



USAF TRAINING SYSTEMS PRODUCT GROUP (TSPG)

COMMON DATASET STANDARD (CDS)

Version 1.0 – 1 DEC 06

Distribution Statement A: Approved for public release; distribution is unlimited. PA Case Number ASC 07-0023.

LIST OF EFFECTIVE PAGES

Dates of issue for original and changes are:

Version 1.0 1 DEC 06

Total number of pages in this document is 10 consisting of the following:

<u>Page No.</u>	<u>Change No.</u>	<u>Page No.</u>	<u>Change No.</u>	<u>Page No.</u>	<u>Change No.</u>
Title.....	V 1.0				
i thru iii.....	V 1.0				
1 to 6.....	V 1.0				

CHANGE HISTORY

[illegible]

Table of Contents

Paragraph	Title	Page
1.0	Purpose, Method, and Limitations	1
1.1	Concept of Operations Overview	1
1.2	Background	2
1.3	Common Dataset Formats	2
1.4	Common Dataset Source Data Preparation	2
1.4.1	Terrain Merge and Editing	2
1.4.2	Features and Models	3
1.4.3	Imagery	3
1.4.3.1	Orthorectification and Geopositioning	3
1.4.3.2	Contrast and Color Balancing	3
1.4.3.3	Colorization	4
1.4.3.4	Feathering	4
1.4.3.5	Imagery Delivery	4
1.4.4	Material Encoding	4
1.4.5	Metadata	4
1.4.6	Source Data Rights	5
1.5	Deliverables	5
1.5.1	Common Datasets	5
1.5.2	Source Data	5
1.5.3	Documentation	5
1.6	Verification and Validation	5
1.7	Changes and Extensions	5

TSPG Common Dataset Standard

1.0 Purpose, Method, and Limitations

The purpose of this standard is to identify a group of database formats commonly used by industry, and reference these formats in contracts as a required dataset output by vendors and deliverable to the Government for reuse by other Government programs. It is the Government's goal to select common formats which are already in-place in vendors' database generation systems, so no development will be required for compliance to this standard. This method focuses on open architecture interfaces in interim format files rather than mandating the run-time product to achieve its purpose. This approach does not favor any single source supplier of databases or database tools.

Since this standard references formats commonly used by Industry, the formats listed here are actually selected by Industry, and not by the Government. Close cooperation and communication between Government and Industry are required to keep this standard up-to-date and useful.

This standard is not intended to result in plug-and-play interoperable databases or exactly identical scenes between programs having different scene requirements or image generators, nor will it eliminate all costs and schedule associated with database generation. This standard is merely a tool to help improve database cost/schedule/performance and reduce correlation differences among USAF TSPG simulation programs.

In order to achieve sufficient scene correlation, some programs may require constraints to be placed on how databases are built, controls placed on database content, or mandate rules regarding how data bases are transformed into run-time scenes. This standard does not impose any such constraints, controls, or rules on any program.

This standard will be contractually invoked by programs that include the requirement to generate terrain databases ("producer programs"), as well as subsequent programs that can use common datasets generated by producer programs to help satisfy their database requirements ("consumer programs"). A "producer program" can also be a "consumer program."

1.1 Concept of Operations Overview

This standard will be referenced in select USAF TSPG simulator acquisition and/or sustainment solicitations, and subsequent contracts or delivery orders that include the requirement to build terrain databases for visual and sensor training applications. Producer programs will be required to export interim format database files and data types in commonly used open formats (defined herein), as well as original digital source materials, and provide those files and materials to the Government for purposes of re-use by consumer programs. The Government will make datasets available for consumer programs to use as a starting point to build their own databases.

Archive contents will be made available as-is, with each producer and consumer program's contractor retaining the responsibility for satisfying its respective run-time data base requirements. Each consumer program's contractor who uses and improves these files may, in turn, be required to export these improved files in the same commonly used open formats back to the Government for further use by future consumer programs. The goals of dataset reuse are to decrease database development cost, engineering redundancy, and database generation schedule, and provide a common foundation of database source materials to help improve correlation between simulators.

A more detailed description of the Concept of Operations (CONOPS) may be found in the complementary TSPG CDS CONOPS document.

1.2 Background

This standard relies heavily upon seminal work conducted by both the TSPG and NAVAIR/Patuxent River's Naval Aviation Simulation Master Plan, Portable Source Initiative (NPSI) to test and validate schedule and cost efficiencies of this method. The group of commercial dataset standards identified here are similar to those used by NPSI. Use of the same set of standards provides the potential of database re-use between USN and USAF programs.

1.3 Common Dataset Formats

The Common Dataset shall retain the maximum amount of non-proprietary, IG-independent, integrated content generated by the producer program, and shall be in the following commonly used formats:

Terrain:	DTED format or GeoTIFF 16-bit signed integer
Features:	ShapeFile format
Imagery:	GeoTIFF format: 8 bit per plane, unsigned integer pixel representation; or lossless JPEG2000
Models and Airfields:	OpenFlight with RGB/RGBa, TIFF, or JPEG 2000 textures

Color palettes shall be included for features, models, lights, and textures.

The Common Dataset shall be delivered at the highest, unfiltered level. For example, if the spatial resolution of the imagery is 1m and for whatever reason an individual program constructs a visual database in which the imagery is down sampled to 5 meters, then the required Common Dataset delivery of this imagery would be both the one meter resolution source imagery and the down sampled five meter dataset derived from it.

The USAF TSPG Common Dataset Standard focuses only on commercial standards that are already in-place in vendor Data Base Generation System (DBGS) tools and commonly used as interim file formats across Industry. Government-unique standards do not currently meet these criteria, and are therefore not included in this standard. Based upon their specific requirements, some programs may be required to follow other Government standards instead of the TSPG Common Dataset Standard.

1.4 Common Dataset Source Data Preparation

The preparation of source data accounts for a significant effort in the generation of a run-time geo-specific database. The following processes address these issues and can be performed using 3rd party tools but can be labor intensive. By requiring the following processes for standardization, the government can more readily reuse the product of an initial delivery. It is important to discriminate which processes should be done only once and which processes are best done on a per runtime database/per publication. The following processes should be done once and the data subsequently reused.

1.4.1 Terrain Merge & Editing

While the terrain data used in simulation and training is often DTED, it's not the only source. Sometimes other digital elevation sources are available and need to be combined with existing DTED. A database supporting aviation training in Alaska is one such example: terrain relief for areas of rapidly changing elevation (inlets or fjords) may not be well represented by level 1, and higher resolutions of DTED for a given region may not be available. Terrain merge and feathering shall be performed as required by each program and ensure that any changes will be captured. The result of this requirement shall be contained within the appropriate interchange format of this standard and registered in Lat/Long to the WGS-84 datum. Changes to terrain data may include

smoothing of sampling noise found within DTED grid posts, flattening the terrain for lakes and airfields, correlating terrain to rivers, incorporating cut and fill for roads, and adding peaks, pits, ridgelines, and valleys that fall outside of the DTED precisions and grid structure. When these alterations of the terrain result in a resolution not represented by DTED, the elevation data shall be exported in GeoTIFF format registered to Lat/Long in the WGS-84 datum as a 16-bit signed intensity map. The original, unaltered terrain source data shall be provided in addition to the modified terrain, to be used for reference or to support alternative smoothing/sampling schemes by subsequent users.

1.4.2 Features and Models

Vector products, such as DFAD and VMAP, designate point, linear and areal features. These vector representations typically do not sufficiently match imagery, and the work associated with correlating features to the imagery is labor-intensive. Corrections are made to source data in applications that map features with imagery, to include all modeled levels of detail and states. Terrain-feature-texture correlation shall be performed as required by each producer program. Features include light points, light faces, and light strings. Lights include rules for color, brightness, extinguishing coefficients, and size. Models include 2D and 3D moving and static models and may also include seasonal representations, damage states, and articulated parts. Models may include geotypical or geospecific phototexture and may include geo-accurate z-elevation values for features and models that are in contact with the terrain. Features, models, and all of their required characteristics shall be contained within the appropriate interchange format of this standard and (for static models) registered to Lat/Long in the WGS-84 datum.

1.4.3 Imagery

Imagery to be included in the Common Dataset may undergo numerous processes that add value to the original source imagery to satisfy program requirements. The intent is to ensure that the value added to imagery is captured by the producer program for reuse by consumer programs.

1.4.3.1 Orthorectification and Geopositioning

Source imagery shall be manipulated to convert it to a vertical viewing perspective, remove the perspective distortion common to lenses, compensate for earth curvature, and remove the distortion caused by changes in terrain elevation within the image. Orthorectification shall be performed, resulting in an image with map-like accuracy and scale. Although most Government furnished imagery already has these corrections made to it, the producer program may need to process other sources of imagery. Orthorectification shall be performed as required by each producer program, and the result of this process shall be contained within the appropriate interchange format of this standard. Geopositioning shall be performed, to take the corrected image and tie it to ground truth control points for precise positioning on the earth's surface. Geopositioning may also be performed to address differences in geometric accuracies, such as a road in one image that is not geometrically continuous with the same road in an adjoining or inset image and requires spatial readjustment. Geopositioning shall be performed as required by each producer program, and the result of this requirement shall be contained within the appropriate interchange format of this standard and registered to Lat/Long in the WGS-84 datum.

1.4.3.2 Contrast and Color Balancing

Contrast and color discontinuities are often observed in photographs taken at different times and can be a distraction in training, especially along the seams between adjoining images. Various methods are available to adjust and minimize these discontinuities. Contrast and color balancing shall be performed as required by each producer program and the result of this requirement shall be contained within the appropriate interchange format of this standard.

1.4.3.3 Colorization

Many forms of imagery source data may be available only in panchromatic/monochrome. Colorization of monochrome imagery shall be performed as required by each producer program and the result of this requirement shall be contained within the appropriate interchange format of this standard. If colorization is achieved through the merging of color source imagery to the monochrome, the monochrome and color source imagery (prior to merge) shall both be saved as separate datasets and shall be contained within the appropriate interchange format of this standard.

1.4.3.4 Feathering

When merging or mosaicking adjoining images, or when inserting high-resolution imagery into an area of lower resolution imagery, a smoothing algorithm for feathering boundaries is often used to avoid a sharp, high-contrast line where one image ends and another begins. If required by the producer program, the result of the feathering process shall be contained within the appropriate interchange format of this standard.

1.4.3.5 Imagery Delivery

All imagery (including the value-added processing described above) shall be delivered in one of two open, commercially standard image formats registered in Lat/Long on the WGS-84 datum:

GeoTIFF. This is a widely used, proven lossless imagery format. There is a file size limitation of 2Gb for GeoTIFF files. In the event an image exceeds this, the image shall be broken into several tiles under the size limit.

JPEG 2000. JPEG 2000 is an emerging open format with widespread Industry support, offers unlimited file size, and various ranges of lossy/lossless compression.

1.4.4. Material Encoding

Some producer programs, particularly those with sensor or Night Vision Goggle (NVG) applications may require material encoded textures applied to features and models and material encoding of overhead imagery at the texture element (texel) level. If required by the producer program, the result of these processes shall be contained within the appropriate interchange format(s) of this standard. Methods of mapping the encoding scheme into material categories shall be described within documentation at paragraph 1.5.3.

1.4.5 Metadata

Metadata is descriptive information about the data, such as source data rights, pedigree, projections, earth datum, feature capture criteria, accuracies, tessellation strategy, rules regarding multiple levels of detail for terrain, culture, models, and additional information regarding surface material types defined for sensor simulation. Metadata shall be generated as design documentation, as described in paragraph 1.5.3.

1.4.6 Source Data Rights

Source data rights and rights to its derivative products must be clearly defined prior to contracted database development. Individual producer programs may need to purchase extended data rights for some forms of data to allow them to be re-used by other programs. Alternatively, they may obtain pricing data for these extensions, to be purchased by consumer programs as the need arises. At a minimum, Government rights to source data and their derivative products shall be clearly defined in database design documentation, as described in paragraph 1.5.3.

1.5 Deliverables

The producer program shall deliver Common Datasets in the digital formats described above, and documentation of that data to the Government for purposes of re-use by consumer programs, ideally without restrictions. If restrictions exist, the producer program shall clearly describe those restrictions and the cost of eliminating them prior to contract award by the Government.

1.5.1 Common Datasets

All value added source data used to build the runtime database or for the purpose of database production shall be delivered in formats as referenced in section 1.3 and shall retain the products of value-adding conducted by the producer program. These delivery items consist of the vendor independent output processes described herein.

1.5.2 Source Data

Copies of original, non-restricted digital source data used to build models, features, or textures shall be delivered in their parent formats and media.

1.5.3 Documentation

The delivered documentation shall include a Data Base Design Document (DBDD) in the format required by each producer program. The DBDD may describe such information as level of detail switching distances, polygon densities, thinning algorithms, a description of the source data that was modified, the hierarchical structure of the models, etc. As a minimum it shall include:

- Security classification.
- Source data rights and restrictions.
- Description of processing steps, sampling intervals, image texel orientation, area, settings, vector/model/image dataset linkages, etc., utilized in the production of the run-time data. (This information is critical for the reconstruction of datasets and comparison with other representations of the same virtual space).
- Identification of 3rd party tools or readers used to prepare datasets, to include version number.
- A listing of the various types of raw source data used for database generation, and their areas of coverage.
- Metadata, with contents and transmittal format in contractor format. Transmittals should include material code definitions and mappings as well as any specific compression formats or color gamuts used.
- Header files with record structures.

1.6 Verification and Validation

The Government will verify that the delivered data includes the results of value-added processing by the producer program. It will be verified that the Common Datasets are in the required formats and that digital source materials are included. Producer program documentation will be inspected to ensure that the information required under paragraph 1.5.3 is included. This task is anticipated

to be conducted by the TSPG team managing the producer program, as part of that program's acceptance testing process and in accordance with that program's Contractor Data Requirements List.

1.7 Changes and Extensions

Changes to this standard are anticipated based upon periodic meetings between Government and Industry and recommendations resulting from those meetings. Also, any changes made to the Navy NPSI set of standards or other similar DoD data base standardization initiatives will be considered for incorporation.

Individual program customers, acquisition organizations, or their vendors may recommend additional data types that are not explicitly referenced in this document, but could be useful for subsequent programs. One such category of new data types is the emerging use of 2D and 3D vector graphics markup languages using XML syntax and enhanced Application Programmers Interfaces to encode terrain and model information rather than the current grid/raster or polygon-based methods. When or if these data types and methods (or any others) become commonly and consistently used by industry to contain data base information, they will be considered for inclusion in this standard.