Ciência Animal Brasileira
Brazilian Animal Science
e-ISSN 1809-6891

Ciência Animal Brasileira

DOI: 10.1590/1809-6891v23e-71763E

Section: Veterinary medicine
Research article

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

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

Anna Julia Portz[1], Nilton Silva[1], Gustavo Lima[1], Leandro Feijó[2], Helder Louvandini[3], Vanessa Peripolli[4*], Renata Vieira[1], Concepta McManus[1]

[1]Universidade de Brasília (UnB), Brasília, Distrito Federal, Brasil.
[2]Ministério da Agricultura, Pecuária e Abastecimento, Brasília, Distrito Federal , Brazil.
[3]Centro de Energia Nuclear na Agricultura, Universidade de São Paulo, Piracicaba, São Paulo, Brazil.
[4]Instituto Federal Catarinense (IFC), Campus Araquari, Araquari, Santa Catarina, Brazil.
[*]Correspondent: vanessa.peripolli@hotmail.com

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, Nicarbazin, Sulfaquinoxaline, Doxycycline and Sulfamethazinein). 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. Nevertless, 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

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

Received: April 2, 2022. Accepted: May 10, 2022. Published: June 20, 2022.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Brazil has a prominent role on the international scene as a producer and exporter of agricultural products. In addition to being the fourth largest grain producer of the world, Brazil is also the second-largest grain exporter, with 19% of the international market[1] suppliying more than 180 countries with agricultural products. In the poultry and pig sector this scenario is no different. Brazil exports chicken meat to 151 countries and pork meat to 97 countries[2].

Effective food safety systems are essential to public health and to the confidence of internal consumer market and international consumers. Chemical and microbiological contamination are the leading causes of foodborne diseases[3]. Residues of veterinary drugs and environmental contaminants entering the production chain can cause adverse effects in the human body such as acute toxicity, allergic reactions, disruption of normal intestinal flora, mutagenicity, teratogenicity and carcinogenicity[4,5].These are controlled using good agricultural practices, which mitigate the risk of these substances reaching levels harmful to human health[5,6]. On-farm identification and mitigation strategies need to be evaluated to understanding their impact on reducing animal and human illnesses, as food has been identified as an important vehicle for the transmission of viruses and bacteria[7,8]. Painter et al.[9] estimated more than 9 million foodborne illnesses caused by major pathogens acquired in the United States every year. The same authors attributed most illness to land animal commodities and more deaths to poultry than any other product. Even organic products can show significant levels of environmental contaminants[10].

Environmental contaminants are difficult to control. Open-air production leaves animals potentially more exposed to environmental contaminants such as dioxins, furans and polychlorinated biphenyls (PCBs)[11] while indoor animals can be exposed to flame retardants[12]. Heavy metals, may also constitute risk. Cadmium[11], copper sulphate and zinc[13], as well as arsenic and lead[14] have been found in food products. Feed may contain phytosanitary products and fertilisers, as well as mycotoxins, with associated consequences in the chemical contamination of meat[11,15]. Residues from veterinary drugs may also occur[11].

Brazil has been monitoring residues and contaminants in animal production since 1986, when the National Residue Control Plan (NRCP) was instituted by the Ministry of Agriculture, Livestock and Supply (MAPA). These data are analysed at the end of each year to enable the development of the following year's monitoring plan[16]. The need for quick and assertive decisions in public and private institutions requires the use of decision-making tools that can help in the decision-making process, this need can be met through the use of data mining techniques. Doyle and Erickson[17] showed that computer simulation and Machine Learning (ML) models have been used with greater frequency in the recent years, including agriculture[18]. These authors found that these methods are still incipient in animal production and we did not find any information on the use of machine learning to predict food contamination in poultry and swine. Therefore, this study looks at the use of decision trees, *Self-Organising Map* (SOM) and *Time-Adaptive Self-Organizing Map* (TASOM) neural networks to predict contamination of pig and poultry tissues with six common drugs in Brazil.

Material and methods

Data source – National Residue Control Plan (NRCP)

Data of the monitoring of residues of veterinary drugs and environmental contaminants in poultry and swine under the MAPA National Residue Control Plan (NRCP) was evaluated. The Maximum Residue Limits (MRLs) reference limits used in the NRCP analyses were adopted by MAPA based on the limits determined by the National Health Surveillance Agency (ANVISA)[16] when they exist. For other substances, the limits suggested by Codex Alimentarius[12] were used.

The samples to be collected were determined by a weekly random selection conducted by the Residue System (SISRES), a system that distributes the samples randomly among the establishments registered with the Federal Inspection Service (SIF). Samples were collected in accordance with the instructions in the Sampling Manual of the National Plan for Waste and Contaminants, which consists of updating the collection procedures[19]. The analyses were executed in Official Brazilian Ministry Laboratory Network, using liquid chromatography-tandem mass spectrometry (LC-MS/MS)[20,21] analysis method and the samples were traceable through a Computerized System (SIGLA), which was connected to SISRES. Limits of detection and quantification were estimated for each analyte, in accordance to MAPA guidelines[22].

Data description

Data were used for the detection of residues and contaminants from the NCPR in poultry and swine species from January 2002 to October 2014. These data were downloaded from the SISRES System database, with formal written authorisation from the Coordinator responsible for this System. In the first stage of this study, which consists on the analysis of data by machine learning, results of analysis of all substances analysed in the framework of the NCPR for the poultry and pig chains were used. Each analysis result is correlated to the information that allows the traceability of the samples, and this information refers to the period in which the samples were collected and the place where this collection took place. Thus, were exported from the SISRES database to Excel the data distributed in the categories listed in Table 1.

The data was classified according to Analysis Status (Table 2).

Table 1. Data taken from SISRES in table form

Data Name	Data Description
Code Analysis Type	Substance chemical group code
UF	Federation unit/state
Species Code	Species code sampled for the analysis
SIF	Number registered in the Federal Inspection Service (SIF) of the establishment where the sample was collected
Analysis Week	Week in which the analysis was done
Tissue	Tissue from which the sample was taken: kidney, liver, fat or urine
Owner	Name of the owner of the farm of origin of the animal(s) from which the sample was collected for analysis
Address Owner	Address of the owner of the animal(s) from which the sample was collected for analysis
Farm	Name of the farm of origin of the animal(s) from which the sample was collected for analysis
Address farm	Address of the farm of origin of the animal(s) from which the sample was collected for analysis
Municipality code	Code of the municipality of the farm from which the sample was collected
UF farm	Unit of the federation/state of the farm of the animal(s) from which the sample was collected for analysis
Municpality	Municipality of the farm from which the sample was collected
Number of Animals	Number of animals in the batch from which the sample(s) was (were) taken
Type of violation	Sample 'compliant' or 'non-compliant' according to status 6 or 7 of the analysis
Status	Evaluation result status
CEP	Postal code of the farm where the sample was collected
CEP 2	Owner's address postcode
Year	Year of analysis

Table 2. Status of the Analysis Results of the values taken from the System

Status	Analysis Results
6	Between zero and the MLR of each substance
7	Equal to or greater than the MLR of each substance

MLR: maximum residue limit

Machine Learning - Decision tree

To develop the decision tree, the pre-processing and transformation stage, or data characteristic engineering, was initially performed, which consisted of cleaning and selecting the data to remove data that could generate noise, interfering with the analysis and filling in missing values. This cleaning and selection consisted of the inclusion of a new analysis status, status 5, which corresponds to analysis results equal to 0 and the exclusion of some information. After this change, Table 2 was modified, as shown in Table 3.

Table 3. Representation after the creation of Status 5

Status	Analysis Results	Results
5	Equal to 0	Negative
6	Between 0.1 and the MLR of each substance	Compliant
6	Equal to or greater than the MLR of each substance	Non-compliant

MLR: maximum residue limit

The data excluded were: the confidential data (name and address of the farms where the samples were collected, as well as their owners, the postcode of the owner and unit of the federation of the farm of origin of the animal(s)). Information considered indifferent to the analysis: species code (since the two species were analysed individually), owner code (natural or legal person), tissue code (information unnecessary since each substance is evaluated in only one tissue in the data used in this study); Duplicate data: municipality code (only the name of the municipality was used), violation situation (information replaced by the data status). Missing data was set to 0. After pre-processing the data, the decision tree was prepared with YADT free software.

For data mining, two files were inserted in the software, one with the database for training and the other with the metadata, that is, the description of the data per column, according to Table 4. A test database was not used, being filled in the YADT the value of 25% of use of the training database to perform the test. Two other factors necessary for the creation of the decision tree were also filled in: SPLIT, which is described as the number of new cases to avoid the creation of new branches, and CONFIDENCE, being this the value used for the pruning of the C4.5 algorithm. The values used are the most cited in literature: 2 and 25%, respectively.

Table 4. Metadata description

Data Name	Data Type	Feature Type
Year	String	Discrete
Week	Integer	Discrete
Waste code	Integer	Discrete
Number os animals	Integer	Discrete
SIF	Integer	Discrete
UF	String	Discrete
Municipality	String	Discrete
CEP (post code)	String	Discrete
Result	Float	Continuous
Status	Integer	Class

Neural Network Analysis

The pig and poultry databases were analysed using *Self-Organising Map* (SOM)[23] and *Time-Adaptive Self-Organizing Map* (TASOM)[24] neural networks. A prototype system was used to perform this analysis on Embarcadero®'s C++Builder® online platform. The data was uploaded to this platform in ".csv" format.

The main interface of the software allows the user to configure which variables should be considered[25], the geometric parameters (height and width of the map) and the initial parameters of the SOM training (initial rates of learning and neighbourhood decay rates). The two algorithms were parameterized with the data detailed in Table 5. The samples from the database used already have the label that classifies the analysis result, called Status. The analysis performed consisted of checking the efficiency of the neurons of the algorithms in grouping the samples according to this label. To measure this efficiency, the percentages of the presence of each class in each of the neurons of the networks were recorded. The ideal result is that a neuron has high density of one class and low density of the others, which means that it classified them correctly.

Table 5. Parameters used in the SOM and TASOM algorithms

Parameter	SOM Value	TASOM Value
Map height	2	2
Map width	2	2
Initial Neighborhood Radius	2	2
Initial Learning Rate	1	1
Learning Rate Decay	1	0.01
Neighborhood Radius Decay	1	0.01
Constant Neighborhood Radius[1]	-	1
Constant Learning Rate[2]	-	1
Alpha	-	0.001
Beta	-	0.1

[1]Constant that changes the Neighborhood Radius calculation; [2]Constant that changes the Learning Rate calculation.

To check the data using the neurons, in the center of the prototype there were two tables, one for each neuron, in which the input values, the position of the neuron and the error rate of the data classification were recorded. To run the test, the algorithms (SOM and TASOM) were supplied first with all the data from the pre-processed table. Later, the algorithms were supplied with the data of the pre-processed table, except when status was equal to 5 (result equal to zero) and finally the analysis results were excluded, given directly linked to the analysis status, in order to evaluate the behavior of the algorithms.

Spatial Analysis Methodology – Quantum GIS (QGIS)

In this step of the data analysis, the results per substance were separated in different spreadsheets, for individual analysis of each historical series, and selected only the substances that presented a significant number of analyses with results of Status 6 and 7, ie, with results different from 0. Table 6 represents the substances selected for this step and their MLRs.

Two new information were inserted: name of the substances analysed and week of analysis, which consists of a numerical sequence starting in the first week of the first year and ending in the last week of the last year of the historical series of each substance. The free software QuantumGIS® was used to produce the distribution maps of the analysis results with status 6 and 7. The geolocation data of the Brazilian municipalities were downloaded from the IBGE[26]. The names of Brazilian municipalities and their geocodes, which consists of a 7-digit number, unique for each municipality, were inserted in a spreadsheet in excel, in which the numbers of results with status 6 and 7 for each city were also inserted, with the number "0" being filled in the cells referring to the municipalities in which no residues of any of these substances were detected. This spreadsheet was converted to the ".dbf" format, which is necessary for the software to read the file, using the free OpenOffice® software.

The data from IBGE[26] containing the geolocation of the municipalities were entered into the QuantumGIS® System and then the data with the waste detection number for each municipality were also included. The two tables were joined with the "joining" function of the System and then the layout of the map was changed to allow the visualisation of the municipalities in which detections of the selected substances occurred according to Table 5.

Table 6. Substances and limits selected for analysis

Species	Substance	Action	No. of compliant and Non-Compliant Results	Maximum Residue Limit – MRL (µg/Kg)
Poultry	Closantel	Halogenated salicylanilide with anti parasitic activity	181	1
	Diclazuril	Triazinone antiprotozoal coccidiostat	30	500
	Nicarbazin	Carbanilide coccidiostat	179	200
	Sulfaquinoxaline	Sulfonamide antibiotic and coccidiostat	108	100
Swine	Doxycycline	Tetracycline antibiotic	13	600
	Sulfamethazine	Sulfonamide antibacterial	133	100

Time Analysis Methodology - Graphs and Linear Regression

Time distribution graphs of the substance analysis have been prepared in order to verify the existence of distribution patterns in these results, which could help to assess the reasons for these non-compliances. These graphs were elaborated with Microsoft Excel® software, using in the database the detected amount of residues of the substances in each analysis week (Table 7).

Table 7. Analysis period of the substances that presented results Compliant and Non-Compliant

Substance/Species	Analysis Period	Last Week's Sequential Number
Closantel in poultry	Week 03, 2002 to 49, 2007	251
Diclazuril in poultry	Week 32, 2008 to 18, 2014	304
Nicarbazin in poultry	Week 30, 2008 to 17, 2014	342
Sulphaquinoxaline in poultry	Week 03, 2002 to 18, 2014	652
Doxycycline in swine	Week 10, 2011 to 28, 2011	131
Sulphamethazine in swine	Week 08, 2002 to 38, 2014	675

In the graphs elaborated in this stage, it is also possible to verify the cases in which there was extrapolation of the Maximum Residue Limit (MLR) for these substances, that is, the cases in which the result of the analyses presents status 7.

Results

Decision trees

The trees extracted with this analysis methodology (Figure 1) showed the two models obtained presented high accuracy. The swine tree (Figure 1a) presented only 0.22% of error and the poultry tree (Figure 1b) 0.09%. The result value was used as branches of the tree, making it impossible to identify patterns in the data by analysis.

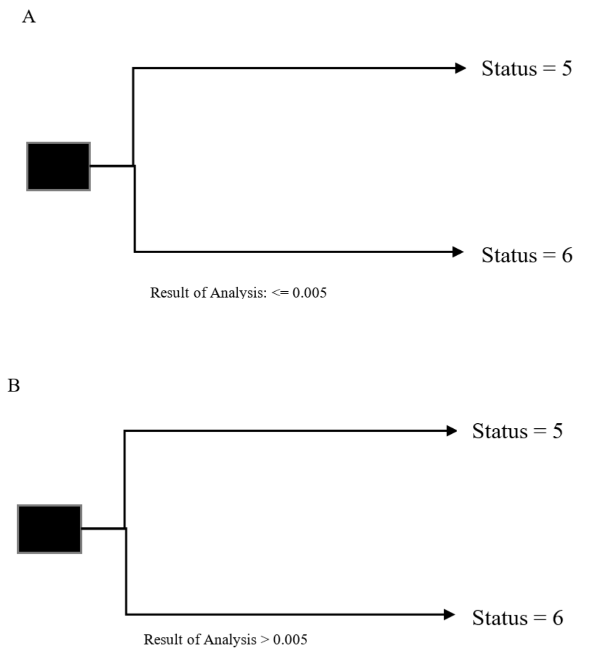

Figure 1. Decision tree for the pig (A) and poultry (B) database.

The decision trees also did not allow the identification of pattern or even analysis results for the non-compliant status. This may be due to the difference in MLRs, values from which non-compliant results occur, for each analysed substance, as well as to the low representativity of this number compared to the total of analysed results, as can be seen in Table 8.

Table 8. Number and percentage of results per analysis status for poultry and swine in Brazil

Species	Analysis Status	N° of results	Results (%)
Poultry	Negative	47,516	98.84
	Compliant	389	0.81
	Non-Compliant	43	0.09
Swine	Negative	16,930	99.11
	Compliant	114	0.67
	Non-Compliant	37	0.23

Neural Networks

The database was evaluated using SOM and TASOM neurons to separate negative, compliant and non-compliant data, according to Table 4, in an attempt to identify common characteristics for results with the same analysis status. The first neural network analysis considered all data from the pre-processed database, i.e., that presented negative, compliant and non-compliant analysis results, according to Table 3, excluding the label (analysis status). The algorithms did not separate the data from the analysis result because the negative results represented the majority of the database (Table 9).

Table 9. Percentage of data grouped in each neuron, using data with negative, compliant and non-compliant results

Species	Neuron	SOM			TASOM		
		Negatives	Compliant	Non-Compliant	Negatives	Compliant	Non-Compliant
Poultry	1	99.54	0.23	0.21	0	0	0
	2	99.45	0.40	0.13	99.40	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
Swine	1	0	0	0	97.70	1.91	0.38
	2	98.60	1.09	0.30	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

This information can be confirmed by noting that the percentage of data with analysis results that are status compliant and non-compliant represents only approximately 1% of the total samples (Table 8). In the second analysis, only the analysis data that presented compliant and non-compliant results were considered (Table 10).

The third analysis was performed without the result values of the analysis, which is a determining factor for the characteristic compliance or non-compliance of results. This analysis was carried out to identify a better separability of the data based on its other characteristics. As a result, there was a better separability of the swine data in neurons 1 and 2 of the SOM algorithm and 2 of the TASOM algorithm, but still little significance was seen (Table 11).

Figure 2a shows the separability of the poultry data by 5X5 map region and Figure 2b shows the same information, but in 9X9 maps. The swine data (not shown) showed separability similar to those in Figure 2.

Table 10. Percentage of data grouped in each neuron, using only data with compliant and non-compliant results

Species	Neuron	SOM		TASOM	
		Compliant	Non- Compliant	Compliant	Non- Compliant
Poultry	1	54.09	45.90	0	0
	2	74.28	25.71	43.93	56.06
	3	62.24	37.75	76.68	23.31
	4	60.20	39.79	65.64	34.35
Swine	1	65.60	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

Table 11. Percentage of data grouped in each neuron, using only data with compliant and non-compliant results excluding the analysis result value

Species	Neuron	SOM		TASOM	
		Compliant	Non-Compliant	Compliant	Non-Compliant
Poultry	1	63.46	36.53	0	0
	2	0	0	69.28	30.73
	3	0	0	39.29	40.70
	4	0	8.11	0	0
Swine	1	91.79	8.2	0	0
	2	91.96	8.03	100	0
	3	65.60	34.39	85.46	14.53
	4	89.24	11.03	87.73	12.26

A

	Class 6 – SOM – 5X5					Class 7 – SOM – 5X5				
Line/Column	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

	Class 6 – TASOM – 5X5					Class 7 – TASOM – 5X5				
Line/Column	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

	Class 6 – SOM – 9X9									Class 7 – SOM – 9X9								
Line/Column	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	0	9	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

	Class 6 – TASOM – 9X9									Class 7 – TASOM – 9X9								
Line/Column	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	0	9	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

Figure 2. 5x5 (A) and 9x9 (B) maps by status for poultry. The different colors represent the dispersion of the dataset on the maps, using the information regarding the position of the cluster of each class.

Spatial analysis (QGIS)

The spatial distribution of the detections of Closantel, Diclazuril, Nicarbazin, and Sulfaquinoxalin in poultry (Figure 3) and Doxycycline and Sulfametazine in swine (Figure 4) showed countrywide distribution, but high concentration in the south, southeast and centerwest. This can be explained by the fact that most of the poultry and swine producing establishments that have Federal Inspection Service are located in one of these three regions (Table 12).

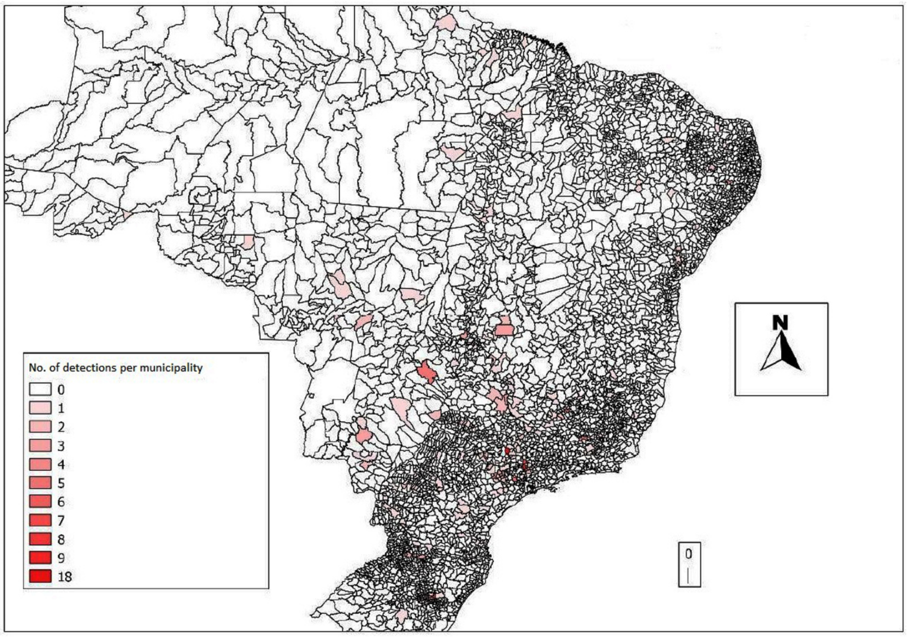

Figure 3. Veterinary drugs residues in poultry from 2002 to 2014 per Federation Unit/municipality.

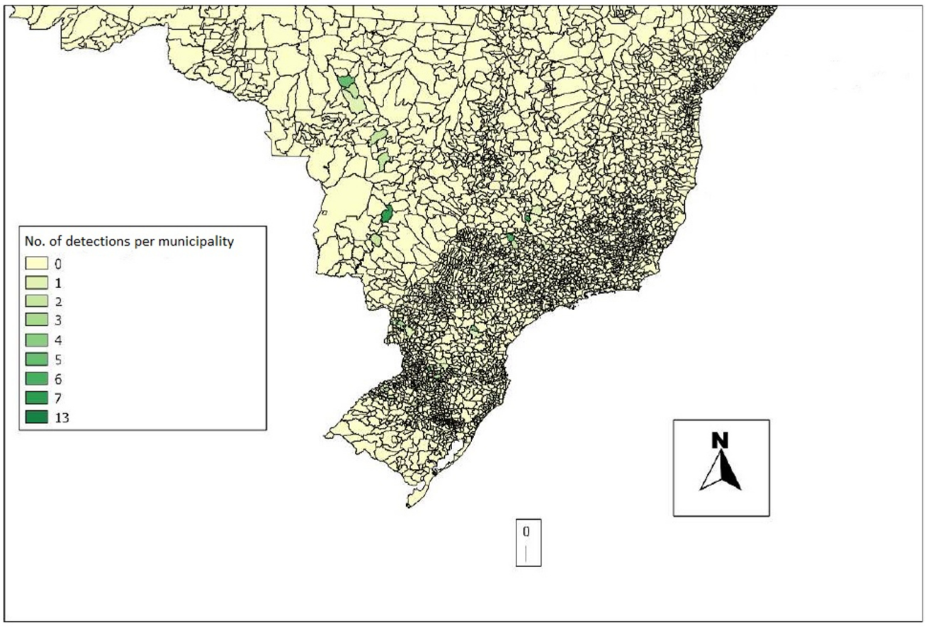

Figure 4. Veterinary drugs residues in swine from 2002 to 2014 per Federation Unit/municipality.

Table 12. Quantitative of poultry and swine slaughtered in an establishment under Federal Inspection per Federation Unit (FU) between January 2002 and October 2014

FU/State	Total Poultry	Total Poultry (%)	Total Swine	Total Swine (%)
Acre	0	0	0	0
Bahia	447,753,015	0.86	437,576	0.14
Distrito Federal	73,7931,707	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	0.00
Mato Grosso	1,715,812,249	3.30	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	0.30	0	0
Paraíba	9,0138,110	0.17	0	0
Paraná	14,007,224,571	26.94	57,737,369	17.89
Pernambuco	210,458,112	0.40	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	0.00
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	14.10	17,001,370	5.27
Sergipe	6,247,908	0.02	70,638	0.02
Tocantins	102,861,778	0.20	0	0
Total	52,003,619,113	100	322,651,691	0

No seasonal pattern of distribution of results was seen (Figure 5). For Doxycycline in swine and Diclazuril in poultry there was a concentration of detection of substances only in a certain period: between weeks 100 and 150 for Doxycycline and 200 and 300 for Diclazuril, but for the other substances there was a constant distribution of the detections until a certain period and after that period no more detections were identified. Nevertheless, regressions showed an increase in levels when these substances were detected. This could be a matter for concern and so monitoring should continue.

When one analyses the MLR of each substance (Table 8), one can see that few results exceeded the MLR, that is, that they were in a concentration that could be harmful to the health of the consumer (Table 13).

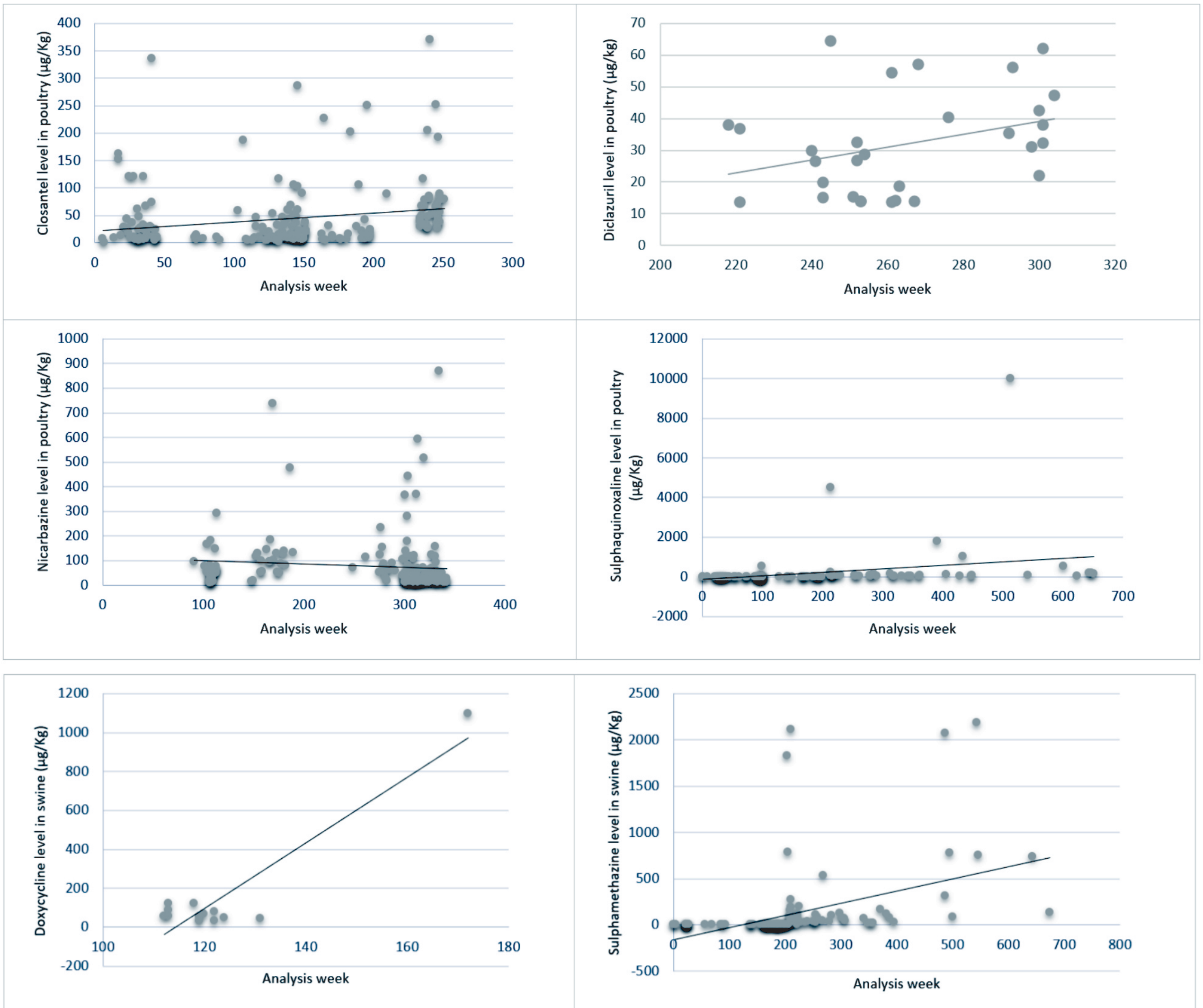

Figure 5. Detection of substances in poultry and swine in the period 2002 to 2014.

Table 13. Percentage of non-compliant residue analysis in poultry and swine

Substance/Species	Number of samples analysed	Number of non-compliant samples	Percentage of non-compliant analyses (%)
Closantel/Poultry	1,938	0	0
Diclazuril/Poultry	1,059	0	0
Nicarbazin/Poultry	2,904	11	0.37
Sulphaquinoxaline/Poultry	5,441	14	0.26
Doxycycline/Swine	2,375	1	0.04
Suphamethazine/Swine	3,182	21	0.66

Discussion

Algorithms for the construction of decision trees are among the best known and used machine learning methods[27]. This is due to their graphic representation, which makes it easier to understand and apply for classification processes[28]. Overfitting is a problem that occurs in machine learning when the algorithm works very accurately on the training data of the model but does not have good accuracy for the new data to be analyzed. Overfitting can occur when the training set is too small or when there is an excess of data that does not add significant information to the analysis, called noise. However, it is also possible that this result is due to another problem that is recurrent in machine learning models, underfitting, which occurs when the model cannot identify the hidden patterns in the training set data because it is not appropriate for that type of problem.

Neural Networks consist of algorithms formed by a set of small processing units, called neurons, which provide inputs and generate inter-connected outputs, allowing them to identify specificities more easily[29]. The SOM network is a neural network that works with unsupervised learning and works only in static environments where no new data is entered during training. TASOM works preferentially in incremental environments, which means it learns continuously, as new inputs enter the System[30]. The main characteristics of a database to use algorithms that operate with incremental learning are: constant need to perform forecasting with the data, database evolves over time, the database has an infinite growth, but the storage resources are finite[31].

In general, the data from pigs showed higher separability than the data from poultry, as can be observed in neurons 2, 3 and 4 for the SOM algorithm and 3 and 4 for the TASOM algorithm (Table 10). This indicates that these neurons identified common characteristics among the data with similar analysis results. Another way to evaluate the analysis results by this methodology is to check the dispersion of the data set in maps, using the information related to the position of each class grouping.

It was seen (Figure 3) that there is not enough information in the data set to obtain a good separation of the data, considering that no neuron was able to classify the data according to its analysis status. This may indicate that the analyzed data has a large linear inseparability. In other words, there were no sufficient features in the data to determine the analysis status. It can also indicate the presence of features that interfere with the analysis of the dataset, i.e. features not required for this analysis that generate noise.

Most of the poultry and swine producing establishments that have Federal Inspection Service are located in south, southeast and centerwest regions. Moreover, SISRES randomly distributes the samples in which the analysis will be carried out, but according to the production size of each slaughter establishment. Thus, an establishment that slaughters several batches of animals per day collects more samples for analysis of residues and contaminants than an establishment that slaughters some batches of animals per week.

According to Mund et al.[32] , intensive poultry farming is common in many developing countries, and as farmers have easy access to veterinary drugs, its use in indiscriminate and inappropriate higher doses is common. As seen here, this is not the case in Brazil. Since Brazil is highly dependent on agricultural exports[33], the existence of these substances in export meat can seriously ham trade, so the farmers also have interest in maintaining these levels low. This may be justified by the official actions of MAPA with the owners of the establishments of origin of these animals or by the interruption of the time series. Closantel and Diclazuril were below MLR in any cases. These results are in line with other authors who found less than 1% of samples contained non-compliant residues[34].

An incidence of non-compliant results is necessary to better instruct the machine learning techniques. Other factors can also affect the success of ML approaches such as noise in the system and sparse data[35], as observed in the present study. According to Sheppard and Cartwright[36], the absence of reliable and systematic historic data is a major obstacle for prediction analyses. This is a sine qua non for statistical, machine learning or calibration of existing models. Knowledge of this noise is necessary for it to be removed in the pre-processing stage of analysis, but the system did not provide this data.

Patterns of temporal distribution in the detection of residues and contaminants in poultry and swine were not evident, using the data available in SISRES until 2014. Most of the detections are concentrated in the centerwest, southeast and southern regions, where the largest number of pork and poultry meat producing establishments were concentrated. The data from this study was not sufficient to develop a decision tree capable of making predictions for the assessment and selection of substances to be officially monitored. The result value was used as branches of the tree, making it impossible to identify patterns in the data by analysis.

Conclusion

Contamination rates with the six substances studied here were very low. While a spatial pattern of distribution was detected (mainly due to the higher concentration of animals in centerwest, southeast and southern regions), no temporal pattern was seen. Nevertless, 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.

Conflict of interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization: A.J. Portz, N. Silva and C. McManus; *Data curation:* A.J. Portz and N. Silva; *Formal analysis:* A.J. Portz, N. Silva, G. Lima and C. McManus; *Funding acquisition:* C. McManus; *Investigation:* A.J. Portz, N. Silva, G. Lima, L. Feijó, H. Louvandini, V. Peripolli, R. Vieira and C. McManus; *Methodology:* A.J. Portz, N. Silva, G. Lima, L. Feijó, H. Louvandini, V. Peripolli, R. Vieira and C. McManus; *Project administration:* A.J. Portz; *Supervision:* N. Silva and C. McManus; *Writing (original draft and review & editing):* A.J. Portz, N. Silva, G. Lima, L. Feijó, H. Louvandini, V. Peripolli, R. Vieira and C. McManus.

Acknowledgements

Thanks are due to CAPES for financial support and the Brazilin Ministry of Agriculture for data.

References

1. FAO. Crops and livestock products [Internet] 2019. [cited 2021 Sep 24]. Available from: https://www.fao.org/faostat/en/#data.

2. ABPA. Relatório Anual 2021 [Internet] 2021. [cited 2021 Nov 19]. Available from: https://abpa-br.org/mercados/#relatorios. Portuguese.

3. WHO. Food safety [Internet] 2020. [cited 2022 Apr 29]. Available from: https://www.who.int/news-room/fact-sheets/detail/food-safety.

4. Kehinde OG, Junaidu 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, Gugsa 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. [cited 2020 Jan 24]. Available from: 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. [cited 2021 Jun 10]. Available from: 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 n° 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. [cited 2022 Apr 28]. Available from: https://www.in.gov.br/en/web/dou/-/instrucao-normativa-n-51-de-19-de-dezembro-de-2019-235414514. Portuguese.

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 n° 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. [cited 2021 Jul 10]. Available from: https://www.gov.br/agricultura/pt-br/assuntos/inspecao/produtos-animal/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. Portuguese.

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 spectrome-

try. 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. [cited 2022 Apr 29]. Available from: http://www.agri cultura.gov.br/assuntos/laboratorios/arquivos-publicacoes-laboratorio/guia-de-validacao-controle-de-qualidade-analitica.pdf. Portuguese.

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. [cited 2020 Feb 14]. Available from:. Available from: https://www.ibge.gov.br/geociencias/informacoes-sobre-posicionamento-geodesico/rede-geodesica.html. Portuguese.

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