ISSN 2410-3438

Volume 8, Issue 22 – January – June - 2021

Journal of Quantitative and Statistical Analysis

ECORFAN®

Journal of Quantitative and Statistical Analysis, Volume 8, Issue 22, January - June 2021, is a journal edited six monthly by ECORFAN-Bolivia. 21 Loa 1179, Cd. Sucre. Chuquisaca, Bolivia. WEB: www.ecorfan.org, revista@ecorfan.org. Chief Editor: MIRANDA - TORRADO, Fernando. PhD. ISSN-On line: 2410-3438. Responsible for the latest update of this number ECORFAN Computer Unit. ESCAMILLA-BOUCHÁN, Imelda, PhD, LUNA-SOTO, Vladimir. PhD. Loa 1179, Cd. Sucre. Chuquisaca, Bolivia, last updated June 30, 2021.

Journal of Quantitative and Statistical Analysis

Definition of Journal

Scientific Objectives

Support the international scientific community in its written production Science, Technology and Innovation in the Field of Physical Sciences Mathematics and Earth sciences, in Subdisciplines of statistical analysis data analysis, multivariate analysis, statistic's and probability, statistical analytical, calculation.

ECORFAN-Mexico SC is a Scientific and Technological Company in contribution to the Human Resource training focused on the continuity in the critical analysis of International Research and is attached to CONACYT-RENIECYT number 1702902, its commitment is to disseminate research and contributions of the International Scientific Community, academic institutions, agencies and entities of the public and private sectors and contribute to the linking of researchers who carry out scientific activities, technological developments and training of specialized human resources with governments, companies and social organizations.

Encourage the interlocution of the International Scientific Community with other Study Centers in Mexico and abroad and promote a wide incorporation of academics, specialists and researchers to the publication in Science Structures of Autonomous Universities - State Public Universities - Federal IES - Polytechnic Universities - Technological Universities - Federal Technological Institutes - Normal Schools - Decentralized Technological Institutes - Intercultural Universities - S & T Councils - CONACYT Research Centers.

Scope, Coverage and Audience

Journal of Quantitative and Statistical Analysis is a Journal edited by ECORFAN-Mexico S.C in its Holding with repository in Bolivia, is a scientific publication arbitrated and indexed with semester periods. It supports a wide range of contents that are evaluated by academic peers by the Double-Blind method, around subjects related to the theory and practice of statistical analysis data analysis, multivariate analysis, statistic's and probability, statistical analytical, calculation with diverse approaches and perspectives, that contribute to the diffusion of the development of Science Technology and Innovation that allow the arguments related to the decision making and influence in the formulation of international policies in the Field of Physical Sciences Mathematics and Earth sciences. The editorial horizon of ECORFAN-Mexico® extends beyond the academy and integrates other segments of research and analysis outside the scope, as long as they meet the requirements of rigorous argumentative and scientific, as well as addressing issues of general and current interest of the International Scientific Society.

Editorial Board

Arbitration Committee

Assignment of Rights

The sending of an Article to Journal of Quantitative and Statistical Analysis emanates the commitment of the author not to submit it simultaneously to the consideration of other series publications for it must complement the Originality Format for its Article.

The authors sign the Authorization Format for their Article to be disseminated by means that ECORFAN-Mexico, S.C. In its Holding Bolivia considers pertinent for disclosure and diffusion of its Article its Rights of Work.

Declaration of Authorship

Indicate the Name of Author and Coauthors at most in the participation of the Article and indicate in extensive the Institutional Affiliation indicating the Department.

Identify the Name of Author and Coauthors at most with the CVU Scholarship Number-PNPC or SNI-CONACYT- Indicating the Researcher Level and their Google Scholar Profile to verify their Citation Level and H index.

Identify the Name of Author and Coauthors at most in the Science and Technology Profiles widely accepted by the International Scientific Community ORC ID - Researcher ID Thomson - arXiv Author ID - PubMed Author ID - Open ID respectively.

Indicate the contact for correspondence to the Author (Mail and Telephone) and indicate the Researcher who contributes as the first Author of the Article.

Plagiarism Detection

All Articles will be tested by plagiarism software PLAGSCAN if a plagiarism level is detected Positive will not be sent to arbitration and will be rescinded of the reception of the Article notifying the Authors responsible, claiming that academic plagiarism is criminalized in the Penal Code.

Arbitration Process

All Articles will be evaluated by academic peers by the Double Blind method, the Arbitration Approval is a requirement for the Editorial Board to make a final decision that will be final in all cases. MARVID® is a derivative brand of ECORFAN® specialized in providing the expert evaluators all of them with Doctorate degree and distinction of International Researchers in the respective Councils of Science and Technology the counterpart of CONACYT for the chapters of America-Europe-Asia- Africa and Oceania. The identification of the authorship should only appear on a first removable page, in order to ensure that the Arbitration process is anonymous and covers the following stages: Identification of the Journal with its author occupation rate - Identification of Authors and Coauthors - Detection of plagiarism PLAGSCAN - Review of Formats of Authorization and Originality-Allocation to the Editorial Board-Allocation of the pair of Expert Arbitrators-Notification of Arbitration -Declaration of observations to the Author-Verification of Article Modified for Editing-Publication.

Instructions for Scientific, Technological and Innovation Publication

Knowledge Area

The works must be unpublished and refer to topics of statistical analysis data analysis, multivariate analysis, statistic's and probability, statistical analytical, calculation and other topics related to Physical Sciences Mathematics and Earth sciences.

Presentation of the content

In the first article we present, *Statistical evolution of clinical laboratories in patients admitted to the CECCS with SARS-COV-2*, by YAÑEZ-VARGAS Israel, PRIETO-GARCÍA Daniela, DOÑATE-ÁLVAREZ Andrea and QUINTANILLA-DOMINGEZ Joel, with adscription in the Universidad Politécnica de Juventino Rosas, in the next article we present, *Analysis of multi-class classification performance metrics for remote sensing imagery imbalanced datasets*, by GONZALEZ-RAMIREZ, Andrea, LOPEZ, Josue, TORRES-ROMAN, Deni and YAÑEZ-VARGAS, Israel, with adscription in the CINVESTAV del IPN, Unidad Guadalajara and Universidad Politécnica de Juventino Rosas, in the next article we present, *Productive performance of York x Landrace sows in a semi-technified farm*, SANCHEZ-CHIPRES, David Román, MORENO-LLAMAS, Gabriel, JIMÉNEZ-PLASCENCIA Cecilia and JIMÉNEZ-CORDERO, Ángel Andrés, with adscription in the Universidad de Guadalajara, in the next article we present, *Use of the effectiveness of cuachalalate from healing plants in Mezquitic, Jalisco*, by GONZÁLEZ-GARCÍA, Arcelia, HERNÁNDEZ-SALAS, Claudia, MARTÍNEZ-ORTIZ, Rosa María and GONZÁLEZ-MARTÍNEZ, Lilia, with adscription in the Universidad Autónoma de Zacatecas and Unidad Académica de Odontología. Campus UAZ.

Content

Article **Journal of Quantitative and Statistical Analysis**
June, 2021 Vol.8 No.22 1-10

Statistical evolution of clinical laboratories in patients admitted to the CECCS with SARS-COV-2

Evolución estadística de laboratorios clínicos en pacientes ingresados en el CECCS con SARS-COV-2

YAÑEZ-VARGAS Israel†*, PRIETO-GARCÍA Daniela, DOÑATE-ÁLVAREZ Andrea and QUINTANILLA-DOMINGEZ Joel

Universidad Politécnica de Juventino Rosas, Departamento de Ingeniería en Redes y Telecomunicaciones.

ID 1st Author: *Juan Israel Yañez Vargas* / **ORC ID:** 0000-0001-5749-8442, **CVU CONACYT ID:** 295711

ID 1er Coauthor: *Daniela Prieto García* / **ORC ID:** 0000-0002-1228-7163

ID 1st Coauthor: *Andrea Abigail Doñate Álvarez* / **ORC ID:** 0000-0003-4005-8533

ID 2nd Coauthor: *Joel Quintanilla Domínguez*/ **ORC ID** 0000-0003-2442-2032, **CVU CONACYT ID:** 165237

DOI: 10.35429/JQSA.2021.22.8.1.10 Received March 14, 2021; Accepted June 29, 2021

Abstract

The evolution of the SARS-CoV-2 Virus in recent months in the world and especially in Mexico has also created changes in the patient's treatment and their laboratory studies, as is the case of patients admitted to the CECCS in Salamanca, Guanajuato, in which after months of caring for COVID-19 patients, new relevant laboratory studies were added (Ferritin and D-Dimer), therefore updates were made in statistical analyzes, and day-to-day behavior. The study is based on updating and cleaning the database, taking a sample of more patients with COVID-19, statistical analysis including admission tables, Body Mass Index, Variance, Normalization, relationship between the first day and the last day of patients who died and lived. The main idea of the work is to learn with more detail those laboratories that contain information to study in the future if there is a relationship between any medication supplied with the improvement or aggravation of the patient.

SARS-COV-2, Statistics evolution, Clinical laboratories

Resumen

La evolución del Virus SARS-CoV-2 en los últimos meses en el mundo y especialmente en México ha creado también cambios en el tratamiento de los pacientes y sus estudios de laboratorio, como es el caso de los pacientes ingresados en el CECCS de Salamanca, Guanajuato, en el que después de meses de atención a pacientes COVID-19 se agregaron nuevos estudios de laboratorio relaventes (Ferritina y Dimero D), por lo anterior se realizaron actualizaciones en análisis estadísticosm, y el comportamiento del día a día. El estudio se basa en la actualización y limpieza de la base de datos, toma de muestra de más pacientes con COVID-19, análisis estadístico incluyendo tablas de ingreso, Índice de Masa Corporal, Varianza, Normalización, relación entre el primer día y último de los pacientes que fallecieron y vivieron. La idea principal del trabajo es conocer más a detalle aquellos laboratorios que contengan mayor información para estudiar a futuro si existe relación entre algún medicamento suministrado con la mejora o agravamiento del paciente.

SARS-CoV-2, Evolución estadística, Laboratorios clínicos

Citation: YAÑEZ-VARGAS Israel, PRIETO-GARCÍA Daniela, DOÑATE-ÁLVAREZ Andrea and QUINTANILLA-DOMINGEZ Joel. Statistical evolution of clinical laboratories in patients admitted to the CECCS with SARS-COV-2. Journal of Quantitative and Statistical Analysis. 2021. 8-22: 1-10

* Correspondence to Author (e-mail: jyanez_ptc@upjr.edu.mx)
† Researcher contributing as first author.

1. Introduction

More than a year has passed in which the country of Mexico started a mandatory quarantine due to the pandemic that still continues to affect the world, the pandemic that originated at the end of 2019 in China and is called Severe Acute Respiratory Syndrome (SARS-CoV-2) or also known as COVID-19.

The increase in cases and deaths has meant a great struggle to be able to contain the virus, for which great efforts have been made to obtain the best vaccines, the best medications, and know the behavior or evolution of patients, despite all the above , the problems derived from COVID-19 and how it affects the body (Coronavirus disease (COVID-19)) are still not fully understood.

The Ministry of Health of Mexico in its daily press conference made a cut on COVID-19 in which they show a figure of 2 million 397 thousand 307 infected, likewise 221 thousand 695 deaths confirmed by the virus (Coronavirus - gob.mx) . For the case study to be carried out, it was determined to extract information from the State Center for Critical Care of Salamanca (CECCS), being a third-level hospital that only receives patients with COVID-19 from the state of Guanajuato and that has the peculiarity of receiving serious patients .

The information that will be extracted from the CECCS will be of great relevance for conducting multiple studies on patients who have or have had COVID-19, since it is still complex to know its behavior in humans ("Coronavirus Guanajuato", 2020). Derived from the above, multiple studies and research of the disease have been carried out through X-ray image analysis or tomography, as well as statistical analyzes that seek a greater understanding of the disease, as is the case of some relevant ones that are mentioned below. continuation.

According to the study (Gao et al., 2020) performs an analysis of laboratory data for a differential diagnosis of COVID-19 patients, the study was for 28 seriously ill adult patients and 19 laboratories for each patient including Dimero-D and Ferritin.

Likewise, Yun et al. (2020, 95 p) in his article he performs a statistical study of clinical laboratories of 2510 patients, in which he describes the tests carried out influenza type A and influenza type B, graphs of the studies carried out and comparisons with others that were performed are shown. have done blood studies. In the study by Rodriguez-Morales et al. (2020) a meta-analysis of multiple research articles on COVID-19 throughout the world is carried out, in which they compare laboratory studies of confirmed patients and X-ray images and Tomographs, the idea is to calculate the prevalences, performing qualitative and quantitative studies of the articles reviewed.

In the article by Liu et al. (2020, p.372) a study was carried out on 12 patients hospitalized with COVID-19, the research aims to show the changes in patients based on their laboratories, medications supplied and X-ray and tomography studies.

Previously, a research article was carried out (Yañez-Vargas, Doñate-Álvarez, Quintanilla-Domínguez, & Aguilera-González, 2020), in which the first statistical studies of the patients admitted to the CECCS were carried out, having studies of the patients with 28 clinical laboratories. The day-to-day study of the patients was carried out with the most outstanding graphics of the patients who died and lived.

Previous investigations carry out statistical studies of the disease but do not have more than 20 laboratory studies, while the present work aims to carry out the study by expanding the laboratory data and focusing on a sample of critically ill patients from the state of Guanajuato.

2. Material and methods

The patients who are admitted to the CECCS must come from second level health institutions, which carry out studies and previous diagnoses to know the severity of each of the patients, so if with the passage of time the patients begin to worsen they are carried out transfers to the CECCS, where upon entering the institution, as a first step, different vital signs and relevant information such as age, weight and previous illnesses must be taken.

ISSN 2410-3438

The information extracted from the CECCS corresponds to 109 patients, who were admitted from March 24 to October 20, 2020, adding that the first patient in the state with symptoms of COVID-19 was admitted to said center.

Based on the information that has been extracted, a study was carried out to find out statistics of the information of the admitted patients, due to the above, the table of total patients who were admitted in the aforementioned time is shown in Table 1, in which A breakdown of how many have died and how many have been discharged is shown, it is important to mention that the CECCS is a third-level center in which the most serious patients arrive in the state, which is why it has a mortality rate higher.

Total patients classification					
	Patients	Deaths	Patient was discharged		Hospitalized
			Counter reference	Reference	
Covid 19	88	40	28	13	7
Negatives	18	8	8	2	0
Pending	3	2	1	0	0
Total	109	50	37	15	7
All the female discharged Totals = 52					

Table 1 Total income in the CECCS
Source: Own

Likewise, Table 2 shows the classification of men who have been admitted at the same time of analysis, you can see the number of patients with COVID-19, those who lived and those who were discharged.

Male patients classification					
	Patients	Deaths	Patient was discharged		Hospitalized
			Counter reference	Reference	
Covid 19	60	27	18	10	5
Negatives	11	6	4	1	0
Pending	2	2	0	0	0
Total	73	35	22	11	5
All the female discharged Totals = 33					

Table 2 Income of men to the CECCS
Source: Own

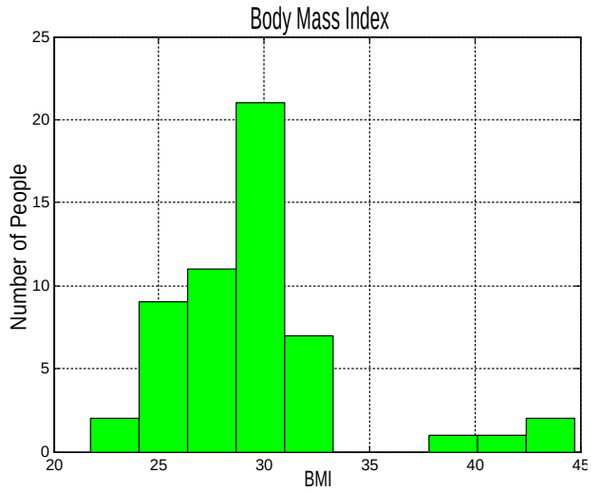
Table 3 shows the classification of women who have been admitted at the same time of analysis, you can see the number of patients with COVID-19, those who lived and those who were discharged, as explained in the article (Yañez-Vargas et al., 2020), which shows the trend of resistance of women to COVID-19.

Female Patients classification					
	Patients	Deaths	Patient was discharged		Hospitalized
			Counter reference	Reference	
Covid 19	28	13	10	3	2
Negatives	7	2	4	1	0
Pending	1	0	1	0	0
Total	36	15	15	4	2
All the female discharged Totals = 19					

Table 3 Income of women to CECCS
Source: Own

According to the fact that when patients are admitted to the CECCS, relevant information such as weight, age and height are collected, a study of the Body Mass Index (BMI) has been carried out, with the intention of obtaining relevant information and statistics from all patients, which can be used for future knowledge processes of the behavior of the virus in humans.

In Graph 1 it is possible to see the general BMI, it is important to mention that the normal value of BMI is between 19 and 24, so very high values are shown in most patients, reaching cases of grade III obesity, which is makes it relevant in the COVID-19 study.



Graph 1 BMI in CECCS patients
Source: Own

3. Clinical laboratories

Unlike the article (Yañez-Vargas et al., 2020) in which 29 laboratories were used, now we work with 34 laboratories that are mentioned below: Glucose, Urea, Ureic Nitrogen, Creatinine, Total Proteins, Albumin, Globulin, A / G Ratio, Alanine Amino Transferase, Aspartate Amino Transferase, Total Bilirubin, Direct Bilirubin, Indirect Bilirubin, Lactic Dehydrogenase, Creatine Phosphokinase, Chlorine, Potassium, Sodium, Hematic Cytometry, Prothrombin Time, Thromboplastin Time,

ISSN 2410-3438

Procalithin, Fibrin Calcium, Phosphorus, Alkaline Phosphate, Magnesium, Protein C. Reactive and Total Cholesterol, HDL Cholesterol, Sedimentation Rate were added, in addition 2 that have gained great relevance in recent months and are Dimero D and Ferritin (Pagana & Pagana, 2015).

4. Methodology

For the development of this statistical study, it is necessary to propose a series of steps that are described in the block diagram of Figure 1 that describes all the steps carried out in the investigation.

The first block corresponds to obtaining / extracting the clinical laboratories of the patients admitted to the CECCS, the second step is the analysis of the information and how it will be its subsequent classification and division (men / women) which is expressed In the third block, having the information classified, an organization of the information is carried out and a subsequent normalization of the data that is shown in blocks 4 and 5, it is important to highlight that the normalization has been carried out based on the maximum values and minimums of each laboratory, taking into account that the normalization will be applied to each laboratory separately, likewise in the organization of the data, the use of the mean and mode has been proposed to complete missing information from laboratories (Gonzalez & Woods, 2008).

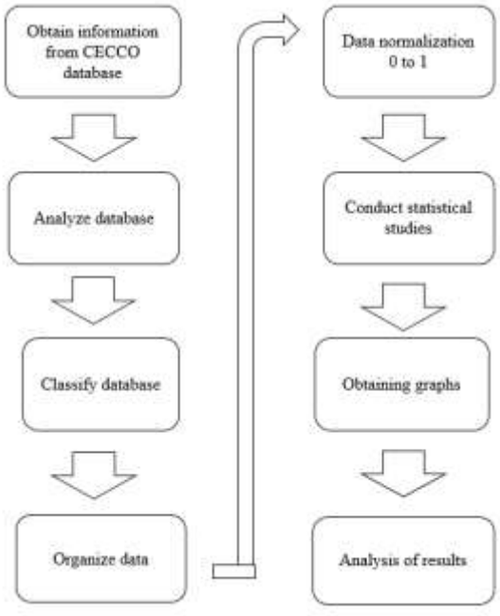


Figure 1 General block diagram
Source: Own

5. Results

Block 6 represents the statistical study of data from clinical laboratories, a study of variance of the data is required to analyze the relevance of the information, as well as day-to-day studies of multiple laboratories, similarity studies between patients and their behaviors during their stay at the CECCS to finally obtain the multiple graphs required for subsequent analysis and conclusions, all under the supervision of a doctor who is working in the hospital (Devore, 2012), (Johnson et al. , 2012).

Number	Laboratory	Variance
1	Ferritin	83395211.83
2	D- Dimer	4083609.155
3	C Reactive Protein	3857759.202
4	Creatin-Fosfocinasa	793769.7153
5	Fibrinogen	115711.6071
6	Lactic Dehydrogenase	33699.79234
7	General Urine Test	27041.14802
8	Glucose mg/dl	7122.691524
9	Alkaline phosphatase	3360.290893
10	Urea mg/dl	2986.325801
11	Chlorine	2088.983047
12	Aspartate aminotransferase	1434.218558
13	Alanine aminotransferase	1293.928146
14	Total Cholesterol	1041.130814
15	Ureic Nitrogen	532.8210967
16	Albumin	314.4155434
17	Sedimentation rate	196.0301266
18	Potassium	176.5350151
19	HDL Cholesterol	101.3912015
20	Sodium	97.37544607
21	Partial thromboplastin time	37.8615637
22	Creatinine mg/dl	35.65593406
23	Hematic Cytometry	21.8262548
24	Indirect bilirubin	19.89173877
25	Prothrombin time	15.65751113
26	Procalcitonin	9.634289823
27	Phosphorus	2.690131457
28	Calcium	2.564005621
29	Total Bilirubin	1.04556895
30	Total Proteins	0.92483412
31	Magnesium	0.568098256
32	Relation A/G	0.508733212
33	Globulin	0.431637386
34	Direct Bilirubin	0.331210161

Table 4 Variance study of clinical laboratories
Source: Own

In the area of medicine, statistical analysis of the behavior of patients in the face of all diseases, symptoms, surgeries and medications is of great relevance, and if at the same time graphs are made that interpret all the processed information, it will be easier to know the COVID -19 and its behavior in a sample of patients from the state of Guanajuato, which also, as an important point, the sample is also taken from critically ill patients with tendencies to die due to their comorbidities, genetics and lifestyle.

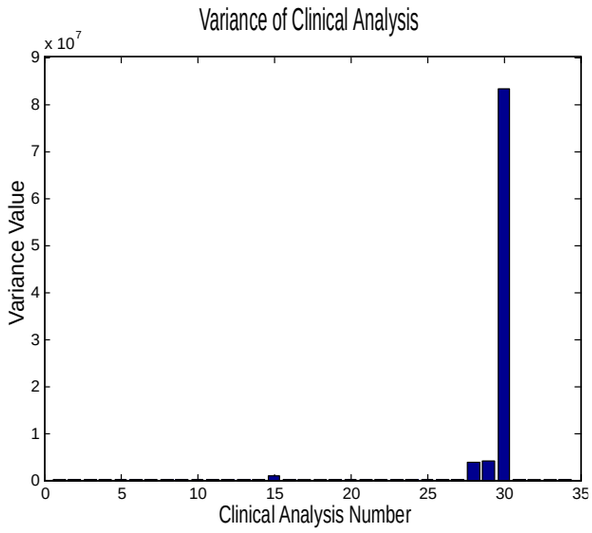
ISSN 2410-3438

Table 1 shows an analysis of variance that has been carried out to the clinical laboratories of patients admitted to the CECCS, it is important to comment that the study of variance will help to measure the greater or lesser dispersion of information or data values Regarding the arithmetic mean that the same data have, the study has been carried out for each of the laboratories. The Table was accommodated from greater variance to less variance, observing that two of the new laboratories added (Ferritina and Dimero D) in the investigation are the ones with the greatest dispersion, this becomes relevant because large changes have been identified in those laboratories between patients who live and die, while the laboratories with the lowest variance correspond to Globulin and Direct Bilirubin.

The study is a part of multiple analyzes that will be used to know what information / laboratories can be relevant as inputs in classifiers and thus be able to reduce errors or computational complexity for the design of artificial neural network architectures.

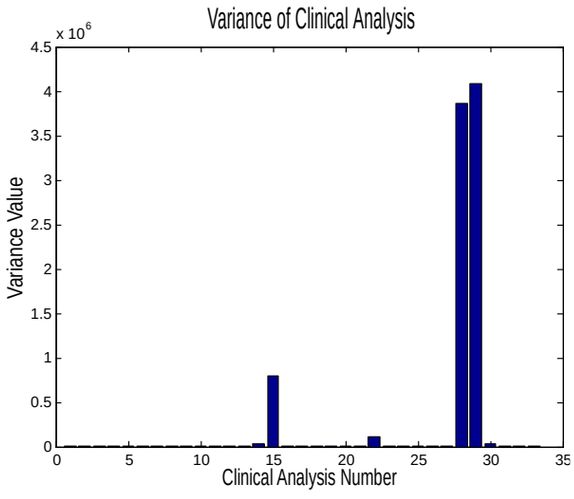
At the same time that Table 1 has been made, several graphs were created that represent the values obtained from the variances of each laboratory, in Graph 2 it is possible to observe the 34 clinical laboratories on the X axis, the order is according to the taking of analysis carried out by the doctors and nurses of the CECCS, while the value obtained from the variance is observed on the Y axis. According to the graph, it is widely noted that laboratory 30 that corresponds to Ferritin is the one with the greatest variance, followed by laboratory 29, which is Dimero D, laboratory 28 continues, which is C-Reactive Protein and finally the creatine phosphosinase that is represented in the lab 15 from the same graph.

In the case of Graph 3, the Ferritin laboratory was eliminated to see the values of the other laboratories, since there are extra tests that can be used for a subsequent classification through some artificial intelligence method. Given the above, the relevance of lactic dehydrogenase, Procalcitotin and general urine examination are visualized in the graph.



Graph 2 Variance graph of clinical laboratories
Source: Own [Matlab]

The analysis carried out previously shows the changes proposed at the international level in which it was proposed to add laboratory studies of Dimero D and Ferritin to patients with COVID-19, observing their relevance through the first study of variance and will serve as a basis for future research in critical patients in the state of Guanajuato.

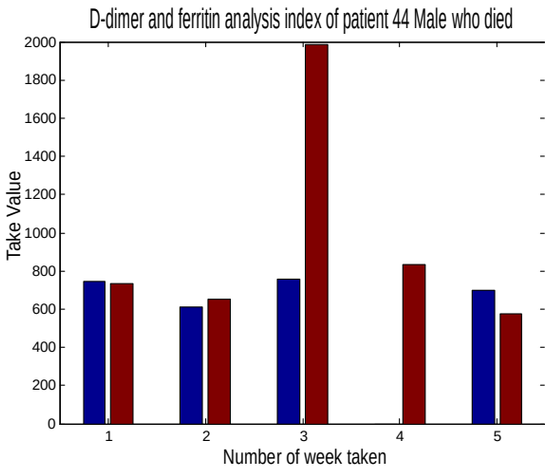


Graph 3 Variance graph of clinical laboratories without Ferritin
Source: Own [Matlab]

The second section of statistical analysis of clinical laboratory data of patients with COVID-19 focused on finding the relevance of the two new laboratories and their behavior in patients admitted to the CECCS.

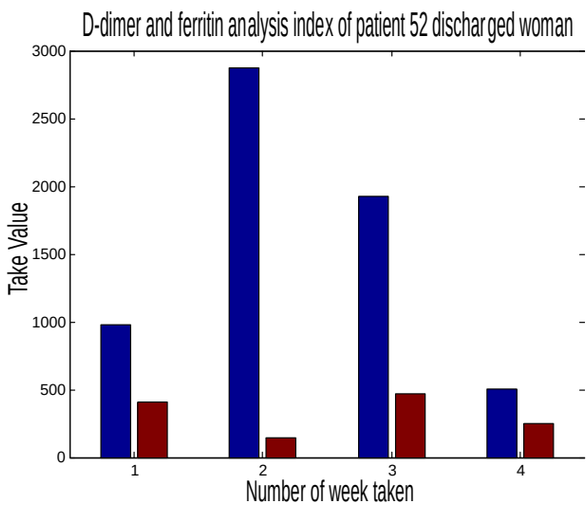
ISSN 2410-3438

As an important note, it should be mentioned that both laboratories are obtained on average every 5 or 6 days, so in patients who were hospitalized for a short time, there are not so many samples, so those patients who were admitted for around 40 days are those of which a study could be carried out in a better way.



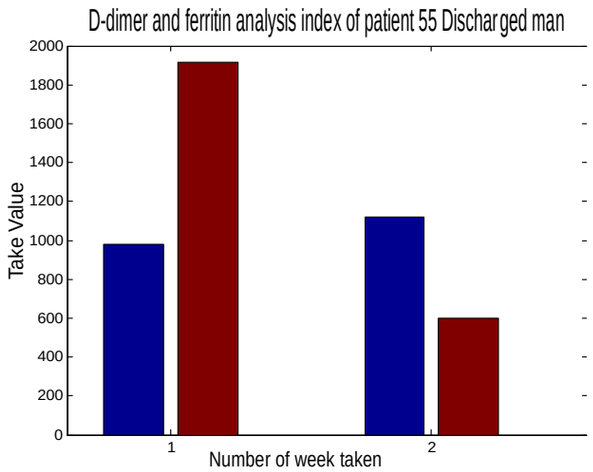
Graph 4 Dimero D and Ferritin analysis in patient 44
Source: Own [Matlab]

Graph 4 represents patient number 44 who had a stay of 22 days, the patient was taken 4 samples of Dimero D and is represented with the blue color in the bar graph, while 6 samples of Ferritin were taken represented in red color throughout his stay, observing that Dimero D remained with slight changes, while Ferritin reached very high values that caused decompensation of the patient. Patient 44 is a man who unfortunately died, the graph shows that the changes shown in Ferritin are important signs in the relapses of patients.



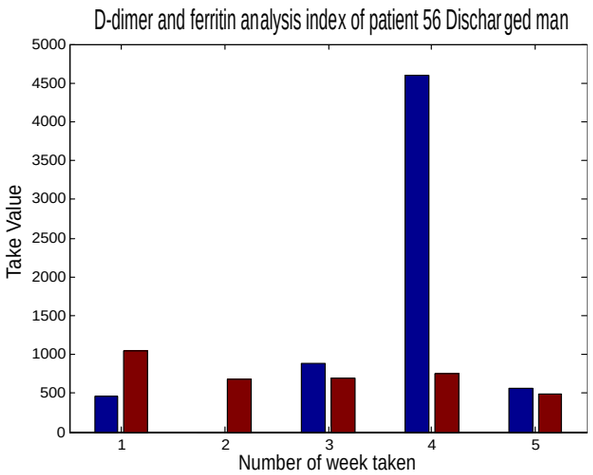
Graph 5 Dimero D and Ferritin analysis in patient 52
Source: Own [Matlab]

For Graph 5 represents patient number 52 who had a stay of 42 days, the patient was taken 4 samples of Dimero D and is represented with the blue color in the bar graph, while 4 samples of Ferritin were taken represented in red color throughout its stay, observing that Dimero D remained with high values in the second sample, reaching a considerable decrease in the last sample in Dimero D and Ferritina. Patient 52 is a woman who was discharged, showing that keeping both laboratories low helps patients, it will be necessary to analyze what causes sudden changes in both.



Graph 6 Dimero D and Ferritin analysis in patient 55
Source: Own [Matlab]

Graph 6 represents patient number 55, who had a stay of 8 days, the patient was taken two samples of Dimero D and is represented with the blue color in the bar graph, while two samples of Ferritin were taken represented in red color throughout their stay, observing that Dimero D and Ferritin had high values in the first sample, while Dimero D rose for the second sample and Ferritin decreased considerably.

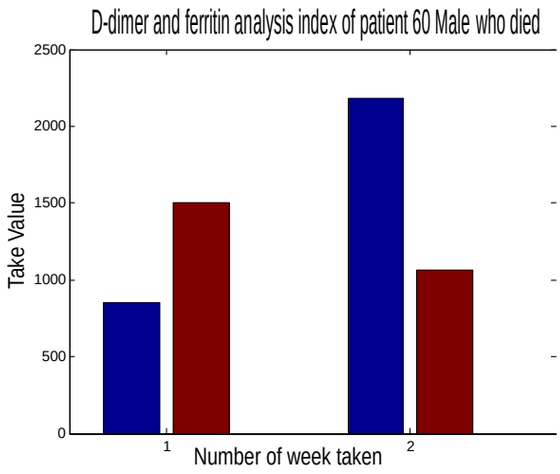


Graph 7 Dimero D and Ferritin analysis in patient 56
Source: Own [Matlab]

ISSN 2410-3438

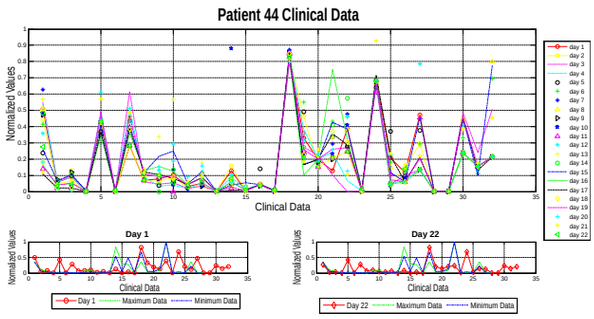
In the case of Graph 7, we have patient number 56, who had a stay of 26 days hospitalized, the patient was taken 4 samples of Dimero D and is represented with the blue color in the bar graph, while took 6 samples of Ferritin represented in red throughout their stay, observing that Dimero D and Ferritin maintain constant and low values, except for sample number 4 of Dimero D that a stronger change was observed but managed to reduce it in the next sample . The male patient was subsequently discharged.

Finally, Graph 8 shows patient 60 who died at the end of his stay, the patient was hospitalized for 11 days, for which there are only two samples taken of Dimero D and Ferritin, in both cases they have values above 800 in average in the first sample and in the second sample Dimero D increases and Ferritin does not fall below 1000, which is still a high value for normal values.



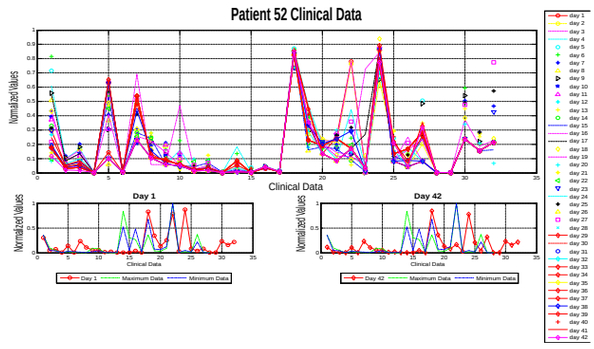
Graph 8 Dimero D and Ferritin analysis in patient 56
Source: Own [Matlab]

An important note in the statistical part of Dimero D and Ferritin is that in Dimero D the normal values must be below 100 ng / ml to be established as normal, above 500 ng / ml will mark blood clotting. In the case of fibrinogen, normal values are between 200 and 400 mg / dl.



Graph 9 Patient behavior 44
Source: Own [Matlab]

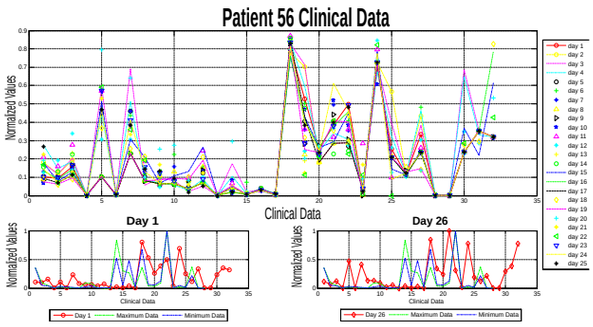
It is important to mention that day-to-day charts were made of the clinical laboratories of the patients admitted to the CECCS to have a comparative study of their progress or setbacks in their stay or to know on which dates more changes were made in the patients, either for right or wrong. The graphs in Graph 9 show the laboratories obtained day by day during his stay, remembering that he died after 22 days, Figure describes 3 graphs, the first shows the day-to-day values of the laboratories, the second graph (lower part left) represents the laboratory values of the first day in the hospital, likewise the normal minimum and maximum values of each laboratory are added, the third graph (lower right) shows the laboratory values of the last day of hospital stay, the values are very different in both graphs to the average values of a normal patient.



Graph 10 Behavior of the patient 52
Source: Own [Matlab]

The diagrams in Graph 10 show the laboratories obtained day by day during her stay, remembering that she was discharged after 42 days, Figure describes 3 graphs, the first shows the day-to-day values of the laboratories, the second graph (lower left) represents the laboratory values of the first day in the hospital, the third graph (lower right) shows the laboratory values of the last day of hospital stay, at some points the values are within the normal range.

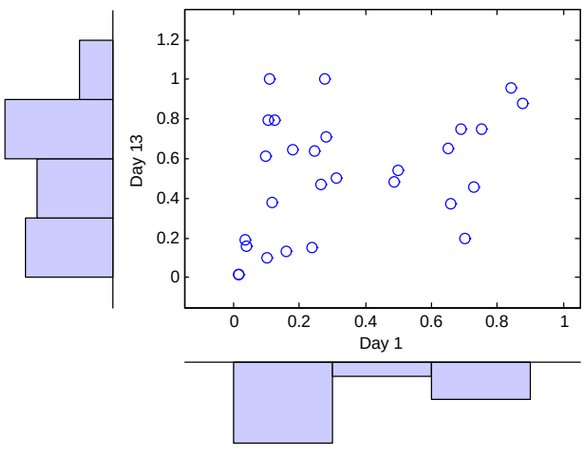
ISSN 2410-3438



Graph 11 Patient behavior 56
Source: Own [Matlab]

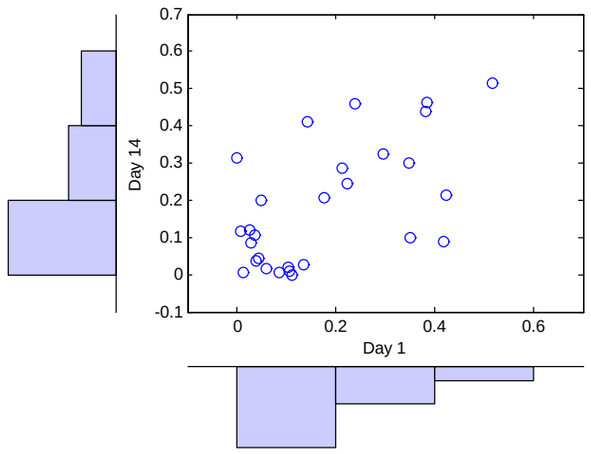
The diagrams in Graph 11 show the laboratories obtained day by day during his stay, remembering that he was discharged after 26 days, Figure shows 3 graphs, in the first one you can see the day-to-day values of the laboratories, the second graph (lower left) represents the laboratory values of the first day in the hospital, the third graph (lower right) shows the laboratory values of the last day of hospital stay.

Based on the information described above, it was proposed to perform a data analysis with that collected from the first day of admission and the last day of hospitalization, the idea is to generate behavioral comparisons through the study of the relationship / correlation of the information.



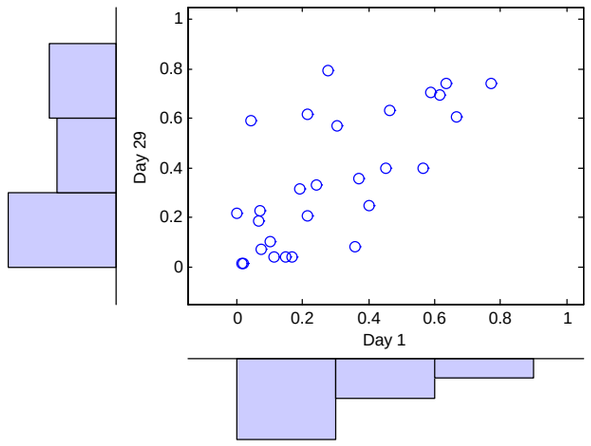
Graph 12 Laboratory relationship analysis of patient 17
Source: Own [Matlab]

Graph 12 shows the comparison of the two days, in which the normalized data are compared, in which the values obtained on that day are taken and its comparison, in which the patient, once he improves, should contain less dispersion of the information, so for patient number 17 (man) who unfortunately died, it is possible to observe the dispersion of the data.



Graph 13 Laboratory relationship analysis of patient 16
Source: Own [Matlab]

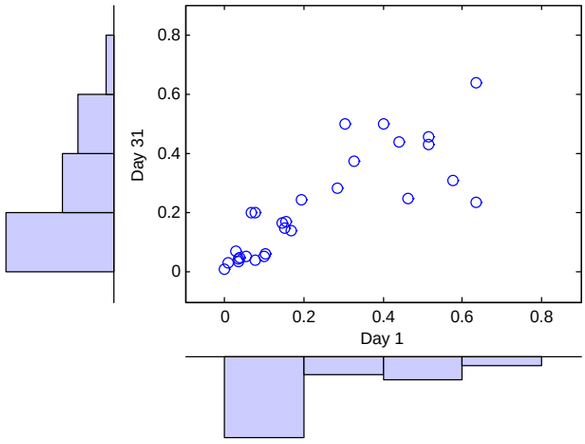
In the case of Graph 13, we have the comparison described in the previous paragraph, for this case patient 16 (man) who was discharged, and the trend is to have a greater convergence / union of the data between the first and last stay day.



Graph 14 Laboratory relationship analysis of patient 2
Source: Own [Matlab]

Observing Graph 14 and as the study was described previously, there is the case of patient 2 (woman) who died, it can be noted that there is a slight relationship / correlation between the data.

ISSN 2410-3438



Graph 15 Behavior of patient 14
Source: Own [Matlab]

In the case of Graph 15 there is the comparison described in the previous paragraphs, for this case patient 14 (woman) was discharged, in her there is a greater relationship / correlation of the data, which has been a trend in the different tests that have been carried out.

It is possible to observe how there are more data that converge to similar results from both the first day and the last day, probably establishing that the patient has not worsened throughout his stay in the CECCS or even had improvements, that is why they begin to overlap multiple values generating a higher correlation between the data.

Acknowledgments

The authors would like to thank the State Center for Critical Care of Salamanca (CECCS), the doctor Dr. Cesar Centeno Fosado and the head of nurses Liliana Yañez Vargas for their support for the preparation of the article and for providing the database, as well as the Polytechnic University of Juventino Rosas.

Conclusions

This article performs an analysis of the clinical laboratories of the patients admitted to the CECCS, with the idea of locating important points that can help to know more about the disease, in addition to looking for a correlation between the new laboratory updates and the statistics, which was confirmed by finding that Dimero D and Ferritin are highly relevant in the disease through the use of variance as a statistical study.

The multiple graphs shown describe the behavior of the disease and that of the patients who died becomes relevant, denoting the laboratories that have the greatest changes and that will benefit in the future when carrying out the statistical study prior to intubation and the supply of some medications, as that it is tried to find if there is a benefit or not of them.

The final approach of the relationship from the first day of stay to the last day shows profound changes in patients and laboratories that do not remain normal, with the above it will be possible to know where a greater emphasis should be placed for possible stronger studies.

References

Coronavirus disease (COVID-19). (2020, 14 diciembre). Coronavirus disease (COVID-19) pandemic. https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Coronavirus – *gob.mx*. (2020, 14 diciembre). https://coronavirus.gob.mx/

Coronavirus Guanajuato. (2020, 10 abril). Recuperado 4 de mayo de 2021, de https://coronavirus.guanajuato.gob.mx/

Devore, J., 2012. Probability And Statistics For Engineering And The Sciences. 8th ed. Australia: Brooks/Cole, pp.28-38.

Gao, Y., Li, T., Han, M., Li, X., Wu, D., Xu, Y., Wang, L. (2020). Diagnostic utility of clinical laboratory data determinations for patients with the severe COVID-19. *Journal of Medical Virology*, *92*(7), 791–796. https://doi.org/10.1002/jmv.25770

Gonzalez, R. C., & Woods, R. E. (2008). *Digital Image Processing* (3ª ed.). USA: Pearson/Prentice Hall.

Johnson, R. A., Miller, I., Freund, J. E., & Olguín, V. C. (2012). Probabilidad y estadística para ingenieros. Pearson.

ISSN 2410-3438

Liu, Y., Yang, Y., Zhang, C., Huang, F., Wang, F., Yuan, J., Wang, Z., Li, J., Li, J., Feng, C., Zhang, Z., Wang, L., Peng, L., Chen, L., Qin, Y., Zhao, D., Tan, S., Yin, L., Xu, J., … Liu, L. (2020). Clinical and biochemical indexes from 2019-nCoV infected patients linked to viral loads and lung injury. Science China Life Sciences, 63(3), 364-374. https://doi.org/10.1007/s11427-020-1643-8

Pagana, K., & Pagana, T. (2015). Laboratorio clínico: Indicaciones e interpretación de resultados (1.a ed.). Editorial El Manual Moderno

Rodriguez-Morales, A. J., Cardona-Ospina, J. A., Gutiérrez-Ocampo, E., Villamizar-Peña, R., Holguin-Rivera, Y., Escalera-Antezana, J. P., Sah, R. (2020). Clinical, laboratory and imaging features of COVID-19: A systematic review and meta-analysis. Travel Medicine and Infectious Disease, 34, 101623. https://doi.org/10.1016/j.tmaid.2020.101623

YAÑEZ-VARGAS Israel, DOÑATE-ÁLVAREZ Andrea, QUINTANILLA DOMINGEZ Joel and AGUILERA GONZÁLEZ Gabriel. Laboratory data statistical analysis from CECCS' patients with or without SARS-CoV-2 (COVID-19). Journal of Quantitative and Statistical Analysis. 2020. 7-21: 9-17

Yun, H., Sun, Z., Wu, J., Tang, A., Hu, M., & Xiang, Z. (2020). Laboratory data analysis of novel coronavirus (COVID-19) screening in 2510 patients. Clinica Chimica Acta, 507, 94-97. https://doi.org/10.1016/j.cca.2020.04.018

ISSN 2410-3438

Article **Journal of Quantitative and Statistical Analysis**
June, 2021 Vol.8 No.22 11-17

Analysis of multi-class classification performance metrics for remote sensing imagery imbalanced datasets

Análisis de métricas de rendimiento de clasificación multi-clase para conjuntos de datos desbalanceados de imágenes de percepción remota

GONZALEZ-RAMIREZ, Andrea†*[1], LOPEZ, Josue[1], TORRES-ROMAN, Deni[1] and YAÑEZ-VARGAS, Israel[2]

[1]*CINVESTAV del IPN, Unidad Guadalajara, Laboratorio de Telecomunicaciones*
[2]*Universidad Politécnica de Juventino Rosas, Departamento de Ingeniería en Redes y Telecomunicaciones*

ID 1st Author: *Andrea, González-Ramírez* / **ORC ID**: 0000-0001-9961-4763, **CVU CONACYT ID**: 1001207

ID 1st Coauthor: *Josué, Lopez* / **ORC ID**: 0000-0001-9703-3973, **CVU CONACYT ID**: 486202

ID 2nd Coauthor: *Deni, Torres* / **ORC ID**: 0000-0002-9813-7712, **CVU CONACYT ID**: 20075

ID 3rd Coauthor: *Israel, Yañez-Vargas* / **ORC ID**: 0000-0001-5749-8442, **CVU CONACYT ID**: 295711

DOI: 10.35429/JQSA.2021.22.8.11.17 Received March 25, 2021; Accepted June 30, 2021

Abstract

Remote sensing imaging datasets for classification generally present high levels of imbalance between classes of interest. This work presented a study of a set of performance evaluation metrics for an imbalance dataset. In this work, a support vector machine (SVM) was used to perform the classification of seven classes of interest in a popular dataset called Salinas-A. The performance evaluation of the classifier was performed using two types of metrics: 1) Metrics for multi-class classification, and 2) Metrics based on the binary confusion matrix. In the results, a comparison of the scores of each metric is developed, some being more optimistic than others due to the bias that they present given the imbalance. In addition, our case study helps to conclude that the Matthews correlation coefficient (MCC) presents the lowest bias in imbalanced cases and is regarded to be robust metric. These results can be extended to any imbalanced dataset taking into account the equations developed by Luque.

Remote sensing images, Metrics, Imbalanced datasets

Resumen

Los conjuntos de datos de imágenes de percepción remota para la clasificación generalmente presentan altos niveles de desbalance entre las clases de interés. Este trabajo presentó un estudio de un conjunto de métricas de evaluación del desempeño para un conjunto de datos desbalanceados. En este trabajo, se utilizó una máquina de vectores de soporte (SVM) para realizar la clasificación de siete clases de interés en un conjunto de datos popular llamado Salinas-A. La evaluación del desempeño del clasificador se realizó utilizando dos tipos de métricas: 1) Métricas para clasificación multiclase y 2) Métricas basadas en la matriz de confusión binaria. En los resultados se desarrolla una comparativa de las puntuaciones de cada métrica, algunas siendo más optimistas que otras por el sesgo que presentan ante el desbalance. Además, nuestro caso de estudio ayuda a concluir que el coeficiente de correlación de Matthews (MCC) presenta el sesgo más bajo en los casos de desbalance y se considera una métrica robusta. Estos resultados se pueden extender para cualquier conjunto de datos desbalanceado teniendo en cuenta las ecuaciones desarrolladas por Luque.

Imágenes de percepción remota, Métricas, Conjuntos de datos desbalanceados

Citation: GONZALEZ-RAMIREZ, Andrea, LOPEZ, Josue1, TORRES-ROMAN, Deni and YAÑEZ-VARGAS, Israel. Analysis of multi-class classification performance metrics for remote sensing imagery imbalanced datasets. Journal of Quantitative and Statistical Analysis. 2021. 8-22: 11-17

* Correspondence to Author (e-mail: andrea.Gonzalez@cinvestav.mx)
† Researcher contributing as first author.

Article **Journal of Quantitative and Statistical Analysis**
June, 2021 Vol.8 No.22 11-17

1. Introduction

Monitoring and analysis of the land surface through remote sensing image processing has become an important task in recent years. (Cheng, 2016, Wang, 2016). Therefore, the use and design of classification algorithms, such as Support Vector Machine (SVM) (Malgani, 2004, Liu, 2017), Random Forest (RF) (Belgiu, 2016) and Artificial Neural Networks (ANN) (Long, 2017), among others, has become an essential task to study these images. One of the main aspects that can influence the performance of a classifier is the imbalance in the input dataset. This happens when one of the classes in the dataset has a greater number of samples than the others. (Chawla 2005, Krawczyk 2016).

Due to the class imbalance in the image data sets, some metrics present high levels of bias, which generates optimistic results in the evaluation of the classifiers (Luque, 2019). Hence, the study of imbalanced data sets has become an important problem in the classification, since it implies the determination of the most appropriate performance metrics that will be used for the performance evaluation of the classifier, therefore, a selection adequate evaluation metrics is an important key (Luque 2019, Boughorbel 2017, Chicco 2020).

There are many metrics that are useful for testing the capability of any single or multi-class classifier, which are useful for comparing performance and analyzing the behavior of the same model (Luque, 2019, Grandini, 2020, Chicco 2020). In this work, an analysis of the bias for a set of classification metrics is presented, both in multi-classes and in those based on the binary confusion matrix. In section 5 a comparison of a set of metrics is made in a case of study of hyperspectral remote sensing images, using SVM.

2. Contribution

This work presents a study for a set of performance evaluation metrics, in which the bias that occurs in multiclass metrics and the information obtained through the binary confusion matrix for an imbalanced data set will be analyzed. For the following study, the Matthews correlation coefficient (MCC) is considered a robust metric, as MCC exhibits the lowest bias in cases of data imbalance (Luque, 2019).

3. Problem statement

Let $\mathcal{X} \in \mathbb{R}^{I\times J\times K}$ be an imbalanced multidimentional array the input to a multi-class classifier, where *I, J* denote the spatial resolution and *K* spectral resolution, that produce a prediction $\widehat{\mathbf{Y}} \in \mathbb{C}^{I\times J}$, where $\mathbb{C}$ the set of the classes of interest, and let $\mathbf{Y} \in \mathbb{C}^{I\times J}$ be the corresponding labels indicating the actual class of each element of $\mathbf{x}$, analyse the behavior of a set of performance evaluation metrics at different imbalance levels.

3.1. Mathematical definition

Given a multi-class confusion matrix $\mathrm{M} \in \mathbb{R}^{C\times C}$, obtained from the true labels Y and the prediction $\widehat{\mathrm{Y}}$, where *C* denotes the number of classes of interest, and the element $\mathrm{m}_{ij} \in \mathrm{M}$ is the number of samples that belong to class *i*-th, but that are classified as members of class *j*-th, i.e., $y_n = i$ and $\hat{y}_n = j$, for $n = 1, \dots, N$.

In this work, we use the mathematical definition of imbalance given by (Luque, 2019). The imbalance coefficient δ_c, for a given class c, is featured with a value in the $[-1,1]$ range, where 0 means that classes are perfectly balanced. The *c*-class imbalance coefficient can be computed by

$$\delta_c = 2\frac{\sum_{j=1}^{C} m_{cj}}{\sum_{i=1}^{C}\sum_{j=1}^{C} m_{ij}} - 1\,. \tag{1}$$

A performance evaluation metric μ is a function that assigns to each confusion matrix, a real value on the set $\mathbb{R}$ i.e.

$$\mu: Y \times \hat{Y} \to \mathbb{R}. \tag{2}$$

In general, the result of this function is in the range[0,1], where 0 means the perfect misclassification, and 1 the perfect classification.

4. Metrics for multi-class classification

4.1. Overall Accuracy

The Overall Accuracy (OA) is the ratio between the number of correctly classified elements and the overall number of samples, and it is computed from multi-class confusion matrix by

$$\mathrm{OA} = \frac{\sum_{i=1}^{C} m_{ii}}{\sum_{i=1}^{C}\sum_{j=1}^{C} m_{ij}}\,. \tag{3}$$

ISSN 2410-3438

GONZALEZ-RAMIREZ, Andrea, LOPEZ, Josue1, TORRES-ROMAN, Deni and YAÑEZ-VARGAS, Israel. Analysis of multi-class classification performance metrics for remote sensing imagery imbalanced datasets. Journal of Quantitative and Statistical Analysis. 2021

This metric is one of the most used for classification performance evaluation. Nevertheless, if the dataset is imbalanced, the OA is not a reliable measure, as it produces optimistic results (Chicco, 2020).

4.2. Balanced Accuracy

An alternative metric to reduce the impact of imbalanced classes in performance evaluation is Balanced Accuracy (BA), computed by the average of another well-known metric, the Sensitivity (SNS) per class, i.e.

$$\mathrm{BA} = \frac{\sum_{c=1}^{C} \mathrm{SNS}_c}{C}, \quad (4)$$

where $\mathrm{SNS}_C = \frac{m_{cc}}{\sum_{j=1}^{C} m_{cj}}$. SNS measures the proportion of the number of elements correctly classified from an individual class.

4.3. Balanced Accuracy Weighted

The Balanced Accuracy Weighted (BAW) is a metric used for imbalanced classes, in which the SNS_C of each class is weighted by its relative frequency w_c. The formula for BAW is

$$\mathrm{BAW} = \frac{\sum_{i=1}^{C} \mathrm{SNSw}_C}{\sum_{c=1}^{C} w_c}, \quad (5)$$

where $\mathrm{SNS}w_C = \frac{m_{CC}}{\sum_{j=1}^{C} m_{cj}} w_c$ and $w_c = \frac{\sum_{i=1}^{C}\sum_{j=1}^{C} m_{ij}}{C \sum_{j=1}^{C} m_{cj}}$ (P, 2021). This metric is an efficient performance indicator, since the recalls are weighted by a relative frequency to the classes size (P, 2021, Grandini, 2020).

4.4. Cohen's Kappa Coefficient

The Kappa Coefficient is currently one of the most popular metrics in machine learning for classification performance evaluation (Chicco, 2020). This metric measures the inter-rater concordance, as the degree of agreement among raters. It is computed by

$$\mathrm{K} = \frac{qs - \sum_{c=1}^{C} p_c t_c}{s^2 - \sum_{c=1}^{C} p_c t_c}, \quad (6)$$

where $q = \sum_{c=1}^{C} m_{cc}$ denotes the overall number of elements correctly predicted, $s = \sum_{i=1}^{C}\sum_{j=1}^{C} m_{ij}$ is the total number of samples, $p_c = \sum_{j=1}^{C} m_{cj}$ the number of times class c was predicted, and $t_c = \sum_{i=1}^{C} m_{ic}$ the number of times class c truly occurs (Grandini, 2020). Kappa coefficient is high sensitive to the marginal totals (Chicco, 2020).

4.5. Matthews Correlation Coefficient

The Matthews Correlation Coefficient (MCC) is generally considered a balanced performance evaluation metric. All the elements of the confusion matrix are included in the numerator and denominator of its formula (Eq. 7), so this metric is less biased by imbalanced datasets than other metrics (Chicco, 2020).

$$\mathrm{MCC} = \frac{qs - \sum_{c=1}^{C} p_c t_c}{\sqrt{(s^2 - \sum_{c=1}^{C} p_c^2)(s^2 - \sum_{c=1}^{C} t_c^2)}} \quad (7)$$

The main disadvantage is that MCC is undefined for extreme cases, for instance, when a whole row or column of the confusion matrix is zero. Generally MCC and Kappa are used in their normalized version to be in the range [0, 1], computed by $\mathrm{MCCn} = \frac{\mathrm{MCC}+1}{2}$ and $\mathrm{Kn} = \frac{\mathrm{K}+1}{2}$ respectively.

5. Metrics based on the binary confusion matrix

For the particular case of binary classification, the elements of the confusion matrix are defined as the True Positives (TP), is the case where the prediction is true and the actual output is also true, $\mathrm{TP} = m_{11}$, the False Negatives (FN), is the case where the prediction is false and the actual output is true, $\mathrm{FN} = m_{12}$, the False Positives (FP), is when the prediction is true and the actual output is false, $\mathrm{FP} = m_{21}$, and the True Negatives (TN), is the cases where the prediction is false and the actual output is false., $\mathrm{TN} = m_{22}$.

In the work of Luque (Luque, 2019), a classical metric based on the binary confusion matrix $\mu(\mathrm{TP}, \mathrm{TN}, \mathrm{FP}, \mathrm{FN})$ is expressed as a function $\mu(\lambda_{PP}, \lambda_{NN}, \delta)$, where λ_{PP} represents the ratio of the correctly classified positive elements $m_{PP} = \mathrm{TP}$ and the total of truly positive elements $m_P = \mathrm{TP} + \mathrm{FN}$, i.e., $\lambda_{PP} = \frac{m_{PP}}{m_P}$.

ISSN 2410-3438

And λ_{NN} denotes the ratio of the correctly classified negative $m_{NN} = \text{FP}$ and the total of truly negative elements $m_N = \text{FP} + \text{TN}$, i.e., $\lambda_{NN} = \frac{m_{NN}}{m_N}$.

Considering the balanced case, $\mu_b = \mu(\lambda_{PP}, \lambda_{NN}, 0)$, it is possible to define the impact of imbalance by the bias of a metric B_μ as

$$B_\mu = \mu - \mu_b. \quad (8)$$

Metrics	$\mu(\text{TP}, \text{TN}, \text{FP}, \text{FN})$
PRC(PPV)	$\frac{\text{TP}}{\text{TP+FP}}$
NPV	$\frac{\text{TN}}{\text{TN+FN}}$
ACC	$\frac{\text{TP+TN}}{\text{TP+FN+TN+FP}}$
F_1	$2\frac{\text{PRC}\cdot\frac{\text{TP}}{\text{TP+FN}}}{\text{PRC}+\frac{\text{TP}}{\text{TP+FN}}}$
GM	$\sqrt{\frac{\text{TP}}{\text{TP+FN}}\cdot\frac{\text{TN}}{\text{TN+FP}}}$
MCCn	$\frac{\text{TP}\cdot\text{TN}-\text{FP}\cdot\text{FN}}{\sqrt{(\text{TP+FP})(\text{TP+FN})(\text{TN+FP})(\text{TN+FN})}}$
MKn	$\text{PPV} + \text{NPV} - 1$

Table 1 Classification performance metrics as classical metric. (*Luque, 2019)*

Metrics	$\mu(\lambda_{PP}, \lambda_{NN}, \delta)$
PRC(PPV)	$\frac{\lambda_{PP}(1+\delta)}{\lambda_{PP}(1+\delta)+(1-\lambda_{NN})(1-\delta)}$
NPV	$\frac{\lambda_{NN}(1-\delta)}{\lambda_{NN}(1-\delta)+(1-\lambda_{PP})(1+\delta)}$
ACC	$\lambda_{PP}\frac{1+\delta}{2}+\lambda_{NN}\frac{1-\delta}{2}$
F_1	$\frac{2\lambda_{PP}(1+\delta)}{(1+\lambda_{PP})(1+\delta)+(1-\lambda_{NN})(1-\delta)}$
GM	$\sqrt{\lambda_{PP}\cdot\lambda_{NN}}$
MCCn	$\frac{1}{2}\left(\frac{\lambda_{PP}+\lambda_{NN}-1}{\sqrt{\left[\lambda_{PP}+(1-\lambda_{NN})\frac{1-\delta}{1+\delta}\right]\left[\lambda_{NN}+(1-\lambda_{PP})\frac{1+\delta}{1-\delta}\right]}}+1\right)$
MKn	$\frac{1}{2}\left(\frac{1+\delta}{(1+\delta)+\frac{1-\lambda_{NN}}{\lambda_{PP}}(1-\delta)}+\frac{1-\delta}{(1-\delta)+\frac{1-\lambda_{PP}}{\lambda_{NN}}(1+\delta)}\right)$

Table 2 Classification performance metrics as a function of imbalance. (*Luque, 2019)*

Metrics	$\mu_b(\lambda_{PP}, \lambda_{NN})$
PRC(PPV)	$\frac{\lambda_{PP}}{\lambda_{PP}+(1-\lambda_{NN})}$
NPV	$\frac{\lambda_{NN}}{\lambda_{NN}+(1-\lambda_{PP})}$
ACC	$\frac{\lambda_{PP}+\lambda_{NN}}{2}$
F_1	$\frac{2\lambda_{PP}}{2+\lambda_{PP}-\lambda_{NN}}$
GM	$\sqrt{\lambda_{PP}\cdot\lambda_{NN}}$
MCCn	$\frac{1}{2}\left(\frac{\lambda_{PP}+\lambda_{NN}-1}{\sqrt{[\lambda_{PP}+(1-\lambda_{NN})][\lambda_{NN}+(1-\lambda_{PP})]}}+1\right)$
MKn	$\frac{1}{2}\left(\frac{1}{1+\frac{1-\lambda_{NN}}{\lambda_{PP}}}+\frac{1}{1+\frac{1-\lambda_{PP}}{\lambda_{NN}}}\right)$

Table 3 Classification performance metrics as a function of balance. (*Luque, 2019)*

We lead this approach to the multi-class case, starting from the multi-class confusion matrix, generating multiple binary confusion matrices M_c using the equations 9a to 9d. It is important to highlight that the results given by the transformation of the multi-class confusion matrix to binary confusion matrices are not necessarily equal to the results when using multiple binary classifiers.

$$\text{TP}_c \cong m_{cc} \quad (9a)$$

$$\text{FP}_c \cong \sum_{j=1}^{C} m_{cj} \text{ for } j \neq c \quad (9b)$$

$$\text{FN}_c \cong \sum_{i=1}^{C} m_{ic} \text{ for } i \neq c \quad (9c)$$

$$\text{TN}_c \cong \sum_{i=1}^{C}\sum_{j=1}^{C} m_{ij} \text{ for } i \neq c \text{ and } j \neq c \quad (9d)$$

Hence, the equations for each metric presented in Tables 2 y 3 can be used to evaluate classification performance given a class of reference. In this analysis, only one of the four metrics that do not present bias, studied in (Luque, 2019), was considered for comparison purposes. In addition, an overall performance can be given by the average of the metric for each class.

The classification metrics studied under this approach are: Precision (PRC), Negative Predictive Value (NPV), Accuracy (ACC), F_1 score (F_1) , Geometric Mean (GM), the Normalized versions of Matthews Correlation Coefficient (MCCn), and Markedness (MKn).

6. Results

In this work, we use an imagery dataset, to study imbalance with the performance evaluation metrics aforementioned. We use popular hyperspectral dataset for experimentation in this work, are available at Hyperspectral Remote Sensing Scenes .

In addition, a Support Vector Machine (SVM) multiclass classifier is used in this work for quantitative analysis of the classification performance evaluation metrics described in Section 4, (Cortes, 2009).

ISSN 2410-3438

6.1. Study Case

6.1.2. SalinasA scene

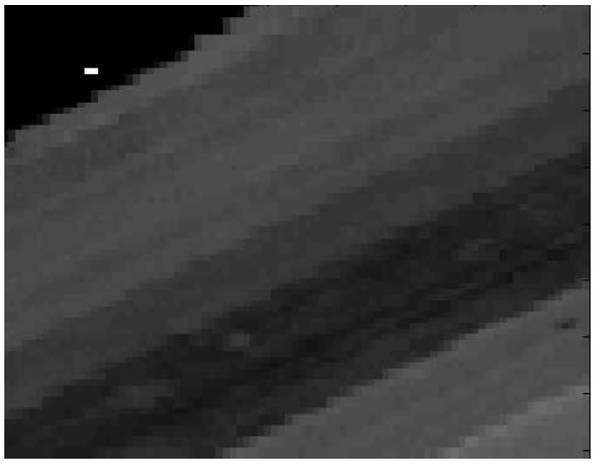


Figure 1 SalinasA scene. (*http://www.ehu.eus/ccwintco /index.php/Hyperspectral_Remote_Sensing_Scenes*).

Figure 1, shows a small sub-scene collected by the AVIRIS sensor over Salinas Valley, California. It has 86×83 pixels and 204 bands.

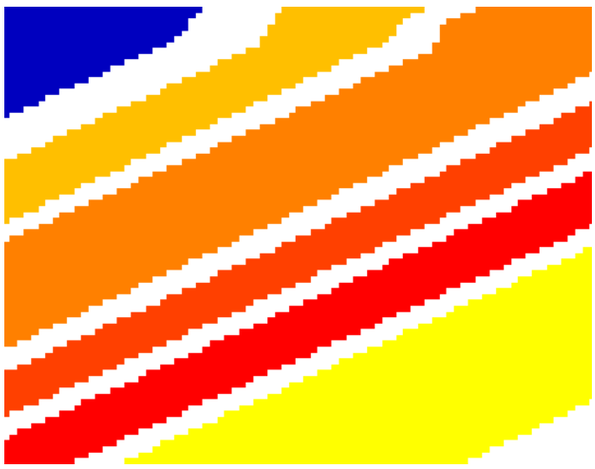


Figure 2 Groundtruth SalinasA scene. (*http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes*).

Figure 2 and Table 4 describe the Salinas ground truth, where each of its 6 classes is shown with their labels and corresponding samples (M Graña, MA Veganzons, B Ayerdi).

	Class	Samples	Training	Test
1	Brocoli-GW1	391	37	354
2	Corn-SGW	1343	143	1200
3	Lettuce-R4	616	70	546
4	Lettuce-R5	1525	130	1395
5	Lettuce-R6	674	69	605
6	Lettuce-R7	799	85	714

Table 4 Ground truth classes for the Salinas-A scene and their respective samples number (*http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes*)

Table 4 describes the classes of the hyperspectral image, where we can observe the samples of each class, with the idea of visualizing that the data is imbalanced, where classes 2 and 4 have a greater number of samples than the others, while class 1 has the smallest number of samples.

Figure 3 shows the multi-class confusion matrix obtained from the SVM classifier implemented in this work, where 90% of the samples were used as tests and 10% of the samples as training, in which we can observe the imbalance of the classes as well as the performance of the classifier.

		Predictions					
Actual	353	0	0	1	0	0
	0	1083	26	20	60	11
	0	152	360	34	0	0
	0	5	0	1390	0	0
	0	0	0	0	603	2
	0	0	0	0	17	697

Figure 3 Confusion matrix. *Python, own.*

With the results of the previously seen multiclass confusion matrix, we obtained the following results (See Table 5), applying the multiclass classification metrics, described in section 3, where we can observe the performance evaluation of the classifier.

Metric	Performance
OA	0.931865393
BA	0.921386044
BAW	0.918517413
AA	0.977288464
Kn	0.957131663
MCCn	0.957576207

Table 5 Metrics multi-class classification. *Python, own.*

ISSN 2410-3438

One of the important points of this work is the analysis of metrics based on the binary confusion matrix, described in section 4. In which we perform a conversion from the multi-class confusion matrix to multiple binary confusion matrices per class using equations (9a), (9b), (9c), (9b), to apply the equations of Table 2 considering equation (1) and we obtained the results shown in Table 6, where the performance evaluation can be observed of each of the metrics.

Class	PRC	NPV	ACC	F1	GM	MCCn	MKn
1	0.9999	0.9998	0.9998	0.9986	0.9985	0.9992	0.9999
2	0.8734	0.9673	0.9431	0.8877	0.9228	0.9249	0.9203
3	0.9326	0.958	0.956	0.7725	0.7168	0.8814	0.9453
4	0.9619	0.9985	0.9875	0.9789	0.9891	0.9852	0.9802
5	0.8868	0.9995	0.9836	0.9385	0.9922	0.9656	0.9431
6	0.9817	0.9959	0.9938	0.9789	0.9868	0.9876	0.9888
Average	0.9394	0.9865	0.9773	0.9259	0.9343	0.9573	0.9629

Table 6 Classification performance evaluation metrics. *Python, own.*

We develop out the same process using the equations in Table 3 where the variable in equation (1) is $\delta_c = 0$, this to apply equation 8 described in section 4, where we can show the result of the bias of each of the metrics in Table 7.

Class	PRC	NPV	ACC	F1	GM	MCCn	MKn
1	$2e^{-6}$	0.0026	0.0012	$1e^{-5}$	-	0.0006	0.0013
2	-0.081	0.0598	0.0136	-0.04	-	-0.005	-0.011
3	-0.058	0.2133	0.1293	-0.019	-	0.0348	0.0775
4	-0.022	0.0021	-0.003	-0.011	-	-0.005	-0.01
5	-0.095	0.0029	-0.006	-0.051	-	-0.024	-0.046
6	-0.015	0.0192	0.0073	-0.007	-	0.001	0.0021
Average	-0.045	0.05	0.0239	-0.021	-	0.0004	0.0024

Table 7 Bias. *Python, own*

7. Conclusions

This paper presents the analysis of a set of classification performance evaluation metrics, under imbalanced data sets.

In this paper, we used SVM as a classifier in different experiments, where 90% of the samples for testing and 10% of the samples for training were considered. As an output from the classifier, we obtained a multiclass confusion matrix observed in Figure 3. Analyzing the results obtained, we can see that for class 2 and 3 there was a higher proportion of errors than in other classes when performing the classification. Since class 2 is one of the classes with highest number of samples, some of the metrics would penalize the result more.

Classes 1,4,5,6 were very high classification (over 0.97) therefore, so a metric such as AA that averages the accuracies for each class presents very positive results. While OA is mostly reflected in the classification errors of classes 2 and 3. On the other hand, the BA and BAW metrics are less optimistic than the OA due to the fact that it mainly considers class 2 errors due to the imbalance. Besides, MCC and K are highly correlated and their scores are higher than OA, BA and BAW. These metrics usually generate high scores only if the classifier was able to correctly predict most of the data (Chicco, 2020).

At the same time, we propose to perform a performance evaluation by class, and from the multiclass confusion matrix, we obtained multiple binary confusion matrices with equations 9a-9d. The metrics based on binary confusion matrices from Table 2 are used. The results are shown in Table 6.

Classes 2 and 3, present a low proportion of TP, hence, the classification scores for classes are penalized by metrics such as PRC and F1, while other metrics such as NPV and ACC do not depend only on TP, but also on TN, which produces higher scores for our case study.

In Table 7, we have the bias of each metric by class, where we can see that the F1 metric presents less bias on average than PRC, NPV and ACC, the F1 metric is a good metric when we perform an evaluation by class. Since GM does not depend on the imbalance, i.e., an unbiased metric, therefore it is a good point of comparison for other metrics.

The MCCn metric turns out to be the least biased and the most reliable for evaluating imbalanced data, (Luque, 2019). In the results, MKn shows to be one of the least biased metrics after MCCn, as a mentioned in (Luque, 2019).

References

Belgiu, Mariana & Drăguţ, Lucian. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing. 114. 24-31. 10.1016/j.isprsjprs.2016.01.011.

ISSN 2410-3438

Boughorbel, S. & Jarray, Fethi & El-Anbari, Mohammed. (2017). Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. PLOS ONE. 12. e0177678. 10.1371/journal.pone.0177678.

Chawla, Nitesh. (2005). Data Mining for Imbalanced Datasets: An Overview. 10.1007/0-387-25465-X_40.

Cheng, Gong & Han, Junwei. (2016). A Survey on Object Detection in Optical Remote Sensing Images. ISPRS Journal of Photogrammetry and Remote Sensing. 117. 11-28. 10.1016/j.isprsjprs.2016.03.014.

Chicco, Davide & Jurman, Giuseppe. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 21. 10.1186/s12864-019-6413-7.

Cortes, Corinna & Vapnik, Vladimir. (2009). Support-vector networks. Chem. Biol. Drug Des.. 297. 273-297. 10.1007/%2FBF00994018.
P. (2021, 6 enero). How to Improve Class Imbalance using Class Weights in Machine Learning. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/10/improve-class-imbalance-class-weights/

Grandini, Margherita & Bagli, Enrico & Visani, Giorgio. (2020). Metrics for Multi-Class Classification: an Overview.

Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. Progress in Artificial Intelligence, 5, 221-232.

Liu, Peng & Choo, Kim-Kwang Raymond & Wang, Lizhe & Huang, Fang. (2017). SVM or deep learning? A comparative study on remote sensing image classification. Soft Computing. 21. 10.1007/s00500-016-2247-2.

Long, Yang & Gong, Yiping & Xiao, Zhifeng & Liu, Qing. (2017). Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks. IEEE Transactions on Geoscience and Remote Sensing. PP. 1-13. 10.1109/TGRS.2016.2645610.

Luque, A. & Carrasco, Alejandro & Martín, Alejandro & de Las Heras, Ana. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. Pattern Recognition. 91. 6829. 10.1016/j.patcog.2019.02.023.

Melgani, Farid & Bruzzone, Lorenzo. (2004). Classification of Hyperspectral Remote Sensing Images with Support Vector Machines. Geoscience and Remote Sensing, IEEE Transactions on. 42. 1778 - 1790. 10.1109/TGRS.2004.831865.

Wang, Weixing & Yang, Nan & Zhang, Yi & Wang, Fengping & Cao, Ting & Eklund, Patrik. (2016). A Review of Road Extraction from Remote Sensing Images. Journal of Traffic and Transportation Engineering (English Edition). 3. 10.1016/j.jtte.2016.05.005.

ISSN 2410-3438

Article **Journal of Quantitative and Statistical Analysis**
June, 2021 Vol.7 No.22 18-24

Productive performance of York x Landrace sows in a semi-technified farm

Comportamiento productivo de cerdas York x Landrace en una granja semi tecnificada

SANCHEZ-CHIPRES, David Román†*, MORENO-LLAMAS, Gabriel, JIMÉNEZ-PLASCENCIA Cecilia and JIMÉNEZ-CORDERO, Ángel Andrés

Animal Production Departament, CUCBA, Universidad de Guadalajara

ID 1st Author: *David Román, Sanchez-Chipres* / **ORC ID**: 0000-0002-5273-0393, **CVU CONACYT ID**: 69431

ID 1er Co-author: *Gabriel, Moreno-Llamas* / **ORC ID**: 0000-0002-1003-1738, **CVU CONACYT ID**: 101392

ID 2do Co-author: *Cecilia, Jiménez-Plascencia* / **ORC ID**: 0000-0003-1705-792X, **CVU CONACYT ID**: 1000339

ID 3er Co-author: *Ángel Andrés, Jiménez-Cordero* / **ORC ID**: 0000-0002-1734-2678, **CVU CONACYT ID**: 947963

DOI: 10.35429/JQSA.2021.22.8.18.24 Received September 05, 2021; Accepted December 05, 2021

Abstract

The aim of this work was to evaluate the productive performance of sows in a semi-technified farm. Sows were F1 York x Landrace inseminated with Pietrain. For the study, we considered 13 litters of the same age. It was recorded number of piglets born, weight at birth and at weaning, as well as number of piglets weaned. The information was recorded in the Pigchamp© program. When analyzing all the information regarding litter size and litter weight, compared with other studies, the data obtained in this work is good. With all the variables recorded, we observed a high and positive correlation, similar to Murillo *et al*. (2017) results. The use of these variables as indicators of maternal productivity increases with each sow calving.

Production, Piglets, Temperature

Resumen

El propósito de este trabajo fue evaluar el comportamiento productivo de cerdas F1 York x Landrace inseminadas con sementales Pietrain, en una granja semi tecnificada. Se tomaron en cuenta 13 camadas de la misma edad. Se obtuvo información de lechones nacidos, peso de lechones al nacimiento y destete, lechones destetados y días de lactancia. Los datos se registraron en el programa Pigchamp©. El análisis de los registros sobre tamaño de camada y peso de camada, comparados con otros estudios, muestra que los resultados del presente trabajo son comparables. Las correlaciones calculadas entre las diferentes variables, indican que hubo valores positivos, similar a los resultados de Murillo *et al*. (2017). El análisis de estas variables como indicadores de productividad materna, indica mejora con cada parto.

Producción, Lechones, Temperatura

Citation: SANCHEZ-CHIPRES, David Román, MORENO-LLAMAS, Gabriel, JIMÉNEZ-PLASCENCIA Cecilia and JIMÉNEZ-CORDERO, Ángel Andrés. Productive performance of York x Landrace sows in a semi-technified farm. Journal of Quantitative and Statistical Analysis. 2021. 8-22: 18-24

† Investigador contribuyendo como primer autor.

Introduction

Annual meat pork consumption in Mexico grew at an annual rate of 3.9%; estimated increase was from 1.4 million tons in 2006 to 2.1 million tons in 2018. It is expected this trend will continue in 2019, to achieve a consumption of 2.2 million tons, 4.3% of annual increase. In the 2014-2018 period, 69% of the pork meat consumption, came from the national production, the rest was imported. Pork meat is a less expensive alternative than other meats in Mexico. Jalisco is an important pork meat producer; it ranked the third nationwide in the last 30 years. Jalisco production in 2018 was 321,735 tons of pork meat (SAGARPA-SIAP, 2019).

There are four specialized regions for swine production: Northern High Lands (Lagos de Moreno, Jalisco, county as the center of the region), Southern High Lands (Tepatitlán County), South (Ciudad Guzmán County) and Central (Zapopan County). Northern and Southern High Lands are the most important regions, with 85.6% of the total pork population, 3,413,513 animals (INEGI 2016). Classification of the swine production systems in Jalisco are intensive, semi-intensive and extensive. The three production types coexist in Jalisco, although the predominant ones are the semi-intensive systems.

This classification depends on the level of technology for environment and animal management. These animal populations are denser and more confined in the intensive production models. Organization and operation of the intensive production systems work according to the capital model, which means more work and profit in the shortest time possible. This is only feasible to the extent that there is standardization in every link in the production chain, to ensure a commercial product uniformity. Such a model needs high tech. The essential resource in the system is the swine, and to guarantee the productive dynamics there are necessary populations with uniform traits. Achievement of this goal requires genetically homogeneous animals, as the result of high selection pressures. It is also necessary a stable and comfortable environment; this way the strength levels of the intensive, semi-intensive or extensive animal production systems, are defined when it is established the levels of environmental control (POET, 1998).

Intensive and semi intensive production is the majority of Jalisco farms system. The usual genetic background in these farms are inter-racial hybrids, although not all come from schemes of planned crossings. Frequently, the breeding grounds of different racial genotypes come from countries as Canada, United States or England.

Jalisco swine farmers are looking for technological changes to achieve higher efficiency in their production processes. They are trying to obtain pigs in 21 weeks with a market weight of 100-110 kg. Improvement of food quality and the breeding systems are the commonly used advances in technified farms. Pork producers pursue a daily increase through genetic breeding, but in Jalisco, there is not a standardized production scheme that take in account the crossing races. Female lines are usually F1 Yorkshire x Landrace, Yorkshire x Hampshire, Large White x Landrace or Landrace x Hampshire. As males, the most frequent races are Duroc, Hampshire, Pietrain, or hybrids of these races (Flores y Gómez, 1995). This leads to producing swine with differences in reproductive performance, feeding efficiency, adaptation to several environmental factors, etc. It is evident when comparing the in country animals with imported ones, there is a difference of 0.6-0.7 kg in piglets at birth and 0.6 kg in weaned piglets (Batista, 2000).

Problem approach

Reproductive traits present a low heritability; to achieve much of the improvement goals it is necessary to reinforce good management practices, as well as housing systems and climatic conditions that influence the sow's productive performance. Not to carry out these activities, result in an underutilization of the sow, and will limit the possibility of genetic improvement of the herd.

Justification

Currently farms require a higher productivity to maintain financial efficiency. The sows are the base to keep high productive parameters, so it is necessary to provide environmental conditions to express their reproductive capacity. Information and analysis thereof shall enable to identify improvement areas in the farm.

ISSN 2410-3438

General Objective

Evaluate the productive performance of sows in a semi-technified farm.

Particular Objectives

Evaluate parameters of live-born piglets, weaned piglets and piglets weight at weaning; correlate parameter effects on litter size at birth and at weaning, the number of births per female, number of weaned piglets per female.

Materials and methods

Location

The study was carried out in December 2017 to January 2018, in a commercial farm in the Zacoalco de Torres County, state of Jalisco. It is a semi-technified farm with 230 bellies.

Animals

It was considered 13 litters of (York x Landrace) x Pietrain.

Studied variables

Five variable records were made of the 13 litters: mortality, weaned piglets, litter weight at weaning and lactating days.

Experimental methodology

The study included thirteen contemporary litters. It was recorded litter size at birth, mortality, litter weight at birth and at weaning, weaned piglets, and lactating days. Data was recorded with Pigchamp© software. It was also recorded four types of temperatures around the litters during the first 28 days of life: maximum, minimum, average and midpoint, with a Fluke® thermograph.

Experimental design

Regression analysis of the productive data was done in order to understand the relation between them, and with temperatures (Steel and Torrie, 1960).

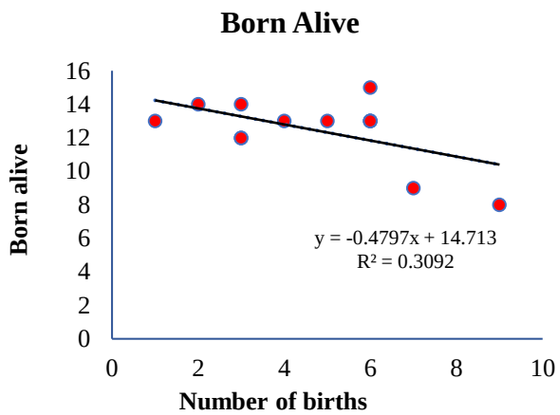
Results

More relevant results are included in this section. Table 1 shows the litter's information from birth to weaning, compared to the average environmental temperature in which the piglets lived the first 28 days.

BA	BD	MU	WL kg	INW kg	DON	ADO	MOR	WEA	WW kg	IWW kg	LAD	AVT°C
15	1	1	17.2	1.15	3	0	0	12	91.2	7.6	30	48.8
13	2	0	16.8	1.29	1	0	0	12	84.2	7.02	30	53.4
12	1	1	15.0	1.25	0	0	2	10	67.2	6.72	30	66.6
13	0	0	13.6	1.05	0	0	2	11	85.2	7.75	34	35.7
12	3	0	16.2	1.35	0	0	3	9	71.2	7.91	29	31.5
8	0	0	12.1	1.51	0	3	1	11	89.4	8.13	30	68.9
13	1	0	17.4	1.34	1	0	3	9	71.0	7.89	34	44.5
12	1	1	14.2	1.18	1	0	3	8	42.4	5.30	34	64.4
12	1	0	19.4	1.62	0	0	5	7	63.8	9.11	31	39.3
14	0	0	18.1	1.29	0	0	1	11	89.8	8.16	31	42.8
3	0	2	5.8	1.93	0	8	0	11	76.4	6.95	30	57.3
14	1	0	16.6	1.19	1	0	0	13	93	7.15	36	32.8
12	0	0	16.0	1.33	0	1	2	11	74.6	6.78	30	64.0

BA = born alive; BD = born dead; MU = number of mummies; WL = litter weight at birth; INW = piglet individual weight; DON = donations; ADO = adoptions; MOR = mortality; WEA = weaned piglets; WWE = litter weight at weaning; IWW = individual weight at weaning; LAD = lactating days; AVT = 28 days average temperature.

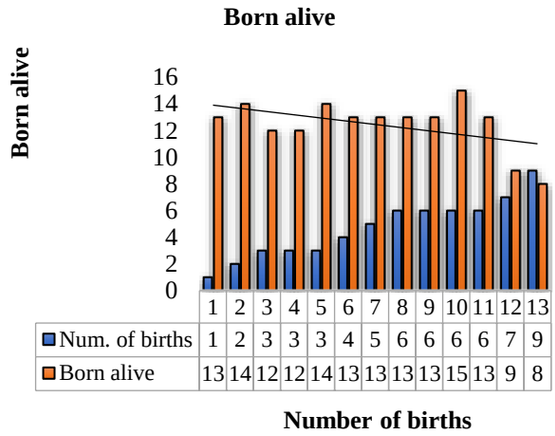
Table 1 Production record of 13 sows from birth to weaning



Graph 1 Regression between the number of births and piglets born alive

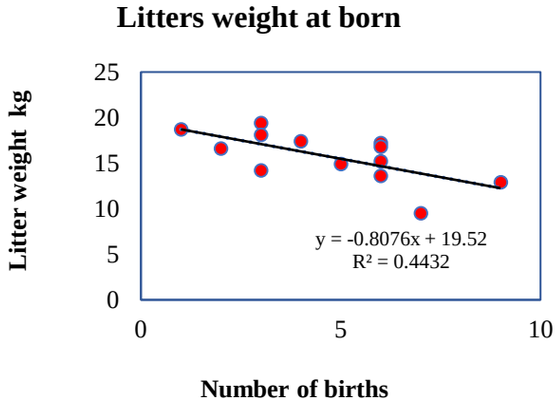
Graph 1 shows the equation and the regression line between the sows' number of births and piglets born alive. The image indicates that as the number of sows' deliveries increases, the amount of live piglets born decreases.

ISSN 2410-3438

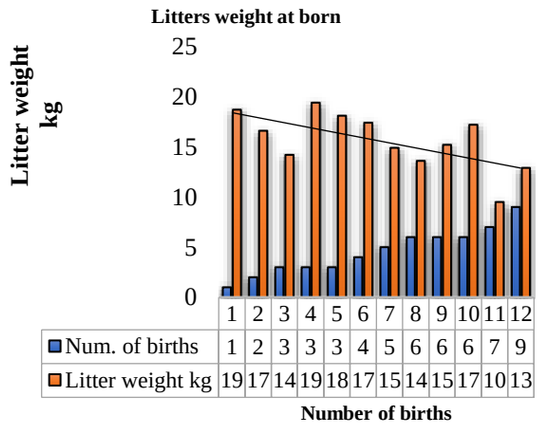


Graph 2 Number of births and piglets born alive

The determination coefficient indicates that the number of births influenced negatively in 31% the amount of piglets born alive. Bars in graph 2 represents the relation between the number of births and the number of piglets alive. Although there is certain variation for piglets born alive and the number of the sow births, there is a negative trend, particularly from the seventh delivery onwards.

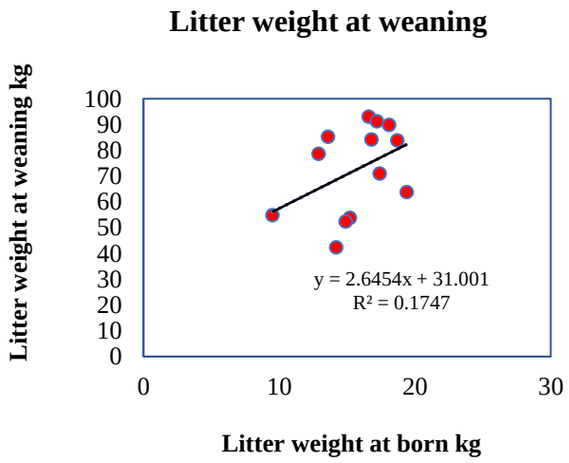


Graph 3 Regression between the number of births and litter weight at born



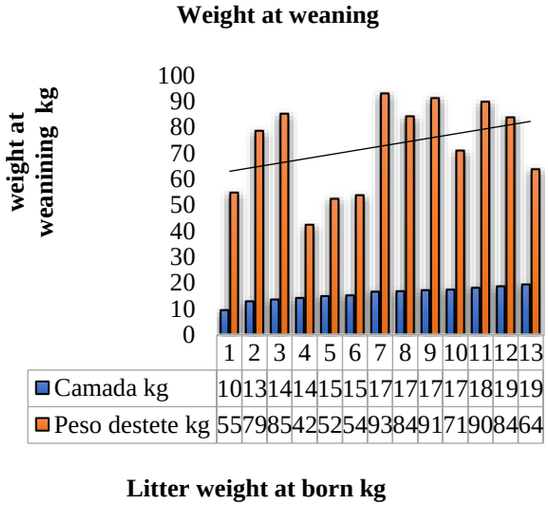
Graph 4 Number of births and litters' weight

Graphs 3 and 4 shows the equation and the regression line between the sows' number of births and the litter weight. This information indicates as the number of the sows' deliveries increases, the piglets' weight decreases. With the exception of litter 10, form the fifth birth onwards the litter weight declines. The coefficient of determination indicates that 44% of the litters' weight received a descendent effect according to the sows' births.



Graph 5 Litter weight at born and at weaning

Graphs 5 and 6 have the equation and the regression line between litters weight and weight at weaning. There is a positive but weak influence of the litters' weight at birth in the piglets' weight at weaning, as shown by the determination coefficient ($R^2 = 0.1747$).

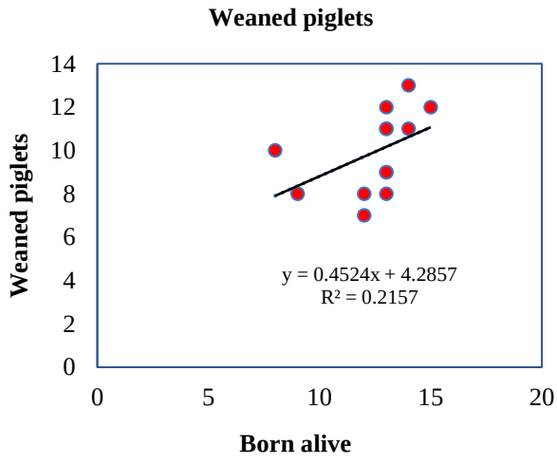


Graph 6 Regression between litters' weight at born and weight at weaning

Graph 6 represents the relation between litters weight and the respective weight at weaning.

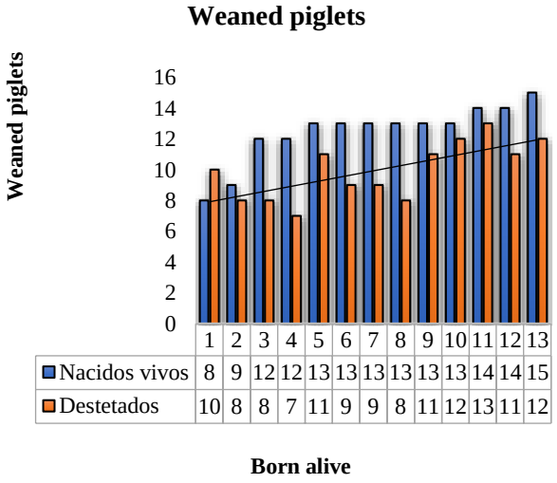
ISSN 2410-3438

Variation observed between litters, particularly litters 4, 5, 6 and 13, makes the R^2 coefficient low, although there is a positive trend, as seen in the adjusted values of the line.



Graph 7 Regression between weaned piglets and born alive

Graph 7 and 8 contains the equation and the regression line between number of piglets born alive and weaned animals. There is a positive trend when observing the number of piglets weaned and the ones born alive, although there is a weak $R^2 = 0.2157$.



	1	2	3	4	5	6	7	8	9	10	11	12	13
Nacidos vivos	8	9	12	12	13	13	13	13	13	13	14	14	15
Destetados	10	8	8	7	11	9	9	8	11	12	13	11	12

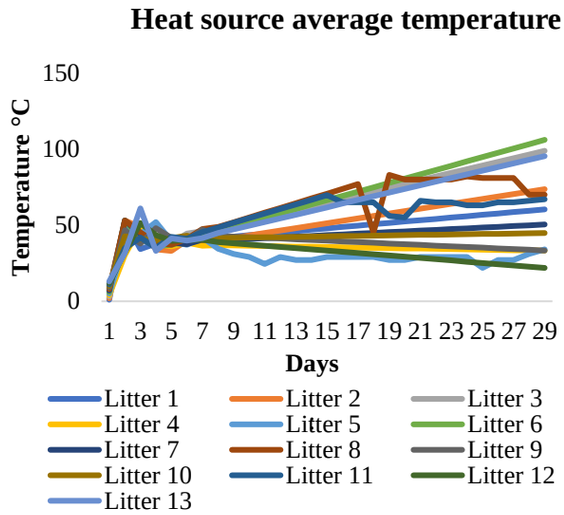
Graph 8 Regression between piglets born alive and weaned animals

Table 2 contains the multiple regression equations and their respective determination coefficients, to illustrate the interrelationship between initial aspects of sows' litters, as indicated in the table, with the rest of the variables recorded. R^2s obtained indicate an influence of the four variables (b1 to b4) at the start of life of piglets on subsequent production characteristics.

y trait	$\hat{y} = a+b1(x)+b2(x)+b3(x)+b4(x)$	R^2
MOR	$\hat{y}$ =19.62-0.18-1.74+1.18 -11.0	0.43
WEP	$\hat{y}$ =-7.48+0.14+1.86-1.21 +10.3+0.32	0.32
IWW	$\hat{y}$ =6.93-0.15-0.31+0.33-0.20 +0.47	0.47
LAC	$\hat{y}$ =13.8-0.11+2.2-1.29+9.16 +0.42	0.42

MOR= mortality; WEP= weaned piglets; IWW= individual weaned weight; LAC= lactating days

Table 2 Multiple regressions between deliveries (b1), piglets born alive (b2), piglets' litter weight at birth (b3), and individual weight at birth (b4) with other productive traits related (y's)



Graph 9 Twenty eight days average temperatures in the maternity area

Graph 9 shows ambient temperatures in the maternity area for the first 28 days of the lactating period. Table 3 contains correlation coefficients between such average temperatures and productive traits recorded.

	Mean temp.	Min. temp.	Mid-point temp.	Max. temp.
MOR	-0.15	-0.04	0.40	0.20
WEP	0.05	-0.23	-0.08	0.35
NW	0.31	-0.25	-0.08	0.40
IWW	-0.52	0.04	0.30	0.45
LAC	-0.40	0.23	0.30	-0.10

MOR= mortality; WEP= weaned piglets; IWW= individual weaned weight; LAC= lactating days.

Table 3 Correlation coefficients between productive traits and ambient temperatures

ISSN 2410-3438

Discussion

Having regard to the information on all sows in this study, litter size and weight at birth, can be considered relatively good results compared to those obtained in other studies (Segura *et al.*, 2007). Although differences are not great, it is evident the greater the number of births, the smaller the litter; in studies where the difference in favor of females with two or more deliveries has been significant, it has been explained due to the largest uterine size capacity, ovulation rate and embryo survival (Ruiz and Johnson, 2001). In a similar way, Koketsu *et al.* (1997), observed a better reproductive performance in sows with equal or more than two deliveries compared to the ones in first birth. PDE was different due to the amount of piglets per sow, nevertheless, their weight allows the pig to be viable after breastfeeding.

This condition can be result of similar management and nutrition during the lactating period, and the homogenization of litters at delivery. When weaning at same age, the litter weight depends primarily on the number of weaned piglets. Ordaz *et al.* (2013) reported an increased PCD ($p>0.05$) as the lactating period was prolonged, in 40.32 ± 1.35 kg with lactations of one to seven days, until 64.01 ± 0.52 kg with lactation periods equal or larger than 29 days.

The influence of temperature in the maternity facility, showed mid-point temperature was the more stable one during the period studied, and in addition to maximum temperature, provided the best correlation with most productive traits. Results indicated there is a definite correlation between deliveries, piglets born alive, piglets' litter weight at birth, and individual weight at birth with other important variables studied, occurring at the end of the studied production cycle. These results are similar to those found by Murillo *et al*. (2007), which means these indicators are useful to evaluate the maternal productivity, increasing in each delivery.

The use of thermography to determine the temperature of heat sources, contributes to minimizing the handling of piglets, allowing them to devote their energy to finding their optimum temperature and food. This allows a faster growth in grams day^{-1}.

It also allows environmental suitability for sows, as an inadequate temperature can decrease milk production, among other things. The knowledge on the environmental condition of the females will allow improving the efficiency of the biological behavior of sows and piglets.

References

Batista, L., 2000. Importancia de la evaluación de parámetros reproductivos. V *Simposium Internacional de Reproducción e Inseminación Artificial en Porcinos*. Ed. Alberto Stephano. 109-114.

Flores, J. J., y Gómez M. A., 1995. Alternativas para el desarrollo de la porcicultura jalisciense. La producción porcícola en México. Contribución al desarrollo de una visión integral. *Luis Kato Coordinador. UAM.*

Kato, L., 1995. Producción porcícola intensiva. En Kato, L. La producción porcícola en México. *Universidad Autónoma Metropolitana*. México 30-34.

Koketsu Y, Dial GD, King VL. 1997. Influence of various factors on farrowing rate on farms using early weaning. J Anim Sci. 75:2580-2587. INEGI (Instituto Nacional de Estadística Geografía e Historia). Censo Agropecuario 2016.

Murillo, C.; Herradora, M. A.; Martínez, R. 2007. Relación entre la pérdida de grasa dorsal de cerdas lactantes con el consumo de alimento, tamaño de la camada, peso de los lechones al destete y días de lactancia.Rev. Cientif. FCV-LUZ. XVII (4): 380-385.

Ordaz-Ochoa, G.; Juárez-Caratachea, A.; García-Valladares, A.; Pérez-Sánchez, R.E.; Ortiz- Rodríguez, R. 2013 Efecto del número de parto sobre los principales indicadores reproductivos de las cerdas. Rev. Cientif. FCV-LUZ. XXIII (6):511-519. 2013.

Proyecto de Ordenamiento Territorial del Estado de Jalisco (POET). 1998. Descripción y diagnóstico de la actividad pecuaria en Jalisco. Grupo pecuario. *Universidad de Guadalajara*. 11-15.

ISSN 2410-3438

Ruiz-Flores A, Johnson RK. 2001. Direct and correlated responses to two-stage for ovulation rate and number of fully formed pigs in swine. J Anim Sci. 79:2286-2297.

Secretaría de Agricultura y Desarrollo Social (SADER) Servicio de Información Agroalimentaria y Pesquera (SIAP) https://nube.siap.gob.mx/cierre_pecuario/

Segura CJC, Alzina-López A, Solorio RJL. 2007. Evaluación de tres modelos y factores de riesgo asociados a la mortalidad de lechones al nacimiento en el trópico de México. Tec Pecu Méx. 45:227-236.

ISSN 2410-3438

Article **Journal of Quantitative and Statistical Analysis**
June, 2021 Vol.8 No.21 25-30

Use of the effectiveness of cuachalalate from healing plants in Mezquitic, Jalisco

Uso de la efectividad de cuachalalate de las plantas curativas en Mezquitic, Jalisco

GONZÁLEZ-GARCÍA, Arcelia, HERNÁNDEZ-SALAS, Claudia, MARTÍNEZ-ORTIZ, Rosa María and GONZÁLEZ-MARTÍNEZ, Lilia

[1]Universidad Autónoma de Zacatecas-Licenciatura en Químico Farmacéutico Biólogo, Licenciatura en Enfermería UAE.
[2]Unidad Académica de Odontología. Campus UAZ. Siglo XXI Carretera Zacatecas-Km.6 Ejido La Escondida Zacatecas Zac.C.P.98160

ID 1st Author: *Arcelia, González-García*

ID 1st Author: *Claudia, Hernández-Salas*

ID 2nd Co-author: *Rosa María, Martínez-Ortiz*

ID 3rd Co-author: *Lilia, González-Martínez*

DOI: 10.35429/JQSA.2021.22.8.25.30 Received January 05, 2021; Accepted June 30, 2021

Abstract

Cuachalalate is (*Amphiterigium adstringens*)an arboreal species of the low deciduous forest whose bark has been used since pre-Hispanic times in traditional medicine because it is one of the products of greatest demand of Mexican herbalism and the economic importance for the communities that live in this ecosystem in the inifap Since 1989, research has been carried out to improve its management, which currently allows its sustainable and legal use. This study was carried out in Jalisco Zacatecas whose objective was to know the use of the curative plants of cuachalalate with traditional medicine as preventive of a scarring of some of the wounds as well as the most frequently consumed and the results indicate that the main areas or fronts of the cortex are from the State of Jalisco for the human

Use of Healing Plants, Medicinal

Resumen

El cuachalalate, (*Amphipterigium adstringens*), es una especie arbórea de la selva baja caducifolia, cuya corteza se utiliza desde tiempos prehispánicos en la medicina tradicional. Por ser uno de los productos de mayor demanda en la herbolaria mexicana y de importancia económica para las comunidades que viven en este ecosistema, en el INIFAP desde 1989 se han realizado investigaciones para mejorar su manejo lo que permite actualmente su aprovechamiento sustentable y legal. Se llevó acabo este estudio en Mezquitic, Jalisco, cuyo objetivo fue conocer el uso de las plantas curativas del cuachalalate como medicina tradicional, como preventivo de una cicatrización de algunas de las heridas así como también el uso más frecuente de consumo. Resultados indican que las principales áreas oferentes de corteza son el estado de Jalisco en el ser humano.

Uso de plantas curativas, Medicinales

Citation: GONZÁLEZ-GARCÍA, Arcelia, HERNÁNDEZ-SALAS, Claudia, MARTÍNEZ-ORTIZ, Rosa María and GONZÁLEZ-MARTÍNEZ, Lilia. Use of the effectiveness of cuachalalate from healing plants in Mezquitic, Jalisco. Journal of Quantitative and Statistical Analysis. 2021. 8-22: 25-30

* Correspondence to Author (e-mail: arcelia2009@ live.com.mx)
† Researcher contributing as first author.

Introduction

Healing plants today have been the traditional alternative medicine, in the effectiveness in the ethnomedicine of Mezquitic, Jalisco, the medicine as an alternative. In line with the prevention, cure and healing, which was carried out by the ethnographic method, to obtain information on perceptions, of the healing plants that today are present in the environment, which are used in traditional medicine, is of great help with an efficiency of consuming them for a health and economic benefit, due to the environment, traditional medicine according to the world health organization (WHO), traditional medicine is a set of knowledge, belief skills , experiences, values in curative preventive medicine as an improvement in the treatment of mental physical illnesses in societies, ethnomedicine refers to the medical study of traditional typical plants, they focus on the interpretation of culture and health, medicinal care processes, healing and health (Williams, 2006) the belief of healing plants for the health of society, making it interdisciplinary.

Traditional medicine has been studied from the medical anthropology that studies the problems of human health and healing systems in their social, cultural and economic contexts; analyzes the mediations that explain the differential ways of falling ill, being cared for and dying among certain individuals and groups, considers the characteristics and peculiarities of the relationships between people and social groups that enable or limit the resolution of their health problems (Freyermuth and Sesia, 2006 , P. According to Aparicio (2004): “five fundamental features define traditional medicines: 1. Validity as ethnomedicine (therapeutic system adapted to a specific socio-cultural and geographical environment and context that responds to health needs.

In Mexico, traditional medicine is recognized by the Ministry of Health, the Mexican Institute for the Study of Medicinal Plants and the Commission for the Use and Conservation of Biodiversity, which has a digital Library on medicinal plants. In traditional Mexican medicine, the use of plants is of great importance, for example, it is practiced by its 60 ethnic groups, who use more than 5,000 plants (González-Stuart and Rivera, 2009 in Juárez-Rosete et al., 2013). In urban areas, traditional medicine is used, mainly through Herbalism (Osuna et al., 2005). Medicinal plants in Mexico are mostly wild (Osuna et al., 2005), which is related to the local environmental knowledge of the country's ethnic groups and that they are not only collected for internal use, but some species are cultivated for export. Traditional medicine is based on local environmental knowledge, it is adaptive to spatio-temporal contexts and its owners are mainly native peoples. This practice is older than any other therapy: the systematic consumption of plants with medicinal attributes probably dates back to 2 million years in Africa (Chifa, 2010).

At a global level, knowledge, relevance and is considered a priority by institutions such as: International Organization for Intellectual Property, World Health Organization, International Labor Organization and various United Nations dependencies (United Nations Food and Drug Administration Organization). Agriculture, United Nations Educational, Scientific and Cultural Organization; United Nations Environment Program; Commission on Rights At the same time they contribute to the conservation of traditional cultures today in the world and the biodiversity, as part of sustainable development.

Currently, society looks for options for good health and physical and psychological well-being, especially through traditional medicine. It is characterized because every therapeutic act contains a merely empirical part, which is accompanied by a magical-religious intention, medicinal healing plants are characterized by an integral approach to health shows that the human being as a cultural and social being is based on local environmental knowledge, includes botanical, zoological, ecological and technological knowledge (Fagetti, 2011) There is interest in traditional medicine due to the search for alternatives to allopathic medicine to treat diseases with natural and non-synthetic remedies, to be cared for in a friendly way and in a trusting relationship between patient and doctor. Cuachalalate. Amphipterygium adstringens Schiede ex Schlech. (*Anacardiaceae*) with popular synonymy chalalate, coachalalate, volador, cuachalalá or cuachinala and botanical synonymy Juliania adstringens Schlechter. It is a tree 10 m high, with a twisted trunk, grayish bark, with scales.

ISSN 2410-3438

The leaves on the front are dull green and on the back they are grayish, grouped at the tips of the branches. The flowers are solitary or in clusters. One of the most important uses that cuachalalate has within Traditional Mexican Medicine is the treatment of diseases related to the digestive system, to corroborate this activity (Navarrete in 1998), carried out the corresponding studies to evaluate the gastro-protective effect of the extract of the bark of A. adstringensis in ulcers, as well as the healing, Currently, Traditional Mexican Medicine is recognized as a resource for the preservation of the health of millions of human beings, it is an important part of the worldview of indigenous peoples, representing the knowledge about mother earth and the use of medicinal plants.

This wisdom, day by day, has been acquiring more and more importance, due to its strengthening in order to preserve its identity. Within this plant treasure, is Cuachalalate, which has been widely used in Traditional Medicine as a gastroprotective agent, where the stem is used in the form of an infusion to treat stomach ulcers, this due to the decrease in the secretion of gastric juices. , helping the renewal of the epithelium. Bark for wound healing Previous studies show that the biological activity is due to the presence of triterpenes, such as masticadienonic acid and some of its derivatives.

The tree highlights the importance of this native tree, in Mezquitic Jalisco as, in its use in the alternative that is curative and preventive for health, these plants as well as for digestive problems, as well as the chemical profile to know the secondary metabolites responsible for their biological activity. The uses of cuachalalate the bark of the tree hardens the gums, relieves fires and mouth ulcers, there is also a decrease in urinary discomfort and affections to the kidneys and wound healing, especially in the skin situation, grains or deep wounds, relieves the stomach inflammation, chronic gastritis, gastric ulcers according to a group of scientists from the department of natural products of the Institute of Chemistry, has been investigating that cuachalalate is capable of inoculating human cancer cells in a less toxic way, more effective than drugs conventional anticancer agents. Patricia Bañuelos told us that the pieces of bark from a tree that they brought were "cuachalalá" (known in other places as cuachalalate). It is for medicinal use.

The UNAM Digital Library of Traditional Mexican Medicine states that this vegetable is used to treat ulcers, wounds and gastritis, among other ailments. Being one of the municipalities in which the Wixárika culture is based, the handicrafts made with beads could not be missing in this sample. Necklaces, earrings, rings and keychains with designs of contrasting figures and colors. Trays made with light sabino wood were also offered wixárika cheered the old mansion in downtown Guadalajara with his music that houses the Museum of Popular Arts in Jalisco. In addition to the lack of economic resources to have access to modern medicine, the change in needs and beliefs contribute to the practice and the rise of alternative, traditional medicine remains (Zhang, 2004), such as the preference for treating diabetes with medicinal plants (Warjeet, 2011).

Figure 1 Cuachalalate
Juan Carlos Núñez

Justification

Ethnomedicine, ethnobotanical medicine, has presented us with a great impact that traditional medicine today is very effective as a preventive of scarring, deep wounds, since it has great effectiveness in traditional alternative medicine.

Methodology

The research was carried out by the longitudinal transversal method. In this, a random sample of 60 people in ages ranging from 50 to 70 years, both sexes, was considered. A questionnaire was applied through an open interview, which was supported by both qualitative and quantitative participatory research. It includes the following questions:

ISSN 2410-3438

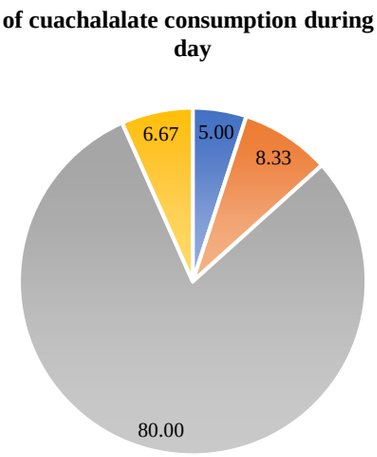
1. How many times do you consume cuachalalate during the day?
2. What is the purpose of consuming the cuachalalate in water?
3. What type of diseases does cuachalalate treat?
4. Do you recommend the use of cuachalalate in people who have a wound or who have undergone surgeries to accelerate the healing process'?
5. What other plants do you use to prevent and cure diseases?

Results

Among the results obtained in the sample considered and that refer to the question about the daytime consumption of cuachalalate, the following was found:

Consumption during the day	Number of people	% of people who use it daily
1 time	3	5,00
2 times	5	8,33
3 times	48	80,00
More than 3 times	4	6,67
Total	60	100,00

Table 1



Graph 1

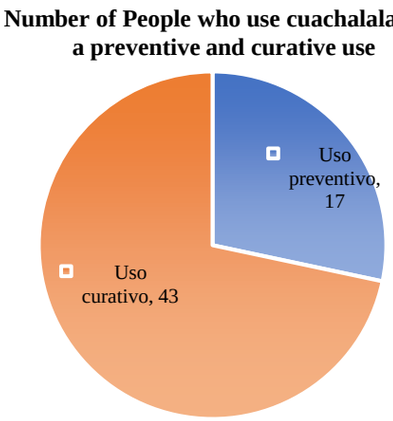
In the graph it can be seen that 80% of people who consume it do so 3 times a day, which suggests that its consumption is like water for use. It is important to mention that the people interviewed suggest its use in a preventive and curative way. They name its use for wound healing and preventing cancer.

Of the total of individuals with whom they spoke, 28. 33% mention making use of cuachalalate in a preventive way, while 71. 66% do it to cure.

	Number of people	% of use
Preventive use	17	28,33
Curative use	43	71,67
Total	60	100

Table 2

In the following graph it can be seen that 43 people of the total that make up the sample, use this medicinal plant to heal.



Graph 2

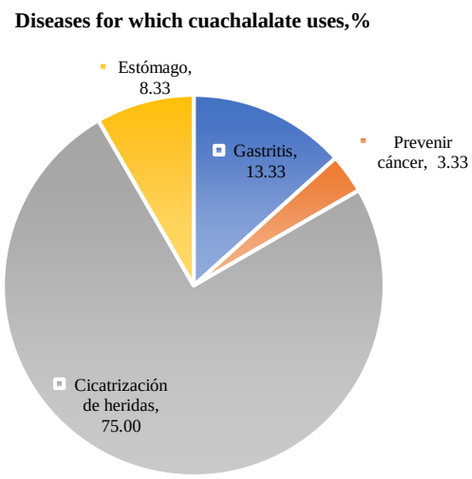
Among the diseases that are treated with cuachalalate, gastritis, preventing cancer, wound healing, stomach, etc. are mentioned.

Diseases	Number of people	% use in diseases
Gastritis	8	13,33
Prevent Cancer	2	3,33
Wound Healing	45	75,00
Stomach	5	8,33
Total	60	100

Table 3

Of these diseases, the majority number is represented by wound healing, with 45 people from the total sample. This implies 75%, which can be seen in the following graph.

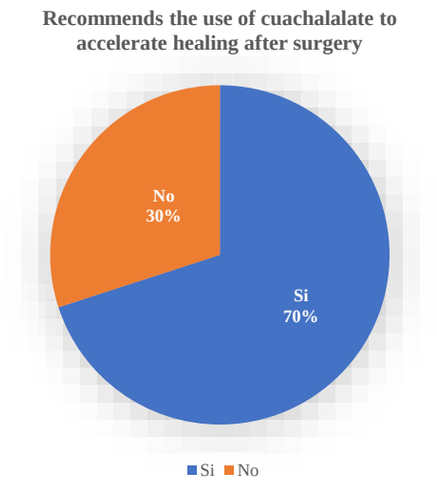
ISSN 2410-3438



Graph 3

It can also be perceived that only 3.33% use it to prevent some type of cancer, so its impact lies in the healing of wounds. The foregoing encourages us to ask, then, if its use is recommended after surgery to accelerate the healing process or in minor wounds.

For this, 42 interviewees indicate that they do recommend it. This is represented by 70% of the sample.

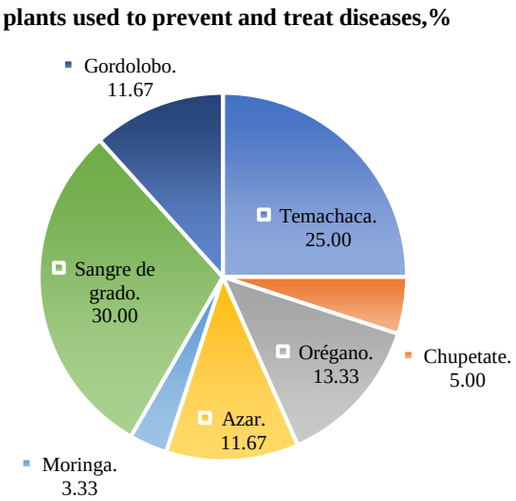


Graph 4

It is important to mention that they also consider other types of plants to prevent and cure diseases, within these are mentioned:

Plant	Persons	% people who use it
Temachaca	15	25,00
Lick yourself	3	5,00
Oregano	8	13,33
Random	7	11,67
Moringa	2	3,33
Grade blood	18	30,00
Mullein	7	11,67
Total	60	100

Table 3



Graph 5

Apart from the cuachalate, they consider the blood of the grade and the temachaca as plants of use.

Conclusions

The bark of the cuachalalate that is commercialized in Mexico, is mainly collected in the municipality of Mezquitic, Jalisco.

Acknowledgments

A thank you to the authorities of the Municipality of Mezquitic Jalisco Zacatecas

References

Caldentey, P. 1986. Comercialización de productos agrarios. Editorial Agrícola Española. Madrid, España. 199p.

Mendoza, G. 1987. Compendio de mercadeo de productos agropecuarios. Instituto Interamericano de Cooperación para la Agricultura. San José Costa Rica, Costa Rica. 343 p.

Olivera O., A. G., M. Soto H., M. Martínez V., T. Terrazas S. y F. Solares A. 1999. Phytochemical study of Cuachalalate (A*mphipterygium adstringens,* Schiede ex Schlecht.) Elsevier Science Ireland. Journal of Ethnopharmacology 68 :109-113.

Solares A., F. 1999. Estudio Químico-Biológico para una producción sostenida de corteza de cuachalalate para uso farmacéutico-industrial en el estado de Morelos. Informe Técnico. INIFAP-SAGARPA. Campo Experimental Zacatepec. Zacatepec, Morelos. 37p.

Solares A., F. y Ma. C. Gálvez C. 2002. Diagnóstico biofísico y socioeconómico de las Áreas Naturales Protegidas de jurisdicción estatal en Morelos. CD1.INIFAP-CEAMA-Morelos. Morelos, México.168p.

Solares A., F. y Ma. C. Gálvez C. 2002. Manual para una producción sustentable de corteza de Cuachalalate (*Amphipterygium adstringens* Schiede ex Schlecht.). Publicación especial No. 34. INIFAP-SAGARPA. Campo Experimental "Zacatepec". Zacatepec, Mor. México. 15 p.

Solares A., F., J. Jasso M., J. Vargas H., M. Soto H. y C. Rodríguez F. 2006. Capacidad de regeneración en grosor y lateral en corteza de cuachalalate (*Amphipterygium adstringens* Schiede ex Schlecht) en el estado de Morelos. RA XIMHAI. 2:481-495.

Solares A., F. 2007. Alineación de la investigación para la conservación y aprovechamiento sustentable del cuachalalate (*Amphipterygium adstringens* Schiede ex Schlecht.). Informe Técnico. INIFAP-SAGARPA. Campo Experimental "Zacatepec". Zacatepec, Mor. México. 30 p.

Solares A., F. 2009. Aprovechamiento sustentable de cuachalalate (*Amphipterygium adstringens* Schiede ex Schlecht.) bajo un marco legal: Tecnología INIFAP. In: Memoria. IV Reunión Nacional de Innovación Agrícola y Forestal. Saltillo, Coah. México. pp 315.

ISSN 2410-3438

Title in Times New Roman and Bold No. 14 in English and Spanish]

Surname (IN UPPERCASE), Name 1st Author†*, Surname (IN UPPERCASE), Name 1st Coauthor, Surname (IN UPPERCASE), Name 2nd Coauthor and Surname (IN UPPERCASE), Name 3rd Coauthor

Institutional Affiliation of Author including Dependency (No.10 Times New Roman and Italic)

International Identification of Science - Technology and Innovation

ID 1st Author: (ORC ID - Researcher ID Thomson, arXiv Author ID - PubMed Author ID - Open ID) and CVU 1st author: (Scholar-PNPC or SNI-CONACYT) (No.10 Times New Roman)

ID 1st Coauthor: (ORC ID - Researcher ID Thomson, arXiv Author ID - PubMed Author ID - Open ID) and CVU 1st coauthor: (Scholar or SNI) (No.10 Times New Roman)

ID 2nd Coauthor: (ORC ID - Researcher ID Thomson, arXiv Author ID - PubMed Author ID - Open ID) and CVU 2nd coauthor: (Scholar or SNI) (No.10 Times New Roman)

ID 3rd Coauthor: (ORC ID - Researcher ID Thomson, arXiv Author ID - PubMed Author ID - Open ID) and CVU 3rd coauthor: (Scholar or SNI) (No.10 Times New Roman)

(Report Submission Date: Month, Day, and Year); Accepted (Insert date of Acceptance: Use Only ECORFAN)

Abstract (In English, 150-200 words)

Objectives
Methodology
Contribution

Keywords (In English)

Indicate 3 keywords in Times New Roman and Bold No. 10

Abstract (In Spanish, 150-200 words)

Objectives
Methodology
Contribution

Keywords (In Spanish)

Indicate 3 keywords in Times New Roman and Bold No. 10

Citation: Surname (IN UPPERCASE), Name 1st Author, Surname (IN UPPERCASE), Name 1st Coauthor, Surname (IN UPPERCASE), Name 2nd Coauthor and Surname (IN UPPERCASE), Name 3rd Coauthor. Paper Title. Journal of Quantitative and Statistical Analysis. Year 1-1: 1-11 [Times New Roman No.10]

* Correspondence to Author (example@example.org)
† Researcher contributing as first author.

Introduction

Text in Times New Roman No.12, single space.

General explanation of the subject and explain why it is important.

What is your added value with respect to other techniques?

Clearly focus each of its features

Clearly explain the problem to be solved and the central hypothesis.

Explanation of sections Article.

Development of headings and subheadings of the article with subsequent numbers

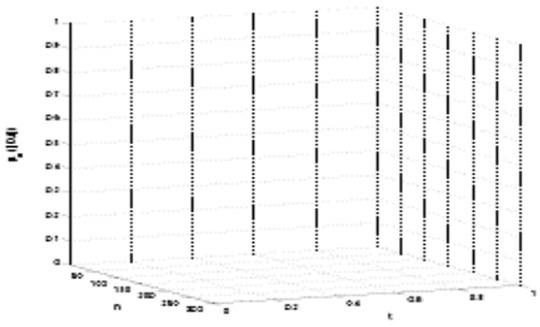
[Title No.12 in Times New Roman, single spaced and bold]

Products in development No.12 Times New Roman, single spaced.

Including graphs, figures and tables-Editable

In the article content any graphic, table and figure should be editable formats that can change size, type and number of letter, for the purposes of edition, these must be high quality, not pixelated and should be noticeable even reducing image scale.

[Indicating the title at the bottom with No.10 and Times New Roman Bold]



Graphic 1 Title and *Source (in italics)*

Should not be images-everything must be editable.

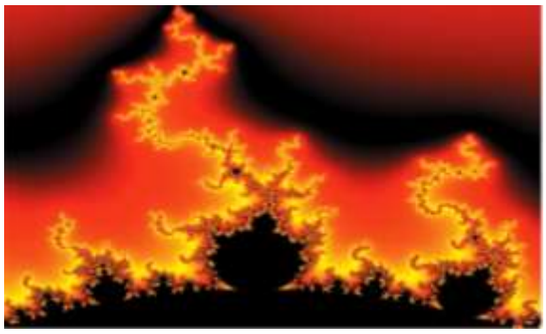


Figure 1 Title and *Source (in italics)*

Should not be images-everything must be editable.

Table 1 Title and *Source (in italics)*

Should not be images-everything must be editable.

Each article shall present separately in **3 folders**: a) Figures, b) Charts and c) Tables in .JPG format, indicating the number and sequential Bold Title.

For the use of equations, noted as follows:

$$Y_{ij} = \alpha + \sum_{h=1}^{r} \beta_h X_{hij} + u_j + e_{ij} \qquad (1)$$

Must be editable and number aligned on the right side.

Methodology

Develop give the meaning of the variables in linear writing and important is the comparison of the used criteria.

Results

The results shall be by section of the article.

Annexes

Tables and adequate sources

Thanks

Indicate if they were financed by any institution, University or company.

Conclusions

Explain clearly the results and possibilities of improvement.

ISSN 2410-3438

References

Use APA system. Should not be numbered, nor with bullets, however if necessary numbering will be because reference or mention is made somewhere in the Article.

Use Roman Alphabet, all references you have used must be in the Roman Alphabet, even if you have quoted an Article, book in any of the official languages of the United Nations (English, French, German, Chinese, Russian, Portuguese, Italian, Spanish, Arabic), you must write the reference in Roman script and not in any of the official languages.

Technical Specifications

Each article must submit your dates into a Word document (.docx):

Journal Name
Article title
Abstract
Keywords
Article sections, for example:

1. Introduction
2. Description of the method
3. Analysis from the regression demand curve
4. Results
5. Thanks
6. Conclusions
7. References

Author Name (s)
Email Correspondence to Author
References

Intellectual Property Requirements for editing:

-Authentic Signature in Color of Originality Format Author and Coauthors

-Authentic Signature in Color of the Acceptance Format of Author and Coauthors

ISSN 2410-3438

Reservation to Editorial Policy

Journal of Quantitative and Statistical Analysis reserves the right to make editorial changes required to adapt the Articles to the Editorial Policy of the Journal. Once the Article is accepted in its final version, the Journal will send the author the proofs for review. ECORFAN® will only accept the correction of errata and errors or omissions arising from the editing process of the Journal, reserving in full the copyrights and content dissemination. No deletions, substitutions or additions that alter the formation of the Article will be accepted.

Code of Ethics - Good Practices and Declaration of Solution to Editorial Conflicts

Declaration of Originality and unpublished character of the Article, of Authors, on the obtaining of data and interpretation of results, Acknowledgments, Conflict of interests, Assignment of rights and Distribution

The ECORFAN-Mexico, S.C Management claims to Authors of Articles that its content must be original, unpublished and of Scientific, Technological and Innovation content to be submitted for evaluation.

The Authors signing the Article must be the same that have contributed to its conception, realization and development, as well as obtaining the data, interpreting the results, drafting and reviewing it. The Corresponding Author of the proposed Article will request the form that follows.

Article title:

– The sending of an Article to Journal of Quantitative and Statistical Analysis emanates the commitment of the author not to submit it simultaneously to the consideration of other series publications for it must complement the Format of Originality for its Article, unless it is rejected by the Arbitration Committee, it may be withdrawn.

– None of the data presented in this article has been plagiarized or invented. The original data are clearly distinguished from those already published. And it is known of the test in PLAGSCAN if a level of plagiarism is detected Positive will not proceed to arbitrate.

– References are cited on which the information contained in the Article is based, as well as theories and data from other previously published Articles.

– The authors sign the Format of Authorization for their Article to be disseminated by means that ECORFAN-Mexico, S.C. In its Holding Bolivia considers pertinent for disclosure and diffusion of its Article its Rights of Work.

– Consent has been obtained from those who have contributed unpublished data obtained through verbal or written communication, and such communication and Authorship are adequately identified.

– The Author and Co-Authors who sign this work have participated in its planning, design and execution, as well as in the interpretation of the results. They also critically reviewed the paper, approved its final version and agreed with its publication.

– No signature responsible for the work has been omitted and the criteria of Scientific Authorization are satisfied.

– The results of this Article have been interpreted objectively. Any results contrary to the point of view of those who sign are exposed and discussed in the Article.

Copyright and Access

The publication of this Article supposes the transfer of the copyright to ECORFAN-Mexico, SC in its Holding Bolivia for it Journal of Quantitative and Statistical Analysis, which reserves the right to distribute on the Web the published version of the Article and the making available of the Article in This format supposes for its Authors the fulfilment of what is established in the Law of Science and Technology of the United Mexican States, regarding the obligation to allow access to the results of Scientific Research.

Article Title:

Name and Surnames of the Contact Author and the Coauthors	Signature
1.	
2.	
3.	
4.	

Principles of Ethics and Declaration of Solution to Editorial Conflicts

Editor Responsibilities

The Publisher undertakes to guarantee the confidentiality of the evaluation process, it may not disclose to the Arbitrators the identity of the Authors, nor may it reveal the identity of the Arbitrators at any time.

The Editor assumes the responsibility to properly inform the Author of the stage of the editorial process in which the text is sent, as well as the resolutions of Double-Blind Review.

The Editor should evaluate manuscripts and their intellectual content without distinction of race, gender, sexual orientation, religious beliefs, ethnicity, nationality, or the political philosophy of the Authors.

The Editor and his editing team of ECORFAN® Holdings will not disclose any information about Articles submitted to anyone other than the corresponding Author.

The Editor should make fair and impartial decisions and ensure a fair Double-Blind Review.

Responsibilities of the Editorial Board

The description of the peer review processes is made known by the Editorial Board in order that the Authors know what the evaluation criteria are and will always be willing to justify any controversy in the evaluation process. In case of Plagiarism Detection to the Article the Committee notifies the Authors for Violation to the Right of Scientific, Technological and Innovation Authorization.

Responsibilities of the Arbitration Committee

The Arbitrators undertake to notify about any unethical conduct by the Authors and to indicate all the information that may be reason to reject the publication of the Articles. In addition, they must undertake to keep confidential information related to the Articles they evaluate.

Any manuscript received for your arbitration must be treated as confidential, should not be displayed or discussed with other experts, except with the permission of the Editor.

The Arbitrators must be conducted objectively, any personal criticism of the Author is inappropriate.

The Arbitrators must express their points of view with clarity and with valid arguments that contribute to the Scientific, Technological and Innovation of the Author.

The Arbitrators should not evaluate manuscripts in which they have conflicts of interest and have been notified to the Editor before submitting the Article for Double-Blind Review.

Responsibilities of the Authors

Authors must guarantee that their articles are the product of their original work and that the data has been obtained ethically.

Authors must ensure that they have not been previously published or that they are not considered in another serial publication.

Authors must strictly follow the rules for the publication of Defined Articles by the Editorial Board.

The authors have requested that the text in all its forms be an unethical editorial behavior and is unacceptable, consequently, any manuscript that incurs in plagiarism is eliminated and not considered for publication.

Authors should cite publications that have been influential in the nature of the Article submitted to arbitration.

Information services

Indexation - Bases and Repositories

LATINDEX (Scientific Journals of Latin America, Spain and Portugal)
RESEARCH GATE (Germany)
GOOGLE SCHOLAR (Citation indices-Google)
REDIB (Ibero-American Network of Innovation and Scientific Knowledge- CSIC)
MENDELEY (Bibliographic References Manager)

Publishing Services:

Citation and Index Identification H.
Management of Originality Format and Authorization.
Testing Article with PLAGSCAN.
Article Evaluation.
Certificate of Double-Blind Review.
Article Edition.
Web layout.
Indexing and Repository
ArticleTranslation.
Article Publication.
Certificate of Article.
Service Billing.

Editorial Policy and Management

21 Santa Lucía, CP-5220. Libertadores -Sucre–Bolivia. Phones: +52 1 55 6159 2296, +52 1 55 1260 0355, +52 1 55 6034 9181; Email: contact@ecorfan.org www.ecorfan.org

Journal of Quantitative and Statistical Analysis

2410 3438
ISSN 2410 - 3438