

Geo-Analytic Functions for UN Field Operations – UN Open GIS: Spiral 3 Geo-Analysis

HaeKyong Kang¹, Timur Obukhov², Minpa Lee³

¹ Korea Research Institute for Human Settlements, Republic of Korea, hkkang@krihs.re.kr

² United Nations, obukhov@un.org

³ Mango System Inc., minpa.lee@mangosystem.com

Commission IV, WG VI/4

KEY WORDS: Open Geospatial Analysis, UN Open GIS Spiral 3, uDig Processing Toolbox, WPS Builder, WPS Request Builder

ABSTRACT:

This paper presents a case of developing open geospatial analysis functions and a processing toolbox as parts of UN Open GIS Spiral 3 activities. A comprehensive study focused on the use of geospatial analysis at the UN operations has been carried out, which covered various UN operations such as situational awareness, security operations, military operations, protection of civilians, epidemiological analysis, and so on. Four requirements on development of 1) primitive geospatial analysis functions, 2) processing toolbox or processing environment like WPS request builder of GeoServer, 3) model builder, and 4) management of the model builder were clarified through the study. By focusing the first and second requirements, a development architecture was proposed based on JAVA and GeoTools. The scope of implementation and its assessment were described. In addition, a use-case introduced a way of using geospatial functions developed in this research to analyse geostatistic factors that affect to occurrence of Ebola disease. Through the use-case study presented in the paper, I would like to share the experiences and lessons about the development of geospatial analysis functions, which would be used as a reference for other developments and applications.

1. INTRODUCTION

Geospatial analysis is an approach to applying statistical analysis and other analytic techniques to data which has a geographical or spatial aspect (Hemakumara, 2015). The geospatial analysis would employ software capable of rendering/visualizing maps, processing spatial data, and applying analytical methods to geospatial datasets (Geospatial, 2012). It provides many advantages to users, for instance, easy to recordkeeping of geographical changes, improvement of communications between organizations as the visualized format is easily understood by all, and easy of managing geographically with knowing what is and will be occurring.

The United Nations also uses geospatial analysis to improve the efficiency of situational awareness, safety and security or monitoring crisis during UN field operations. UN Geospatial Services (UNGIS) provide overall supports in using the geospatial analysis at the UN operations including development of strategies, technologies, professional capacity and etc (Eom, 2018).

Free and open source software for geospatial (FOSS4G) is a geospatial software that freely licensed to use, copy, study, and change the software in any way, and the source code is openly shared so that people are encouraged to voluntarily improve the design of the software (FOSS, 2018). In the past, it was mentioned often disadvantage such as technical support weakness or SW incompleteness relatively in comparison with commercial SWs. However, the SW quality is getting robust and other advantages such as easy to develop/add a new function due to flexibility and scalability of FOSS4G, and easy to share/transfer technologies between UN and host nations due to free-license increased the demands in using FOSS4G at UN operations.

In 2016, UN Open GIS Initiative, led by UNGIS under UN Department of Operational Support (UNDOS, 2019), was established to leverage free and open source-based geospatial technologies in support of UN peace operations. The third technical working group of the UN Open GIS initiative, named Spiral 3, aims to develop and deliver geo-analytic functions to the UN operations.

This paper presents a development practice of geospatial analysis services in UN Open GIS Spiral 3. For this, a comprehensive study on the use of geospatial analysis at the UN operations has been carried out, which covers the following UN operations; situational awareness, security operations, military operations, protection of civilians, human rights violation investigation, epidemiological analysis, and so on. Based on these requirements for geospatial analysis, development architecture was designed in section 3 including implementation. In section 4, a use-case called *Ebola Predictive Model* was presented to show how geospatial analysis function developed by UN Open GIS Spiral 3 can be applied on Ebola crisis.

2. GEO-ANALYSIS FOR UN OPERATIONS

2.1 UN operations and geo-analysis

UNGIS (UN Geospatial Services) is responsible of geospatial information services to the wide range of UN operations from sharing situational awareness and monitoring activities to assuring the interoperability among the UN missions, the Secretariat, and other UN agencies. Geo-analytical tools have been widely used by UNGIS experts to support UN operations.

Specifically, geo-analytic tools are used to support the fulfilment of core mission mandates, to assist the senior

managers in decision making process, to support planning and operations, to enhance situational awareness and so on. More UN operations using geo-analysis are listed in Table 1 (Timur 2017).

UN operations	Use of geo-analysis
Situational Awareness	Incident Analysis, Destruction Detection
Safety and Security	Security and Threat Analysis
Military Operations and Protection of Civilians	Intelligence for Early Detection of Threats
Human Rights Violation Investigation	Atrocities Assessment, Mass Graves Identification
Epidemiological Analysis	Epidemiological Analysis
UN Mission Support	Planning for Camps' Establishment, Management of Mission's Utilities, Groundwater Exploration, Water pipeline monitoring
Environmental Assessment	Compliance of Rules and Regulations of Host country and UN Departmental guidelines, Waste Management
Cross-cutting Analysis	Threat Analysis of Road Operations

Table 1. UN operations and geo-analysis

Situational Awareness: UNGIS provides various incident analysis products to support situational awareness and informed decision making. Due to the different natures of crisis, each mission has different requirements for incident data collection, situational awareness, and incident analysis. It is therefore required to customize analytical models. UNGIS also provides various destruction detection products in response to natural and man-made disasters to support search/rescue operations and reconstruction or to confirm the reported events. Rapid assessment of the extent and severity of natural and man-made disasters is often critical. Change detection analysis using optical/radar image is required.

Safety and Security: UNGIS provides simulation of different threat scenarios for UN Premises in the field. Each security situation on the ground requires different type of analysis and modeling tools.

Military Operations and Protection of Civilians: UNGIS provides GeoINT(Geospatial Intelligence) products for early detection of threats to support military operations and Protection of Civilians (POC). Military operations require various types and aspects of terrain and road analysis for patrol route planning, evacuation, identification of possible ambush areas, etc.. Protection of civilians requires analysis of possible threats to the population of concerned, preparedness and management of POC camps.

Human Rights Violation Investigation: UNGIS provides atrocities assessment to identify and assess the extent and severity of various atrocities and mass graves identification to detect locations of possible mass graves in conflict areas.

Epidemiological Analysis: UN responds various epidemiological crisis such as Ebola crisis. Analyzing spread, extent, occurrence, and severity depends on the type disease. Different epidemics have different spread of infection, so that the analysis and prediction of each epidemic are different. As

a consequence, various analytical functions and customizable tools are required.

UN Mission Support: When UN starts a new operation, the mission support colleagues need to identify the best suitable location for a UN Camp, or IDP or Refugee camps, so that it has easy access, limited threats considering the terrain in the vicinity, sustainability during the raining season, close proximity to suitable places for helicopter landing sites, etc.. UNGIS responds to these requests by using analytical products and services such as suitability analysis, 3D feature simulation, accessibility, tracking for assets movement, helicopter landing sites, runway establishment, and detecting a location of electricity lines and optical-fiber cables damage. In various missions such as Darfur, Mali and Somalia, UNGIS provides assistance in identification of possible reserves of ground water. Analyses and processed information provides location possible locations of ground water reserves not only for the UN peacekeepers, but also for the local communities. Also, in order not to waste water in the mission, UN assists out Mission Support colleagues with monitoring and analysis of damages water pipes.

Environmental Assessment: UN has an obligation to compliance of rules and regulations published by host countries or provided by UN itself to protect environment of the area where UN serves. So it is required reductions of the footprint, facility management, water-waste management such as proper collection and disposal of waste water, biohazard materials, and etc. UNGIS provides GIS analysis services to the Environmental Officers in the missions and the HQ which would assist them to conduct an environmental assessment.

2.2 Geo-analysis requirements

Based on the observation of geo-analysis in UN Operations, six requirements were summarized:

- Primitive analysis functions should be developed for both vector and raster data.
- Customized generic analysis functions are required in a form of "building block" consisting of mission-specific geo-analysis functions since each UN operation has different requirements for analysis tools.
- Capability is a crucial factor to build analysis models based on a specific analysis task.
- Analysis models shall be supported by desktop and web environment.
- It is required to provide support and research capability to the UN upon specific requests.
- It is also necessary to conduct a research and determine a possibility to deploy analytical functions and models on hybrid model (proprietary and open source platforms)

3. DEVELOPMENT OF SPIRAL 3 - GEO ANALYSIS

3.1 Architecture

Spiral 3 has four major components; primitive geospatial analysis functions, processing of the primitive geospatial analysis functions, model-builder and management of the model builder. Among the four components, Spiral 3 focused on developing primitive geospatial analysis functions and its processing. In current, the model-builder and its management

were excluded from the development scope of Spiral 3. As alternative solutions, there are WPS Model Builder and WPS Request Builder of GeoServer for chaining one of geospatial analysis functions.

The system development environment was determined as Java using GeoTools for the compatibility and portability. In Figure 1, a library of geospatial analysis functions (①) was developed by using Java and GeoTools library. It is required that java compatible processing application environment to execute the geospatial analysis functions, to handle parameters, to import/export of input datasets and so on.

For processing the library of geospatial analysis functions through Internet, OGC WPS extension should be developed. So the WPS extension of GeoServer (②) was developed by Spiral 3. For processing it on a desktop environment, a processing toolbox need to be developed. In Spiral 3, the processing toolbox was developed as a plug-in (③) of uDig SW which is a Java-based open source desktop GIS. It means that the preparation and pre-processing of data for geo-analysis and the visualization of the analysis are supported by uDig.

For processing the library of geospatial analysis functions on Spiral 1 (④) of UN Open GIS platforms which is a geoportals service, a processing toolbox or an extension should be developed. However, the Spiral 1 is not compatible with Java language, the fully-integrated development was excluded from the scope of Spiral 3 development.

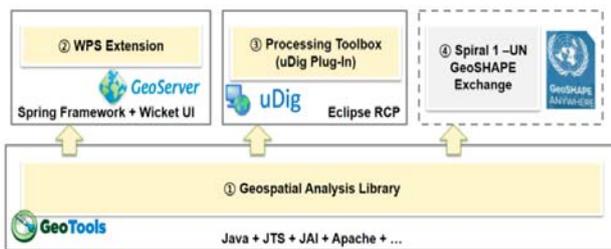


Figure 1. Key components of Spiral 3 development

3.2 Implementation

Development of around 140 primitive geospatial analysis functions was done by Spiral 3. Parts of the analysis functions in comparison among GIS SWs are presented in Figure 2. Each function in the library (① in Figure 1) is executable by plug-in the library to uDig SW. All of the primitive functions will be shown up on the Processing toolbox of uDIG in Figure 3.

Figure 2. Comparison parts of geospatial analysis functions among GIS SWs

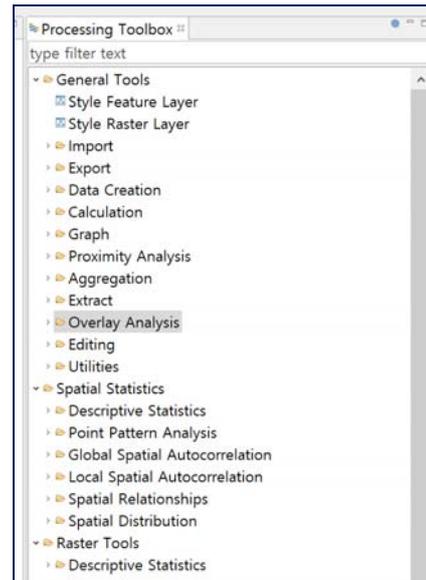


Figure 3. Geo-analysis functions of Spiral 3 on uDig SW

WPS Extension of GeoServer is available to download and use. When the WPS Extension is installed successfully, the WPS will be shown up on a list of service capabilities of Welcome page of GeoServer in Figure 4. All of Spiral 3 geospatial analysis functions can be checked through WPS request builder menu of GeoServer in Figure 5.



Figure 4. GeoServer WPS extension installed

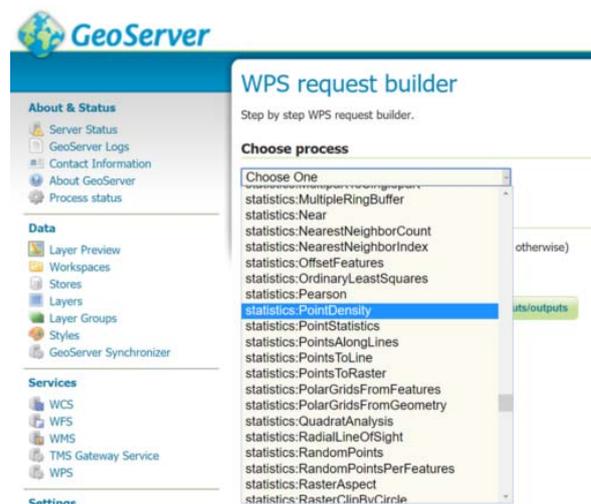


Figure 5. Spiral 3 geospatial analysis functions on GeoServer

3.3 Assessment of the implementations

Geospatial analysis functions of Spiral 3 implemented open standards or open algorithms published as well as ArcGIS documents. The correctness of each geospatial analysis function of Spiral 3 was verified empirically. Some of functions were verified in comparison with ones of other SW like ArcGIS.

Although the name of functions is equal, the results of geospatial analysis functions are often inequivalent depending on algorithms or parameters implemented. For instance, the results from *Quadratic* function of ArcGIS and one of Spiral 3's are different in Figure 6. So it does not mean that either one is wrong even though the results of the two functions are different.

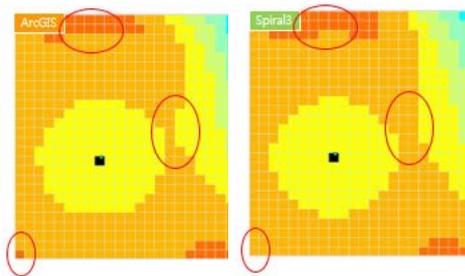


Figure 6. Quadratic Comparison: ArcGIS (left), Spiral 3 (right)

4. USE-CASE OF USING SPIRAL 3 GEO-ANALYSIS: EBOLA PREDICTIVE MODEL

4.1 Ebola Dataset

A sample dataset used in this Ebola Predictive Model covering parts of Guinea, Sierra Leone and Liberia in Africa consists of six layers; *Districts*, *Airports*, *Community Care Center (CCC)*, *Ebola Treatment Center/Units (ETC)*, *Logistic Base (LBS)* and *Roads*. Metadata including Geometric type and Coordinate Reference System (CRS) is summarized in Table 2. The geometric features of Ebola dataset displayed in Figure 7.

Layer	Description	CRS	Type
<i>districts</i>	Guinea, Liberia, Sierra Leone	3857	Polygon
<i>airports</i>	Airports in the countries	3857	Point
<i>ccc</i>	Community Care Center	3857	Point
<i>etc</i>	Ebola Treatment Center and Units	3857	Point
<i>lbs</i>	logistic bases	3857	Point
<i>Roads</i>	Roads	3857	Line
<i>Maplink</i>	OpenStreetMap	3857	TMS

Table 2. Metadata of Ebola dataset

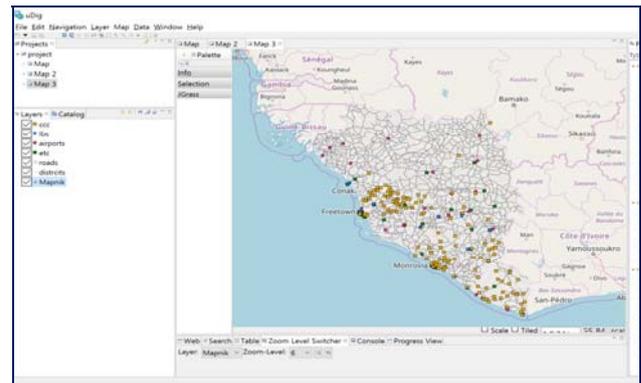


Figure 7. Ebola Dataset displayed on uDig SW

4.2 Ebola Predictive Model

The *Ebola Predictive Model* in Figure 15 used in this paper was provided by the UN. Geospatial analysis functions of ArcGIS SW written in the *Ebola Predictive Model*. Each function in the *Ebola Predictive Model* corresponds to a geospatial analysis function of Spiral 3. For instance, Select, Merge, Split, and Clip correspond to the same of geospatial analysis functions of Spiral 3. Some of these operations are integrated into one operation. While

Whereas analysis sequence of a road layer consists of three steps (Select → Clip → Line Density) in the Ebola Predictive Model in Figure 15, it consists of two steps (Point in Polygon → Line Density) in Feature 8. It is because Line Density function of Spiral 3 automatically executes Clip operation as a parameter of Line Density. Since the result can be generated by executing a less number of geospatial analysis functions, it becomes an advantage of the Spiral 3 approach in terms of user convenience.

The *Ebola Predictive Model* is simplified as shown in Figure 8 by using geospatial analysis functions of Spiral 3. To execute the *Ebola Predictive Model* in Figure 8, Point in Polygon operation for vector dataset, Line Density and Zonal Statistics for raster dataset and Ordinary Least Squares Regression (OLS) for spatial statistics in Spiral 3 were applied.

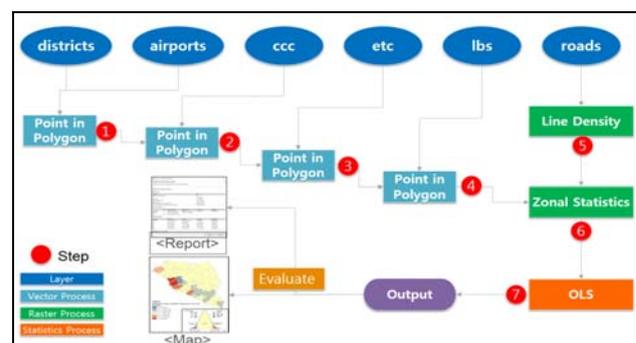


Figure 8. Simplified Workflow of Ebola Predictive Model

4.3 Analytic Process

The first step of simplified Ebola Predictive Model in Figure 8 is Point in Polygon overly function. It creates a new district layer by adding a new column with a number of *airports* in each district.

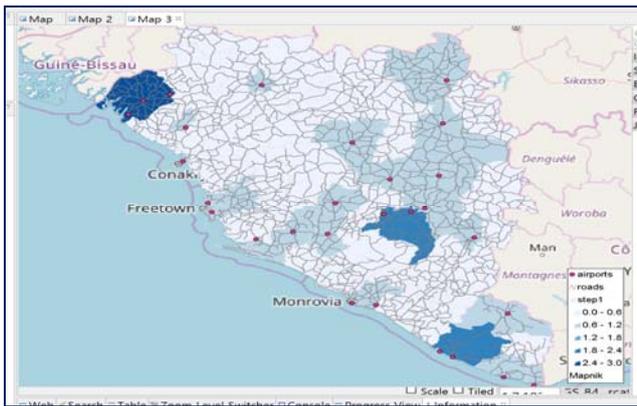


Figure 9. Point in Polygon Result: *districts*

The processes from the second to the fourth steps are the same as the first one except parameters. Once the fourth step is done, each district will have columns with each number of *airports*, *LBSs*, *CCCs*, and *ETCs*.

The fifth step is for the Line Density of Spiral 3. When the Line Density function is selected from the Processing Toolbox, which contains the geo-analysis functions of Spiral 3, it requests line feature as an input data, which is a road layer in this use case. The Line Density function automatically executes Clip operation by parameterizing. For instance, *districts* were set as an extent for clipping during the Line Density process in the Figure 10.

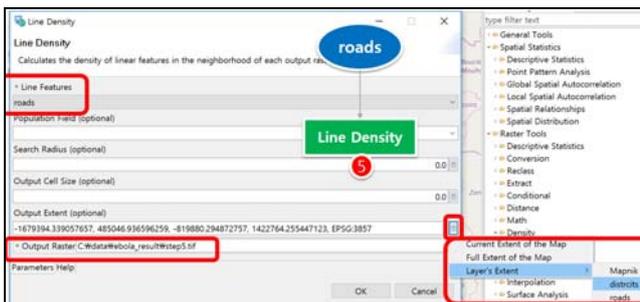
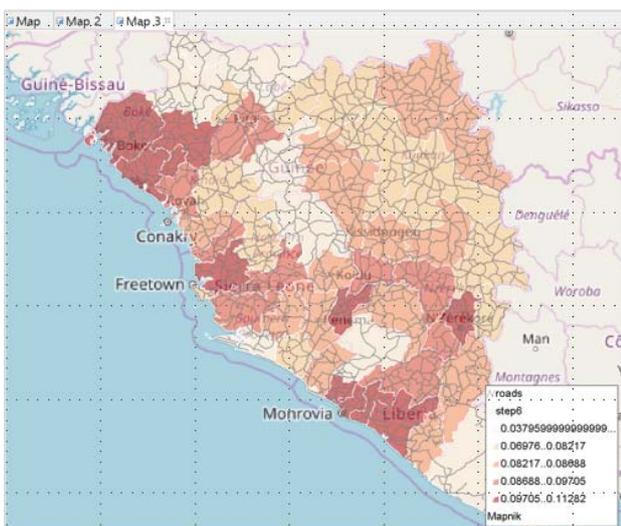


Figure 10. Parameters for Line Density

The sixth step is Zonal Statistics. It calculates the road density of each district.



country	whocs...	location	last7d...	previo...	total14	airports	ccc	etc	lbs	roads	Cell_A...
Sierra ...	BO	BO	2	4	6	1	0	1	0	0.0854...	5.5555...
Sierra ...	TONK...	TONK...	3	9	12	0	13	1	0	0.0965...	6.6311...
Sierra ...	KOIN...	KOIN...	0	1	1	0	2	0	0	0.0669...	1.2683...
Sierra ...	KONO	KONO	13	14	27	1	5	0	0	0.0842...	5.4964...
Sierra ...	KAILA...	KAILA...	0	0	0	0	0	1	1	0.1098...	4.0779...
Sierra ...	KENE...	KENE...	1	0	1	1	0	1	1	0.0844...	6.5129...
Sierra ...	PUJEH...	PUJEH...	0	0	0	0	0	0	0	0.0820...	4.2316...
Sierra ...	BONT...	BONT...	0	0	0	1	0	0	0	0.0697...	3.5578...
Sierra ...	MOYA...	MOYA...	4	4	8	0	0	1	0	0.0967...	6.9857...

Figure 11. Result of Zonal Statistics: districts with road density

The last step is Ordinary Least Squares (OLS) Linear Regression. The OLS calculates how the facilities such as *airports*, *logistic bases* (lbs), *Ebola Treatment Centers* (etc), *community care centers* (ccc), and *roads* affect Ebola occurrence. The facilities therefore become explanatory variables in the OLS step. Ebola occurrence is coded as a dependent variable *total14* in Figure 12.

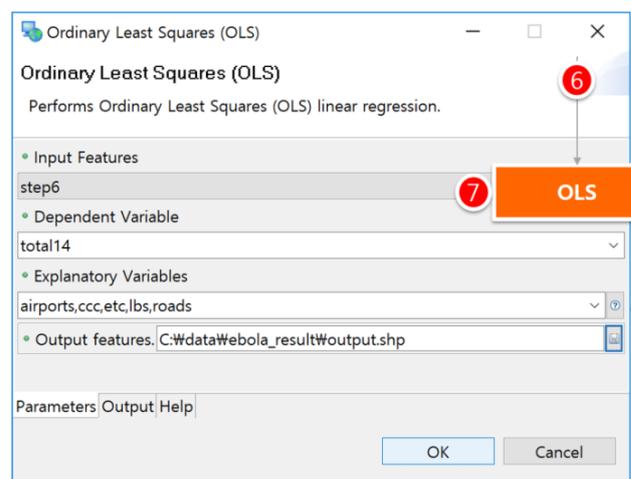


Figure 12. Parameters for OLS

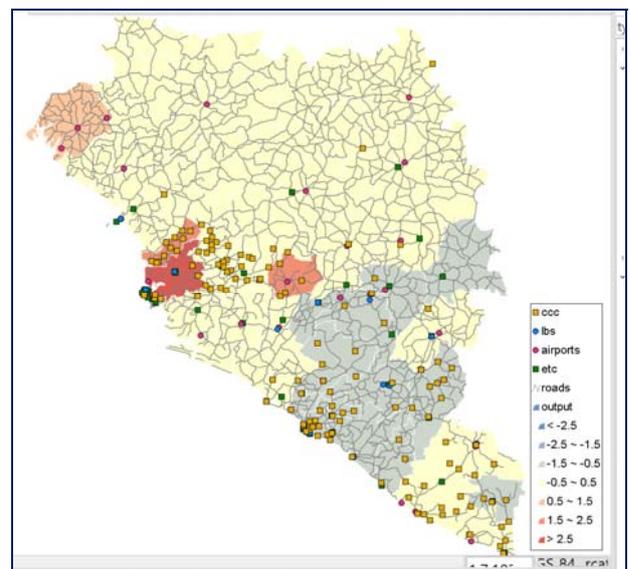


Figure 13. Output of OLS: districts re-classified with residual standard deviation

Category	Value
Number of Observations	63
R	0.61134641
R-Squared	0.37374444
Adjusted R-Squared	0.31880974
Standard Error	13.0316903
Akaike's Information Criterion (AIC)	509.97139471
Corrected Akaike's Information Criterion (AICc)	512.00775834

	Dof	Sum of squares	Square mean	F-Statistic	Probability
Regression	5	5776.96185836	1155.39237167	6.80343117	0.0004889
Residual	57	9680.02226863	169.82495208	-	-
Sum	62	15456.98412698	-	-	-

Variable	Coefficient	StdError	t-Statistic	Probability
Intercept	-0.6777681	7.21943548	-0.09388224	0.92553019
airports	-2.3747474	2.6635323	-0.89157835	0.37636373
ccc	-0.26228967	0.47090173	-0.5569945	0.57971288
etc	6.86737869	1.23616146	5.55540593	0.00000085
lbs	-0.18247019	2.6191792	-0.06966694	0.94470069
roads	24.06584497	66.15313771	0.36378992	0.71736428

Figure 14. Statistical Report of OLS

4.4 Lessons Learned from the Use-Case

From this use-case, we learned the following lessons. First, the clean data set is essential for accurate results of the analysis. Noisy and incorrect data not only degrade the accuracy of the results but also make the analysis process time consuming due to the data cleaning steps.

Second, user-friendly interface and functions facilitate the geo-analytical process. Considering most users of the geo-analytic tool are not GIS experts, user-friendly interfaces such as model builder make the tool more accessible to wider user group.

Third, we have to validate the result of our geo-analytic tool in comparison with proprietary commercial tool. During this use-case, we found that the results of some analysis by our tool slightly differ from those with a proprietary commercial tool. Unfortunately, the internal approach and methods used for geo-analytic functions in the proprietary commercial tool are not accessible.

5. CONCLUSIONS

In this paper, a practice developing a JAVA library for open geospatial analysis functions and a processing toolbox was presented. For this, it was reviewed how geospatial analysis can be used appropriately in different UN operations such as situational awareness, security operations, military operations, and epidemiological analysis through a comprehensive study. It successfully contributed to clarify four user requirements which were development of 1) primitive geospatial analysis functions, 2) processing toolbox or processing environment like WPS request builder of GeoServer, 3) model builder, and 4) management of the model builder.

By focusing the first and second requirements, a development architecture was proposed based on JAVA and GeoTools. More than 140 primitive geospatial analysis functions were implemented as a java library in compliant with OGC WPS standard. In addition, the use of Processing Toolbox for uDig desktop SW and WPS Extension of GeoServer was proposed for user convenience.

Through the use case of Ebola Predictive Model, we found advantage of our geospatial analysis functions in terms of user

convenience. For instance, the Line Density function automatically executes Clip operation by parameterizing in its process. It reduces numbers of analysis functions to be executed and consequently becomes to improve user convenience.

We will keep trying to develop other user-friendly components and as many use-cases as possible and to improve the quality of analysis results to be the same with the results of commercial SWs.

ACKNOWLEDGEMENTS

Parts of this research were supported by a grant (14NSIP-B080144-01) from National Land Space Information Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

REFERENCES

- Eom K.S. (2018). Challenge & Opportunity: UN open GIS initiative. Keynote Presentation at FOSS4G 2018 Dar es Salaam.
- FOSS. (2018). What is free software? The Free Software Definition". The GNU Project -- GNU.org. 2018-06-12. Retrieved 2018-09-15.
- Geospatial. (2012). Collins English Dictionary - 11th Edition. Retrieved on 5th August 2012 from CollinsDictionary.com at <http://www.collinsdictionary.com/dictionary/english/geospatial>
- Hemakumara, GPTS, & Rainis, Ruslan. (2015). Geo-statistical modeling to evaluate the socio-economic impacts of households in the context of low-lying areas conversion in Colombo metropolitan region-Sri Lanka. Paper presented at the AIP Conference Proceedings.
- Spiral 3 Manual, (2019). GeoServer WPS 1.0 User Manual from github.com at https://github.com/mappplus/spatial_statistics_for_geotools_udig/blob/master/docs/manual/GeoServer_WPS_1.0_User_Manual_en_v.2.latest.pdf
- Spatial 3 library, (2019). OpenGXT –Spatial Analysis Tools from sourceforge.net at <https://sourceforge.net/projects/mango-spatialstatistics>.
- Timur O. Eom K.S., Koh D.J., Florent G. Alphonse A., (2017). UN Open GIS Spiral 3 Requirements. Presentation at UN Open GIS Initiative Spiral 3 (GeoAnalysis) Workshop.
- UNDOS. (2019). What We Do: Technology: Partnership Initiatives: Ongoing Initiatives: UN Open GIS Initiative at <https://operationalsupport.un.org/en/partnership-initiatives>.

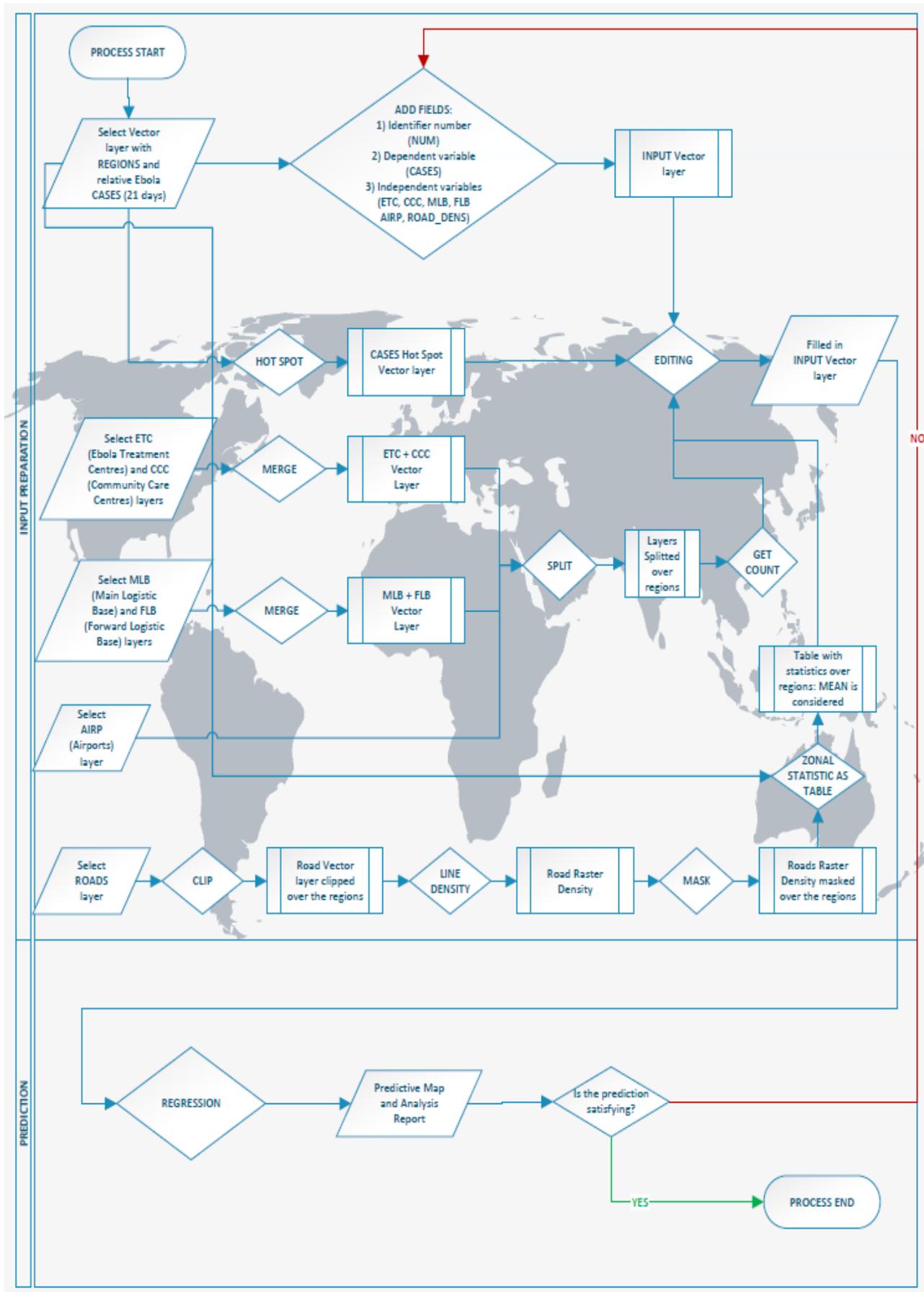


Figure 15. Ebola Predictive Model