayey material by Casagrande method, fall cone test and EBS parameter. **Materials Technology**,

[S.L], v. 29, n. 3, p. B82-B87, jul. 2014. Informa UK Limited. doi: <http://dx.doi.org/10.1179/1753555714y.0000000153>

QUEIROZ DE CARVALHO, J. B. **Algumas propriedades físico-químicas e de engenharia de uma argila orgânica do Recife**. 1975. 139 f. Dissertação (Mestrado em Ciências) – Universidade Federal da Paraíba, Campina Grande, 1975.

REIS, E.A, REIS, I.A. **Análise Descritiva de Dados**. 5. Ed. Relatório Técnico do Departamento de Estatística da UFMG: 2002. Disponível em: www.est.ufmg.br

REHMAN, H. U., POULADI, N., PULIDO-MONCADA, M., ARTHUR, E. Repeatability and agrément between methods for determining the Atterberg limits of fine-grained soils. **Soil Science Society Of America Journal**, [S.L], v. 84, n. 1, p. 21-30, jan. 2020. Wiley. <http://dx.doi.org/10.1002/saj2.20001>

SANTOS, M. D. **Correlações entre sondagens de simples reconhecimento e resultados de ensaios de campo (SPT, CPT, DP) para diferentes solos arenosos**. 2017. 175 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Espírito Santo, Vitória, 2017.

SHAHIN, M. A., JAKSA, M. B., MAIER, H. R. Artificial Neural Network Applications in Geotechnical Engineering. **Australian Geomechanics**, v.36, n. 1, p. 49-62, mar. 2001.

SHAHIN, M. A., JAKSA, M. B., MAIER, H. R. State of the Art of Artificial Neural Networks in Geotechnical Engineering. **Eletronic Journal of Geotechnical Engineering**, v. 8, p. 1-26, 2008.

SHIMOBÉ, S., SPAGNOLI, G. A global database considering Atterberg limits with Casagrande and fall-cone tests. **Engineering Geology**, [S.L], v. 260, p. 1-11, out. 2019. Elsevier BV. <https://doi.org/10.1016/j.enggeo.2019.105201>.

SHIMOBÉ, S., SPAGNOLI, G. Fall cone tests considering water content, cone penetration index, and plasticity angle of fine- grained soils. **Journal of Rock Mechanics And Geotechnical Engineering**, [S.L], v. 12, p. 1347-1355, dez. 2020. Elsevier BV. <https://doi.org/10.1016/j.jrmge.2020.02.005>.

SILVA, A. C. S. **Definição de padrões de plasticidade de solos através do cone de penetração, por comparação com o método de Casagrande**. 2013. 119 f. Dissertação (Mestrado em Engenharia Geotécnica e Geoambiente) – Instituto Superior de Engenharia do Porto, Porto, Portugal, 2013.

SILVEIRA, L. C. S. **Cone de Penetração: Limites de Liquidez Ajustados a Solos Tropicais**. 1998.

SILVEIRA, L. C. S. O cone de penetração como ensaio alternativo na determinação da plasticidade de solos. In. XXIX COBRENTE. **Anais...**, Porto Alegre, p. 16-19, 2001.

SITTON, J. D., ZEINALI, Y., STORY, B. A. Rapid soil classification using artificial neural networks for use in constructing compressed earth blocks. **Construction and Building Materials**, [S.L], v. 138, p. 214-221, maio 2017. Elsevier BV <http://dx.doi.org/10.1016/j.conbuildmat.2017.02.006>.

SOUSA, P. M. L. P. **Limite de liquidez – Correlações e Comparações entre os Métodos de Fall Cone e da Concha de Casagrande**. 2011. 148 f. Dissertação (Mestrado em Engenharia Geológica - Geotecnia) – Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, Lisboa, Portugal, 2011.

SOWERS, G. B., SOWERS, G. F. **Introducción a la Mecánica de Suelos y Cimentaciones**. Tradução de José Menendez. Revisão de Alfonso Rico Rodrigues. 1. ed. México: Editorial Limusa-Wiley S. A., 1972.

SPAGNOLI, G. Comparison between Casagrande and dro-cone methods to calculate liquid limit for purê clay. **Canadian Journal Of Soil Science**, [S.L], v. 92, n. 6, p. 859-864, nov. 2012. Canadian Science Publishing. <http://dx.doi.org/10.4141/cjss2012-011>.

SRIDHARAN, A., PANDIAN, N. S., PRASAD, P. S. Liquid Limit Determination of Class F Coal Ash. **Journal of Testing and Evaluation**, v. 28, n. 6, p. 455-461, 2000.

SRIDHARAN, A.S., PRAKASH, K. Percussion and Cone Methods of Determining the Liquid Limit of Soils. **Geotechnical Testing Journal**, v. 23, n. 2, p. 242-250, jun. 2000.

THERON, E., STOOT, P. R., VOSLOO, P., LANGROUDI, A. A. Assessment of the suitability of the fall cone method to replace the Casagrande cup for liquid limit determination of South African Soils. In: **Proceedings of the 17th African Regional Conference on Soil Mechanics and Geotechnical Engineering** At: Cape Town, South Africa, v. 1, p. 175-179, out. 2019.

TOTOLA, L. B. **Aplicação de redes neurais artificiais na estimativa de curvas de retenção de solos brasileiros**. 2020. 126 f. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal do Espírito Santo, Vitória, 2020.

VARDANEGA, P. J., HAIGH, S. K., O'KELLY, B. C. Use of fall-cone flow index for soil classification: a new plasticity chart. **Géotechnique**, [S.L], v. 72, n. 7, p. 610-617, jul. 2022. Thomas Telford Ltda. <http://dx.doi.org/10.1680/jgeot.20.p.132>.

VERÁSTEGUI-FLORES, R. D., EMIDIO, G. Di. Assessment of clay consistency through conventional methods and indirect extrusion tests. **Applied Clays Science**, [S.L], v. 101, p. 632-636, nov. 2014. Elsevier BV. <http://dx.doi.org/10.1016/j.clay.2014.09.033>.

WASTI, Y., BEZIRCI, M. H. Determination of the consistency limits of soils by the fall cone test. **Canadian Geotechnical Journal**, v. 23, p. 241-246, 1986.

WIDJAJA, B., KURNIAWAN, K. Development of Relationship between Liquid Limit and Clay Content for West Java Soils. **Iop Conference Series: Materials and Engineering**, [S.L], v. 1007, n. 1, p. 1-7, dez. 2020. IOP Publishing. <http://dx.doi.org/10.1088/1757-899x/1007/1/012102>.

YOUSEFPOUR, N., MEDINA-CETINA, Z., HERNANDEZ-MARTINEZ, F. G., AL-TABBAA, A. Stiffness and Strength of Stabilized Organic Soil – Part II/II: parametric analysis and modeling with machine learning. **Geosciences**, [S.L], v. 11, n. 5, p. 218, maio 2021. <https://doi.org/10.3390/geosciences11050218>.

ZENTAR, R. ABRIAK, N., DUBOIS, V. Effects os salts and organic matter on Atterberg limits of dredged marine sediments. **Applied Clay Science**, [S.L], v. 42, n. 3-4, p. 391-397, jan. 2009. Elsevier BV. <http://dx.doi.org/10.1016/j.clay.2008.04.003>.

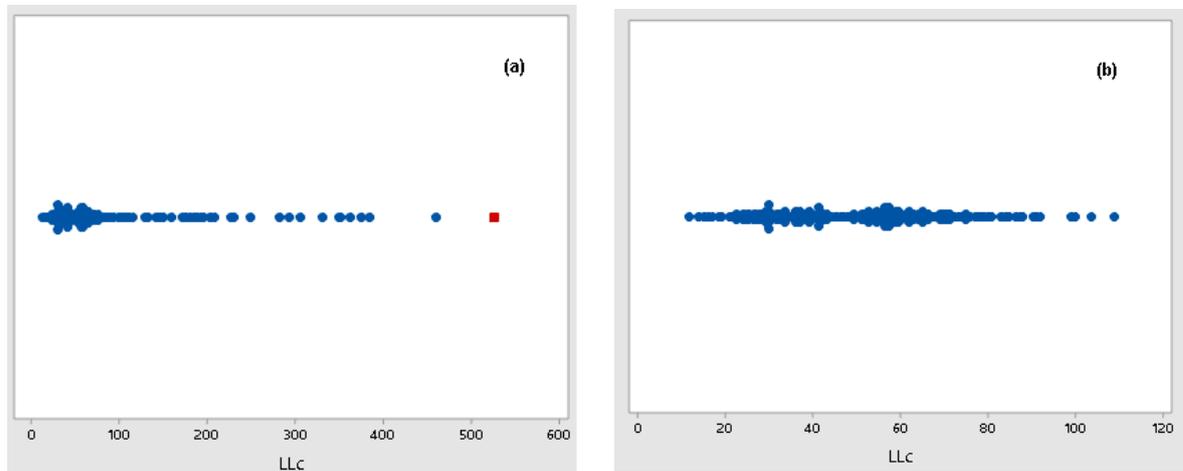
ZHANG, W., LI, H., LI, Yongqin, LIU, Hanlong, CHEN, Yumin, DING, X. Application of deep lerarning algorithms in geotechnical engineering: a short critical review. **Artificial Intelligence Review**, [S.L], v. 54, n. 8, p. 5633-5673, fev. 2021. Spring Science and Business Media LLC. <https://doi.org/10.1007/s10462-021-09967-1>.

ZOLFAGHARI, Z., MOSADDEGHI, M. R., AYOUBI, S. ANN-based pedotransfer and soil prediction functions for predicting Atterberg consistency limits and indices from easily available properties at the watershed scale in western Iran. **Soil Use And Management**, [S.L], v. 31, n. 1, p. 142-154, jan. 2015. Wiley. <http://dx.doi.org/10.1111/sum.12167>.

APÊNDICE A

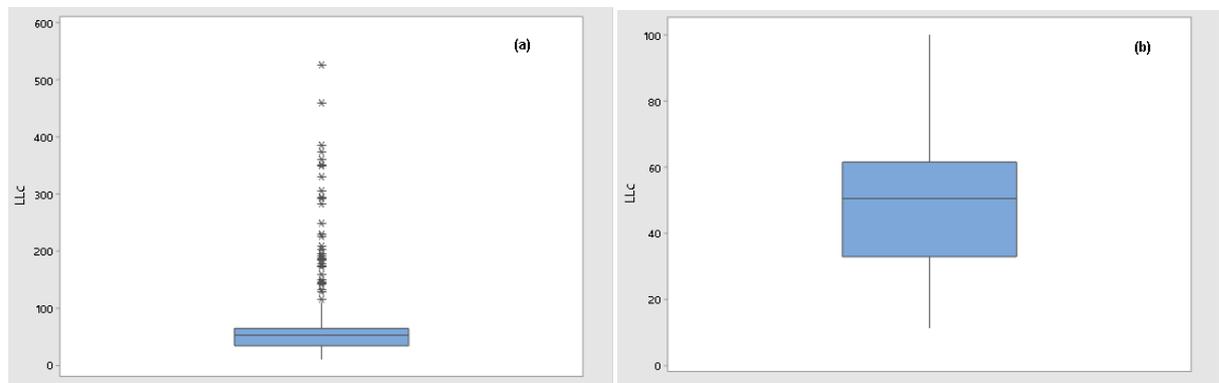
Gráficos de *outliers* e diagramas *boxplot* para os dados do grupo B-01

Figura 1 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_c do grupo B-01.



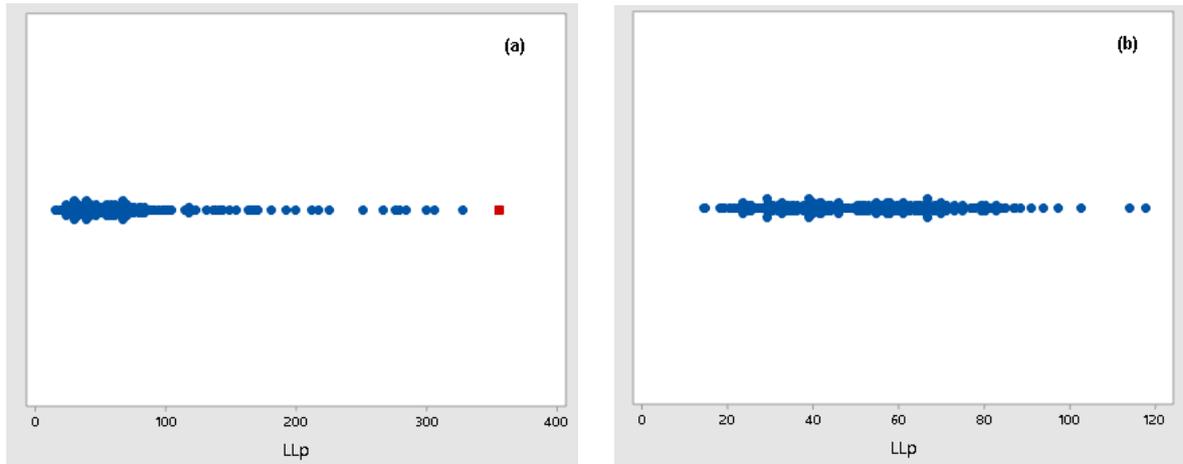
Fonte: Autor (2022)

Figura 2 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_c do grupo B-01.



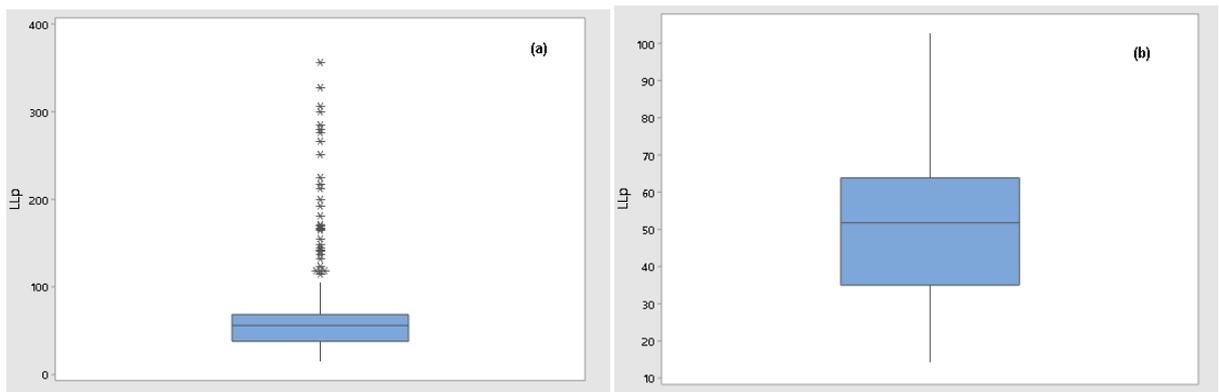
Fonte: Autor (2022)

Figura 3 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_p do grupo B-01.



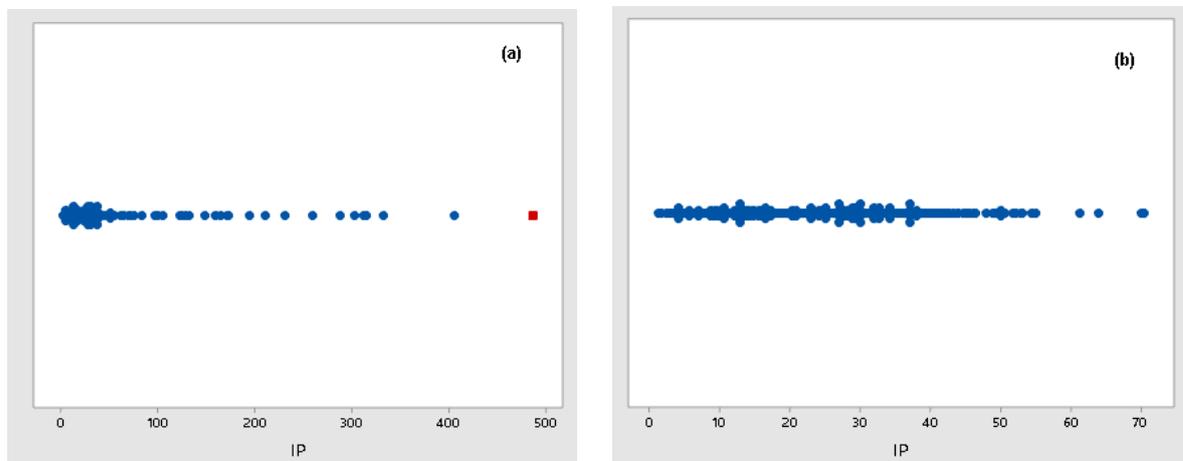
Fonte: Autor (2022)

Figura 4 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_p do grupo B-01.



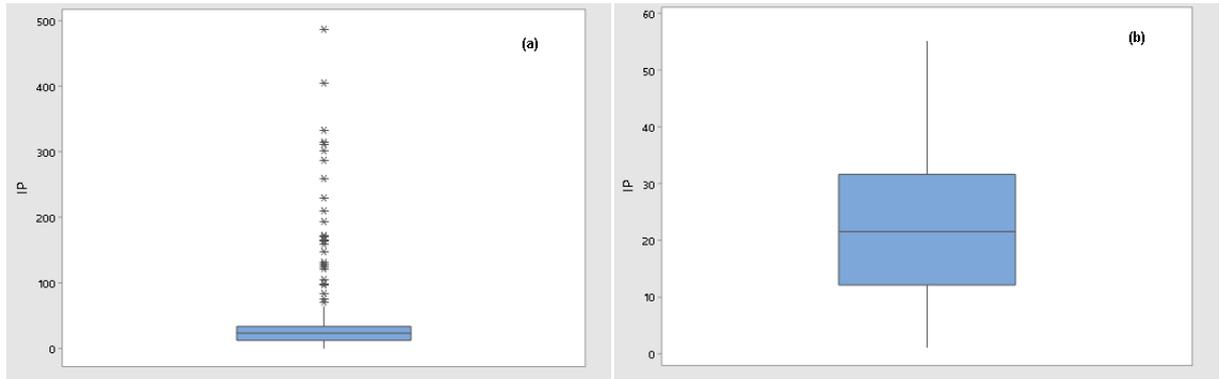
Fonte: Autor (2022)

Figura 5 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade IP do grupo B-01.



Fonte: Autor (2022)

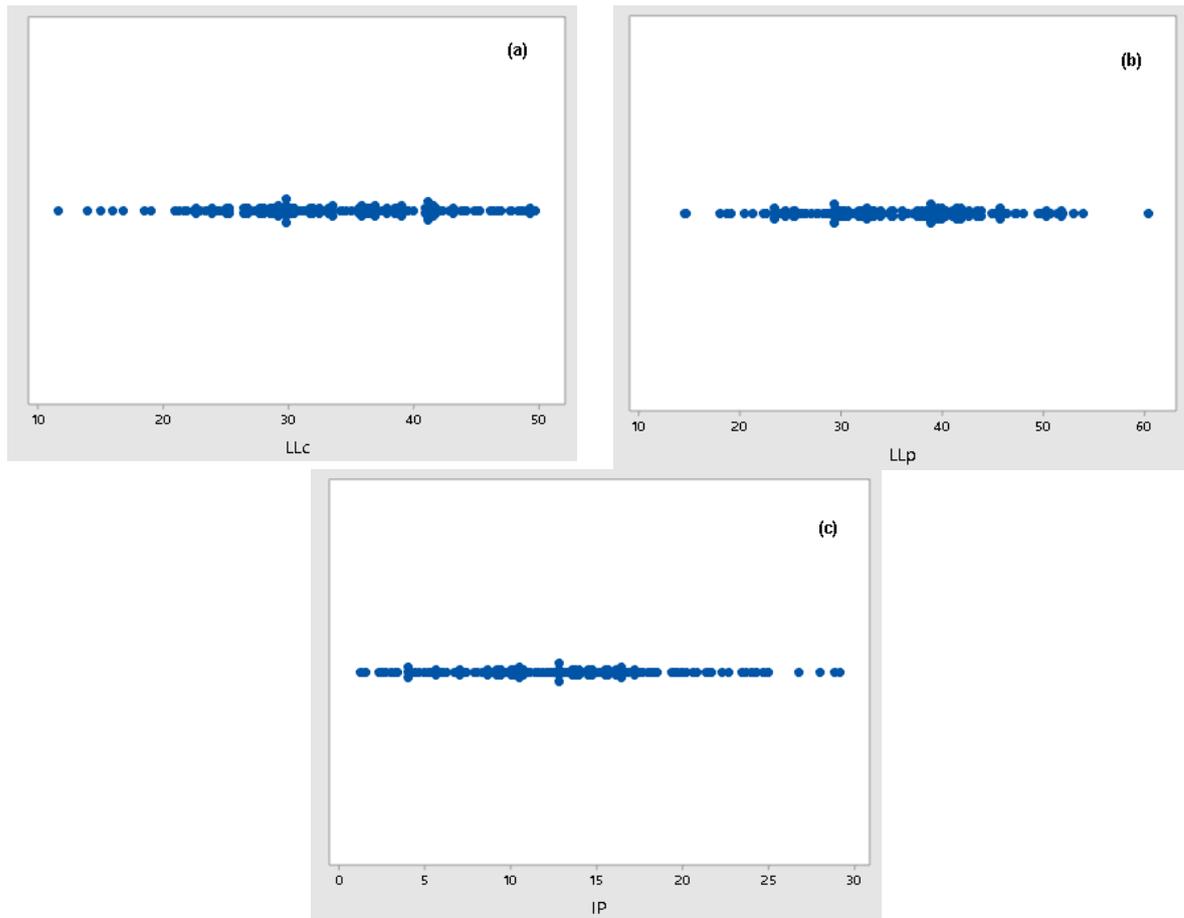
Figura 6 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-01.



Fonte: Autor (2022)

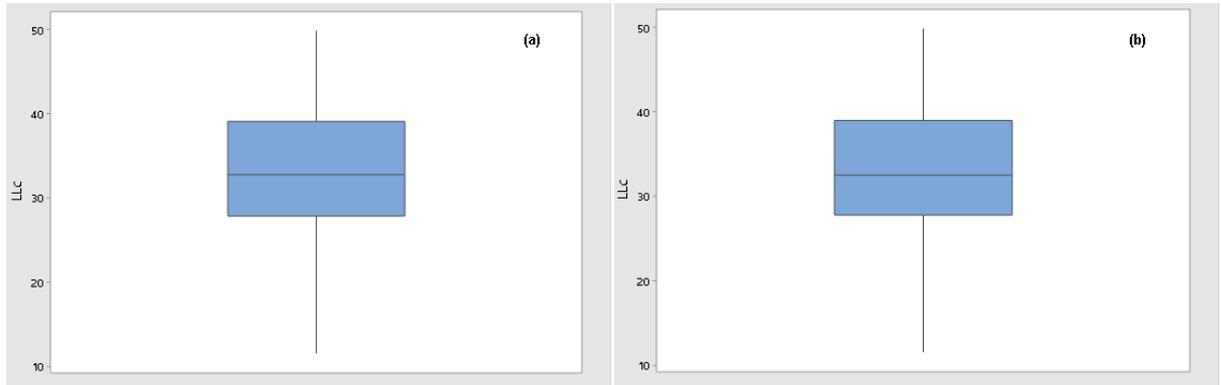
Gráficos de outliers e diagramas *boxplot* para os dados do subgrupo B-01i

Figura 7 – Gráfico de outliers para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-01i.



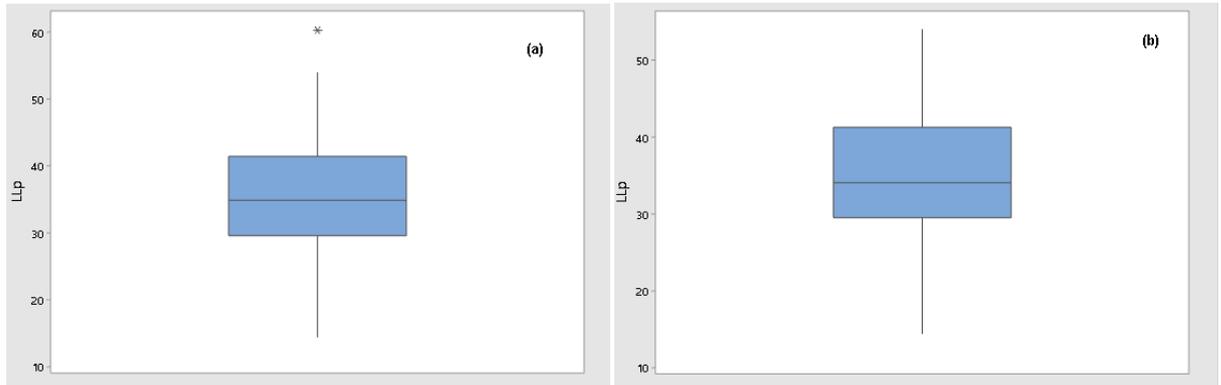
Fonte: Autor (2022)

Figura 8 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_c do grupo B-01i.



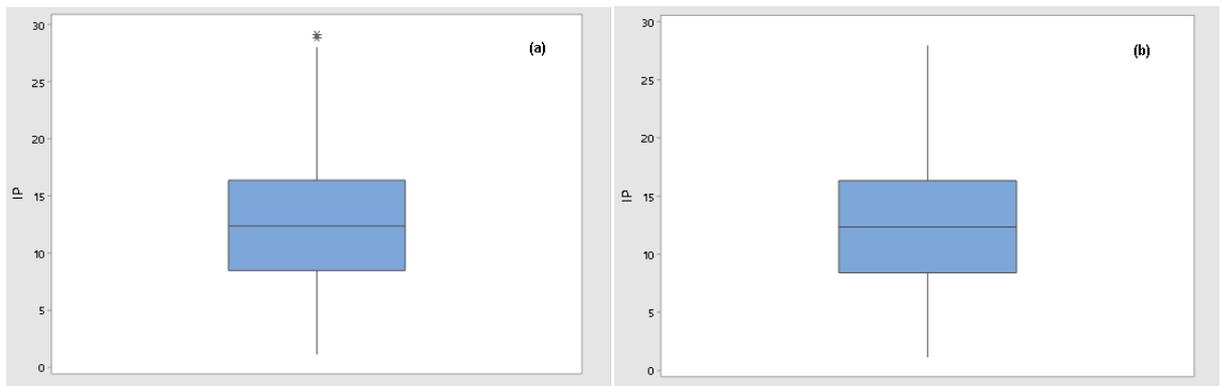
Fonte: Autor (2022)

Figura 9 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_p do grupo B-01i.



Fonte: Autor (2022)

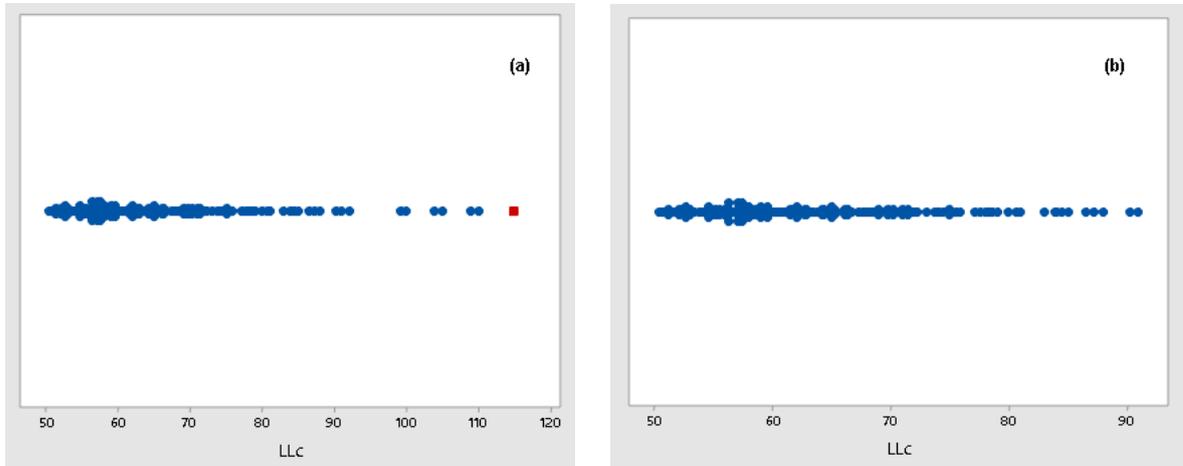
Figura 10 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-01i.



Fonte: Autor (2022)

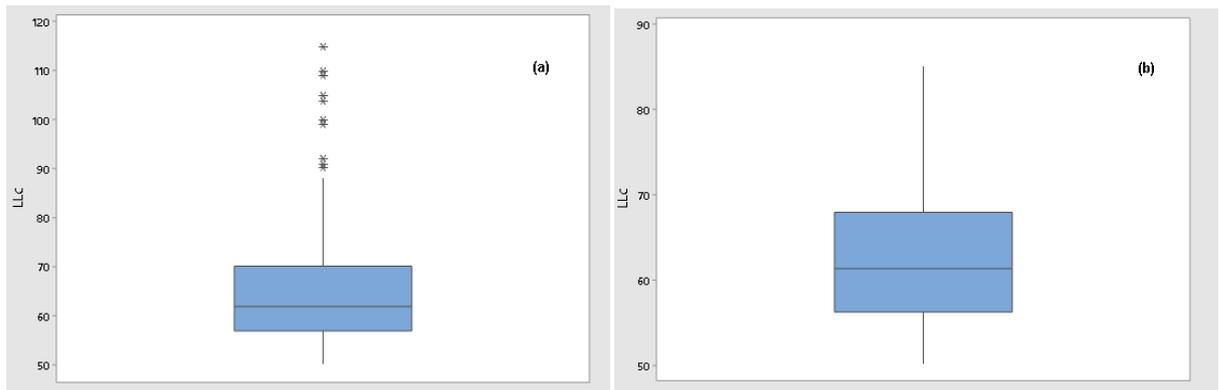
Gráficos de outliers e diagramas *boxplot* para os dados do subgrupo B-01ii

Figura 11 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_c do grupo B-01ii.



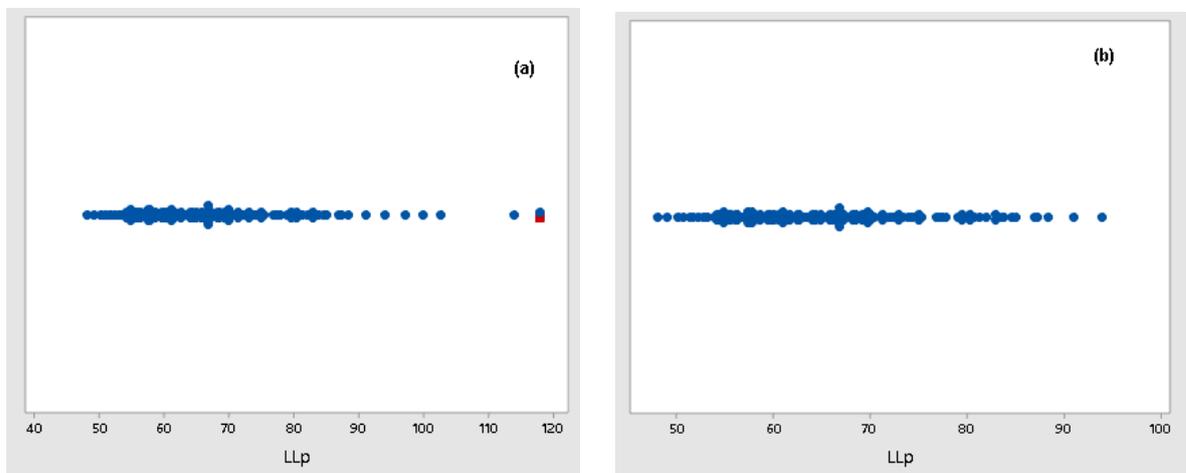
Fonte: Autor (2022)

Figura 12 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no *boxplot* para a propriedade LL_c do grupo B-01ii.



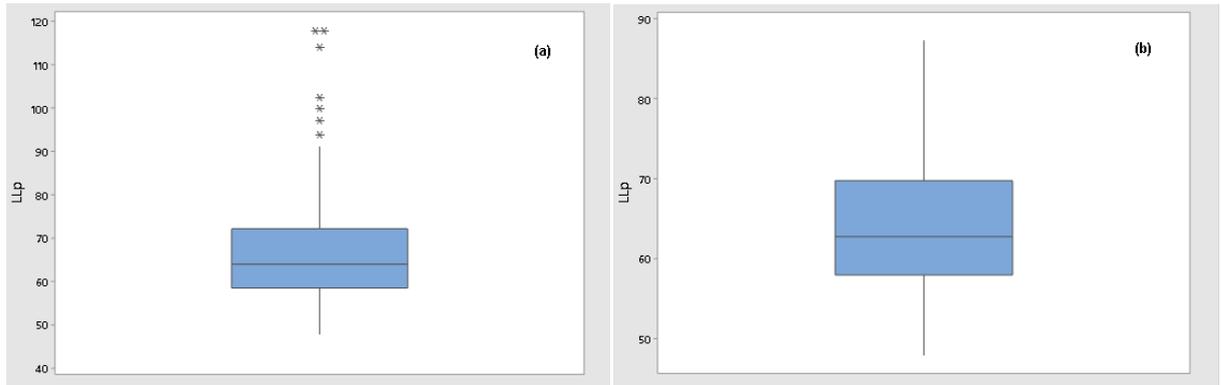
Fonte: Autor (2022)

Figura 13 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_p do grupo B-01ii.



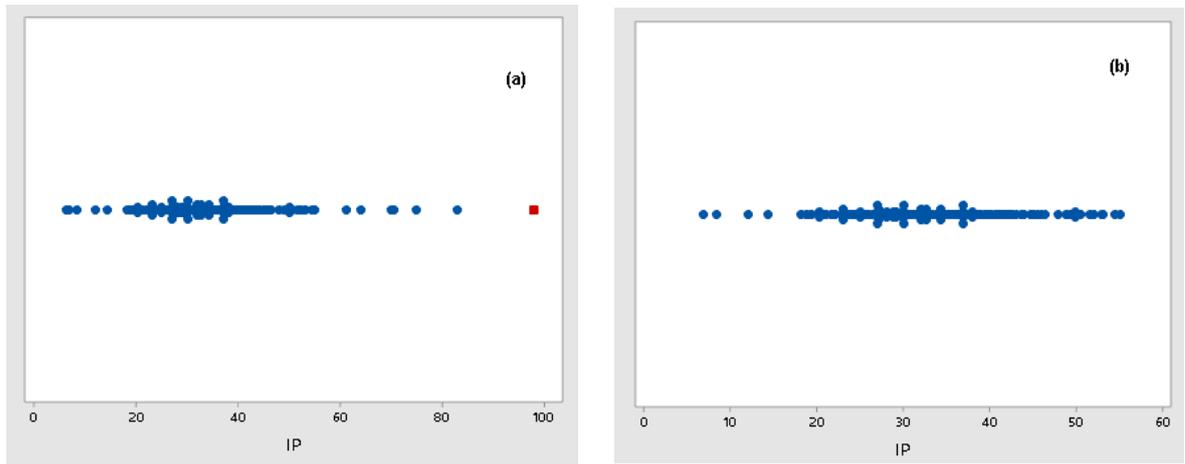
Fonte: Autor (2022)

Figura 14 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_p do grupo B-01ii.



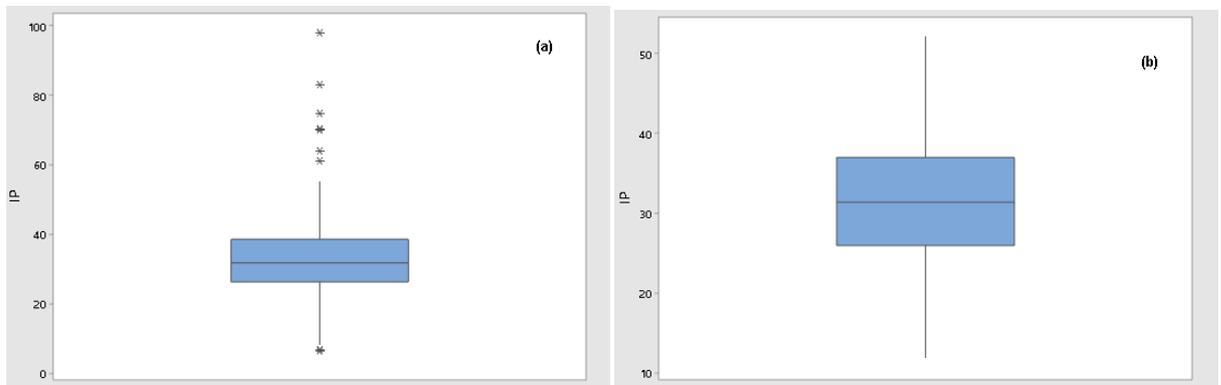
Fonte: Autor (2022)

Figura 15 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade IP do grupo B-01ii.

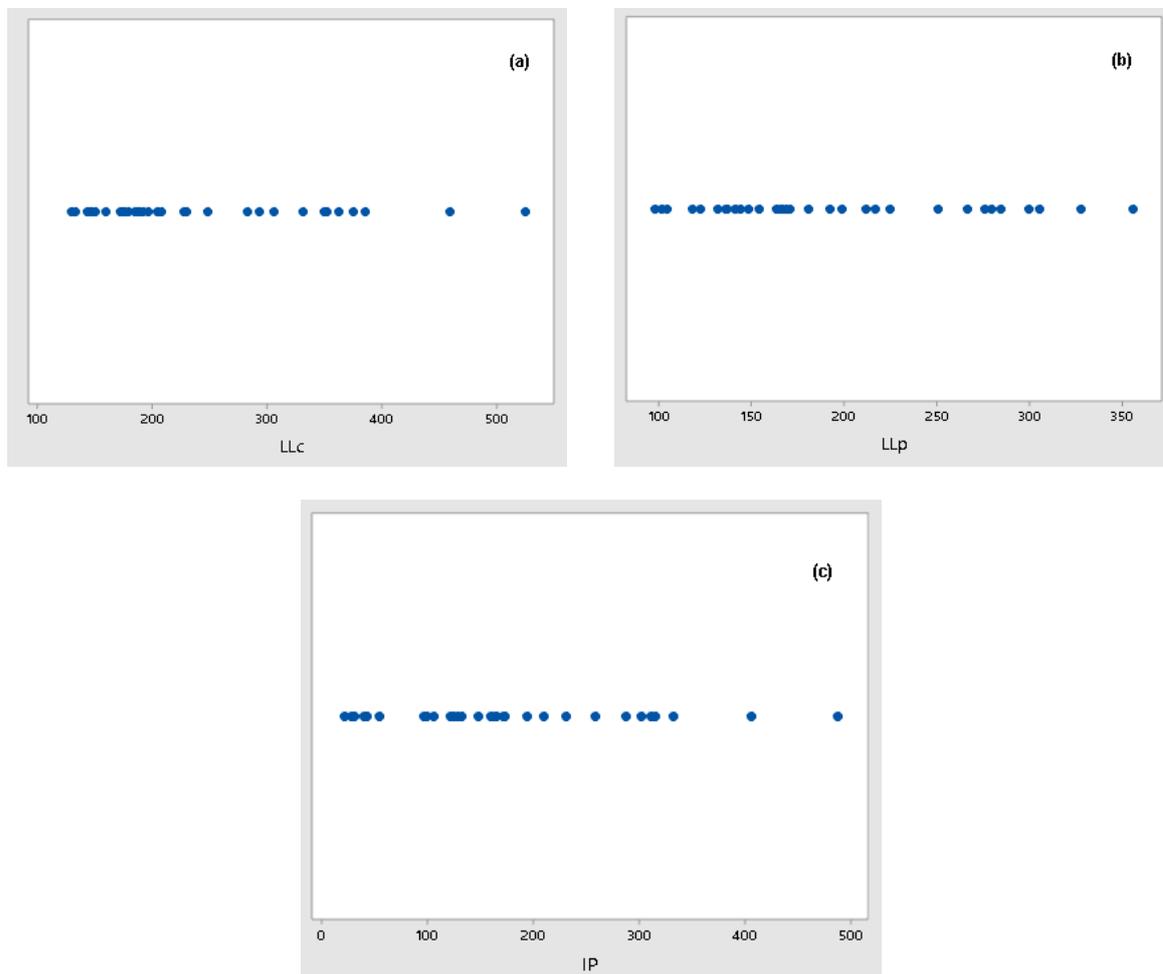


Fonte: Autor (2022)

Figura 16 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-01ii.

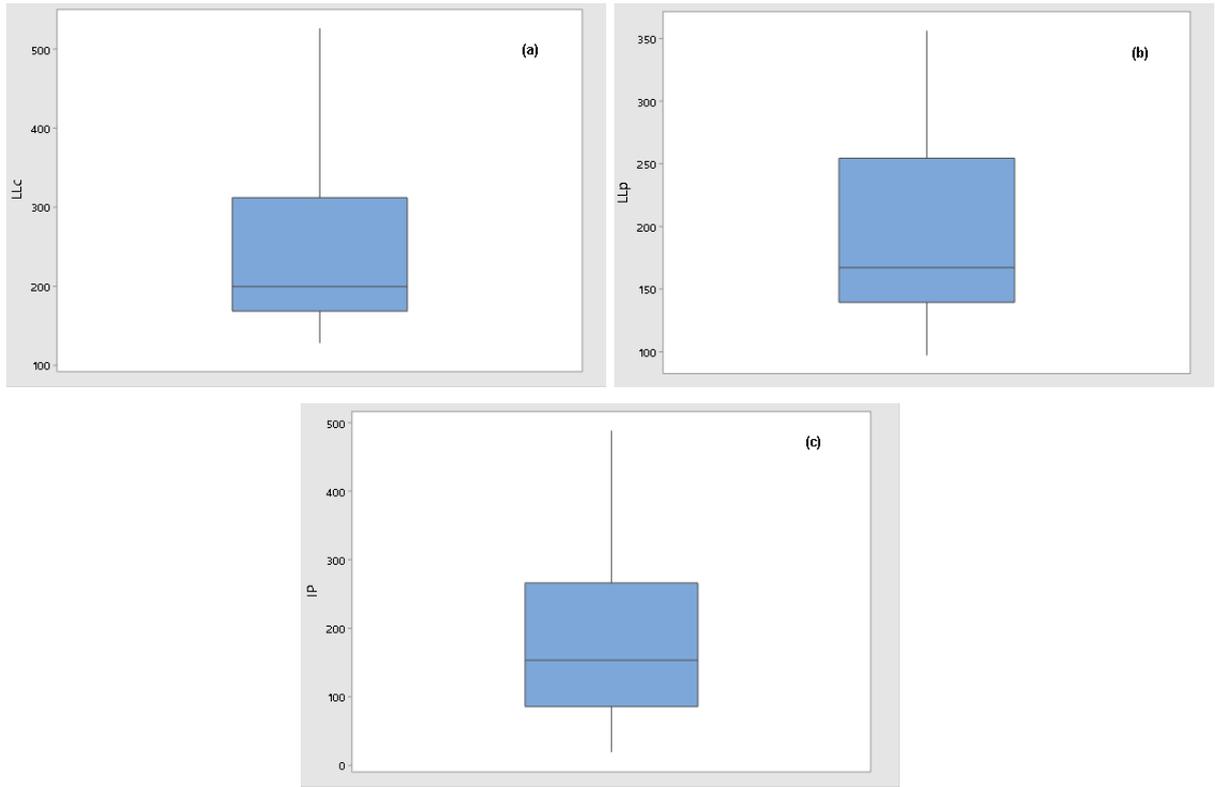


Fonte: Autor (2022)

Gráficos de *outliers* e diagramas *boxplot* para os dados do subgrupo B-01iii**Figura 17 – Gráfico de outliers para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-01iii.**

Fonte: Autor (2022)

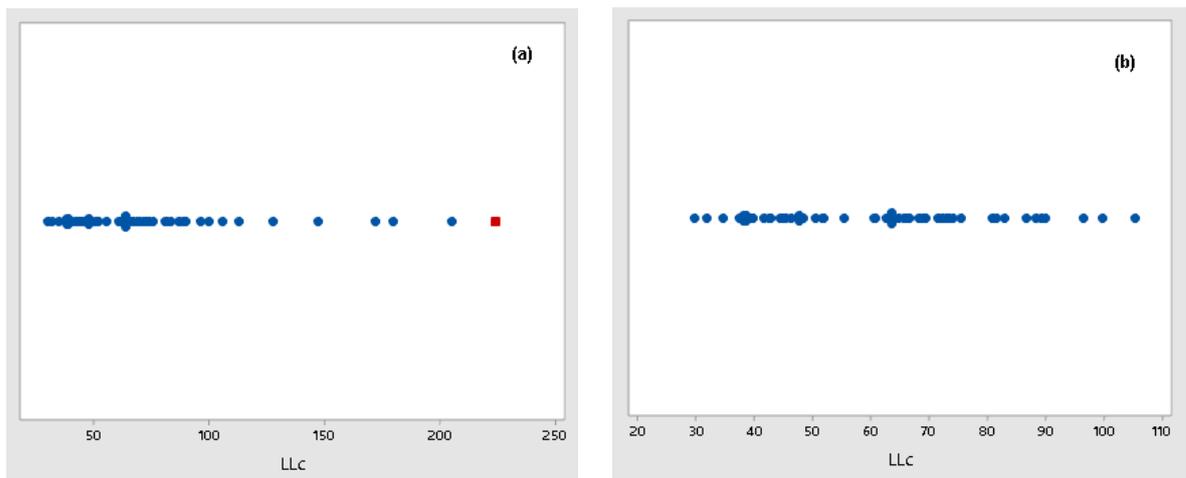
Figura 18 - Diagrama boxplot para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-01iii.



Fonte: Autor (2022)

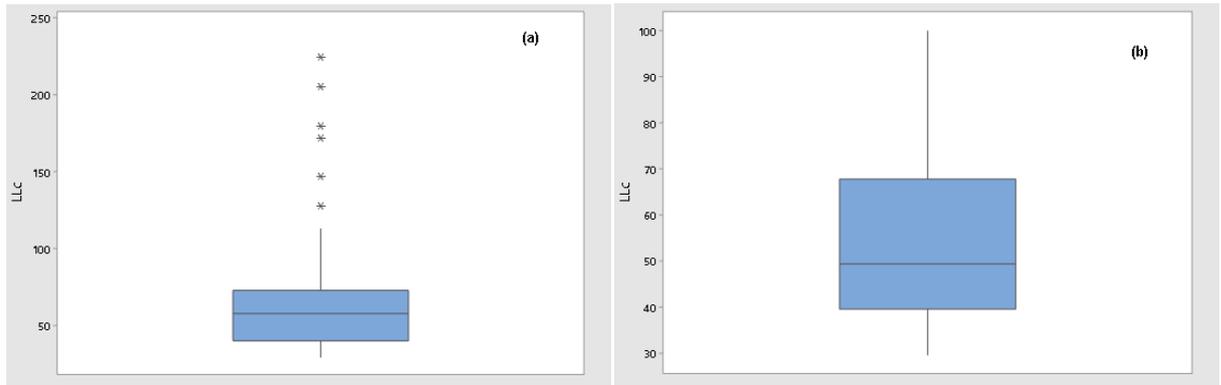
Gráficos de outliers e diagramas *boxplot* para os dados do grupo B-02

Figura 19 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_c do grupo B-02.



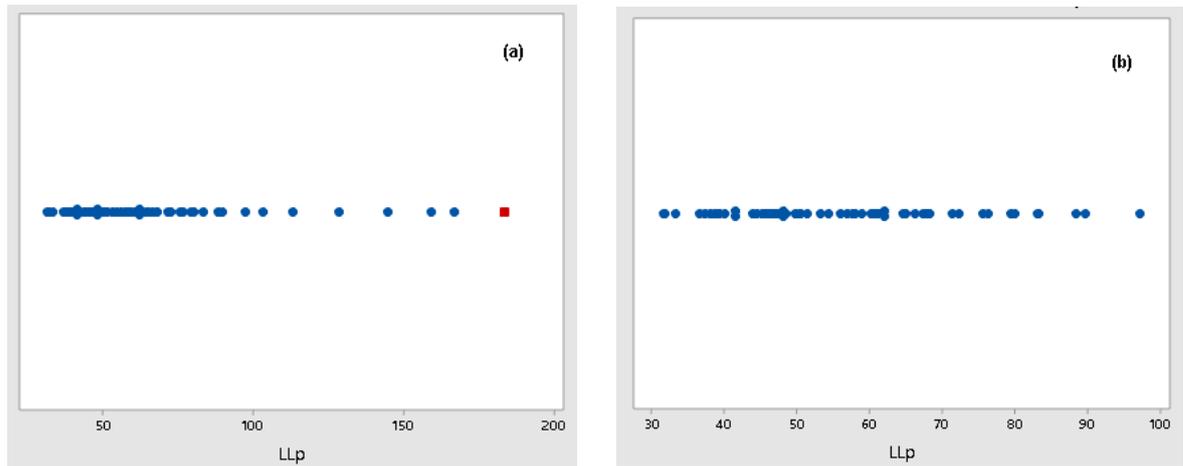
Fonte: Autor (2022)

Figura 20 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_c do grupo B-02.



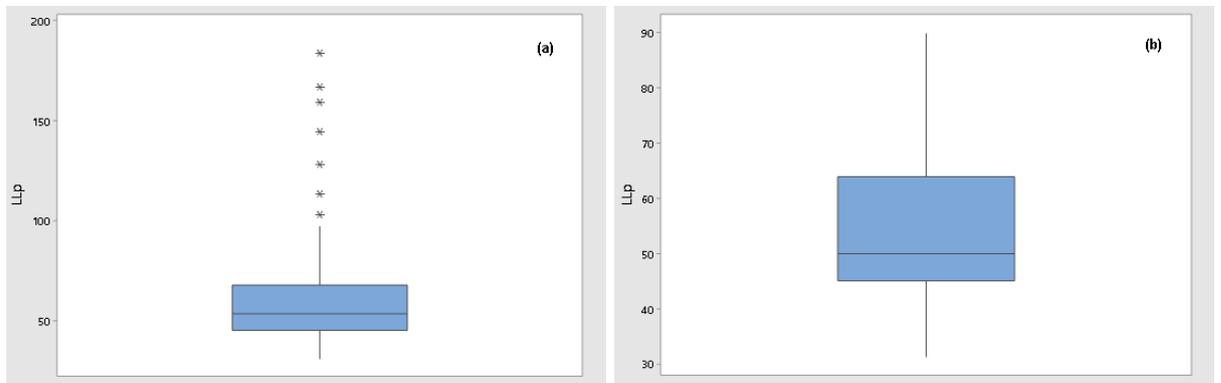
Fonte: Autor (2022)

Figura 21 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_p do grupo B-02.



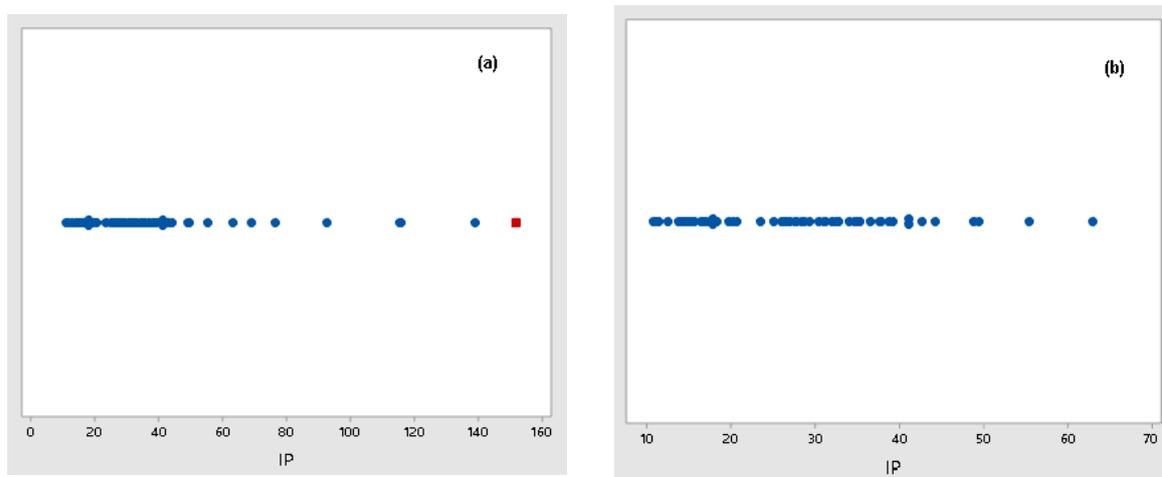
Fonte: Autor (2022)

Figura 22 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_p do grupo B-02.



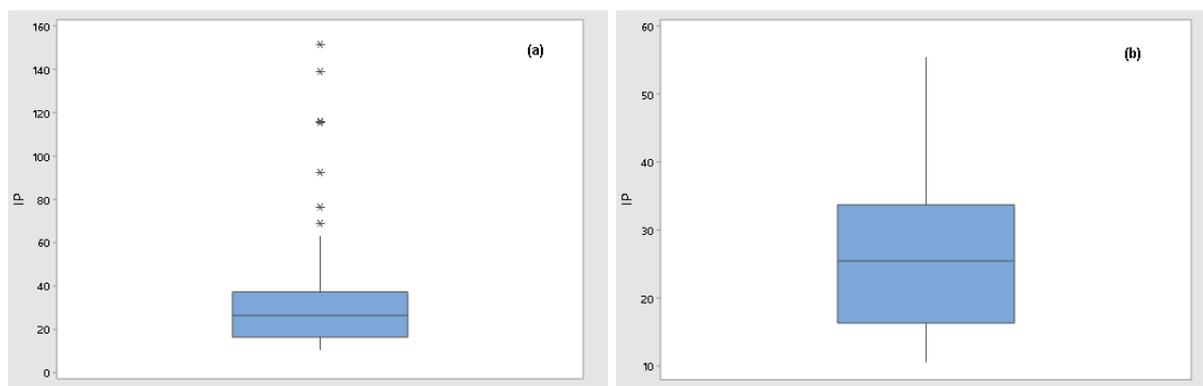
Fonte: Autor (2022)

Figura 23 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade IP do grupo B-02.



Fonte: Autor (2022)

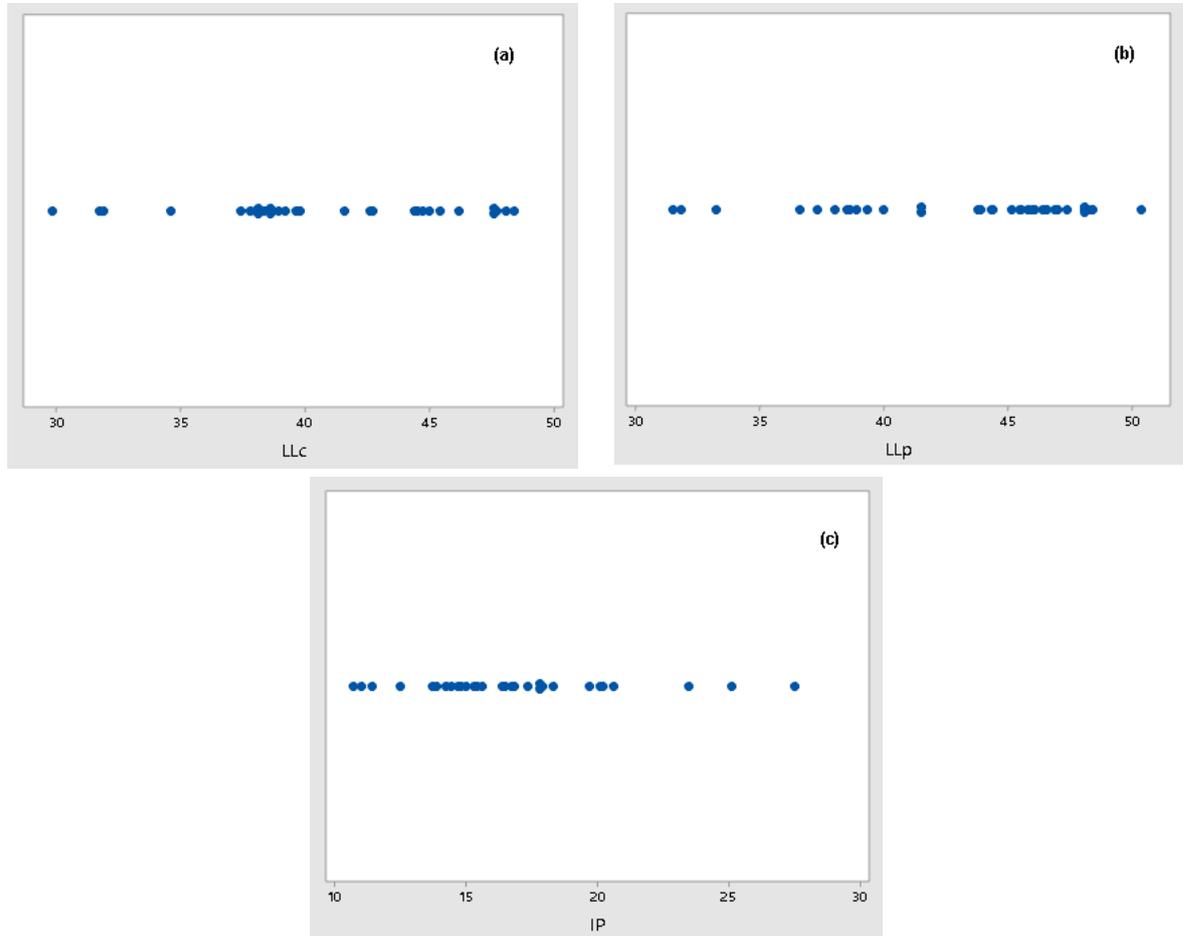
Figura 24 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-02.



Fonte: Autor (2022)

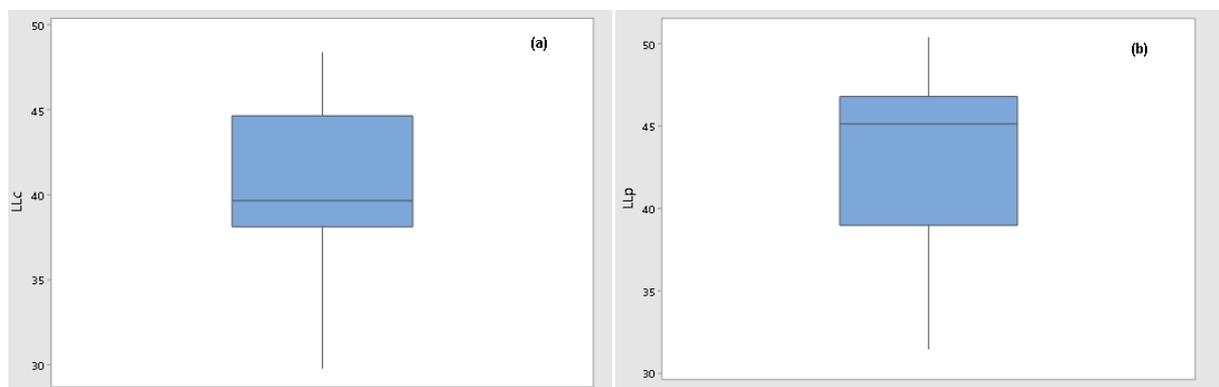
Gráficos de *outliers* e diagramas *boxplot* para os dados do subgrupo B-02i

Figura 25 – Gráfico de outliers para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-02i.



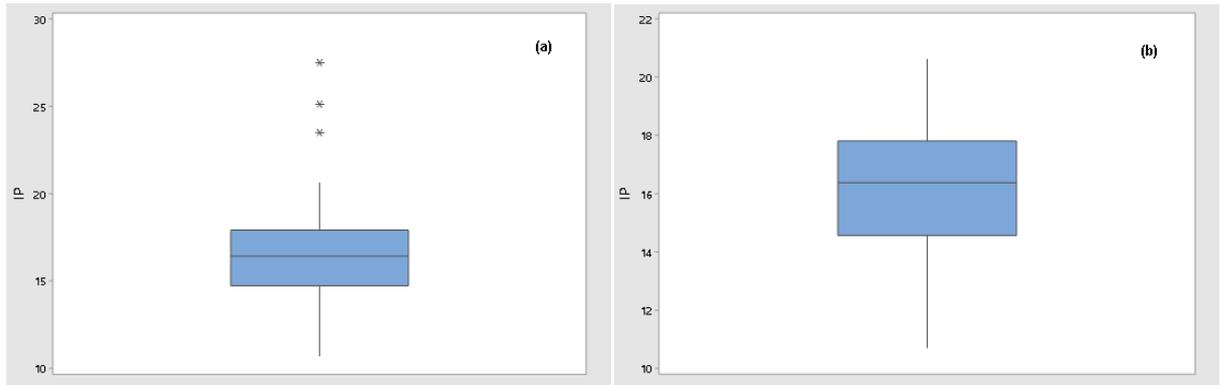
Fonte: Autor (2022)

Figura 26 - Diagrama *boxplot* para as propriedades LL_c (a) e LL_p (b) do subgrupo B-02i.



Fonte: Autor (2022)

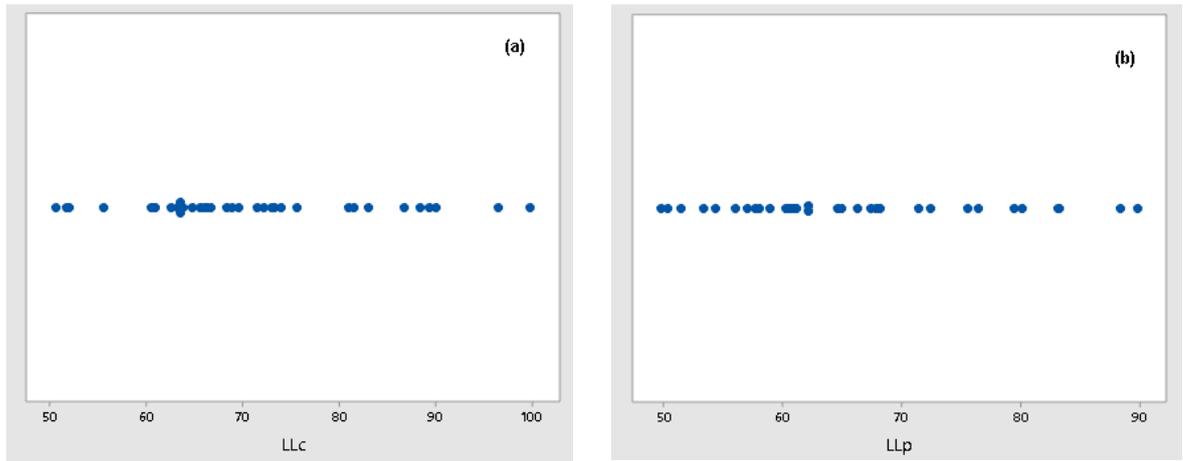
Figura 27 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-02ii.



Fonte: Autor (2022)

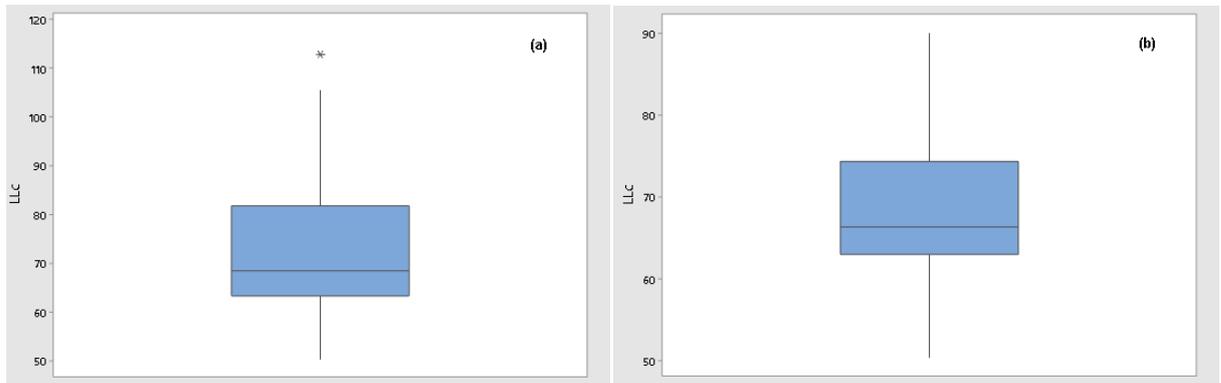
Gráficos de outliers e diagramas *boxplot* para os dados do subgrupo B-02ii

Figura 28 – Gráfico de outliers para as propriedades LL_c (a) e LL_p (b) do subgrupo B-02ii.



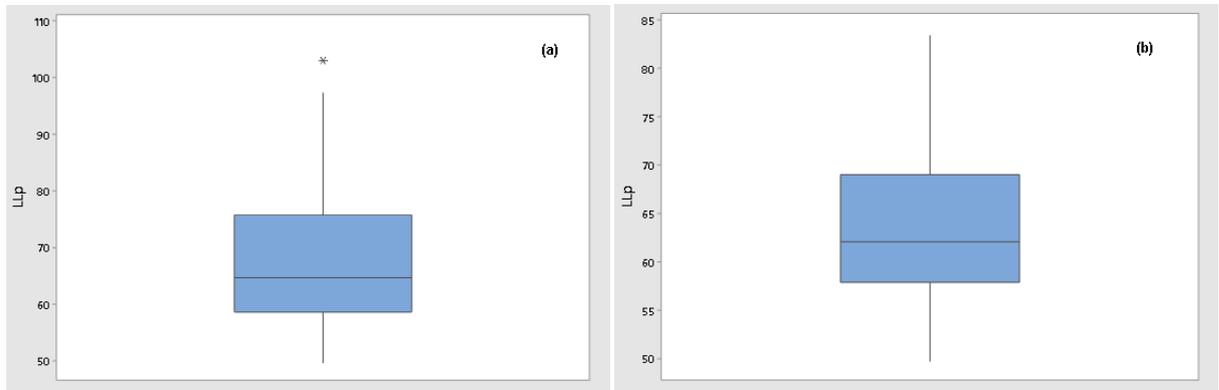
Fonte: Autor (2022)

Figura 29 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_c do grupo B-02ii.



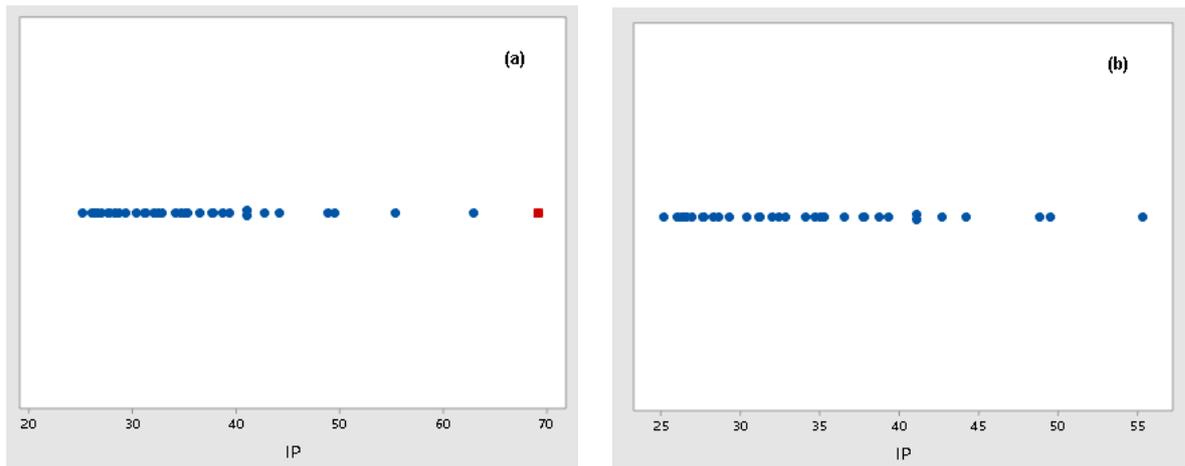
Fonte: Autor (2022)

Figura 30 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_p do grupo B-02ii.



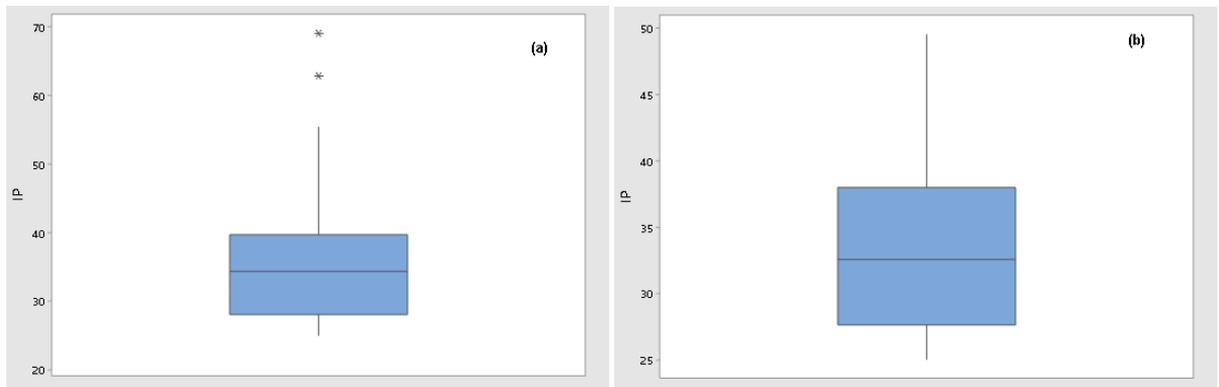
Fonte: Autor (2022)

Figura 31 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade IP do grupo B-02ii.

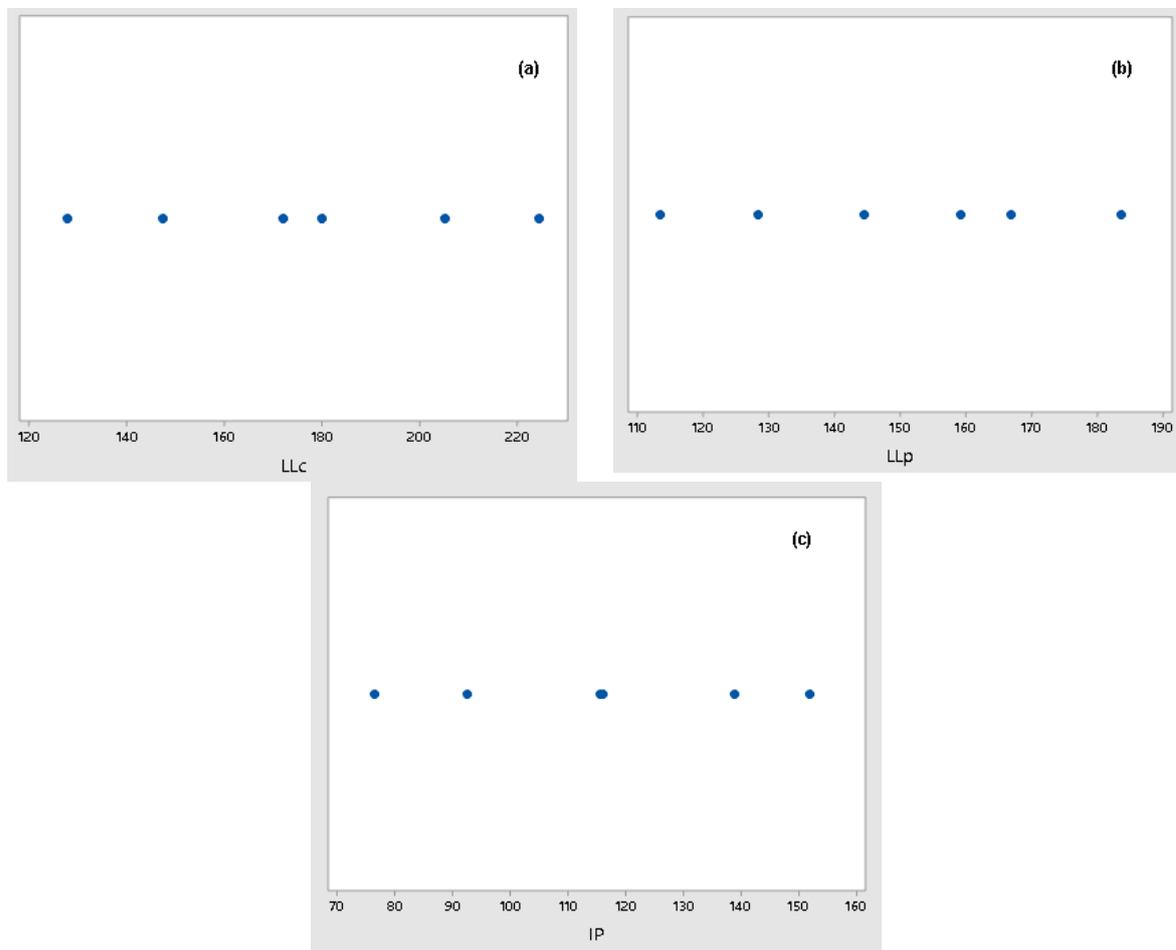


Fonte: Autor (2022)

Figura 32 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-02ii.

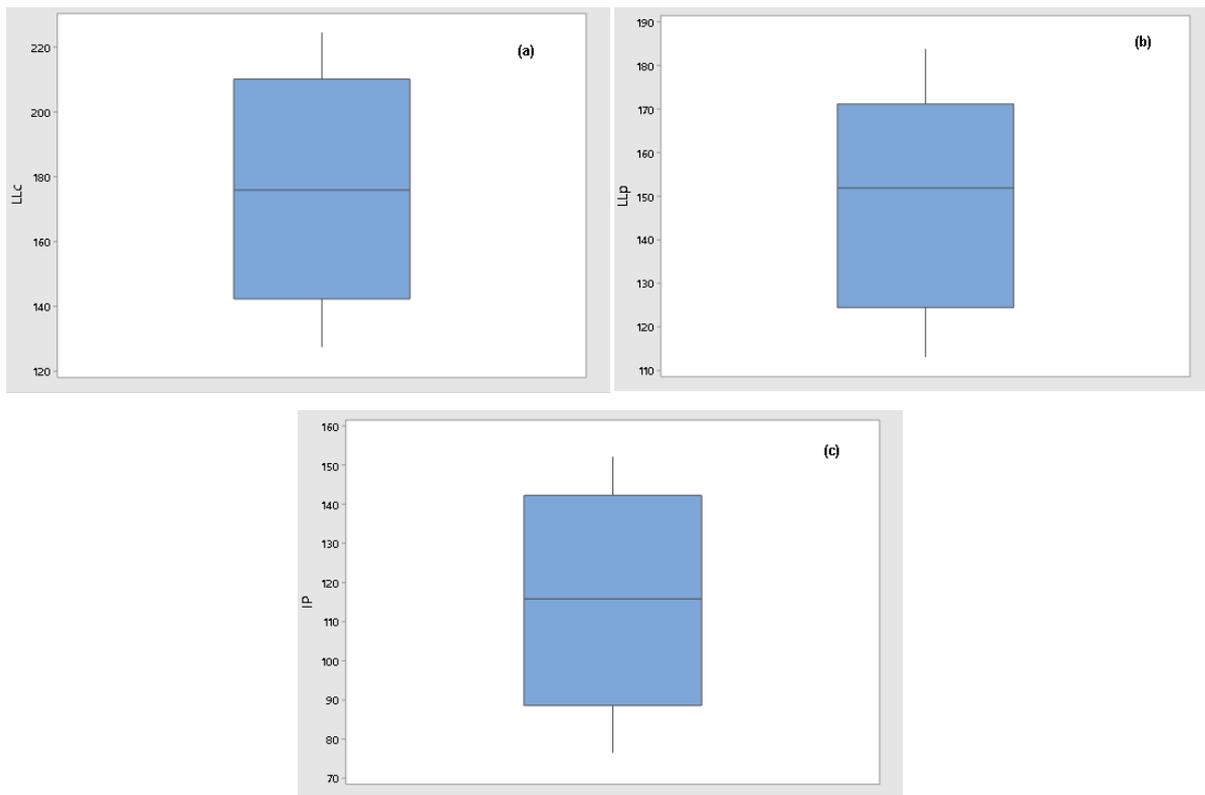


Fonte: Autor (2022)

Gráficos de *outliers* e diagramas *boxplot* para os dados do subgrupo B-02iii**Figura 33 – Gráfico de outliers para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-02iii.**

Fonte: Autor (2022)

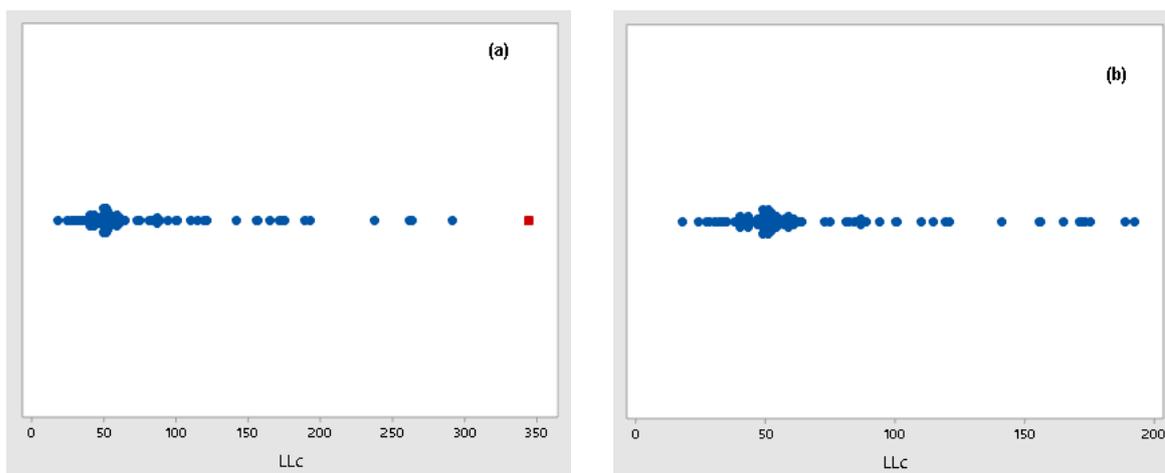
Figura 34 - Diagrama boxplot para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-02iii.



Fonte: Autor (2022)

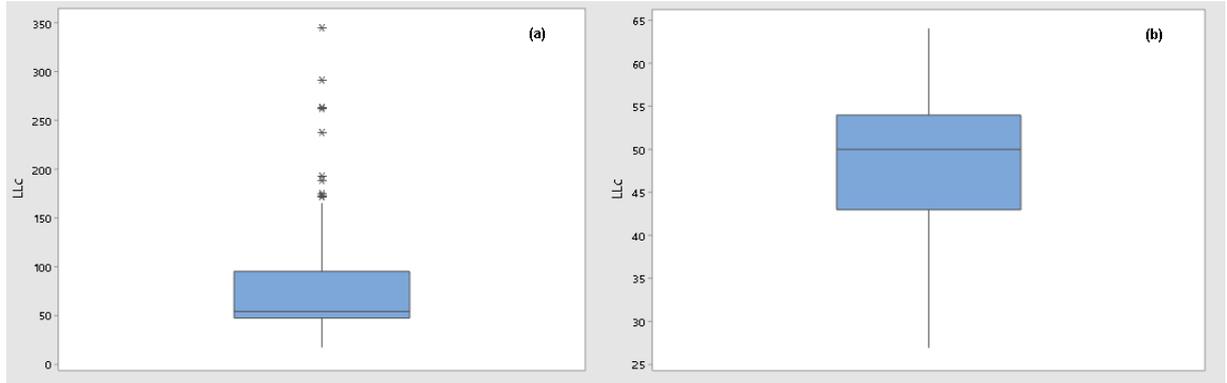
Gráficos de outliers e diagramas boxplot para os dados do grupo B-03

Figura 35 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_c do grupo B-03.



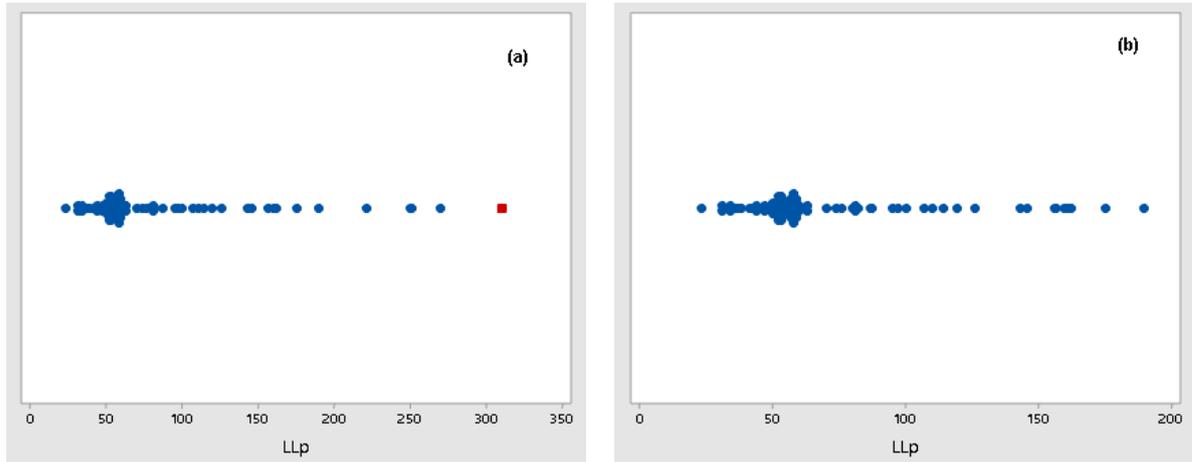
Fonte: Autor (2022)

Figura 36 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_c do grupo B-03.



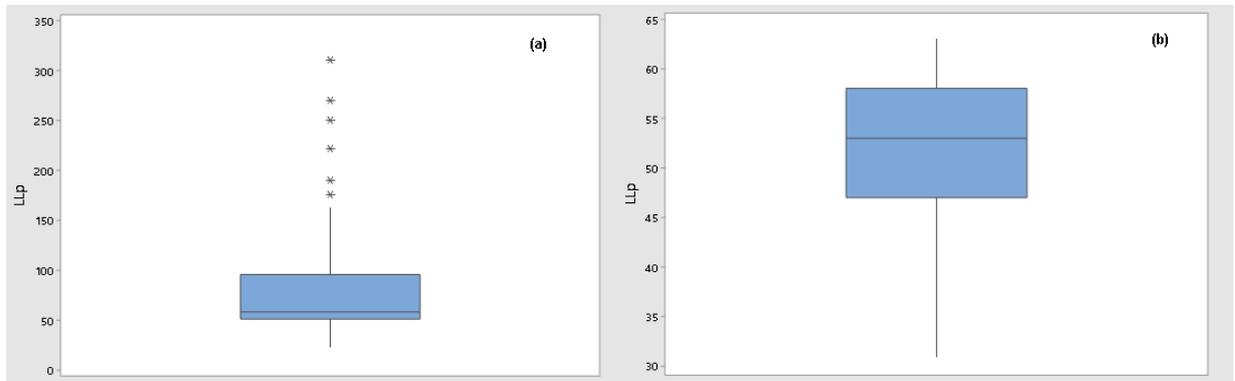
Fonte: Autor (2022)

Figura 37 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade LL_p do grupo B-03.



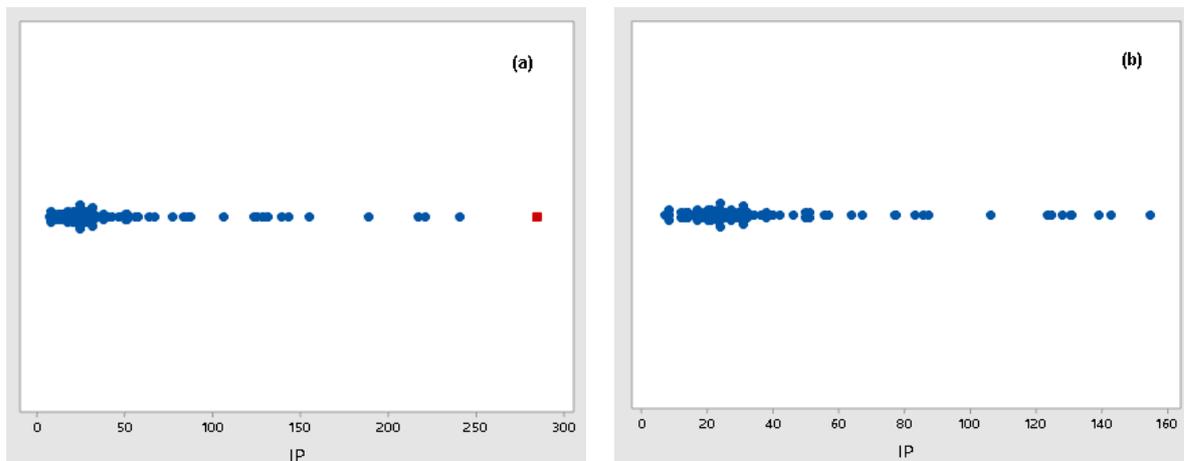
Fonte: Autor (2022)

Figura 38 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_p do grupo B-03.



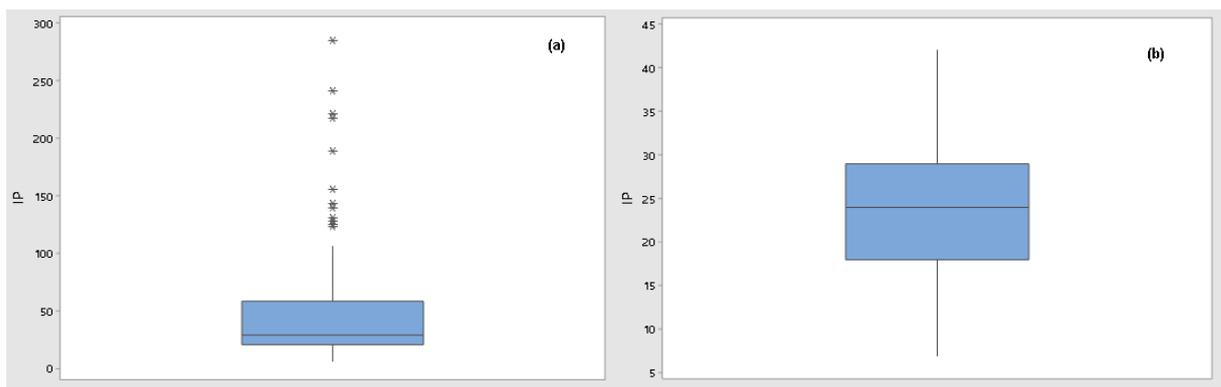
Fonte: Autor (2022)

Figura 39 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no gráfico de outliers para a propriedade IP do grupo B-03.

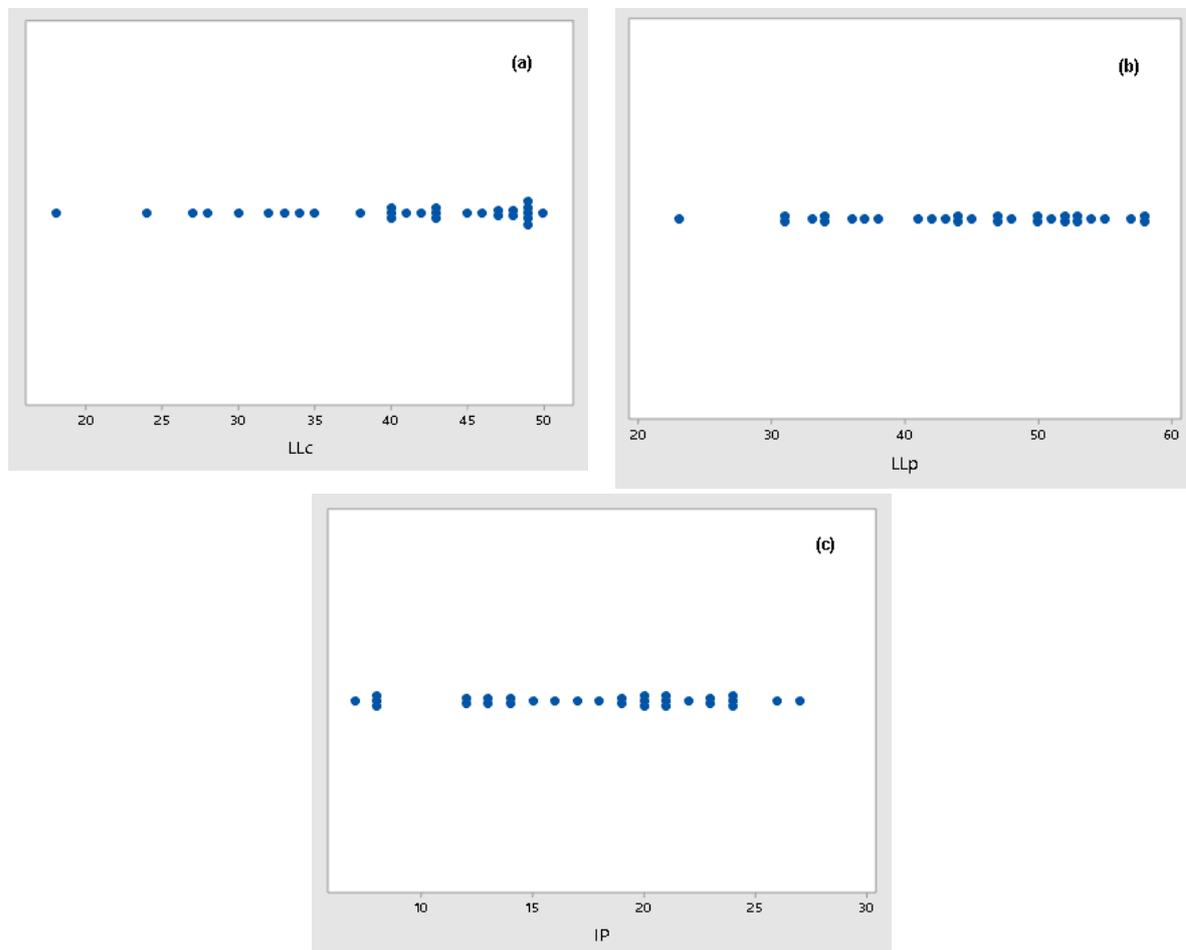


Fonte: Autor (2022)

Figura 40 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-03.

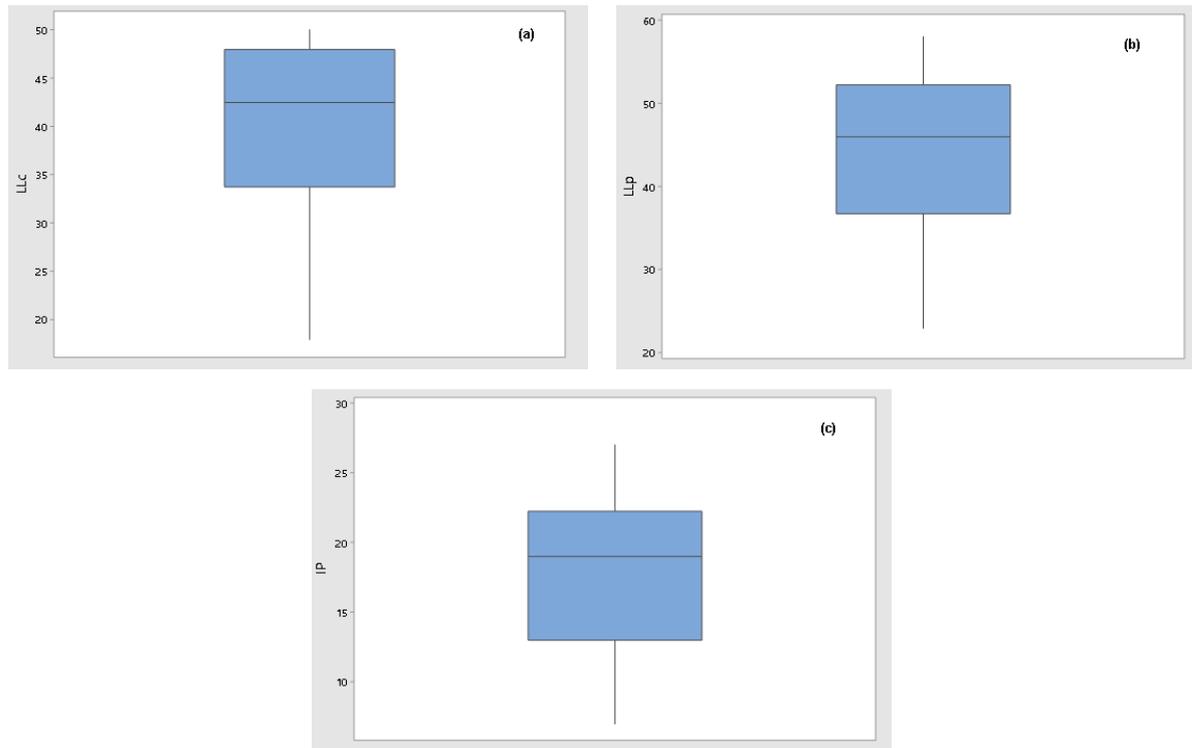


Fonte: Autor (2022)

Gráficos de *outliers* e diagramas *boxplot* para os dados do subgrupo B-03i**Figura 41 – Gráfico de outliers para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-03i.**

Fonte: Autor (2022)

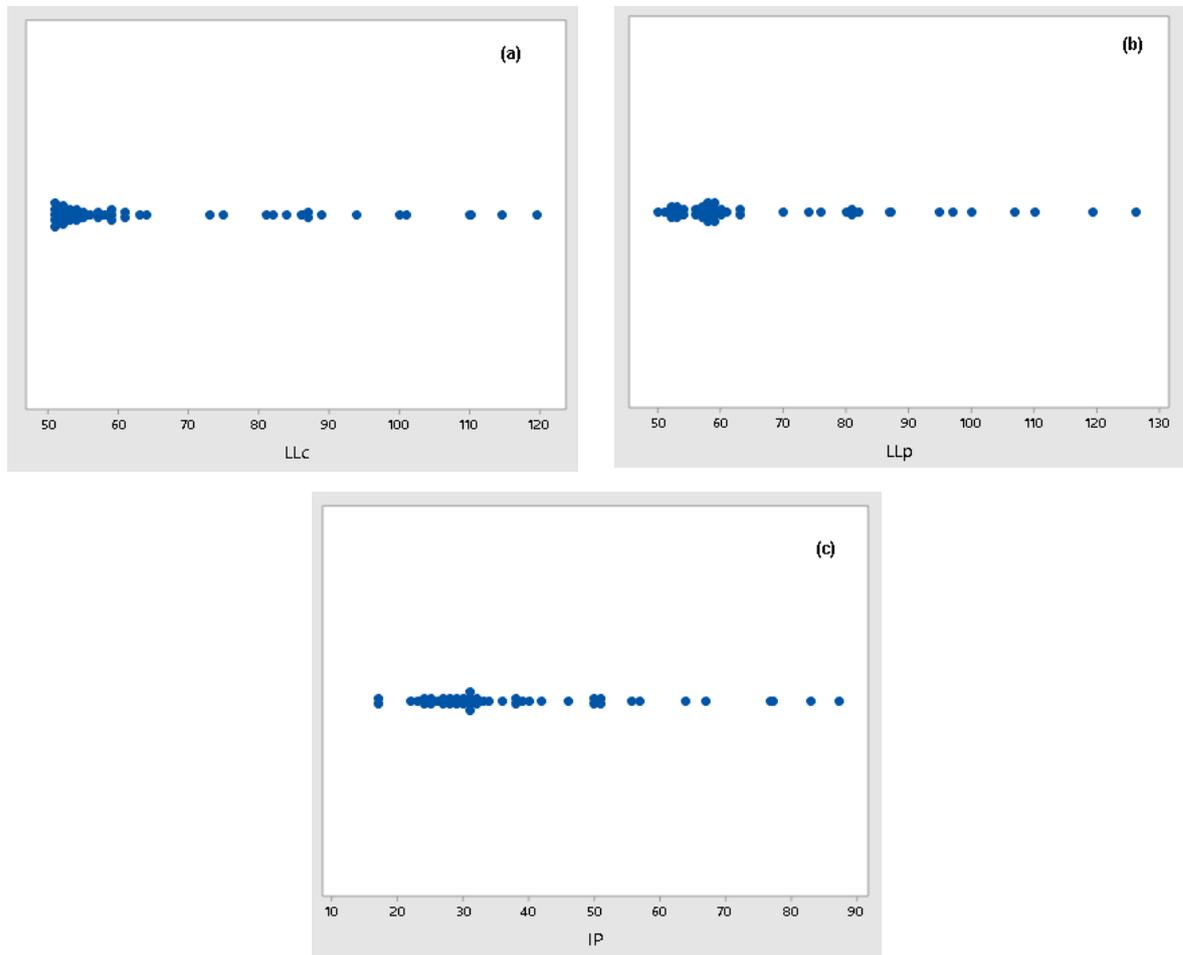
Figura 42 - Diagrama boxplot para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-03i.



Fonte: Autor (2022)

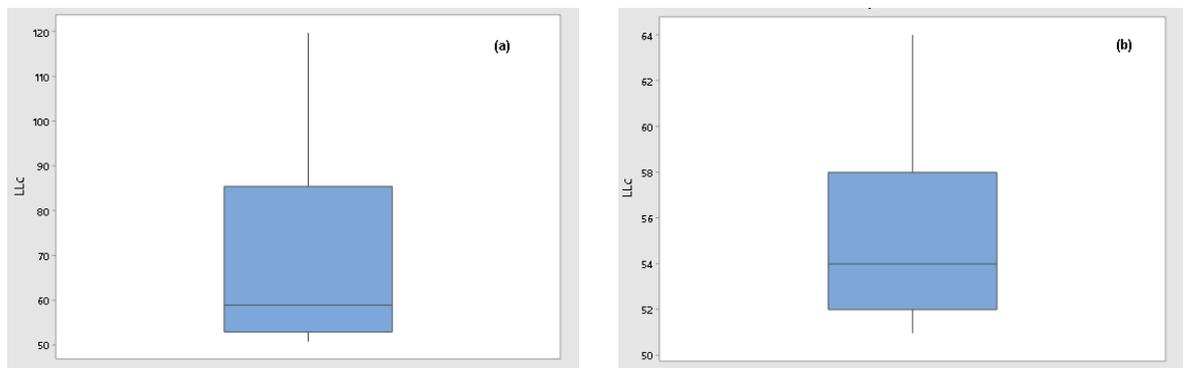
Gráficos de *outliers* e diagramas *boxplot* para os dados do subgrupo B-03ii

Figura 43 – Gráfico de outliers para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-03ii.



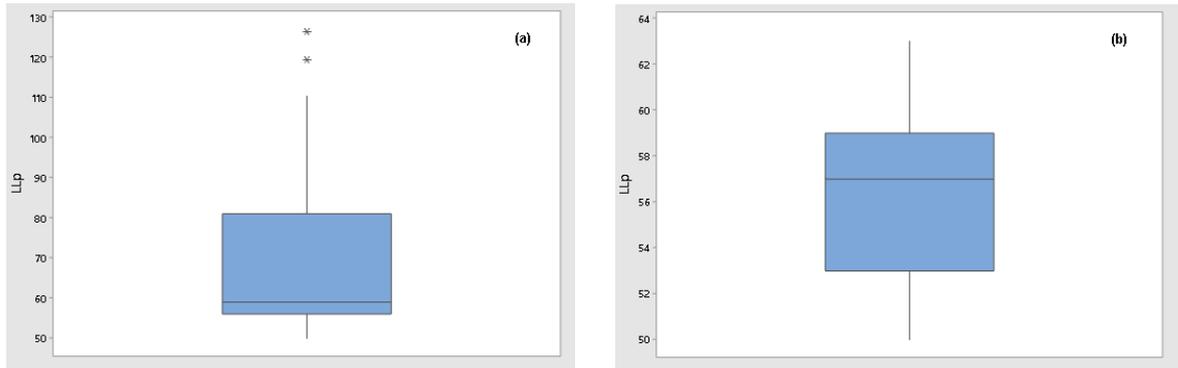
Fonte: Autor (2022)

Figura 44 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_c do grupo B-03ii.



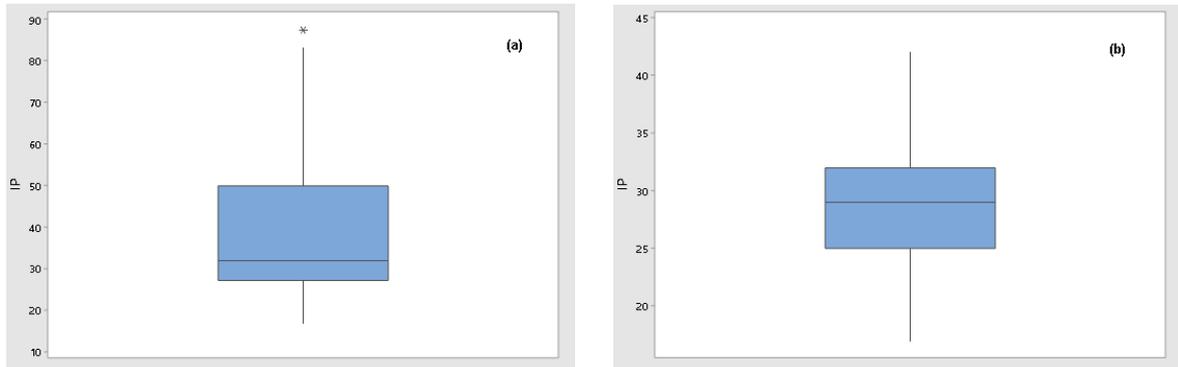
Fonte: Autor (2022)

Figura 45 – Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade LL_p do grupo B-03ii.

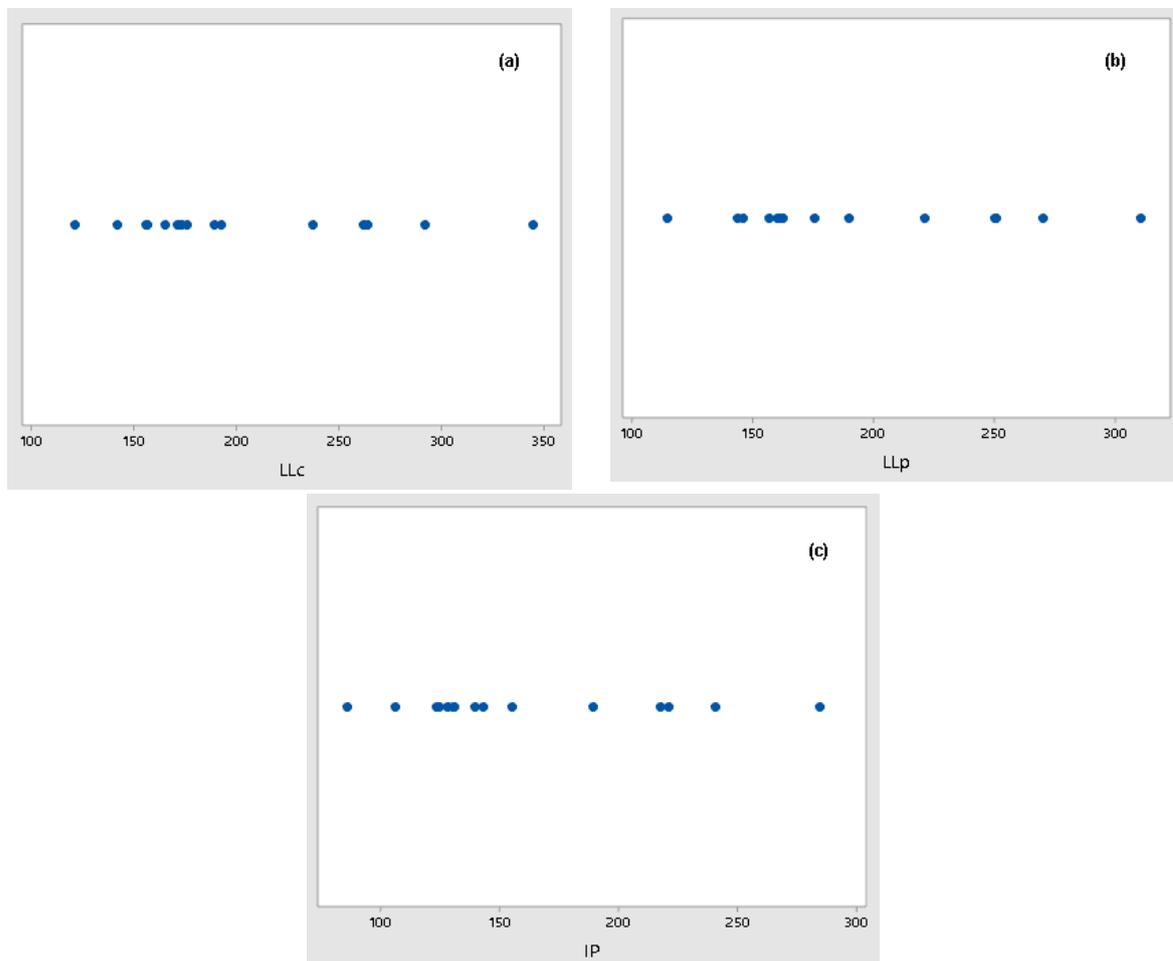


Fonte: Autor (2022)

Figura 46 - Amostras de solo antes (a) e após (b) a remoção dos outliers identificados no boxplot para a propriedade IP do grupo B-03ii.

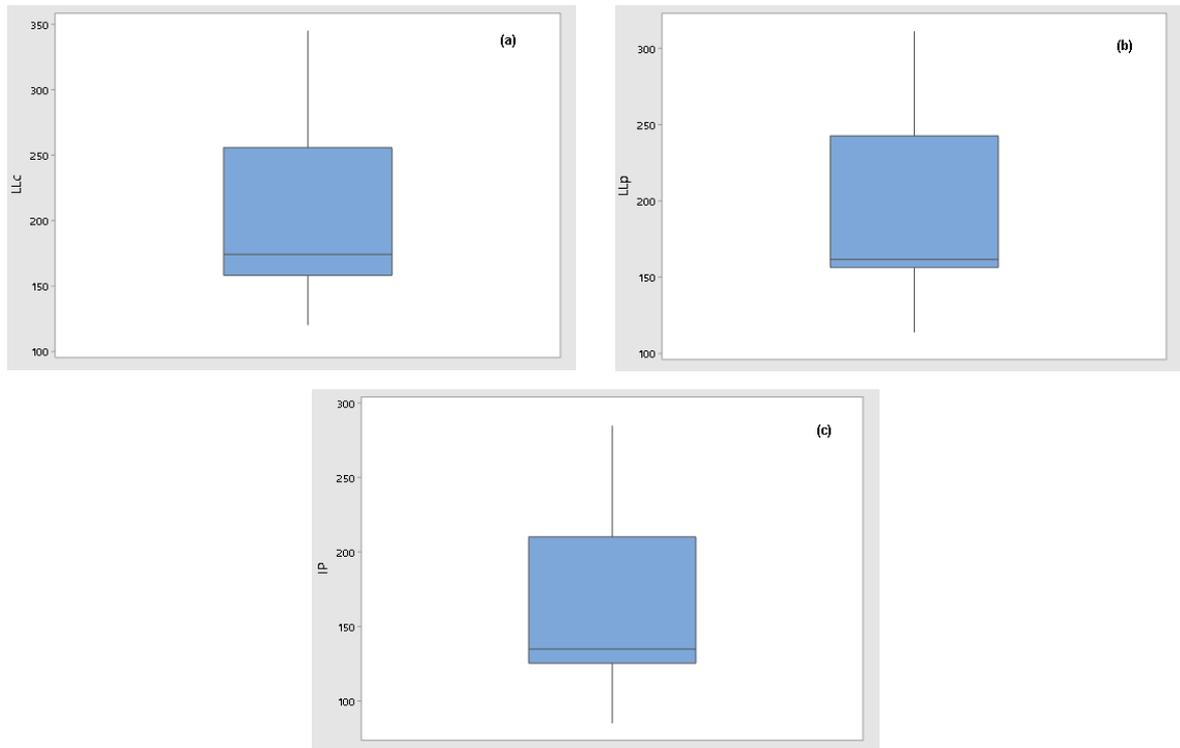


Fonte: Autor (2022)

Gráficos de *outliers* e diagramas *boxplot* para os dados do subgrupo B-03iii**Figura 47 – Gráfico de outliers para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-03iii.**

Fonte: Autor (2022)

Figura 48 – Diagrama *boxplot* para as propriedades LL_c (a), LL_p (b) e IP (c) do subgrupo B-03i.



Fonte: Autor (2022)