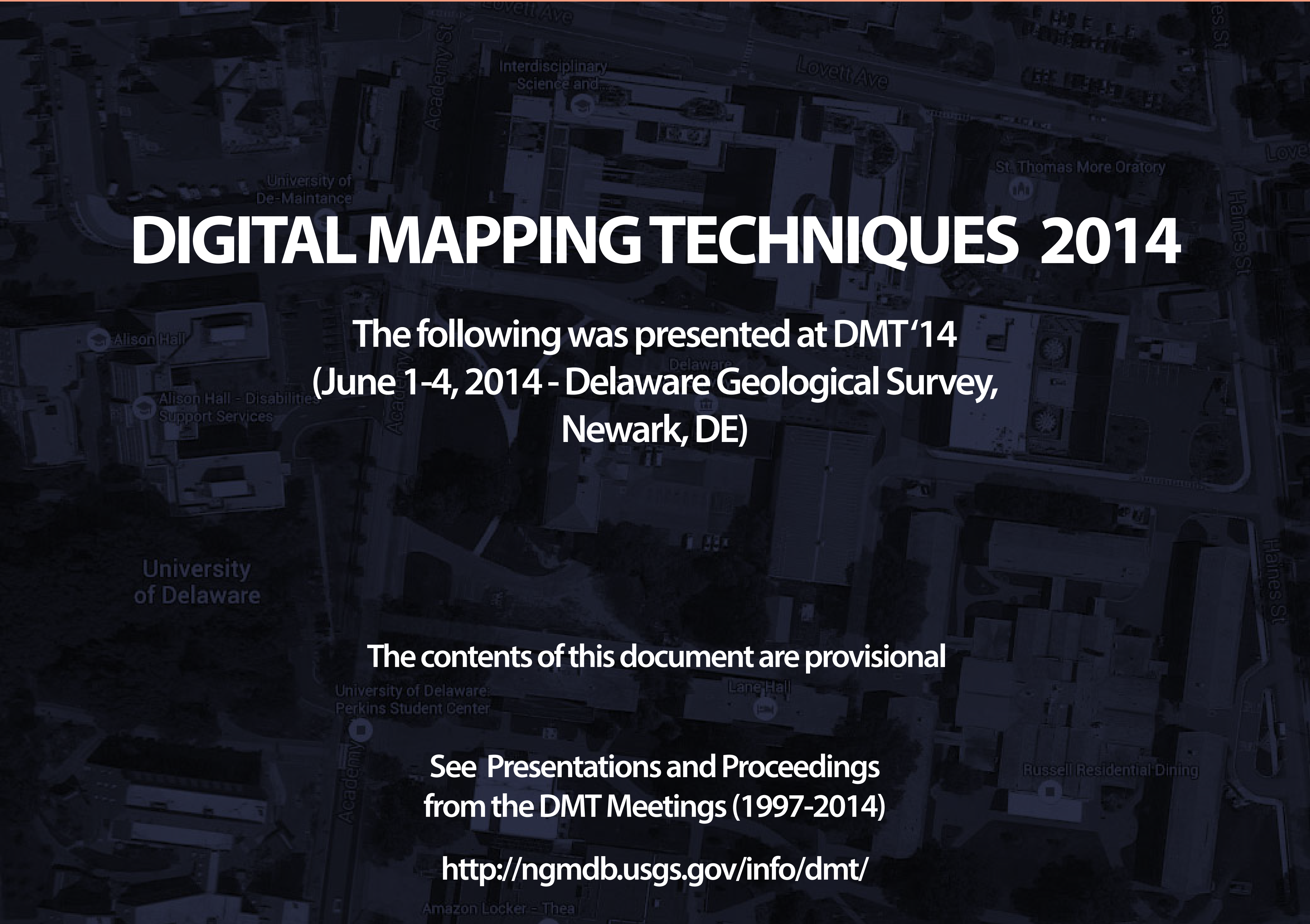
DMT 2014



DIGITAL MAPPING TECHNIQUES 2014

The following was presented at DMT'14
(June 1-4, 2014 - Delaware Geological Survey,
Newark, DE)

The contents of this document are provisional

See Presentations and Proceedings
from the DMT Meetings (1997-2014)

http://ngmdb.usgs.gov/info/dmt/

Deriving Traverse Paths for Scientific Fieldwork with Multi-criteria valuation and Path Modeling in a Geographic Information System

By [1]Ryan R. Reeves, P. Kyle House[2], Jordan T. Hastings[3]

[1] Independent consultant (formerly student at University of Southern California)
email: ryan-reeves@hotmail.com

[2] Research Geologist
U.S. Geological Survey
2255 N. Gemini Drive
Flagstaff, AZ 86001
Telephone: (928) 556-7179
email: khouse@usgs.gov

[3] Assistant Professor
University of Southern California
P.O. Box 13166
Reno, NV 89507
Telephone: (775) 771-3166
email: jordan.hastings@gmail.com

INTRODUCTION

Field research is a necessary component of many realms of geoscientific practice since it provides the primary data crucial for understanding geologic, phenomena. Unlike work in an office or laboratory, fieldwork has additional cost related to travel, lodging, and per diem expenses. Field scientists must therefore ensure they make efficient and effective field navigational decisions that result in expedient execution of field campaign objectives.

Technologies and analytical approaches such as decision analysis, path modeling, and geographic information systems (GIS) offer assistance to navigational decision-making while in the field as do analytical techniques such as weighted linear combination and analytical hierarchy process. These tools are often underutilized, however. This paper, based on a poster presented at the 2014 DMT Workshop and master's thesis project (Reeves, 2015), describes a methodology by which these technologies and analytical procedures may assist field scientists with navigational decision-making. Specifically, the paper documents development of a model that uses a spatial multi-criteria decision evaluation to derive favorability values. These values are then used to determine the placement of traverse paths that are suggested routes to be followed by field researchers. The paper includes an introduction to the methodology, a description of its underlying concepts , and a brief summary of its benefits and limitations.

Fieldwork Planning and Technology

Field researchers from many scientific disciplines typically use archival data such as government records, maps, and remotely sensed imagery to inform fieldwork planning. Geologists, in particular, use archival data to conduct preliminary inspections of study areas prior to visitation (Compton, 1985; Coe, 2012). Published maps and remotely sensed data (e.g., aerial photography, LiDAR, satellite imagery, etc.) are primary data sources. Preparatory work may involve analysis and digital processing of aerial and satellite imagery to characterize and distinguish varying rock types and landforms (e.g. Mars, 2013); study of existing geologic maps and reports; and communication with other scientists familiar with the area in question. Scientists often rely on these types of data prior to (and in conjunction with) their fieldwork to minimize superfluous efforts and reduce the size of the sampling area necessary to describe a study area.

The tools that may be used to plan a traverse and augment it while in the field are quickly advancing. House et al. (2013) describe how technological advancements regarding geographic information systems, light detection and ranging, virtual globes, mobile hardware and software, and geocoded field data are changing the practice of geologic mapping. Nevertheless, fieldwork is still an essential and costly activity, so the incorporation of such advanced technology to improve the process of traverse planning is important.

Multi-criteria Evaluation

Multi-criteria evaluation (MCE) is defined as decision support tool that combines multiple map layers, each weighted to indicate relative importance, to produce an output map contain favorability values. It comprises a set of analytical procedures that may be thought of as a sub-discipline of multi-criteria decision analysis or multi-criteria decision making (Carver, 2008). Many of the methods applied in MCE originated from the field of operations research in the 1960s and 1970s (Carver, 2008). They arose in response to critiques of early techniques in decision-making and site location analysis (Carver, 1991). The MCE approach combines multiple datasets representing various criteria and/or objectives, assigns them with a weight indicating their relative importance, and produces a multi-valued output (e.g. a raster data model with a grid of georeferenced cells with different values) indicating the degree to which an objective(s) has been met.

The term ‘criterion’ is often used generically to refer to concepts of both criterion attributes and objectives. It is used here to refer to attributes of entities or phenomena that may be used to measure the fulfillment of a certain objective, or various objectives. This process may be done for geographic space by designing such an evaluation around spatial data. A GIS is often used for this due to its ability to store, display, and analyze these data relevant to many decision problems (Carver, 1991).

Geographic information systems alone, while advantageous for working with various types of spatial data in a wide variety of applications, were not originally designed to handle analyses involving a complex value structure consisting of conflicting objectives and varying priorities (Malczewski, 1999). In 1991, Carver described a GIS as a data management framework for the spatial data used in a MCE. He noted that a MCE provides a GIS with the ability to handle conflicting objectives that encompass multiple criteria and multiple decision makers. Now, two decades later, most geographic information systems do provide a means by which at least some MCE techniques may be implemented directly within the GIS framework. By incorporating the technologies associated with MCE and GIS, decision makers are able to analyze spatial problems containing multiple criteria and objectives.

Carver (2008) outlines the main steps involved in a multi-criteria evaluation as: Problem definition, criterion selection, standardization of criterion scores, allocation of weights, and implementation of an aggregation algorithm. Additional steps such as a sensitivity analysis and making decisions with the processed information may also be included. Problem definition involves identifying the difference between existing and desired states of a system (Malczewski, 1999). Once the problem has been identified, it can be determined how the achievement of a certain objective(s) may bring the system closer to the desired state. After the attribute values of multiple criteria have been standardized, weighted, and aggregated, they may then be used to determine the degree to which an objective(s) has been met. A sensitivity analysis is performed to discover error or uncertainty that may be contained within the derived values. Once confident that the values attained are of sufficient quality, they may be used to make decisions (see Figure 1).

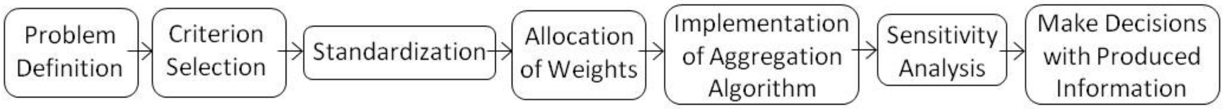


Figure 1. Flowchart of typical multicriteria evaluation procedures.

Path Modeling

Paths in the field may be modeled to suggest the best way to move between locations (Mitchell, 2012). The two common types of paths that are generated include network paths and overland paths. Network paths follow a predetermined network (e.g. transportation network) and overland paths follow a model-determined path between two points. The former is performed in a vector GIS environment while the later is performed in a raster GIS environment. The placements of paths are often determined by associated costs. These costs may be expressed as money, time, distance, etc. Network costs are associated with edges, intersections, and turns, while overland costs are associated with raster cells values (Mitchell, 2012).

A cost-path analysis can be performed within ArcGIS using a raster to determine the cost values associated with traveling across particular cells. Using an evaluation process such as MCE is an example of how such a cost surface, or weighted surface, may be created. The accumulative cost is calculated on a cell by cell basis by starting at the origin cell and traveling towards a destination cell, sampling all of its adjacent cells, and recording the value associated with each edge. Once the cost distance rasters have been generated, they may be used as inputs to derive a path.

METHODOLOGY

The methodology described here derives traverse paths for fieldwork in a nine-step process (Table 1). Most of these steps follow the MCE workflow described previously. Additional steps cover the processes of data assembly and construction of evaluation criteria layers. The final step involves the derivation of the origin and destination points and the traverse paths that cross them. The location of the origin and destination points and the traverse paths are determined by the values contained within the final weighted surface derived by the MCE. These steps, along with a summary of the work they require, are shown below in Table 1. There are many options available with regard to the specific techniques that may be employed at each step. Only a generic framework is presented here. Those adopting this type of approach must determine which techniques best suit their needs.

Table 1 - Methodology Steps and Required Work

Steps	Required Work
1. Identify objectives and criteria — Problem definition and criteria selection steps of MCE	Literature examination, analytical study, and/or attainment of expert opinions
2. Assemble relevant data	Acquire data that may be used to create criteria layers for measure of objective(s) fulfillment
3. Sketch and derive features of interest relevant to meeting objective(s)	Delineate and/or create criteria layers and ensure they share a common coordinate system and extent
4. Apply necessary manipulations or analysis to derived features of interest	For example, apply distance calculations
5. Standardize the non-standardized criteria layers — Standardization step of MCE.	Transform non-standardized criteria layer values to a common scale
6. Establish field campaign priorities — Weighting step of MCE.	Derive and assign weights to each criteria and objective
7. Produce weighted surface layers of study area — Aggregation step of MCE	Perform map overlay
8. Review results of MCE — Sensitivity analysis step of MCE	Perform sensitivity analyses and/or analysis of results
9. Define traverse paths on basis of time availability	Derive origin and destination points and traverse paths

Identify Objectives and Criteria

This step of the methodology relates to the problem definition and criteria selection. As field research is conducted to attain data that is unavailable via remote means, the problem facing scientists preparing to go in the field will often be determining what data should be acquired and how. This problem should be divided into multiple objectives. These objectives may be, for example, to attain a scientific return and to avoid obstacles that impede travel across the study area. A group consensus, or individual decision, establishing the overall decision problem and separating this problem into applicable objectives is required to proceed to the subsequent steps.

This step also involves deciding which criteria will be used to measure the fulfillment of the determined objectives. Research may be needed to determine which criteria affect a given objective in order to establish a scientific foundation for the remaining steps. Analytical studies or an opinion survey are additional options that may be used to make this decision regarding appropriate criteria. Once the set of criteria is determined, they should be separated into factors or constraints. If a criterion is considered a constraint, the constraining attribute values should be noted. For example, a criterion layer containing slope values may be considered a constraint where all values are greater than some threshold, for example, greater than fifty degrees. It also should be determined whether a given criterion's attribute values will have a favorable or unfavorable influence on meeting the objective to which it is applicable. This will assist in determining an appropriate algorithm to use during standardization. While a criterion may contain attribute values that are not favorable with regards to meeting a particular objective, they may not necessarily act as a constraint (i.e., a hard limitation).

Assemble Relevant Data

Once it has been determined which criteria will be used to measure the degree to which a particular objective(s) is being met, data are sought that may be used to represent these criteria. Each data set used must share a common coordinate system and have positional accuracy sufficient for the research at hand.

Sketch and Derive Features of Interest Relevant to Meeting Objective(s)

Data that provide information on criteria influencing the degree to which a given objective(s) is being met will often not be suitable for the remaining steps of this methodology. In raw form, the data may not provide a suitable representation of the criteria or objectives being evaluated. Such data should be brought into a suitable form through various manipulations or analyses. Such techniques may include the derivation of slope or visibility from a DEM, or buffering of features. It may also include the manual delineation of various features of interest based upon image or map interpretations.

Apply Necessary Manipulations or Analysis to Derived Features of Interest

The delineated features of interest may be further analyzed so that they contain information more directly related to measuring the degree to which a particular objective(s) has been met. An example of this is determining the distance from these delineated features of interest to all other locations within the study area. This would be relevant to situations where a scientist's proximity to various features of interest relate to the ability to attain a scientific return from the features.

Standardization

This step relates to the standardization step of a MCE. Criteria layers that do not share a common scale must be converted to a common scale before they may be aggregated. This is done through the process of standardization. Various standardization techniques are shown in Table 2. During this process, each layer is transformed into a common scale containing floating point values ranging from zero to one. The aim of this methodology is to create traverse paths derived from a cost surface, based on cost-criteria. It is fitting therefore, to represent favorable characteristics with a low value and unfavorable characteristics with a high value. Care should be taken to ensure these values are not reversed erroneously in the subsequent steps.

Table 2 - Summary of Standardization Techniques: Summary of four common standardization techniques (after Malczewski, 1999)

Standardization Technique	Description
Linear scale transformation	Divides the raw attribute values within a given criterion layer by the layer's maximum value for this same attribute.
Value/Utility function approach	Uses input from decision makers to assist in defining a function that identifies the relationship between a non-standardized criterion layer and a standardized criterion layer.
Probability	Uses probability theory to determine the likelihood of a given outcome, which is then used to determine standardized values.
Fuzzy set membership	Process of assigning standardized values based on a membership function.

Allocation of Weights

This step relates to the weighting step of a MCE. It must be determined which weight assessment technique is appropriate for the given decision problem. Various weighting techniques are shown in Table 3.

Table 3 - Summary of Weighting Techniques: Summary of four common weighting techniques (after Malczewski, 1999)

Weighting Technique	Description
Ranking	The decision maker uses their preference to place the set of chosen criteria in order based on their relative importance. Then, numerical weights may be derived by inserting these ordinal values into a mathematical formula.
Rating	The decision maker estimates criteria weights relative to a predetermined scale. Each criterion is then allocated a number of points across a predetermined scale with a set range, where the collective points allocated equate to a fixed number.
Analytical Hierarchy Process (AHP)	Use of pair-wise comparison to create a matrix, which is then subject to calculations to derive the right eigenvector of the largest eigenvalue of this matrix. The derived eigenvectors become the criterion and objective weights.
Trade-off analysis	Assess trade-offs between pairs of alternatives.

Implementation of Aggregation Algorithm

After all the relevant criteria factoring into the navigational decision making process had been standardized and once weights for these layers had been determined, these layers may then be aggregated. Various aggregation techniques are shown in Table 4. If more than one objective is necessary to assess the decision problem, then multiple weighted surfaces will be created. These surfaces may be combined to create one final weighted surface to be used during the subsequent steps.

Table 4 - Summary of Aggregation Techniques: Summary of three common aggregation techniques (after Malczewski, 1999)

Aggregation Technique	Description
Weighted linear combination	Takes predetermined weights, multiplies them by normalized values given to criterion attributes, and then sums the products over all criteria.
Ideal Point methods	Derives values that represent amount of separation from an ideal value.
Concordance methods	Based on a pair-wise comparison of alternatives and a mathematical function applied to a concordance/discordance matrix derived from these comparisons. Differs from AHP in that criteria may only be compared as having preference over another criteria, but without indication of how much.

Review Results of Multicriteria Evaluation

This step relates to the sensitivity analysis and analysis of outcome steps of a MCE. If appropriate, an error propagation analysis or the construction of an error matrix may be performed. Otherwise, one should inspect the values contained within the weighted surface to ensure its values contain the desired meaning. Once confident that these values are reliable, one may proceed to the subsequent steps.

Define Traverse Paths on Basis of Time Availability

This step involves deriving the set of points that the traverse path must visit during the field campaign and the path to be followed between them. Note that this methodology does not describe how to determine visitation sites, but produces suggested traverse paths that position field scientists within close proximity to features they have deemed to be of interest. The description below explains how this step may be performed within ArcGIS. In order to derive a traverse path using this software, users must determine origin and destination points and a weighted surface. These are used as inputs to the ArcGIS Cost Distance and Cost Path tools. The tools are run once for each segment of the traverse path. All origin points also act as destination points and will thus be referred to hereafter as destination points.

Since the lowest values in the weighted surface (i.e. the cost layer) indicate favorability with regard to meeting a particular objective, this layer is used to determine the destination points. If time is limited for a particular field campaign, a traverse path may be prioritized by delineating only the most favorable locations (i.e. those with the lowest values on the weighted surface). For example, only locations with the top 5% most favorable values of the weighted layer may be used when a short time duration is available for fieldwork. Once it is determined which percent to use, the weighted layer is reclassified so that these most favorable values are represented as some value (e.g., 1) while all greater values (i.e., less favorable) are represented as NoData. This reclassified layer is then converted to a polygon feature class. The centroids of these polygons are then derived and serve as the destination points. Derived points that are in close proximity should be manually deleted to avoid excessive calculations that will not greatly alter the location of the derived traverse path.

In order to use these derived points as individual destination points in the multiple iterations of the Cost Distance and Cost Path tools, all points in the feature class created above must be separated into distinct feature classes. The sequence in which these points are used in the iterations of the Cost Distance and Cost Path tools will determine the connectivity of the destination points of the derived traverse path. Thus, these points should be manually ordered so the resulting path will contain a logical sequence. For example, in Figure 2, the destination points have been arranged so that no segments of the traverse path cross. Rather, the traverse path makes a loop around this portion of the study area. Figure 2 also illustrates that each point acts as both an origin and destination point. These locations can either contain one point feature class that acts as both and origin and destination point or contain two point feature classes with one representing an origin and the other, a destination.

Once the destination points have been derived they can be used in combination with the weighted surface to derive a traverse path using the ArcGIS tools mentioned above. Finally, the ArcGIS Raster to Polyline tool is used to convert the raster output of the Cost Path tool to a polyline feature class. This will reduce the size of the files representing the traverse path and will convert it to a format that may be easier for use in the field.

Figure 3 shows the steps (in ArcGIS ModelBuilder) used to derive one segment of the traverse path. In order to make a traverse path with multiple segments this sequence of steps must be repeated multiple times. Conducting this work within ModelBuilder expedites this process.

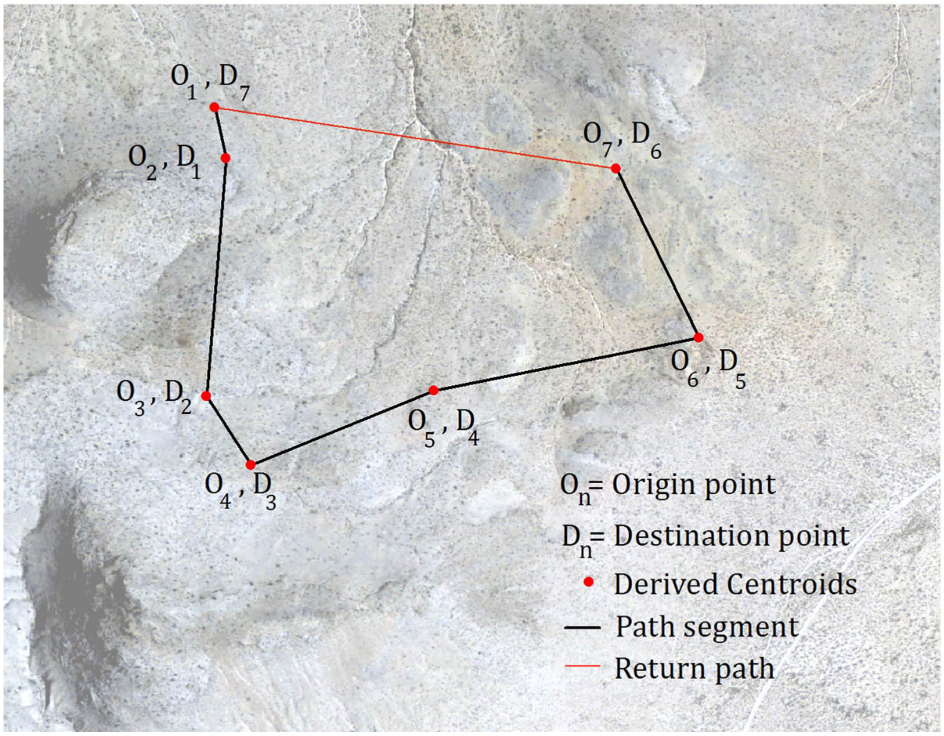


Figure 2. Example of how the origin and destination points factor in to the derivation of the traverse paths. Note that the segments between points are drawn here as simple straight lines, not as final derived traverse paths.

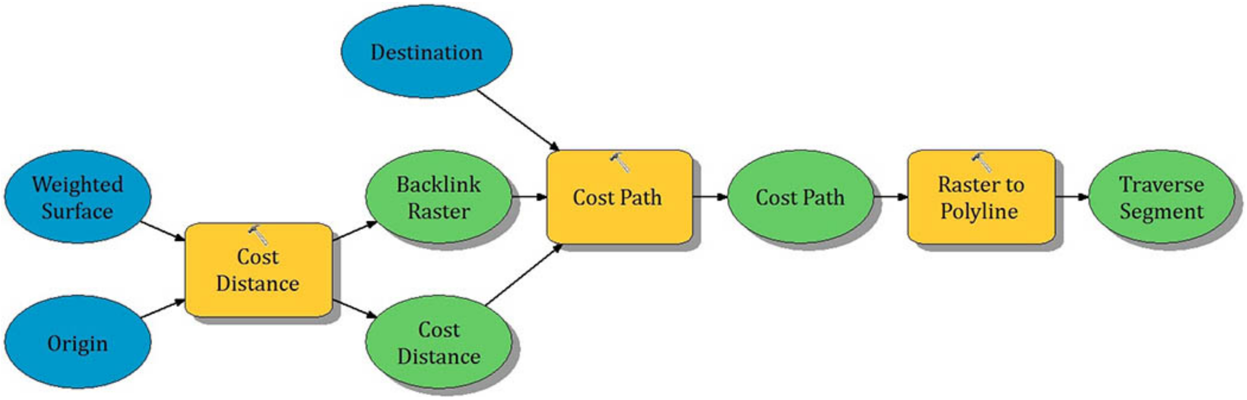


Figure 3. ArcGIS ModelBuilder steps used to derive traverse path segment.

CONCLUSIONS

The methodology described in this paper allows field scientists to obtain benefits of the MCE and path modeling processes. It also utilizes a GIS's ability to manage, manipulate, and analyze spatial data. Of these three technologies, a GIS is the only tool that has been extensively adopted by the field research community. While there are many benefits to the methodology described here, we also recognize that there are also several limitations.

Benefits Gained by Using this Methodology

Field scientists are driven by a variety of objectives, but all are faced with the problem of determining where to go in the field. Voogd (1983) explains that a MCE provides an opportunity to classify a problem and allows an examination of the form, controls, and cost of a decision making process. The MCE process divides this problem into its various components. These components are then assessed to determine their form and control on the decision making process. Classifying the problem has been the first component of the MCE process described here. Characterization of criteria relevant to a problem require each to be considered both individually (i.e. with the analysis of digital data and derivation of criteria layers) and collectively (i.e. with the prioritization and aggregation steps). The developed methodology also makes possible an examination of form, control, and cost of the decision-making process by explicitly defining each criteria and objective, considering their relative importance, and assessing their overall contribution to meeting the campaign objectives.

The advantages of using path modeling include its suitability to analytically address the problem of field navigation. The final weighted surface produced by the MCE process explicitly indicates which areas within the analysis area will best account for a campaign's objectives. The traditional fieldwork approach would involve a more intuitive and iterative determination of these areas. Path modeling also provides a means to incorporate considerations of distance and time into the field navigational decision-making process. The largest controls on the traverse path placement are the values within the final weighted surface produced during the MCE. Also, the path modeling process accounts for distance and by proxy, time, by determining direct routes between origin and destination points.

Lastly, this methodology requires scientists to thoroughly investigate a study area prior to its visitation. The steps required implement the methodology leads scientists through work that has the potential to elicit valuable new information. This information may then be used as an advantage when the actual fieldwork begins.

Limitations of Methodology and Suggestions for Improvement

Many of the disadvantages of the MCE process relate to its complexity and the lack of a simplified framework for its use. While the process of deriving destination points and traverse paths is straightforward, it too lacks a framework for quick and easy implementation. ArcGIS ModelBuilder can be used to expedite many of the techniques required to develop the traverse paths. The development of a program to automate the entire process, however, would make this methodology more appealing and better suited to widespread application. This methodology also requires the repetition of many of its components once new parameters are introduced. For example, introducing new objectives, changing feature boundaries, or changing weights would all result in the need to repeat many steps.

Importantly, the methodology lacks the ability to quickly incorporate the information provided by field observations. An ultimate goal is the development of this methodology so that it may be employed while in the field to quickly generate new traverses once new observations have been made. Thus, the traverse would continually adapt to a scientist's understanding of a study area. This, of course, would require the development of a streamlined program that would allow the quick and easy incorporation of new information. The methodology developed here provides an example of a workflow that may be developed into such a model or into a mobile application.

Conclusion and Opportunities for Future Work

This paper has described how technologies and analytical approaches such as decision analysis, path modeling, and GIS can support navigational decision-making while preparing for and conducting scientific fieldwork. It demonstrates an alternative approach to traditional fieldwork planning and makes explicit key aspects of the navigational decision making process. While the intuitive and artistic aspect of field research will likely always remain, this work demonstrates the value of utilizing technologies that can provide meaningful assistance to its practice.

REFERENCES

ArcGIS, Version 10.2, Esri, Redlands, California, U. S.

Carver, S.J., 1991, Integrating multi-criteria evaluation with geographical information systems: International Journal of Geographical Information Systems. 5, 321–339.

Carver, S.J., 2008, Multicriteria Evaluation. In Encyclopedia of Geographic Information Science: ed. K. K. Kemp. Los Angeles, CA: SAGE Publications, 291-294.

Coe, A.L., 2010, Geological Field Techniques: Hoboken, NJ: Wiley-Blackwell.

Compton, R.R., 1985, Geology in the field: New York, NY: John Wiley & Sons.

House, P.K., Clark, R., and Kopera, J., 2013, Overcoming the momentum of anachronism: American geologic mapping in a twenty-first-century world: In Geological Society of America Special Papers, ed. V. R. Baker. 502, 103-125.

Malczewski, J., 1999, GIS and Multicriteria Decision Analysis: New York, NY: Wiley & Sons.

Mars, J. C., 2013, Hydrothermal alteration maps of the central and southern Basin and Range province of the United States compiled from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data (ver. 1.1, April 8, 2014). U.S. Geological Survey Open-File Report 2013–1139.

Mitchell, A., 2012, The Esri guide to GIS analysis: Modeling Suitability, Movement, and Interaction: Redlands, CA: ESRI Press.

Reeves, R., 2015, Deriving Traverse Paths for Scientific Fieldwork with Multicriteria Evaluation and Path Modeling in a Geographic Information System: Master's Thesis, University of Southern California.

Voogd, H. 1983. Multicriteria Evaluation for Urban and Regional Planning: London, England: Pion.