Degree of Site Suitability Measurement in a GIS: The Effect of Standardisation Method

Basnet, B.B. [1] **and A.A. Apan** [1]

[1] Faculty of Engineering & Surveying and Australian Centre for Sustainable Catchments, University of Southern Queensland, Queensland
Email: basnetb@usq.edu.au

Keywords: *GIS, degree of site suitability, classification, standardisation, weighted linear combination*

EXTENDED ABSTRACT

Site suitability analysis is performed to identify suitable land units (i.e. grid cells) for a specific purpose so that management decisions can be made in a site-specific manner. However, these grid cells are rarely equally suitable in the real world. They may vary substantially in their degree (or level) of suitability. Yet, the discrimination between suitable cells is often beyond the scope of conventional site suitability analysis. Widening the scope of conventional site suitability analysis to include a degree of site suitability (DoSS) measurement is therefore crucial for managing sites in a truly site-specific manner. Conventionally, site suitability analysis involves weighted linear combination (WLC) of standardised input factors (e.g. land use, slope, distance from stream, etc.) within a Geographic Information Systems (GIS) framework. In a conventional site suitability analysis, factor attributes are standardised using discrete classification method. Yet, the effect of this standardisation method on the DoSS measurement is unknown. Therefore, the objective of this study was to quantify the effect of the discrete classification methods of input factor attribute standardisation on the DoSS measurement.

In this study, seven input factors affecting the suitability of an agricultural land for site-specific application of animal waste as fertiliser were selected, pre-processed and standardised. Discrete classification method of standardisation, which replaced continuous or discrete factor attributes with a fixed number of differentially weighted classes, was employed. Three different classification and weighting schemes were adopted. Firstly, the attributes of each input factor were classified in up to five *equal-sized classes* to examine the effect of class number on the DoSS measurement. These classes were weighted with equally incremented weights that added up to 100. Secondly, they were classified into three sets of three classes each using *equal area*, *equal interval* and *defined interval methods* of classification to examine the effect of the class size on the DoSS measurement. These classes were also weighted with equally incremented weights that added up to 100. Thirdly, the attributes of each input factor were classified into two sets of three classes each, using *equal area method of classification* to examine the effect of differential weighting on the DoSS measurement. These sets were respectively weighted with equally and unequally incremented weights that added up to 100. Finally, the standardised input factors were correspondingly combined within GIS framework to produce 10 *different composite maps* (i.e. five for varying class number, three for varying class size and two for varying class weight). The DoSS measurements of each of the composite maps was quantified using the descriptive statistical parameters such as weighted average (WA), coefficient of variation (CV), value range (VR), and coefficient of skewness (CS) to make them comparable.

The conventional discrete classification method of standardisation resulted in a series of suitability maps that varied widely depending on the class number, the class size, and the method of weighting the classes. The WA varied between 700 (CV=0 & VR=0) and 221.9 (CV=6.31 & VR=100) for class number ranging between one and five. The WA for various class sizes and weight distribution between classes were less dramatic. However, they have resulted in DoSS measurements that were clustered and skewed.

The comparisons of results from these tests have highlighted the inconsistencies in the DoSS measurement when using various discrete classification methods of input factor attribute standardisation. It was found that the variations in terms of the class number, the class size, and the weight distribution between classes were the major contributing elements towards measurement inconsistencies. Therefore, it was concluded that the usefulness of this method of standardisation is limited for obtaining a comparable and repeatable DoSS measurement unless a more robust

technique could be developed through further research.

1. INTRODUCTION

Site suitability analysis can use input datasets (or factors), formatted in a raster (grid cell) data structure, to delineate suitable sites. The number of input factors required in a particular study may vary depending on purpose, location, and circumstances surrounding the analysis. Each input factor used in a suitability analysis imposes constraints through its attributes. These constraints have effect on the magnitude and the degree (or level) of site suitability. For instance, a soil input factor may have soil types as constraints which could range from unsuitable through less suitable to highly suitable for a specific purpose. Logically, an input factor with many cells with totally “unsuitable” attribute values will reduce the extent of suitable area, whereas an input factor with greater proportion of “less suitable” attributes may only lower the degree of site suitability. Since most input factors have attributes varying widely in their level of suitability (i.e. “low”, “medium”, or “high”, or on a numeric scale like 0-100), an outcome with different degrees (or levels) of site suitability is a possibility.

The degree of site suitability (DoSS) is a parameter of interest because the suitability of a site is not usually discrete or Boolean (i.e. suitable or unsuitable) in nature. Instead, it expresses varying degrees of fuzziness or set membership (Jiang and Eastman, 2000). The DoSS measurement is therefore an approach of practical significance to make management decisions in a truly site-specific manner. However, this measurement has received very little attention in the past. This is largely because the degree of site suitability measurement is an outcome of a complex relationship between the number of input factors included in the analysis, the differential weighting of input factors, and the method of factor attributes standardisation adopted in the process (Basnet, 2002).

The spatial variation of attributes within each factor is not uncommon because most datasets come with inherent natural variability. Standardisation is therefore necessary to make it commensurable for a site suitability analysis. Standardisation is a data reduction process that simplifies the data structure (Burrough et al., 1992). In a suitability analysis, the input factors may be standardised using a Boolean logic, a continuous rescaling, or a discrete classification method. Many datasets used in a site suitability analyses are inherently categorical (e.g., land use) or recorded in a categorical format (e.g., soil type). These datasets are typically standardised using a discrete classification method. The discrete classification method of standardisation involves replacing the continuous or discrete attributes with discrete classes and weighting the classes appropriately (Burrough et al., 1992). This method of standardisation is conventional in a site suitability analysis.

In a discrete classification method, the input factors are brought to a common numeric range by classifying their attributes into discrete classes (e.g., Banai-Kashani 1989; and Jain et al., 1995) of similar or different class sizes (Chrisman, 1997). These classes are conventionally weighted or scored (Banai-Kashani, 1989; Hendrix and Buckley 1992; Siddiqui et al., 1996) for a site suitability analysis (Eastman, 2000). Both the classification and the weighting schemes determine the cell values of the input factors. The cell values of an input factor have an affect on the DoSS measurements since the conventional site suitability analysis is a process of deriving a composite map through the linear combination of input factors (Chrisman, 1997). Thus, the uniformity in the classification and the weighting schemes is crucial for the DoSS measurement. However, classification and weighting uniformity is beyond the scope of conventional site suitability analysis.

It is generally agreed that too many classes are not desirable in a suitability analysis. However, there is no optimum number recommended for an analysis. Consequently, the attributes of an input factor could be categorised into any number of classes. It is also common to have a different number of classes for each of the input factor used in an analysis. The varying number of classes, within and between the input factors, may have a substantial effect on the magnitude of DoSS measurement; but this has not been the subject of investigation in the past.

In a discrete classification system, the attributes of an input factor may be classified using various classification schemes. Most modern GIS have built-in classification schemes such as equal area, equal interval, natural break, quantile and standard deviation functions (Mitchell, 1999). Each of these schemes may produce classes of different sizes (i.e. number of pixels) by splitting factor attributes differently (Basnet, 2002). Variations in the class size (e.g. equal-sized or unequal-sized classes) could have effect on the DoSS measurement. However, the effect of class size on the DoSS measurement is not yet reported in scientific literature.

The differential weighting (or scoring) of classes is also an important aspect of a site suitability analysis. Differential weighting can be assigned in many different ways. Some of the common choices may include the following: distributing weights arbitrarily to a sum of one, 100 or 255 (Burrough, 1996 & Eastman, 2000); increasing weights with a constant interval while maintaining a sum of one or 100 (Basnet, 2002); or distributing weights using analytic hierarchy process to ensure consistency in weight distribution (e.g., Siddiqui et al., 1996 & Eastman, 1999). The selection of any one of these methods of weight distribution could have effect on the DoSS measurement. However, effect is yet to be determined.

The classification and the weighting methods available within a discrete classification system of standardisation are unlimited. Yet, the choice of a particular classification and weighting scheme may determine the class number, class size and weight distribution between classes. Variations in the class numbers, class sizes, and weight distribution between classes may have consequential effect on the DoSS measurement. In effect, it is possible that the DoSS measurement of a selected site depends on the classification and weighting scheme adopted. However, there has been no attempt in the past to compare and contrast these effects. Therefore, the objective of this study was to assess, understand and quantify the effect of class number, class size, and weight distribution between classes on the DoSS measurement.

2. METHOD

In this study, analyses were performed to identify suitable sites and to determine their degree of suitability for site-specific application of animal waste as fertiliser in the agricultural fields. The Westbrook sub-catchment in the south-east Queensland, Australia was selected as the study area. The 24 903 ha area of this sub-catchment is drained by the Westbrook Creek system. It encompassed 22 dairies, 4 feedlots, 9 piggeries, and 4 poultry farms at the time of this study. This is a relatively flat (i.e., mostly less than 10% slope) sub-catchment with some undulating hills. Most flat areas with fertile self-mulching Vertosols are used for extensive farming.

Seven input datasets that are influential on the social, economical, environmental, and/or agricultural suitability of a site for animal waste application were selected based on the literature. These datasets were pre-processed within Arc/Info GIS software to create raster grids of 10m × 10m cell resolution prior to the analysis. Unsuitable attributes for animal waste application were excluded by assigning no-data value from each of the seven input factors. The exclusionary criteria adopted by Basnet (2002) were employed to identify the unsuitable attributes. Potentially suitable area for animal waste application was calculated for each input factor from the remaining attributes (Table 1).

Table 1. Potentially suitable area within each input factor for animal waste application

Selected input factors	Unsuitable		Suitable	
	attributes*	ha	ha	%
Sub-catchment	None	0	24903	100
Land cover	Non-crop or pasture	14183	10720	43.0
Proximity to town	Within 250 m radius	7516	17387	69.8
Proximity to stream	Within 100 m distance	7071	17832	71.6
Soils	Shallow or stony soils	5895	19008	76.3
Slopes	With >10 % slope	2453	22450	90.1
Proximity to road	Within 25 m distance	1787	23116	92.8
Proximity to IAI	Within 100 m radius	120	24783	99.5

*IAI = Intensive Animal Industry, * from Basnet (2002)*

The suitable attributes of all but the land cover input factor were standardised using discrete classification method. The land cover input factor contained only one suitable attribute (i.e., crop/pasture) with no further detail on crop or pasture type. Thus, it was retained as a single-classed input factor with a class weight of 100 throughout the test. The remaining six input factors were firstly classified in up to five classes each of approximately equal class sizes using equal area method of classification (Table 2a).

Table 2(a). Area in hectares under each class using equal area method of classification

Class no.	Proximity to		Soil	Slope	Proximity to	
	town	stream			road	IAI
Single class						
I	17387	17832	19008	22450	23116	24783
Two classes						
I	8695	8967	9502	11279	11559	12401
II	8692	8865	9506	11171	11557	12382
Three classes						
I	5800	5945	6389	7637	7741	8272
II	5796	5947	6150	7343	7675	8254
III	5791	5940	6469	7470	7700	8257
Four classes						
I	4373	4564	4650	5613	5919	6209
II	4322	4403	5331	5666	5640	6193
III	4353	4431	4728	5627	5785	6187
IV	4339	4434	4299	5544	5772	6194
Five classes						
I	3500	3608	3432	4568	4644	4964
II	3459	3532	5331	4568	4640	4955
III	3482	3579	4728	4378	4613	4958
IV	3475	3549	2775	4511	4596	4951
V	3471	3564	2742	4425	4623	4955

These classes were weighted using an arbitrarily selected increment of five (Table 2b) to distribute weights evenly (i.e., at equal interval) between classes. The highest weight was assigned to the most suitable class and the sum of all the class weights was maintained to 100.

Table 2(b). Weight distribution between classes

No. of factor attribute classes	Weight distribution to the class					Sum of weights	Increment
	I	II	III	IV	V		
1	100.0					100	-
2	52.5	47.5				100	5
3	38.3	33.3	28.3			100	5
4	32.5	27.5	22.5	17.5		100	5
5	30.0	25.0	20.0	15.0	10.0	100	5

Example: Increment = (38.3 – 33.3) = (33.3 - 28.3) = 5.0

For the second test, the attributes of each of the six input factors were classified into three classes each using three different methods of classification (i.e., equal area, equal interval and defined interval). The equal area method classified factor attributes into three classes of approximately equal sizes by finding appropriate break points in the data (ESRI, 1996). The equal interval method divided attributes into equal sized sub-ranges (ESRI, 1996). The defined interval method employed information from the literature to classify the attributes. For example; fields with 6-10, 3-6, and 0-3 % slopes are considered as good, better, and best for waste application (NSW A&F, 1989) and therefore classified into class III, II, and I, respectively. This has resulted in three sets of classified data with varying class sizes (Table 3). These classes were weighted as before.

Table 3. Area within each class as determined by three classification methods

Factor	Class sizes (ha) using three classification schemes								
	Equal Area			Equal Interval			Defined interval		
	I	II	III	I	II	III	I	II	III
Town	5800	5796	5791	13601	3409	377	5936	10053	1398
Stream	5945	5947	5940	14380	3197	255	7528	10239	65
Soil	6389	6150	6469	12279	4854	1875	3636	12835	2537
Slope	7637	7343	7470	11504	8453	2493	5294	11705	5451
Road	7741	7675	7700	19838	2910	368	4455	18329	332
IAI	8272	8254	8257	18098	5260	1425	2491	18181	4111

IAI: Intensive animal industries

Finally, the attributes of each of the six input factors were classified into two sets of three classes each, using the equal area method of classification, to examine the effect of differential weighting on the DoSS measurement. The first set was weighted as above with equal incremented weight. The second set was assigned unequally incremented weight (Table 4) using the analytic hierarchy process (AHP).

Table 4. Weights derived using AHP method

Class	Proximity to		Soil	Slope	Proximity to	
	Town	Stream			Road	IAI
I	07.02	06.39	62.67	66.08	59.36	62.67
II	37.08	13.83	27.97	20.81	24.93	27.97
III	55.90	79.78	09.36	13.11	15.71	09.36
Sum	100.0	100.0	100.0	100.0	100.0	100.0

Unequally incremented weights were derived separately for each input factor using the AHP developed by Saaty (1980). The class weights added up to 100 for both equally and unequally incremented weights.

$$S_i = \sum_{j=1}^{n} \left(f_{ji}.suit \times w_j\right) \quad [Eq.\ 2]$$

The standardised input factors were combined spatially using the weighted linear combination (WLC) model (Equation 2) within Arc/Info GRID.

Where,

S_i = suitability value at i_{th} cell locations

f_{ji}.suit = grid dot notation for factor attribute classes for j_{th} factor with class weights at i_{th} cell locations, and

w_j = respective weight for factor f_j (all factors treated equally in this case)

Separate tests were conducted to evaluate the effects of the number of classes, class sizes, and the method of weight distribution between classes on the DoSS measurement. All seven input factors were weighted equally in this analysis.

The WLC produced a suitability map with composite values through cell-wise summation of corresponding cell values from each input factor. Cells identified with no-data in any one of the input factor made this area unsuitable for animal waste application. Thus, higher and lower cell value in a composite map indicated higher and lower DoSS, respectively, while no-data indicated not suitable. Descriptive statistical parameters such as weighted average (WA), coefficient of variation (CV), values range (VR) and coefficient of skewness (CS) was calculated from the composite map to assess the DoSS measurement. The WA quantified the central tendencies of the cell values, while CV and VR measured their dispersions. The CS determined their degree of symmetry.

3. RESULTS

In this analysis, 10 different composite maps (i.e. five for varying class number, three for varying class size, and two for varying class weight) were produced. A sample suitability map for three classes with equal class size and weight increment is given in Figure 1. The DoSS values were grouped into low, medium and high suitability in this case.

The variation in class number is found to have substantial effect on the magnitude of the expected DoSS values. These effects were evident from a comparison of the five composite maps that were produced using factors classified in one to five classes (Table 5).

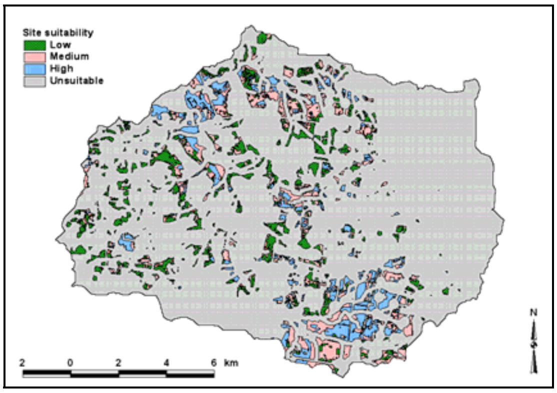

Figure 1. A sample map showing low, medium and high degree of site suitability.

The WA of the suitability value ranged between 700 (with CV = 0 and VR = 0) and 221.9 (with CV = 6.3 and VR = 100) depending on the number of classes. Clearly, the increase in class number has decreased the weighted average while increasing the coefficient of variation and the value range (Table 5).

Table 5. Effects of the number of factor attribute classes on the DoSS measurement

No. of attribute class	Weighted average (WA)	Weighted standard deviation	Coeff. of variation (CV) %	Value range (VR)
1	700.0	0.00	0.00	0
2	399.5	5.31	1.33	30
3	298.8	8.34	2.79	60
4	250.3	11.43	4.57	80
5	221.9	13.96	6.29	100

Seven input factors used in the analysis

The effect of class size on the WA, CV and VR measurements were not so dramatic but there was an indication of clustering of the suitability values in case of unbalanced class sizes. Nevertheless, the skewness measurements were noticeably different between various class sizes (Table 6). In this instance, the unbalanced class sizes have resulted in upper bound of the suitability values as indicated by the negative skewness coefficient (or left skewing) of the suitability values (Table 6).

Table 6. Effects of the class size distribution on the DoSS measurement

Class size	Weighted average (WA)	Weighted standard deviation	Coeff. of variation (CV) %	Value Range (VR)	Coeff. of skewness (CS)
Equal area	298.78	8.34	2.79	60	0.15
Defined interval	298.23	5.70	1.91	45	-0.04
Equal interval	306.85	4.89	1.60	35	-0.31

The method of weight distribution between classes is found to have remarkable effect on the DoSS measurement. There has been a sizeable increase in the CV and VR measurements due to uneven weighting (Table 7). It has also resulted in lower bounded suitability values as indicated by the positive coefficient of skewness measurement.

Table 7. Effects of the method of class weight distribution on the DoSS measurement

Method of weight distribution	Weighted average (WA)	Weighted standard deviation	Coeff. of variation (CV) %	Value Range (VR)	Coeff. of skewness (CS)
Equal increment *	298.8	8.34	2.79	60	0.15
Uneven weighting #	292.1	54.72	18.73	337	+ 0.27

* As per Table 2b for 3 classes; # As per Table 4

4. DISCUSSION

Discrete classification methods of standardisation are being used conventionally as a process of data reduction to make complex data sets understandable (Burrough et al., 1992). However, the options available within this method of standardisation are unlimited. The input factors may be standardised into many classes of different sizes that could be weighted differently. This study has revealed that the choice of a particular class number, class size, or weight distribution between classes has effect of various extents on the DoSS measurement.

The composite maps of the input factors classified in up to five classes (Table 2a) have resulted in the DoSS measurements that are substantially different to each other. The WA of the suitability value ranged between 700.0 and 221.9 depending on the number of classes (Table 5). The CV and VR also varied accordingly. A clear trend has emerged in the sense that the weighted average decreased and the coefficient of variation and the value range increased with an increase in the class number. The decrease in the WA and the increase in the CV and VR are mainly due to the split of suitable area into smaller classes and the fragmentation of weights assigned to those classes during the classification. These occurrences can be seen in an individual input factor (Table 2a and 2b) where an increase in class number is associated with corresponding decrease in class size and weight distribution between classes. Increase in the class number and the decrease in class weight have resulted in a composite map with lower but wide ranging cell suitability values. Therefore, the increase in class number is associated with the decrease in WA and increase in CV and VR. In this instance, the increase in the class number may appear to be beneficial in terms of differentiating suitable areas into various DoSS by recognising subtle differences. However, the inconsistency in the DoSS measurement becomes an issue when there is no limit to the number of classes to be used in an analysis. Therefore, the effect of class number on the DoSS measurement remains inexplicable.

The composite maps produced from balanced and unbalanced class sizes have also shown differences in the DoSS measurements. The unbalanced class sizes, created by equal interval and defined interval methods of classification, have resulted in lower CV and VR measurements. Lower values of coefficient of variation and value range denote that these measurements are close to the WA or less dispersed or more clustered. These outcomes are expected because the larger classes have the dominating effects when the class sizes are unbalanced. The unbalanced class sizes have also been responsible in skewing the suitability values towards the left (Table 6). It is left skewed when the higher suitability values are on the right hand side of the mean (i.e. upper bound). This situation occurs when a larger class has small suitability values. This is again a function of disproportion in class sizes. In a nutshell, the unbalanced class sizes have varying effect in the DoSS measurement. Yet, the standardisation of input factors using discrete classification method does not necessarily produce balanced class sizes. Thus, the class size effect on the DoSS measurement also remains unresolved.

The effect of the method of weight distribution between classes on the DoSS measurement has been quite remarkable. The CV and VR have increased substantially and the suitability values have become lower bounded as a result of uneven weighting (Table 7). This outcome is not unusual given the role of weighting in differentiating suitable sites. In a site suitability analysis, differential weightings are assigned to the classes to be able to distinguish them in terms of their suitability. Differential weightings can be assigned in many ways to the suitability classes. In this instance, they were assigned evenly by maintaining equal increment between classes and unevenly by weighting some classes heavier than others. The distinction between classes (or class effect) may remain consistent when the weights are assigned with an equal increment. However, the class effect may become exaggerated or understated depending on class weighting when the weights are not assigned evenly. These effects may have resulted in the increase of CV and VR measurements in this case. It is also the case that the heavily weighted classes show greater influence to the overall outcome. This is probably the reason for right skewing of the suitability values. Thus, the method of weighting determines the DoSS measurement. Yet, in practice, the class weights can be assigned in many different ways.

5. CONCLUSION

The effect of discrete classification methods of input factor attribute standardisation on the degree of site suitability measurements was examined. It was revealed that the class number, class size, and weight distribution between classes have effects of various extents on the degree of site suitability measurements. The measurements were found to be dependent on the choice of a classification option. It was also highlighted that there is no single approach in achieving optimum class number, class size, and weight distribution between classes. Therefore, this study concludes that the usefulness of the discrete classification method of standardisation is limited in obtaining a comparable and repeatable degree of site suitability measurements.

6. REFERENCES

Banai-Kashani, R. (1989), A new method for site suitability analysis: The analytic hierarchy process, *Environmental Management*, 13(6), 685-693.

Basnet, B.B. (2002), Geographic information system based manure application planning, Ph.D. dissertation, University of Southern Queensland, Australia.

Burrough, P.A. (1996), Geographic objects with intermediate boundaries, Chapter 1: Natural objects with indeterminate boundaries, Talylor and Francis, 3-28 pp., UK

Burrough, P.A., R.A. MacMillan, and W. van Deursen (1992), Fuzzy classification methods for determining land suitability from soil profile observations and topography. *Journal of Soil Science*, 43(2), 193-210.

Chrisman, N. (1997), Exploring geographic information systems, John Wiley & Sons, 91-103 pp., Canada.

Eastman, J.R. (1999), Idrisi32: Guide to GIS and image processing, Clarks Labs, Clark University, Worcester.

Eastman, J.R. (2000), Multi-criteria evaluation and GIS, In Geographical Information Systems – Chapter 35: Principles and technical issues, edited by P.A. Longley, M.F. Goodchild, D.J. Maguire, and D.W. Rhind, John Wiley & Sons Inc.

ESRI, (1996), Using ArcView GIS: The geographic information system for everyone, ESRI, Redlands, California.

Hendrix, W.G. and D.J.A. Buckley (1992), Use of a geographic information system for selection of sites for land application of sewage waste, *Journal of Soil and Water Conservation*, May-June, 271-275.

Jain, D.K., U.S. Tim, and R.W. Jolly (1995), A spatial decision support system for livestock production planning and environmental management, *Applied Engineering in Agriculture*, 11(5), 711-719.

Jiang, H. and J.R. Eastman (2000), Application of fuzzy measures in multi-criteria evaluation in GIS, *International Journal of Geographical Information Science*, 14(2), 173-184.

Mitchell, A. (1999), The ESRI guide to GIS analysis, ESRI Press, Redlands, California.

NSW Ag & Fish. (1989), Guidelines for the use of sewage sludge on agricultural land, New South Wales Agriculture and Fisheries, NSW, Australia.

Saaty, T.L. (1980), The analytic hierarchy process, McGraw-Hill, Inc., New York.

Siddiqui, M.Z., J.W. Everett and B.E. Vieux (1996), Landfill siting using geographic information systems: a demonstration, *Journal of Environmental Engineering*, 122(6), 515-523.