Recent Trends in Civil Engineering and Built Environment Vol. 3 No. 1 (2022) 167-178
© Universiti Tun Hussein Onn Malaysia Publisher's Office

RTCEBE

Homepage: http://publisher.uthm.edu.my/periodicals/index.php/rtcebe
e-ISSN :2773-5184

The Development of Data Management in Geographic Information System (GIS) for Glenmarie Heights, Johor Bahru, Malaysia

Siti Khadijah Salihin[1], **Zaihasra Abu Talib**[2]*

[1]Faculty of Civil Engineering and Built Environment,
Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, 86400, MALAYSIA

*Corresponding Author Designation

DOI: https://doi.org/10.30880/rtcebe.2022.03.01.019
Received 4 July 2021; Accepted 13 December 2021; Available online 15 July 2022

Abstract: Geographic Information System (GIS) mapping helps visualizing and identifying patterns that are hard to determine if the data elements are in table format. To design a map required a wide range of data and information file that come from difference types of sources. Dealing with a big data can be difficult and complicated to manage and organize. Thus, a proper management is very importance in order to defeat the risks of data missing, time consuming and much easier to make comparison with existing data if required since all the data are store safely. The main focus of this study is to develop an effective and efficient data management in GIS. An effective and efficient management of data is very importance because it can cause more productivity since shorten time finding, cost efficiency, reduced data loss when all the data are stored safely in place and make accurate decisions by analyze the correct and most recent data and information. There are two types of software involved in this research which are Quantum GIS(QGIS) and Microsoft Office Excel. With this combination of two software, the GIS data management workflows have been developed to be better in term of collect, store, maintain, prepare, and share data not only to the organization but also for those who are under construction or development of study area.

Keywords: Geographic Information System, Quantum GIS, Data Management, GIS Mapping

1. Introduction

Geographic Information System (GIS) is a system that manage the information, analyse and display a range of spatial and geographic data and visualize mapping. Previous researchers have defined Geographic Information system (GIS) in many ways. Geographic Information System (GIS) is developed to manage and analyze spatial data, which is based on geomatics technologies [1]. GIS as a technology or system allows the storage of spatial information in the relational database, and, as a science, is also beyond data storage system. The attribute information associated with spatial features stored in the database allows for further spatial analysis using both the spatial and non-spatial attributes

*Corresponding author: zaihasra@uthm.edu.my

[1]. Geographic information system (GIS) has emerged as a powerful tool for storing, analyzing, and displaying spatial data and using these data for decision making in several areas including engineering and environmental fields[2] and a system can provide an intelligent inter-face for data storage, database access, and data display using the GIS[3].

With this system, the jobs as researcher, engineer and contractor much easier because all the data and information needed is provide in the system. GIS mapping helps visualizing and identifying patterns that are difficult to see if the data elements are in table format. It also helps identify patterns that appear when viewing two or more data sets together. This is a particular type of mapping technology that enables information bound by geographic points to be organised. This GIS mapping allows to see customised data layer combinations with dynamic tools instead of only looking at a few main features on a static map. To produce a complete map using GIS required data from various sources. There are many previous project that have been done using GIS mapping. For example select a site for the building of a new hospital in Iskandar Malaysia using GIS-based Multi-Criteria Analysis (MCA) with consideration of various factor criteria, and constraint criteria in order to balance the medical resource there[4], the current study utilized Kepong Herbarium records, GIS-based multi-criteria decision making and analytical hierarchy process to identify IPA in the state of Johor, Malaysia[5] and the cross-validation of a multivariate logistic regression model using remote sensing data and GIS for landslide hazard analysis on the Penang, Cameron, and Selangor areas in Malaysia[6].

However, in order to produce a map, this system have to accommodate with a wide range of data and information file that come from various types of sources. The large volumes of data can be difficult and complicated to manage, and organize. A proper management for data collected is very importance to overcome highly risk of lost or missing data, reduce time taken to find the data and much easier to analyse the current and existing data. The previous study had mention that large volumes of subsurface data can be difficult to manage, and analysing it can be even more difficult. Trying to find spatial patterns simply by viewing tabular data can be nearly impossible [7]. Data management involves the collection, storage, organize, protection, verifying, and processing of important data. An effective and efficient management of data is very importance because it can cause (1) productivity since shorten time finding, understanding and relaying information, (2) cost efficiency by avoiding unnecessary extra cost to conduct same research again, (3) reduced data loss when all the data are stored safely in place and (4) make accurate decisions by analyze the correct and most recent data and information. Thus in this research, the objective is to develop an effective and efficient data management in GIS that can help to keep everything under control.

2. Materials and Methods

The materials used to obtain the results of the study are data mapping of the study area and software to analyse data. Meanwhile, the methods that have been implemented in order to carry out this research is data classification method.

2.1 Materials

The materials used to obtain the results of the study are data mapping of the study area and software to analyse data.

Data types and sources

The data used in this study are taken from the Geographic Information Systems (GIS) Mapping for soil classifications data in Glenmarie Heights on lot 99396, Ptd 68905 and Ptd 68903, Mukim Pulai, Johor Bahru, Johor Darul Takzim Malaysia with coordinate of 1° 56' 3.8400" N and 103° 21' 31.4172" E. Figure 1 show the study area in Glenmarie Heights.

Figure 1: Glenmarie Heights Johor

To design the mapping, data were collected from JMG Johor and JUPEM. There are many types of data that have been taken from the GIS mapping as in Figure 2 that shown the attributes data table. For this research only four main types of data were analyzed for data classification which are moisture content, Atterberg limit, particle size distribution and pH value.

BOREHOLE — Features Total: 26, Filtered: 26, Selected: 0

	NO BH	LATITUDE	LONGITUDE	BULK DENSITY	DRY DENSITY	MOISTURE CONTENT	LIQUID LIMIT	PLASTIC LIMIT	PLASTIC INDEX	GRAVEL	SAND	SILT	CLAY	pH VALUE
1	1	1.588016389	103.7516386	1.859	1.4075	32.5	43.5	26.5	17	7	46.5	19	27.5	4
2	2	1.587643333	103.7550203	2.061	1.512	36	53.5	28.5	25	1	32	35	32	4.3
3	3	1.590133056	103.7507883	2.0435	1.4075	46.5	65.5	32	33.5	5	35.5	23	36.5	4.4
4	4	1.593096944	103.7527758	1.936	1.507	28	39.5	25	14.5	19	28.5	21	31.5	4.2
5	5	1.573571111	103.7787044	1.7645	1.309	35	61	31	30	34	23.5	16	26.5	4.7
6	6	1.572335278	103.7805636	1.846	1.339	38	61	31	29	18	43.5	19	19.5	4.9
7	7	1.591184444	103.7564022	1.914	1.335	43.5	61	30.5	30.5	1.5	19	32	47.5	4.4
8	8	1.588909167	103.7597061	1.819	1.246	46.5	60	30.5	29.5	2	46.5	22.5	29	4
9	9	1.590791111	103.7622247	1.9425	1.511	29	39.5	25	14.5	10	50.5	26	13.5	4.3
10	10	1.592414444	103.7600906	1.7755	1.3395	35.5	71	33	38	13.5	49.5	25	18.5	4.8
11	11	1.576238056	103.7814189	1.9105	1.3665	40	59	31	28	2	18	34.5	45.5	4.5
12	12	1.5928925	103.7624025	1.867	1.441	30	45	26.5	18.5	15.5	36	25	23.5	4.1
13	13	1.589253333	103.7661286	1.8565	1.4385	29	48	27	21	4.5	39	28	28.5	4
14	14	1.592166667	103.7657319	1.7605	1.354	30	42.5	26	16.5	14.5	41	20.5	24	4
15	15	1.589006111	103.7700986	1.9185	1.5575	23	33	23.5	9.5	31.5	28	21.5	19	4.2

Show All Features

BOREHOLE — Features Total: 26, Filtered: 26, Selected: 0

	NO BH	LATITUDE	LONGITUDE	BULK DENSITY	DRY DENSITY	MOISTURE CONTENT	LIQUID LIMIT	PLASTIC LIMIT	PLASTIC INDEX	GRAVEL	SAND	SILT	CLAY	pH VALUE
16	16	1.590320833	103.7717725	1.915	1.5485	24	31	23	8	28	43.5	20.5	14	4.2
17	17	1.594315556	103.7707492	1.749	1.4285	22.5	32.5	23.5	9	26	33.5	23.5	17	4.1
18	18	1.594551111	103.7678286	1.7735	1.4115	22	30	43	7331.5	35	22	22	11.5	4.1
19	19	1.595215278	103.7653342	1.7735	1.804	25.5	35.5	24	11.5	16.5	46	18	19.5	4
20	20	1.599041111	103.7589133	1.975	1.5005	31.5	50.5	18	32.5	5.5	32.5	26.5	35.5	4
21	21	1.599750278	103.7634467	1.8005	1.4565	23.5	35	24	11	23.5	42	21	13.5	4.2
22	22	1.57968	103.7814875	2.0815	1.6525	26	40	25	15	19.5	33.5	16	31	4.5
23	23	1.604023333	103.7628417	1.91	1.61	18.5	28.5	23	5.5	17.5	43	24	15.5	4.2
24	24	1.605601389	103.7598778	2.128	1.6595	27	38	24.5	13.5	17	37	21	25	4.2
25	25	1.604801667	103.7571583	1.813	1.3845	31	46	26.5	19.5	19.5	34	26.5	20	4
26	26	1.605351667	103.7518717	1.896	1.485	27.5	37.5	24.5	13	15	43	20	22	4.4

Show All Features

Figure 2: The attributes data table

<u>Software</u>

In this research, there are two types software that have been used to accomplish the objective of the study which are:

- QGIS
- Microsoft Excel.

2.2 Methods

Producing a map required a huge amount of data. A proper data management is very importance to have a better data organization and storage. In this research, the data have been managed and classified in a certain classification. .Classification method is used to simplify from a group of huge amount of data into several small group that only consist of specific data. In previous study said that classification technique is capable of processing a wider variety of data than regression and is growing in popularity [8]. This method is the most suitable for data management when it comes to deal with numerous volume of data because it is much more easier to identify the required data in a small group compared to search throughout a group that consist with many unnecessary data. There are four main types of data involved which are:-

- Soil moisture content conditions
- Types of soils based on plasticity index
- Soil texture
- pH value of soil

(i) Moisture Content

Malaysian Meteorological Department, MMD in 2007 classified soil moisture content into several categories [9] as shown in Table 1. By using the table below, the data of moisture content are classify into the type soil condition.

Table 1: Classification of Soil Moisture Content Conditions

Soil Moisture Content	Condition
>30%	Very Wet
25 - 30%	Wet
20 - 25%	Moderate
15 - 20%	Dry
<15%	Very Dry

(ii) Atterberg limit

The plasticity index is used as a basic feature of soil in soil classification and in different correlations with other soil properties. In the previous research, based on the plasticity index, the soils were classified by Atterberg, Table 2 shows the classification of soil type based on the percentage of plasticity index.

Table 2: Types of soils based on plasticity index [10]

Plasticity index (%)	Soil type	Degree of plasticity	Degree of cohesiveness
0	Sand	Non-plastic	Non-cohesive
<7	Silt	Low plastic	Partly cohesive
7-17	Silt clay	Medium plastic	Cohesive
>17	Clay	High plastic	cohesive

(iii) Particle size distribution

The texture of the soils by plotting the percentage ratio of sand, silt and clay using the soil texture triangle [11]. Figure 3 shows one of the method that used to classify the texture of soil by using the percentage present of silt, sand and clay in the study area.

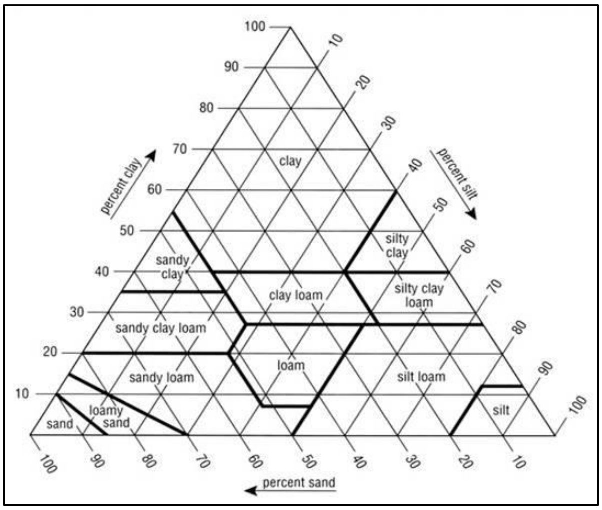

Figure 3: The Soil Texture Triangle

(iv) pH value

Soil can be acidic or alkaline naturally, and by measuring its pH value, this can be determined. The soil pH value classification in the study area used the same range as in the previous study below. that soils can be classified according to their pH value [12] :

- over 7.5 - alkaline
- 6.5 to 7.5 - neutral
- less than 6.5 - acidic

3. Results and Discussion

Data classification method help users to analyze and identify which location of the borehole is involved in that required data group. This method categorized the attribute data layers into a certain categories based on their data properties. As the result, the features of borehole location points were appeared in a different color on the map according to their group of classification.

3.1 Moisture Content

The soil moisture content was classified into several types of soil conditions which are very wet, wet, moderate, dry, and very dry. Table 3 represents the classification of data for moisture content based on the soil condition.

Table 3: The classification of moisture content

No. BH	Soil Moisture Content (%)		Condition
	SI report	Classification	
1	32.5	> 30%	Very Wet
2	36.0		
3	46.5		
5	35.0		
6	38.0		
7	43.5		
8	46.5		
10	35.5		
11	40.0		
20	31.5		
25	31.0		
4	28.0	25 - 30%	Wet
9	29.0		
12	30.0		
13	29.0		
14	30.0		
19	25.5		
22	26.0		
24	27.0		
26	27.5		
15	23.0	20 - 25%	Moderate
16	24.0		
17	22.5		
18	22.0		
23	18.5	15 - 20%	Dry
-	-	<15%	Very Dry

Figure 4 shows the map with the borehole location that has been categorized based on the moisture content classification. From Table 3 and Figure 4, the analyzed data samples recorded that 11 out of 26 contained more than 30% of moisture, 9 samples in between 25 to 35%, 4 samples in 20 to 25%, 1 sample in the range of 15 to 20% and none of the data samples has less than 15% of moisture content. Thus, it can be concluded that the soil condition of the study area was very wet since the data samples contributed about 42% in that category of more than 30%.

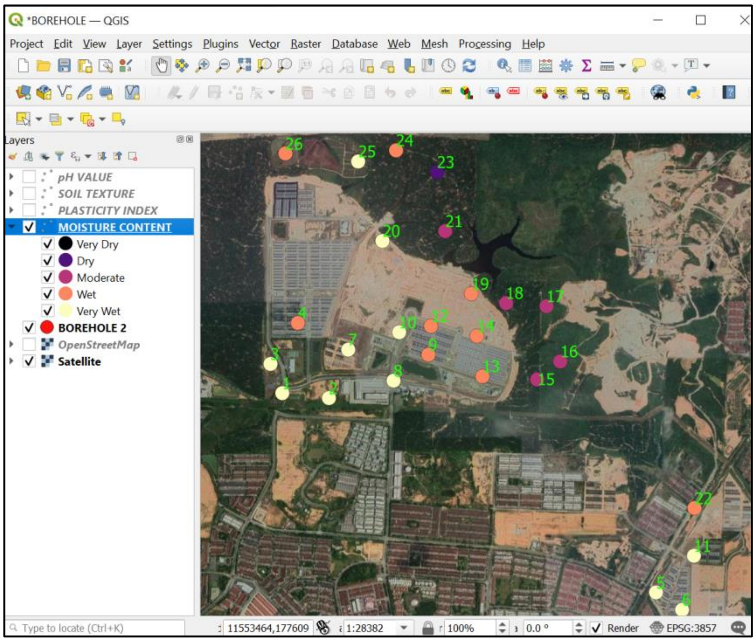

Figure 4: The soil condition of soil mapping

3.2 Plasticity Index(PI)

Table 4 shows the classification of data of plasticity index. After the attribute data for plasticity index categorized in QGIS software, the result of data mapping for borehole location was shown in Figure 5. From the analysis result, none of the data samples was classified as sand since there is no plasticity index zero and the highest data were classified in clay type of soil which recorded 12 out of 26 data of borehole samples with a PI of more than 17%. Thus the type of soil for the study area was clay.

Table 4: Classification of type of soil based on plasticity index

No BH	Plasticity Index (%)		Type of soil
	SI Report	Classification	
-	-	0	Sand
23	5.5	< 7	Silt
1	17	7 - 17	Silt Clay
4	14.5		
9	14.5		
14	16.5		
15	9.5		
16	8		
17	9		
19	11.5		
21	11		
22	15		
24	13.5		
26	13		
2	25	> 17	Clay
3	33.5		
5	30		
6	30		
7	30.5		
8	29.5		
10	38		
11	28		
12	18.5		
13	21		
18	73		
20	32.5		
25	19.5		

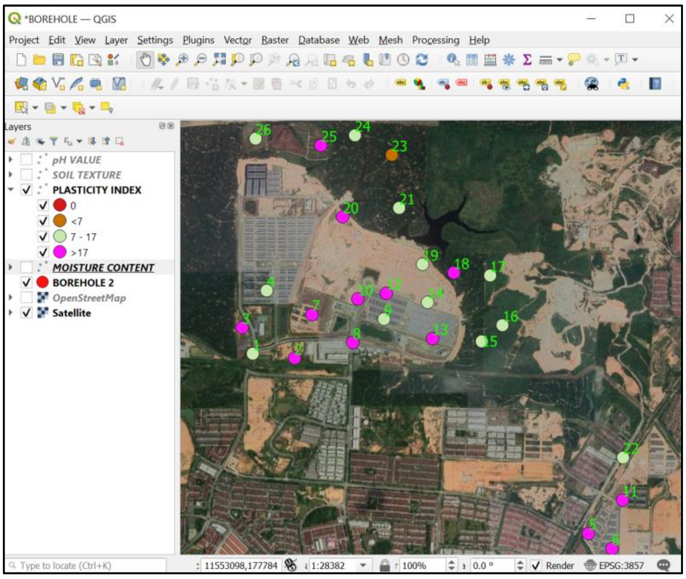

Figure 5: The data mapping for types of soils based on plasticity index

3.3 Soil Texture

In Table 5 below, five types of soil texture have been analyzed which are sandy clay loam, clay loam, sandy loam, clay, and loam. The distribution of soil texture of the study area is shown in Figure 6. Results from the analysis showed that 6 out of 26 samples were classified as sandy clay loam texture of the soil, 10 samples were clay loam, 3 samples were clay, 5 samples were sandy loam and only 2 samples classify as loam. Mostly, the soil texture for the study area were classified as clay loam which contributed about 38% of the total data samples.

Table 5: The classification of soil texture

No BH	Soil Texture Triangle			Soil Texture
	Sand	Silt	Clay	
1	50.0	20.4	29.6	Sandy Clay Loam
2	32.3	35.4	32.3	Clay Loam
3	37.4	24.2	38.4	Clay Loam
4	35.2	25.9	38.9	Clay Loam
5	35.6	24.2	40.2	Clay
6	53.0	23.2	23.8	Sandy Clay Loam
7	19.3	32.5	48.2	Clay
8	47.4	23.0	29.6	Sandy Clay Loam
9	56.1	28.9	15.0	Sandy Loam
10	57.2	28.9	21.4	Sandy Loam
11	18.4	35.2	46.4	Clay
12	42.6	29.6	27.8	Clay Loam
13	40.8	29.3	29.8	Clay Loam
14	48.0	24.0	28.1	Sandy Clay Loam
15	40.9	31.4	27.7	Clay Loam
16	60.4	28.5	19.4	Sandy Loam
17	45.3	31.8	23.0	Loam
18	33.8	33.8	17.7	Clay Loam
19	55.1	21.6	23.4	Sandy Clay Loam
20	34.4	28.0	37.6	Clay Loam
21	54.9	27.5	17.6	Sandy Loam
22	41.6	19.9	38.5	Clay Loam
23	52.1	29.1	18.8	Sandy Loam
24	44.6	25.3	30.1	Clay Loam
25	42.2	32.9	24.8	Loam
26	50.6	23.5	25.9	Sandy Clay Loam

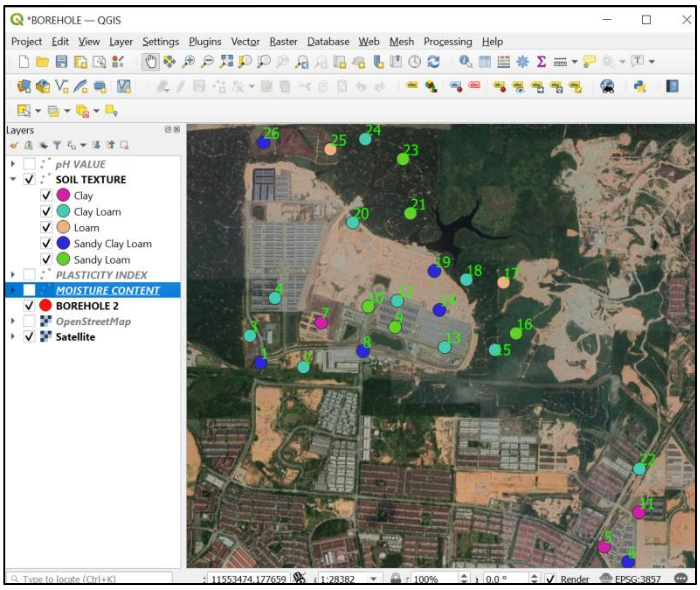

Figure 6: The distribution of soil texture at study area

3.4 pH Value

Soil can be classify based on pH value either neutral, acidic or alkaline. Table 6 represent the data analysis of type of soil according to the reading of pH value. As in the table above shows that all of 25 number of the borehole location have the reading of pH value below than 6.5 which referring to acidic soil condition. In Figure 7, there is none of the borehole location under categories of neutral (6.5 - 7.5) nor alkaline (> 7.5) conditions. Thus the study area have an acidic pH condition of soil.

Table 6: Type of soil condition according to the reading of pH value

No BH	pH Value		pH condition in soil
	SI Report	Classification	
1	4	< 6.5	Acidic
2	4.3		
3	4.4		
4	4.2		
5	4.7		
6	4.9		
7	4.4		
8	4		
9	4.3		
10	4.8		
11	4.5		
12	4.1		
13	4		
14	4		
15	4.2		
16	4.2		
17	4.1		
18	4.1		
19	4		
20	4		
21	4.2		
22	4.5		
23	4.2		
24	4.2		
25	4		

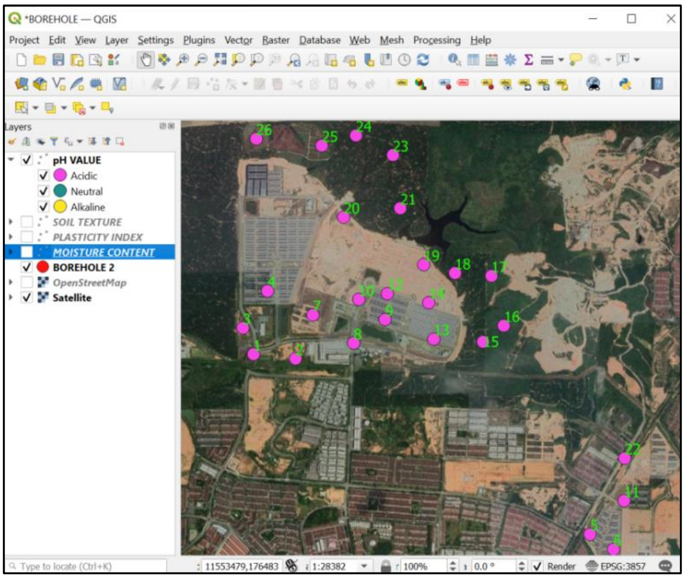

Figure 7: Classification of soil condition based on pH value

From the overall analysis results, it can be summarized that the study area has a wet acidic loam clay soil condition. Thus, we can see that classification method help the user to analyse the soil properties of study area more effectively compared to only depend on the tabulated attribute data.

Conclusion

The classification method is actually a method where we classify data with lots of information into some categories. This method can improve and develop the system of data management in GIS software especially when involved in mapping. Without this method, the process of data analysis required many processes and consumes a longer time. Classification methods can also assist in managing, organizing, and storing Site Investigation data information for the Glenmarie Heights area more systematically. This is because all the data has been organized by category and is much easier to handle as compared to dealing and managing with a bunch of data with a lot of information without the right solution. When the data is organized systematically, the data can be stored safely without the problem of loss or misplacement.

Acknowledgement

The authors would also like to thank the Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia for its support.

References

[1] Cheng, J. C., Liu, X., Li, X., Liu, R., Wang, X., & Wright, G. (2017). ISPRS International Journal of Geo-Information. A State-of-the-art Review on the Integration of Building Information Modeling (BIM) and Geographic Information System (GIS), 6(2), 53.

[2] Stafford DB(1991). Civil engineering applications of remote sensing and geographic information systems. New York: ASCE.

[3] Droste, R. L., Lukasheh, A. F., & Warith, M. A. (2001). Waste Management & Research. Review of Expert System (ES), Geographic Information System (GIS), Decision Support System (DSS), and Their Applications in Landfill Design and Management, 19(2), 177-185.

[4] Rezayee, M. (2020). International Journal of Basic Sciences and Applied Computing (IJBSAC). Hospital Site Selection in Iskandar Malaysia Using GIS-Multi Criteria Analysis, 2(10).

[5] Chua, L. S. L., Hamidah, M., Hasmadi, I. M., Lau, K. H., & Yong, W. S. Y. (2020, July).GIS Based Multi-Criteria Decision Making for Identification of Important Plant Areas: A Case Study in Johor, Malaysia. In IOP Conference Series: Earth and Environmental Science (Vol. 540, No. 1, p. 012080). IOP Publishing.

[6] Biswajeet Pradhan. (2010). Advances in Space Research. Remote Sensing and GIS based Landslide Hazard Analysis and Cross-validation Using Multivariate Logistic Regression Model on Three Test Areas in Malaysia, 45(10), 1244-1256.McCarthy, J. D., & Graniero, P. A. (2006). A GIS-based borehole data management and 3D visualization system. Computers & Geosciences, 32(10), 1699-1708.

[7] Graniero, P. A., & McCarthy, J. D. (2006). A GIS-based borehole data management and 3D visualization system. Computers & Geosciences, 32(10), 1699-1708.

[8] Phyu, T. N. (2009, March). Survey of classification techniques in data mining. In Proceedings of the International MultiConference of Engineers and Computer Scientists (Vol. 1, No. 5).

[9] Hamdan., & Omar. (2010). Slope Stability Using Remote Sensing And Geographic Information System Along Karak Highway, Malaysia.

[10] Bhalla, S. K., & Roy, S. (2017). Role of geotechnical properties of soil on civil engineering structures.Resources and Environment,7(4), 103-109.

[11] Abd Rahim, S., Gasim, M. B., Mir, S. I., Sahid, I., & Toriman, M. E. (2015). Prediction of soil and nutrient losses from the lake Chini watershed, Pahang, Malaysia.Journal of Physical Science,26(1), 53

[12] Queensland;, C. O. S. O. (2013, September 24). Soil pH. Environment, Land and Water | Queensland Government. https://www.qld.gov.au/environment/land/management/soil/soil-properties/ph-levels