TASK-BASED METRICS FOR THE EVALUATION OF TERRAIN INTEROPERABILITY

David R. Pratt, David Dryer, Charles Campbell, and Richard Dunn-Roberts
Applied Software Systems Engineering Technology Group
Science Applications International Corporation
Orlando, Florida

Historically, computer generated forces (CGF) have been developed using the system-level approach. That is, they have been developed as a unit with the terrain they operate on and use for behavioral reasoning tightly coupled to the rest of the system. Over the last several years, the evident trend has been towards making various CGF systems interoperable. To do this, system developers and integrators have sought to use the same Terrain Databases (TDBs) and/or live with the inconsistencies that arise when linking different systems together.

This paper will present a set of grid, polygon and task-based metrics that addresses the perceived behavioral differences of CGF when operating in TDBs of differing fidelities (by fidelity, we mean how closely does the terrain match the real world) and formats. The metrics focus on atomic behavior tasks and those that are shared across other higher level behaviors, such as planning and following a route. From the quantifiable differences in the low-level tasks, we can quantify the effect of the underlying terrain representation on the higher level behaviors. The results provide new insight into terrain interoperability issues and their effects on CGF behaviors.

Dr. David R. Pratt is a Chief Scientist/Fellow at SAIC ASSET Group. Prior to joining SAIC, he had been the JSIMS Technical Director and a tenured Associate Professor of Computer Science at the Naval Postgraduate School. Dr. Pratt received a Ph.D. in Computer Science in 1993 and a Masters of Science degree in Computer Science in 1988 from Naval Postgraduate School.

Dr. David A. Dryer has over 11 years of combined industry and U.S. Army experience in training and simulation analysis, specializing in virtual environments, mission visualization design, simulation analysis, and human-computer interaction. He received a Ph.D. in Industrial Engineering from the University of Central Florida, a M.S. in Operations Research from the Naval Postgraduate School and a B.S. from the United States Military Academy. Dr. Dryer also has 7 additional years of U.S. Army experience in armored combat tactical operations as both a line and staff officer.

Chuck Campbell is a senior software engineer assigned to SAIC's ASSET Research Programs division. His responsibilities include design and development of Synthetic Natural Environment representations and services for CGF systems. Mr. Campbell holds a Masters of Science in Computer Science from the University of Central Florida and a Bachelor of Science in Computer Science from Indiana University. His interests include simulation, software engineering, computational geometry, and computer graphics.

Richard Dunn-Roberts is a senior software engineer assigned to SAIC's ASSET Research Programs division. He is responsible for design and development of software for Synthetic Environments and Computer Generated Forces. Mr. Dunn-Roberts holds a Bachelor of Science in Computer Science from the University of Central Florida. He has 11 years of professional experience in software engineering, specializing in virtual and constructive simulations, and computer graphics.

TASK-BASED METRICS FOR THE EVALUATION OF TERRAIN INTEROPERABILITY

David Pratt, David Dryer, Charles Campbell, and Richard Dunn-Roberts
Applied Software Systems Engineering Technology Group
Science Applications International Corporation
Orlando, Florida

INTRODUCTION

An evolving capability requirement for future Computer Generated Forces (CGF) systems involves the use of multi-resolution components, including Terrain Databases (TDBs). This concept is touted to having the ability to represent a large number of forces and still provide the detailed portrayal of a particular aspect of the simulation. To date many of these efforts have involved linking two existing simulations that operate one TDB of different resolution. For many reasons, these experiments have met only limited success. One of the key reasons is the complexity of the interacting complexity of the elements of the simulation. This results in a form of emerging differences that are more than the simple additions of the known differences in the simulations. Addressing these differences at that system-level is just too complex; there are too many unknowns [1]. To address a portion of this interoperability problem, this paper will address a series of experiments focusing on the terrain generation and low-level terrain services for a limited range of data sources and locations. This is not meant as an exhaustive and definitive study of what to do to achieve interoperability, rather it is a step in converting the "base of sand" to a base of insights and understanding.

The use of multiple levels of Digital Terrain Elevation Data (DTED) datasets as source data and differing parameters to the TDB surface generation process in current and future CGF systems raises issues concerning both the consistency of TDB representations and, more importantly, the consistency of the CGF across derived TDBs. In order to investigate the effect of multi-resolution TDB[1] on model behaviors, difficulties involving the quantification of model behaviors and terrain characterizations need to be addressed. Due to the large number of interacting attributes, behavior-based metrics that attempt to measure combat functionality at an aggregate level have proven to be difficult to quantify. There has been additional difficulty in investigating a terrain database's effect on behavior differences due to a lack of quantitative techniques to fully characterize the terrain. Such terrain characterization metrics have previously focused on global descriptions, such as average terrain roughness, which can hide the influence of more localized terrain factors critical to the low-level model task. For the purpose of this discussion, a behavior is a complex action that involves an ordered sequence of low-level tasks and a cognitive process. For example, the targeting behavior involves the detection and weapon tasks and the recognition and weapon selection cognitive processes. To help quantify the effects of the differing fidelities has on the CGF, this study will focus on the effect the terrain has on the tasks.

In order to address the above issues, this project has developed a task-based experimental methodology to investigate and quantify perceived behavioral differences of CGF when operating on TDBs of differing terrain densities and source resolutions. Such investigations are needed to provide clearer insight into CGF behavior sensitivities due to multi-resolution Synthetic Natural Environment (SNE) factors. A rigorous experimental approach is used to evaluate task-based terrain interoperability by examining task primitives of simulated terrain effects on entity behavior using a comprehensive characterization of various terrain resolutions and representations. A key point in task-based SNE analysis is to tailor the methodology and selected metrics to an operational domain. The operational domain being investigated in this study deals with armored vehicle and rotary wing aircraft combat, including ground mounted maneuver and rotary wing geometric intervisibility tasks.

The experimental methodology consists of two analysis components, (see Figure 1 and Figure 2). Experimental areas representing a variety of geomorphologic types are selected from standard format 1° by 1° DTED cells [2]. The first part of the experimental analysis is quantitative terrain characterization and comparison of experimental areas at multiple data resolutions, using selected *grid* and *polygon*[2] terrain metrics. The terrain *grid* metric comparisons (see Figure 1) directly characterize the various resolutions of source DTED cells

[1] Only the terrain surface configuration component of a Terrain Database (TDB) is being addressed by this experimental plan.

[2] While we refer to polygons, the underlying polygonal representations are limited to triangles as part of a Triangular Irregular Network (TIN) representation.

across three DTED levels, 0, 1 and 2. Grid metric comparisons are meant to vary in spatial extent[3] by characterizing terrain cells with both a global metric area statistic (cell wide bin of 3600 arc sec^2 area) statistic and multiple local metric area statistics (subdivided cell bins of 1200, 300, 120, 30, and 3 arc sec^2). The other terrain category of terrain characterization deals with *polygon* metrics. Multiple Triangulated Irregular Network (TIN) representations of each DTED source data resolution are created using various TIN-generation process input parameters. Polygon metric bin values, such as vertex intervisibility range, are used to vary the spatial extent of polygon metric comparisons.

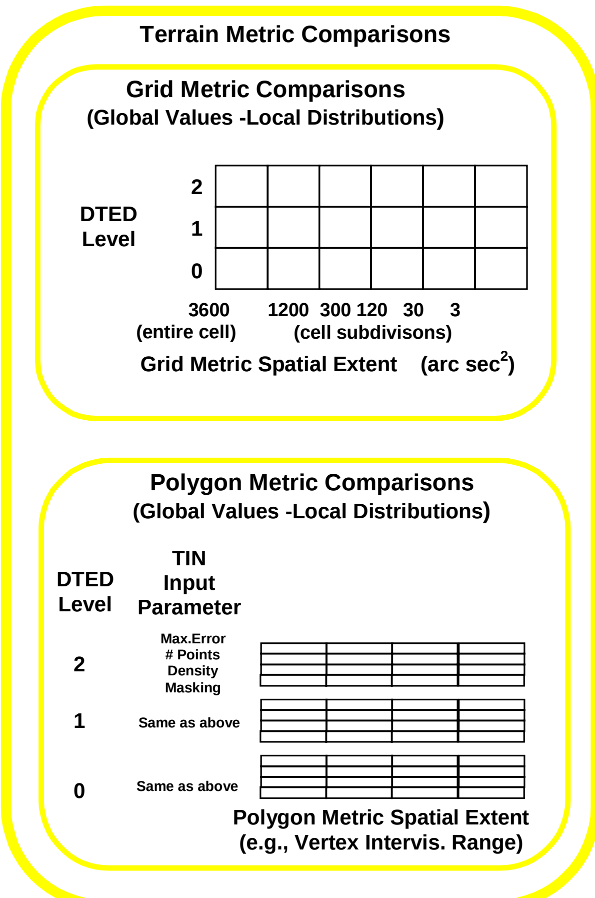

Figure 1. Terrain Metrics Components

The second part of the experimental analysis deals with task-based metric comparisons (see Figure 2). As shown in this figure, DTED source data resolution, TIN input parameter(s), task-based metric spatial extent (e.g., distance between endpoints used to generate routes), and experimental area regions, are varied during these task-based SNE comparisons. For task-based metric analysis, experimental areas are evaluated as entire regions and subset into terrain Regions of Interest (ROI) based on localized terrain characterizations. Such ROI subsets should be regions appropriate to the domain of ground maneuver and rotary wing intervisibility being used in this analysis. The localized terrain characterization metric planned to describe such ROI is mean elevation of local sub-cell areas. However, other grid metrics will be assessed during terrain characterization experiments for potential use in identifying task-based ROI. Further details on the research methodology involved in these terrain metric and task-based metric comparisons are contained in following sections. Results from this series of experimental comparisons can address critical unknown SNE sensitivity issues and provide new insight into multi-resolution terrain interoperability and its effects on CGF behaviors.

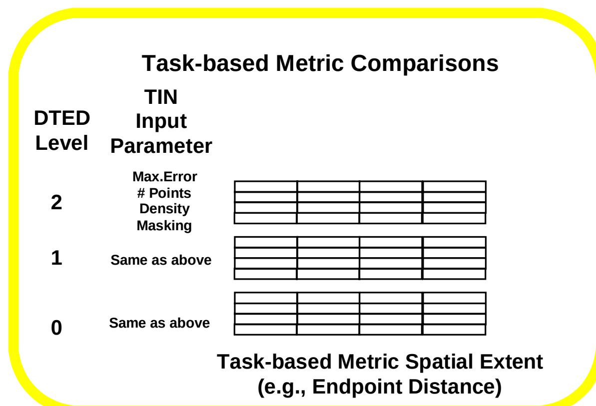

Figure 2. Task Based Metrics

Three terrain areas were used for initial experimentation. The experimental areas selected are standard 1^o by 1^o DTED cells representing a range of geomorphology. Initial direction from the technical representative was to target low roughness (smooth/rolling terrain), medium roughness (hilly), and high roughness (mountainous) as a range of characteristic areas of potential armored and rotary-wing combat maneuver. This gives a range of representative terrain types with respect to roughness and geomorphology that can be expressed in both statistical and operational terms to validate the extensibility of the experimentation results. Investigation is currently ongoing as to the availability of DTED Level 0, 1, and 2 products for selected 1^o by 1^o cells in these areas. In addition, the regions that are selected are restricted to contain no more than 25% water coverage to avoid large "flat" areas skewing the results. From this initial guidance, we selected three DTED

[3] Spatial extent of terrain metrics refers to the geographic area or parameter length used in calculating various grid and polygon metric statistics.

cells based upon qualitative measures. Elevation shaded images were made of candidate areas. These were then projected side-by-side on a large wall display. This allowed the team to conduct a subjective analysis of the roughness of the terrain. This was done by visually assessing the size and quantity of the elevation bands. The more narrow the bands, the steeper the terrain. The closeness of ridges and valleys provided insight in the roughness of the terrain. It was based upon this qualitative analysis that the initial areas were selected. The final areas are Desert (rolling terrain, low roughness, (see Figure 3)), Hilly (medium roughness, (see Figure 4)), and Rugged (mountainous, high roughness, (see Figure 5)).

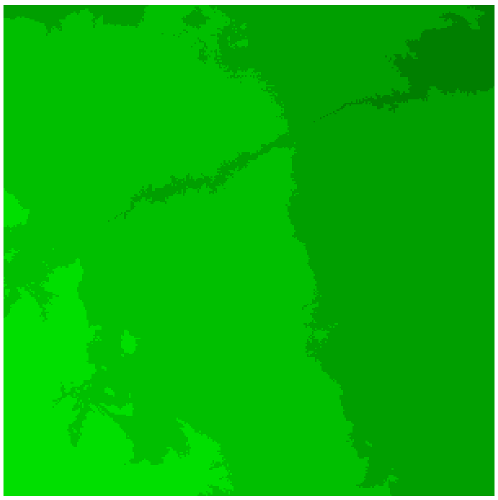

Figure 3. Smooth Roughness Sample

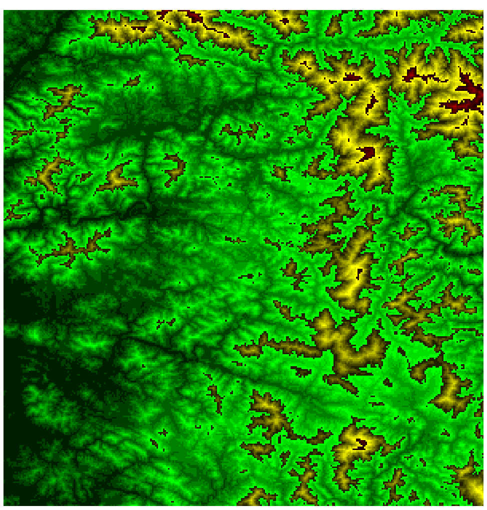

Figure 4. Medium Roughness Sample

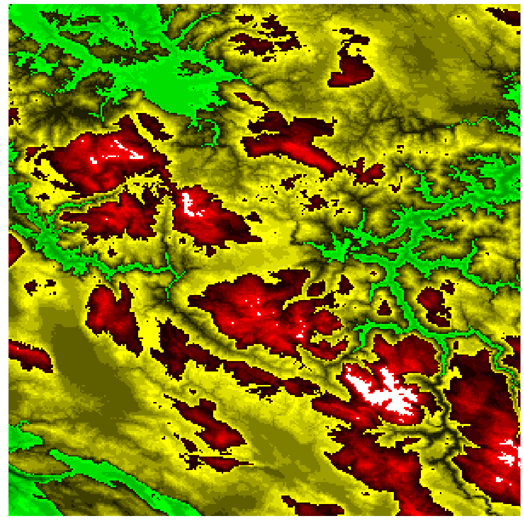

Figure 5. High Roughness Sample

TERRAIN GENERATION PIPELINE

The terrain database generation process uses a series of conversion routines to derive a Warfighter's Simulation 2000 (WARSIM)-compatible formatted terrain database from the source DTED data. Graphically portrayed below (see Figure 6), the process is representative of most TDB generation processes with the exception of the manual editing steps required for feature additions. The terrain generation capability used in this project uses DTED Level 0, 1, and 2 data as the source datasets. This data is triangulated and imported into the WARSIM terrain compiler to produce a WARSIM terrain surface database with no real surface type attribution. As a result of our desire to minimize the amount of new code developed for this program, the data undergoes several format and datum transformations along the way. While the number of transformations might seem excessive, it is not uncommon in the TDB generation process. This is the critical problem that the Synthetic Environment Data Representation Interchange Specification (SEDRIS) program is addressing.

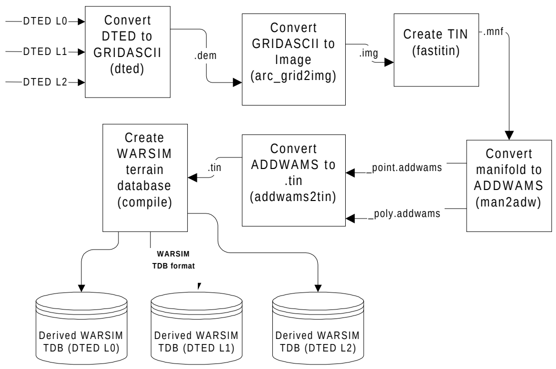

Figure 6. TDCS-SNE Database Generation Process

The grid metrics are derived before the first stage of the transformation pipeline. The first stage of the pipeline converts the DTED source data to ARC/INFO GRIDASCII format. This conversion is required because the Carnegie Mellon University (CMU) Integrated Triangular Irregular Network (iTIN) tool does not read DTED source data directly [3]. Instead, the TIN tool uses CMU Image format as the input format. The CMU tool set provides a GRIDASCII to Image converter (*arc_grid2img*), but does not provide a DTED to Image converter. This required that SAIC develop a DTED to GRIDASCII converter (*dted*). This converter reads DTED Level 0, 1, and 2 data files and writes GRIDASCII files (.dem). The *dted* utility runs under SunOS.

The second stage of the pipeline is the conversion from GRIDASCII to CMU Image format (.img). This is done using the CMU converter *arc_grid2img*. This utility runs, as do all tools in the CMU tool set, under IRIX on SGI workstations.

The third stage of the pipeline is the generation of the CMU TIN. The iTIN tool[4] developed at CMU (*fastitin*) imports a CMU Image file, constructs the TIN, and exports manifold file format (.mnf). At this stage of the pipeline, there are several options that may be used in processing the elevation data into the TIN structure. These parameters include max Root Mean Square (RMS) elevation error (-*e*) or max 3D error (-*s3*), number of vertices in TIN (-*p*), overload control (limiting polygon density in a load module), and masking (-*m*) to force more polygons to flatter regions. The CMU tool also outputs TIN analysis information, which is used for further comparisons. For instance, if the number of vertices input parameter is held constant while the input datasets are of varying resolution, the error output information can be subsequently compared.

The fourth stage of the processing pipeline is the conversion of the manifold format to MOSS Export format (ADDWAMS). The CMU TIN tools include a converter from manifold format to ADDWAMS (*man2adw*).

In the next stage of the processing pipeline, the ADDWAMS file format is converted to the Centric Software TIN (.tin) format. An SAIC developed tool (*addwams2tin*) is used to convert ADDWAMS to the Centric Software TIN format. The *addwams2tin* tool runs under SunOS.

Finally, the Centric Software TIN file is compiled by the WARSIM terrain database compiler into the format that will be read by the locally developed software. The WARSIM terrain compiler directly imports the Centric Software TIN format. It is at this stage that the polygon metrics are computed.

While this process contains a series of seemingly unnecessary format conversions, it was our desire to treat both the CMU TIN generation software and the WARSIM terrain database compiler as black boxes that required these conversions. The alternative would have required a significant amount of software development effort on the two tools that would not have benefited the goals of this study.

METRICS

As discussed above, the terrain areas were selected using qualitative judgements by a team that has been involved in the TDB process for many years. While this is fairly representative of the state-of-the-art, it is far from an automatable process. To address this, we rely on metrics. Metrics are objective measurements of a process or characteristic. Very seldom is a single metric of much use. Rather, they are used in sets to give a fuller picture of the item under study. As discussed above, we use three sets of metrics, collected at different points along the TDB pipeline, to give us comprehensive insights into the effects of the TDB on the basic tasks a CGF performs.

Grid Metrics

The first set of metrics measure the source DTED cells. Grid metrics, mean elevation, standard deviation/sigma - t (σ - t), and surface roughness index, can be calculated on both a global (cell wide) and local (grided subset) basis.

[4] One of the more powerful features of the iTIN tool is the ability to integrate a road network into the underlying TIN structure. This results in an Integrated Triangular Network (iTIN). Since the input DTED data does not have any roads, we will not be using this feature.

Mean Elevation Metric

Mean elevation for partitions of varying spatial extent can be used for DTED source data resolution comparisons and terrain characterizations. The distribution of the elevations will provide insight into the overall shape of the terrain. These measures can be used for assessing mean elevation distributions for selected terrain areas and to help identify regions of interest for future task-based metric analysis. As shown below, mean elevation is the summation of a set of grid post elevations, divided by the number of grid points. The grid post elevations used in the calculation are DTED vertical datum representing Mean Sea Level (MSL). The possible range of these elevations in practice is -12,000 meters to +9,000 meters. The absolute accuracy of the input National Imagery and Mapping Agency (NIMA) DTED vertical datum is 90% Linear Error (L.E.) MSL $\leq$ 30 meters. The relative accuracy (point-to-point) of the input NIMA DTED vertical datum is 90% L.E. MSL $\leq$ 20 meters.[2]

$$\overline{e_{i,j}} = \frac{\sum_{i=1}^{nr}\sum_{j=1}^{nc} e_{i,j}}{(nr * nc)}$$

$\overline{e_{i,j}}$ = Mean elevation
nr = Number of rows in selected area
nc = Number of columns in selected area
$e_{i,j}$ = Grid post elevation at location i,j

Standard Deviation Analysis (σ - t) Metric Description

Standard deviation analysis, also called sigma - t or σ - t (t indicating terrain) is a measure of terrain roughness[5]. U.S. Army Topographic Engineering Center (USA TEC) provided a formula to Evans & Sutherland (E&S) to help quantify the surface of the terrain as part of the Close Combat Tactical Trainer (CCTT) program. σ - t, which is the standard deviation of a set of grid post elevations, was one of several metrics used by TEC and E&S. Defined in a [4], the equation is a formulation of standard deviation. Using formula, σ - t can be described for a set of grid post elevations, as the square root of the summation of the square of the difference between the grid point elevations and the mean elevation, divided by the number of grid points minus 1.

$$S = \sqrt{\frac{\sum_{i=1}^{nr}\sum_{j=1}^{nc} e_{i,j}^{\,2} - \frac{\left(\sum_{i=1}^{nr}\sum_{j=1}^{nc} e_{i,j}\right)^2}{(nr * nc)}}{(nr * nc) - 1}}$$

S = Sigma–t (σ - t)
nr = Number of rows in selected area
nc = Number of columns in selected area
$e_{i,j}$ = Elevation grid post height at location i,j

The CCTT program found certain limitations of σ - t when used in "flat" areas. Specifically, areas with identical σ - t values can have noticeably different topography and polygon counts. This led to the method of identifying differences in such flat areas, defined as Surface Roughness, and is covered in the next metric section.

Surface Roughness Index Description

The terrain or surface roughness index uses neighboring elevation grid posts to measure the changes in elevation inflection around each grid post. This E&S surface roughness metric is defined in [4]. Small roughness index values indicate small changes in elevation inflections and large roughness index values indicate large changes in elevation inflections, which can be defined as rough terrain. A key point about this metric is that a small number is calculated for a set of grid post elevations representing an area of uniformly steep slope. Even though adjacent grid posts have large elevation differences for uniformly steep terrain, there are few inflection points on the terrain and this is not considered rough. A set of grid posts where adjacent grid posts have large elevation differences in different directions will result in a large roughness index because there are frequent changes in elevation inflection. A limitation of this metric is that it considers only the points directly north-south and east-west from the point under consideration, the diagonals are ignored. However, much like σ-t, this metric can be applied to the entire DTED cell, *nc* = max_columns and *nr* = max_rows, or just a gridded subset.

$$R = \frac{\sum_{i=1}^{nr}\sum_{j=1}^{nc}\left\{\frac{\left|2e_{i,j} - (e_{i,j+1} + e_{i,j-1})\right| + \left|2e_{i,j} - (e_{i+1,j} + e_{i-1,j})\right|}{4}\right\}}{(nr * nc)}$$

R = Roughness Index
nr = Number of Rows in sample
nc = Number of Columns in sample
$e_{i,j}$ = Elevation grid post at location i,j
$e_{i,j+1}$ = Elevation grid post immediately east of $e_{i,j}$
$e_{i,j-1}$ = Elevation grid post immediately west of $e_{i,j}$
$e_{i+1,j}$ = Elevation grid post immediately north of $e_{i,j}$
$e_{i-1,j}$ = Elevation grid post immediately south of $e_{i,j}$

[5] In this context, "roughness" is a qualitative term indicating the irregular variations between the elevation posts.

Polygon Metrics

Once the gridded surface has been converted to a polygonal database, a new set of metrics has to be used. Rather then measuring the differences in the regular spaced grid posts, these metrics examine the relationships of the polygons to each other.

Adjacent Polygon Variance

A TIN metric computed on a per polygon basis which compares the normal of the polygon to the three adjacent polygons using weighted average based on polygon area in the plane of each polygon. A polygon normal, by definition, has length of 1 and is orthogonal to the polygon. When comparing the two normals, we compare the magnitude of the X, Y, and Z components. We have chosen to use the Root Mean Square (RMS) to capture the absolute differences between the components. If there is no difference between the normals (RMS error is 0), the polygons are co-planar and, therefore, very smooth. The maximum RMS error value will occur when the polygons are at right angles, maximum roughness. A weighted average based on the sum of the polygon areas was used to reduce the impact of small sliver polygons. Adjacent Polygon Variance is a measure of local roughness and is shown below:

$$A_t = A_p + A_a + A_b + A_c$$
$$P = \sum_{i=a}^{a,b,c} \frac{A_i * RMS(N_p, N_i)}{A_t}$$

P = Adjacent polygon variance metric
A_t = Total polygon areas
A_p = Polygon area
A_a = Adjacent polygon A area
A_b = Adjacent polygon B area
A_c = Adjacent polygon C area
N_p = Polygon normal
N_i = Adjacent polygon normal

Vertex Intervisibility

This is a modification of a method suggested by Dr. Robert Richbourg [5]. An observation point is placed at a specified elevation above the surface of the polygon at its center. The elevation can be determined by pre-experimentation to give the widest distribution of ***V*** possible or derived based upon a particular weapon system's sensor height. Setting the range to a fixed radial distance value (i.e., 1, 5, 10 km) to vary spatial extent of the metric helps to eliminate any local bias. Geometric intervisibility is then computed between this point and all of the polygon vertices within range. The resultant metric is a ratio between the number of visible vertices and the total number of vertices.

$$V = \frac{\#of\_visable\_Pnts}{\#of\_Pnts\_within\_range}$$

Projected Area

This is a local polygon metric, which can be used by localized area or globally for experimental areas. It measures steepness of the polygon (omni-directional slope) in isolation and is shown below. By calculating projected area with both a global metric area (cell wide bin of 3600 arc sec^2 area) statistic and multiple local metric area (subdivided cell bins of 1200, 300, 120, 30 and 3 arc sec^2) statistic, the steepness of derived TDBs can be characterized and compared to grid steepness metrics, such as σ –t.

$$A = \frac{Area_{2d}}{Area_{3d}}$$

Task-based Metrics

The task-based metric comparisons focus on atomic CGF behavior tasks, such as ground model mobility and hover model geometric intervisibility. These two metrics were chosen since they represent two of the fundamental actions a CGF performs. Performance metrics, such as access time as a function of polygon density, are also collected for comparison.

Task-based Metrics Definition and Generation

There are several descriptors that quantify the "degree" of interoperability achieved. Ultimately, when two simulations interoperate, commanders should expect certain behaviors from interoperating CGFs in support of the assigned tasking. While there are several ways of doing this, ensuring that the atomic tasks performed by the entities are consistent across the entire simulation is a first step in achieving the interoperability the commander expects. If the commander is looking at a single simulation, he should not need to know, or care, that there are different models underneath. This led us to constructing metrics based upon the task entities performed.

As part of this research project, we are using modifications to a subset of the task primitives developed under the Composable Behaviors Technology (CBT) effort as the foundation for the task-based metrics[6]. The table below (See Table 1) lists the experimental task primitives, actions performed, and dependent variable metrics that will be captured for subsequent analysis. As atomic tasks, these provide the ability to evaluate the interactions with the terrain skin in isolation without dealing excessively with the differences in cognitive representations of the systems.

Table 1. Experimental Task Based Metrics

Exp. Task	Action Performed	Evaluated Dependent Variables
1. Straight-line Path	Straight-line path generation and entity movement along path	-3D distance of straight-line path -Ground model time to traverse
2. Trafficable Route	Entity trafficable path generation and movement	-Surface length of path -Ground model time to traverse -Location of path
3. Geometric Intervisibility (GI)	Hover primitive until unmask	-Vertical height

Geometric Intervisibility Metric

The geometric intervisibility metric is designed to simulate a helicopter popping up from a concealed position.

1. Select a point at random within the Experimental Area/ROI.
2. Select another point at random within the Experimental Area/ROI using the designated 3D Euclidean distance bin size (1, 5, or 10 km).
3. Place an eyepoint at the first location at an altitude of 0.5m.
4. The geometric intervisibility between the two points is checked in "hover mode." If intervisibility is achieved, the altitude at which intervisibility occurs is noted. The target point is the 50% height of a typical vehicle, 1 meter above the ground.
5. If intervisibility is not achieved, the altitude of the entity at the first point is incremented by 0.5m increments and Step 4 is repeated.

Straight Line and Trafficable Route

The second set of metrics is designed to simulate the movement of a very simple vehicle traversing the terrain.

Path/Route Generation

The goal of the path selection is to obtain a number of straight-line paths and trafficable routes on each terrain type. There will be a finite set of paths of each distance within each terrain type. Each path will consist of two endpoints with an Euclidean 3D distance between the endpoints of 5km, 15km, or 30km. These endpoints should be chosen at random.

- Select a particular experimental ROI to be used.
- Select a distance to be used for this particular pair of endpoints (5km, 15km, or 30km).
- Select a point at random within the area of interest. This point is rejected if it falls within a no-go region (water or terrain too steep). Per the selection of the three basic areas above, the only large bodies of water are oceans/seas. Since they have elevation 0 (and are the only ones that do), any point selected with zero elevation is considered to be in the water. For our purposes, slope is too steep when it exceeds 60 degrees in any direction. In the case of a rejection, the point is discarded and a new point is determined.
- The second point is determined by randomly selecting an angle from the first point and projecting a vector out by the fixed distance (5km, 15km, or 30km). This second point is rejected if it falls outside the terrain area selected or if it lies in a no-go region. In the case of a rejection, a new point is selected.
- (Experiment 2: Straight line path) From Euclidean 3D line between endpoints, project surface path onto the terrain as a straight-line path from the start point to the endpoint.
- (Experiment 3: Trafficable route) An A* search routine is used to determine a trafficable path from the start point to the endpoint. The cost function is the actual time of traversal. The predictor function is the time to traverse the 3D distance at max velocity. The evaluation function is the summation of the two.

Data Collected

A generic entity model is used to traverse a path between the two endpoints that have been selected. The entity model moves in fixed time step increments and determines the following polygon parameters at each time step:

- Location
- Surface Normal

In addition to the parameters collected at each time step, the simulation also collects the following aggregate parameters for both the straight line and trafficable paths:

- Time to traverse
- Number of time steps elapsed
- 3D path length

Route Traversal

A fixed time step will be used. A time step of 0.5 second was chosen. This results in a maximum traversal distance of 4.167 meters. This is a factor of the unclassified advertised cross-country speed of an M1A2 Main Battle Tank of 30 kph. The maximum traversal distance is approximately an order of magnitude smaller than the DTED Level 2 post spacing. Likewise, it is consistent with the minimum update rate required for stability in the current Modular SemiAutomated Forces (ModSAF)-based systems.

- At the first point determined, an entity is placed.
- The location (x, y, z) of the entity is retrieved from the database. At this point, additional checks are made to determine if the entity has left the ROI in an effort to move around a no-go area (in the trafficable path case). If the entity has left the Area of Interest (AOI), the path is rejected.
- The surface normal for the location of the entity is retrieved from the database. Based on the entity orientation, it should be determined if the slope of the terrain polygon is too great for the entity to traverse (trafficable path case).
- The trafficability code for the terrain polygon that the entity is over is retrieved from the terrain database. This is used to determine if the entity can traverse this terrain polygon (trafficable path case). Since we are only dealing with DTED derived databases, polygons with elevations of 0 are assumed to be untrafficable water. All other polygons will have a fixed trafficability code.
- The speed of the vehicle is determined, using the slope and the trafficability code of the terrain polygon upon which the vehicle is currently located. Note that this will differ depending on which mode of operation is used, straight-line path or trafficable path.
- The vehicle position is updated.
- The number of time steps is incremented.
- This process is repeated until the endpoint (Point 2) is reached or until the vehicle leaves the terrain database in an effort to reach the endpoint (Point 2). Note that the second case will only happen if the entity is searching for a trafficable path.

PRE-EXPERIMENTATION

Pre-experimentation is needed to derive an estimated maximum grid-based metrics error ratio from the DTED relative vertical accuracy of 20m at 90% linear error. A sample DTED Level 1 database is selected and random vertical datum points are modified by collapsing them 20m toward the mean elevation. The metrics are then calculated on the original and modified data. Then, a ratio is calculated between the square of the original and modified database values, with the larger value on top. This gives an estimate of the maximum σ-t ratio variance due to within DTED relative vertical accuracy error. This variance will be used to determine what constitutes "significant differences."

METRICS EXPERIMENTATION

Experimentation was comprised of the generation and collection of the metrics. To ensure consistency and reduce the manpower required and possible contamination of the results, a set of scripts were developed to run the metric collection run in a fully automated mode. This also allowed the processing to be done overnight when the computation resources were readily available.

As expected, the grid metrics were collected first. The metrics were taken globally (one 3600 arc sec^2 sample), and in a series of progressively finer grids (1200 arc sec^2, 300 arc sec^2, 120 arc sec^2, and 30 arc sec^2 sample sizes). The finest grid was limited by the spacing of the DTED level 0 data.

The polygon metrics were recorded globally, based on the same grid sizes as the grid metrics, and at the polygon level. They were initially taken on a per polygon basis. An area based averaging was used for the generation of the global value. To determine the subsampled grid values, the polygons were first clipped to the geographic boundaries. The metrics were then weight averaged based upon the resulting polygon's geographic coverage.

The final set of metrics collected were the task based metrics. Since each of the metrics encompasses a geographic extent, they are dealt with on a global basis for each of distance bins. Task metric levels, such as path and route endpoint lay down distances (3D Euclidean) and target-to-hover model shooter distance, are used to vary the spatial resolution of selected experimental task metric comparisons. Only one set of endpoints was generated for each DTED cell. This ensured that the metrics collected could be compared against others collected on TDBs that covered the same geographic area but were generated from different sources or TIN parameters.

EXPERIMENTAL ANALYSIS

Three basic sets of analysis were conducted. There are three distinct axes for comparisons and correlation (see Figure 7). The first of these is along the horizontal axis, the metric type. The purpose of this set of correlation analysis was to ensure that the attributes that the metrics measured were preserved through out the TDB generation process. For example, an area that had a high value for the grid surface roughness had a comparable value for the polygon variance metric. Likewise, the relationship between roughness and stepness were preserved.

The second set of analysis was along the vertical axis, the different sources and TIN parameters. This was the heart of the experimental effort. The different TDB represented the multi-resolution aspect of the problem. The study focused on the consistency of the metrics across the databases. For example, while we would expect to see differences in the individual time to traverse metrics, we would expect the distribution of times to be comparable across the sources and TIN parameters.

The final axis of comparisons was the depth, across the different geographic areas. In this case were looking for constancy of the metrics across the geographic samples. For example, the relationship between the roughness metric and the steepness metric on the mountainous terrain should be the comparable to the relationship in the medium and smooth terrain.

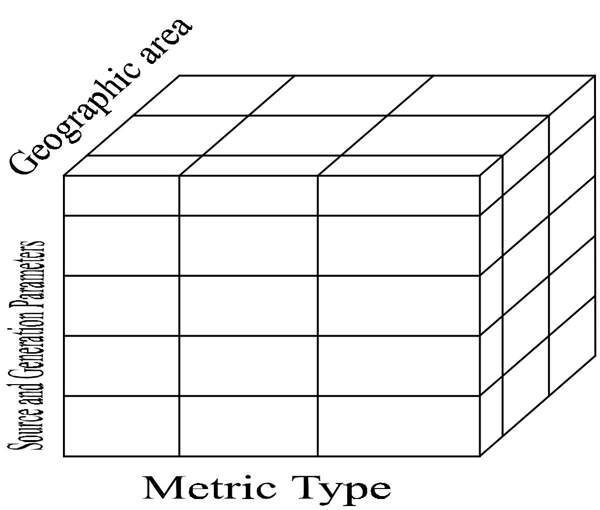

Figure 7. The Relationship of the Metrics

ANALYTIC TECHNIQUES

By themselves, metrics do not mean much. It is only when they are analyzed that they take on value. Depending on the metric, a range of techniques was used. The following represents some of the methods applied in determining the relationships between the metrics, source databases, and the tasks.

- Exploratory data analysis of metric statistics (mean, variance, consistency MOC)
- Image Generation using pixelation of the metric under consideration.
- Large-sample (1-a)100% Confidence Interval for $m_d = (m_1 - m_2)$: Matched Pairs
- Bartlett's test of homogeneity of variance
- One way ANOVAs of experimental task cases
- Schuirmann's Two One-Sided Test for equivalence using confidence intervals to assess consistency (proven in bioequivalence studies)

With a sample size of one, this entire cell metric is not adequate for conventional statistical hypothesis testing. However, the following comparison test can be done using previously described DTED datum accuracy requirements. For each hypothesis, a difference is calculated between the two metric's values. This difference is compared to the result determined in the pre-experimentation phase. If the difference is less than the error sample, a determination that no difference exists between the two metrics. It is only when the difference exceeds that of the error threshold do we determine that correlation does not exists.

RESULTS

The results to date have shown a strong consistency across the DTED levels, metrics, and subregion samples, as expected. However, there have been two major revelations from this set of experiments concerning the source data and the metrics. As stated above, three levels of DTED source data were used as inputs for the experiments. Given the geographic areas that were used, it was expected that there would be minor differences in the DTED levels 1 and 2, which are known to come from different sources. The DTED level 0 was expected to be a "thinned" version of the DTED level 1, but oddly enough, the results indicate this is not the case. In two of the three samples, the DTED levels 1 and 2 exhibited higher correlation than the DTED levels 0 and 1. The third case performed as expected. Though the differences between the various levels are within the error limits of the DTED specification [2], the differences were a source of variation. We have decided to use the DTED level 2 as the source and perform our own thinning as a means of controlling the terrain generation process and reducing the number of independent variables in the study.

The other revelation dealt with the Surface Roughness Index (SRI) metric. We noticed that the DTED level 0 datasets were always rougher (i.e., had a higher SRI value) than the DTED level 1 datasets which, in turn, were rougher than the DTED level 2 datasets. This was contrary to what we had expected. Yet, the SRI values were consistent with expectations when we compared the different geographic areas to each other. This caused us to go back and reexamine the metric.

The SRI metric is basically the measure of the variation of the slope between the grid points. Upon reexamination, we found that the "4" in the denominator assumes that a constant distance between the grid posts exists. This is a valid simplifying assumption when the datasets under consideration are uniformly spaced grids. With minor variations accounting for latitude, this was acceptable when comparing the same level datasets representing the differing geographic regions. However, when this metric was used to compare the different DTED levels, the results became skewed. Accounting for the difference in post spacing gave us the following equation:

$$R = \frac{\sum_{i=1}^{nr} \sum_{j=1}^{nc} \left\{ \frac{\left| 2e_{i,j} - (e_{i,j-1} - e_{i,j+1}) \right|}{4d_{ewi}} + \frac{\left| 2e_{i,j} - (e_{i-1,j} - e_{i+1,j}) \right|}{4d_{ns}} \right\}}{(nr * nc)}$$

R = Roughness Index
nr = Number of rows in sample
nc = Number of columns in sample
$e_{i,j}$ = Elevation grid post at location i,j
$e_{i,j+1}$ = Elevation grid post immediately east of $e_{i,j}$
$e_{i,j-1}$ = Elevation grid post immediately west of $e_{i,j}$
$e_{i+1,j}$ = Elevation grid post immediately north of $e_{i,j}$
$e_{i-1,j}$ = Elevation grid post immediately south of $e_{i,j}$
d_{ns} = Distance between two posts in the north-south direction
d_{ewi} = Distance between two posts in the east-west direction

We are continuing to run experiments to validated both our hypothesis and the validity of the metrics approach using both geospecific and geotypical terrain datasets.

CONCLUSIONS

Consistent with our expectations, preliminary results suggest a correlation of the metrics within a local geographic area. However, there are variations in single metrics that strongly indicate that the use of any one metric in isolation can lead to a false impression of the underlying terrain skin characteristics. However, a tuple of the grid, polygon, and/or task metrics does present a viable means of characterizing the terrain skin.

As the geographic area increases, there is a tendency for all the metrics to become less indicative of the type of terrain or of the source. Some of this results from the normalizing effect of using larger data samples. But when looking at larger geographic areas, metrics also tend to "average" out. For example, when combined, the metrics for a meadow and a steep hillside are equivalent to those of a set of rolling hills. Like all approaches to reduce the amount of data and characterize data, care must be taken not only in the generation of the metrics but also in their interpretation.

We have taken a significant step forward in our ability to quantify the characteristics of the terrain surface. A significant amount of work remains to be done in interpreting the results and analyzing their effects on simulations to include CGF behaviors, interoperability among heterogeneous systems, and ultimately, the end results of simulation runs.

ACKNOWLEDGMENTS

This work was performed under contract number N61399-99-C-0040 as part of the Defense Advanced Research Project Agency (DARPA) Advanced Simulation Technology Thrust (ASTT) Program.

REFERENCES

[1] Bundy, Gary N., David W. Seidel, Ben C. King, and Carl D. Burke. “An Experiment in Simulation Interoperability.” Proceedings of the 1996 Winter Simulation Conference. Edited by J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain.

[2] Defense Mapping Agency, Performance Specification: Digital Terrain Elevation Data (DTED), MIL-PRF-89020A, 19 April 1996.

[3] Polis, M., Gifford, S., McKeown, D. "Integrated TIN Generation User's Manual", Carnegie Mellon University Working Document, Revision 1.0, 15 April, 1996

[4] "CCTT Terrain Component Design Report for The Central U.S. Database - Terrain Skin Analysis" Version 1.0, July 26, 1994.

[5] Richbourg, R. F., C. K. Ray, and L. W. Campbell. “Tactical Operations Planning as a Metric for Terrain Fidelity.” 95-13-012. 13^{th} DIS Workshop, Orlando, Florida.

[6] von der Lippe, Sonia R., et. al., "An Architectural Overview of Composable Behavioral Technologies", Proceedings of the 7th Conference on Computer Generated Forces & Behavioral Representation, May 12-14, 1998. Orlando, Florida.