











Padrões temporais e espaciais na detecção de resíduos de medicamentos veterinários em aves e suínos no Brasil

Temporal and spatial patterns in the detection of veterinary drug residues in poultry and swine in Brazil

Anna Julia Portz¹ , Nilton Silva¹ , Gustavo Lima¹ , Leandro Feijó² , Helder Louvandini³ , Vanessa Peripolli^{4*} , Renata Vieira¹ , Concepta McManus¹ 

¹Universidade de Brasília (UnB), Brasília, Distrito Federal, Brasil.

²Ministério da Agricultura, Pecuária e Abastecimento, Brasília, Distrito Federal, Brasil.

³Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, São Paulo, Brasil.

⁴Instituto Federal Catarinense (IFC), Campus Araquari, Araquari, Santa Catarina, Brasil.

*Correspondente: vanessa.peripolli@hotmail.com

Resumo

A Segurança Alimentar é um tema importante para a saúde pública e o comércio internacional de alimentos. Resíduos de medicamentos veterinários e contaminantes ambientais em produtos de origem animal podem causar doenças e toxicidade aguda em organismos expostos a essas substâncias. Este estudo avaliou dados oficiais de monitoramento de resíduos de medicamentos veterinários do Ministério da Agricultura, Pecuária e Abastecimento em tecidos de aves e suínos no período de 2002 a 2014 para verificar padrões ocultos na ocorrência de seis medicamentos comuns (Closantel, Diclazuril, Nicarbazina, Sulfaquinoxalina, Doxiciclina e Sulfametazina). A análise dos dados foi realizada por meio de dois métodos de aprendizado de máquina: árvore de decisão e redes neurais, além da avaliação visual por meio de gráficos e mapas. As taxas de contaminação foram baixas, variando de 0 a 0,66%. Foi identificado um padrão de distribuição espacial das detecções de substâncias por região, mas nenhum padrão de distribuição temporal foi observado. No entanto, as regressões mostraram um aumento nos níveis quando essas substâncias foram detectadas, portanto, o monitoramento deve continuar. No entanto, os resultados mostram que os produtos monitorados durante o período do estudo apresentaram baixo risco à saúde pública.

Palavras-chave: Aprendizagem de máquina; saúde pública; segurança alimentar; resíduos

Abstract

Food Safety is an important topic for public health and international trade in food. Residues of veterinary drugs and environmental contaminants in animal products can cause diseases and acute toxicity in organisms exposed to these substances. This study evaluated official monitoring data of veterinary drug residues from the Brazilian Ministry of Agriculture, Livestock and Supply in tissues of poultry and swine in the period between 2002 and 2014 to check for hidden patterns in the occurrence of six common drugs (Closantel, Diclazuril, Nicarbazine, Sulfaquinoxaline, Doxycycline and Sulfamethazine). The analysis of data was performed by using two machine learning methods: decision tree and neural networks, in addition to visual evaluation through graphs and maps. Contamination rates were low, varying from 0 to 0.66%. A spatial distribution pattern of detections of substances by region was identified, but no pattern of temporal distribution was observed. Nevertheless, regressions showed an increase in levels when these substances were detected, so monitoring should continue. However, the results show that the products monitored during the study period presented a low risk to public health.

Keywords: Machine learning; food safety; public health; residues

Recebido: 2 de abril de 2022. Aceito: 10 de maio de 2022. Publicado: 20 de junho de 2022.



Introdução

O Brasil tem papel de destaque no cenário internacional como produtor e exportador de produtos agrícolas. Além de ser o quarto maior produtor de grãos do mundo, também é o segundo maior exportador de grãos, com 19% do mercado internacional⁽¹⁾ abastecendo mais de 180 países com produtos agrícolas. No setor de aves e suínos esse cenário não é diferente. O Brasil exporta carne de frango para 151 países e carne suína para 97 países⁽²⁾.

Sistemas eficazes de segurança alimentar são essenciais para a saúde pública e para a confiança do mercado consumidor interno e dos consumidores internacionais. Contaminação química e microbiológica são as principais causas de doenças transmitidas por alimentos⁽³⁾. Resíduos de medicamentos veterinários e contaminantes ambientais que entram na cadeia produtiva podem causar efeitos adversos no organismo humano como toxicidade aguda, reações alérgicas, ruptura da flora intestinal normal, mutagenicidade, teratogenicidade e carcinogenicidade^(4,5). Estes são controlados com boas práticas agropecuárias, que mitigam o risco dessas substâncias atingirem níveis prejudiciais à saúde humana^(6,6). As estratégias de identificação e mitigação na fazenda precisam ser avaliadas para entender seu impacto na redução de doenças em animais e humanos, pois os alimentos têm sido identificados como um importante veículo de transmissão de vírus e bactérias^(7,8). Painter et al.⁽⁹⁾ estimaram mais de 9 milhões de doenças transmitidas por alimentos causadas pelos principais patógenos adquiridos nos Estados Unidos a cada ano. Os mesmos autores atribuíram a maioria das doenças a commodities de animais terrestres e mais mortes a aves do que qualquer outro produto. Mesmo os produtos orgânicos podem apresentar níveis significativos de contaminantes ambientais⁽¹⁰⁾.

Os contaminantes ambientais são difíceis de controlar. A produção ao ar livre deixa os animais potencialmente mais expostos a contaminantes ambientais, como dioxinas, furanos e bifenilos policlorados (PCBs)⁽¹¹⁾, enquanto os animais produzidos em sistemas fechados podem ser expostos a retardadores de chama⁽¹²⁾. Metais pesados, também podem constituir risco. Cádmio⁽¹¹⁾, sulfato de cobre e zinco⁽¹³⁾, assim como arsênio e chumbo⁽¹⁴⁾ foram encontrados em produtos alimentícios. A ração pode conter produtos fitossanitários e fertilizantes, além de micotoxinas, com consequências associadas na contaminação química da carne^(11,15). Resíduos de medicamentos veterinários também podem ocorrer⁽¹¹⁾.

O Brasil monitora resíduos e contaminantes na produção animal desde 1986, quando foi instituído o Plano Nacional de Controle de Resíduos (PNCR) pelo Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Esses dados são analisados ao final de cada ano para possibilitar a elaboração do plano de monitoramento do ano seguinte⁽¹⁶⁾. A necessidade de decisões rápidas e assertivas em instituições públicas e privadas exige o uso de ferramentas de tomada de decisão que possam auxiliar

no processo decisório, essa necessidade pode ser suprida através do uso de técnicas de mineração de dados. Doyle e Erickson⁽¹⁷⁾ mostraram que a simulação computacional e os modelos de Aprendizado de Máquina (AM) têm sido utilizados com maior frequência nos últimos anos, inclusive na agricultura⁽¹⁸⁾. Esses autores constataram que esses métodos ainda são incipientes na produção animal e não foram encontrados nenhuma informação sobre o uso de aprendizado de máquina para prever a contaminação de alimentos em aves e suínos. Portanto, este estudo analisa o uso de árvores de decisão e redes neurais Mapa auto-organizado (SOM) e Mapa auto-organizado adaptável ao tempo (TASOM) para prever a contaminação de tecidos de suínos e aves com seis drogas comuns no Brasil.

Material e métodos

Fonte de dados - Plano Nacional de Controle de Resíduos (PNCR)

Foram avaliados dados do monitoramento de resíduos de medicamentos veterinários e contaminantes ambientais em aves e suínos no âmbito do Plano Nacional de Controle de Resíduos (PNCR) do MAPA. Os limites de referência dos Limites Máximos de Resíduos (LMRs) utilizados nas análises do PNCR foram adotados pelo MAPA com base nos limites determinados pela Agência Nacional de Vigilância Sanitária (ANVISA)⁽¹⁶⁾, quando existentes. Para as demais substâncias, foram utilizados os limites sugeridos pelo Codex Alimentarius⁽¹²⁾.

As amostras a serem coletadas foram determinadas por sorteio aleatório semanal realizado pelo Sistema de Resíduos (SISRES), sistema que distribui as amostras aleatoriamente entre os estabelecimentos cadastrados no Serviço de Inspeção Federal (SIF). As amostras foram coletadas de acordo com as instruções do Manual de Amostragem do Plano Nacional de Resíduos e Contaminantes, que consiste na atualização dos procedimentos de coleta⁽¹⁹⁾. As análises foram realizadas na Rede de Laboratórios do Ministério Oficial do Brasil, utilizando-se como método de análise a cromatografia líquida-espectrometria de massas em tandem (LC-MS/MS)^(20,21) e as amostras foram rastreadas por meio de um Sistema Informatizado (SIGLA), conectado ao SISRES. Os limites de detecção e quantificação foram estimados para cada analito, de acordo com as diretrizes do MAPA⁽²²⁾.

Descrição dos dados

Foram utilizados dados de detecção de resíduos e contaminantes do PNCR em aves e suínos de janeiro de 2002 a outubro de 2014. Esses dados foram coletados do banco de dados do Sistema SISRES, com autorização formal por escrito do Coordenador responsável por este Sistema. Na primeira etapa deste estudo, que consiste na análise dos dados por aprendizado de máquina, foram utilizados resultados de análise de todas as substâncias

analisadas no âmbito do PNCR para as cadeias de aves e suínos. Cada resultado da análise foi correlacionado com as informações que permitem a rastreabilidade das amostras, e essas informações referem-se ao período em

que as amostras foram coletadas e ao local onde essa coleta ocorreu. Assim, foram exportados da base de dados do SISRES para Excel os dados distribuídos nas categorias listadas na Tabela 1.

Tabela 1. Dados retirados do SISRES em formato de tabela

Nome do Dado	Descrição do dado
Tipo de Código de análise	Código do grupo químico da substância
UF	Unidade da Federação/Estado
Código da Espécie	Código da espécie amostrada para análise
SIF	Número do estabelecimento onde a(s) amostras foi/foram coletada(s) registrado no Serviço de Inspeção Federal (SIF)
Semana da Análise	Semana em que cada análise foi realizada
Tecido	Tecido do qual a amostra foi retirada: rim, fígado, gordura ou urina
Proprietário	Nome do proprietário da granja de origem do(s) animal/animais dos quais foi coletada amostra para análise
Endereço do Proprietário	Endereço do proprietário do(s) animal/animais do(s) qual/quais foi coletada amostra para análise
Granja	Nome da propriedade de origem do(s) animal/animais do(s) qual/quais foi coletada amostra para análise
Endereço da granja	Endereço da propriedade de origem do(s) animal/animais de cada amostra coletada
Código do Município	Código do município da propriedade de cada amostra coletada
UF da propriedade	Unidade da Federaçãestado da propriedade em que estavam o(s) animal/animais de cada amostra coletada
Município	Município da propriedade de cada amostra coletada
Número de animais	Número de animais no lote do qual a(s) amostra(s) foram coletadas
Tipo de violação	Amostra conforme ou não-conforme de acordo com o status 6 ou 7 da análise
Status	Avaliação do status do resultado
CEP	Código postal da propriedade de onde as amostras foram coletadas
CEP 2	Código postal do proprietário
Ano	Ano da análise

Os dados foram classificados de acordo com o Status da Análise (Tabela 2).

Tabela 2. Status dos resultados das análises dos valores consultados no sistema.

Status	Resultado da análise
6	Entre zero e o LMR de cada substância
7	Igual ou maior que o LMR de cada substância

LMR: Limite Máximo de Resíduo

Aprendizagem de máquina – Árvore de Decisão

Para desenvolver a árvore de decisão, inicialmente foi realizada a etapa de pré-processamento e transformação, ou engenharia de características dos dados, que consistiu em limpar e selecionar os dados para remover dados que pudessem gerar ruídos, interferir na análise e preencher os valores

ausentes. Essa limpeza e seleção consistiu na inclusão de um novo status de análise, status 5, que corresponde a resultados de análise iguais a 0 e a exclusão de algumas informações. Após essa alteração, a Tabela 2 foi modificada, conforme mostrado na Tabela 3.

Os dados excluídos foram: os dados confidenciais (nome e endereço das propriedades onde as amostras foram coletadas, bem como seus proprietários, CEP do proprietário e unidade da federação da propriedade de origem do(s) animal(es)). Informações consideradas indiferentes à análise: código da espécie (já que as duas espécies foram analisadas individualmente), código do proprietário (pessoa física ou jurídica), código do tecido (informação desnecessária, pois cada substância é avaliada em apenas um tecido nos dados utilizados neste estudo); Dados duplicados: código do município (foi utilizado apenas o nome do município), situação de infração (informação substituída pela situação dos dados). Dados ausentes foram definidos como 0. Após

o pré-processamento dos dados, a árvore de decisão foi preparada com o software livre YADT.

Tabela 3. Representação após a criação do Status 5

Status	Resultado da Análise	Results
5	Igual a 0	Negativo
6	Entre 0,1 e o LMR de cada substância	Conforme
7	Igual ou maior que o LMR de cada substância	Não-conforme

LMR: Limite Máximo de Resíduo

Para mineração de dados, foram inseridos no software dois arquivos, um com o banco de dados para treinamento e outro com os metadados, ou seja, a descrição dos dados por coluna, conforme Tabela 4. Não foi utilizado banco de teste, sendo preenchido no YADT o valor de 25% de utilização da base de treinamento para realização do teste. Também foram preenchidos dois outros fatores necessários para a criação da árvore de decisão: SPLIT, que é descrito como o número de novos casos para evitar a criação de novos ramos, e CONFIDENCE, sendo este o valor utilizado para a poda do algoritmo C4.5. Os valores utilizados são os mais citados na literatura: 2 e 25%, respectivamente.

Tabela 4. Descrição dos metadados

Nome do dado	Tipo de dado	Tipo da característica
Ano	Fragmento	Discreto
Semana	Inteiro	Discreto
Waste code	Inteiro	Discreto
Número de animais	Inteiro	Discreto
SIF	Inteiro	Discreto
UF	Fragmento	Discreto
Município	Fragmento	Discreto
CEP (Código postal)	Fragmento	Discreto
Resultado	Flutuador	Contínuo
Status	Inteiro	Classe

Análise da rede neural

Os bancos de dados de suínos e aves foram analisados por meio das redes neurais Mapa auto-organizado (SOM)⁽²³⁾ e Mapa auto-organizado adaptável ao tempo (TASOM)⁽²⁴⁾. Um sistema protótipo foi utilizado para realizar esta análise na plataforma online C++Builder® da Embarcadero®. Os dados foram carregados para esta plataforma no formato ".csv".

A interface principal do software permite ao usuário configurar quais variáveis devem ser consideradas⁽²⁵⁾, os parâmetros geométricos (altura e largura do mapa) e os parâmetros iniciais do treinamento do SOM (taxas iniciais de aprendizado e

taxas de decaimento da vizinhança). Os dois algoritmos foram parametrizados com os dados detalhados na Tabela 5. As amostras da base de dados utilizadas já possuem o rótulo que classifica o resultado da análise, denominado Status. A análise realizada consistiu em verificar a eficiência dos neurônios dos algoritmos em agrupar as amostras de acordo com este rótulo. Para medir essa eficiência, foram registradas as porcentagens da presença de cada classe em cada um dos neurônios das redes. O resultado ideal é que um neurônio tenha alta densidade de uma classe e baixa densidade das outras, o que significa que as classificou corretamente.

Tabela 5. Parâmetros usados nos algoritmos SOM e TASOM

Parâmetro	Valor SOM	Valor TASOM
Altura do mapa	2	2
Largura do mapa	2	2
Raio de vizinhança inicial	2	2
Taxa de aprendizagem inicial	1	1
Decaimento da taxa de aprendizagem	1	0,01
Decaimento do raio da vizinhança	1	0,01
Raio de vizinhança constante ¹	-	1
Taxa de aprendizagem Constante ²	-	1
Alfa	-	0,001
Beta	-	0,1

¹Constante que altera o cálculo do raio de vizinhança; ²Constante que altera o cálculo da taxa de aprendizagem.

Para verificar os dados utilizando os neurônios, no centro do protótipo havia duas tabelas, uma para cada neurônio, nas quais foram registrados os valores de entrada, a posição do neurônio e a taxa de erro da classificação dos dados. Para executar o teste, os algoritmos (SOM e TASOM) foram fornecidos primeiramente com todos os dados da tabela pré-processada. Posteriormente, os algoritmos foram fornecidos com os dados da tabela pré-processada, exceto quando o status foi igual a 5 (resultado igual a zero) e por fim os resultados da análise foram excluídos, dados diretamente vinculados ao status da análise, para avaliar a comportamento dos algoritmos.

Metodologia de Análise Espacial – Quantum GIS (QGIS)

Nesta etapa da análise dos dados, os resultados por substância foram separados em planilhas diferentes, para análise individual de cada série histórica, e selecionadas apenas as substâncias que apresentaram um número significativo de análises com resultados de Status 6 e 7, ou

seja, com resultados diferente de 0. A Tabela 6 representa as substâncias selecionadas para esta etapa e seus LMRs.

Duas novas informações foram inseridas: nome das substâncias analisadas e semana de análise, que consiste em uma sequência numérica que inicia na primeira semana do primeiro ano e termina na última

semana do último ano da série histórica de cada substância. O software livre QuantumGIS® foi usado para produzir os mapas de distribuição dos resultados da análise com status 6 e 7.

Tabela 6. Substâncias e limites selecionados para as análises

Espécie	Substância	Ação	Nº. de resultados conformes e não-conformes	Limite Máximo de Resíduo – LMR (µg/Kg)
Aves	Closantel	Salicilanilida halogenada com atividade antiparasitária	181	1.000
	Diclazuril	Triazinona coccidiostático antiprotozoário	30	500
	Nicarbazina	Carbanilida coccidiostático	179	200
	Sulfaquinoxalina	Sulfonamida antibiótico e coccidiostático	108	100
Suínos	Doxycycline	Tetraciclina antibiótico	13	600
	Sulfametazina	Sulfonamida antibacteriano	133	100

Os dados de geolocalização dos municípios brasileiros foram baixados do IBGE⁽²⁶⁾. Os nomes dos municípios brasileiros e seus geocódigos, que consiste em um número de 7 dígitos únicos para cada município, foram inseridos em uma planilha em excel, na qual também foram inseridos os números dos resultados com status 6 e 7 para cada município, sendo preenchido o número “0” nas células referentes aos municípios em que não foram detectados resíduos de qualquer uma dessas substâncias. Essa planilha foi convertida para o formato “.dbf”, necessário para que o software lesse o arquivo, utilizando o software gratuito OpenOffice®.

Os dados do IBGE⁽²⁶⁾ contendo a geolocalização dos municípios foram inseridos no Sistema QuantumGIS® e, em seguida, os dados com o número de detecção de resíduos de cada município também foram incluídos. As duas tabelas foram unidas com a função “juntar” do Sistema e então o layout do mapa foi alterado para permitir a visualização dos municípios em que ocorreram as detecções das substâncias selecionadas conforme Tabela 5.

Metodologia de Análise de Tempo - Gráficos e Regressão Linear

Gráficos de distribuição temporal da análise das substâncias foram elaborados para verificar a existência de padrões de distribuição nesses resultados, o que poderia ajudar a avaliar as razões dessas não conformidades. Esses gráficos foram elaborados com o software Microsoft Excel®, utilizando no banco de dados a quantidade detectada de resíduos das

substâncias em cada semana de análise (Tabela 7).

Tabela 7. Período de análise das substâncias que apresentaram resultados conforme e não conforme

Substância/Espécies	Período da Análise	Número Sequencial das últimas semanas
Closantel em aves	Semana 03, 2002 to 49, 2007	251
Diclazuril em aves	Semana 32, 2008 to 18, 2014	304
Nicarbazina em aves	Semana 30, 2008 to 17, 2014	342
Sulfaquinoxalina em aves	Semana 03, 2002 to 18, 2014	652
Doxiciclina em suínos	Semana 10, 2011 to 28, 2011	131
Sulfametazina em suínos	Semana 08, 2002 to 38, 2014	675

Nos gráficos elaborados nesta etapa também foi possível verificar os casos em que houve extrapolação do Limite Máximo de Resíduos (LMR) para essas substâncias, ou seja, os casos em que o resultado das análises apresentaram status 7.

Resultados

Árvores de Decisão

As árvores extraídas com esta metodologia de análise (Figura 1) mostraram que os dois modelos obtidos apresentaram alta acurácia. A árvore para suínos

(Figura 1a) apresentou apenas 0,22% de erro e a para aves (Figura 1b) 0,09%. O valor do resultado foi utilizado como galhos da árvore, impossibilitando a identificação de padrões nos dados por análise.

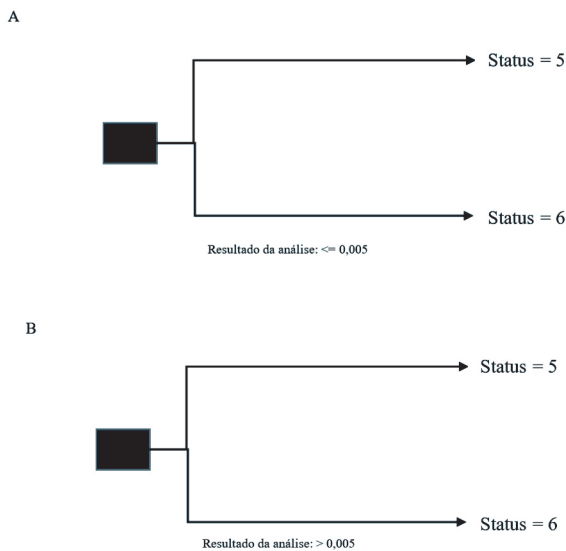


Figura 1. Árvore de decisão da base de dados para suínos (A) e aves (B).

As árvores de decisão também não permitiram a identificação de padrões ou mesmo resultados de análises para a situação de não conformidade. Isso pode ser devido à diferença de LMRs, valores a partir dos quais ocorrem os resultados não conformes, para cada substância analisada, bem como à baixa representatividade desse número em relação ao total de resultados analisados, como pode ser observado na Tabela 8.

Tabela 8. Número e percentual de resultados por status de análise para aves e suínos no Brasil

Espécies	Status da Análise	Nº de resultados	Resultados (%)
Aves	Negativo	47.516	98,84
	Conforme	389	0,81
	Não-conforme	43	0,09
Suínos	Negativo	16.930	99,11
	Conforme	114	0,67
	Não-conforme	37	0,23

Redes Neurais

A base de dados foi avaliada utilizando os neurônios SOM e TASOM para separar dados negativos, conformes e não conformes, conforme Tabela 4, na tentativa de identificar características comuns para resultados com o mesmo status de análise. Na primeira análise de rede neural consideraram-se todos os dados da base de dados pré-processada, ou seja, que apresentaram resultados de análise negativos, conformes e não conformes, conforme Tabela 3, excluindo o rótulo (status da análise). Os algoritmos não separaram os dados do resultado da análise, pois os resultados negativos representaram a maior parte do banco de dados (Tabela 9).

Tabela 9. Porcentagem dos dados agrupados em cada neurônio, usando dados com resultados negativos, conformes e não-conformes

Espécies	Neurônio	SOM			TASOM		
		Negativos	Conformes	Não-Conformes	Negativos	Conformes	Não-conformes
Aves	1	99,54	0,23	0,21	0	0	0
	2	99,45	0,4	0,13	99,4	0,41	0,18
	3	99,67	0,16	0,16	99,54	0,29	0,16
	4	99,46	0,37	0,15	99,52	0,26	0,21
Suínos	1	0	0	0	97,7	1,91	0,38
	2	98,6	1,09	0,3	99,25	0,35	0,38
	3	98,59	0,92	0,48	96,37	3,28	0,24
	4	97,29	2,48	0,22	97,47	2,32	0,21

Essas informações podem ser confirmadas observando que a porcentagem de dados com resultados de análise com Status conforme e não conforme representa apenas aproximadamente 1% do total de amostras (Tabela 8). Na segunda análise, foram considerados apenas os dados de análise que apresentaram resultados conformes e não conformes (Tabela 10).

A terceira análise foi realizada sem os valores dos resultados da análise, o que é um fator determinante para a característica conformidade ou não conformidade dos resultados. Essa análise foi realizada para identificar uma melhor separabilidade dos dados com base em suas demais características. Como resultado, houve uma melhor separabilidade dos dados suínos nos neurônios 1 e 2 do algoritmo SOM e 2 do algoritmo TASOM, mas ainda assim foi observada pouca significância (Tabela 11).

A Figura 2a mostra a separabilidade dos dados de aves por região do mapa 5X5 e a Figura 2b mostra a

mesma informação, mas em mapas 9X9. Os dados de suínos (não apresentados) mostraram separabilidade semelhante aos da Figura 2.

Análise Espacial (QGIS)

A distribuição espacial das detecções de Closantel, Diclazuril, Nicarbazina e Sulfaquinoxalina em aves (Figura 3) e Doxiciclina e Sulfametazina em suínos (Figura 4) mostra distribuição nacional, mas concentração no sul, sudeste e centro-oeste foi elevada.

Isso pode ser explicado pelo fato de a maioria dos estabelecimentos produtores de aves e suínos que possuem Serviço de Inspeção Federal estarem localizados em uma dessas três regiões (Tabela 12).

Tabela 10. Porcentagem de dados agrupados em cada neurônio, usando apenas dados com resultados conformes e não-conformes

Espécies	Neurônio	SOM		TASOM	
		Conforme	Não Conforme	Conforme	Não Conforme
Aves	1	54,09	45,9	0	0
	2	74,28	25,71	43,93	56,06
	3	62,24	37,75	76,68	23,31
	4	60,2	39,79	65,64	34,35
Suínos	1	65,6	24,29	0	0
	2	86,74	13,25	62,42	37,57
	3	91,79	8,2	88,16	11,83
	4	91,88	8,12	95,66	4,33

Tabela 11. Porcentagem de dados agrupados em cada neurônio, usando apenas dados com resultados conformes e não conformes, excluindo o valor do resultado da análise

Espécies	Neurônio	SOM		TASOM	
		Conforme	Não Conforme	Conforme	Não Conforme
Aves	1	63,46	36,53	0	0
	2	0	0	69,28	30,73
	3	0	0	39,29	40,7
	4	0	8,11	0	0
Suínos	1	91,79	8,2	0	0
	2	91,96	8,03	100	0
	3	65,6	34,39	85,46	14,53
	4	89,24	11,03	87,73	12,26

A

		Classe 6 – SOM – 5X5					Classe 7 – SOM – 5X5				
Linha/Coluna		1	2	3	4	5	1	2	3	4	5
1		7	1	30	13	20	4	21	1	5	16
2		39	14	8	12	14	24	17	1	0	9
3		34	14	15	6	9	12	1	1	10	8
4		5	16	32	1	0	10	2	0	0	0
5		9	4	4	0	0	12	0	0	0	25

		Classe 6 – TASOM – 5X5					Classe 7 – TASOM – 5X5				
Linha/Coluna		1	2	3	4	5	1	2	3	4	5
1		0	0	0	0	0	0	0	0	0	0
2		0	43	0	0	0	0	24	0	0	0
3		0	93	36	56	22	0	56	12	40	26
4		7	12	0	3	0	12	0	0	3	0
5		0	27	0	5	0	0	0	0	5	0

B

		Classe 6 – SOM – 9X9									Classe 7 – SOM – 9X9								
Linha/Coluna		1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
1		14	13	1	7	11	12	6	6	6	18	2	25	0	0	12	0	0	0
2		15	20	9	8	2	7	5	1	8	9	5	1	0	0	0	0	0	0
3		10	16	5	0	2	2	1	2	0	5	1	4	0	12	0	0	0	0
4		20	11	3	4	5	1	1	1	0	16	1	1	0	0	12	0	0	0
5		7	3	2	2	0	2	2	3	0	0	0	0	0	0	0	0	0	0
6		4	1	0	0	1	1	0	4	0	0	0	1	0	0	0	0	0	0
7		0	3	2	0	1	1	2	0	0	0	0	9	0	0	0	0	0	0
8		4	1	1	2	0	0	0	0	0	0	21	0	0	0	4	0	0	0
9		4	0	1	1	2	2	0	0	0	12	0	0	0	0	0	0	0	8

		Classe 6 – TASOM – 9X9									Classe 7 – TASOM – 9X9								
Linha/Coluna		1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
1		0	5	56	12	30	12	23	9	38	0	0	16	31	12	1	22	5	2
2		7	0	4	0	0	0	0	25	1	0	0	0	0	0	0	0	8	0
3		1	0	0	1	0	1	1	0	0	25	0	0	0	0	0	0	0	0
4		1	5	7	0	0	6	25	0	0	0	10	12	0	0	0	0	0	0
5		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6		0	1	0	0	0	40	0	0	0	0	0	0	0	0	24	0	0	0
7		0	0	0	9	0	0	0	0	0	0	0	9	0	0	0	0	0	0
8		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figura 2. Mapas 5x5 (A) e 9x9 (B) por status para aves. As diferentes cores representam a dispersão do conjunto de dados nos mapas, utilizando as informações referentes à posição do cluster de cada classe.

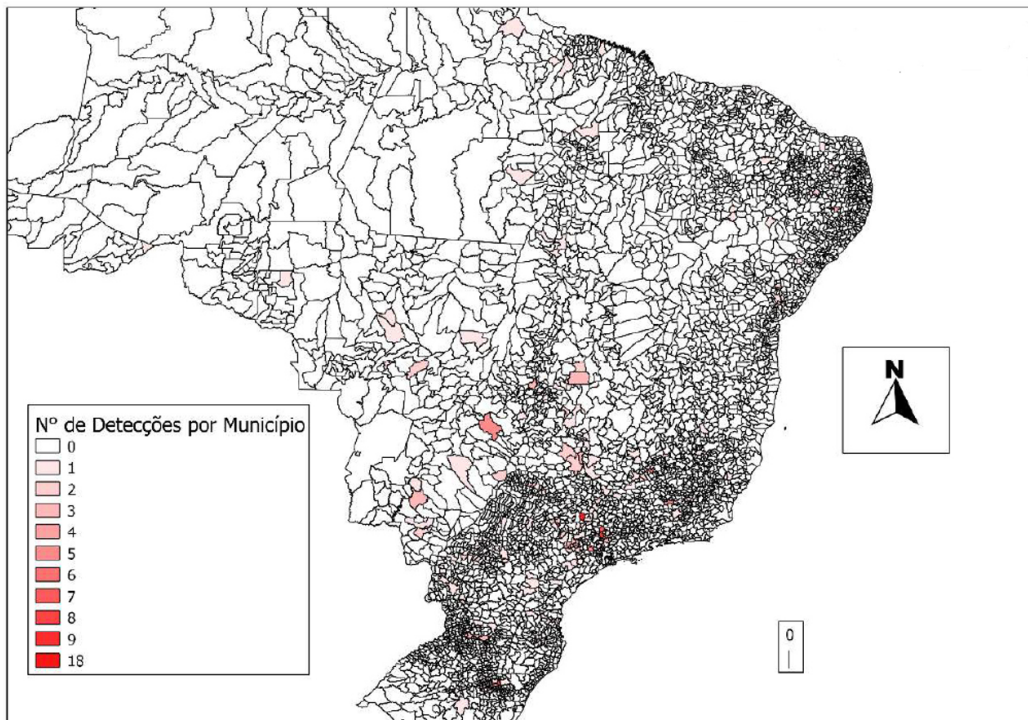


Figura 3. Resíduos de medicamentos veterinários em aves de 2002 a 2014 por Unidade da Federação/município.



Figura 4. Resíduos de medicamentos veterinários em suínos de 2002 a 2014 por Unidade da Federação/município.

Tabela 12. Quantitativo de aves e suínos abatidos em estabelecimento sob Inspeção Federal por Unidade da Federação (UF) entre janeiro de 2002 e outubro de 2014

UF/Estado	Total de Aves	Total de Aves (%)	Total de Suínos	Total de Suínos (%)
Acre	0	0	0	0
Bahia	447.753.015	0,86	437.576	0,14
Distrito Federal	737.931.707	1,42	1.393.434	0,43
Espírito Santo	175.987.005	0,34	39.928	0,01
Goiás	2.943.868.392	5,66	18.247.231	5,66
Maranhão	0	0	4.215	0
Mato Grosso	1.715.812.249	3,3	14.999.395	4,65
Mato Grosso do Sul	1.566.502.066	3,01	10.688.699	3,31
Minas Gerais	3.579.682.963	6,88	30.722.967	9,52
Pará	154.729.386	0,3	0	0
Paraíba	90.138.110	0,17	0	0
Paraná	14.007.224.571	26,94	57.737.369	17,89
Pernambuco	210.458.112	0,4	0	0
Piauí	7.312.661	0,01	0	0
Rio de Janeiro	19.896.591	0,04	0	0
Rio Grande do Norte	0	0	4.756	0
Rio Grande do Sul	8.970.131.584	0,18	76.720.640	23,78
Rondônia	92.790.950	0,18	0	0
Roraima	0	0	15.581	0,01
Santa Catarina	9.844.767.235	18,94	94.567.892	29,31
São Paulo	7.329.522.830	14,1	17.001.370	5,27
Sergipe	6.247.908	0,02	70.638	0,02
Tocantins	102.861.778	0,2	0	0
Total	52.003.619.113	100	322.651.691	0

Não foi observado padrão sazonal de distribuição dos resultados (Figura 5). Para Doxiciclina em suínos e Diclazuril em aves houve concentração de detecção das substâncias apenas em um determinado período: entre as semanas 100 e 150 para Doxiciclina e 200 e 300 para Diclazuril, mas para as demais substâncias houve uma distribuição constante das detecções até um determinado período e após esse período não foram identificadas mais detecções. No entanto, as regressões mostraram um aumento nos níveis quando essas substâncias foram detectadas. Isso pode ser motivo de preocupação e,

portanto, o monitoramento deve continuar.

Quando se analisa o LMR de cada substância (Tabela 8), percebe-se que poucos resultados ultrapassaram o LMR, ou seja, que estavam em uma concentração que poderia ser prejudicial à saúde do consumidor (Tabela 13).

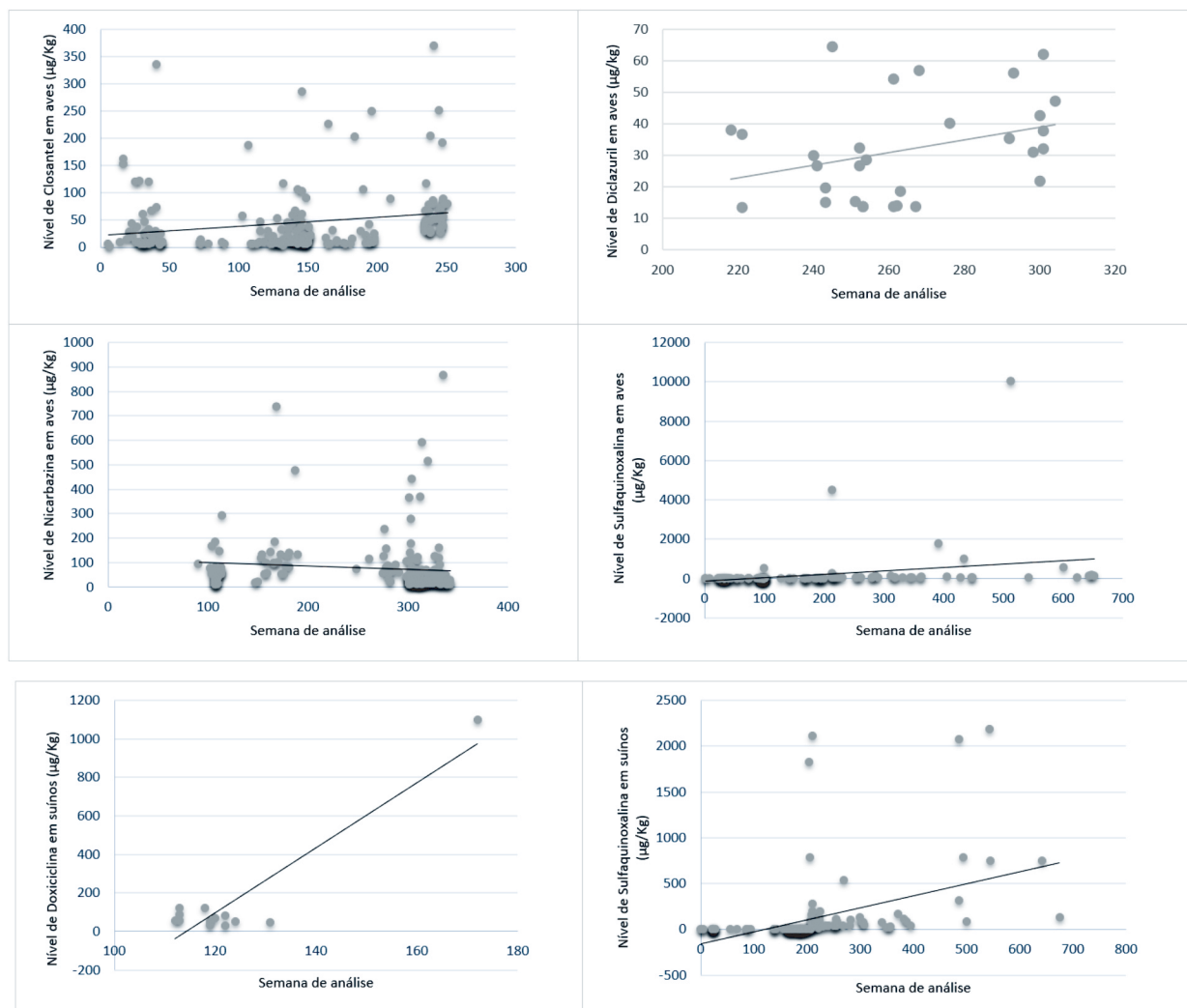


Figura 5. Detecção de substâncias em aves e suínos no período de 2002 a 2014.

Tabela 13. Percentual de análises de resíduos não conformes em aves e suínos

Substâncias/Espécies	Número de amostras analisadas	Número de amostras não conformes	Percentual de análises não conformes (%)
Closantel/Aves	1.938	0	0
Diclazuril/Aves	1.059	0	0
Nicarbazina/Aves	2.904	11	0,37
Sulfaquinoxalina/Aves	5.441	14	0,26
Doxiciclina/Suínos	2.375	1	0,04
Sulfametazina/Suínos	3.182	21	0,66

Discussão

Algoritmos para construção de árvores de decisão estão entre os métodos de aprendizado de máquina mais conhecidos e utilizados⁽²⁷⁾. Isso se deve à sua representação gráfica, o que facilita a compreensão e aplicação dos processos de classificação⁽²⁸⁾.

O sobreajuste é um problema que ocorre no aprendizado de máquina quando o algoritmo trabalha com

muita precisão nos dados de treinamento do modelo, mas não tem boa precisão para os novos dados a serem analisados. O sobreajuste pode ocorrer quando o conjunto de treinamento é muito pequeno ou quando há excesso de dados que não adicionam informações significativas à análise, chamado de ruído. No entanto, também é possível que esse resultado se deva a outro problema recorrente em modelos de aprendizado de máquina, o subajuste, que ocorre quando o modelo não consegue identificar os

padrões ocultos nos dados do conjunto de treinamento por não ser adequado para esse tipo de problema.

As Redes Neurais consistem em algoritmos formados por um conjunto de pequenas unidades de processamento, denominadas neurônios, que fornecem entradas e geram saídas interconectadas, permitindo identificar com mais facilidade as especificidades⁽²⁹⁾.

A rede SOM é uma rede neural que trabalha com aprendizado não supervisionado e funciona apenas em ambientes estáticos onde nenhum dado novo é inserido durante o treinamento. A TASOM trabalha preferencialmente em ambientes incrementais, o que significa que aprende continuamente, à medida que novas entradas entram no Sistema⁽³⁰⁾. As principais características de um banco de dados utilizar algoritmos que operam com aprendizado incremental são: necessidade constante de realizar previsões com os dados, banco de dados evolui ao longo do tempo, banco de dados tem um crescimento infinito, mas os recursos de armazenamento são finitos⁽³¹⁾.

Em geral, os dados de suínos apresentaram maior separabilidade do que os dados de aves, como pode ser observado nos neurônios 2, 3 e 4 para o algoritmo SOM e 3 e 4 para o algoritmo TASOM (Tabela 10). Isso indica que esses neurônios identificaram características comuns entre os dados com resultados de análise semelhantes.

Outra forma de avaliar os resultados da análise por essa metodologia é verificar a dispersão do conjunto de dados em mapas, utilizando as informações relacionadas à posição de cada agrupamento de classes.

Notou-se que não há informações suficientes no conjunto de dados para obter uma boa separação dos dados (Figura 3), considerando que nenhum neurônio foi capaz de classificar os dados de acordo com seu status de análise. Isso pode indicar que os dados analisados possuem uma grande inseparabilidade linear. Em outras palavras, não havia características suficientes nos dados para determinar o status da análise. Isto também pode indicar a presença de características que interferem na análise do conjunto de dados, ou seja, características não necessárias para esta análise que geram ruído.

A maioria dos estabelecimentos produtores de aves e suínos que possuem Serviço de Inspeção Federal estão localizados nas regiões Sul, Sudeste e Centro-Oeste. Além disso, o SISRES distribui aleatoriamente as amostras em que a análise será realizada, mas de acordo com o tamanho da produção de cada estabelecimento frigorífico. Assim, um estabelecimento que abate vários lotes de animais por dia coleta mais amostras para análise de resíduos e contaminantes do que um estabelecimento que abate alguns lotes de animais por semana.

Segundo Mund et al.⁽³²⁾, a avicultura intensiva é comum em muitos países em desenvolvimento e, como os agricultores têm fácil acesso a medicamentos

veterinários, é comum seu uso em doses mais elevadas indiscriminadas e inadequadas. Como visto no presente estudo, este não é o caso no Brasil. Como o Brasil é altamente dependente das exportações agrícolas⁽³³⁾, a existência dessas substâncias na carne de exportação pode prejudicar seriamente o comércio, de modo que os produtores também têm interesse em manter esses níveis baixos. Isso pode ser justificado pelas ações oficiais do MAPA junto aos proprietários dos estabelecimentos de origem desses animais ou pela interrupção da série temporal.

Closantel e Diclazuril estavam abaixo do LMR em todos os casos. Esses resultados estão de acordo com outros autores que encontraram menos de 1% das amostras contendo resíduos não conformes⁽³⁴⁾. Uma incidência de resultados não conformes é necessária para melhor instruir as técnicas de aprendizado de máquina. Outros fatores também podem afetar o sucesso das abordagens de aprendizado de máquina, como ruído no sistema e dados esparsos⁽³⁵⁾, como observado no presente estudo. Segundo Sheppard e Cartwright⁽³⁶⁾, a ausência de dados históricos confiáveis e sistemáticos é um grande obstáculo para as análises de previsão. Esta é uma condição sine qua non para estatística, aprendizado de máquina ou calibração de modelos existentes. O conhecimento desse ruído é necessário para que ele seja removido na etapa de pré-processamento da análise, mas o sistema não forneceu esses dados.

Padrões de distribuição temporal na detecção de resíduos e contaminantes em aves e suínos não foram evidentes, utilizando os dados disponíveis no SISRES até 2014. A maioria das detecções concentra-se nas regiões centro-oeste, sudeste e sul, onde o maior número de suínos e os estabelecimentos produtores de carne de aves estavam concentrados. Os dados deste estudo não foram suficientes para desenvolver uma árvore de decisão capaz de fazer previsões para a avaliação e seleção de substâncias a serem monitoradas oficialmente.

O valor do resultado foi utilizado como galhos da árvore, impossibilitando a identificação de padrões nos dados por análise.

Conclusão

As taxas de contaminação com as seis substâncias aqui estudadas foram muito baixas. Enquanto um padrão espacial de distribuição foi detectado (principalmente devido à maior concentração de animais nas regiões centro-oeste, sudeste e sul), nenhum padrão temporal foi observado. Apesar disso as regressões mostraram um aumento nos níveis quando essas substâncias foram detectadas, portanto, o monitoramento deve continuar. No entanto, os resultados mostram que os produtos monitorados durante o período do estudo apresentaram baixo risco à saúde pública.

Conflito de interesses

Os autores declaram não ter conflito de interesses.

Contribuições do autor

Conceituação: A.J. Portz, N. Silva e C. McManus; *Curadoria de dados:* A.J. Portz e N. Silva; *Análise formal:* A.J. Portz, N. Silva, G. Lima e C. McManus; *Aquisição de financiamento:* C. McManus; *Investigação:* A.J. Portz, N. Silva, G. Lima, L. Feijó, H. Louvandini, V. Peripolli, R. Vieira e C. McManus; *Metodologia:* A.J. Portz, N. Silva, G. Lima, L. Feijó, H. Louvandini, V. Peripolli, R. Vieira e C. McManus; *Gerenciamento do projeto:* A.J. Portz; *Supervisão:* N. Silva e C. McManus; *Redação (revisão e edição):* A.J. Portz, N. Silva, G. Lima, L. Feijó, H. Louvandini, V. Peripolli, R. Vieira e C. McManus.

Agradecimentos

Agradecimentos à CAPES pelo apoio financeiro e ao Ministério da Agricultura do Brasil pelos dados.

Referências

1. FAO. Crops and livestock products [Internet] 2019. [citado 2021 Sep 24]. Disponível em: <https://www.fao.org/faostat/en/#data>.
2. ABPA. Relatório Anual 2021 [Internet] 2021. [citado 2021 Nov 19]. Disponível em: <https://abpa-br.org/mercados/#relatorios>. Português.
3. WHO. Food safety [Internet] 2020. [citado 2022 Apr 29]. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/food-safety>.
4. Kehinde OG, Junaidi K, Mohammed M, AbdulRahman AM. Detection of antimicrobial drug residues in commercial eggs using Premi® Test. *International Journal of Poultry Science*. 2012; 11: 50-54. <https://doi.org/10.3923/ijps.2012.50.54>
5. Ture M, Fentie T, Regassa B. Veterinary Drug Residue: The Risk, Public Health Significance and its Management. *Journal of Dairy & Veterinary Sciences*. 2019; 13(2): 555856. <https://doi.org/10.19080/JDVS.2019.13.555856>
6. Wang J, Yang C, Diao H. Determinants of breeding farmers' safe use of veterinary drugs: a theoretical and empirical analysis. *International Journal of Environmental Research and Public Health*. 2018; 15(10):2185. <https://doi.org/10.3390/ijerph15102185>
7. Miranda RC, Schaffner DW. Virus risk in the food supply chain. *Current Opinion in Food Science*. 2019; 30:43-48. <https://doi.org/10.1016/j.cofs.2018.12.002>
8. Abebe E, Gugsu G, Ahmed M. Review on major Food-Borne zoonotic bacterial pathogens. *Journal of Tropical Medicine*. 2020; 2020: 4674235. <https://doi.org/10.1155/2020/4674235>
9. Painter JA, Hoekstra RM, Ayers T, Tauxe RV, Braden CR, Angulo FJ, Griffin P.M. Attribution of foodborne illnesses, hospitalisations, and deaths to food commodities by using outbreak data, United States, 1998-2008. *Emerging Infectious Diseases*. 2013; 19: 407-415. <https://doi.org/10.3201/eid1903.111866>
10. Dervilly-Pinel G, Guérin T, Minvielle B, Travel A, Normand J, Bourin M, Engel E. Micropollutants and chemical residues in organic and conventional meat. *Food Chemistry*. 2017; 232: 218-228. <http://doi.org/10.1016/j.foodchem.2017.04.013>
11. Guéguen L, Pascal G. An update on the nutritional and health value of organic foods. *Cahiers de Nutrition et de Diététique*. 2010; 45: 130-143.
12. Cariou R, Venisseau A, Amand G, Marchand P, Marcon M, Huneau A, Le Bouquin S. Codex Alimentarius, 2007. CAC/GL 62. Working principles for Risk Analysis for Food Safety for application by governments [Internet] 2007. [citado 2020 Jan 24]. Disponível em: <http://www.fao.org/fao-who-codexalimentarius/standards/list-ofstandards/en/?provide=standards&orderField=fullReference&sort=asc&num1=CAC/GL>.
13. Hummes AP, Bortoluzzi EC, Tonini V, Silva LP, Petry C. Transfer of copper and zinc from soil to grapevine-derived products in young and centenarian vineyards. *Water Air Soil Pollut*, 2019; 230:150 <https://doi.org/10.1007/s11270-019-4198-6>
14. Filazi A, Yurdakok-Dikmen B, Kuzukiran O, Sireli UT. Chemical contaminants in poultry meat and products [Internet] 2017. [citado 2021 Jun 10]. Disponível em: <https://www.intechopen.com/books/poultry-science/chemical-contaminants-in-poultry-meat-and-products>. <http://doi.org/10.5772/64893>
15. Van Loo EJ, Alali W, Ricke SC. Food safety and organic meats. *Annual Review of Food Science and Technology*. 2012; 3: 203-225. <https://doi.org/10.1146/annurev-food-022811-101158>
16. Brasil. Agência Nacional de Vigilância Sanitária. Instrução Normativa nº 51, de 19 de dezembro de 2019. Estabelece a lista de limites máximos de resíduos (LMR), ingestão diária aceitável (IDA) e dose de referência aguda (DRfA) para insumos farmacêuticos ativos (IFA) de medicamentos veterinários em alimentos de origem animal [Internet] 2019. [citado 2022 Apr 28]. Disponível em: <https://www.in.gov.br/en/web/dou/-/instrucao-normativa-n-51-de-19-de-dezembro-de-2019-235414514>. Português.
17. Doyle MP, Erickson MC. Opportunities for mitigating pathogen contamination during on-farm food production. *International Journal of Food Microbiology*. 2012; 152: 54-74. <http://doi.org/10.1016/j.ijfoodmicro.2011.02.037>
18. Liakos KG, Busato P, Moshou D, Pearson S, Bochtis D. Machine Learning in Agriculture: A Review. *Sensors*. 2018; 18: 2674. <https://doi.org/10.3390/s1808267419>
19. Brasil. Ministério da Agricultura Pecuária e Abastecimento. Instrução Normativa nº 42, de 20 de dezembro de 1999. Altera o Plano Nacional de Controle de Resíduos em produtos de origem animal – NCPR e os Programas de Controle de Resíduos em Carne – PCR, Mel – PCRM, Leite – PCRL e Pescado – PCRP e dá outras providências [Internet] 1999b. [citado 2021 Jul 10]. Disponível em: <https://www.gov.br/agricultura/pt-br/assuntos/inspecao/produtos-animais/plano-de-nacional-de-controle-de-residuos-e-contaminantes/documentos-da-pncrc/instrucao-normativa-sda-n-o-42-de-20-de-dezembro-de-1999.pdf/view>. Português.
20. Bittencourt MS, Martins MT, Albuquerque FGS, Barreto F, Hoff R. High-throughput multiclass screening method for antibiotic residue analysis in meat using liquid chromatography-tandem mass spectrometry: a novel minimum sample preparation procedure. *Food Additives & Contaminants: Part A*. 2012; 29(4): 508-516. <http://doi.org/10.1080/19440049.2011.606228>
21. Almeida MP, Rezende CP, Souza LF, Brito RB. Validation of a quantitative and confirmatory method for residue analysis of aminoglycoside antibiotics in poultry, bovine, equine and swine kidney through liquid chromatography-tandem mass spectrometry. *Food Additives & Contaminants: Part A*. 2012, 29(4): 517-525. <http://doi.org/10.1080/19440049.2011.623681>
22. Brasil. Ministério da Agricultura Pecuária e Abastecimento. Guia de validação e controle da qualidade analítica: fármacos

- em produtos para alimentação animal e medicamentos veterinários [Internet] 2011. [citado 2022 Apr 29]. Disponível em: <http://www.agricultura.gov.br/assuntos/laboratorios/arquivos-publicacoes-laboratorio/guia-de-validacao-controle-de-qualidade-analitica.pdf>. Português.
23. Kohonen T. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*. 1982; 43, 59-69. <https://doi.org/10.1007/BF00337288>
24. Shah-Hosseini H, Safabakhsh R. TASOM: a new time adaptive self-organizing map. *IEEE Transactions on Systems, Man, and Cybernetics - Part B*. 2003;33:271-82. <https://doi.org/10.1109/TSMCB.2003.810442>
25. Hermuche PM, Silva NC, Giuimarães RF, Carvalho Jr OA, Gomes RAT, Paiva SR, McManus CM. Dynamics of sheep production in Brazil using principal components and auto-organization features maps. *Revista Brasileira de Cartografia*. 2012; 64, 821-832.
26. IBGE. Redes Geodésicas [Internet] 2014. [citado 2020 Feb 14]. Disponível em: Disponível em: <https://www.ibge.gov.br/geociencias/informacoes-sobre-posicionamento-geodesico/rede-geodesica.html>. Português.
27. Salzberg SL. Book Review: C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993. *Machine Learning*. 1994; 16: 235-240. <https://doi.org/10.1023/A:1022645310020>
28. Stiglic G, Kocbek S, Pernek I, Kokol P. Comprehensive decision tree models in bioinformatics. *Plos One*. 2012; 7:1-14. <https://doi.org/10.1371/journal.pone.0033812>
29. Mohri M, Rostamizadeh A, Talwalkar A. *Foundations of Machine Learning*. 2nd ed. Cambridge: MIT Press; 2018. 427p.
30. He H, Chen S, Li K, Xu X. Incremental Learning From Stream Data. *IEEE Transactions on Neural Networks*. 2011; 22: 1901-1914. <https://doi.org/10.1109/TNN.2011.2171713>
31. Read J, Bifet A, Pfahringer B, Holmes G. Batch-Incremental versus Instance-Incremental Learning in Dynamic and Evolving Data. In: Hollmén J, Klawonn F, Tucker A. *Advances in Intelligent Data Analysis XI*. Berlin: Springer; 2012. pp. 313-323. https://doi.org/10.1007/978-3-642-34156-4_29
32. Mund MD, Khan UH, Tahir U, Mustafa B-E, Fayyaz A. Antimicrobial drug residues in poultry products and implications on public health: A review. *International Journal of Food Properties*. 2017; 20: 1433-1446. <https://doi.org/10.1080/10942912.2016.1212874>
33. Gouvêa R, Santos FF, Aquino MHC, Pereira VLA. Fluoroquinolones in industrial poultry production, bacterial resistance and food residues: a review. *Brazilian Journal of Poultry Science*. 2015; 17: 1-10. <https://doi.org/10.1590/1516-635x17011-10>
34. Cordle MK. USDA regulation of residues in meat and poultry products. *Journal of Animal Science*. 1988; 66, 413-433. <https://doi.org/10.2527/jas1988.662413x>
35. Zhu X, Wu X. Class Noise vs. Attribute Noise: A Quantitative Study. *Artificial Intelligence Review*. 2004; 22: 177-210. <https://doi.org/10.1007/s10462-004-0751-8>
36. Shepperd M, Cartwright M. Predicting with sparse data. *IEEE Transactions on Software Engineering*. 2001; 27: 987-998. <https://doi.org/10.1109/32.965339>