



**FUNDAÇÃO UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS VETERINÁRIAS
CURSO DE MESTRADO**



**AVALIAÇÃO DE TÉCNICA DE ESPECTROSCOPIA
ÓPTICA E APRENDIZADO DE MÁQUINA PARA
DIAGNÓSTICO SOROLÓGICO DE *Brucella abortus* E
Mycobacterium bovis EM BOVINOS**

BRUNO SILVA DE REZENDE

**Campo Grande – MS
2023**

BRUNO SILVA DE REZENDE

**AVALIAÇÃO DE TÉCNICA DE ESPECTROSCOPIA
ÓPTICA E APRENDIZADO DE MÁQUINA PARA
DIAGNÓSTICO SOROLÓGICO DE *Brucella abortus* E
Mycobacterium bovis EM BOVINOS**

*Evaluation of optical spectroscopy technique and machine learning for
serological diagnosis of *Brucella abortus* and *Mycobacterium bovis* in
cattle*

BRUNO SILVA DE REZENDE

Orientador: Carlos Alberto do Nascimento Ramos

Dissertação apresentada ao programa de
Pós-Graduação em Ciências Veterinárias da
Universidade Federal de Mato Grosso do Sul,
como requisito parcial para a obtenção do
título de Mestre em Ciências Veterinárias.

**Campo Grande – MS
2023**

DEDICATÓRIA

Dedico esta vitória a minha madrinha Elizabeth da Silva Benites, ao meu padrinho Porfirio da Silva Benites, a minha avó Margarete Almeida de Rezende e ao meu avô Antonio Carlos Ferreira de Rezende. Estas pessoas foram fundamentais ao longo deste caminho. A minha gratidão é eterna.

AGRADECIMENTOS

Quero agradecer primeiramente a Deus por todo o amparo para que eu tivesse determinação em prol dos meus objetivos.

Agradeço ao CNPq pelo apoio financeiro. Além disso, fica aqui o meu agradecimento, pela parceria de pesquisa, ao Prof. Dr. Cícero Cena e Thiago França. Indubitavelmente, minha gratidão aos alunos de iniciação científica Maykko e Arthur, por toda ajuda com os experimentos.

Aos meus familiares, em especial a minha mãe Patricia, que sempre esteve comigo: obrigado!

Agradeço também aos meus amigos e a todos que torceram por mim!

E o meu principal agradecimento será para o meu orientador Prof. Dr. Carlos Alberto do Nascimento Ramos. Foi um verdadeiro orientador e um grande amigo! Nada disso seria possível se eu não tivesse todo o suporte deste grande ser humano. Agradeço de coração por toda ajuda, ensinamentos e conselhos!

*“Não existe um caminho para a felicidade. A felicidade é o caminho”
Mahatma Gandhi*

RESUMO

A tuberculose e a brucelose bovina são doenças importantes no cenário da bovinocultura do país, haja vista os impactos econômicos devido ao sacrifício de animais positivos, condenação de produtos de origem animal e o risco à saúde pública. Ademais, *B. abortus* e *M. bovis* são microrganismos muito resistentes no ambiente o que facilita a sua disseminação. O Programa Nacional de Controle e Erradicação de Brucelose e Tuberculose (PNCEBT) promove além de vacinação para brucelose, testes de diagnóstico para *B. abortus* e *M. bovis*, sacrifício ou abate sanitário de animais positivos e ações de adesão voluntária com a finalidade de fomentar a prevenção e controle destas doenças. Entretanto, os testes atualmente empregados podem apresentar alguns problemas seja de especificidade e ou sensibilidade, ou demanda de estrutura laboratorial complexa, alto custo e demora no tempo de execução. Dessa forma, justifica-se as pesquisas por métodos baseados em espectroscopia como FTIR, UV-Vis e outras, que associadas a métodos de análise multivariada e algoritmos de aprendizado de máquina (*machine learning*) têm demonstrado serem promissoras no diagnóstico de doenças, apresentando índices de sensibilidade-especificidade aceitáveis, custos mais baixos e rapidez na obtenção de resultados. Dessa forma, foi avaliado nessa pesquisa o método UV (Ultra violeta), associado ao aprendizado de máquina, para diagnóstico de brucelose e tuberculose em bovinos. Avaliou-se a metodologia de UV para ambas as enfermidades, com diferentes antígenos (antígeno de *B. abortus* para soroaglutinação lenta Tecpar®; antígenos recombinantes P27, MPB83 e MPB70 de *M. bovis*). Foram utilizados um total de 106 amostras de soro bovino (53 positivas e 53 negativas) para brucelose, e 88 amostras (44 positivas e 44 negativas) para tuberculose. As proporções de antígeno-soro foram avaliadas para cada antígeno e identificou-se, por meio da análise de componentes principais (PCA), que a proporção 1:1 para brucelose e, 1:16 (MPB83), 1:2 (MPB70) e 1:2 (P27), para tuberculose, apresentaram melhor separação de positivos e negativos. Ademais, com a proporção definida, foi realizado a coleta dos espectros UV do total de amostras. Em seguida os espectros foram submetidos a análise de componentes principais com o objetivo de observar a tendência da formação de agrupamentos, sendo que para brucelose, não apresentou nenhuma tendência de agrupamento, enquanto para tuberculose, somente o antígeno P27 apresentou bons resultados de agrupamento. Finalmente, usando algoritmos de aprendizado de máquina, foi alcançada uma acurácia geral de 92,5% para brucelose e 96,3% para tuberculose (usando antígeno P27).

Palavras-chave: Aprendizado de máquina, Bovino, Diagnóstico, Sorologia.

Abstract

Tuberculosis and bovine brucellosis are important diseases for cattle farming in Brazil, given the economic impacts due to the sacrifice of positive animals, condemnation of products of animal origin and the risk to public health. In addition, *B. abortus* and *M. bovis* are very resistant microorganisms in the environment, which facilitates their dissemination. The National Program for the Control and Eradication of Brucellosis and Tuberculosis (PNCEBT) promotes vaccination against brucellosis, diagnostic tests for *B. abortus* and *M. bovis*, sacrifice or sanitary slaughter of positive animals and voluntary adherence actions with the purpose of promoting the prevention and control of these diseases. However, the tests currently employed may have some specificity and/or sensitivity problems, or demand for a complex laboratory structure, high cost and excessive execution time. Thus, research on methods based on spectroscopy such as FTIR, UV-Vis and others is justified, which associated with multivariate analysis methods and machine learning algorithms have shown to be promising in the diagnosis of diseases, presenting indices acceptable sensitivity-specificity, lower costs and faster results. Thus, the UV (Ultraviolet) method associated with machine learning was evaluated in this research for the diagnosis of brucellosis and tuberculosis in cattle. The UV methodology was evaluated for both diseases, with different antigens (*B. abortus* antigen for serum agglutination test; recombinant antigens P27, MPB83 and MPB70 of *M. bovis*). A total of 106 bovine serum samples (53 positive and 53 negative) for brucellosis and 88 samples (44 positive and 44 negative) for tuberculosis were used. The antigen-serum ratios were evaluated for each antigen, and it was identified, through principal component analysis (PCA), that the ratio 1:1 for brucellosis and, 1:16 (MPB83), 1:2 (MPB70) and 1:2 (P27), for tuberculosis, showed better separation of positives and negatives. Furthermore, with the proportion defined, the collection of the UV spectra of the total samples was carried out. Then the spectra were submitted to principal component analysis in order to observe the tendency of cluster formation. For brucellosis, no clustering tendency was observed, while for tuberculosis, only the P27 antigen showed good clustering results. Finally, using machine learning algorithms, an overall accuracy of 92.5% for brucellosis and 96.3% for tuberculosis (using P27 antigen) was achieved.

Keywords: Bovine, Diagnosis, Machine learning, Serology.

LISTA DE FIGURAS

- Figura 1. Prevalência de casos de brucelose bovina por UF.
- Figura 2. Prevalência de focos de brucelose bovina por UF.
- Figura 3. Prevalência de casos de tuberculose bovina por UF.
- Figura 4. Prevalência de focos de tuberculose bovina por UF.
- Figura 5. PCA realizada com os dados de espectroscopia UV-Vis de amostras positivas (preto) e negativas (azul) para *Mycobacterium bovis*.
- Figura 6. Descrição hierárquica e exemplos de técnicas de IA e suas subáreas ML e *Deep Learning* (DL).
- Figura 7. Os quatro componentes básicos do processo de *machine learning*.
- Figura 8. Average antigen/serum UV-SNV spectra for bovine brucellosis positive (POS - red color) and negative (NEG – blue color) groups. The standard deviation is evidenced by the colored shadow around the mainline. Main electronic bands centered around 206, 220, and 273 nm, assigned to $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions in carbonyl groups from proteins and amino acids.
- Figura 9. Principal Component Analysis for UV-SNV spectra: (a) score plot for PC1 versus PC2 from bovine blood serum UV-SNV spectra, [red square] bovine brucellosis positive group, and [blue circle] negative group. (b) Loadings for PC1 and PC2 in the 200 to 300 nm range.
- Figura 10. Violin plot for three main PCs normalized scores selected by RFE in 200 to 300 nm from bovine brucellosis positive group (red color), and negative group (blue color). A monomodal distribution is shown with the main peak center dislocated for both groups when projected over the y-axis direction. The data variance percentage of each PC is described below each plot. Respective loading plots for PC3 (b), PC4 (d), and PC14 (f) are also shown.
- Figura 11. Principal Component Analysis for UV-SNV spectra after RFE analysis. 3D Score plot for PC3 versus PC4, and PC14 from bovine blood serum UV-SNV spectra in the 200 to 300 nm range, [red

circle] bovine brucellosis positive group, and [blue square] negative group.

- Figura 12 (a) Overall accuracy obtained in the LOOCV tests (light blue bars with right inclined line pattern) for Discriminant analysis (DA), K-nearest neighbor (KNN), and Support vector machine (SVM) algorithms by using PC3, PC4, and PC14 data in the 200 to 300 nm range. Confusion matrix for the performance of (b) SVM algorithm with a linear function, and (c) KNN algorithm with weighted function. The highest overall accuracy was 92.5%.
- Figura 13 Scatter plot based on Principal component analysis for UV-SNV spectra of bovine blood serum. Bovine tuberculosis positive samples (red) and negative (blue). The (a) correspond to the P27 antigen, (b) correspond to the MPB70 antigen, and (c) to the MPB83 antigen.
- Figura 14 (a) Overall accuracy obtained in LOOCV tests (light blue bars) for the machine learning algorithms. (b) Confusion matrix for the performance of the KNNM algorithm with an accuracy of 96.3%.

SUMÁRIO

RESUMO.....	V
Abstract.....	VI
LISTA DE FIGURAS.....	VII
1. INTRODUÇÃO.....	10
2. REVISÃO DE LITERATURA.....	14
2.1 Brucelose bovina no Brasil.....	14
2.1.1 Etiologia e patogenia.....	14
2.1.2 Técnicas de diagnóstico.....	17
2.1.3 Distribuição geográfica.....	20
2.2 Tuberculose bovina no Brasil.....	22
2.2.1 Etiologia e patogenia.....	22
2.2.2 Técnicas de diagnóstico.....	23
2.2.3 Distribuição geográfica.....	26
2.3 Espectroscopia UV-Vis.....	27
2.4 PCA e <i>Machine Learning</i>.....	28
3. REFERÊNCIAS BIBLIOGRÁFICAS.....	32
4. CAPÍTULO I: Turning chaotic sample group clusterization into organized ones by feature selection: Application on photodiagnosis of <i>Brucella abortus</i> serological test.....	42
5. CAPÍTULO II: Evaluation of UV spectroscopy and machine learning for serological diagnosis of bovine tuberculosis.....	58
6. CONSIDERAÇÕES GERAIS.....	72
7. IMPACTO ECONÔMICO, SOCIAL, TÉCNOLÓGICO E/OU INOVAÇÃO.....	72

1. INTRODUÇÃO

A brucelose e a tuberculose são enfermidades infectocontagiosas, causadas por microorganismos do gênero *Brucella* e *Mycobacterium*, respectivamente. Às duas doenças têm grande impacto econômico por gerar inúmeros prejuízos à bovinocultura além de serem zoonoses importantes (BRASIL, 2006). A prevalência de focos de brucelose em bovinos pelo país depende de cada região, variando entre 0,32% e 41,5% (CHATE et al., 2009). Ademais, a tuberculose bovina, em um estudo entre 1989 e 1998, apresentou prevalência média em todo o território nacional de 1,3% (FILHO et al., 2017). No estado de Mato Grosso do Sul (MS), a partir de um levantamento epidemiológico no ano de 1998, observou-se prevalência para brucelose de 6,3% (BRASIL, 2006). Além disso, em um estudo realizado entre 2012 e 2019, o estado de MS apresentou prevalência para tuberculose bovina de 0,15% (GARCIA et al., 2021).

A ocorrência de brucelose no rebanho bovino brasileiro pode ocasionar um impacto significativo haja vista a condenação de produtos alimentícios como: carne e leite. Além disso, há grandes gastos para a implementação dos programas de controle e erradicação da doença somado aos prejuízos na bovinocultura devido aos abortos, bezerros subdesenvolvidos, menor índice reprodutivo, menor produção de leite e derivados (POESTER et al., 2009). Indubitavelmente, tais fatores mencionados somam-se com as mortalidades e intervenções veterinária. A brucelose bovina no Brasil gera perdas estimadas em R\$ 420,12 ou R\$ 226,47, para cada fêmea infectada acima de 24 meses em rebanhos de leite ou corte, respectivamente. O prejuízo total anual estimado para tal doença no rebanho bovino brasileiro, foi de R\$ 892 milhões, correspondendo a 0,32% do PIB do Brasil em 2011 (SANTOS et al., 2013).

O homem pode ser infectado por meio dos produtos de origem animal contaminados ou por contato direto ou indireto com animais infectados (PAULIN et al., 2008). Dores articulares, febre, tosse, náusea, cefaleia, linfonodos reativos, mialgia, depressão e perda de peso são os principais sinais clínicos observados em humanos. Além disso, pode ocorrer supurações em órgão como fígado e baço (LAWINSKY et al., 2010).

A tuberculose bovina é uma doença que pode acometer diferentes espécies de animais e humanos, caracterizando assim sua importância para a bovinocultura e saúde pública (SALAZAR, 2005). O agente da tuberculose se caracteriza por gerar uma doença principalmente respiratória, pois quando há a inalação dos bacilos pelo hospedeiro, macrófagos alveolares fagocitam o microorganismo, causando lesões denominadas tubérculos (MARCONDES, 2002; PRITCHARD, 1988). Além disso, a transmissão ocorre principalmente por via aerógena, mas também há a inclusão de via oral, congênita e venérea (NEIL et al., 1994). Em seres humanos, comumente, a transmissão pode ocorrer pelo consumo de leite e derivados contaminados (ACHA et al., 2001). Ademais, segundo o estudo de Ruggiero et al. (2022), a disseminação do bacilo pode ocorrer por meio de corrimento nasal, leite, fezes, urina, secreção vaginal e uterina e pelo sêmen.

Mycobacterium bovis pode causar uma grande variedade de sinais clínicos tanto em humanos como em animais, associando-se a variação da localização das lesões (COSIV et al., 1998). Comumente a tuberculose em bovinos é inaparente, apresentando poucos ou ausência de sintomas, mas em casos avançados os hospedeiros podem apresentar dispneia (MATTHIAS, 1988).

Em 2001, o Ministério da Agricultura, Pecuária e Abastecimento (MAPA) instituiu no Brasil o Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose Animal (PNCEBT), buscando diminuir a prevalência das enfermidades a partir de medidas sanitárias. Este programa visa o controle dos animais em risco, certificação das propriedades livres ou monitoradas, aprimoramento para médicos veterinários que atuem nessa área, educação sanitária e diagnósticos laboratoriais (BRASIL, 2006).

O PNCEBT se baseia no controle através do diagnóstico, sacrifício ou abate sanitário dos animais infectados e vacinação para brucelose. Ademais, este programa sugere a desinfecção de currais, estábulos, manejo de pastagens, desinfecção de solos e veículos de transporte de animais, como medidas preventivas.

O diagnóstico de *B. abortus* é baseado nos aspectos clínicos, epidemiológicos e principalmente nos laboratoriais, sendo estes últimos divididos

em diretos e indiretos. Os métodos diretos incluem o isolamento do agente, identificação por imuno-histoquímica ou PCR (BRASIL, 2006).

Os métodos indiretos para brucelose bovina são bastante amplos. O PNCEBT define como oficiais os seguintes testes: Antígeno acidificado tamponado (AAT), Teste do Anel em Leite (TAL), 2-Mercaptoetanol (2-ME), Soroaglutinação Lenta (SAL), Fixação de Complemento (FC) e Teste de Polarização Fluorescente (FPA), sendo que os dois primeiros testes são realizados como triagem; e os quatro últimos são considerados confirmatórios (BRASIL, 2006). Embora os testes acima descritos apresentem em geral bons parâmetros de sensibilidade e especificidade, o diagnóstico definitivo de um animal suspeito pode levar alguns dias, devido a necessidade de realização de testes de triagem e confirmatório.

O diagnóstico clínico da tuberculose bovina torna-se difícil haja vista, muitas vezes, sinais inaparentes e condição crônica do portador. Portanto, o uso de métodos laboratoriais é de extrema importância. Após a morte do animal, a necrópsia pode ser realizada e associada a coleta de material, para assim ser realizada a histopatologia, PCR e isolamento bacteriano. Entretanto, o teste tuberculínico é a principal metodologia recomendada pelo MAPA para o diagnóstico de tuberculose bovina in vivo. Tal teste in vivo, se caracteriza pela aplicação de tuberculina e verificação da reação de hipersensibilidade (para os animais positivos). Com isso, esta última ferramenta pode ser dividida em três tipos: Teste da Prega Caudal (TPC), Teste Cervical Simples (TCS) e Teste Cervical Comparativo (TCC) (BRASIL, 2006). Embora de grande importância para identificação de animais infectados, os testes baseados em tuberculinização apresentam inconvenientes, sejam relacionados a sensibilidade e especificidade, seja relacionado a praticidade de execução.

Diante do exposto, a busca por novas metodologias de diagnóstico, é justificada, para o aprimoramento do controle das enfermidades. Neste cenário, técnicas de espectroscopia óptica associadas a metodologias de análise baseadas em aprendizado de máquina surgem como promissoras. Estas se baseiam na interação da radiação com a matéria gerando informações que podem ser analisadas para classificar e determinar as propriedades do respectivo material. O estímulo da luz é capaz de fazer com que o material em

análise mude de um estado fundamental para excitado. Com isso, a partir da medição da radiação emitida ou a quantidade de radiação absorvida, é possível obter resultados e informações importantes sobre a substância em questão (SKOOG et al., 2005).

A espectroscopia óptica são um conjunto de técnicas que podem ser executadas a partir da radiação por luz infravermelha, ultravioleta ou visível, entre outras (MARTINHO, 1994). A partir disso, a espectroscopia pode ter inúmeras aplicações, com a análise de variados materiais assim como sistemas mais complexos que podem envolver células e tecidos. Esta técnica pode ser muito útil, pois é capaz de analisar amostras nos três estados físicos (líquido, sólido e gasoso) e estando a matéria pouco ou sem prévia preparação. Ademais, há um baixo custo tanto para suporte quanto para a manutenção dos equipamentos por esta técnica (SILVA, 2007).

Assim a técnica de espectroscopia pode ser útil no diagnóstico de inúmeras doenças, assim como doenças causadas por bactérias. Com isso há a hipótese de que tal tecnologia poderia contribuir no diagnóstico de brucelose e tuberculose bovina resultando em ótima performance na confiabilidade dos resultados bem como na rapidez e baixo custo para a sua funcionalidade.

Após a obtenção dos espectros de uma determinada amostra, pode ser utilizado uma ferramenta de grande valia: análise de componentes principais (PCA). Esta análise é um método fatorial no qual os fatores são relacionados a variância total. O objetivo da PCA é utilizar muitas variáveis e encontrar combinações destas para produzir índices que descrevam a variação dos dados (ARAUJO et al., 2009). Este sistema, em outras palavras, permite reter as informações relevantes e mostrar, de modo mais simples, cada variável individualmente. Ou seja, um espaço multidimensional é transformado em um subespaço menor, criando assim, em termos geométricos das componentes principais, a covariância das variâncias usando baixo número de autovetores (GODINHO et al., 2008). E quando a PCA é associada a métodos de aprendizado de máquina, temos uma metodologia poderosa para identificação de padrões que podem auxiliar na distinção entre amostras positivas e negativas.

2. REVISÃO DE LITERATURA

2.1 Brucelose bovina no brasil

2.1.1 Etiologia e patogenia

Brucella é o gênero da bactéria que pode causar a doença brucelose, comumente associada a diversos hospedeiros (suínos, bovinos, caninos etc.) incluindo humanos (MORENO et al., 2006). *Brucella abortus* é o principal agente causador de brucelose em bovinos, podendo gerar inúmeros prejuízos a bovinocultura, comumente relacionado a sua capacidade de causar infertilidade no hospedeiro, aborto, diminuição na produção de leite e outros prejuízos de ordem reprodutiva (HARMON et al., 1988; XAVIER, 2009; NETA et al., 2010).

B. abortus é uma bactéria gram-negativa, não formadora de esporos, não móvel, intracelular facultativa, aeróbia e flagelada (LAGE et al., 2008; BRASIL, 2006). Tal bactéria apresenta morfologia de colônia lisa, ou seja, há a presença de lipídeo A, núcleo oligossacarídeo e cadeia O na composição do LPS (lipopolissacarídeo) que constitui a parede celular. Ademais, tal característica morfológica tem relação direta com a virulência do microrganismo, sendo que o grupo das bactérias lisas são mais patogênicas (PAULIN et al., 2003; LAGE et al., 2008). Embora os gêneros sejam aeróbios, uma atmosfera com concentração de 5 a 10 % de CO₂ propicia o crescimento de algumas espécies. A multiplicação deste patógeno acontece na faixa de 20 a 40°C, sendo a temperatura ideal de 37°C, e um pH entre 6.6 e 7.4 (PAJUABA, 2006; OIE, 2009). Este microrganismo resiste bem a inativação no ambiente, levando a um alto potencial de transmissibilidade, podendo permanecer por vários meses na água, em vestimentas, feno, locais secos, temperaturas baixas e determinados materiais biológicos (OIE, 2016).

As espécies do gênero são sensíveis à acidez e ao calor, podendo ser eliminadas em até 15 minutos quando em contato com compostos fenólicos a 2,5%, permanganato de potássio (1:5000) e soluções de formaldeído a 2%. Além disso, o carbonato de cálcio (1:10) consegue eliminar os microrganismos em 30 minutos e o álcool a 70% elimina de imediato (PAULIN et al., 2003; OIE, 2009).

Em alimentos como carne, *Brucella spp.* pode manter-se viável durante meses, sendo resistentes ao congelamento e refrigeração, bem como a

acidificação do meio. Portanto, a eliminação, neste caso, só ocorre com o uso do calor e em condições de pH inferior a 4,0 (PESSEGUEIRO et al., 2005).

Em produtos lácteos e no leite, tal bactéria, pode sobreviver de 15 a 90 dias, dependendo do pH, presença de outros microrganismos e temperatura. A fervura do leite bem como os métodos de esterilização e processos de pasteurização contribuem para a erradicação de *Brucella spp.* (CARVALHO et al., 1995; PAULIN et al., 2003; BRASIL, 2006).

A mais comum porta de entrada de *B. abortus* é a mucosa orofaríngea, entretanto há de se considerar as mucosas nasofaríngeas, conjuntiva ocular, oral e em soluções de continuidade de pele (BRASIL, 2006). Geralmente a transmissão ocorre quando o bovino está em contato com fetos abortados, líquidos vaginais, envoltórios fetais, fômites e ingestão de água e alimentos contaminados (OIE, 2009).

Após a entrada da bactéria no hospedeiro, os macrófagos fagocitam estes microrganismos levando-os aos linfócitos regionais, e estes, por sua vez, sofrem multiplicação do agente podendo originar um quadro de hiperplasia e linfadenite (BISHOP et al., 1994; LAGE et al., 2008; NETA et al., 2010). Ademais, após a circulação nos primeiros linfonodos, *B. abortus* pode, através do sangue ou linfa, acometer outros linfonodos, como os supramamários, e órgãos como o fígado e baço, albergando-se por longos períodos e sobrevivendo a resposta imune do hospedeiro (HARMON et al., 1988; GORVEL et al., 2002; CAMPAÑA et al., 2003; LAGE et al., 2008; LIRA, 2008; MATRONE et al., 2009; XAVIER, 2009).

Indubitavelmente, um grande fator responsável pela patogenicidade do microrganismo, é sua capacidade de escape da resposta imune do hospedeiro. *B. abortus* é capaz de permanecer no interior dos macrófagos sintetizando enzimas antioxidantes, as quais somada à produção de guanosina 5' monofosfato-GMP e adenina, impedem a degranulação dos macrófagos, e consequentemente a destruição do agente (ARÉSTEGUI et al., 2001; BALDWIN et al., 2002; NETA et al., 2010).

Além disso, este patógeno pode causar reações inflamatórias diversas que podem resultar em esplenomegalia, hepatomegalia e até hiperplasia linfóide, podendo ocorrer sinais clínicos no aparelho locomotor associado a acometimento de músculos, bursas, tendões e articulações (MASCARENHAS,

2017). Os tecidos acometidos, devido a inflamação, além da modulação por macrófagos, há a infiltração de plasmócitos e linfócitos, podendo resultar em focos de necrose somado ao desenvolvimento de capsulas devido a proliferação de tecido conjuntivo. Contudo, a preferência do microrganismo é por órgãos que oferecem elementos para o desenvolvimento do seu metabolismo, como o eritritol. Este elemento, por sua vez, é encontrado principalmente no útero gravídico, tecidos osteoarticulares, tecido mamário e órgãos reprodutivos do macho (RIBEIRO et al., 2008; XAVIER, 2009).

De acordo com Silva et al. (2005), as *Brucellas* possuem afinidade pelos trofoblastos devido à grande quantidade de eritritol e progesterona produzida pela placenta. Ademais, a infecção no útero através da via hematogênica, pode causar diferentes intensidades de alterações associado ao fator do tempo de gestação. A concentração de eritritol no bovino varia de acordo com o período gestacional, sendo que há maior quantidade de tal componente próximo ao parto, e com isso aumentando a chance de infecção e multiplicação do microrganismo (CARTER et al., 1991; LAGE et al., 2008). Com isso, as lesões necróticas e inflamatórias na placenta, causam descolamento de cotilédones somado a lise de vilosidades, comprometendo a circulação entre mãe e filhote, e consequentemente impossibilidade de troca de nutrientes e oxigênio, resultando em variados danos incluindo o subdesenvolvimento do feto e até o aborto. Além disso, outras manifestações materno-fetal podem ocorrer, tais como: retenção de placenta, nascimento de bezerros fracos e endometrite crônica, que pode acarretar problemas de fertilidade (BISHOP et al., 1994; PAULIN et al., 2003; LAGE et al., 2008; XAVIER, 2009).

Os machos infectados por *B. abortus*, podem apresentar aumento de volume testicular, além de comprometimento de vesículas seminais, ampolas e epidídimo. As lesões inflamatórias podem evoluir para necrose causando atrofia dos testículos e consequentemente levando o animal a quadro de infertilidade ou subfertilidade (GORVEL et al., 2002; PAULIN et al., 2003; LAGE et al., 2008; NOZAKI, 2008).

As lesões de aparelho locomotor causadas por *B. abortus*, estão principalmente relacionadas aos tendões, bursa, músculos e articulações. Ademais, as artrites carpianas e tarsianas são comuns, assim como espondilites

301 e bursite, podendo atingir bainha dos tendões e medula óssea (PAULIN et al.,
302 2003).

303 **2.1.2 Técnicas de diagnóstico**

304 Os métodos indiretos ou sorológicos constituem importantes ferramentas
305 para o diagnóstico da brucelose bovina (POESTER et al., 2005).

306 O PNCEBT definiu como oficiais os seguintes testes: Teste do Antígeno
307 Acidificado Tamponado (AAT), Teste do Anel do Leite (TAL), Teste do 2-
308 Mercaptoetanol (2-ME), Soroaglutinação Lenta (SAL), Fixação do Complemento
309 (FC) e Fluorescência de Polarização (FPA), sendo os dois primeiros definidos
310 como testes de triagem e os quatro últimos como confirmatórios (BRASIL, 2006).
311 Sem dúvidas, tais testes não apresentam total sensibilidade, e por isso há a
312 necessidade, muitas vezes, da associação entre eles em busca de melhor
313 eficiência na detecção de animais positivos (COSTA et al., 2011; OLIVEIRA,
314 2003).

315 O AAT é um teste indireto de triagem na qual se caracteriza pela
316 simplicidade da sua realização, além de que é sensível e rápido. Entretanto, a
317 aplicação da vacina B19, pode resultar em animais falso positivos quando
318 submetidos a tal teste (HUBER et al., 1986; TANYA et al., 1993; ABIMERHI et
319 al., 1998; MEGID et al., 2000; FOSGATE et al., 2002).

320 Já o TAL consiste em um teste bastante sensível no qual quando há a
321 combinação de antígeno-anticorpo, forma-se um anel azulado indicando assim
322 resultado positivo. Este teste é realizado para diagnóstico e monitoramento de
323 rebanho de gado de leite, sendo que os positivos, devem ser testados
324 individualmente por AAT, 2-ME e SAL, para se descobrir qual ou quais estão
325 infectados. Contudo, o teste de anel em leite pode apresentar falsos positivos
326 em animais portadores de mamites, presença de leites ácidos ou em colostros
327 (BRASIL, 2006).

328 O 2-Mercaptoenol (2-ME) é um teste confirmatório altamente específico,
329 pois devido a degradação de anticorpos IgM, há a detecção somente de IgG o
330 que indica, quando presente, uma infecção crônica. Além disso, deve ser
331 executada em paralelo com a prova lenta em tubos. Ademais, a prova lenta,

também chamada de Soroaglutinação em Tubos (SAT), embora possa identificar uma alta proporção de animais positivos, pode apresentar resultados falso-negativos em infecções crônicas somado a demora (aproximadamente 48 horas) para leitura (BRASIL, 2006).

A Organização Mundial de Saúde Animal (OIE) recomenda a FC como teste referência para o trânsito internacional de animais (DAJER et al., 1999). Este teste tem sido utilizado em muitos países como ferramenta de controle e erradicação (NIELSEN, 2002), pois além da alta especificidade, há menor interferência de anticorpos vacinais quando comparado aos testes do AAT e 2-ME (ALTON et al., 1988).

A FPA fundamenta-se na diferença de rotação entre a molécula de antígeno (marcado com fluorocromo) e a conjugação antígeno-anticorpo. Moléculas individualizadas (somente o antígeno) giram a uma velocidade maior, resultando em maior despolarização da luz, e moléculas maiores (antígeno ligado ao anticorpo) giram a uma menor velocidade, produzindo-se menor despolarização luminosa. A detecção de diferenças de polarização consequentemente consegue diferenciar amostras positivas e negativas (NIELSEN, 2002). Ademais, tal método indireto tem boa sensibilidade e especificidade, pode ser realizado a campo e com pouco tempo de execução (2 minutos) (NIELSEN, 2002; POESTER et al., 2005; OIE, 2008). Segundo Mathias et al., (2010) a FPA apresentou sensibilidade de até 97,3% e especificidade de até 98,3 %. Contudo, tal técnica apresenta a desvantagem do alto custo de equipamentos e reagentes somado ao fato de que é necessária a importação dos itens.

De acordo com Meirelles-Bartoli et al. (2010) foi realizado um estudo comparativo entre os testes sorológicos preconizados pelo PNCEBT para o diagnóstico de brucelose bovina. Então, foram analisadas 1061 amostras de soros bovinos e comparou-se as provas de AAT, a combinação de SAL e 2-ME e FC. Constatou-se sensibilidade relativa de 99,6%, 98,8% e 91,1%, respectivamente, para o AAT, 2-ME e a FC, e especificidade relativa de 83,9%, 96,2% e 100%. Portanto, observou-se boa concordância entre os testes de triagem e os testes confirmatórios, e ótima concordância quando os confirmatórios foram comparados entre si. Contudo, houve soros com titulação

elevada em um dos testes confirmatórios e resultado negativo em outro, ratificando a suposição de que o diagnóstico de brucelose bovina é mais confiável quando são realizados vários testes.

Os diferentes métodos diretos (isolamento bacteriano, imuno-histoquímica e PCR) para diagnóstico de *B. abortus*, podem apresentar vantagens e desvantagens (BRASIL, 2006). A realização de isolamento bacteriano através de material de aborto, sêmen e líquido sinovial de articulações comprometidas apresentam resultados muito bons. Porém, é necessário colheita, transporte e processamento adequado (LAGE et al., 2006). Ademais, tal técnica possui como vantagem a alta especificidade, sendo capaz de identificar diferentes espécies, contribuindo assim para avanços na epidemiologia da doença (LAGE et al., 2008). Entretanto, além da baixa sensibilidade, o isolamento requer tempo de execução de 3 a 5 dias de incubação, em alguns casos 21 dias antes que um resultado esteja disponível (WALKER, 2003). Além disso, há como desvantagens a necessidade de instalações e equipamentos de bioproteção de nível 3, pois, tal técnica possui riscos de infecção aos profissionais que manuseiam material altamente contaminado e infeccioso (POESTER et al., 2005).

A imuno-histoquímica, por sua vez, é considerada versátil e prática, pois apenas requer a adição de um anticorpo específico com um determinado marcador, na amostra preparada para histopatologia. Com isso, é possível além da identificação do agente, a visualização das lesões ocasionadas pelo microrganismo (XAVIER, 2009). Todavia, este método direto é de execução complexa e laboriosa, gerando falhas significativas e sendo necessário extrema cautela para interpretação dos resultados (BRANDTZAEG, 1998).

A PCR tem se popularizado nos últimos anos como útil na identificação de *Brucella spp.*, a partir de material de aborto, sangue e sêmen (QUIEPO-ORTUNO et al., 1997). Possui como vantagens a alta sensibilidade e especificidade, permite trabalhar com a bactéria inativada além de demandar baixo tempo de execução quando comparado aos métodos citados anteriormente (FEKETE et al., 1990; BRICKER, 2002). As principais desvantagens são a presença de fatores inibidores de amplificação nas amostras (proteínas, lipídeos e carboidratos), o alto custo de equipamentos e

reagentes, e demanda por profissionais qualificados (LAGE et al, 2006). Leyla et al., (2003) utilizaram a PCR baseado no gene IS711 para detectar *Brucella spp.* em material de conteúdo estomacal de fetos ovinos, com sensibilidade de 97,4%, e especificidade de 100% quando comparada ao isolamento microbiológico. Ademais, Seónadh et al. (2006) avaliaram o método convencional e a PCR em tempo real (qPCR) para detecção de *B. abortus* em sangue, linfonodos e leite de vacas infectadas, concluindo que é possível detectar *B. abortus* a partir de amostras de linfonodo e leite com eficiência. No entanto, a partir de sangue, não se observou grande eficiência.

2.1.3 Distribuição geográfica

Indubitavelmente, o conhecimento epidemiológico acerca da distribuição da doença sobre estados e regiões, torna-se muito importante para a implementação de um programa que objetiva o controle e a erradicação. Ademais, a razão para tal afirmativa é a de que, com o conhecimento da distribuição da morbidade, permite-se escolher as melhores estratégias, relacionadas especificamente com as diferentes frequências de prevalência das subpopulações estudadas. Além disso, o acompanhamento das diferentes localidades, fomenta para um melhor estudo, queixando-se de forma racional a necessidade de correções bem como evitando o desperdício de tempo e recursos (CHATE et al., 2009).

Em 2001 começou o inquérito epidemiológico para conhecer a prevalência de brucelose bovina nos estados da BA, DF, ES, GO, MA, MT, MS, MG, PB, PR, PE, RJ, RS, RO, SC, SP, SE e TO, entretanto AC, AL e PA tinham previsão de finalizar os estudos até o ano de 2020 e ficando sem inquéritos apenas AP, AM, CE, PI, EM e RR (BRASIL, 2020). A seguir, a partir das figuras 1 e 2, observa-se os resultados dos inquéritos realizados pelas unidades federativas (UF).

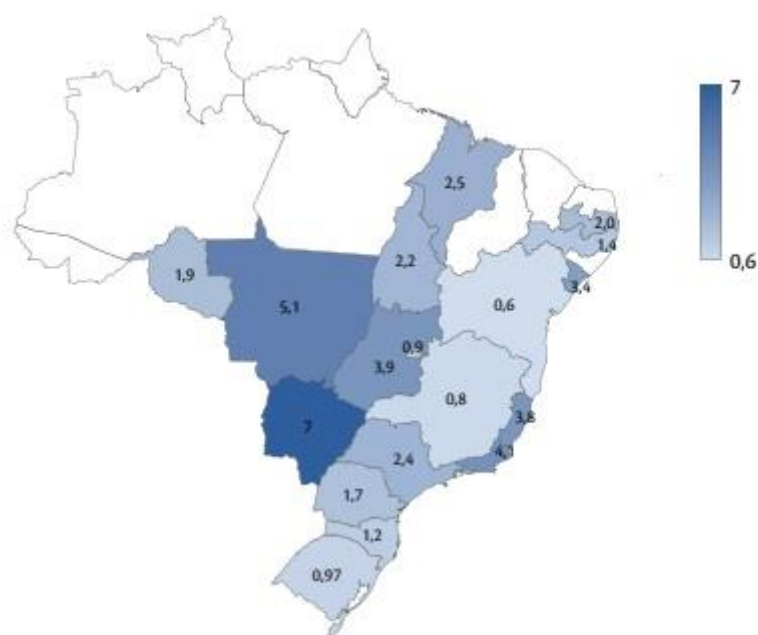


Figura 1. Prevalência de casos de brucelose bovina por UF.
 Fonte: (BRASIL, 2020).

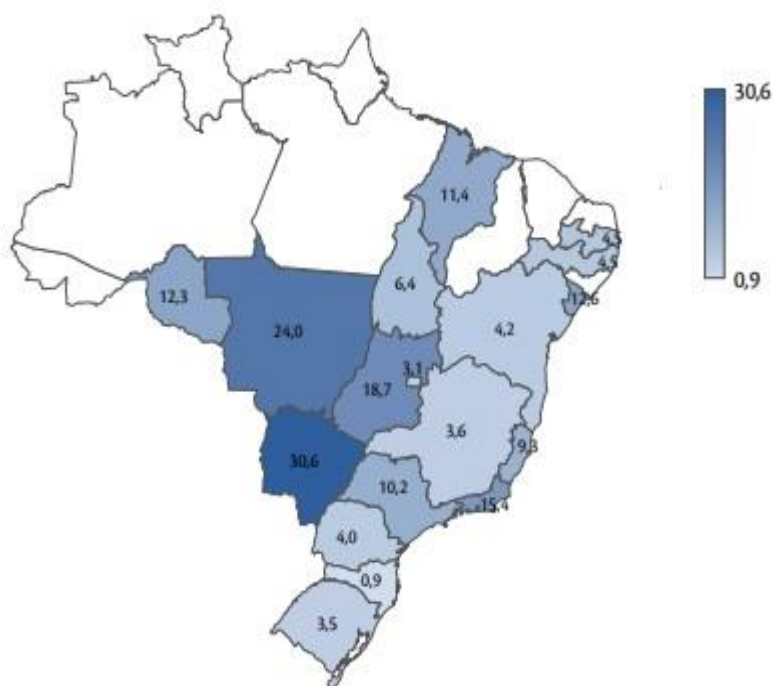


Figura 2. Prevalência de focos de brucelose bovina por UF.
 Fonte: (BRASIL, 2020).

2.2 Tuberculose bovina no Brasil

2.2.1 Etiologia e patogenia

A tuberculose bovina é uma importante doença infecciosa de bovinos, causada pela bactéria *Mycobacterium bovis* e que tem prevalência relevante principalmente em países em desenvolvimento. Geralmente a doença se caracteriza por quadro crônico, raramente apresentando evolução aguda. Este bacilo acomete principalmente bovinos, mas também pode infectar outras espécies de animais silvestres e domésticos, assim como o ser humano (CENTRO PANAMERICANO DE ZOONOSIS, 1988).

Mycobacterium bovis é um microrganismo que se caracteriza por ser imóvel, não capsulado, intracelular facultativo, bacilos curtos e não flagelados (SILVA et al., 2011). Além disso, este agente é muito resistente, sendo capaz de sobreviver em pasto, esterco e estábulo por até 2 anos, e por até 10 meses em materiais biológicos contaminados. Ademais, é sensível a pasteurização do leite e podem ser eliminados por alguns agentes desinfetantes, como álcool, hipoclorito de sódio e formol (JORGE, 2010).

A tuberculose bovina pode causar lesões do tipo granulomatosas nodulares, denominadas tubérculos, e isso se caracteriza como principal sinal clínico (SILVA et al., 2011). Mas também, vale ressaltar que a localização das lesões corresponde, de modo geral, as rotas de infecção do animal pelos bacilos. A infecção por via aerógena (inalação de aerossóis contaminados) é a mais comum (BRASIL, 2006), correspondendo assim a lesões que se restringem a cavidade torácica. Entretanto, quando a infecção ocorre por via digestiva (ingestão de pastagem, água e fômites contaminados), lesões podem aparecer em linfonodos mesentéricos e faríngeos e posteriormente a disseminação para todo o organismo. Ademais, bezerros podem ser infectados ao ingerir o leite materno contaminado (JORGE, 2010).

De acordo com (THOMAZ, 2006), as lesões iniciadas nas junção bronquíolo-alveolar disseminam-se para os alvéolos e linfonodos brônquicos, podendo persistir no local, progredir ou regredir. A queda da imunidade do animal pode estar diretamente relacionada a disseminação da infecção para outros órgãos. Além disso, a dispersão da infecção pode ocorrer de forma miliar

(quando ocorre de maneira intensa, com grande quantidade de bacilos na circulação) e protraída (através de vasos linfáticos e sanguíneos atingindo linfonodos, sistema nervoso, fígado, rim, baço, ossos, entre outros tecidos) (BEER, 1998).

Geralmente os animais que estão infectados por *M. bovis* apresentam uma evolução crônica da doença, permanecendo assintomáticos até o estágio terminal. Entretanto, em algum momento, o animal infectado pode apresentar sintomatologia tais como linfonodos aumentados, emagrecimento e dificuldade respiratória. Além disso, pode ocorrer diarreia, mastite, lesões na genitália e infertilidade (CARVALHO, 2013; JORGE, 2010; SALAZAR, 2005).

2.2.2 Técnicas de diagnóstico

Para o diagnóstico de *M. bovis*, os testes são extremamente importantes haja vista que o diagnóstico clínico, muitas vezes, torna-se difícil pois com a evolução crônica da doença há uma ausência de sinais clínicos identificáveis. Dessa forma, os testes de tuberculização (TCC, TCS e TPC), são os principais métodos para triagem e confirmação de casos in vivo de tuberculose em bovinos. Para diagnóstico post-mortem, tecidos como linfonodos de cabeça e tórax, pulmão e fígado, podem ser utilizados para histopatologia, testes moleculares e isolamento bacteriano (SALES, 2012; BRASIL, 2006).

O MAPA recomenda os testes tuberculínicos, e estes por sua vez, são os métodos de diagnóstico mais comuns no Brasil (BRASIL, 2006). Os testes consistem na inoculação intradérmica de tuberculina, e posterior aferição da resposta de hipersensibilidade do tipo IV, por meio do aumento da espessura da dobra de pele no local da inoculação. Quando os animais estão infectados, há uma intensa reação que aumenta significativamente a espessura da pele alguns dias após a inoculação (MONAGHAN et al., 1994). O procedimento de tuberculinização pode ser dividido em três testes, sendo que o teste da prega caudal (TPC) é empregado em gado de corte; o teste cervical simples (TCS), é também utilizado como teste de triagem, mas para gado de leite; e há o teste confirmatório, para animais com resultados positivos no TCS e TPC, compreendido pelo teste cervical comparativo (TCC) (BRASIL, 2006).

Entretanto, os testes intradérmicos podem apresentar algumas desvantagens, pois, por se tratar de testes *in vivo*, há dificuldade na aplicação em rebanhos com alto contingente, e com isso há dificuldades na realização de estudos epidemiológicos (MENZIES, 1999; LIU et al., 2007). Ademais, tais métodos não são 100% específicos, podendo ocorrer reações falso-positivas, devido a sensibilização dos animais com micobacterias ambientais, podendo acarretar sacrifício desnecessário do animal (WOOD et al., 1991; MONAGHAN et al., 1994). Somado a isso, em caso de reações indeterminadas, os animais devem ser testados novamente apenas após 60 dias a partir da primeira inoculação, sendo necessário novamente a contenção dos animais. (BRASIL, 2004; AAGAARD et al., 2006). Reações falso-negativas, para tal método indireto, podem acontecer em animais anérgicos cronicamente infectados, tornando estes animais possíveis fontes de infecção para o rebanho.

A técnica de ELISA (*Enzyme-Linked Immunosorbent Assay*), para diagnóstico de tuberculose bovina, se mostra útil para identificação de animais positivos anérgicos, e como ferramenta complementar aos testes que avaliam a imunidade celular (WORLD HEALTH ORGANIZATION, 1994). De acordo com Waters et al., (2011) um teste ELISA utilizando antígenos recombinantes MPB70 e MPB83 forneceu sensibilidade de 63,0% e especificidade de 98,0% em bovinos. Ademais, Meneses et al., (2015) avaliaram proteínas recombinantes para detecção de anticorpos contra *M. bovis* por ELISA, e concluíram que a quimera com fragmentos hidrofílicos de MPB70, MPB83 e ESAT-6 de *M. bovis*, apresentou melhor resultado (sensibilidade de 83,2% e especificidade de 86,5%) quando avaliaram os soros bovinos previamente testados pelo TCC. Um outro estudo, foi realizado por Farias et al., (2012), no qual avaliou-se o teste ELISA baseado em proteínas recombinantes P27 e MPB70 contra anticorpos de *M. bovis*. Neste último estudo citado, a sensibilidade e especificidade do teste foram respectivamente 88,7% e 94,6% para MPB70 e 98,1% e 91,9% para P27. Entretanto, tal teste pode apresentar algumas desvantagens como necessidade de muito tempo para a sua execução, alto custo e necessidade de laboratório especializado (FRÁGUAS et al., 2008).

O teste de interferon gama bovino objetiva detectar a reação imune celular a infecção por *M. bovis* (CENTRO PANAMERICANO ZOONOSIS, 1988). Tal

método é definido pela dosagem de uma citocina (interferon gama) presente em amostras de sangue de animais infectados, quando há a estimulação prévia dos linfócitos com PPD bovino e aviário, sendo que, animais não infectados não produzem a citocina (JONES et al., 1992; WOOD et al., 1991). As vantagens de tal técnica baseiam-se em ser um procedimento não invasivo, permitir a realização várias vezes e sem intervalo de tempo, e necessidade de contenção do animal apenas uma vez. Porém, é um diagnóstico de alto custo, restrito tempo para o processamento das amostras (8 horas), e possibilidade de reações cruzadas com outras microbactérias (ABRAHÃO, 1998; KANTOR et al., 1994).

Outras ferramentas como a cromatografia líquida de alta resolução (HPLC) para análise do ácido micólico, foi sugerida como técnica de identificação de microbactérias em culturas (BUTLER et al., 1991). Mas, esse ensaio não é capaz de diferenciar por exemplo *M. bovis* de *M. tuberculosis*, possuindo então pouca especificidade (FLOYD et al., 1992).

Os exames bacteriológicos e histológicos são importantes na comprovação do diagnóstico realizado a campo, assim como a tuberculinização, e com isso requerem que as amostras de animais suspeitos sejam coletadas de forma rápida e adequada (CENTRO PANAMERICANO DE ZOONOSIS, 1988; ORGANIZACIÓN PANAMERICANA DE LA SALUD, 1995). O isolamento bacteriano apresenta baixa sensibilidade, exigindo grande quantidade de bacilos viáveis e demora no tempo de execução (tempo de crescimento da cultura varia entre 1 e 12 semanas) (CORNER, 1994). A histologia, por sua vez, embora seja um método rápido e barato, requer grande quantidade de microrganismos e algumas lesões causadas por outras bactérias podem ser semelhantes as causadas por *M. bovis* (HAAGSMA, 1995; WARDS et al., 1995; ZANINI et al., 2001).

A PCR como técnica de diagnóstico para tuberculose pode ser realizada usando isolados de cultivos e materiais clínicos (WIT et al., 1990; SAKAMOTO et al., 1999). Embora tal técnica apresente ótimos resultados de sensibilidade e especificidade, há um alto custo somado a complexidade de execução. Esses fatores têm sido relatados como restrição para o uso no diagnóstico de tuberculose bovina (KLAFTER et al., 1998; VITALE et al., 1998; ZANDEN, 2002). Os primeiros estudos utilizando PCR para diagnóstico de *M. bovis* oriundos de

amostras de bovinos, denotam de 1995 e com isso tal ferramenta mostrou-se eficaz na identificação de bacilos inviáveis bem como na rapidez de execução, mas as desvantagens se restringiam aos métodos de extração (COLLINS et al., 1994; RORING et al., 2000).

2.2.3 Distribuição geográfica

Em 2003 iniciou-se os inquéritos epidemiológicos com o objetivo de identificar a prevalência de tuberculose bovina nos estados do Brasil. Até o ano de 2018 foram realizadas as avaliações nas seguintes Unidades Federativas (UF): BA, DF, ES, GO, MT, MS, MG, PR, PE, RS, RO, SC, SP e TO, entretanto no presente estudo as UF de AC, PA e SE estavam com inquéritos em andamento (Figuras 3 e 4) (BRASIL, 2020).

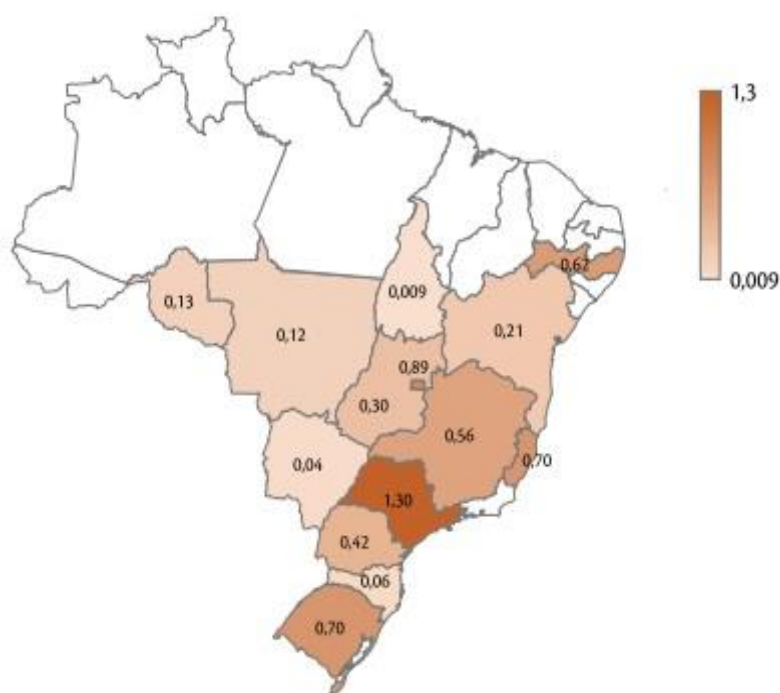


Figura 3. Prevalência de casos de tuberculose bovina por UF.

Fonte: (BRASIL, 2020).

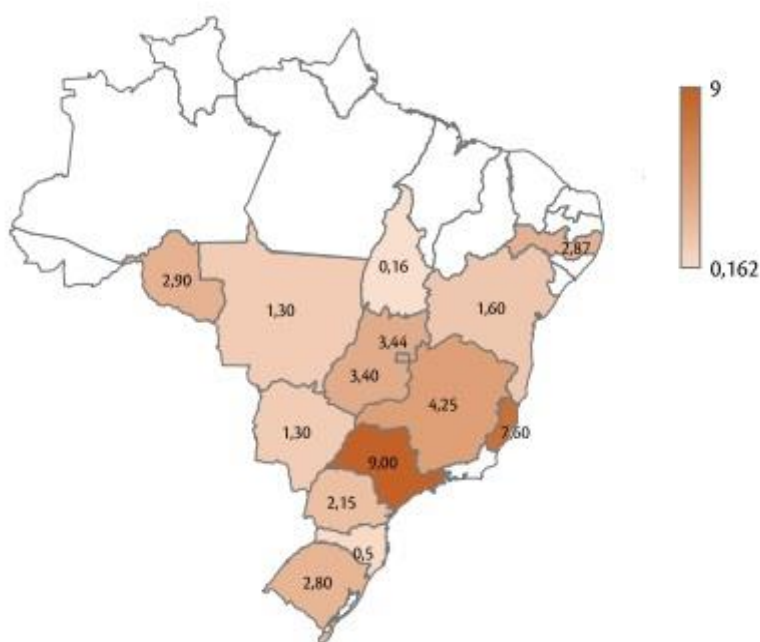


Figura 4. Prevalência de focos de tuberculose bovina por UF.

Fonte: (BRASIL, 2020).

2.3 Espectroscopia UV-Vis

A espectroscopia UV-Vis é uma técnica capaz de medir a absorção de radiação UV e/ou luz visível a partir das moléculas que compõe a substância analisada. Tal técnica funciona sob um intervalo de longitude de onda de 200 a 800nm, resultando em informações sobre os estados de energia da amostra em questão (PERKAMPUS, 1992).

Espectrofotômetro é o instrumento usado para obter espectros de uma determinada amostra (OWEN, 2000). Além disso, dentre as espectroscopias, a UV-Vis é a mais utilizada no mundo, tendo aplicações diversas, contribuindo por exemplo em áreas biológicas na quantificação de proteínas determinação de atividades enzimáticas (BACON et al., 2004), avaliação de viabilidade celular (STOCKERT et al., 2012) e entre outras atividades.

Dentre as variadas aplicações, o uso de UV-Vis associado ao infravermelho, podem ser utilizadas na medicina resultando em efeitos gerais de fôtonica no corpo, cabelos, ossos e pele (ANDERSON et al., 1981; LIEBOLD et al., 2000). Ademais, a utilização dessa tecnologia pode ser útil no tratamento de eritema que é causado por determinadas doenças (KOLLIAS et al., 1994). Mas

também vale destacar que diagnósticos prévios tem sido consequência do uso desta tecnologia, por exemplo na detecção de pré-câncer e câncer de colo de útero (BURGHARDT, 1984).

Neste cenário, o uso da técnica de espectroscopia ultravioleta-visível a partir de biofluidos para diagnóstico de doenças infecciosas tais como a brucelose e tuberculose bovina, pode resultar em um método de análise sensível, específico e rápido, que poderá ser explorado em uma nova etapa do programa nacional sanitário do Brasil (PNCEBT), contribuindo para identificação de infectados em cenários de baixas prevalências, e erradicação da doença.

2.4 PCA e Machine Learning

Análise de componentes principais (PCA) é uma ferramenta que possibilita a partir de um conjunto de dados, a redução da sua dimensionalidade em uma forma interpretável possibilitando a manutenção da maioria das informações e com isso preservar o máximo de variabilidade possível (JOLLIFFE et al., 2016). Em outras palavras, a PCA é uma técnica estatística de análise multivariada que é capaz de utilizar e transformar um montante de dados originais (variáveis) em um conjunto menor de variáveis (HONGYU et al., 2016).

A PCA tem como vantagens a exclusão da multicolinearidade das informações, objetivando a mudança de dados originais intercorrelacionados em um novo conjunto de dados não correlacionados (componentes principais). Ademais, reduz as variáveis a eixos perpendiculares que explicam a variação dos dados de forma independente e decrescente. As desvantagens de tal técnica se restringem a sensibilidade a *outliers* não recomendada quando se tem muitos zeros na matriz e dados ausentes. A seguir, na figura 5, é possível observar uma representação gráfica de análise de componentes principais de amostras positivas e negativas para tuberculose bovina obtidas através de espectros UV-Vis.

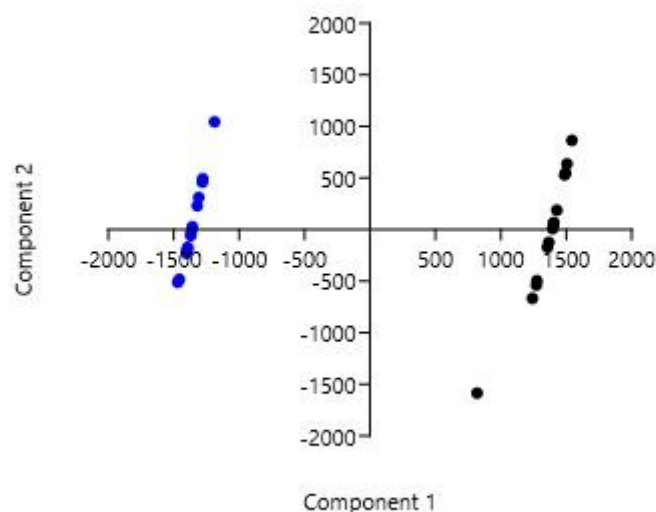


Figura 5. PCA realizada com os dados de espectroscopia UV-Vis de amostras positivas (preto) e negativas (azul) para *Mycobacterium bovis*.
Fonte: Arquivo pessoal.

O aprendizado de máquina (*machine learning*, ML), é uma ferramenta estatística que auxilia na busca de conhecimento científico através da identificação automatizada em diferentes conjuntos de dados e assim, em outras palavras, busca detectar padrões dos dados (HEY et al., 2009). Portanto, com o reconhecimento de diferentes padrões dos dados, é possível prever informações desconhecidas ou ajudar no processo de tomada de decisões em situações de incertezas (MURPHY, 2012). Além disso, segundo outras definições, ML melhora com a experiência em tarefas determinadas, mas sem ser totalmente colocada para efetuar tal função (SAMUEL, 1959; HASTIE et al., 2001). Tal técnica evoluiu a partir da inteligência artificial (IA), e basicamente, esta última, por sua vez, é qualquer auxílio que permite os computadores imitar a inteligência humana. A figura 6 exemplifica melhor a relação entre IA e ML (SCHLEDER et al., 2021).

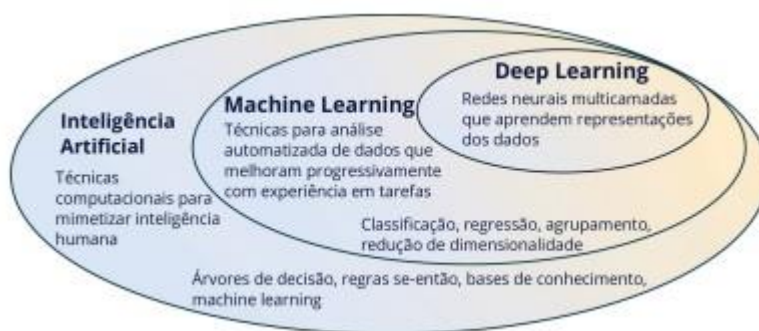


Figura 6. Descrição hierárquica e exemplos de técnicas de IA e suas subáreas ML e Deep Learning (DL).
Fonte: (SCHLEDER et al., 2021).

Assim, ML pode ser usada para duas finalidades: a primeira é com objetivo de solucionar problemas os quais outros métodos, como os tradicionais não conseguem resolver; a segunda seria para melhorar a resolução de questões que já foram resolvidas, seja de uma forma melhor e mais robusta, mais econômica, mais rápida, ou tudo simultaneamente. A seguir, na figura 7, há a demonstração, de uma forma básica, de como funciona um ML (SCHLEDER et al., 2021).



Figura 7. Os quatro componentes básicos do processo de *machine learning*.
Fonte: (SCHLEDER et al., 2021).

Portanto, o aprendizado de máquina pode ser utilizado seguindo os quatro passos fundamentais: definição do problema; dados; representação; e algoritmos, validação e aplicação (SCHLEDER et al., 2021). Assim, ML pode ser usada como complemento, otimização e validação dos dados obtido através da

651 PCA, sendo que os dados desta última são rotineiramente empregados em
652 algoritmos de ML (HEEL et al., 2016).

3. REFERÊNCIAS BIBLIOGRÁFICAS

AAGAARD, C.; GOVAERTS, M.; MEIKLE, V. et al. Optimizing antigen cocktails for detection of *Mycobacterium bovis* in herds with different prevalences of bovine tuberculosis: ESAT6-CFP10 mixture shows optimal sensitivity and specificity. **Journal of Clinical Microbiology**, v.44, n.12, p.4326-4335, 2006.

ABIMERHI, D.; GUTIÉRREZ, A.F.; VILLALOBOS, E.J. Uso de las pruebas de ensayo inmunoabsorbente ligado a enzimas y aglutinación con rivanol para el diagnóstico de brucelosis bovina en Yucatán. **Veterinária México**, v.29, n.2, p.167-171, 1998.

ABRAHÃO, R.M.C.M. **Tuberculose humana causada pelo *Mycobacterium bovis*: considerações gerais e a importância dos reservatórios animais**. 1998. 273p. Dissertação (Mestrado) – Faculdade de Saúde Pública da Universidade de São Paulo, São Paulo, 1998.

ACHA, P.N.; SZYFRES, B. **Zoonosis y enfermedades transmissibles comunes al hombre y a los animales**. Washington: O.P.S. p.28-25. 2001.

ALTON, G.G.; JONES, L.M.; ANGUS, R.D. et al. **Techniques for the brucellosis laboratory**. Paris: Institut Nacional de la Recherche Agronomique, 1988. 190p.

ANDERSON, R.R.; PARRISH, J.A. The optics of human skin. **The Journal of Investigative Dermatology**, v.77,1, n.77(1), p.13-9, 1981.

ARAUJO, O.W.; COELHO, C.J. Análise de componentes principais (PCA). **University Center of Anápolis**, 2009.

ARÉSTEGUI, M.B.; GUALTIERI, S.C.; DOMÍNGUEZ, J. et al. El género *Brucella* y su interacción com el sistema mononuclear fagocítico. **Veterinária México**, v.32, n 2, p.131-139, 2001.

BACON, C.P.; MATTLEY, Y.; DEFRECE, R. Miniature spectroscopic instrumentation: Applications to biology and chemistry. **Review of Scientific instruments**, v.75, n.1-1, p.1-16, 2004.

BALDWIN, C.L.; PARENT, M. Fundamentals of host immune response against *Brucella abortus*: what the mouse model has revealed about control of infection. **Veterinary Microbiology**, v.90, n.1-4, p.367-382, 2002.

BEER, J. **Doenças Infecciosas em Animais Domésticos**. São Paulo: Roca, p.178-183, 1998.

BISHOP, G.C.; BOSMAN, O.; HERR, S. et al. **Infectious diseases of livestock**, Austin, Texas: AeM University Press, p.1053-1066, 1994.

BRANDTZAEG, P. The increasing power of immunohistochemistry and immunocytochemistry. **Journal of Immunological Methods**, v.216, n.1-2 p.49-67. 1998.

- 691 **Brasil. Ministério da Agricultura, Pecuária e do Abastecimento.**
 692 Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose
 693 Animal, 2001.
- 694 **Brasil. Ministério da Agricultura, Pecuária e Abastecimento.**
 695 Secretaria da Defesa Sanitária. Instrução normativa SDA nº 06, de 08 de janeiro
 696 de 2004, aprova o regulamento técnico do Programa nacional de controle e
 697 erradicação da brucelose e tuberculose animal. Diário Oficial da União, Brasília,
 698 DF, Seção 1, p.6-10, 2004.
- 699 **Brasil. Ministério da Agricultura, Pecuária e Abastecimento.**
 700 Programa Nacional de Controle e Erradicação da Brucelose e Tuberculose
 701 Animal, 2006.
- 702 **Brasil. Ministério da Agricultura, Pecuária e Abastecimento.**
 703 Diagnóstico Situacional do PNCEBT: Programa Nacional de Controle e
 704 Erradicação da Brucelose e Tuberculose Animal, 2020.
- 705 BRICKER, B.J. PCR as a diagnostic tool for brucellosis. **Veterinary**
 706 **Microbiology**, v.90, n1-4, p.435-446, 2002.
- 707 BURGHARDT, E. **Colposcopy Cervical Pathology Textbook and Atlas.**
 708 Stuttgart: Georg Thieme Verlag, 1984.
- 709 BUTLER, W.R.; JOST, K.C.; KILBURN, J.O. Identification of mycobacteria
 710 by high-performance liquid chromatography. **Journal of Clinical Microbiology**,
 711 v.29, n.1, p.268-272, 1991.
- 712 CAMPANÃ, R.N.; GOTARDO, D.J.; ISHIZUCA, M.M. **Epidemiologia e**
 713 **Profilaxia da Brucelose Bovina e Bubalina.** Campinas, São Paulo:
 714 Coordenadoria de Defesa Agropecuária CDA/SAA, 2003, 20p.
- 715 CARTER, G.R.; CHENGAPPA, M.M. **Essentials of Veterinary**
 716 **Bacteriology and Mycology.** Philadelphia, London, p.196-200, 1991.
- 717 CARVALHO, M.S.; BARROSO, M.R.; PINHAL, F. et al. Brucelose: Alguns
 718 aspectos epidemiológicos. **Medicina Interna**, v.2, n.4, p.259-261. 1995.
- 719 CARVALHO, R. **Uso da Nested PCR para diagnóstico de tuberculose**
 720 **em carcaças de bovinos.** 2013. Tese (Mestrado) - Faculdade de Agronomia,
 721 Medicina Veterinária e Zootecnia, Universidade Federal de Mato Grosso,
 722 Cuiabá, 2013.
- 723 **Centro panamericano de zoonosis.** Manual de normas y procedimientos
 724 tecnicos para la bacteriologia de la tuberculosis. Buenos Aires: Cepancho, 30p,
 725 1988.
- 726 CHATE, S.C.; DIAS, R.A.; AMAKU, M. et al. Situação epidemiológica da
 727 brucelose bovina no Estado do Mato Grosso do Sul. **Arquivos Brasileiros de**
 728 **Medicina Veterinária e Zootecnia**, v.61, n.1, p.46-55. 2009.

729 COLLINS, D.M.; ADFORD, A.J.; LISLE, G.W. et al. Diagnosis and
730 epidemiology of bovine tuberculosis using molecular biological approaches.
731 **Veterinary Microbiology**, v.40, n.1-2, p.83-94, 1994.

732 CORNER, L.A. Post mortem diagnosis of *Mycobacterium bovis* infection
733 in cattle. **Veterinary Microbiology**, v.40, n.1-2, p.53-63, 1994.

734 COSIVI, O.; GRANGE, J.M.; DABORN, C.J. et al. Zoonotic tuberculosis
735 due to *Mycobacterium bovis* in developing countries. **Emerging Infectious**
736 **Diseases**, v.4, n.1, p.59-70, 1998.

737 COSTA, A.C.; REIS, M.C.G.; SILVA, B.D.S. et al. Resposta imune
738 humoral ao antígeno rGroES do *Mycobacterium tuberculosis* em pacientes com
739 tuberculose e seus contatos domiciliares. **Revista de Patologia Tropical**, v.40,
740 n.1, p.23-34, 2011.

741 DAJER, A.; LUNA-MARTÍNEZ, E.; ZAPATA, D. et al. Evaluation of a
742 fluorescence polarization assay for the diagnosis of bovine brucellosis in México.
743 **Preventive Veterinary Medicine**, v.40, n.1, p.67-73, 1999.

744 FARIAS, T.A.; ARAÚJO, F.R.; OSÓRIO, A.L.A.R. et al. ELISA based on
745 recombinant MPB70 and P27 for detection of antibodies against *Mycobacterium*
746 *bovis*. **Revista de Patologia Tropical**, v.41, n.2, p.155-162, 2012.

747 FEKETE, A.; BANTLE, J.A.; HALLING, S.M. et al. Development of a
748 diagnostic Test for *Brucella* using Polymerase Chain Reaction. **Journal of**
749 **Applied Bacteriology**, v.69, n.2, p.216-227, 1990.

750 FILHO, G.L.P.; ALVES, A.J.S.; CARVALHO, L.G. et al. Occurrence of
751 brucellosis and bovine tuberculosis and risk perception in Mato Grosso do Sul,
752 Brazil. **Arquivos do Instituto Biológico**, v.84, e0472016, 2017.

753 FLOYD, M.M.; SILCOX, V.A.; JONES, W.D. et al. Separation of
754 *Mycobacterium bovis* BCG from *Mycobacterium tuberculosis* and *Mycobacterium*
755 *bovis* by using high-performance liquid chromatography of mycolic acids. **Journal**
756 **of Clinical Microbiology**, v.30, n.5, p.1327-1330, 1992.

757 FOSGATE, G.T.; ADESIYUN, A.A.; HIRD, D.W. et al. Comparison of
758 serologic test for detection of *Brucella* infections in cattle and water buffalo
759 (*Bubalus bubalis*). **American Journal of Veterinary Research**, v.63, n.11,
760 p.1598-1605, 2002.

761 FRÁGUAS, S.; CUNHA-ABREU, M.S.; FERREIRA, A.M. et al. Estudo
762 comparativo de métodos complementares para o diagnóstico da tuberculose
763 bovina em animais reagentes à tuberculinização. **Revista Brasileira de Ciência**
764 **Veterinária**, v. 15, n. 3, p.117-121, 2008.

765 GARCIA, M.; MELO, A.F.; CARVALHO, G.F. et al. Epidemiologia da
766 tuberculose bovina na América do Sul. **Research, Society and Development**,
767 v.10, n.9, p.e8610917936-e8610917936, 2021.

- 768 GODINHO, M.S.; PEREIRA, R.O.; RIBEIRO, K.D.O. et al. Classificação
769 de refrigerantes através de análise de imagens e análise de componentes
770 principais (PCA). **Química Nova**, v.31, n.6, p.1485-1489, 2008.
- 771 GORVEL, J.P.; MORENO, E. *Brucella* intracellular life: from invasion to
772 intracellular replication. **Veterinary Microbiology**, v.90, n.1-4, p.281-297, 2002.
- 773 HAAGSMA, J. **Bovine tuberculosis**. Geneve: Office International des
774 Épizooties, 1995. 11p. (Manual Amendment 2).
- 775 HARMON, B.G.; ADAMS, L.G.; FREY, M. Survival of rough and smooth
776 strains of *Brucella abortus* in bovine mammary gland macrophages. **American**
777 **Journal of Veterinary Research**, v.49, n.7, p.1092– 1097, 1988.
- 778 HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J.H. **The elements of**
779 **statistical learning: data mining, inference, and prediction**. New York:
780 Springer, 2001.
- 781 HEEL, M.V.; PORTUGAL, R.V; SCHATZ, M. et al. Multivariate statistical
782 analysis of large datasets: single particle electron microscopy. **Open Journal of**
783 **Statistics**, v.6, n.04, p.701, 2016.
- 784 HEY, T.; TANSLEY, S.; TOLLE, K. et al. **The Fourth Paradigm: Data-**
785 **Intensive Scientific Discovery**, Redmond, WA: Microsoft research, 2009.
- 786 HONGYU, K.; SANDANIELO, V.L.M.; OLIVEIRA, J. et al. Análise de
787 componentes principais: resumo teórico, aplicação e interpretação. **E&S**
788 **Engineering and Science**, v.5, n.1, p.83-90, 2016.
- 789 Huber, J.D.; Nicoletti, P. Comparison of the results of card, rivanol,
790 complement-fixation, and milk ring tests with the isolation rate of *Brucella abortus*
791 from cattle. **American Journal of Veterinary Research**, v.47, n.7, p.1529-1531,
792 1986.
- 793 JONES, S.L.; COX, J.C.; SHEPHERD, J.M. et al. Removal of false-
794 positive reactions from plasma in an enzyme immunoassay for bovine interferon-
795 gamma. **Journal of Immunology Methods**, v.155, n.2, p.233-240, 1992.
- 796 JOLLIFFE, I.T.; CADIMA, J. Principal component analysis: a review and
797 recent developments. **Philosophical transactions of the royal society A:**
798 **Mathematical, Physical and Engineering Sciences**, v.374, n.2065,
799 p.20150202, 2016.
- 800 JORGE, K. **Identificação de Mycobacterium bovis em bovinos e sua**
801 **importância na ocorrência de tuberculose zoonótica**. 2010. Tese de
802 doutorado - Universidade Federal de Mato Grosso do Sul, Campo Grande, 2010.
- 803 KANTOR, I.N.; RITACCO, V. Bovine tuberculosis in Latin America and
804 Caribbean: current status, control and eradication programs. **Veterinary**
805 **Microbiology**, v.40, n.1-2, p.5-14, 1994.

KLASTER, P.R.; KUIJPER, S.; INGEN, C.W.V. et al. Stabilized, freeze-dried PCR mix for detection of mycobacteria. **Journal of Clinical Microbiology**, v.36, n.6, p.1798-1800, 1998.

KOLLIAS, N.; BAQER, A.; SADIQ, I. Minimum erythema dose determination in individuals of skin type V and VI with diffuse reflectance spectroscopy. **Photodermatology, Photoimmunology & Photomedicine**, v.10, n.6, p.249–254, 1994.

LAGE, A.P.; ROXO, E.; MULLER, E. et al. **Programa nacional de controle e erradicação da brucelose e da tuberculose animal (PNCEBT)**. Brasília: Ministerio da Agricultura, Pecuaria e Abastecimento, 2006. 188p. (Manual técnico).

LAGE, A.P.; POESTER, F.P.; PAIXÃO, T.A. et al. Brucelose bovina: uma atualização. **Revista Brasileira de Reprodução Animal**, v.32, n.3, p.202-212, 2008.

LAWINSKY, M.L.J.; OHARA, P.M.; ELKHOURY, M.R. et al. Estado da arte da brucelose em humanos. **Revista Pan-Amazônica de Saúde**, v.1, n.4, p.10-10, 2010.

LEYLA, G.; KADRI, G.; UMRAN, O.K. Comparison of polymerase chain reaction and bacteriological cultura for the diagnosis of sheep brucellosis using aborted fetus samples. **Veterinary Microbiology**, v.93, n.1, p.53-61, 2003.

LIEBOLD, K.; FASSLER, D.; SCHMIDT, W.D. et al. In vivo spectroscopy in dermatology: methods and new fields of application. **Journal of the European Academy of Dermatology Venereology**, v.14, n.1, p.1-4, 2000.

LIRA, N.S.C. **Lesões anatomopatológicas e detecção da *Brucella ovis* cepa REO em ovinos inoculados experimentalmente pelas vias intraprepucial e conjuntival simultaneamente**. 2008. Tese (Doutorado) - Faculdade de Medicina Veterinária e Zootecnia, Universidade Estadual Paulista, FMVZ/UNESP - Campus de Botucatu/SP, Botucatu, São Paulo, 2008.

LIU, S.; GUO, S.; WANG, C. et al. A novel fusion protein-based indirect enzyme-linked immunosorbent assay for the detection of bovine tuberculosis. **Tuberculosis**, v 87, n.3, p.212-217, 2007.

MARCONDES, A.G. **Padronização da técnica de cultivo em camadadelgada de agar milddlebrook 7 H11 para isolamento de *Mycobacterium bovis***. 2002. 115p. Dissertação (Mestrado) – Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo, São Paulo, 2002.

MARTINHO, J. M. G. Espectroscopia de Absorção no Ultravioleta e Visível. **Boletim da Sociedade Portuguesa de Química**, v.52, n.3, p.44-46, 1994.

MASCARENHAS, D.R.M. **Validação da técnica de PCR em tempo real (qPCR) para detecção de *Mycobacterium bovis* e *Brucella abortus* em**

- 847 **amostras de leite cru.** Dissertação (Mestrado) - Universidade Federal de
848 Viçosa, Viçosa, 2017.
- 849 MATHIAS, L.A.; CORBELLINI, L.G.; MAIA, L. et al. Validação
850 interlaboratorial do teste de polarização fluorescente para o diagnóstico
851 sorológico da brucelose bovina. **Ciência Rural**, v.40, n.10, p.2135-2140, 2010.
- 852 MATRONE, M.; KEID, L.B.; ROCHA, V.C.M. et al. Evaluation of DNA
853 extraction protocols for *Brucella abortus* PCR detection in aborted PCR detection
854 in aborted fetuses or calves born from cows experimentally infected with strain
855 2308. **Brazilian Journal of Microbiology**, v.40, p. 480-489, 2009.
- 856 MATTHIAS, D. Infecções por microbactérias. In: Beer, J. **Doenças**
857 **infecciosas em animais domésticos**. São Paulo: Roca, p.261-289, 1988.
- 858 MEGID, J.; RIBEIRO, M.G.; MARCOS, J.G. et al. Evaluation of rapid
859 agglutination, tube agglutination, buffered plate antigen and 2- mercaptoethanol
860 tests in the diagnosis of bovine brucellosis. **Brazilian Journal of Veterinary**
861 **Research and Animal Science**, v.37, n.4, p.395-399, 2000.
- 862 MEIRELLES-BARTOLI, R.B.; MATHIAS, L.A. Estudo comparativo entre
863 os testes adotados pelo PNCEBT para o diagnóstico sorológico da brucelose em
864 bovinos. **Arquivos do Instituto Biológico**, v.77, n.1, p.11-17, 2010
- 865 MENESES, I.D.S. **Avaliação de proteínas recombinantes para**
866 **detecção de anticorpos contra Mycobacterium bovis por ELISA**. 2015.
- 867 MENZIES, D. Interpretation of repeated tuberculin tests. Boosting,
868 conversion, and reversion. **American Journal of Respiratory and Critical Care**
869 **Medicine**, v.159, n.1, p.15-21, 1999.
- 870 MONAGHAN, M.; DOHERTY, M.L.; COLLINS, D. et al. The tuberculin test.
871 **Veterinary Microbiology**. v.40, n.1-2, p.111-124, 1994.
- 872 MORENO, E.; MORIYÓN, I. et al., The genus *Brucella*. In: Kersters, K.;
873 Paul, M.G.V.; SWINGS, J. et al. (eds). **The Prokaryotes**. New York, NY:Springer,
874 p.315–456, 2006.
- 875 MURPHY, KP. **Machine Learning: A Probabilistic Perspective**. London:
876 MIT Press, 2012. 421p.
- 877 NEILL, S.D.; POLLOCK, I.M.; BRYSON, D.B. et al. Pathogenesis of
878 *Mycobacterium bovis* infection in cattle. **Veterinary Microbiology**, v.40, n.1-2,
879 p.41-52, 1994.
- 880 NETA, A.V.C.; MOL, J.P.; XAVIER, M.N. et al. Pathogenesis of bovine
881 brucellosis. **The Veterinary Journal**, v.184, n.2, p.146–155, 2010.
- 882 NIELSEN, K. Diagnosis of brucellosis by serology. **Veterinary**
883 **Microbiology**, v.90, p.447-459, 2002.

- 884 NOZAKI, C.N. **Aspectos epidemiológicos, clínicos e avaliação de**
 885 **métodos de diagnósticos nas fases da evolução da brucelose em ovinos**
 886 **inoculados experimentalmente com *Brucella ovis***. 2008. 109p. Tese
 887 (Doutorado) - Faculdade de Medicina Veterinária e Zootecnia, Universidade
 888 Estadual Paulista "Júlio Mesquita Filho", Botucatu, 2008.
- 889 **OIE - OFFICE INTERNATIONAL DES EPIZOOTIES**. Manual of
 890 Standards for Diagnostic Tests and Vaccines. 5 ed. Paris: Office International
 891 des Epizooties, p. 328-345, 2008.
- 892 **OIE - Organização Mundial de Saúde Animal**. Terrestrial manual -
 893 access online. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals.
 894 Version adopted by the World Assembly of Delegates of the OIE. Bovine
 895 tuberculosis, v.1, pt 2, chap. 2.4.7. Bovine tuberculosis, 2009.
- 896 **OIE- Organização mundial de saúde animal**. Listed Diseases, Infections
 897 and Infestations in Force in 2016; 2016.
- 898 OLIVEIRA, JP. **Estudo das lesões sugestivas de brucelose em**
 899 **bovinos e bubalinos abatidos para consumo**. 2003. 53p. Dissertação
 900 (Mestrado) - Universidade Federal do Pará, Pará, 2003.
- 901 **Organización panamericana de la salud**. Situación de la tuberculosis
 902 bovina en las Américas. Washington, D.C., 1995. 6p. (RIMS 9, Informativo 26).
- 903 OWEN, T. **Fundamentals of UV-Visible Spectroscopy: A Primer**.
 904 Agilent Technologies, Germany, 2000.
- 905 PAJUABA, A.C.A.M. **Avaliação de frações hidrofóbicas e hidrofílicas**
 906 **de *Brucella abortus* em ensaios imunoenzimáticos para caracterizar o perfil**
 907 **de anticorpos produzidos por bovinos vacinados e não-vacinados**. 2006.
 908 64 p. Dissertação (Mestrado) - Instituto de Ciências Biomédicas, Universidade
 909 Federal de Uberlândia, Uberlândia, 2006.
- 910 PAULIN, L.M.; NETO, J.S.F. **O combate à brucelose bovina: situação**
 911 **brasileira**. Jaboticabal: Funep, 2003. 154p.
- 912 PAULIN, L.M.; NETO, J.S.F. Brucelose em búfalos. **Arquivos do**
 913 **Instituto Biológico**, v.75, n.3, p.389-401, 2008.
- 914 PERKAMPUS, H.H. **Introduction**. In: **UV-VIS Spectroscopy and Its**
 915 **Applications**. Berlin: Springer Science & Business media, 1992, 29p.
- 916 PESSEGUEIRO, P.; BARATA, C.; CORREIA, J. Brucelose – uma revisão
 917 sistematizada. **Medicina Interna**, v.10, n.2, p.91-100. 2005.
- 918 POESTER, F.P.; SAMARTINO, L.E.; LAGE, A.P. Diagnóstico da
 919 brucelose bovina. **Caderno Técnico de Veterinária e Zootecnia**, n.47, p.13-29,
 920 2005.
- 921 POESTER, F.; FIGUEIREDO, V.C.F.D.; LÔBO, J.R. et al. Estudos de
 922 prevalência da brucelose bovina no âmbito do Programa Nacional de Controle e

- 923 Erradicação de Brucelose e Tuberculose: Introdução. **Arquivo Brasileiro de**
924 **Medicina Veterinária e Zootecnia**, v.61, p.01-05, 2009.
- 925 PRITCHARD, D.G.A. A century of bovine tuberculosis 1888 - 1988,
926 conquest and controversy. **Journal of Comparative Pathology**, v.99, n.4,
927 p.357-399, 1988.
- 928 QUEIPO-ORTUNO, M.I.; MORATA, P.; OCÓN, P. et al. Rapid diagnosis
929 of human brucellosis by peripheral-blood pcr assay. **Journal of Clinical**
930 **Microbiology**, v.35, n.11, p.2927-2930. 1997.
- 931 RIBEIRO, M.G.; MOTTA, R.G.; ALMEIDA, C.A.S. Brucelose equina:
932 aspectos da 646 doença no Brasil. **Revista Brasileira de Reprodução Animal**,
933 v.32, 647 n. 2, p.83-92, 2008.
- 934 RORING, S.; HUGHES, M.S.; SKUCE, R.A. et al. Simultaneous detection
935 and strain differentiation of *Mycobacterium bovis* directly from bovine tissue
936 specimens by spoligotyping. **Veterinary Microbiology**, v.74, p.227-236, 2000.
- 937 RUGGIERO, A.P.; IKUNO, A.A.; FERREIRA, V.C.A. et al. Tuberculose
938 bovina: alternativas para o diagnóstico. **Arquivos do Instituto Biológico**, v.74,
939 p.55-65, 2022.
- 940 SAKAMOTO, S.M.; HEINEMANN, M.B.; TELLES, M.A.S. et al. Detecção
941 e identificação de *Mycobacterium bovis* pela reação em cadeia da polimerase
942 (PCR). **Arquivos do Instituto Biológico**, v.66, n.2, p.44-58, 1999.
- 943 SALAZAR, F.H.P. **Ocorrência de tuberculose causada por**
944 ***Mycobacterium bovis* em bovinos abatidos em frigoríficos no Estado de**
945 **Mato Grosso**. 2005. 68p. Dissertação (Mestrado) - Universidade Federal de
946 Mato Grosso do Sul, Faculdade de Medicina Veterinária e Zootecnia, Campo
947 Grande, 2005.
- 948 SALES, M.L. **Identificação de *Mycobacterium bovis* e *Mycobacterium***
949 **tuberculosis por PCR**. 2012. 62p. Tese (Mestrado) - Escola de Veterinária da
950 Universidade Federal de Minas Gerais, Belo Horizonte, 2012.
- 951 SAMUEL, A.L. Some studies in machine learning using the game of
952 checkers. **IBM Journal of Research and Development**, v.3, n.3, p.210-229,
953 1959.
- 954 SANTOS, RL.; MARTINS, T.M.; BORGES, Á.M. et al. Perdas econômicas
955 devidas à brucelose bovina no Brasil. **Pesquisa Veterinária Brasileira**, v.33, n.
956 3, p.759-764, 2013.
- 957 SCHLEDER, G.R.; FAZZIO, A. Machine Learning na Física, Química, e
958 Ciência de Materiais: Descoberta e Design de Materiais. **Revista Brasileira de**
959 **Ensino de Física**, v.43, 2021.
- 960 SEÓNADH, O.L.; MICHAEL, S.; TORRES, S. *Brucella abortus* detection
961 by PCR assay in blood, milk and lymph tissue of serologically positive cows.
962 **Research in Veterinary Science**, v.81, n.2, p.170-176, 2006.

- 963 SILVA, C.S. **Espectroscopia no Infravermelho para Aplicações**
 964 **Forenses: documentoscopia e identificação de sêmen em tecidos**. 2007.
 965 Tese (doutorado) - Departamento de Química Fundamental da Universidade
 966 Federal de Pernambuco, Recife, 2007.
- 967 SILVA, F.L.; PAIXÃO, T.A.; BORGES, A.M. et al. Brucelose Bovina.
 968 **Caderno Técnico de Veterinária e Zootecnia**, v.47, p.1-12, 2005.
- 969 SILVA, M.C.; MOURA, M.S.; REIS, D.O. **Tuberculose – Revisão de**
 970 **literatura**. Londrina: PUBVET, v.5, n.17, 2011.
- 971 SKOOG, D.; WEST, D.; HOLLER, F.J. **Fundamentos de Química**
 972 **Analítica**. Mexico: Thompson, 2005.
- 973 SOUZA, I.I.; MELO, E.S.; RAMOS, C.A.N. et al. Screening of recombinant
 974 proteins as antigens in indirect ELISA for diagnosis of bovine
 975 tuberculosis. **Springerplus**, v.1, n.1, p.1-6, 2012.
- 976 STOCKERT, J.C.; BLÁZQUEZ-CASTRO, A.; CAÑETE, M. et al. MTT
 977 assay for cell viability: Intracellular localization of the formazan product is in lipid
 978 droplets. **Acta Histochemica**, v.114, n.8, p.785-796, 2012.
- 979 TANYA, V.; NICOLETTI, P.V. Comparison of enzymelabeled
 980 immunosorbent assay and particle concentration fluorescence immunoassay
 981 with standard serologic methods and bacteriologic culture for detection of
 982 *Brucella* sp.-infected cows in herds with brucellosis. **Journal of the American**
 983 **Veterinary Medical Association**, v.202, n.12, p.1975-1977, 1993.
- 984 THOMAZ, L.W. Apostila: **Conhecimentos Específicos para Médico**
 985 **Veterinário**. Editora Vesticon, p.48-82, 2006.
- 986 VITALE, F.; CAPRA, G.; AXIA, L. Detection of *Mycobacterium tuberculosis*
 987 complex in cattle by PCR using milk, lymph node aspirates and nasal swabs.
 988 **Journal of Clinical Microbiology**, v.36, n.4, p.1050-1055, 1998.
- 989 WALKER, R.L. *Brucella*. In: HIRSH, D.C.; ZEE, Y.C. **Microbiologia**
 990 **Veterinária**. Rio de Janeiro: Editora Guanabara Koogan S. A., p.185-191, 2003.
- 991 WARDS, B.J.; COLLINS, D.M.; LISLE, G.W. Detection of *Mycobacterium*
 992 *bovis* in tissues by polymerase chain reaction. **Veterinary Microbiology**, v.43,
 993 n.2-3, p.227-240, 1995.
- 994 WATERS, W.R.; BUDDLE, B.M.; VORDERMEIER, H.M. et al.
 995 Development and evaluation of an enzyme-linked immunosorbent assay for use
 996 in the detection of bovine tuberculosis in cattle. **Clinical and Vaccine**
 997 **Immunology**, v.18, n.11, p.1882-1888. 2011.
- 998 WIT, D.; TEYN, L.; SHOEMAKER, S. et al. Direct detection of
 999 *Mycobacterium tuberculosis* in clinical specimens by DNA amplification. **Journal**
 1000 **of Clinical Microbiology**, v.28, n.11, p.2437-2441, 1990.

- 1001 WOOD, P.R.; ORNER, L.A.; ROTHEL, J.S. et al. Field comparison of the
1002 interferon gamma assay and the intradermal tuberculin test for the diagnosis of
1003 bovine tuberculosis. **Australian Veterinary Journal**, v.68, n.9, p.286-290, 1991.
- 1004 **World health organization.** Zoonotic tuberculosis (*Mycobacterium bovis*)
1005 memorandum from a WHO meeting with the participation of FAO. Bulletin of
1006 World Health Organization, v.72, p.8510-8857, 1994.
- 1007 XAVIER, M.N. **Desenvolvimento de PCR espécie-específico para o**
1008 **diagnóstico da infecção por Brucella ovis e avaliação comparativa de**
1009 **métodos sorológicos.** 2009. 68p. Dissertação (Mestrado) - Escola de
1010 Veterinária, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.
- 1011 ZANDEN, A.G.M. **Spoligotyping, a tool in epidemiology, diagnosis**
1012 **and control of tuberculosis.** 2002. p.12-45. Tese (Doutorado) - Katholieke
1013 Universiteit Nijmegen, Nijmegen, 2002.
- 1014 ZANINI, M.S.; MOREIRA, E.C.; LOPES, M.T.P. et al. *Mycobacterium*
1015 *bovis*: polymerase chain reaction identification in bovine lymph node biopsies and
1016 genotyping in isolates from southeast Brazil by spoligotyping and restriction
1017 fragment length polymorphism. **Memórias do Instituto Oswaldo Cruz**, v.96, n.6,
1018 p.2809-2813, 2001.

4. CAPÍTULO I

Turning chaotic sample group clusterization into organized ones by feature selection: Application on photodiagnosis of *Brucella abortus* serological test

Bruno Silva de Rezende^a, Thiago França^b, Herbert Patric Kellermann Cleveland^a, Cícero Cena^{*b} and Carlos Alberto do Nascimento Ramos^a

Bovine brucellosis diagnosis is a major problem to be solved, the disease has a great economic impact with significant losses in meat and dairy products, besides the fact that it can be transmitted to humans. The sanitary measures instituted in Brazil are based on disease control through diagnosis, animal sacrifice, and vaccination. Although the currently available diagnostic tests show good quality parameters, they are time-consuming, and the incidence of false-positive and/or false-negative results is still observed, which difficult the effectively control the disease. The development of a low-cost, fast, and accurate brucellosis diagnosis test remains a need for proper sanitary measures at a large-scale analysis. In this context, spectroscopy techniques associated with machine learning tools have shown great potential for use in diagnostic tests. In this study, bovine blood serum was investigated by UV-vis spectroscopy and machine learning algorithms to build a prediction model for *Brucella abortus* diagnosis. Here we first apply principal component analysis to observe the group formation tendency, the first results showed no clustering tendency with a messy sample score distribution, then we properly select the main principal components to improve clusterization. Finally, by using machine learning algorithms (SVM and KNN), the predicting models achieved a 92.5% overall accuracy. The present methodology provides a test result in an average time of 5 minutes, while the standard diagnosis, with the screening and confirmatory tests, can take up to 48 hours. The present result demonstrates the viability of the method for diagnosing bovine brucellosis, which can significantly contribute to disease control programs in Brazil and other countries.

Keywords: photodiagnosis; diagnostic test; UV-vis spectroscopy; machine learning; bovine brucellosis

Introduction

Bovine brucellosis is a chronic disease that has a great economic impact with significant losses in meat and dairy products, besides the fact that it can be transmitted to humans¹. The prevalence of bovine brucellosis outbreaks in Brazil is variable according to the region ranging from 0.32% to 41.5%². The economic impact of brucellosis in cattle herds occurs mainly due to abortions, underdeveloped calves, lower reproductive index, and lower production of milk and dairy products³. Bovine brucellosis in Brazil generated important economic losses. In some regions, the annual economic losses per herd can reach US\$ 5.890,16⁴.

Brucella abortus, the causative agent of brucellosis in cattle, is a gram-negative, non-spore-forming, non-motile, facultative intracellular, aerobic, and flagellated bacterium⁵. This microorganism is resistant to inactivation in the environment being able to remain active for several months with a high transmission potential⁶. The infection can cause various inflammatory reactions that can result in splenomegaly, hepatomegaly, and even lymphoid hyperplasia, in addition, clinical signs in the locomotor system can affect associated muscles, bursae, tendons, and joints⁷.

The bacterial infection settles preferentially in the pregnant uterus, breast tissue, bones, organs, and male reproductive system. In general, infected pregnant females may present placental deficiency, endometritis, abortions, or even give birth to underdeveloped calves^{8,9}. Humans can also be infected by contact with contaminated animal products or animals, showing clinical symptoms of joint pain, fever, cough, nausea, headache, reactive lymph nodes, myalgia, depression, and weight loss. In addition, suppurations may occur in organs such as the liver and spleen^{10,11}.

In 2001, sanitary measures were instituted in Brazil by the National Program for the Control and Eradication of Animal Brucellosis and

Tuberculosis (PNCEBT), seeking to reduce the prevalence and incidence of brucellosis and tuberculosis. The PNCEBT is based on disease control through diagnosis, animal sacrifice, and vaccination for brucellosis. The diagnosis of *B. abortus* consists of clinical, epidemiological, and mainly laboratory tests, the latter being divided into direct and indirect. The PNCEBT defines as official laboratory tests (i) screening tests: Buffered Acidified Antigen (BAAT), Milk Ring Test (MRT); and (ii) confirmatory tests: 2-Mercaptoethanol (2-ME), Seroagglutination in Tubes (SAT), Complement Fixation Test (CFT) and Fluorescent Polarization Test (FPA)¹².

The BAAT test is based on the agglutination of the antigen with specific IgG antibodies, although it is considered a low-specificity test once the same reaction can also occur due to the vaccine (B19) antibodies, and infected animals by other bacteria genera. While the MRT test consists of a high-sensitivity test, in which a positive result is obtained when a bluish ring is formed if an antigen-antibody combination occurs. However, the MRT can present false positive results due to mastitis, the presence of acidic milk, and colostrum¹². Although BAAT and 2-ME associated with SAT tests are routinely used in Brazil as screening and confirmatory tests, a small number of false-negative and false-positive results are still observed¹³. If an animal was considered negative in the screening or confirmatory test, it will remain in the herds as a source of infection for other animals, making it difficult to effectively control the disease.

Then, the development of fast and accurate methods for practical diagnoses arises to improve disease control at a low-cost and easy implementation. In recent years, optical spectroscopy has been evaluated in the diagnosis of different diseases^{14,15,16}. Although smaller values for sensitivity and specificity have been observed compared to standard tests, the easy implementation, low cost, fast response, lack of multiple processing steps, or the use of specific reagents call attention. In recent studies, by using canine blood serum *Leishmania infantum* infected dogs were identified with 85% accuracy by using Fourier-transformed infrared spectroscopy (FTIR), and 75% accuracy by using UV-vis spectroscopy¹⁷.

Besides the different data analyses, the main difference in the final accuracy result may be related to each technique, FTIR provides information related to molecular vibrational modes, while UV-vis provides information related to the electronic transition in molecular orbitals.

FTIR can be able to access more important information from a molecular point of view, but UV-Vis spectroscopy is much more common in laboratories, easy to handle, economic, and technically more viable. In this study, bovine blood serum was investigated by UV-Vis spectroscopy and machine learning algorithms to build a prediction model for *Brucella abortus* diagnosis. Here the inclusion of an antigen source in the serum may promote an antigen-antibody interaction, in positive samples, causing changes in the sample molecular electronic orbitals. Whereas in negative samples, antigen and antibody molecules would remain separate and no significative changes may occur. These differences may allow more efficient identification of the spectra with the aid of machine learning algorithms.

Materials and Methods

Sample description, preparation, and UV-Vis spectroscopy

A total of 106 bovine blood serum samples were kindly provided by the National Agricultural Laboratory of Pedro Leopoldo, Minas Gerais (LANAGRO – MG), Brazil. The samples were divided into negative (53 samples) and positive (53 samples) for antibodies against *B. abortus* groups according to reference tests recommended by Brazilian legislation (BAAT, 2-ME, and SAT). Also, commercial antigens composed of an inactivated suspension of *B. abortus* 1119/3 diluted to 4.5% in 0.85% saline containing 0.5% of phenol and manufactured by Instituto Biológico (São Paulo, Brazil) were used. The antigen was kept under refrigeration (4°C) before use, according to the manufacturer's instructions.

First, an exploratory test was performed by using 6 positives and 6 negative serum samples to determine the better antigen/serum proportion

rate for group discrimination in our analysis. Different antigen-serum dilutions (1:1 to 1:16) were evaluated, and the best ratio (1:1) was found and used to investigate the other samples. Then, the UV sample spectrum was obtained from 2 μ L of serum antigen mixture, in a NanoDrop One C (Thermo Fisher Scientific) spectrophotometer, in the 200 to 300 nm range, after 3 minutes of incubation at room temperature.

Data analysis and sample classification

The data analysis was implemented and performed in the Python programming language (version 3.9.12), using the Scikit-learn package (version 1.1.2) ¹⁸. First, each UV-Vis spectra were subjected to pre-processing using the Standard Normal Variate (SNV). This method allows the removal of the variation from the baseline and rescales the absorbances, preventing random experimental variations from interfering with the result ¹⁹. Then, UV-SNV spectra, in the 200 nm to 300 nm range, was submitted to principal component analysis (PCA) ²⁰, an unsupervised method that converts multiple variables into a few principal components through dimension reduction, retaining most of the information from the original variables and express the data variation in a score plot. It also, allows us to analyze the main data variance in the analyzed spectral range through the loading plot. PCA is an important step to visualize the group classification tendency, and the output data have shown better performance to build a prediction model based on machine learning algorithms than raw spectral data ^{17,21}.

Finally, a sample classification test was performed by using PCA data submitted to machine learning (ML) algorithms. We use ML algorithms based on three different methods: (i) Discriminant Analysis (DA), which classifies the sample based on the distance between the sample data and the contour built by using a linear or quadratic function to separate the classes (group) ²²; (ii) k-Nearest Neighbor (KNN) uses the Euclidean distance between k (k= 1, 5, 10, 100) closest neighbors to classify the sample ²³; and (iii) Support Vector Machine (SVM), which organizes each

sample class through the optimization of a hyperplane – the hyperplane can be linear or nonlinear, being optimized to reach high performance – between the classes ²⁴.

Each ML algorithm can use a determined number of PCs (input data) to build the predicting model, but to avoid overfitting and underfitting is mandatory to determine the ideal number of Principal Components (PCs). Here we used the Feature Selection Recursive Feature Elimination (RFE), which may select the main PCs that most contribute to achieve high accuracy and remove other PCs with the weakest contribution. The use or removal of a determined PC was made based on the accuracy achieved by using Linear Discriminant Analysis (LDA) to classify the samples in a Leave One Out Cross-validation (LOOCV) test.

In LOOCV, one sample is taken from the data set, and the others are used to build the prediction model (training). Then, the sample data draw is used to assess the quality of the model (test). The procedure is repeated until all sample data have been tested. Model quality is measured by accuracy, and the average percentage of correct answers remains in each test ²⁵.

Then, a ranking of PCs is established and the PCs with the weakest contribution for accuracy are gradually removed from the data set until the highest accuracy is obtained with the lowest number of PCs ²⁶. Finally, once the main PCs were determined and selected by RFE, a new LOOCV test was performed by using DA, KNN, and SVM algorithms.

Results and Discussion

The average antigen/serum UV-SNV spectra for bovine brucellosis positive and negative groups, Figure 8, exhibit two wide electronic bands centered around 206 and 220 nm, followed by a small shoulder around 230 nm, and a wide and weak band centered around 273 nm, probably assigned to $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions in carbonyl groups from proteins and amino acids [ref]. The data standard deviation is shown as a shadow around the main average spectra, the first three absorption bands carry the

largest contribution to the standard deviation, despite this, the data obtained showed a small and usual deviation between them. No difference between both UV-SNV spectra could be identified by the naked eye, then we use a multivariate analysis approach to differentiate both groups.

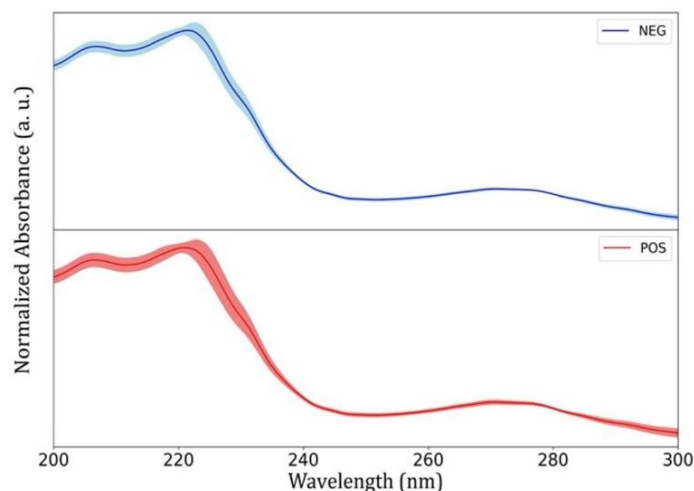


Figure 8: Average antigen/serum UV-SNV spectra for bovine brucellosis positive (POS - red color) and negative (NEG – blue color) groups. The standard deviation is evidenced by the colored shadow around the mainline. Main electronic bands centered around 206, 220, and 273 nm, assigned to $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions in carbonyl groups from proteins and amino acids.

Figure 9 shows the principal component analysis results for the UV-SNV spectral data from bovine brucellosis positive and negative groups. Figure 9^a exhibits the PC1 versus PC2 score plot, which was responsible for 92.1% of data variance with no cluster formation or group separation evidenced – exhibiting a messy sample score distribution. The loading plot, Figure 9^b, suggests that the main contribution for data variance in PC1 and PC2 is assigned to the 220-230 nm range followed by the 235-280 nm range. To improve group separation and clustering formation with success, usually we divide the data set into different ranges ^{27,28}, searching for spectral information that most contributes to group separation. But here the antigen/serum UV bands are too wide, and the data variance is present in the entire range for PC1 and PC2, so as an alternative, we have to find the

best PCs – it means, the best data projection – that most contribute to data separation.

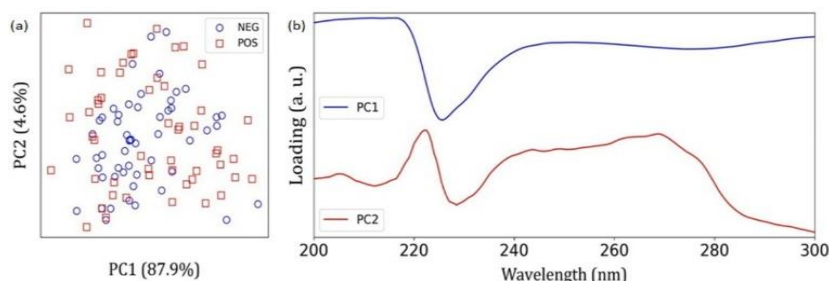


Figure 9: Principal Component Analysis for UV-SNV spectra: (a) score plot for PC1 versus PC2 from bovine blood serum UV-SNV spectra, [red square] bovine brucellosis positive group, and [blue circle] negative group. (b) Loadings for PC1 and PC2 in the 200 to 300 nm range.

First, the Feature Selection Recursive Feature Elimination (RFE) was used to find the main PCs that most contribute to achieving high accuracy. To perform such selection the overall accuracy for sample classification achieved by one specific PC was determined by using Linear Discriminant Analysis (LDA) in a Leave One Out Cross-validation (LOOCV) test. The main PCs found for our data were PC3, PC4, and PC14. Figure 10 shows the violin plot and loading plot for PC3, PC4, and PC14, respectively. The violin plot exhibits the score data distribution project over the respective PC for the bovine brucellosis positive and negative groups. A monomodal distribution was found for both PCs, and the median value (around the peak center) assumes a better distinct position projection over the axis for these PCs compared to the others.

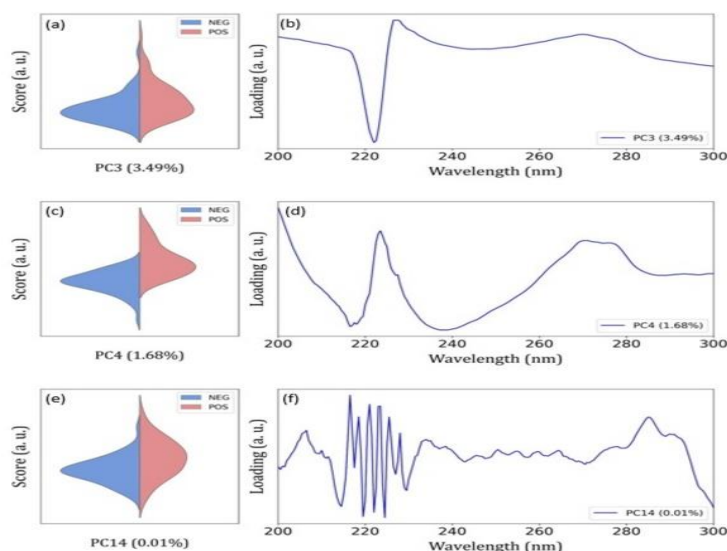


Figure 10: Violin plot for three main PCs normalized scores selected by RFE in 200 to 300 nm from bovine brucellosis positive group (red color), and negative group (blue color). A monomodal distribution is shown with the main peak center dislocated for both groups when projected over the y-axis direction. The data variance percentage of each PC is described below each plot. Respective loading plots for PC3 (b), PC4 (d), and PC14 (f) are also shown.

The PC3, PC4, and PC14 were responsible for only 5.18% of data variance. The PC14 was responsible for only 0.01%, but the median for score distribution over the vertical axis for both groups drive us to a different position, which contributes to sample classification in the LDA algorithm. Besides that, the loading plot for each PC, Figure 10 (right column), shows a high contribution for data variance in the 220 to 230 nm range in accordance with the main bands identified in the UV-SNV spectra (Figure 8). Then, low data variance PCs can be also useful for sample classification, not only PC with high accuracy must influence the algorithm performance. These aspects were confirmed by the score plot for PC3 x PC4 x PC14, Figure 11, which exhibited a clear tendency for clustering and group separation.

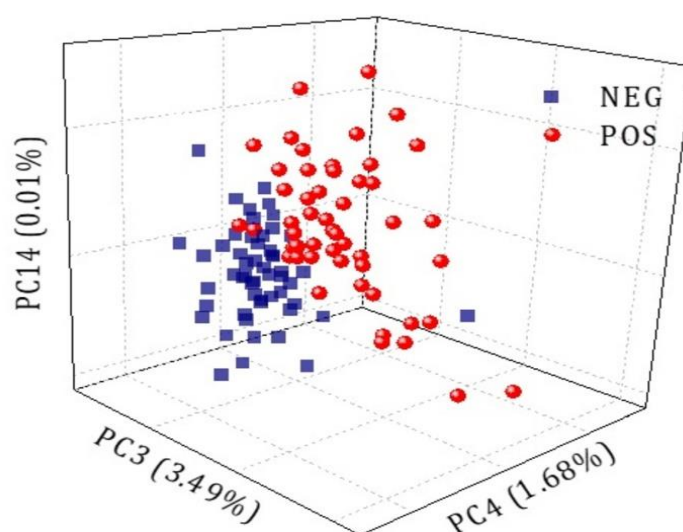


Figure 11: Principal Component Analysis for UV-SNV spectra after RFE analysis. 3D Score plot for PC3 versus PC4, and PC14 from bovine blood serum UV-SNV spectra in the 200 to 300 nm range, [red circle] bovine brucellosis positive group, and [blue square] negative group.

Finally, the PC3, PC4, and PC14 were submitted to machine learning algorithms (DA, SVM, and KNN) to build a prediction model for sample classification. Figure 12^a shows the maximum overall accuracy achieved by the algorithms in a LOOCV test by using different functions for classification. The maximum overall accuracy achieved by the predicting models was 92.5% for Linear SVM and Weighted KNN.

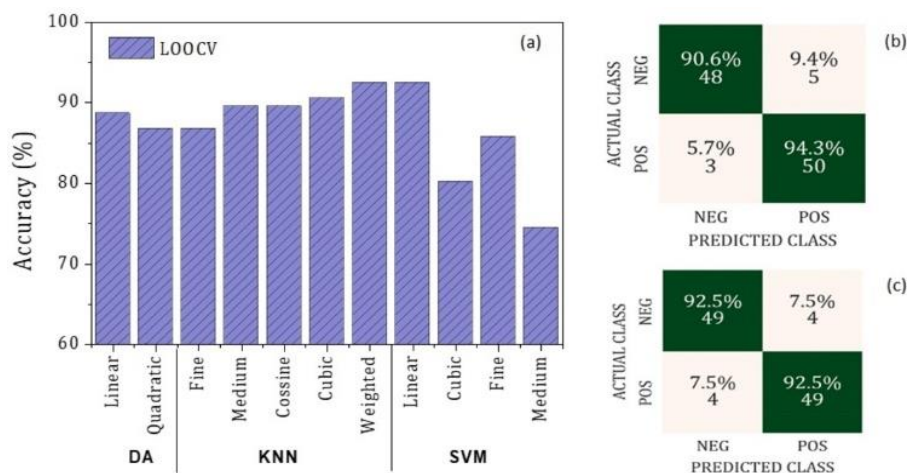


Figure 12: (a) Overall accuracy obtained in the LOOCV tests (light blue bars with right inclined line pattern) for Discriminant analysis (DA), K-nearest neighbor (KNN), and Support vector machine (SVM) algorithms by using PC3, PC4, and PC14 data in the 200 to 300 nm range. Confusion matrix for the performance of (b) SVM algorithm with a linear function, and (c) KNN algorithm with weighted function. The highest overall accuracy was 92.5%.

The confusion matrix, Figure 12^{b,c} shows the sample classification for Linear SVM and Weighted KNN. From the sample classification values was possible to determine for Linear SVM a 94.3% sensitivity, and 90.6% specificity, Figure 12^b, against 94.2% sensitivity, and specificity for Weighted KNN, Figure 12^c. To compare, the relative sensitivity and specificity parameters for the main tests recommended in the PNCBET were respectively 99.6% and 83.9% (BAAT), 98.8% and 96.2% (2-ME / SAT), 91.1 % and 100% (CFT) ¹³. Other confirmatory diagnostic tests for brucellosis, recently introduced in the PNCBET, such as the FPA ¹², have shown sensitivity and specificity around 94% and 95%, respectively ²⁹.

The present study (UV-Vis + ML) was able to identify positive and negative samples for antibodies against *B. abortus* in cattle with sensitivity and specificity similar to that observed in tests currently used. Our best prediction model build with the KNN algorithm and weighted function was able to correctly classify 49 samples from positive and negative groups, only 4 samples from each group were misclassified according to the standard serological procedures recommended by

PNCBET/Brazil ¹². These results are very promising for laboratory implementation as standard screening tests, take around only 5 minutes to provide the result, while the current standard diagnosis, with the screening and confirmatory tests, can take up to 48 hours. In addition, the need for sample processing, in the present methodology was greatly reduced, which enables a reduction in laboratory costs and an increase in sample processing capacity, in addition to the possibility of automation.

Conclusions

The methodology proposed in the present study (UV-Vis + ML) was able to identify positive and negative samples for antibodies against *B. abortus* in cattle with 92.5% overall accuracy, with sensitivity and specificity above 90%. The initial data analysis usually applied in photodiagnosis studies was not able to provide a promising clustering formation for future sample classification. The use of the Feature Selection Recursive Feature Elimination (RFE) algorithm, suggested only 3 main PCs as greatly responsible for group formation, eliminating the usual PCs with high data variance (PC1 and PC2). The score plot of RFE-PCs (PC3xPC4xPC14) demonstrated a clear tendency for clustering and group separation/classification, greatly improving the final sample classification. Finally, our method can provide a test result in an average time of 5 minutes, while the standard diagnosis, with screening and confirmatory tests, can take up to 48 hours. In addition, the sample processing was greatly reduced, which enables a reduction in laboratory costs and an increase in sample processing capacity with great quality control.

Acknowledgments:

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), code 001. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), code 403651/2020-5; 302525/2022-0; 440214/2021-1. Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado

de Mato Grosso do Sul (FUNDECT), code 007/2019; 360/2022.
Universidade Federal de Mato Grosso do Sul – UFMS/MEC – Brasil.

References

- [1] LAGE, A. P., ROXO, E., MÜLLER, E., POESTER, F., CAVALLÉRO, J. C. M., FERREIRA NETO, J. S., ... & GONÇALVES, V. S. P. Programa nacional de controle e erradicação da brucelose e da tuberculose animal (PNCEBT). Brasília: Ministério da Agricultura, Pecuária e Abastecimento, 2006.
- [2] CHATE, S. C.; DIAS, R. A.; AMAKU, M.; FERREIRA, F.; MORAES, G. M.; COSTA NETO, A. A.; MONTEIRO, L. A. R. C.; LÔBO, J. R.; FIGUEIREDO, V. C. F.; GONÇALVES, V. S. P.; FERREIRA NETO, J. Situação epidemiológica da brucelose bovina no Estado do Mato Grosso do Sul. Arquivos Brasileiros de Medicina Veterinária e Zootecnia, v. 61, supl. 1, p. 46-55. 2009.
- [3] POESTER, F.; FIGUEIREDO, V. C. F. D.; LÔBO, J. R.; GONÇALVES, V. S. P.; LAGE, A. P.; ROXO, E.; ... & FERREIRA NETO, J. S. Estudos de prevalência da brucelose bovina no âmbito do Programa Nacional de Controle e Erradicação de Brucelose e Tuberculose: Introdução. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, v. 61, p. 01-05, 2009.
- [4] POSSA, M. G; BERNARDI, F.; NETO, A. P.; CATTELAM, J.; JÚNIOR, I. A. N.; TRENKEL, C. K. G.; MOTA, M. F. Epidemiology and economic impact of brucellosis in the municipality of Santa Catarina State, Brazil. RSD [Internet]. 2021Mar.8 [cited 2023Feb.20];10(3):e12610313208. Available from: <https://rsdjournal.org/index.php/rsd/article/view/13208>
- [5] LAGE, A. P.; POESTER, F. P.; PAIXÃO, T. A.; SILVA, T. M. A.; XAVIER, M. N.; MINHARRO, S.; MIRANDA, K. L.; ALVES, C. M.; MOL, J. P. S.; SANTOS, R. L. Brucelose bovina: uma atualização. Revista Brasileira de Reprodução animal, Belo Horizonte, v. 32, p. 202-212, 2008. Disponível em:<<http://www.cbpa.org.br/pages/publicacoes/rbra/download/RB206%20Lage%20vr2%20pag202-212.pdf>>.

- [6] OIE. Organização Mundial de Saúde Animal. Terrestrial manual - access online. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. Version adopted by the World Assembly of Delegates of the OIE in May 2009, vol 1, pt 2, chap. 2.4.7 Bovine tuberculosis. Disponível em: http://web.oie.int/eng/normes/MANUAL/2008/pdf/2.04.07_BOVINE_TB.pdf.
- [7] MASCARENHAS, D. R. Validação da técnica de PCR em tempo real (qPCR) para detecção de *Mycobacterium bovis* e *Brucella abortus* em amostras de leite cru. Dissertação (Mestrado em Medicina Veterinária) - Universidade Federal de Viçosa, Viçosa. 2017.
- [8] RIBEIRO, M. G.; MOTTA, R. G.; ALMEIDA, C. A. S. Brucelose equina: aspectos da doença no Brasil. Revista Brasileira de Reprodução Animal. Belo Horizonte, v.32, n. 2, p.83-92, abr./jun. 2008. Disponível em: <http://www.cbra.org.br/pages/publicacoes/rbra/download/RB155%20Ribeiro%20pag83-92.pdf>.
- [9] XAVIER, M. N. Desenvolvimento de PCR espécie-específico para o diagnóstico da infecção por *Brucella ovis* e avaliação comparativa de métodos sorológicos. 68f. Dissertação (Mestrado em Ciência Animal) - Escola de Veterinária, Universidade Federal de Minas Gerais. Belo Horizonte. 2009.
- [10] PAULIN, L. M.; NETO, J. S. F. O Combate à Brucelose Bovina: situação brasileira. Jaboticabal. Funep, 2003. 154p. Disponível em: http://www.biologico.sp.gov.br/docs/arq/V69_2/paulin.pdf.
- [11] LAWINSKY, M. L. J. et al. Estado da arte da brucelose em humanos. Revista Pan-Amazônica de Saúde, v. 1, n. 4, p. 10-10, 2010.
- [12] BRASIL. Ministério da Agricultura, Pecuária e Abastecimento (2017). Programa Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal (PNCEBT). Brasília: IN MAPA/SDA/DSA, 10/2017.
- [13] MEIRELLES-BARTOLI, R. B.; MATHIAS, L. A. Estudo comparativo entre os testes adotados pelo PNCEBT para o diagnóstico sorológico da brucelose em bovinos. Arquivos do Instituto Biológico, v. 77, n. 1, p. 11-17, 2010
- [14] SAKUDO, A.; SUGANUMA, Y.; SAKIMA, R.; IKUTA, K. Diagnosis of HIV-1 infection by near-infrared spectroscopy: Analysis using molecular clones

- of various HIV-1 subtypes, *Clinica Chimica Acta*, v. 413, n. 3–4, p. 467-472, 2012.
- [15] AUNER, G. W.; KOYA, S. K.; HUANG, C.; BROADBENT, B.; TREXLER, M.; AUNER, Z.; BRUSATORI, M. A. Applications of Raman spectroscopy in cancer diagnosis. *Cancer Metastasis Rev* 37, 691–717 (2018).
- [16] BRITO, E.C.A; FRANÇA, T.; CANASSA, T.; WEBER, S.S.; PANIAGO, A.M.M.; CENA, C. Paracoccidioidomycosis screening diagnosis by FTIR spectroscopy and multivariate analysis, *Photodiagnosis and Photodynamic Therapy*, v.39, 102921, 2022.
- [17] LARIOS, G.; RIBEIRO, M.; ARRUDA, C.; OLIVEIRA, S. L.; CANASSA, T.; BAKER, M. J.; CENA, C. A new strategy for canine visceral leishmaniasis diagnosis based on FTIR spectroscopy and machine learning. *Journal of Biophotonics*, v. 14, n. 11, p. e202100141, 2021.
- [18] PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, v. 12, p. 2825-2830, 2011.
- [19] RINNAN, Å.; BERG, F. V. D; ENGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, v. 28, n. 10, p. 1201-1222, 2009.
- [20] JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 374, n. 2065, p. 20150202, 2016.
- [21] LARIOS, G.; NICOLODELLI, G.; RIBEIRO, M.; CANASSA, T.; REIS, A. R.; OLIVEIRA, S.L.; CENA, C. Soybean seed vigor discrimination by using infrared spectroscopy and machine learning algorithms. *Anal Methods*. 2020;12(35):4303–9
- [22] WU, W.; MALLETT, Y.; WALCZAK, B.; PENNINGCKX, W.; MASSART, D. L.; HEUERDING, S.; ERNI, F. Comparison of regularized discriminant analysis linear discriminant analysis and quadratic discriminant analysis applied to NIR data. *Analytica Chimica Acta*, v. 329, n. 3, p. 257-265, 1996.

- 1422 [23] MUCHERINO, A.; PAPAJOGEI, P. J.; PARDALOS, P. M. K-nearest
1423 neighbor classification. In: Data mining in agriculture. Springer, New York,
1424 NY, 2009. p. 83-106.
- 1425 [24] NOBLE, W. S. What is a support vector machine? Nature biotechnology,
1426 v. 24, n. 12, p. 1565-1567, 2006.
- 1427 [25] WONG, T.T. Performance evaluation of classification algorithms by k-fold
1428 and leave-one-out cross validation. Pattern Recognition, v. 48, n. 9, p.
1429 2839-2846, 2015.
- 1430 [26] THEERTHAGIRI, P. Predictive analysis of cardiovascular disease using
1431 gradient boosting-based learning and recursive feature elimination
1432 technique. Intelligent Systems with Applications, v. 16, p. 200121, 2022.
- 1433 [27] RIOS, T.G.; LARIOS, G.; MARANGONI, B.; OLIVEIRA, S. L.; CENA, C.;
1434 RAMOS, C. A. N. FTIR spectroscopy with machine learning: A new
1435 approach to animal DNA polymorphism screening. Spectrochim Acta A Mol
1436 Biomol Spectrosc. 2021.
- 1437 [28] FRANCA, T.; GONCALVES, D.; CENA, C. ATR-FTIR spectroscopy
1438 combined with machine learning for classification of PVA/PVP blends in
1439 low concentration. Vibrational Spectroscopy 120
- 1440 [29] KALLESAMURTHY, T.; SHEKAR, R.; NIRANJANAMURTHY, H. H.;
1441 NATESAN, K.; SHOME, B. R.; BAMBAL, R. G.; ... & SHOME, R. Assessment
1442 of fluorescence polarization assay: a candid diagnostic tool in Brucella abortus
1443 strain 19 vaccinated areas. Microbiology and immunology, v. 62, n. 11, p. 694-
1444 701, 2018.

5. CAPÍTULO II

Evaluation of UV spectroscopy and machine learning for serological diagnosis of bovine tuberculosis

Bruno Silva de Rezende¹, Thiago Franca², Cícero Cena², Carlos Alberto do Nascimento Ramos^{1*}

[*carlos.nascimento@ufms.br](mailto:carlos.nascimento@ufms.br)

- 1- UFMS – Universidade Federal de Mato Grosso do Sul, Faculdade de Medicina Veterinária e Zootecnia (FAMEZ), Campo Grande, MS, Brasil
- 2- UFMS –Universidade Federal de Mato Grosso do Sul, Optics and Photonic Lab (SISFOTON-UFMS), Campo Grande, MS, Brasil

Abstract

Bovine tuberculosis is a disease of great importance for cattle, causing significant economic losses and risks to human health. The sanitary measures currently used in Brazil are limited to the diagnosis and sacrifice of infected animals. Furthermore, it is known that the diagnostic methodologies that are used may present some problems, either due to the high cost, delay in execution time or the presentation of false positives and false negatives. Thus, further studies are needed to develop more efficient diagnostic methodologies. In this context, as an alternative and promising measure for the diagnosis of diseases, optical spectroscopy associated with machine learning has shown great potential. In the present study, bovine blood serum was used for evaluation, through UV spectroscopy and machine learning algorithms, with the objective of making possible a standard for the diagnosis of *Mycobacteriym bovis* in cattle. As sources of antigens, recombinant proteins P27, MPB70 and MPB83 were used. Only with the P27 protein was it possible to observe the formation of well-defined clusters of positive and negative samples. Applying the KNNM machine learning methodology, it was possible to obtain an accuracy of 96.3%, with a sensitivity of 100% and specificity of 92.6%. With the P27 protein as an antigen, the methodology evaluated in the present study proved to be promising as a

diagnostic tool for bovine tuberculosis. But studies will be necessary to elucidate the poor performance of the MPB70 and MPB83 antigens.

Keywords: photodiagnosis, diagnostic test, UV-vis spectroscopy, machine learning.

Introduction

Bovine tuberculosis, caused by *Mycobacterium bovis*, is responsible for great economic impact on livestock. In addition, *M. bovis* is an important zoonotic agent (DABORN et al., 1995). *M. bovis* is an alcohol-acid resistant bacilli, gram-positive, aerobic, non-capsulated, non-spore forming, non-flagellated, immobile and slow-growing (CORSI, 2021). Furthermore, this pathogen has a high concentration of lipids in its cell wall, thus contributing to resistance to most disinfectants, antimicrobials and the host's immune system (PALOMINO et al., 2007).

The disease in cattle is characterized by lower milk production, drop in weight gain, cachexia, eventually death and condemnation of carcasses in slaughterhouses. In Brazil, the last major prevalence study carried out between 1989 and 1998, found a prevalence of 1.3%, and reduction of the productive efficiency of the herds between 10% and 25% (BRASIL, 2006).

Transmission occurs mainly through aerobic route, but hosts can also become infected by ingesting contaminated water and food such as grass and milk (DSAIP, 2021). According to Neill et al. (1994), even though it is of little importance, transplacental transmission can occur, in addition to infection through intercourse by contaminated semen.

After infection, the microorganisms are captured by macrophages. The cellular immune response that is triggered, associated with neutrophils and macrophages, results in lesions with the appearance of a granuloma characterized by caseous necrosis and calcification in the center of the tubercle (JONES et al., 2000). When the infection occurs via the aerobic route, the lungs and regional lymph nodes are the main organs affected. However, when the animal ingests the bacillus, the mesenteric and pharyngeal lymph nodes can be

affected, as well as in generalized infection all organs can be affected (SALAZAR, 2005).

The clinical signs of the disease will depend on the affected organs, but in general it can be said that bovine tuberculosis is a pathology of a chronic nature and that eventually it can progress to an acute condition. Thus, the clinical signs that may be present are hyperplasia of superficial and/or deep lymph nodes, endometritis, infertility, mastitis, dyspnea, cough, cachexia, fever, pneumonia, lymphadenopathy, diarrhea, among others. In addition, there are cases of congenital tuberculosis, which in turn can progress and cause death in newborn calves (DSAIP, 2021; CORSI, 2021).

Therefore, due to the importance of the disease, with the objective of reducing the prevalence and incidence of brucellosis and tuberculosis, in 2001, the National Program for the Control and Eradication of Brucellosis and Animal Tuberculosis (PNCEBT) was instituted by the Ministry of Agriculture, Livestock and Food Supply (MAPA), of Brazil. And among the health measures employed by the program, the detection of infected animals and removing them from herds is critical (BRASIL, 2006). For this, the tuberculin test has been the main diagnostic tool used for in vivo diagnosis of bovine tuberculosis in Brazil.

Tuberculin tests consist of intradermal application of tuberculin (an extract of proteins from *Mycobacterium*) and subsequent evaluation of the type IV hypersensitivity response. Animals that are infected have an exacerbated response and have a significant increase in the skin fold at inoculation site (MONAGHAN et al., 1994). The variations of the tuberculin test currently in use in Brazil are: Single Intradermal Comparative Cervical Tuberculin (SICCT) Caudal Fold Tuberculin Test (CFTT) and Comparative Intradermal Tuberculin Test (CITT). The first two tests are used as screening tools, while the last one is a confirmatory test (BRASIL, 2006).

Although tuberculin tests have several qualities, they still have disadvantages that require research to develop new diagnostic strategies. For example, the test requires handling each animal in the herd at two different times, which makes it difficult to apply in large herds (MENZIES, 1999; LIU et al., 2007). In addition, false-positive reactions may occur due to sensitization of the animals with environmental microbacteria, and false-negative reactions may occur due

chronically infected animals already in a state of anergy (WOOD et al., 1991; MONAGHAN, 1994; BRASIL, 2004; AAGAARD et al., 2006). In this context, optical spectroscopy, associated with powerful data analysis tools such as machine learning, has emerged as alternative techniques that can be exploited as diagnostic tools (Sakudo et al., 2012; Auner et al., 2018; Brito et al. al., 2022).

Therefore, in this study, UV spectroscopy associated with machine learning was evaluated for serological diagnosis of bovine tuberculosis, using proteins MPB70, MPB83 and P27 as antigens.

Materials and Methods

Sample and antigens

In this study were used a total of 88 samples of bovine blood serum, 50% of them from infected animals, and another 50% from non-infected animals. The infectious status of the animals was based on the presence or absence of a positive result in the tuberculin test and identification of compatible lesions in carcasses at slaughter.

The antigen sources were recombinant proteins P27, MPB70 and MPB 83 (all with a concentration of 1 mg/ml), produced as described by SOUZA et al. (2012) and provided by Embrapa Gado de Corte, Campo Grande, MS.

Reaction and acquisition of UV spectra

The best reaction conditions were evaluated with 14 positive and 14 negative samples for *Mycobacterium bovis*. Initially, different antigen-serum proportions (1:1 – 1:16), for each antigen, was performed. After homogenization at room temperature, 2uL of each dilution was submitted to collection of UV spectrum, in NanoDrop One C (Thermo Fisher Scientific), between 200 and 300 nm. The spectra of each serum/antigen combination were exported to a spreadsheet and submitted to multivariate analysis (Principal Component Analysis - PCA), with the aid of the software (PAST 4.03) using the scatter plot methodology. The best serum antigen proportion was defined as the one with the greatest distance between the clusters of the groups (positive and negative).

Thus, better results were observed: P27 and MPB70 (1:2 antigen/serum proportion), and MPB83 antigen (1:16 antigen/serum proportion). With the best

proportions for each antigen defined, the total number of samples were evaluated.

Data analysis and sample classification

Data analysis was implemented and performed in the Python programming language (version 3.9.12), using the Scikit-learn package (version 1.1.2) (PEDREGOSA et al., 2011). First, each UV-Vis spectrum was subjected to pre-processing using Standard Normal Variation (SNV). This method allows removing the baseline variation and resizing the absorbances, preventing random experimental variations from interfering with the result (RINNAN et al., 2009). Then, the UV-SNV spectra, in the range of 200 nm to 300 nm, were subjected to principal component analysis (PCA) (JOLLIFFE et al., 2016), an unsupervised method that converts multiple variables into some principal components through dimension reduction, retaining most of the original variable information and expressing the data variance in a score chart. It also allows us to analyze the variation of the main data in the analyzed spectral range through the loading graph. The PCA is an important step to visualize the group classification trend, and the output data showed better performance to build a prediction model based on machine learning algorithms than on raw spectral data.

Finally, a sample classification test was performed using PCA data subjected to machine learning (ML) algorithms. We use ML algorithms based on three different methods: (i) Discriminant Analysis (DA), which classifies the sample based on the distance between the sample data and the constructed contour using a linear or quadratic function to separate the classes (group) (Wu et al., 1996); (ii) k-Nearest Neighbor (KNN) uses the Euclidean distance between k (k= 1, 5, 10, 100) nearest neighbors to classify the sample (MUCHERINO et al., 2009); and (iii) Support Vector Machine (SVM), which organizes each sample class through the optimization of a hyperplane – the hyperplane can be linear or non-linear, being optimized to achieve high performance – between classes (NOBLE, 2006).

Each ML algorithm can use a certain number of PCs (input data) to build the prediction model, but to avoid overfitting and underfitting it is mandatory to determine the ideal number of Principal Components (PCs). Here we use Feature

Selection Recursive Feature Elimination (RFE), which can select the top PCs that contribute the most to achieving high accuracy and remove other PCs with the weakest contribution. The use or removal of a given PC was based on the accuracy obtained using Linear Discriminant Analysis (LDA) to classify samples in a Leave One Out Cross Validation Test (LOOCV).

In LOOCV, a sample is taken from the dataset and the others are used to build the prediction (training) model. Then, the sample data draw is used to assess the quality of the model (test). The procedure is repeated until all sample data has been tested. The quality of the model is measured by precision, remaining the average percentage of correct answers in each test (WONG, 2015).

Then, a ranking of CPs is established and the CPs with the weakest contribution to accuracy are gradually removed from the dataset until the highest accuracy is obtained with the smallest number of CPs (THEERTHAGIRI, 2022). Finally, once the main PCs were determined and selected by the RFE, a new LOOCV test was performed using the DA, KNN and SVM algorithms.

Results and Discussion

Figure 13 shows the scatter plot based on principal component analysis results for the UV-SNV spectral data of the bovine tuberculosis positive and negative groups. Figure 13^a shows the result using P27 antigen, Figure 13^b to MPB70 and Figure 13^c to MPB83. Only with the P27 antigen was it possible to observe the formation of two clusters (positive and negative). For the MPB70 and MPB83 antigens, additional clusters were formed, containing samples previously classified as positive or negative.

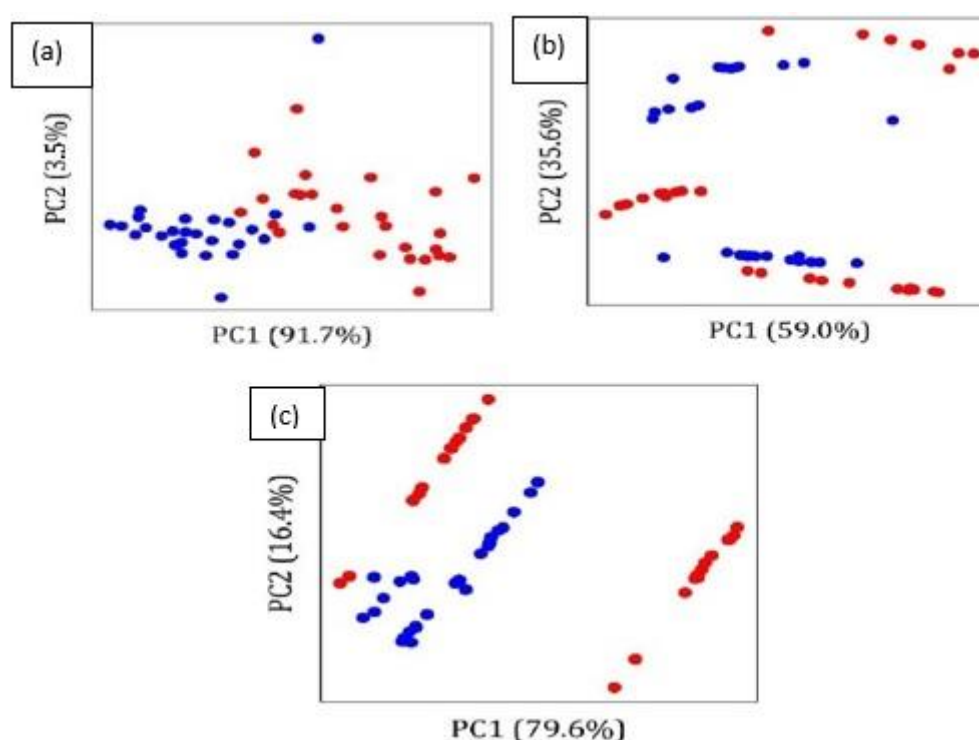


Figure 13: Scatter plot based on Principal component analysis for UV-SNV spectra of bovine blood serum. Bovine tuberculosis positive samples (red) and negative (blue). The (a) correspond to the P27 antigen, (b) correspond to the MPB70 antigen, and (c) to the MPB83 antigen.

Proteins P27, MPB70 and MPB83 has been explored as an antigen for serological diagnosis of bovine tuberculosis for more than a decade (Farias et al., 2012; Souza et al., 2012). In evaluation of antigens (P27 and MPB70) in ELISA tests, FARIAS et al. (2012) observed that the sensitivity and specificity of the tests were respectively 98.1% and 91.9% for P27 and 88,7% and 94.6% for MPB70, Already Souza et al. (2012), in a chimeric construction composed of MPB70, MPB83 and ESAT6, also observed in the ELISA test, sensitivity of 83.2% and specificity of 86.5%. In the present study, satisfactory results were observed only using P27 as antigen (Figure 13^a).

In Figures 13^b and 13^c, it is possible to observe the formation of at least three clusters, composed of mixture of positive and negative samples. Although we still do not have a definitive answer to explain the results with the MPB70 and MPB83 antigens, one hypothesis is that there were false-negative and false-

positive results in the reference tests. And thus, the training of the machine learning algorithm was compromised, making it difficult to accurately classify the samples.

The infectious status of the animals from which the serum samples were used in the present study was based on tuberculin tests and identification of lesions compatible with tuberculosis (LST). However, both methods have high rates of misdiagnosis. In a study conducted by Souza et al. (2012), 24 of the 32 CITT-negative cattle from infected herds exhibited LST, from which *M. bovis* was isolated. On the other hand, in LST collected from 39 carcasses of cattle slaughtered in a region of Brazil, and subjected to specific microbiological isolation for *M. bovis*, only in 17.94% (7/39) was observed positive bacteriological diagnosis for *Mycobacterium* spp (Ferreira et al., 2021).

The performance differences observed between P27 and the other antigens in the present study may also be related to the samples that were used in each evaluation. Although the total number of samples used in the study was 88 (44 positive and 44 negative), the applied outliers identification methodology excluded different samples for each antigen from the analyses. And effectively, 54 samples were used for P27 analysis, 60 for MPB70 and 56 for MPB83. This, associated to the fact that the kinetics of antibody production in animals infected with *M. bovis* varies according to the stage of infection and due the individual genetic of each animal, may have contributed to the large differences observed (Waters et al., 2017; Souza et al., 2019). In addition, kinetics of the antibody responses to different *M. bovis* antigens is variable upon infection. Anti-MPB83 antibodies can be detected relatively early, typically around four weeks post-infection (Waters et al., 2006). Anti-ESAT6 antibodies can be detected 12 weeks after experimental infection (Lyashchenko et al., 1998). In contrast, antibodies for MBP70 usually develop 18 to 22 months after experimental infection (Harboe et al., 1990; Fifiş et al., 1992). PC1 and PC2 were subjected to machine learning algorithms to build a prediction model for sample classification. For the P27 antigen, data were submitted to algorithms (LDA, QDA, KNNF, KNNM, KNNG, KNNC, KNN3, KNNW, SVML, SVM2, SVM3, SVMF, SVMM, SVMG). Figure 14^a shows the maximum overall accuracy achieved by the algorithms in a LOOCV

test using different classification functions. The maximum overall accuracy achieved by the precision models was 96.3% for KNNM.

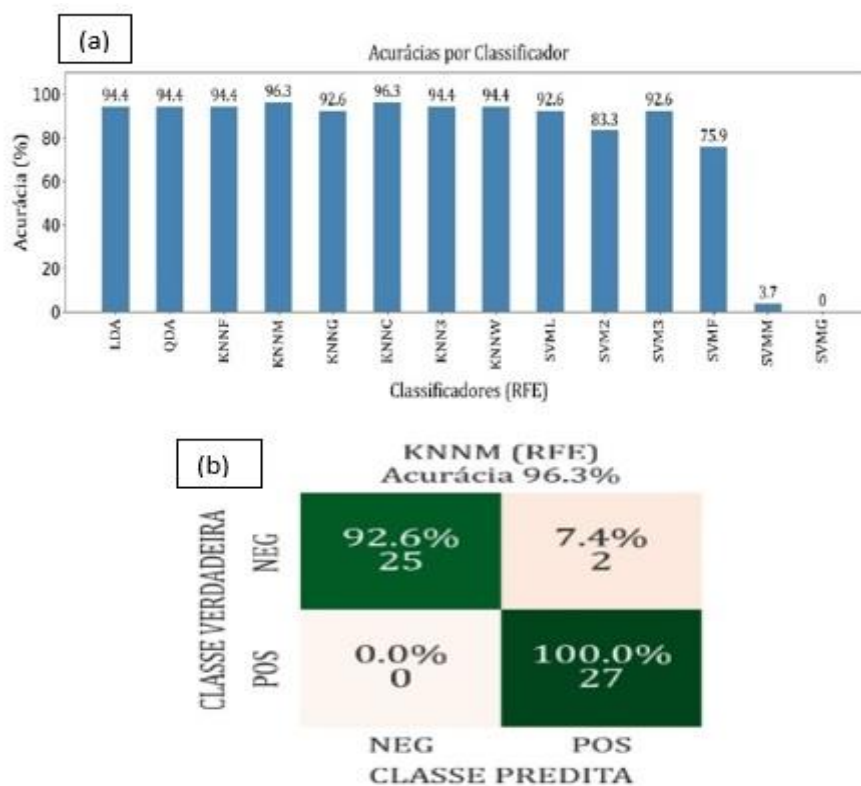


Figure 14: (a) Overall accuracy obtained in LOOCV tests (light blue bars) for the machine learning algorithms. (b) Confusion matrix for the performance of the KNNM algorithm with an accuracy of 96.3%.

The confusion matrix, figure 14^b, shows the sample classification for KNNM. From the sample classification values, it was possible to determine for KNNM a sensitivity of 100.0% and a specificity of 92.6%. To compare, in a study carried out by Jones et al. (1992), they performed the interferon gamma test using bovine PPD and observed 81.8% sensitivity and 99.1% specificity. Furthermore, this same author used bovine PPD in the tuberculin test and observed sensitivity of 68.1% and a specificity of 96.7%. Furthermore, other methodologies such as ELISA using chimera with hydrophilic fragments of MPB70, MPB83 and ESAT-6 from *M. bovis*, showed a result of sensitivity of 83.2% and specificity of 86.5% (Souza et al., 2012).

In this study (UV spectroscopy and machine learning), it was possible to identify positive and negative samples for *M. bovis* with sensitivity and specificity similar to those observed in the tests usually employed, using P27 antigen. In addition, it is worth mentioning that the results can be obtained in an average time of 3 minutes, which is faster than the tests currently employed. Another positive point was the lower complexity for processing samples, requiring less labor and possibility of automation.

Conclusion

The proposed methodology, using UV-Vis associated with machine learning, was able to identify and separate positive and negative samples for *Mycobacterium bovis* antibodies, using the P27 antigen, with a total accuracy of 96.3%, with sensitivity and specificity above 90%. For the other antigens, it was not possible to observe separation of two groups according to the cluster results, and therefore it was not possible to clearly identify positives and negatives. Although the methodology has proven to be useful and practical, further studies and adjustments in machine learning algorithms and sample classification are still needed, so that we can better understand the poor performance of the MPB70 and MBP83 antigens.

Acknowledgments:

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), code 001. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), code 403651/2020-5; 302525/2022-0; 440214/2021-1. Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT), code 007/2019; 360/2022. Embrapa Gado de Corte. Universidade Federal de Mato Grosso do Sul – UFMS/MEC – Brasil.

References

AAGAARD, C.; GOVAERTS, M.; MEIKLE, V. et al. Optimizing antigen cocktails for detection of *Mycobacterium bovis* in herds with different prevalences of bovine tuberculosis: ESAT6-CFP10 mixture shows optimal sensitivity and specificity. **Journal of Clinical Microbiology**, v.44, n.12, p.4326-4335, 2006.

- 1720 AUNER, G.W.; KOYA, S.K.; HUANG, C. et al. Applications of Raman
1721 spectroscopy in cancer diagnosis. **Cancer Metastasis Review**, v.37, p.691–717,
1722 2018.
- 1723 **Brasil. Ministério da Agricultura, Pecuária e Abastecimento.** Secretaria da
1724 Defesa Sanitária. Instrução normativa SDA nº 06, de 08 de janeiro de 2004,
1725 aprova o regulamento técnico do Programa nacional de controle e erradicação
1726 da brucelose e tuberculose animal. *Diário Oficial da União*, Brasília, DF, Seção
1727 1, p. 6-10, 2004.
- 1728 **Brasil. Ministério da Agricultura, Pecuária e Abastecimento.** Programa
1729 Nacional de Controle e Erradicação da Brucelose e da Tuberculose Animal
1730 (PNCEBT). Brasília: MAPA/SDA/DSA, 2006. 188 p.
- 1731 BRITO, E.C.A.; FRANÇA, T.; CANASSA, T. et al. Paracoccidioidomycosis
1732 screening diagnosis by FTIR spectroscopy and multivariate analysis.
1733 **Photodiagnosis and Photodynamic Therapy**, v.39, p.102921, 2022.
- 1734 CORSI, C. **Tuberculose.** Disponível em:
1735 https://www.crmvsp.gov.br/arquivo_zoonoses/TUBERCULOSE.pdf. Acesso em:
1736 05 abr. 2021.
- 1737 DABORN, C.J.; O'REILLY, L.M. **The epidemiology of *Mycobacterium bovis***
1738 **infections in animals and man: a review.** *Tuber Lung Dis*, v.76, p.1-46, 1995.
- 1739 **DSAIP - Departamento de Saúde Animal e Insumos Pecuários.** Ficha
1740 Técnica. Disponível em:
1741 http://sistemasweb.agricultura.gov.br/pages/fichas_tecnicas/Ficha_T%C3%A9cnica_TUBERCULOSE_jan20.pdf. Acesso em: 04 de abril de 2021.
- 1743 FARIAS, T.A.; ARAÚJO, F.R., OSÓRIO, A.L.A.R. et al. ELISA based on
1744 recombinant MPB70 and P27 for detection of antibodies against *Mycobacterium*
1745 *bovis*. **Revista de Patologia Tropical**, v.41, n.2, p.155-162, 2012.
- 1746 FERREIRA, F.F.; OLIVEIRA, M.L.M.; FILHO, F.A. et al. Isolation and genotyping
1747 of *Mycobacterium bovis* in suggestive lesions of tuberculosis in cattle slaughtered

- 1748 in the state of Ceará, Brazil. **Pesquisa Veterinária Brasileira**, v.40, p.863-870,
1749 2021.
- 1750 FIFIS, T.; COSTOPOULOS, C.; CORNER, L.A. et al. Serological reactivity to
1751 Mycobacterium bovis protein antigens in cattle. **Veterinary Microbiology**, v30,
1752 n.4, p.343–354, 1992.
- 1753 HARBOE, M.; WIKER, H.G.; DUNCAN, J.R. et al. Protein G-based enzyme-
1754 linked mmunosorbent assay for anti-MPB70 antibodies in bovine tuberculosis.
1755 **Journal of Clinical Microbiology**, v.28, n.5, p.913–921, 1990.
- 1756 JOLLIFFE, I.T.; CADIMA, J. Principal component analysis: a review and recent
1757 developments. **Philosophical transactions of the royal society A:**
1758 **Mathematical, Physical and Engineering Sciences**, v.374, n.2065,
1759 p.20150202, 2016.
- 1760 JONES, T.C.; HUNT, R.D.; KING, N.W. **Patologia Veterinária**. 6.ed. São Paulo:
1761 Manole, 2000. 1415p.
- 1762 JONES, S.L.; COX, J.C.; SHEPHERD, J.M. et al. Removal of false-positive
1763 reactions from plasma in an enzyme immunoassay for bovine interferon-gamma.
1764 **Journal of Immunology Methods**, v.155, n.2, p.233-240, 1992.
- 1765 LIU, S.; GUO, S.; WANG, C. et al. A novel fusion protein-based indirect enzyme-
1766 linked immunosorbent assay for the detection of bovine tuberculosis.
1767 **Tuberculosis**, v.87, n.3, p.212–217, 2007.
- 1768 LYASHCHENKO, K.P.; POLLOCK, J.M.; COLANGELI, R. et al. Diversity of
1769 antigen recognition by serum antibodies in experimental bovine tuberculosis.
1770 **Infection and Immunity**, v66, n.11, p.5344–5449, 1998.
- 1771 MONAGHAN, ML, DOHERTY, M.L, JD COLLINS. et al. 1994. The tuberculin test.
1772 **Veterinary Microbiology**, v.40, n.1-2, p.111-124, 1994.
- 1773 MENZIES, D. Interpretation of repeated tuberculin tests. Boosting, conversion,
1774 and reversion. **American journal of respiratory and critical care medicine**,
1775 v.159, n.1, p.15–21, 1999.

- 1776 MUCHERINO, A.; PAPAJORGJI, P.J.; PARDALOS, P.M. K-nearest neighbor
1777 classification. In: **Data mining in agriculture**. Springer: New York, p. 83-106,
1778 2009.
- 1779 NEILL, S.D.; POLLOCK, J.M.; BRYSON, D.G. et al. Pathogenesis of
1780 Mycobacterium bovis infection in cattle. **Veterinary Microbiology**, v. 40, n.1-2,
1781 p. 41-52, 1994.
- 1782 NOBLE, W.S. What is a support vector machine? **Nature biotechnology**, v.24,
1783 n.12, p. 1565-1567, 2006.
- 1784 PALOMINO, J.C.; LEÃO, S.C.; RITACCO, V. **Tuberculosis 2007: From basic**
1785 **science to patient care**. 1.ed. Bourcillier Kamps, 2007. 687p.
- 1786 PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A. et al., **Journal of**
1787 **Machine Learning Research**, v.12, p.2825-2830, 2011.
- 1788 RINNAN, Å.; BERG, F.V.D.; ENGELSEN, S.B. Review of the most common pre-
1789 processing techniques for near-infrared spectra. **TrAC Trends in Analytical**
1790 **Chemistry**, v.28, n.10, p.1201-1222, 2009.
- 1791 SAKUDO, A.; SUGANUMA, Y.; SAKIMA, R.; et al. Diagnosis of HIV-1
1792 infection by near-infrared spectroscopy: Analysis using molecular clones of
1793 various HIV-1 subtypes. **Clinica Chimica Acta**, v.413, n 3–4, p.467-472,
1794 2012.
- 1795 SALAZAR, F.H.P. **Ocorrência de tuberculose causada por Mycobacterium**
1796 **bovis em bovinos abatidos e frigoríficos no estado de Mato Grosso**. 2005.
1797 68 p. Dissertação (Mestrado) - Curso de Medicina Veterinária, Universidade
1798 Federal do Mato Grosso do Sul, Campo Grande, 2005.
- 1799 SOUZA, I.I.F.; RODRIGUES, R.A.; JORGE, K.S.G. et al. ELISA using a
1800 recombinant chimera of ESAT-6/MPB70/MPB83 for Mycobacterium bovis
1801 diagnosis in naturally infected cattle. **Journal of Veterinary Medicine Science**,
1802 v.81, n.1, p.9-14, 2009.

- 1803 SOUZA, I.I.; MELO, E.S.; RAMOS, C.A.N. et al. Screening of recombinant
1804 proteins as antigens in indirect ELISA for diagnosis of bovine
1805 tuberculosis. **Springerplus**, v.1, n.1, p.1-6, 2012.
- 1806 THEERTHAGIRI, P. Predictive analysis of cardiovascular disease using gradient
1807 boosting-based learning and recursive feature elimination technique. **Intelligent**
1808 **Systems with Applications**, v. 16, p. 200121, 2022.
- 1809 WATERS, W.R.; PALMER, M.V.; THACKER, T.C. et al. Early antibody responses
1810 to experimental Mycobacterium bovis infection of cattle. **Clinical and Vaccine**
1811 **Immunology**, v.13, n.6, p.648–654, 2006.
- 1812 WATERS, W.R.; VORDERMEIER, H.M.; RHODES, S. et al. Potential for rapid
1813 antibody detection to identify tuberculous cattle with non-reactive tuberculin skin
1814 test results. **BMC Veterinary Research**, v.13, p.1-7, 2017.
- 1815 WONG, T.T. Performance evaluation of classification algorithms by k-fold and
1816 leave-one-out cross validation. **Pattern Recognition**, v. 48, n. 9, p. 2839-2846,
1817 2015.
- 1818 WOOD, P.R.; ORNER, L.A.; ROTHEL, J.S. et al. Field comparison of the
1819 interferon gamma assay and the intradermal tuberculin test for the diagnosis of
1820 bovine tuberculosis. **Australian Veterinary Journal**, v.68, n.9, p.286-290, 1991.
- 1821 WU, W.; MALLET, Y.; WALCZAK, B. et al. Comparison of regularized
1822 discriminant analysis linear discriminant analysis and quadratic discriminant
1823 analysis applied to NIR data. **Analytica Chimica Acta**, v.329, n.3, p.257-265,
1824 1996.

6. CONSIDERAÇÕES GERAIS

Neste estudo evidenciou-se o grande impacto que as doenças brucelose e tuberculose bovina ocasionam e, com isso, a técnica de espectroscopia associada ao aprendizado de máquina mostra-se útil e prática, colaborando para bons resultados para o diagnóstico de brucelose. Ademais, para *Brucella abortus*, a metodologia empregada foi capaz de identificar amostras positivas e negativas, com acurácia global de 92,5%, com sensibilidade e especificidade acima de 90%. Para *Mycobacterium bovis*, utilizando antígeno P27, foi possível identificar amostras positivas e negativas, com acurácia total de 96,3%, com sensibilidade e especificidade acima de 90%.

Tal técnica foi capaz de fornecer resultados com tempo médio de 5 minutos e exige menos processamento de amostras, o que possibilita a redução de custos laboratoriais e mais qualidade da amostragem. Contudo, esta ferramenta, embora possa ser automatizada e sem risco biológico, requer material complexo e pessoas com experiência em bioinformática.

Diante do exposto, portanto, a espectroscopia associada ao aprendizado de máquina se mostra uma ferramenta promissora. É importante destacar que para diagnóstico de *B. abortus*, tal ferramenta apresentou bons resultados colaborando assim como uma alternativa de diagnóstico. Entretanto, para *M. bovis*, o mesmo não pode ser dito, haja vista que para apenas um antígeno (P27), foi obtido bons resultados. Assim, serão necessários mais ajustes e estudos.

7. IMPACTO ECONÔMICO, SOCIAL, TÉCNOLÓGICO E/OU INOVAÇÃO

Os resultados alcançados com a presente pesquisa apresentam grande relevância social e econômica, pois, podem contribuir em um futuro próximo com o diagnóstico preciso e rápido das enfermidades brucelose e tuberculose. O diagnóstico mais preciso, por sua vez, permitirá um controle mais efetivo das enfermidades, reduzindo as barreiras econômicas importadas aos produtos em decorrência da ocorrência das doenças no rebanho brasileiro.

Ademais, o método desenvolvido e avaliado como ferramenta de diagnóstico, pode também ser explorado como ferramenta para diagnóstico de outras enfermidades animais. Inclusive com grande potencial para exploração

1857 industrial, por meio da produção de equipamentos portáteis para diagnóstico
1858 sorológico de doenças.

1859 Especificamente relacionado ao diagnóstico de tuberculose e brucelose,
1860 a metodologia aqui apresentada, com um pouco mais de ajustes, permitirá
1861 reduzir o tempo médio de diagnóstico de 72 horas para alguns minutos, e 48
1862 horas para alguns minutos, respectivamente.