IAC-17,B1,5,5,x41262

EPIDEMIOLOGY STUDY OF THE CHAGAS DISEASE IN BOLIVIA USING REMOTE SENSING DATA

Natalia I. Vargas-Cuentas
Beihang University (BUAA), China; Universidad de Ciencias y Humanidades – UCH, Perú,
natalia.i.vargascuentas@ieee.org

Avid Roman-Gonzalez
Universidad de Ciencias y Humanidades - UCH, Perú, avid.roman-gonzalez@ieee.org

Alicia Alva Mantari
Universidad de Ciencias y Humanidades - UCH, Perú, aalva@uch.edu.pe

Luis AnthonyAucapuma Muñoz
Universidad de Ciencias y Humanidades - UCH, Perú, laucapuma@uch.edu.pe

Remote sensing is the technology that has enabled us to obtain information about the Earth's surface without directly contacting it. For this reason, currently, the Bolivian state has considered a list of interesting applications of remote sensing in the country, including the following: biodiversity and environment monitoring, mining and geology, epidemiology, agriculture, water resources and land use planning. The use of satellite images has become a great tool for epidemiology because with this technological advance we can determine the environment in which transmission occurs, the distribution of the disease and its evolution over time. In that context, one of the important diseases related to public health in Bolivia is Chagas disease, also known as South American Trypanosomiasis.

Chagas is caused by a blood-sucking bug or Vinchuca, which causes serious intestinal and heart long term problems and affects 33.4% of the Bolivian population. This disease affects mostly humble people, so the Bolivian state invests millions of dollars to acquire medicine and distribute it for free.

Due to the above reasons, the present research aims to analyze some areas of Bolivia using satellite images for developing an epidemiology study. The primary objective is to understand the environment in which the transmission of the disease happens, and the climatic conditions under which occurs, observe the behavior of the blood-sucking bug, identify in which months occur higher outbreaks, in which months the bug leaves its eggs, and under which weather conditions this happens. All this information would be contrasted with information extracted from the satellite images and data from the Ministry of Health, and the Institute of Meteorology in Bolivia. All this data will allow us to have a more integrated understanding of this disease and promote new possibilities to prevent and control it.

I. INTRODUCTION

Vectors are living organisms that can transmit diseases, from one infected person or animal to another [1].

The vectors cause serious diseases in the human being and are more prevalent in tropical and subtropical areas and in places with problems of access to drinking water and sanitation.

For the World Health Organization (WHO) over 17% of infectious diseases are transmitted by vectors [2]. Chagas disease (CH) is a serious public health problem because of its impact and complicated control [3].

In the Latin American region, Chagas disease is present in 20 countries, as we can observe in the next figure and currently affects 8 to 9 million people and more than 90 million people are at direct risk of contagion [4].

Fig. 1: Incidence of Chagas in South America. Source: Corazones Unidos contra el Chagas

CH is caused by the parasite Trypanosoma cruzi (TC) [2]. The vinchucas nest and multiply in houses built with precarious materials, such as adobe, mud or straw, for this reason in rural areas where this type of constructions predominate, people are at higher risk of contracting the disease [3].

The main CH infection is carried out by contact of mucous or wounded skin with the feces of the infected host animal, usually a triatomine insect known in the region as vinchuca [5]. There are other forms of contagion, for example: transfusional, congenital, organ transplants, oral and laboratory accident.
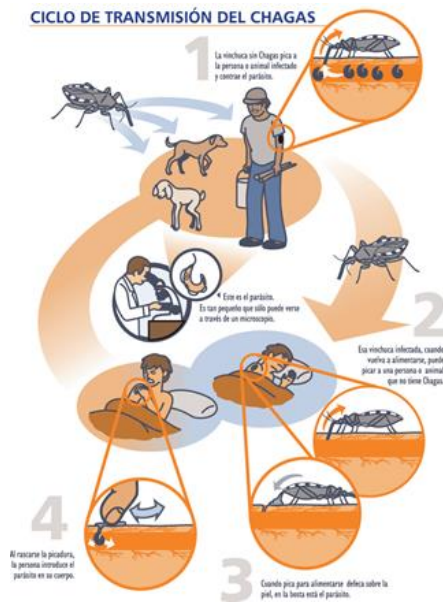

Fig. 2: Chagas disease transmission cycle
Source: Argentina Chagas Initiative

Chagas disease has two phases: acute and chronic; the first may present without symptoms or with very mild symptoms such as fever, general malaise, swelling of the eyes, swollen area and redness at the site of the bite. If Chagas is not treated in time it can cause cardiovascular, gastrointestinal or neurological disorders. 30% of patients develop heart disease that can trigger death.

Although the outbreaks of Chagas are a concern for the Ministry of Health, so far has not been possible to characterize the environmental macrofactors that condition the reproduction, development and maintenance of the mosquito vector in the country.

Therefore, it has not been possible to identify the risk areas in the country to apply preventive measures to mitigate the effects of the disease in the population and to make a better decision-making.

Seeing the problematic situation of the country with respect to Chagas disease, this project wants to analyze the relationship between the environmental factors and

the prevalence of Chagas cases in Bolivia, using meteorological data and Landsat 8 satellite imagery of 2016, to identify the environmental conditions in which the disease is transmitted.

II. CHAGAS IN BOLIVIA

There are 140 species of vinchucas in the world, in Bolivia there are 21 types of vinchucas identified, the most common vinchuca is the Triatoma Infestans, which is responsible for the largest number of Chagas cases recorded in the country.
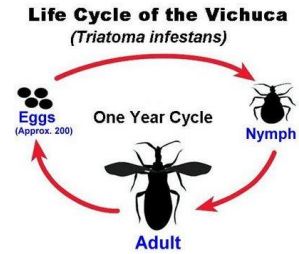

Fig. 3: Life Cycle of the Vinchuca.
Source: World Health Organization (WHO) 2010

Bolivia is the country with the largest dispersion area of the CH vector (Triatoma Infestans). It is dispersed in approximately 60% of the Bolivian territory. According to figures calculated by WHO in 2010, Bolivia is the country with the highest incidence of Chagas in South America [6].
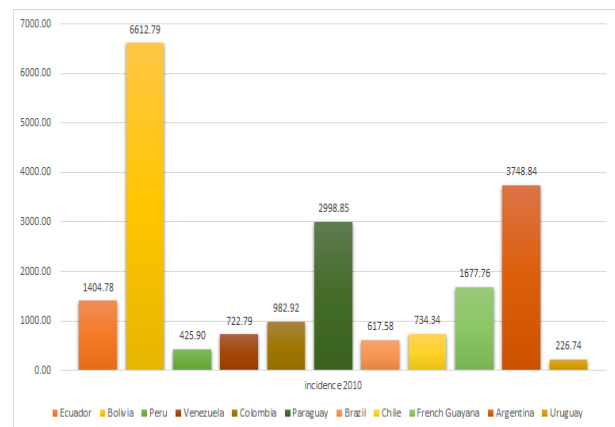

Fig. 4: Dispersion of the Chagas vector in South America. Source: World Health Organization (WHO) 2010

In the country, three endemic zones were identified: the valley area comprised by the departments of Cochabamba, Chuquisaca, Tarija and Potosí, the Chaco area comprised by the departments of Santa Cruz, Chuquisaca and Tarija, and finally the Amazon area departments of Beni, Pando, part of Santa Cruz, north of La Paz and north of Cochabamba.
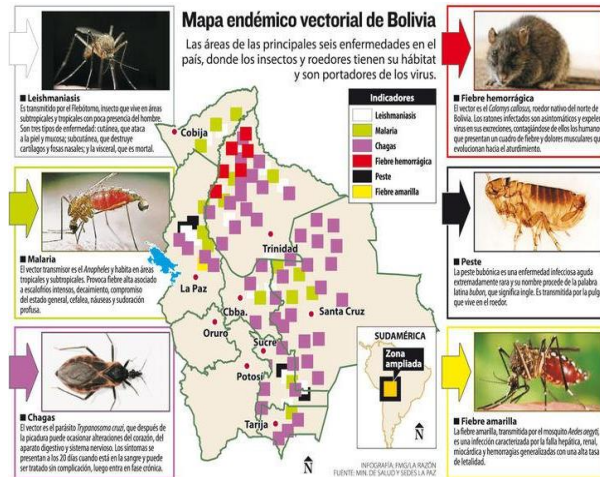
IAC-17,B1,5,5,x 41262

Fig. 5: Endemic vector map of Bolivia
Source: Ministry of Health of Bolivia

As we can see in the image, the purple squares symbolize the endemic areas where incidence of Chagas, we can see that the vector disease has comprised part of the departments of Santa Cruz, Beni, Cochabamba, La Paz, Sucre and Tarija.

The departments of Tarija, Chuquisaca, show an almost complete dispersion of the vector and the departments of Cochabamba, Santa Cruz, La Paz and Beni show a partial dispersion [7].

In Bolivia there are approximately 3 million people infected with Chagas. In 2016 the Ministry of Health of Bolivia recorded 17,892 new cases.

In the department of Santa Cruz, 57.72% of the total cases of chagas were concentrated, followed only by the department of Cochabamba, which concentrated a total of 23.13% of infected persons.

III. SELECTION OF THE STUDY AREA

As described above the Chagas has become an emerging disease in Bolivia, as it has found in the warm departments of the country, such as Santa Cruz, Tarija, Chuquisaca and Cochabamba, the adequate conditions to prevail and transmit the disease.

For the other hand we can also observe that Chagas disease has affected six of the nine departments of Bolivia.

For this reason we must select among the departments of the country the most affected by this viral disease and where there is a high prevalence of cases. [8]

For this, the National Program of Preventative and Control of Chagas in Bolivia was revised, where the epidemiological situation of the country in 2016 can be observed.

DEPARTMENTS	TOTAL		
	TOTAL-M	TOTAL-F	TOTAL
BENI	82	85	167
CHUQUISACA	350	483	833
COCHABAMBA	1705	2433	4138
LA PAZ	79	131	210
ORURO	4	17	21
PANDO	6	5	11
POTOSI	408	448	856
SANTA CRUZ	4334	5994	10328
TARIJA	545	783	1328
Total:	7513	10379	17892

Table 1: Acute Chagas incidence in Bolivia - 2016. Source: National Program of Preventative and Control of Chagas in Bolivia

As we can see in the table the departments of Bolivia with the highest incidence of Chagas cases are three: Santa Cruz, Cochabamba and Tarija. During the year 2016, Santa Cruz presents 10328 cases of Chagas, so we will choose a study area belonging to this department.

PROVINCES	TOTAL		
	TOTAL-M	TOTAL-F	TOTAL
ANDRES IBAÑEZ	3768	4958	8726
CHIQUITOS	11	47	58
CORDILLERA	13	11	24
FLORIDA	6	5	11
GERMAN BUSCH	1	1	2
GUARAYOS	1	1	2
ICHILO	11	15	26
MANUEL MARIA CABAL	0	1	1
ÑUFLO DE CHAVEZ	42	49	91
OBISPO SANTISTEVAN	391	458	849
SARA	3	1	4
VALLEGRANDE	1	0	1
VELASCO	2	4	6
WARNES	84	443	527
Total:	4334	5994	10328

Table 2: Acute Chagas incidence in Santa Cruz - 2016. Source: National Program of Preventative and Control of Chagas in Bolivia

As we can see in the table above, there are three provinces of the department of Santa Cruz that record the highest incidence rates of acute Chagas in 2016.

Fig. 6: Map of provinces of Santa Cruz, Bolivia Source: Military Geographic Institute Bolivia

IAC-17,B1,5,5,x 41262

As we can see in the figure, the area selected by the shape in red is our area of interest, where we have three provinces, these are: Andrés Ibañez, Obispo Santistevan and Warnes. Therefore these three provinces will be selected as our study area.

IV. SATELLITE IMAGE PROCESSING

The general block diagram of the present project is shown below:
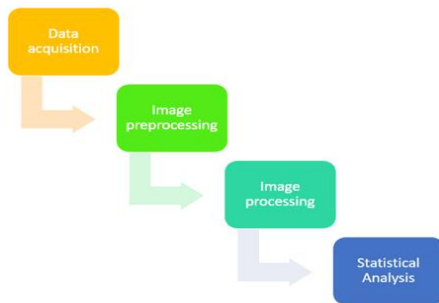

Fig. 7: General block diagram.
Source: Own elaboration

In the previous diagram we can observe the sequence of ordered steps that will help us to achieve the objective of the present research project.

Data acquisition

For the development of this research, different types of data must be acquired, the most crucial is, the number of acute Chagas cases in the study area region during the study period, on the other hand, it will be necessary to supplement these data with data from the meteorological stations located in the provinces of study and finally the satellite images of the department of Santa Cruz.

A. Incidence of Acute Chagas

The Ministry of Health of Bolivia, through its Departmental Health Service (SEDES), in a joint effort of the nine departments of Bolivia, has developed an informatic tool for develope the epidemiological surveillance of vector-borne diseases, including: Dengue, Zika, Chikungunya and the acute Chagas disease. [9]

Fig. 8: Epidemiological surveillance tool
Source: Ministry of Health of Bolivia

This tool allows us to know the number of cases of acute Chagas by department, province and municipality during all the year, in this case the data of 4 months have been collected from July to October of 2016, since these are the months selected to carry out the study.

Province	Capital	Municipality	Total cases of acute Chagas			
			July	August	September	October
Andrés Ibáñez	Santa Cruz de la Sierra	Cotoca	5	6	12	2
		El Torno	4	8	5	2
		La Guardia	5	5	2	7
		Porongo	0	0	0	0
		Santa Cruz de la Sierra	875	636	521	521
Obispo Santistevan	Montero	Fernandez Alonso	2	0	2	4
		General Saavedra	0	0	1	0
		Mineros	0	0	0	1
		Montero	64	114	58	44
		San Pedro	3	7	2	3
Warnes	Warnes	Warnes	57	54	31	37
		Okinawa	0	0	0	0

Table 3: Epidemiological surveillance tool
Source: Ministry of Health of Bolivia

As can be seen in the above table, in total our area of study comprises three provinces and twelve municipalities. It can be observed that the incidence rate of acute Chagas throughout the four months of study is of 3100 confirmed cases during that period.

B. Meteorological data

Through the National Service of Meteorology and Hydrology of Bolivia it was possible to identify the meteorological stations installed in the study area. [10] They can be observed in the following figure:
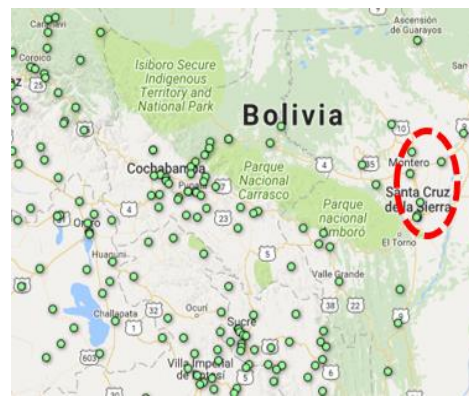

Fig. 9: Location of the weather stations in Bolivia
Source: National Service of Meteorology and Hydrology - SENAMHI

In total there are six meteorological stations located in the provinces of Obispo Santistevan, Warnes and Andres Ibañes, as we can observe in the next table:

Name	Province	Municipality	Altitude (masl)
San Pedro_Sc	Obispo Santistevan	Mineros	450
Minero (Unagro)	Obispo Santistevan	Mineros	245
Montero_Muyurina	Obispo Santistevan	Montero	300
Viru Viru	Warnes	Warnes	378
Trompillo Aeropuerto	Andrés Ibañes	Santa Cruz de la Sierra	424
Santa Cruz SENAMHI	Andrés Ibañes	Santa Cruz de la Sierra	416

Table. 4: Selection of weather stations
Source: National Service of Meteorology and Hydrology - SENAMHI

From the six meteorological stations selected, data of the minimum, maximum and average temperature were obtained; in addition to the average precipitation data of the study area in the selected period of time.

C. Satellite Images

For this research project we use the satellite imagery from Landsat 8, which is an American terrestrial observation satellite launched on February 11, 2013. It is the eighth and most recent satellite of the Landsat project operated by NASA and the United States Geological Survey (USGS) since 1972. [11]

The Landsat 8 satellite carries two instruments: Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). Of the total of 11 spectral bands, 9 bands are taken by the OLI instrument and 2 bands are taken by the TIRS instrument, in the next table we can see the specifications:

Band	Name	Wavelength (µm)	Resolution (m)
1	Ultra Blue (coastal/aerosol)	0.435 - 0.451	30
2	Blue	0.452 - 0.512	30
3	Green	0.533 - 0.590	30
4	Red	0.636 - 0.673	30
5	Near Infrared (NIR)	0.851 - 0.879	30
6	Shortwave Infrared 1 (SWIR 1)	1.566 - 1.651	30
7	Shortwave Infrared 2 (SWIR 2)	2.107 - 2.294	30
8	Panchromatic	0.503 - 0.676	15
9	Cirrus	1.363 - 1.384	30
10	Thermal Infrared (TIR 1)	10.60 - 11.19	100
11	Thermal Infrared (TIR 2)	11.50 - 12.51	100

Table 5: Landsat 8 Satellite Sensor Specifications.
Source: USGS-NASA

A level 1 scene consists of a radiometrically and geometrically corrected image with 13 files, where 11 files are in GeoTIFF and correspond to each of the 11 recorded bands, one metadata file (MTL) and one more file with a scene quality evaluation (QA).

For download the Landsat 8 images we must enter in the website of USGS (https://earthexplorer.usgs.gov/), for download the satellite images we do not need to use special software, as mentioned above is free and allows you to download the images that you want. Initially we look for our area of interest, in this case Santa Cruz, Bolivia:

Fig. 10: Downloading Landsat 8 images.
Source: USGS Website

We can filter the images by different criteria, mainly by year, month, percentage of cloudiness and others.

The selected study months are: July, August, September and October of 2016, due to the fact that the other months of the year have a high percentage of cloud coverage of more than 70%, which can lead to erroneous results.
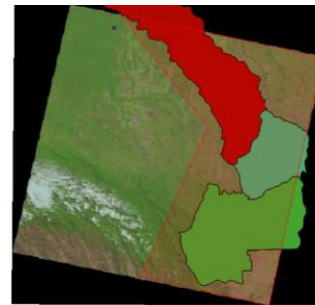

Fig. 11: Making a shape of the study area.
Source: Own elaboration

As we can see in the previous image, it was decided to create a shape and crop the satellite image in order to obtain the data only from the area of interest.

Image preprocessing

Due to satellite instruments, different effects can occur that can affect the images acquired, which is why, in order to acquire valid information from our satellite images, different types of corrections must be made in the pre-processing stage, since we need correct the errors that our satellite image acquired when it was exposed to different interferences.
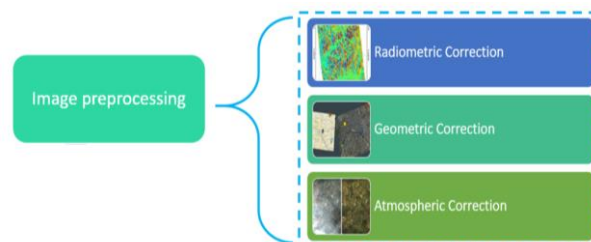

Fig. 12: Preprocessing step block diagram.
Source: Own elaboration

The corrections that will be made in the present project will be explained in the following section.

D. Radiometric correction

This correction is the technique to reconstruct the physically calibrated values by correcting the spectral distortions caused by different reasons like: the instrumental effects due to the calibration, the seasonal changes, the type of terrain that influences the radiation, the sunlight that affects the radiometric values and the scattering effects of the atmosphere, among others. [12]

The main purposes of the radiometric correction is to perform a sensor calibration for reduce the red-out effects and improve the sensitivity of the detector, also the calibration wants to reduce the sun angle and topographic effect for reduce the surface slope, finally wants to correct the phenomena caused for the absorption and scattering caused for the Earth's atmosphere.

For this reasons the radiometric correction is really important for this project as we will perform a quantitative image analysis and we know that the improperly calibrated Landsat 8 images may lead us to a wrong conclusion.

E. Geometric correction

The geometric correction is the technique to re-project the radiometrically corrected image onto a common map projection. Also an important step is to georeferencing each pair of pixel in the image with a geographic coordinate of the map. Finally applying these two techniques we will achieve a geocoded product. [13]

For this project we will need a totally geometrically corrected image, because for this application we required the absolute locations of the data extracted from the image.

F. Atmospheric correction

The atmospheric correction is the technique of evaluate and eliminate the atmospheric and terrain distortions that are introduced in the values of radiance that arrive at the sensor from the Earth's surface. [14]

This with the main objective of recovering the physical parameters of the terrestrial surface including the reflectance of the surface, the visibility of the ground and the temperature. This correction is especially needed for the application of the present project, because it is necessary to compare and analyze images of different dates and taken in different conditions.

Image processing

In the image processing step we will need to extract environmental parameters that are relevant indicators for the study of the epidemiological chain of Chagas disease. For this we need to extract from the satellite images the next parameters:

Fig. 13: Processing step block diagram.
Source: Own elaboration

As we can observe we need to compute the Land Surface Temperature (LST), Relative Humidity, Normalized Difference Water Index (NDWI), Normalized Difference Soil Index (NDSI), Normalized Difference Vegetation Index (NDVI), Tasseled Cap index, which is an indicator of greenery, brightness and humidity.

An algorithm was programmed in MATLAB to calculate the above mentioned parameters, the general block diagram of the algorithm was as follows:
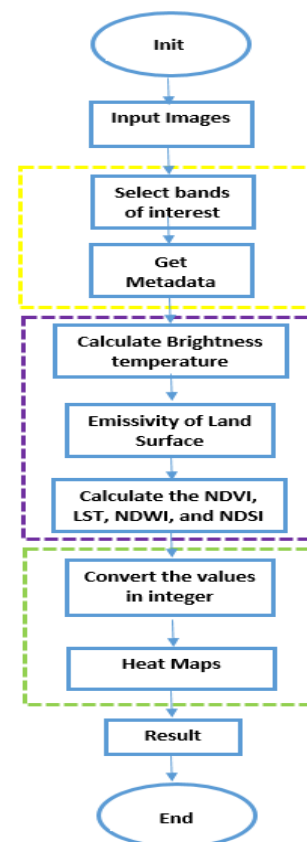

Fig. 14: General flowchart of the algorithm
Source: Own elaboration

According to the flow diagram, the programmed algorithm first selects the bands of interest, then reads the metadata file from which it will extract certain constants,

then perform the calculations with the programmed equations and finally show heat maps and graphics to be able to analyze the behaviour of the environmental parameters in the different areas of the image.

After compute the environmental parameters of the satellite images of the study area we create a database with all the information acquired in the different stages of the study:
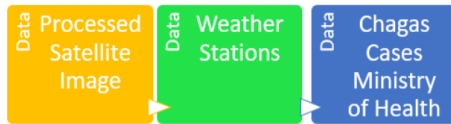

Fig. 15: Data base diagram.
Source: Own elaboration

As can be observed, the environmental parameters calculated from the image processing step, the meteorological data of the SENAMHI stations located in the study area and finally the data of the incidence of acute Chagas were included.

After the creation of the database, a statistical analysis was performed. The results will be displayed in the next section.

V. ANALYSIS OF RESULTS

The data obtained have been analyzed by statistical tools that allow synthesizing the information, minimizing the loss of information. For this purpose the principal components model was selected to analyze the data.

The variables included in the analysis are detailed in the following table:

Variable	Definition
Incidence	Number of new cases of Chagas per 100,000 hab. in the study region
NDWI	Index that represents the presence of water bodies in the image
NDSNI	Index that represents the presence of snow in the image
NDVI	Index that represents green vegetation in the analyzed image
NDMI	Index that represents the amount of soil moisture
NDWILEAVES	Index that represents the water stress of the vegetation
Temperature	Average temperature per month (°C)
Tempdesv	Standard deviation of the average temperature per month (°C)
Desvvariacitemp	Standard deviation of variation of daily temperatures per month (°C)
Precipitacionmm	Monthly accumulated precipitation (mm)
Temmin	Minimum temperature recorded (°C)
Temmax	Maximum temperature recorded (°C)

Table 6: Definition of the measured variables.
Source: Own elaboration

These variables have been calculated to investigate their effect within the variation of number of cases during the months of analysis. An important effect to which we gave greater presence is the variation of temperature, for this reason we analyzed the maximum and minimum temperature of the period, and the standard deviation of the daily variation of temperature, this parameter is a cause of the variations in the production of the vector and therefore the contagion of the same.

From the analysis of principal components the following table of results was obtained:

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	5.00896	0.284543	0.4554	0.4554
Comp2	4.72441	3.45778	0.4295	0.8849
Comp3	1.26663	1.26663	0.1151	1
Comp4	0	0	0	1
Comp5	0	0	0	1

Table 7: Result of principal component analysis.
Source: Own elaboration

To maximize the explained variance, the first three components that explain 100% of the variance have been selected. If we add a new component no differences are observed, if we decrease one of the components we lose 12% of the explained variance.

In the following table we show the relationships of each principal component with each variable:

Variables	Comp1	Comp2	Comp3
Incidence	0.4311	-0.1163	-0.0639
NDWI	0.1664	0.4255	-0.0696
NDSNI	-0.0731	0.4193	0.3357
NDVI	0.038	-0.4575	0.0543
NDWILEAVES	0.342	0.2931	0.0805
Temperature	0.4438	-0.0498	0.0348
Tempdesv	-0.1173	0.3413	-0.5483
Desvvariacitemp	-0.4174	0.1116	0.2326
Precipitación mm	0.13	0.34	0.5398
Temmin	0.3606	-0.2205	0.3064

Table 8: Relationship between components and variables. Source: Own elaboration

From the table we can see that the three components depend on all the measured variables.

Component 1 has a strong positive influence on incidence, temperature, and water stress of vegetation.

Second component has positive influence in important percentage with a NDWI, NDSNI, and temperature variation.

Finally, component 3 is positively related to precipitation and minimum temperature.

VI. CONCLUSIONS AND RECOMMENDATIONS

● Taking into account that in Bolivia the occurrence of Chagas outbreaks shows an increasing intensity since 2015, this study allow us to have a more integrated understanding of this disease, understand the environmental factors in which the transmission of the disease happens.

- With this proof of concept we can indicate that the incidence of Chagas cases is linked to certain parameters such as: temperature, the water stress of vegetation, water bodies and accumulated precipitation.
- A future project pretend create a system that allow us to identify areas with ideal environmental conditions to house the vinchuca in order to plan and apply preventive measures in those places to mitigate the disease, resources and actions should also be planned, impact assessments and even projections can be made in the future to protect the vulnerable population living close to areas where mosquitoes develop.

VII. REFERENCES

[1] M. Rojas-Cortez et al., "Trypanosoma cruzi-infected Panstrongylus geniculatus and Rhodnius robustus adults invade households in the Tropics of Cochabamba region of Bolivia," Parasit. Vectors, vol. 9, Mar. 2016.

[2] M. das G. V. Barbosa et al., "Chagas disease in the State of Amazonas: history, epidemiological evolution, risks of endemicity and future perspectives," Rev. Soc. Bras. Med. Trop., vol. 48, pp. 27–33, Jun. 2015.

[3] J. R. Coura and A. C. Junqueira, "Risks of endemicity, morbidity and perspectives regarding the control of Chagas disease in the Amazon Region," Mem. Inst. Oswaldo Cruz, vol. 107, no. 2, pp. 145–154, Mar. 2012.

[4] L. A. Messenger et al., "Toward Improving Early Diagnosis of Congenital Chagas Disease in an Endemic Setting," Clin. Infect. Dis., vol. 65, no. 2, pp. 268–275, Jul. 2017.

[5] M. L. Díaz and C. I. González, "Enfermedad de Chagas agudo: transmisión oral de Trypanosoma cruzi como una vía de transmisión re-emergente," Rev. Univ. Ind. Santander Salud, vol. 46, no. 2, pp. 177–188, Aug. 2014.

[7] S. F. Brenière, E. W. Condori, R. Buitrago, L. F. Sosa, C. L. Macedo, and C. Barnabé, "Molecular identification of wild triatomines of the genus Rhodnius in the Bolivian Amazon: Strategy and current difficulties," Infect. Genet. Evol. J. Mol. Epidemiol. Evol. Genet. Infect. Dis., vol. 51, pp. 1–9, Jul. 2017.

[8] Ministry of Health Bolivia, "Sistema Nacional de Información en Salud-Vigilancia Epidemiológica", Jul-Oct 2016.

[9] Ministry of Health Bolivia, "Boletín de vigilancia epidemiológica, Chagas Agudo", Form. 302 a. 2016.

[10] National Service of Meteorology and Hydrology Bolivia, "Ubicación de Estaciones Meteorológicas Bolivia".

[11] Roy, D. P., Wulder, M. A., Loveland, T. R., Woodcock, C. E., Allen, R. G., Anderson, M. C & Scambos, T. A. (2014). Landsat-8: Science and product vision for terrestrial global change research. Remote sensing of Environment, 145, 154-172.

[12] Markham, B., Barsi, J., Kvaran, G., Ong, L., Kaita, E., Biggar, S. & Helder, D. (2014). Landsat-8 operational land imager radiometric calibration and stability. Remote Sensing, 6(12), 12275-12308.

[13] Storey, J., Choate, M., & Lee, K. (2014). Landsat 8 Operational Land Imager on-orbit geometric calibration and performance. Remote sensing, 6(11), 11127-11152.

[14] Kruse, F. A. (2004, March). Comparison of ATREM, ACORN, and FLAASH atmospheric corrections using low-altitude AVIRIS data of Boulder, CO. In Summaries of 13th JPL Airborne Geoscience Workshop, Jet Propulsion Lab, Pasadena, CA.