

INTRODUÇÃO À ANÁLISE DE SOLOS

SUMÁRIO

1-	POLUENTES NO SOLO	3
2-	TIPOS DE SOLOS	6
3-	A IMPORTÂNCIA DO SOLO	14
4-	SALINIZAÇÃO	18
5-	MÉTODO DE CALIBRAÇÃO MULTIVARIADA	20
REFERÊNCIAS		

1- POLUENTES NO SOLO

A **Poluição do Solo** é toda e qualquer mudança em sua natureza (do solo), causada pelo contato com produtos químicos, resíduos sólidos e resíduos líquidos, os quais causam sua **deterioração** ao ponto de tornar a terra inútil ou até gerar um risco a saúde.

Ora, devemos saber também que o solo é repleto de vida, especialmente sua camada inicial (15 centímetros), onde encontramos os fungos, bactérias, protozoários e vermes decompositores, responsáveis pelo equilíbrio entre os diversos níveis tróficos.

Vale citar que o solo se forma pela desagregação de rochas e a decomposição de restos vegetais e animais, por meio da ação dos referidos agentes decompositores e outras intempéries (chuva, ventos, etc.).

Por sua vez, é justamente essa camada a mais afetada pelos resíduos sólidos e líquidos, fertilizantes químicos, pesticidas e herbicidas, a maioria frutos da química inorgânica desenvolvida após a Segunda Guerra Mundial.

Principais Causas



Depósito de lixo não autorizado

De modo mais geral, podemos afirmar que solventes, detergentes, lâmpadas fluorescentes, componentes eletrônicos, tintas, gasolina, diesel e óleos automotivos, bem como fluídos hidráulicos, hidrocarbonetos e o chumbo são os principais agentes poluidores do solo.

Sabemos, ainda, que o condicionamento inadequado do lixo doméstico, esgoto e resíduos sólidos industriais degradam a superfície, além de produzirem **gases tóxicos** e **chuva ácida** (a qual também se infiltra no solo).

Principais Tipos de Poluição do Solo

- **Detritos da vida urbana** - Em quantidade é a principal fonte causadora da poluição dos solos. É responsável pela produção exacerbada de lixo nas grandes cidades.
- **Depósitos ilegais de despejos industriais** - É fato conhecido que as indústrias fazem uso desse recurso e descartam indevidamente metais pesados, produtos químicos de alto risco, além de dejetos sólidos.
- **Agrotóxicos e adubação incorreta** - Nas áreas rurais, por sua vez, os principais vilões são a utilização indiscriminada de defensivos agrícolas, bem como a

Principais Consequências

Assim, por meio de processos como **lixiviação** (dissolução de componentes sólidos em meio líquido), **impregnação** e **infiltração** no subsolo de substâncias poluentes, ocorre não só a contaminação do terreno, tornando-o inútil, mas também acarreta na contaminação do **lençol freático**.

As principais consequências da poluição do solo são a perda da **fauna**, a esterilização da **terra** para plantação e a contaminação da **água**.

A poluição que se infiltra transmite doenças como infecundidade, hipersensibilidades alérgicas, bem como disfunção hepática e renal ou até câncer.

A contaminação transcende o solo e chega aos alimentos, pois os vegetais se contaminam com aquelas substâncias, as quais, por seu turno, são ingeridas por humanos e outros seres, tornando a alimentação mais tóxica na medida em que vamos expandindo a **cadeia alimentar**.

Outros fatores, como destruição de espécies animais e vegetais úteis (as abelhas polinizadoras, por exemplo), causam um descontrole no **meio ambiente**, permitindo o surgimento de pragas cada vez mais resistentes aos agrotóxicos e inseticidas, o que leva à produção de venenos cada vez mais potentes.

Principais Medidas de Combate

Para reverter à poluição dos solos, medidas simples como a **coleta seletiva e reciclagem** são essenciais.

Outras, como o tratamento de resíduos domésticos e industriais, uso de materiais biodegradáveis, cultivo orgânico, reflorestamento e a proteção das matas nativas, medidas de saneamento básico e controle biológico de pragas, são as práticas mais efetivas de combate a degradação das superfícies.

2- TIPOS DE SOLOS

O **Solo** é a camada superficial da crosta terrestre. Trata-se de um complexo composto de materiais minerais e orgânicos.

Formação e composição do solo

O solo é resultado da ação de vários elementos: água, clima, organismos vivos, relevo, tipo de rocha e o tempo de atuação desses fatores. Em função da ação conjunta dos diversos fatores, originam-se diversos tipos de solo.

A decomposição das rochas por ação dos agentes físicos, químicos ou biológicos dão origem aos componentes minerais. A incorporação e a decomposição de elementos orgânicos animais e vegetais (húmus), dão fertilidade ao solo.

Classificação dos solos

Com relação a **cor**, a maior parte dos solos podem ser agrupadas em três tipos:



- **avermelhados e amarelos** - indicam forte presença de óxido de ferro
- **escuros** - indicam forte presença de materiais orgânicos
- **claros** - indicam a fraca presença ou ausência de materiais orgânicos.

Com relação a **textura** os solos, são classificados:



- **arenoso** - retém pouca água e nutrientes, pois possuem grandes poros, facilitando o escoamento da água
- **argiloso** - o solo argiloso retém mais água e nutrientes (cálcio, potássio, ferro)
- **orgânico** - é composto de materiais orgânicos em processo de decomposição, além de areia e argila

Solo no Brasil

Entre os solos mais comuns encontrados no Brasil, destacam-se o massapê e a terra roxa:

- **Massapê** - é um solo escuro, argiloso e orgânico, originado da desagregação e decomposição da rocha gnaisse. Aparece em grande trecho do Nordeste Brasileiro, na região chamada de Zona da Mata, onde desde o século XVI se cultivava a cana-de-açúcar, que se adapta muito bem a esse tipo de solo.

- **Terra roxa** - é um solo avermelhado e vulcânico, originado da decomposição do basalto. Aparece no oeste do estado de São Paulo e no norte do Paraná. É excelente para a agricultura e, desde o século passado é utilizado para a cultura do café.

Curiosidade: Você Sabia?

A terra roxa, na verdade é vermelha e não roxa. O nome surgiu com os colonos italianos das lavouras do café que falavam "terra rossa", que em italiano quer dizer vermelho. O povo confundiu rossa com roxa.

Solo Arenoso

O **Solo Arenoso**, chamado também de "solo leve", é um tipo de solo muito presente na região nordeste do Brasil.

Ele possui uma textura leve e granulosa, sendo composto, em grande parte, por **areia** (70%) e, em menor parte, por argila (15%).



Por esse motivo, construções em terreno de solo arenoso próximas aos lençóis freáticos, podem provocar rachaduras na estrutura, devido às principais características desse tipo de solo: poroso e permeável.

Interessante notar que estradas construídas nos terrenos de solo arenoso não formam poeira na época seca e não atolam na época das chuvas.

Por outro lado, terrenos com solo argiloso os grãos de areia se aglutinam, tornando as estradas sem pavimentação, barrentas na época das chuvas e com solo duro no tempo de seca.

Principais Características do Solo Arenoso

As principais características desse tipo de solo são:

- Consistência granulosa (grãos grossos, médios e finos)
- Alta porosidade e permeabilidade
- Pouca umidade
- Seca rapidamente
- Pobre em nutrientes e água
- Deficiência em cálcio
- pH ácido e baixo teor de matéria orgânica
- Presença de grandes poros (macroporos) entre os grãos de areia
- Dificulta a sobrevivência de plantas e organismos
- Altamente suscetíveis à erosão

Medidas para Utilização dos Solos Arenosos

Para a utilização sustentável do solo arenoso na agricultura, faz-se necessário a adoção de práticas de conservação como o cuidado do solo por meio de técnicas de manejo, utilização do sistema de plantio direto, integração lavoura-pecuária, rotação de culturas, adubação verde (adubo orgânico), dentre outros.

Visto que esse tipo de solo apresenta grande porosidade e, conseqüentemente, perigo de erosão, especialistas apontam para a importância da rotação de culturas.

No tocante à pobreza de nutrientes dos solos arenosos, é indicado a aplicação de resíduos vegetais e adubos orgânicos (bagaço de cana, bagaço de coco e esterco de animais) com fosfato e potássio, já que para a acidez do solo, recomenda-se a adição de calcário.

Solo Argiloso

O **solo argiloso**, chamado de “**solo pesado**”, é uma terra úmida e macia, composto por mais de 30% de argila, alumínio e ferro.

Após a chuva, os terrenos de solos argilosos, que absorvem bastante água, ficam encharcados. Por outro lado, na época de seca, esse tipo de solo tende a formar uma camada dura e pouco arejada do terreno, prejudicando o desenvolvimento da vegetação.



Características do Solo Argiloso

- Grãos pequenos (microporos) e compactos
- Impermeável a líquidos
- Grande retenção de água
- Alta impermeabilidade
- Grande concentração de nutrientes
- Pouca acidez
- Propício para o cultivo e atividade agrícola
- Mais resistentes à erosão

Terra Roxa



A terra roxa é um dos principais tipos de solo argiloso, de coloração vermelho-roxeada, encontrada nos estados de São Paulo, Mato Grosso do Sul, Paraná, Santa Catarina, Rio Grande do Sul; e também na Argentina chamada de “*tierra colorada*”.

Considerado bastante fértil (presença de diversos minerais), este tipo de solo é muito indicado para a prática da agricultura, principalmente para a cultura de café.

Nesse ínterim, vale lembrar que o nome “terra-roxa” surgiu da influência dos imigrantes italianos que trabalham nas fazendas de café no Brasil nos séculos XIX e XX. Assim, o termo “*rosso*”, que na língua italiana significa “vermelho”, ficou traduzido no português como “terra roxa”.

Curiosidade: Você Sabia?

- Muitos povos antigos, utilizavam a argila (um tipo de barro remodelável) como matéria-prima para a fabricação de utensílios (vasos, recipientes, panelas), telhas, tijolos, cerâmicas, objetos de arte e adorno.

Compartilhar

Solo Orgânico e Inorgânico

O **solo** é a camada que acoberta a superfície terrestre sendo formado basicamente por matéria orgânica e matéria inorgânica (componentes sólidos) através da ação de fatores climáticos e biológicos.

Vale lembrar que, além dos elementos sólidos, o solo é formado por componentes líquidos (água) e gasosos (gás carbônico, hidrogênio, nitrogênio, oxigênio, etc.)

intimamente relacionados com a porosidade necessária para o desenvolvimento do solo.

Poluição do Solo

O solo é muito importante para a sobrevivência do ser humano, dos animais e das plantas, uma vez que todos retiram dele os alimentos necessários para sobreviverem.

No entanto, a poluição do solo a partir da utilização de produtos químicos (fertilizantes, agrotóxicos) e da quantidade de resíduos sólidos e líquidos, têm gerado diversos problemas ambientais, desde a perda de espécies e, consequentemente dos ecossistemas.

Composição do Solo

De acordo com sua composição, há dois tipos de solo: o orgânico e o inorgânico.

Solo Orgânico

Os solos orgânicos são compostos de matéria orgânica, ou seja, formados através da decomposição de vegetais, animais e microrganismos.

O **húmus**, responsável pela fertilidade do solo, é o nome dado a matéria orgânica de coloração escura que fica depositada nesse tipo de solo através das condições aeróbicas, ou seja quando há presença de oxigênio, por exemplo, nos animais vertebrados e invertebrados.

Por sua vez, a **turfa** é o nome dado à matéria orgânica formada por processos anaeróbicos, que ocorrem com a ausência de oxigênio, por exemplo, os fungos e bactérias. É o solo mais apropriado para o desenvolvimento de plantas, sendo muito utilizado na agricultura.

Solo Inorgânico

Ao contrário do solo orgânico, o inorgânico é formado por matéria inorgânica, ou seja, os minerais os quais se formam sobretudo pela desagregação das rochas ao longo do tempo seja pela ação do vento, da chuva, e das alterações de temperatura.

Esses elementos são chamados de coloides inorgânicos, os quais tem funções muito importantes para a desintoxicação do solo.

Os principais minerais que surgem nesse tipo de solo são o calcário, quartzo, mica, argila, dentre outros. Esse tipo de solo não é muito apropriado para agricultura, sendo encontrado, por exemplo, no deserto.

Os compostos inorgânicos são mais abundantes que os compostos orgânicos e ambos são importantes para o desenvolvimento e equilíbrio dos ecossistemas.

Adubos Orgânicos e Inorgânicos

Com a interferência humana e as mudanças climáticas na atualidade, muitos locais do mundo apresentam solos pobres em nutrientes. Para tanto, os adubos, compostos de matéria orgânica e inorgânica, devolvem ao solo os nutrientes necessários.

Assim, os adubos orgânicos são aqueles provenientes da matéria orgânica de origem vegetal ou animal, enquanto os adubos inorgânicos são obtidos através da extração de minerais.

3- A IMPORTÂNCIA DO SOLO

O **solo** corresponde a camada superficial da crosta terrestre, sendo muito importante para o desenvolvimento da vida na terra, visto que dele retiramos os alimentos necessários para nossa sobrevivência.

Note que utilizamos o solo não somente para a produção da alimentação, mas também como matéria prima para diversas construções.

Além disso, o solo possui importantes funções, desde o armazenamento e escoamento e infiltração da água na superfície, sendo um componente fundamental para o desenvolvimento de diversos ecossistemas.



Por esse motivo, o manejo adequado e a preservação do solo tornam-se tarefas essenciais, já que é um recurso natural não renovável, ou seja, é limitado, e a exploração desenfreada pode acarretar muitos problemas futuros.

A Lei Federal 7867 de 13 de novembro de 1989, implementou no Brasil o “Dia Nacional da Conservação Solo” que é comemorado todo dia 15 de abril.

A data alerta para a importância desse recurso tão importante para o desenvolvimento da vida na Terra, assim como a água, o ar, a fauna e a flora.

Composição e Tipos de Solo

O solo é um complexo composto de minerais e matéria orgânica oriundo de um lento processo decorrente da degradação de rochas e da decomposição de diversos animais e plantas.

Existem diversos tipos de solo, resultantes da ação de elementos como a água, o clima e o relevo. Assim, os principais tipos de solo, são classificados em:

A Importância do Solo para o Ser Humano

Na vida humana, o solo participa quase que inteiramente pois dele retiramos os alimentos necessários para nossa sobrevivência.

Além disso, utilizamos esse recurso na construção civil, ou seja, na construção de casas, edifícios, dentre outros.

A Importância do Solo para os Animais

Para os animais, tanto quanto para os seres humanos, o solo é um recurso muito importante de desenvolvimento, pois é dele que eles retiram os alimentos para sobreviverem.

A Importância do Solo para a Agricultura

O solo é fundamental na composição do ecossistema terrestre, pois é dele que as plantas retiram todos os nutrientes necessários para se desenvolverem.

O tipo de solo é muito importante para as plantações e o desenvolvimento da agricultura. Nesse sentido, não são todos os solos que auxiliam na reprodução de plantas. Isso porque há solos pobres de nutrientes, os quais impedem o desenvolvimento da flora.

Para melhorar os problemas ambientais causados no solo, a produção sustentável de alimentos através da agricultura biológica tem sido uma boa alternativa.

Os alimentos orgânicos são produzidos pela ausência de produtos químicos (agrotóxicos, inseticidas, adubos químicos). O sistema de agricultura mais utilizado afeta não somente o solo, empobrecendo-o, mas também podem gerar diversas doenças nos seres humanos.

Poluição do Solo

A poluição do solo pode ocorrer de diversas maneiras, sendo que a ação humana tem sido um importante fator de degradação.

As queimadas, o desmatamento, o desenvolvimento de pastos (para animais) ou plantações e a contaminação dos recursos hídricos (água) geram diversos problemas ambientais, como a erosão, que afeta diretamente o solo, desequilibrando os ecossistemas.

Algumas medidas importantes devem ser tomadas, por exemplo:

- diminuição de queimadas e do uso de agrotóxicos na agricultura;
- reflorestamento de determinadas zonas;
- não jogar lixo e produtos químicos (geralmente causado pelas indústrias) em locais inapropriados;
- fazer a separação dos resíduos segundo a coleta seletiva.

4- SALINIZAÇÃO

A **salinização do solo** é um processo de acumulação de sais minerais (Na^+ , Ca_2^+ , Mg_2^+ , K^+ , etc.) na terra. Esse aumento de concentração de sais prejudica as propriedades do solo e consequentemente o crescimento das plantas.



A salinização gera muitos impactos negativos no meio ambiente, impedindo o desenvolvimento da agricultura e a proliferação das espécies que habitam o local, diminuindo assim, a biodiversidade.

Causas

Embora muitas causas de salinização do solo sejam de ordem natural, como a baixa pluviosidade ou ação das marés em áreas costeiras, esse processo pode ser intensificado com a atividade humana e o manejo incorreto do solo, desde o uso de fertilizantes, irrigação com água rica em sais e contaminação do solo.

O alto índice de evaporação, por exemplo, em locais áridos e semiáridos, os quais apresentam altas temperaturas, pode acelerar esse processo. Note que a água possui uma quantidade de sais e quando ocorre a evaporação, ela evapora, no entanto, os sais ficam retidos no solo.

Consequências

Com o excesso de salinização do solo a terra torna-se imprópria, infértil e improdutiva para o desenvolvimento das espécies vegetais e animais.

A salinização, como dito acima, reflete diretamente no equilíbrio do ecossistema, levando a perda da biodiversidade local, tornando a terra imprópria para uso e diminuindo as áreas de produção agrícola. Esse fator causa grande impacto no ambiente bem como nas populações que o habitam.

Outros processos prejudicam diretamente a fertilidade com o uso inadequado do solo tais como: a compactação, a erosão, a desertificação e a sedimentação.

Compactação do Solo

A compactação do solo corresponde a perda de porosidade natural da terra, dificultando a entrada da água. Da mesma maneira que a salinização o solo, o processo de compactação torna-o impróprio para a prática da agricultura.

Entretanto, esse processo é causado sobretudo, pelo uso de máquinas agrícolas e pela presença de animais, de forma que o peso deles acaba compactando a terra cada vez mais. O processo de compactação pode levar a erosão das áreas afetadas.

Desertificação do Solo

O processo de salinização do solo pode levar a desertificação do local, ou seja, formação e expansão de desertos. Por isso, as regiões mais afetadas pelo processo de desertificação são as zonas áridas e semiáridas, donde o índice pluviométrico é baixo.

O uso inadequado do solo, o desmatamento e as queimadas para o uso agrícola tem sido as principais atividades que resultam na desertificação.

Erosão do Solo

A erosão é um processo natural provocado pela ação das chuvas e do vento. Ela ocorre da seguinte forma: pelo desgaste do solo, o transporte de partículas pela água e, por fim, a deposição desses sedimentos nas áreas mais baixas do relevo, tal qual o leito dos rios.

Sedimentação do Solo

A sedimentação revela o processo de desgaste das rochas e dos solos que ocorre sobretudo, pela ação das águas e das massas de ar. Nesse sentido, está

intimamente relacionado com a erosão, no entanto, os sedimentos são os produtos da atividade erosiva.

Salinização do Solo no Brasil

No Brasil, o processo de salinização tem afetado diretamente as áreas da região nordeste do país, as quais estão inseridas no clima semiárido que apresenta baixa pluviosidade facilitando o acúmulo de sais.

Além disso, áreas litorâneas que sofrem com a ação das marés têm potencializado ainda mais esse processo. Vale lembrar que esse processo é natural, todavia, as ações humanas têm aumentado as áreas de solos inférteis.

5- MÉTODO DE CALIBRAÇÃO MULTIVARIADA

A medida quantitativa de uma propriedade de interesse em um sistema químico é uma das práticas mais frequentes em química analítica. Na grande maioria dos casos, essa propriedade trata da concentração de uma espécie presente no meio. Contudo, quando se utilizam técnicas instrumentais de análise, a concentração é uma propriedade determinada indiretamente por meio de relação com outra grandeza, como absorção ou emissão de luz, medidas de condutividade ou potencial elétrico, numa prática conhecida como calibração¹.

O processo de calibração pode ser definido como uma série de operações que estabelecem, sob condições específicas, uma relação entre medidas instrumentais e os valores correspondentes à propriedade de interesse realizados em padrões². Entre os métodos de calibração existentes, sem dúvida, os mais difundidos são os métodos univariados, onde se tem apenas uma medida instrumental para cada uma das amostras de calibração. Esses métodos são relativamente fáceis de serem aplicados, validados e se encontram descritos na literatura em vários trabalhos³⁻⁶. Porém, a aplicação de modelos univariados é restrita a situações em que a grandeza que é medida diretamente no sistema (ex.: absorbância de um composto) é livre de interferentes que possam provocar desvios entre sua relação linear com a propriedade de interesse.

Em métodos de Calibração Multivariada duas ou mais respostas instrumentais são relacionadas com a propriedade de interesse. Esses métodos possibilitam análises mesmo na presença de interferentes, desde que estejam presentes nas amostras de calibração, determinações simultâneas, análises sem resolução etc. Isso faz com que, modelos de Calibração Multivariada sejam uma alternativa quando métodos univariados não podem ser aplicados. Diversos modelos de calibração multivariada vêm sendo utilizados, tais como Regressão Linear Múltipla (MLR), Regressão por Componentes Principais (PCR) e Regressão por Mínimos Quadrados Parciais (PLS)⁷. Esses modelos têm apresentado ótimos resultados, com aplicações em diversas áreas. Contudo, sua aceitação e implementação têm sofrido restrições em

alguns setores, como as áreas farmacêutica e médica, devido às exigências de validação. Essa validação é feita com base no cálculo e análise das chamadas figuras de mérito, que certificam que o modelo multivariado proposto é confiável e atende às especificações impostas pela indústria e órgãos de fiscalização nacionais e internacionais como "United States Pharmacopoeia" (USP), "International Conference on Harmonisation's" (ICH), "American Society for Testing and Materials" (ASTM), Instituto Nacional de Metrologia, Normalização e Qualidade Industrial (INMETRO), Agência Nacional de Vigilância Sanitária (ANVISA) etc.

OBJETIVO

Neste trabalho pretende-se apresentar procedimentos que podem ser utilizados para determinação de figuras de mérito em modelos de Calibração Multivariada, construídos pelo método de Mínimos Quadrados Parciais (PLS) e exemplificar sua utilização na validação de um modelo de calibração para a determinação de pureza polimórfica de Carbamazepina em misturas binárias dos polimorfos I e III, construído com base em dados de reflectância difusa na região do infravermelho próximo (NIR).

MODELO DE CALIBRAÇÃO MULTIVARIADA

Como na calibração univariada, a primeira etapa da construção do modelo de calibração multivariada trata da aquisição de respostas instrumentais para uma série de padrões (amostras de calibração), em que o valor da propriedade de interesse é conhecido. As respostas instrumentais para as amostras de calibração são arranjadas em uma matriz $\mathbf{X}_{(i \times j)}$ e os valores de referência da propriedade de interesse em um vetor $\mathbf{y}_{(i \times 1)}$, onde "i" é o número de amostras e "j" o número de variáveis. Utilizando o trabalho de Sijmen de Jong⁸ o modelo PLS pode ser construído, segundo o algoritmo SIMPLS, pela decomposição em componentes principais ("PCA - Principal Component Analysis") das matrizes $\mathbf{X}$ e $\mathbf{y}$ simultaneamente, levando a:

$$\hat{X}_A = T_A P_A^T + E \quad (1)$$

$$\hat{y}_A = T_A q_A + f \quad (2)$$

onde T_A são os escores, P_A e q_A são os "loadings" e E e f , as matrizes de erro referentes à decomposição de X e y respectivamente, " A " é o número de variáveis latentes utilizado no modelo, o índice " T " sobrescrito indica a operação de transposição de uma matriz ou vetor e o símbolo " $\wedge$ " indica que as grandezas são estimativas.

Paralela à decomposição de X e y é calculada uma matriz de pesos R , de forma que quando multiplicada pela matriz de dados X apresenta como resultado os escores T .

$$T_A = X R_A \quad (3)$$

A matriz de pesos R é determinada de forma a cumprir as seguintes condições:

1. maximizar a covariância entre X e y , isto é: $q_A^T (y^T X) r_A = \max$
2. R e q são normalizados, isto é, para cada variável latente A : $q_A^T q_A$ e $r_A^T r_A = 1$.
3. os vetores escores são ortogonais entre si, assim: para as variáveis latentes 1 e 2, $t_1^T t_2 = 0$.

Dessa forma, uma estimativa do vetor y poder ser obtida como descrito na Equação

2. Para um novo conjunto de amostras de previsão, podemos escrever:

$$\hat{y} = T_{\text{prev}} q_A = X_{\text{prev}} R_A q_A = X_{\text{prev}} R_A T_A^T y_{\text{cal}} \quad (4)$$

onde y_{cal} são as concentrações das amostras utilizadas na etapa de calibração.

Dessa forma o vetor de coeficientes de regressão $\hat{b}$ pode ser estimado por:

$$\hat{b} = R_A T_A^T y_{\text{cal}} \quad (5)$$

e o modelo de regressão representado por:

$$\hat{y}_{\text{prev}} = x_{\text{prev}} \hat{b} + \epsilon \quad (6)$$

onde $\hat{y}$ é o vetor de concentrações estimado através do modelo PLS e $\hat{\epsilon}$ são os erros de previsão.

SINAL ANALÍTICO LÍQUIDO

A validação de modelos de calibração multivariada pode ser feita com base no cálculo de parâmetros que assegurem que o modelo apresenta performance adequada e dentro dos objetivos desejados. Em calibração multivariada o conceito de Sinal Analítico Líquido, do inglês "Net Analyte Signal" (NAS), exerce uma importante função na determinação de figuras de mérito.

Lorber⁹ propôs um método para o cálculo do NAS para modelos multivariados de calibração inversa, definindo NAS para uma propriedade de interesse k , como sendo a parte do sinal instrumental que é ortogonal às contribuições de outros possíveis constituintes presentes na amostra. A [Figura 1](#) mostra a representação geométrica da propriedade de ortogonalidade do NAS¹⁰. Ferré *et al.*¹¹ fizeram correções ao modelo proposto por Lorber, de forma a possibilitar o cálculo exato do NAS para modelos de calibração construídos a partir de PCR e PLS. No cálculo do NAS, primeiramente a matriz X é reconstruída com A variáveis latentes gerando a matriz $\hat{X}_A$ (segundo a Equação 1), em seguida é determinada a matriz $\hat{X}_{A-k}$ que é a matriz que contém a informação de todas espécies presentes na amostra exceto da espécie de interesse k , como¹²:

$$\hat{X}_{A-k} = [I - \hat{y}_{A,k} \hat{y}_{A,k}^+] \hat{X}_A \quad (7)$$

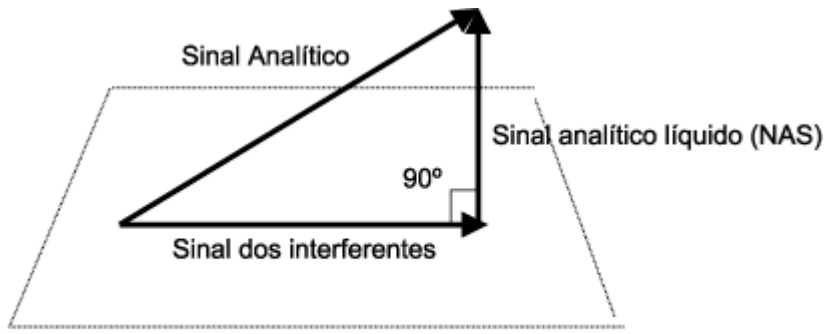


Figura 1. Representação geométrica do vetor de sinal analítico

onde $\hat{y}_{A,k}$ é o vetor de concentrações da espécie de interesse k estimado com A variáveis latentes segundo a equação 2, ou por $\hat{y}_{A,k} = \hat{X}_A \hat{X}_{A,k}^+$ e o índice "+" sobrescrito indica a pseudoinversa do vetor ou da matriz em questão. Isso faz com que a matriz $\hat{X}_{A,k}$ fique livre de qualquer contribuição da espécie k . O vetor NAS é então obtido como:

$$\hat{x}_{A,k}^{nas} = (I - \hat{X}_{A,k}^T (\hat{X}_{A,k}^T)^+) \hat{x}_A \quad (8)$$

onde $\hat{x}_A$ é o vetor de respostas instrumentais de uma amostra estimado com A variáveis latentes, calculado por $\hat{x}_A = P_A R_A^T x$.

Uma vez que $\hat{x}_{A,k}^{nas}$ é livre de interferentes, é possível substituí-lo por uma representação escalar sem perda de informação. Assim temos:

$$nas = \| \hat{x}_{A,k}^{nas} \| \quad (9)$$

onde $\| \cdot \|$ representa a norma Euclidiana do vetor $\hat{x}_{A,k}^{nas}$.

Com a possibilidade de calcular um valor escalar livre de interferentes, a partir de um vetor contendo contribuições de constituintes desconhecidos, torna-se possível a construção de uma nova forma de calibração multivariada, em que o modelo pode ser representado em uma forma univariada. Primeiro o cálculo do NAS é feito para

as i amostras de calibração, em seguida o coeficiente de regressão é determinado por mínimos quadrados entre o vetor $\hat{n}as$ e o vetor de concentrações y :

$$\hat{b}_{nas} = (\hat{n}as^T \hat{n}as)^{-1} \hat{n}as^T y \quad (10)$$

E o modelo de regressão pode, então, ser representado por:

$$\hat{y} = \hat{b}_{nas} \hat{n}as \quad (11)$$

onde os resultados obtidos por meio das Equações 6 e 11 são equivalentes.

Se os dados foram centrados na média, antes da determinação do coeficiente de regressão $\hat{b}_{nas}$, o vetor nas precisa ser corrigido de forma a evitar um erro de sinal que é introduzido pelo uso da norma Euclidiana. Esta correção pode ser feita pela multiplicação de cada elemento do vetor nas pelo seu sinal correspondente no vetor $(y - \bar{y})$, onde $\bar{y}$ é a média do vetor y , que contém os valores de referência¹³.

FIGURAS DE MÉRITO

No presente trabalho as definições para as figuras de mérito apresentadas seguem as recomendações presentes nos protocolos emitidos pela "International Union of Pure and Applied Chemistry" (IUPAC)¹⁴, "International Conference on Harmonisation (ICH)"^{15,16} e "American Society for Testing and Materials (ASTM)", norma E1655–00¹⁷.

Exatidão

Expressa o grau de concordância entre o valor estimado ou medido e o valor tido como verdadeiro ou de referência¹⁶. Comumente em aplicações com Calibração Multivariada a exatidão é estimada através da raiz quadrada do erro médio quadrático de previsão (RMSEP – "Root Mean Squares Error of Prediction"):

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (12)$$

onde n é o número de amostras de validação.

Segundo o protocolo emitido pela ICH¹⁵, a exatidão deve ser determinada usando um mínimo de 9 determinações em no mínimo três níveis de concentração que cubram a faixa de aplicabilidade do modelo. No presente trabalho ela foi determinada com base nas 39 amostras utilizadas como conjunto de validação.

Precisão

Expressa o grau de concordância entre os resultados uma série de medidas feitas para uma mesma amostra homogênea em condições determinadas. Em geral é considerada em diversos níveis, tais como repetibilidade, precisão intermediária (interdias, interanalista, etc.), reprodutibilidade¹⁶ e precisão média¹⁷.

- *Repetibilidade*: é a precisão do método em um curto intervalo de tempo. Pode ser determinada a partir de nove determinações (três concentrações e três replicatas) cobrindo a faixa útil do modelo de calibração, ou um mínimo de seis determinações de uma amostra.
- *Precisão intermediária*: a extensão em que a precisão intermediária deve ser determinada depende das circunstâncias em que o método será aplicado. Ela é o desvio devido à variação de uma determinada condição experimental. Variações típicas estudadas incluem dias, analistas, equipamento, etc.
- *Reprodutibilidade*: é acessada por meio de ensaios interlaboratoriais. É requerida apenas em casos de padronização de procedimentos analíticos como, por exemplo, inclusão de procedimentos em farmacopéias.

No presente trabalho foram determinadas a repetibilidade, a precisão média e interdias, calculada como¹⁸:

$$\text{precisão} = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (\hat{y}_{ij} - \bar{\hat{y}}_i)^2}{n(m-1)}} \quad (13)$$

onde m é o número de replicatas feitas, n o número de amostras, $\hat{y}_i$ é a média dos valores previstos de cada replicata $\hat{y}_{ij}$.

Ajuste do modelo

O grau de ajuste do modelo pode ser estimado a partir da correlação entre os valores de referência e os valores estimados da propriedade de interesse pelo modelo para as amostras de calibração. Isso é feito determinando-se, por mínimos quadrados, a reta que melhor se ajusta aos valores de referência e os estimados pelo modelo^{18,19}. Uma vez que o escalar "nas" é determinado, o ajuste do modelo também pode ser determinado através da melhor reta que se ajusta ao gráfico do "nas" contra a concentração, para as amostras de calibração.

Linearidade

Em modelos de calibração multivariada uma medida quantitativa da linearidade não corresponde a uma tarefa simples, ou mesmo possível. Qualitativamente, gráficos de resíduos e dos escores contra a concentração, os quais devem ter comportamento aleatório e linear, respectivamente, podem indicar se os dados seguem ou não o comportamento linear.

Sensibilidade

É definida como fração de sinal responsável pelo acréscimo de uma unidade de concentração da propriedade de interesse. Para modelos de calibração multivariada, como PLS, pode ser determinada como^{9,11}

$$SEN = \frac{1}{\|b\|} \quad (14)$$

ou, quando o NAS é determinado, o vetor de sensibilidade líquida para cada amostra i presente no conjunto de calibração pode ser determinado a partir do vetor $\hat{x}_{A,k}^{nas}$ como:

$$\hat{s}_k^{nas} = \frac{\hat{x}_{A,k}^{nas}}{y} \quad (15)$$

onde o vetor de sensibilidades $\hat{s}_k^{nas}$ deve ser igual para todas as amostras de calibração. O escalar SÊN pode ser determinado por:

$$SÊN = \|\mathbf{s}^{nas}\| \quad (16)$$

Sensibilidade analítica

É definida como a razão entre a sensibilidade e uma estimativa do desvio padrão para a flutuação do sinal analítico (dx)²⁰:

$$\gamma = \frac{SÊN}{\|\delta x\|} \quad (17)$$

onde o inverso desse parâmetro (γ^{-1}) permite estabelecer a menor diferença de concentração entre amostras, a qual pode ser distinguida pelo método e dx é estimado por meio do desvio padrão do valor de NAS obtido para 15 espectros do sinal do branco (espectro de uma superfície metálica utilizada como referência).

Intervalo de confiança

O intervalo de confiança para o valor estimado de y para uma amostra i pode ser definido como o intervalo no qual se pode afirmar, com certo grau de confiança, ou probabilidade, que inclui o valor verdadeiro da propriedade de interesse³. É importante que os limites de confiança estimados sejam consistentes e cubram o intervalo esperado para aquele nível de probabilidade, isto é, para o nível de 95% de confiança, 95 em 100 amostras devem ter o valor verdadeiro da propriedade dentro do intervalo de confiança calculado. Sua utilidade não se restringe apenas à determinação da incerteza do valor estimado para a propriedade de interesse. Se existir um modelo de calibração que apresenta erros de previsão que possam ser bem representados por um modelo de distribuição de probabilidade, como por exemplo *normal* ou *t-student*, esses intervalos podem ser utilizados como um método na identificação de amostras anômalas.

O cálculo desses intervalos depende de uma estimativa razoável da variância dos erros de previsão ($V(EP)$) para amostras desconhecidas. De acordo com a norma E1655-00 da ASTM¹⁷, a variância dos erros de previsão para amostras desconhecidas, em modelos de calibração multivariada, pode ser calculada por:

$$V(EP_i) = (1 + h_i)MSEC \quad (18)$$

onde n_{cal} é o número de amostras de calibração, h_i o "leverage" de uma amostra desconhecida, definido como:

$$h_i = \sum_{a=1}^A \frac{\hat{t}_{i,a}^2}{\hat{t}_a^T \hat{t}_a} \quad (19)$$

onde $\hat{t}$ e $\hat{t}_i$ são os escores do conjunto de calibração e de uma amostras desconhecida, respectivamente, e MSEC é a média quadrática dos erros de calibração, definida como:

$$MSEC = \sum_{i=1}^{n_{cal}} \frac{(y_i - \hat{y}_i)^2}{v} \quad (20)$$

onde n é o número de graus de liberdade.

Essa aproximação de $V(EP)$ leva em consideração que as variâncias das respostas instrumentais no conjunto de calibração e para amostras de previsão são iguais e que a variância dos erros devido ao método de referência é insignificante. Estudos feitos por Faber *et al.*²¹ mostram que a primeira suposição é consistente com a maior parte das aplicações práticas, contudo em ocasiões em que a variância do método de referência for significativa, esta deve ser adicionada à Equação 18.

Outra dificuldade no cálculo dos intervalos de confiança é estabelecer o número efetivo de graus de liberdade (n) envolvidos no cálculo do MSEC. Nesse sentido, Voet²² definiu o conceito de pseudograus de liberdade (PDF), que leva em consideração a diferença entre o erro médio quadrático de calibração estimado por validação cruzada (MSECV) e pela previsão das próprias amostras de calibração (MSEC), de modo que quanto maior é essa diferença, n é menor:

$$v = n - n \left[1 - \sqrt{\frac{\text{MSEC}}{\text{MSECV}}} \right] \quad (21)$$

Após o cálculo da variância, com o número apropriado de graus de liberdade, os limites de confiança f podem ser obtidos por:

$$\phi_i = \pm t_{1-\alpha/2, v} \sqrt{V(\text{PE})_i} \quad (22)$$

onde $t_{1-\alpha/2, n}$ é o parâmetro estatístico da distribuição t-student com probabilidade $(1-\alpha)/2$ de recobrimento.

Enfim, admite-se que os erros correspondentes às concentrações estimadas não são correlacionados, têm variância constante na faixa estudada e seguem uma distribuição normal.

Seletividade

É a medida do grau de sobreposição entre o sinal da espécie de interesse e os interferentes presentes na amostra e indica a parte do sinal que é perdida por essa sobreposição. Para modelos de calibração multivariada esse parâmetro é definido como^{9,11}:

$$\text{SEL}_i = \frac{n \hat{s}_i}{\|x_i\|} \quad (23)$$

onde $n \hat{s}_i$ é o escalar NAS estimado para amostra 'i' e x_i o vetor de dados original. O valor de seletividade estimado a partir da expressão 21 informa quanto do sinal original não é retirado por não ser ortogonal à propriedade de interesse. Dessa forma, a seletividade calculada a partir da Equação 23 não se refere ao sentido geralmente empregado para o termo em química analítica, com modelos univariados, e sim a uma forma de estimar quanto do sinal é perdido por ortogonalidade.

Razão sinal ruído

Indica quanto da intensidade do NAS da espécie de interesse está acima da estimativa do desvio padrão para a flutuação do sinal analítico (dx). É definido através do NAS como^{9,11}:

$$S/R_i = \frac{n\hat{s}_i}{\|\delta x\|} \quad (24)$$

Teste para erro sistemático ("bias")

Segundo a IUPAC, "bias" é definido como a diferença entre a média populacional e o seu valor verdadeiro. Erros sistemáticos são todas as componentes de erros que não são randômicas, logo, eles correspondem à parcela fixa do "bias" em uma medida química¹⁴. A existência desse tipo de erro afeta a precisão, a exatidão e a determinação dos intervalos de confiança. A norma E1655-00 da ASTM¹⁷ sugere um teste-t para avaliar se o "bias" incluso no modelo é significativo. O "bias" médio pode ser calculado como:

$$\text{bias} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (25)$$

Em seguida, o desvio padrão dos erros de validação (SDV, "standard deviation of validation") é estimado por:

$$\text{SDV} = \sqrt{\frac{\sum [(y - \hat{y}) - \text{bias}]^2}{n - 1}} \quad (26)$$

O valor de t é então obtido:

$$t = \frac{|\text{bias}|\sqrt{n}}{\text{SDV}} \quad (27)$$

Caso o valor de t calculado seja menor que o seu valor crítico para n graus de liberdade com 95% de confiança, o "bias" incluído no modelo pode ser considerado insignificante e desprezado.

PROCEDIMENTO EXPERIMENTAL

O polimorfo III, que é utilizado na formulação de medicamentos para tratamento de epilepsia, foi fornecido pela Indústria Cristália Produtos Químicos Farmacêuticos Ltda. O laudo de análises fornecido pela empresa indicou que a amostra enviada estava compatível com o padrão USP de concentração 100%. O polimorfo I foi obtido a partir da forma I por aquecimento à temperatura de 160 °C por 3 h e 30 min. Os polimorfos I e III, em suas formas puras, foram caracterizados por meio de análises em Infravermelho em pastilha de KBr e Calorimetria de Varredura Diferencial (DSC)^{23,24}.

As amostras foram preparadas por meio de misturas das formas I e III na faixa de 80 a 100%(m/m). As pesagens foram realizadas em uma balança analítica com precisão de 0,00001 g e misturadas em recipientes de 5 mL com ajuda de três esferas metálicas.

Os espectros de reflectância difusa na região do infravermelho próximo (NIR) foram registrados em um espectrofotômetro Bomem Hartman & Braun DA-8 Transformada de Fourier, equipado com lâmpada de quartzo halogênio (2000-25000 cm⁻¹), divisor de feixe quartzo (4000-25000 cm⁻¹), detector de InGaAs IPH 6300L resfriado com nitrogênio líquido e resolução de 4 cm⁻¹. Foi utilizado o acessório de reflectância difusa Jasco modelo DR81. Cada espectro foi obtido como uma média de 128 varreduras e foram obtidos 4 espectros de cada amostra. O elevado número de replicatas foi adotado por problemas com a homogeneidade da amostra, visto que os polimorfos não se misturavam com facilidade. Foram feitas replicatas apenas das medidas instrumentais, portanto os erros com origem na pesagem dos polimorfos não foram levados em consideração nos cálculos de incerteza.

MODELO DE CALIBRAÇÃO

Para a construção do modelo de calibração foram preparadas 41 amostras com concentração entre 80 a 100% e 40 amostras para validação, distribuídas na mesma faixa de concentração.

Os espectros obtidos no infravermelho próximo apresentaram um acentuado desvio de linha base entre as replicatas, que pode ser visualizado na [Figura 2](#). As principais causas desse desvio de linha de base são espalhamentos aditivos, que provocam o deslocamento do espectro na vertical e espalhamentos multiplicativos, que provocam uma inclinação da linha base do espectro. Além dos efeitos de espalhamento, observaram-se duas regiões com alto nível de ruído, entre 1350-1420 nm e 1800-1950 nm, que são devidas principalmente ao 1º sobretom e à banda de combinação da água. Como forma de correção foi adotada, como pré-processamento, a 1ª derivada das absorbâncias, seleção da melhor faixa espectral e dados centrados na média. Foram selecionadas duas faixas espectrais: 1440-1800 nm e 1925-2185 nm, que são devidas principalmente aos 1º e 3º sobretons de ligações N-H em amidas, 1º sobretom de ligações C-H em aromáticos (na primeira região) e 2º sobretom de ligações N-H em amidas, 3º sobretom de ligações C-C em aromáticos, 1º sobretom de ligações C=O em amidas e bandas de combinação de ligações C-H em aromáticos (segunda região).

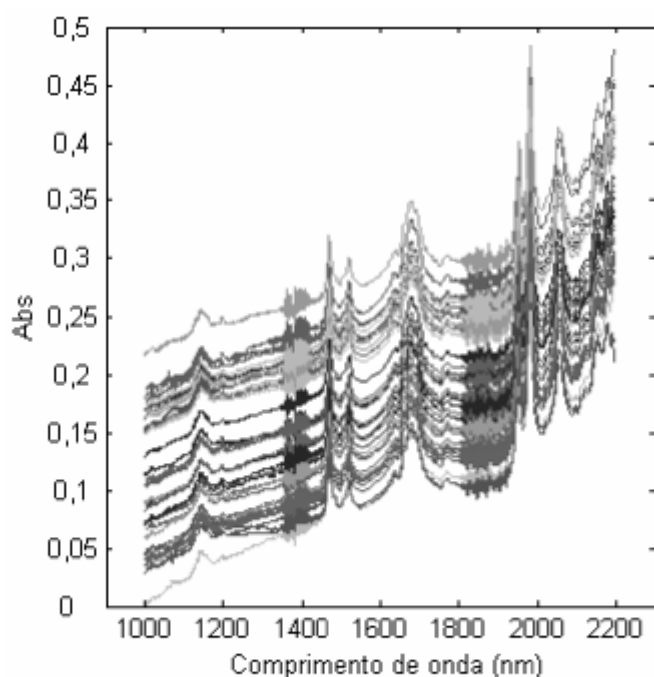


Figura 2. Espectros NIR de reflectância difusa para as replicatas do conjunto de calibração

Para a construção do modelo de calibração foi utilizado o espectro médio das quadruplicatas e a dimensionalidade do modelo foi determinada através do procedimento de validação cruzada "deixe uma fora" ("leave one out"). Para validação foram previstas individualmente cada replicata e a concentração da amostra obtida como a média dos resultados de cada replicata.

RESULTADOS E DISCUSSÃO

A Carbamazepina (CBZ) é um conhecido princípio ativo de medicamentos antiepiléticos, que apresenta quatro formas polimórficas, três anidras (formas I, II e III) e uma di-hidratada (IV)²³. A [Figura 3](#) mostra a fórmula estrutural da CBZ. Suas principais formas são I e III, sendo a forma I meta estável à temperatura ambiente com estrutura trigonal e a forma III estável com estrutura monoclinica²⁴. O medicamento produzido à base de CBZ é feito com a forma III que é absorvida pelo organismo mais rapidamente, devido à sua solubilidade, sendo que a forma I é tida como impureza na matéria-prima que é adquirida pela indústria. O modelo de calibração PLS foi construído para a determinação quantitativa da forma III da CBZ em misturas binárias das formas I e III, entre a faixa de concentração de 80 a 100% (m/m), do polimorfo III em relação ao I.

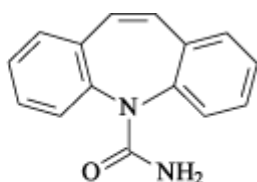


Figura 3. Fórmula estrutural da Carbamazepina

Os termogramas e espectros de infravermelho referentes às formas I e III utilizados na etapa de caracterização são mostrados na [Figura 4](#). Através da [Figura 4a](#) pode-se notar que o termograma da forma III apresenta dois sinais endotérmicos de fusão, o primeiro em aproximadamente 175 °C correspondente à fusão da forma III, seguidos por um sinal exotérmico indicando a cristalização do polimorfo I em 180 °C e sua subsequente fusão em aproximadamente 192 °C. O termograma da forma I apresenta apenas um pico endotérmico que corresponde à sua fusão. Os espectros de infravermelho em pastilha de KBr, na região do infravermelho médio apresentaram picos característicos da forma I em 955 cm^{-1} e em 1273 cm^{-1} (indicados). Foi feita a medida da intensidade desses sinais em relação a picos constantes para as duas formas²³, o que apresentou como resultado que ambas as formas estavam em seu estado puro. Essa consideração foi admitida no cálculo dos valores de referência das concentrações das amostras preparadas.

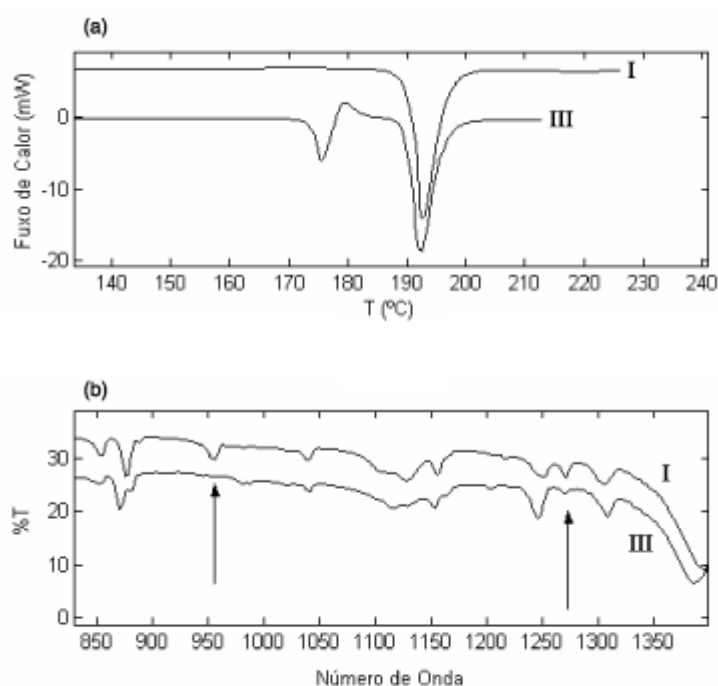


Figura 4. Espectros de DSC (a) e infravermelho (b) para os polimorfos I e III de CBZ

Para construção do modelo, foram usadas 38 amostras na etapa de calibração e 39 para validação. As demais amostras foram apontadas como anômalas, segundo os testes descritos na norma E1655 da ASTM¹⁷.

O modelo PLS construído utilizou três variáveis latentes, que correspondem a mais de 95% da variância dos dados originais (espectrais e concentrações de referência). O maior erro percentual observado para as amostras de validação foi de 1,7%, com média igual a 0,6% (m/m) com valores absolutos seguindo aproximadamente uma distribuição normal.

O resultado do cálculo das figuras de mérito é apresentado na [Tabela 1](#). O modelo linear parece se ajustar bem aos dados. As [Figuras 5](#) e [6](#) mostram os gráficos dos valores de referência contra os valores estimados, e gráfico do escalar NAS contra a concentração, respectivamente, que demonstram o bom ajuste do modelo linear e a apresentação do modelo PLS na sua forma pseudounivariada.

Tabela 1. Resultados de figuras de mérito estimadas para o modelo PLS

Figuras de mérito NIR			
Exatidão ^a	RMSEC		0,6
	RMSEP		0,7
Precisão ^a	Média		0,6
	Repetibilidade		0,9
	Intermediária		0,5
Sensibilidade ^b			0,0004
Inverso da Sensibilidade Analítica ^a			0,2
Razão Sinal-Ruído	Max		791
	Min		3
Seletividade Média	Calibração		0,5
	Validação		0,4
Ajuste (referência x estimado)	Inclinação		0,992 ± 0,015
	Intersecção		0,5 ± 1,4
	Corr. Coef.		0,992
Ajuste (NAS)	Inclinação		2550 ± 39
	Intersecção		90,0 ± 0,1
	Corr. Coef.		0,996
Erro Sistemático	ASTM	bias ^a	-0,1
		SDV	0,9
		tbias	0,67
		DF	43

^a Valores expressos em % (m/m) de CBZ III em relação a CBZ I;

^b valor expresso como a razão entre a derivada da absorbância e a % (m/m).

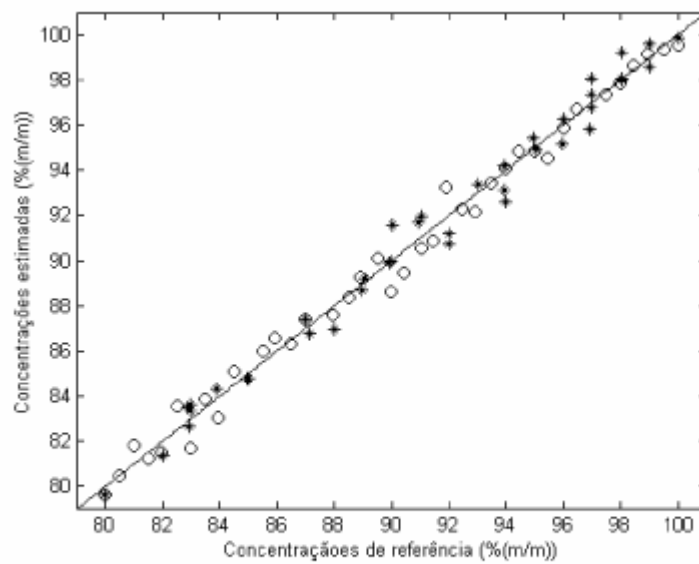


Figura 5. Valores de referência versus estimados pelo modelo NIR, (○) amostras de calibração, (*) validação e (—) reta ajustada

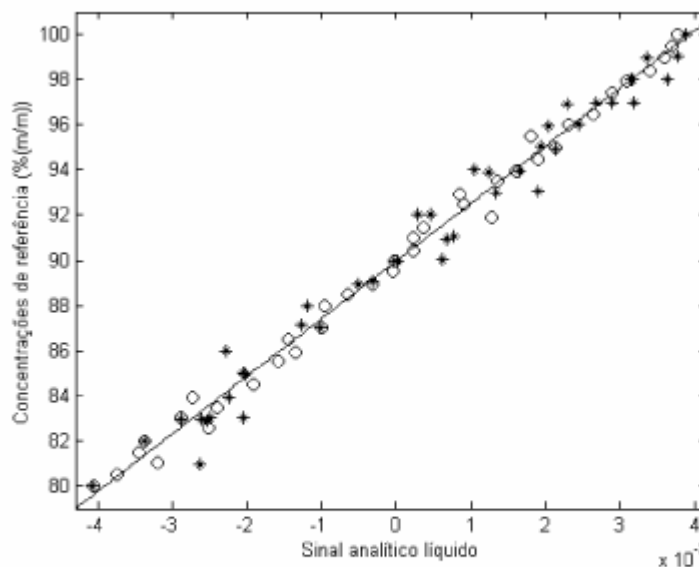


Figura 6. Escalar NAS contra as concentrações de referência, (○) amostras de calibração, (*) validação e (—) o modelo PLS com 3 variáveis latentes em sua forma pseudounivariada

A [Figura 7](#) mostra o gráfico dos escores da primeira variável latente contra a concentração e a [Figura 8](#), o gráfico de resíduos. Claramente observa-se um comportamento linear na [Figura 7](#) e erros com comportamento aleatório na [Figura 8](#). No gráfico de escores, os pontos que apresentam maior desvio do comportamento linear não devem ser encarados como indício de falta de linearidade dos dados, uma vez que o modelo PLS utiliza três variáveis latentes e não apenas uma, desprezando quatro pontos que apresentam maior desvio, o coeficiente de correlação dos escores contra os valores de referência calculado é 0,953.

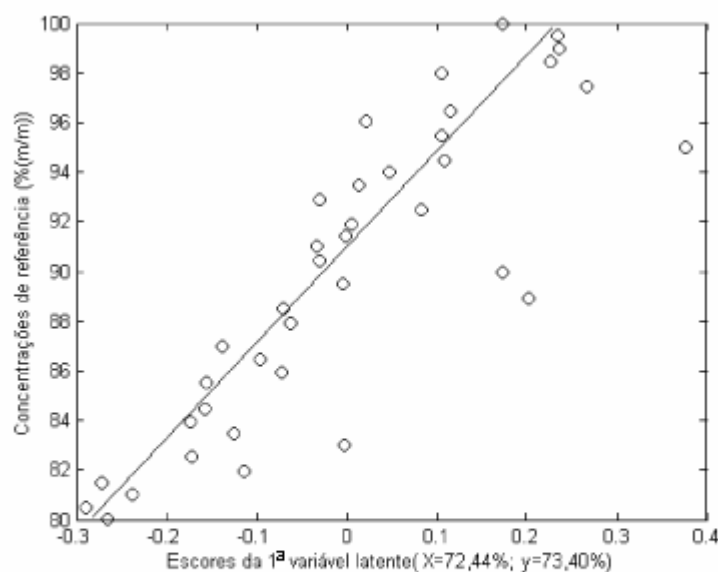


Figura 7. (○) Escores da 1ª variável latente (65,68% de variância em X e 91,95% em Y) contra as concentrações de referência para as amostras de calibração, (—) reta ajustada

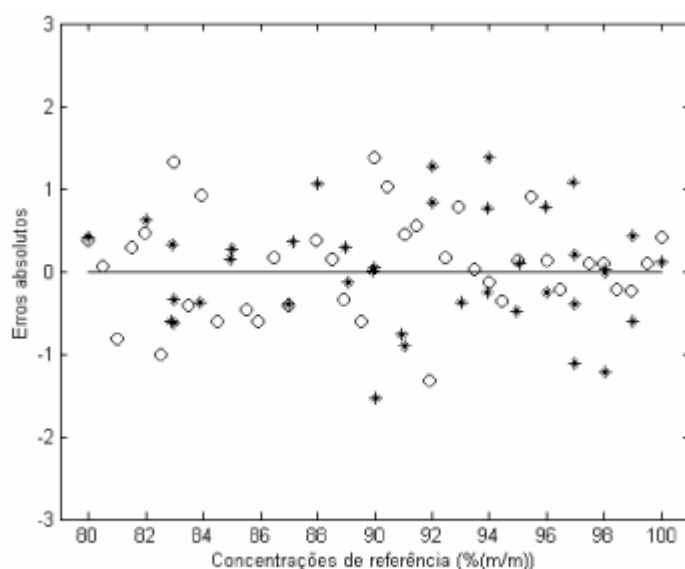


Figura 8. Gráfico de erros absolutos contra as concentrações de referência

Os indicadores de exatidão, mostrados na [Tabela 1](#), apresentam um nível aceitável de dispersão e uma boa concordância entre si, implicando que o valor de MSEC é uma boa estimativa do desvio padrão dos erros de previsão observados no conjunto de validação. Segundo o teste-t para ocorrência de "bias", suas possíveis influências no modelo e nos parâmetros estimados podem ser desprezadas, uma vez que o valor de t calculado é menor que o valor de $t_{\text{crítico}}$ tabelado da distribuição de t-student.

O modelo apresentou um baixo valor numérico de sensibilidade, que em parte é decorrência do pré-processamento utilizado. O valor para o inverso da sensibilidade analítica, apresentado na [Tabela 1](#), pode ser interpretado de forma mais clara, por sua relação direta com a concentração. Segundo esse valor, o modelo é capaz de distinguir entre amostras com diferença de concentração da ordem de 0,2% (m/m).

Os valores de precisão apresentam um nível de dispersão aceitável. Os resultados para precisão intermediária e repetibilidade foram obtidos pelo desvio padrão das médias de 6 medidas de uma amostra de concentração 95%, feitas em 6 dias consecutivos e 6 vezes no mesmo dia, respectivamente. Para precisão média foi calculado o desvio padrão médio de três amostras de concentrações 83, 90, 97%,

que foram medidos seis vezes em três dias consecutivos. Os resultados apresentados na [Tabela 1](#) mostram que o valor obtido para precisão intermediária é maior que o obtido para repetibilidade, apresentando uma ordem inversa da que é esperada teoricamente. Essa diferença pode ser atribuída a possíveis variações instrumentais que ocorreram ao longo do dia em que foram realizadas as medidas de repetibilidade. Contudo, se um teste F for realizado, a partir dos quadrados dos respectivos valores de precisão intermediária e repetibilidade, observa-se que o valor de F calculado é de 3,24, enquanto que o valor de $F_{95\%}$ tabelado (com 5 e 5 graus de liberdade) é de 5,05. Uma vez que o valor calculado não excede o valor crítico, pode-se concluir que não há diferença estatisticamente significativa entre as duas estimativas de precisão.

Os valores de seletividade que são mostrados na [Tabela 1](#) representam a fração média do sinal do espectro da amostra que não é retirado, por não ser ortogonal em relação à propriedade de interesse. Logo, esses valores não se referem à seletividade no seu significado físico, que é em geral empregado em Química Analítica. Eles indicam que cerca de metade do sinal é retirado na etapa de aniquilação de posto realizado durante o cálculo do NAS.

Os valores para razão sinal ruído apresentados mostram o quanto o escalar NAS está acima do desvio padrão da flutuação do sinal instrumental. Segundo os resultados, a menor razão observada foi 3, que representa um limite mínimo aceitável para determinações quantitativas.

Os resultados para as percentagens de recobrimento dos intervalos de confiança, que representam a percentagem de amostras que possuem o valor verdadeiro dentro dos limites de confiança estimados os níveis de probabilidade 99,0, 95,0, e 90,0%, foram 100,0; 92,3 e 87,2%, respectivamente. Pode-se perceber que as variâncias estimadas através da equação presente na ASTM E1655¹⁷ conseguem fornecer intervalos de recobrimento próximos aos esperados teoricamente. O histograma de resíduos estudantizados, apresentado na [Figura 9](#), aproxima-se bem de uma distribuição t-student com desvio padrão igual a 1,13, sendo que os limites médios estimados ficam em torno de 1,7; 1,2 e 1,0 para os níveis de 99,0 95,0 e 90,0% de confiança, respectivamente. No gráfico de barras de erro para o nível de

95% de confiança, mostrado na [Figura 10](#), pode-se visualizar as 3 amostras, que correspondem a 7,7% do conjunto de validação, que têm seu valor de referência fora do intervalo de confiança estimado. O maior erro observado para os intervalos de recobrimento foi de 2,7%, que corresponde ao erro no nível de 95% de confiança, que é um erro totalmente aceitável considerando o número reduzido de amostras de validação utilizadas nesse estudo.

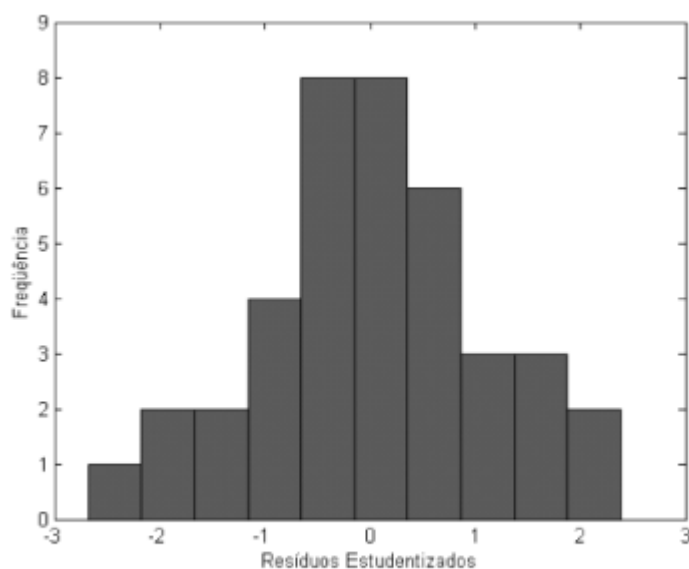


Figura 9. Histograma dos resíduos estudentizados para as amostras de validação

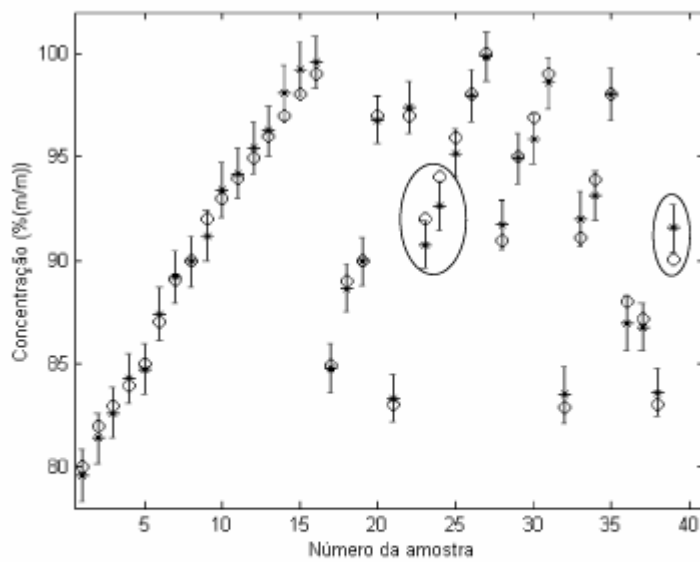


Figura 10. Gráfico de barras de erro para as amostras de validação no nível de probabilidade 95%. Em círculos são mostradas as 3 amostras cujos limites não cobrem o valor experimental, (○) concentração experimental, (*) concentração estimada

REFERÊNCIAS

<https://www.todamateria.com.br/poluicao-do-solo/>>acesso em 21/05/2020

<https://www.todamateria.com.br/tipos-de-solo/>>acesso em 21/05/2020

https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-40422004000600027>acesso em 21/05/2020