# Rationalizing Source Data: Tailoring Geospatial Data for Advanced Distributed Simulation

*David M. McKeown*
*TerraSim, Inc.*
*Pittsburgh, Pennsylvania, USA*

*William Starmer*
*TerraSim, Inc.*
*Pittsburgh, Pennsylvania, USA*

*Siddharth Shankar*
*TerraSim, Inc.*
*Pittsburgh, Pennsylvania, USA*

*Evan Klauda*
*TerraSim, Inc.*
*Pittsburgh, Pennsylvania, USA*

## ABSTRACT

*The production of geospatial environments for modeling, simulation, and training starts with the acquisition of suitable source data, imagery, digital elevation models, various vector data layers, and representative or geospecific 3D models. Whether the resulting modeling and simulation runtimes are fully compiled or created "on the fly" for presentation, significant effort must be expended to rationalize disparate data sources prior to the database production. In this paper, we describe a set of source data preparation tools that have been developed to address the generation of derived data sets. These tools are tailored to the requirements of the database generation process and have been adapted to the production of datasets which support correlated runtimes across visual, constructive, and serious game runtimes. Specifically, we will discuss issues in the preparation of large area digital elevation models (DEMs) and the generation of road networks and surface material maps from remotely sensed imagery. We will focus on the unique requirements of advanced ground simulation, such as including expanded environmental data models that contain detailed attribution and tailoring databases to address differences and limitations inherent across multiple simulation runtimes. These issues are particularly important in networked simulations where different simulations are running in an overall distributed exercise and are expected to be correlated in terms of mobility, sensor simulation, cultural content, and intrinsic properties of the terrain environment.*

*Specific examples will be taken from a range of constructive (i.e., OneSAF, JointSAF, JCATS) and serious game (i.e., Steel Beasts, VBS2, Havok) runtimes.*

## PAPER SUMMARY

One of the first steps in the generation of runtime environments for modeling and simulation is determining the available geospatial source data that meets the basic content requirements of the simulation. As most readers are well aware, basic source datasets generally need to be enhanced, augmented, or merged with disparate datasets in order to meet exercise requirements. When database generation is required across multiple disparate runtimes, particularly combinations of visual, constructive, and serious game runtimes, the source data selection problem is compounded since the requirements quickly become the superset of each individual simulation.

Traditional commercial geospatial data processing products such as ESRI ARC-X, ERDAS Imagine, ENVI, Global Mapper, are applied to manually by skilled GIS analysts to mitigate many of the shortfalls in source data content, areas of missing content, resolving scale related conflicts due to the merge of various data layers from different sources. An example of this is the overlap of features such as buildings and roads from datasets that were not compiled from the same source materials. Likewise combining a contiguous road network from a variety of low-medium-high resolution sources often leads to geometry issues resulting in gaps and overshoots in the network and missing 2D topology necessary to ensure planning and routing during a simulation.

In this paper we discuss four issues that we believe are representative of source data preparation problems:

(1) Rationalizing digital elevation models (DEM's) from disparate geospatial sources. Challenges include creating one consistent DEM to represent the surface of the terrain for all simulations. Problems include the need to project (or re-project) DEM data with different elevation post spacing, using different vertical datums (ellipsoid / geoid), overlapping coverages with different quality source, and integration of LIDAR coverages, a digital terrain model (DTM) to support high resolution insets.

(2) Rationalizing road networks derived from disparate geospatial sources. Challenges include the generation of complete road networks within complex simulation areas. Problems include cases where the vector source data is out of date with the imagery being used in the simulation, multiple cartographic sources are used, leading to road feature duplication,and road feature overlap and horizontal displacement is evident from visual inspection across the entire database area.

(3) Rationalizing surface material information to meet the needs of various simulation runtimes including trafficability, mobility, visual effects, and sensor modeling. Challenges include the generation of surface material maps at a spatial resolution that is consistent with runtime requirements which vary wildly for visual, constructive, and serious game runtimes. Surface material maps can be created at a spatial resolution that far exceeds that is necessary or can be represented in most simulation systems. Techniques to generalize and aggregate detailed material maps while maintaining consistency of material attribution across disparate runtimes is discussed.

(4) Rationalizing the use of geospatial source data which is typically embedded in legacy simulation runtime databases and where the original source data is not available to the end user. Challenges include the limitations of recovering all of the actual source data, particularly attribution, due to the lack of representation or the limits of the runtime formats. Typical investments include formats such as OpenFlight, OpenSceneGraph, JointSAF/ModSAF CTDB and OneSAF OTF. This problem is exacerbated when the requirement is to correlate the legacy runtime with one or more new simulation formats which will be constructed from the result of the source data reuse process.

In summary, rationalizing source data is one of the most overlooked components of the modeling and simulation runtime generation process. Yet like most engineering solutions, the quality of any simulation is no better than the source data that is used to construct the runtime. The goal of this paper is to argue that significantly more attention needs to be paid to this process and that COTS software products developed specifically for addressing this problem are a necessary, but not sufficient, solution.

## AUTHOR BIOGRAPHIES

*David M. McKeown, President, TerraSim, Inc.*
David M. McKeown has been involved in environmental issues for modeling and simulation since the early '90s, beginning with his involvement with the DARPA SIMNET program. He has actively participated in technical conferences for IEEE, IMAGE Society, etc. and gave a well-received tutorial on environmental considerations for pattern of life (POL) at GameTech 2012. He is President of TerraSim, Inc and an Adjunct Research Professor in Computer Science at Carnegie Mellon University in Pittsburgh.

*William Starmer, Manager of Database Engineering Group, TerraSim, Inc.*
Mr. Starmer received a BA in Geography in 1994 from the University of Maryland, Baltimore. He has over fifteen years of experience in GIS and Remote Sensing. He developed and presented a well-received three hour specialist course entitled "Issues in Urban Database Construction", presented over three successive years (2002-2004) at the IMAGE Society Conference in Scottsdale, AZ. He is responsible for TerraSim's process architecture for creating correlated visual and constructive databases using TerraTools, including the generation of correlated OneSAF OTF and VBS2 databases.

*Siddharth Shankar, Software Engineer, TerraSim, Inc.*
Mr. Shankar received a MS in Computer Science with a specialization in Computer Graphics from University of Utah in December 2009. As a graduate student at Utah, he worked towards the development of a Large-Scale Parallel Visualization System - VisIt. Prior to that he received a BS in Information Technology from Cochin University (India) in May 2005, where he graduated in First Class with Distinction. Since joining TerraSim in 2010, he has contributed to the optimization of the TerraTools data processing pipeline, particularly in the area of imagery display and 3D rendering. He has also made significant improvements to RoadMAP, one of TerraSim's source data preparation products. He serves as the principal developer for our Havok Vision Engine exporter and TerraSim's Vision Engine preview application.

*Evan Klauda, GIS Analyst, TerraSim, Inc.*
Mr. Klauda received a BS in Geography with a specialization in Geographic Information Science from Penn State University in December 2011. Since joining TerraSim, he has contributed to the automated generation of large area databases, correlated across multiple runtimes and environments. He has contributed to the extensive documentation library that supports the TerraTools customer base as well as provided quality control for the TerraSim suite of products.