NDVI and LST extraction of MODIS data under a GIS open source application - Rickettsia study case in Angola

Ana Cláudia Teodoro
Dep. Geosciences, Environment and Land Planning Faculty of Sciences, Rua Campo Alegre, 4169-007, Porto, University of Porto, Porto, Portugal; Earth Sciences Institute (ICT), Faculty of Sciences, Rua Campo Alegre, 4169-007, Porto, University of Porto, Porto, Portugal

Lia Duarte
Dep. Geosciences, Environment and Land Planning Faculty of Sciences, Rua Campo Alegre, 4169-007, Porto, University of Porto, Porto, Portugal; Earth Sciences Institute (ICT), Faculty of Sciences, Rua Campo Alegre, 4169-007, Porto, University of Porto, Porto, Portugal

Patrícia Barradas
Dep. of Pathology and Molecular Immunology, Institute of Biomedical Sciences Abel Salazar (ICBAS), Rua Jorge Viterbo, 4050-313, Porto, University of Porto, Porto, Portugal

T. L. Mateus
Escola Universitária Vasco da Gama

Zoraima Neto
Laboratório de Biologia Molecular, Instituto Nacional de Saúde Pública, Ministério da Saúde, Maianga-Luanda, Angola

NDVI and LST extraction of MODIS data under a GIS open source application - Rickettsia study case in Angola

Optional Cover Page Acknowledgements
Joana Afonso and Jocelyne Neto, Instituto Nacional de Saúde Pública, Ministério da Saúde, Maianga-Luanda, Angola for the support. USGS for satellite data.

NDVI and LST extraction of MODIS data under a GIS open source application - Rickettsia study case in Angola

Ana Cláudia Teodoro[a,b*], Lia Duarte[a,b], Patrícia Barradas[c], Teresa L. Mateus[d,e,f], Zoraima Neto[g]

[a] *Dep. Geosciences, Environment and Land Planning Faculty of Sciences, Rua Campo Alegre, 4169-007, Porto, University of Porto, Porto, Portugal*

[b] *Earth Sciences Institute (ICT), Faculty of Sciences, Rua Campo Alegre, 4169-007, Porto, University of Porto, Porto, Portugal*

[c] *Dep. of Pathology and Molecular Immunology, Institute of Biomedical Sciences Abel Salazar (ICBAS), Rua Jorge Viterbo, 4050-313, Porto, University of Porto, Porto, Portugal*

[d] *Departamento de Medicina Veterinária, Escola Universitária Vasco da Gama, Lordemão, Coimbra, Portugal*

[e] *Escola Superior Agrária, Instituto Politécnico de Viana do Castelo, Refóios do Lima, Portugal*

[f] *EpiUnit, Instituto de Saúde Pública da Universidade do Porto, Porto, Portugal*

[g] *Laboratório de Biologia Molecular, Instituto Nacional de Saúde Pública, Ministério da Saúde, Maianga-Luanda, Angola*

Abstract: Fevers of unknown origin can have different aetiologies. The overlapping symptomatology of rickettsial infection and other endemic diseases that cause fever leads to a misdiagnosis or under-diagnosis of spotted fever group of Rickettsia (SFGR).

To better understand the epidemiology of this vector-borne disease in Angola, a comprehensive seroprevalence study was conducted investigating the exposure to SFGR in a sample of 92 febrile, Malaria and Yellow Fever negative human plasma specimen, collected to the study of the national surveillance of febrile syndromes between 2016 and 2017, in Angola.

The seroprevalence of IgG antibodies against SFG Rickettsia in humans was calculated by gender, and aimag (province). All data were analyzed through a logistic regression. Spatial data sources included Normalized Differential Vegetation Index (NDVI) and Land Surface Temperature (LST) products acquired by Moderate Resolution Imaging Spectroradiometer (MODIS).

The main objective of this work was the development of a GIS open source application – QMOD - to automatize the extraction of LST and NDVI products from MODIS data. The application was created as a simple graphic interface composed by two input fields (the text file with the coordinates of the sampling points (in sinusoidal coordinate system and the folder with the MODIS images), the field to define the buffer distance, and the output file. The application was tested considering MOD11A1 (LST product), MOD13Q1 and MYD13Q1 (NDVI product), free download from the USGS.

---

*Corresponding author
Email address: amteodor@fc.up.pt (First Author)
URL: https://first-author.com (First Author)

QGIS 2.18.17 was used for geospatial operations and the Python language was employed for the development of the GIS open source application under QGIS software. The process includes the circumscription of the major clusters where human data were collected. Then, a convex hull (minimum convex bounding geometry) was created around each sampling site with a 10 km buffer zone to accommodate the mobility among the nomadic people being samples. Counts of seropositive and seronegative humans were calculated within each of these sampling clusters along with the mean, maximum, and minimum values of NDVI and LST, and percent area of each land cover class.

The application was tested in a set of 92 points in Angola and a buffer of 10 km considering the Universal transverse Mercator (UTM) Zone 33S projection (EPSG:32733) was applied for each point. The LST and NDVI statistical values were extracted for each sampling cluster.

Variations in ecological niches, abundance of vegetation and land surface temperature, for ticks and fleas between different provinces could be in part responsible for the geographic differences in seroprevalence observed with SFGR.

1. Introduction

Rickettsiae is a group of vector-borne organisms which cause fever. It is transmitted by ticks, body lice, and fleas and the transmission vary depending on the region and organism. For the spotted fever group Rickettsiae, humans are a possible receptor leading to death (Walker and Fishbein, 1991). The infection can cause morbidity and mortality in humans with no vaccines available (Sahni et al., 2013).

Spotted fevers are diseases caused by Gram negative, obligate intracellular bacterial agents belonging to the spotted-fever (SF) group of the genus Rickettsia, transmitted by arthropods.

There are several circulating Rickettsia species. The main clinical manifestations of a rickettsial syndrome in humans are fever, rash and eschar however, they are not pathognomonic. When there is a rickettsiosis suspicion, antimicrobial treatment should be started. Confirmatory laboratory results contribute to the epidemiological knowledge of the circulating species.

Remote sensing data are increasingly used to study vector-borne diseases relating environmental variables such as temperature, humidity, land cover type, and elevation data (Heinrich et al., 2015; Kalluri et al., 2007). The information derived from the satellite data including vegetation indices, soil types, soil moisture, land surface temperature, etc. integrated in a geographical information system (GIS) is a powerful strategy in the study of vector-borne diseases (Palaniyandi, 2012).

2. Literature Review

Dogs in tick-infested areas are at risk for contracting this infection (Stiles, 2000). Consequently, humans can contract the disease. Dogs can develop rickettsiemia, however the knowledge on the susceptibility of dogs to different Rickettsia is still limited (Barradas et al., 2017). In Africa, the spotted fever group of Rickettsia (SFGR) has been detected in ticks and humans. A study was already performed to understand the epidemiology and clinical importance of vector-borne agents in Luanda, Angola (Barradas et al., 2017).

In the last decades remote sensing data and GIS established several important and crucial links between landscape ecology and diseases. The combination of remote sensing data and GIS has allowed the analysis of geographical distribution of several vector-borne diseases.

Several review articles have been published regarding the usefulness of remote sensing data and GIS in the analysis of vector-borne diseases (Viana et al., 2017; Khan et al., 2010; Kitron, 1998). Most of the studies involving remote sensed data were related to parasitic diseases and refer to the pathogen, studying the presence/absence of the specific vectors (Beck et al., 2000). Malaria is one of the most studied diseases from the remote sensing point of view (e.g., Masuoka et al., 2003; Li et al., 2016). From our knowledge, few studies are addressed considering the Rickettsiae and remote sensing data/GIS analysis.

3. Study Area and Dataset Used

3.1. Study Area

Angola is located in the southern hemisphere, in the African continent. Is a large country, with an area of 1 246 700 km^2. The population in 2017 is estimated in 29 310 273 hab.

Angola has a sub-tropical climate almost everywhere, with a cool and dry season from May to August (called Cacimbo), and a hot and rainy season, from September to April in the north-east, from mid-October to April in the centre, from November to March in the south, and from February to April in the south-west. Luanda is the Angola capital, the major city and the most populated (apx. 10 000 000 hab).

3.2. Dataset

Selected specimens provided from patients that lived in urban neighbourhoods in provinces located across Angola, without access to potable water and with contact with domestic animals were analysed. Table 1 provides the demographic, sanitary and climacteric characteristics of the provinces were data were collected.

Table 1: Demographic, sanitary and climacteric characteristics of the provinces were data were collected.

Provinces	Population Density (Census 2014)	Zone type	Sanitary conditions	Mean Annual Average Temperature	Mean Annual Average Humidity
Cuango	800 000 hab	Suburban Area	Latrines (WC outside the house)	31ºC	20%
Cacuaco	2 500 000 hab	Suburban Area	Latrines (WC outside the house)	27 ºC	72%
Benguela	510 000 habi	Urban Area	WC inside the house	28 ºC	55%
Viana	200 000 hab	Urban Area	WC inside the house	26 ºC	70%
Maianga	789 000 hab	Urban Area	WC inside the house	26 ºC	72%
Capenda	80 000 hab	Suburban Area	Latrines (WC outside the house)	31 ºC	20%
Huambo	2 300 000 hab	Suburban Area	Latrines (WC outside the house)	24 ºC	25%
Cacuaco	2 500 000 hab	Suburban Area	Latrines (WC outside the house)	27 ºC	72%

The 92 samples were defined by their coordinates (longitude, latitude) and the sample collection date, according to Table 2.

Table 2: Field samples collected in Angola (extraction).

Nº	County	Coordinates (Longitude, Latitude)	Sample collection date	Nº	County	Coordinates (Longitude, Latitude)	Sample collection date
2	Cuango	8°39'21.22"S, 18° 8'33.98"E	21-05-16	48	Cacuaco	8°48'20.13"S, 13°20'49.84"E	10-03-16
3	Cacuaco	8°48'20.13"S, 13°20'49.84"E	07-03-16	49	Benguela	12°35'25.70"S, 13°24'59.55"E	22-04-16
9	Benguela	12°35'25.70"S, 13°24'59.55"E	22-05-16	55	Gambos	14°49'44.40"S, 14°33'1.12"E	06-03-16
10	Viana	8°53'49.70"S, 13°20'58.63"E	N/A	56	Benguela	12°35'25.70"S, 13°24'59.55"E	22-04-16
11	Maianga	8°51'5.87"S, 13°14'50.45"E	N/A	57	Maianga	8°51'5.87"S, 13°14'50.45"E	18-02-16
22	Capenda	9°25'3.39"S, 18°26'6.08"E	11-06-16	68	Cacuaco	8°48'20.13"S, 13°20'49.84"E	04-03-16
23	Huambo	12°46'26.31"S, 15°44'48.67"E	05-06-16	69	Ganda	13° 1'28.34"S, 14°37'46.64"E	15-02-16
42	Cacuaco	8°48'20.13"S, 13°20'49.84"E	10-03-16	88	Cazenga	8°50'3.81"S, 13°17'42.75"E	02-03-16

The spatial data sources included Normalized Differential Vegetation Index (NDVI) and Land Surface Temperature (LST) products retrieved by Moderate Resolution Imaging Spectroradiometer (MODIS). MOD11A1 (MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1 km SIN Grid V006), MOD13Q1 (MODIS/Terra Vegetation Indices 16-Day L3 Global 250 m SIN Grid V006) and MYD13Q1 (MODIS/Aqua

Vegetation Indices 16-Day L3 Global 250 m SIN Grid V006), freely downloaded from the USGS site, were considered (USGS, 2018). The products with the dates closest to the field sample collection dates were downloaded. Also, the study case considers two distinct tiles (19h, 10v) and (19h, 9v). The MODIS products are currently produced at 250 m, 500 m, 1 km and 5600m (0.05 deg) spatial resolutions for the NDVI and 1000 km for the LST. MODIS are mapped in the Sinusoidal (SIN) grid projection.

The MOD13Q1 Version 6 product provides a Vegetation Index (VI) value at a per pixel basis. There are two primary vegetation layers. The first is the NDVI which is referred to as the continuity index to the existing National Oceanic and Atmospheric Administration-Advanced Very High Resolution Radiometer (NOAA-AVHRR) derived NDVI. The second vegetation layer is the Enhanced Vegetation Index (EVI), which has improved sensitivity over high biomass regions. The grid consists of 4,800 rows and 4,800 columns of 250 m pixels. The algorithm chooses the best available pixel value from all the acquisitions from the 16-day period. The criteria used is low clouds, low view angle and the highest NDVI/EVI value (Didan, 2015).

The MOD11A1 Version 6 product provides daily, per-pixel LST in a 1200 x 1200 km grid. The pixel temperature value is derived from the MOD11_L2 swath product. Above 30 degrees' latitude, some pixels may have multiple observations where the criteria for clear-sky are met. When this occurs, the pixel value is a result of the average of all qualifying observations. Provided along with both the day-time and night-time surface temperature bands and their quality indicator layers, are MODIS bands 31 and 32 and six observation layers (Wan et al., 2015). The study area and the location of the field samples is presented in Figure 1.

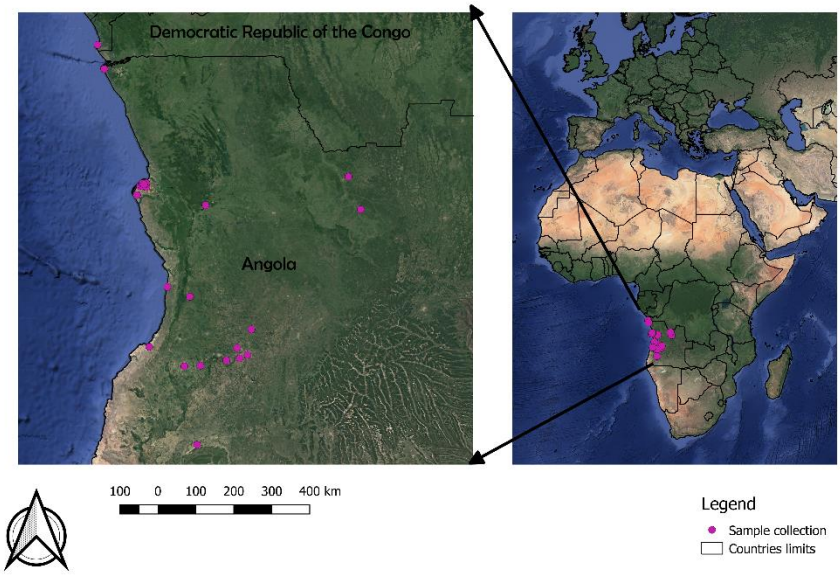


Figure 1: Study area and field samples location.

4. Methodology

The 92 samples were defined by their coordinates (longitude, latitude) and the date of collection. The main objective was to compute the statistics (minimum, maximum, sum, count, mean, standard deviation, range, variance, median and mode) of the NDVI and the LST values from MODIS products, considering a buffer of 10 km to each point/sample. This

methodology is time consuming when performed manually. So, a GIS open source application - QMOD - which automatizes the procedure, was developed.

4.1. Data uniformization

MODIS products are in sinusoidal projection, so all the images used were projected to World Geodetic System 1984 (WGS84) UTM 33S (EPSG: 32733), the coordinate system correspondent to the field samples. The field sample coordinates (WGS84) were projected to WGS84 UTM 33S coordinate system through Proj4 (Proj, 2018).

Proj is a standard command line filter function which allows to convert geographic longitude and latitude coordinates into cartesian coordinates (and vice versa) (Proj, 2018). The following command was used: *Gdalwarp -t_srs epsg:32733* (Proj, 2018).

4.2. MODIS images

The MOD11A1, MOD13Q1 and MYD13Q1 products were downloaded from USGS (USGS, 2018). A factor of 0.0001 was applied to the NDVI products (MOD13Q1 and MYD13Q1), and a factor of 0.02 was applied to the LST product. To each day where field samples were taken, the closest image (NDVI or LST) was identified and downloaded. In the NDVI products, the algorithm chooses the best available pixel value from all the acquisitions from the 16-day period. Because the MODIS sensors aboard on Terra and Aqua satellites are identical, the algorithm generates each 16-day composite 8-days apart (phased products) to allow a higher temporal resolution product by combining both MODIS sensors. The MOD11A1 LST product provides daily information.

4.3. NDVI_LST application

The QMOD application was developed under QGIS software (version 2.18), an open source software (QGIS, 2018). The Python programming language was used (Python, 2018) and the graphic interface was built through QtDesigner (QtDesigner, 2018). The application is composed by a single button which opens a window.

The window is composed by five parameters: (1) the text file input with the coordinates of the samples; (2) the input folder containing the MODIS images considered according the dates; (3) the value of the buffer to consider; (4) the spatial reference system selector allowing to choose the coordinate system; and (5) the output parameter whit the resulting statistics for each zone (Figure 2).

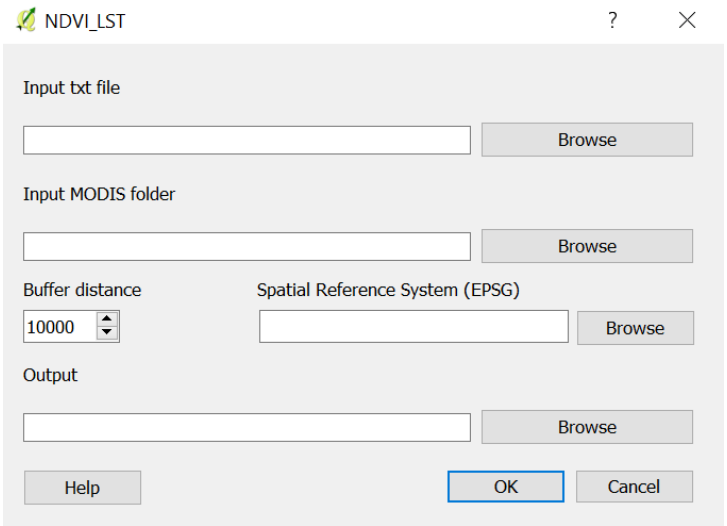


Figure 2: Plugin graphic interface.

To develop the application, the first step was the graphic interface creation through QtDesigner. Then, the main code was developed considering the function *run()*.

Several algorithms from *Processing Toolbox* framework were used to develop the application. For instance, *reprojectlayer*, *buffervectors* and *zonalstatistics* from GDAL/OGR (GDAL, 2018). Figure 3 presents a flowchart of the methodology implemented.

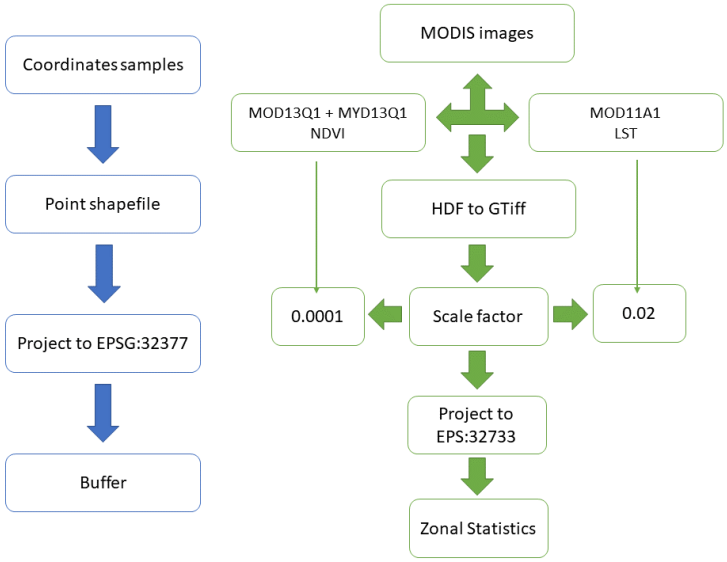


Figure 3: Flowchart of the methodology implemented.

The first parameter was the text file where the coordinates are inserted, that allowed to find files which ends with *txt* format and automatically creates a point shapefile using *QgsVectorFile* class. This shapefile considers the coordinates in sinusoidal projection. These points are then automatically projected to the coordinate system selected by the user. In this case, to WGS84 UTM zone 33S (EPSG: 32733). A buffer, with the specified value (10 000 m), is created in the point shapefile.

Then, the application run raster by raster (in MODIS folder) and interprets the dates in the format day, month and year; using functions from Python, such as *datetime()*. Since NDVI and LST MODIS products are in the same folder, the product name is also interpreted and two conditions were created: (1) if the product is NDVI; or (2) if the product is LST.

If the product is LST, the *hdf* format is converted to *GTiff* format through *gdal_translate* executable (GDAL, 2018). Through *rastercalculator* algorithm from GDAL, the scale factor is applied (in this case 0.02). The resulting raster is then projected to UTM 33S coordinate system using *gdalwarp*, algorithm from GDAL (GDAL, 2018). Finally, considering the buffer shapefile the *zonalstatistics* algorithm from QGIS, is used to estimate the statistics of the raster (LST) in the defined buffer zones.

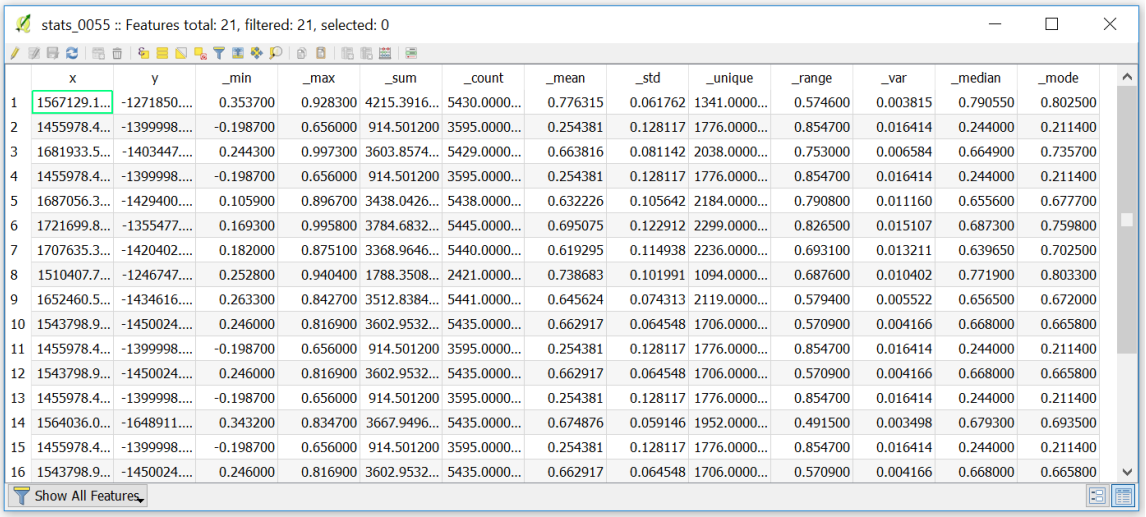
The same procedure is applied in the case of the NDVI product considering a factor of 0.0001. This procedure is automatized inside the application.

The coordinate system selector allows to open a window with several coordinate systems (geographic and projected). This functionality was developed using *QgsGenericProjectionSelector* class.

5. Results and Discussion

The plugin can take some time to run given the fact that it analyses all the MODIS images. Then, in a specific folder, the shapefile with the buffer area is created for each MODIS product and specific date. An excel file composed by the statistics is also created.

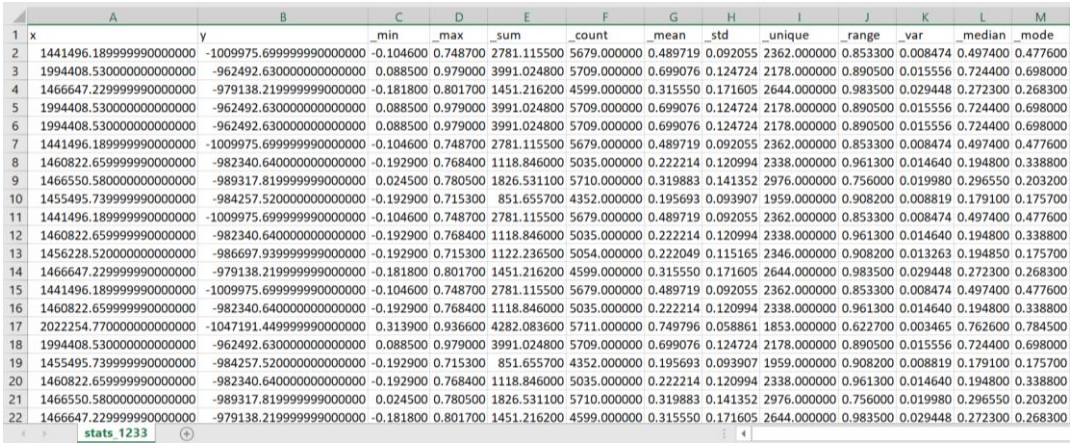
Moreover, the statistics are returned in the attribute table of the buffer shapefile. In this case, the statistics are: minimum, maximum, sum, count, mean, standard deviation, range, variance, median and mode. Figure 4 presents the attribute table of the buffer shapefile with the statistics computed. Figure 5 presents an extraction of the excel table resulting from the NDVI values for 29st March.


stats_0055 :: Features total: 21, filtered: 21, selected: 0

	x	y	_min	_max	_sum	_count	_mean	_std	_unique	_range	_var	_median	_mode
1	1567129.1...	-1271850....	0.353700	0.928300	4215.3916...	5430.0000...	0.776315	0.061762	1341.0000...	0.574600	0.003815	0.790550	0.802500
2	1455978.4...	-1399998....	-0.198700	0.656000	914.501200	3595.0000...	0.254381	0.128117	1776.0000...	0.854700	0.016414	0.244000	0.211400
3	1681933.5...	-1403447....	0.244300	0.997300	3603.8574...	5429.0000...	0.663816	0.081142	2038.0000...	0.753000	0.006584	0.664900	0.735700
4	1455978.4...	-1399998....	-0.198700	0.656000	914.501200	3595.0000...	0.254381	0.128117	1776.0000...	0.854700	0.016414	0.244000	0.211400
5	1687056.3...	-1429400....	0.105900	0.896700	3438.0426...	5438.0000...	0.632226	0.105642	2184.0000...	0.790800	0.011160	0.655600	0.677700
6	1721699.8...	-1355477....	0.169300	0.995800	3784.6832...	5445.0000...	0.695075	0.122912	2299.0000...	0.826500	0.015107	0.687300	0.759800
7	1707635.3...	-1420402....	0.182000	0.875100	3368.9646...	5440.0000...	0.619295	0.114938	2236.0000...	0.693100	0.013211	0.639650	0.702500
8	1510407.7...	-1246747....	0.252800	0.940400	1788.3508...	2421.0000...	0.738683	0.101991	1094.0000...	0.687600	0.010402	0.771900	0.803300
9	1652460.5...	-1434616....	0.263300	0.842700	3512.8384...	5441.0000...	0.645624	0.074313	2119.0000...	0.579400	0.005522	0.656500	0.672000
10	1543798.9...	-1450024....	0.246000	0.816900	3602.9532...	5435.0000...	0.662917	0.064548	1706.0000...	0.570900	0.004166	0.668000	0.665800
11	1455978.4...	-1399998....	-0.198700	0.656000	914.501200	3595.0000...	0.254381	0.128117	1776.0000...	0.854700	0.016414	0.244000	0.211400
12	1543798.9...	-1450024....	0.246000	0.816900	3602.9532...	5435.0000...	0.662917	0.064548	1706.0000...	0.570900	0.004166	0.668000	0.665800
13	1455978.4...	-1399998....	-0.198700	0.656000	914.501200	3595.0000...	0.254381	0.128117	1776.0000...	0.854700	0.016414	0.244000	0.211400
14	1564036.0...	-1648911....	0.343200	0.834700	3667.9496...	5435.0000...	0.674876	0.059146	1952.0000...	0.491500	0.003498	0.679300	0.693500
15	1455978.4...	-1399998....	-0.198700	0.656000	914.501200	3595.0000...	0.254381	0.128117	1776.0000...	0.854700	0.016414	0.244000	0.211400
16	1543798.9...	-1450024....	0.246000	0.816900	3602.9532...	5435.0000...	0.662917	0.064548	1706.0000...	0.570900	0.004166	0.668000	0.665800

Show All Features

Figure 4: Attribute table with the statistics (extraction).



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	x	y	_min	_max	_sum	_count	_mean	_std	_unique	_range	_var	_median	_mode
2	1441496.189999990000000	-1009975.699999990000000	-0.104600	0.748700	2781.115500	5679.000000	0.489719	0.092055	2362.000000	0.853300	0.008474	0.497400	0.477600
3	1994408.530000000000000	-962492.630000000000000	0.088500	0.979000	3991.024800	5709.000000	0.699076	0.124724	2178.000000	0.890500	0.015556	0.724400	0.698000
4	1466647.229999990000000	-979138.219999990000000	-0.181800	0.801700	1451.216200	4599.000000	0.315550	0.171605	2644.000000	0.983500	0.029448	0.272300	0.268300
5	1994408.530000000000000	-962492.630000000000000	0.088500	0.979000	3991.024800	5709.000000	0.699076	0.124724	2178.000000	0.890500	0.015556	0.724400	0.698000
6	1994408.530000000000000	-962492.630000000000000	0.088500	0.979000	3991.024800	5709.000000	0.699076	0.124724	2178.000000	0.890500	0.015556	0.724400	0.698000
7	1441496.189999990000000	-1009975.699999990000000	-0.104600	0.748700	2781.115500	5679.000000	0.489719	0.092055	2362.000000	0.853300	0.008474	0.497400	0.477600
8	1460822.659999990000000	-982340.640000000000000	-0.192900	0.768400	1118.846000	5035.000000	0.222214	0.120994	2338.000000	0.961300	0.014640	0.194800	0.338800
9	1466550.580000000000000	-989317.819999990000000	0.024500	0.780500	1826.531100	5710.000000	0.319883	0.141352	2976.000000	0.756000	0.019980	0.296550	0.203200
10	1455495.739999990000000	-984257.520000000000000	-0.192900	0.715300	851.655700	4352.000000	0.195693	0.093907	1959.000000	0.908200	0.008819	0.179100	0.175700
11	1441496.189999990000000	-1009975.699999990000000	-0.104600	0.748700	2781.115500	5679.000000	0.489719	0.092055	2362.000000	0.853300	0.008474	0.497400	0.477600
12	1460822.659999990000000	-982340.640000000000000	-0.192900	0.768400	1118.846000	5035.000000	0.222214	0.120994	2338.000000	0.961300	0.014640	0.194800	0.338800
13	1456228.520000000000000	-986697.939999990000000	-0.192900	0.715300	1122.236500	5054.000000	0.222049	0.115165	2346.000000	0.908200	0.013263	0.194850	0.175700
14	1466647.229999990000000	-979138.219999990000000	-0.181800	0.801700	1451.216200	4599.000000	0.315550	0.171605	2644.000000	0.983500	0.029448	0.272300	0.268300
15	1441496.189999990000000	-1009975.699999990000000	-0.104600	0.748700	2781.115500	5679.000000	0.489719	0.092055	2362.000000	0.853300	0.008474	0.497400	0.477600
16	1460822.659999990000000	-982340.640000000000000	-0.192900	0.768400	1118.846000	5035.000000	0.222214	0.120994	2338.000000	0.961300	0.014640	0.194800	0.338800
17	2022254.770000000000000	-1047191.449999990000000	0.313900	0.936600	4282.083600	5711.000000	0.749796	0.058861	1853.000000	0.622700	0.003465	0.762600	0.784500
18	1994408.530000000000000	-962492.630000000000000	0.088500	0.979000	3991.024800	5709.000000	0.699076	0.124724	2178.000000	0.890500	0.015556	0.724400	0.698000
19	1455495.739999990000000	-984257.520000000000000	-0.192900	0.715300	851.655700	4352.000000	0.195693	0.093907	1959.000000	0.908200	0.008819	0.179100	0.175700
20	1460822.659999990000000	-982340.640000000000000	-0.192900	0.768400	1118.846000	5035.000000	0.222214	0.120994	2338.000000	0.961300	0.014640	0.194800	0.338800
21	1466550.580000000000000	-989317.819999990000000	0.024500	0.780500	1826.531100	5710.000000	0.319883	0.141352	2976.000000	0.756000	0.019980	0.296550	0.203200
22	1466647.229999990000000	-979138.219999990000000	-0.181800	0.801700	1451.216200	4599.000000	0.315550	0.171605	2644.000000	0.983500	0.029448	0.272300	0.268300

stats_1233

Figure 5: Excel file with the statistics (extraction).

6. Conclusion

The goal in using this tool was to examine the association between Rickettsia seroprevalence and the environmental variables circumscribing provinces where human data were collected. The results may alert for physicians consider testing for SFGR infections and treating clinically compatible cases as well as, alert public health authorities for this emerging infection. The application developed allowed to automatize: (1) the spatial data downloaded; (2) the products selection (NDVI and LST); (3) the NDVI and LST computation and; (4) the statistical analysis.

In the future, this application could be optimized in order to incorporate other remote sensed derived products that prove useful in this study case. In addition, this application can be used not only for this propose, but for many other applications where the optimization of these procedures is a demand.

The application is free, open source and available in www.fc.up.pt/pessoas/liaduarte/NDVI_LST.rar.

Acknowledgement Joana Afonso and Jocelyne Neto, Instituto Nacional de Saúde Pública, Ministério da Saúde, Maianga-Luanda, Angola for the support. USGS for satellite data.

References

Barradas, P.F., Vilhena, H., Oliveira, A.C., Granada, S., Amorim, I., Ferreira, P., Cardoso, L., Gartner, F., Sousa, R., 2017. Serological and molecular detection of spotted fever group Rickettsia in a group of pet dogs from Luanda, Angola. Parasites & Vectors 10(271). Doi:10.1186/s13071-017-2216-3.

Didan, K., 2015. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V006 [Data set]. NASA EOSDIS LP DAAC. doi: 10.5067/MODIS/MOD13Q1.006.

Heinrich, N., Dill, T., Dobler, G., Clowes, P., Kroidl, I., Starke, M., Ntinginya, N.E., Maboko, L., Loscher, T., Hoelscher, M., Saathoff, E., 2015. High Seroprevalence for Spotted Fever Group Rickettsiae, Is Associated with Higher Temperatures and Rural Environment in Mbeya Region, Southwester Tanzania. PLoS Negl Trop Dis 9(4): e0003626. Doi: 10.1371/journal.pntd.0003626.

Kalluri, S., Gilruth, P., Rogers, D., Szczur, M., 2007. Surveillance of Arthropod Vector-Borne Infectious Diseases Using Remote Sensing Techniques: A Review. Pathog 3(10): e116. doi: 10.1371/journal.ppat.0030116.

Khan, O.A, Davenhall, W., Ali, M., Castillo-Salgado, C., Vazquez-Prokopec, G., Kitron, U., Magalhaes, R. J. Soares, Clements, A.C.A., 2010. Geographical information systems and tropical medicine. Annals of Tropical Medicine and Parasitology 104(4), 303-318.

Kitron, U., 1998. Landscape ecology and epidemiology of vector-borne diseases: Tools for spatial analysis. Journal of Medical Entomology (4), 35, 435-445.

Palaniyandi, M., 2012. The role of remote sensing and GIS for spatial prediction of vector-borne diseases transmission: a systematic review. J Vector Borne Dis 49(4), 197-204.

PROJ contributors, 2018. PROJ coordinate transformation software library. Open Source Geospatial Foundation. Online, accessed 20-March-2018. URL https://proj4.org/.

Sahni, S.K., Narra, H.P., Sahni, A., Walker, D.H., 2013. Recent molecular insights into rickettsial pathogenesis and immunity. Future Microbiol 8(10), 1265–1288. doi:10.2217/fmb.13.102.

Stiles, J., 2000. Canine rickettsial infections. Vet Clin North Am Small Anim Pract 30(5), 1135-49.

Viana, J., Santos, J.V., Neiva, R.M., Souza, J., Duarte, L., Teodoro, A.C., Freitas, A., 2017. Remote Sensing in Human Health: A 10-Year Bibliometric Analysis. Remote Sens 9(12), 1225. doi:10.3390/rs9121225.

Walker, D.H., Fishbein, D.B., 1991. Epidemiology of rickettsial diseases. Eur J Epidemiol 7(3), 237-45.

Wan, Z., Hook, S., Hulley, G., 2015. MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006 [Data set]. NASA EOSDIS LP DAAC. doi: 10.5067/MODIS/MOD11A1.006.