NC STATE UNIVERSITY

Enterprise GIS

GIS Data Collection Standards

and

Geospatial Data Standards

January, 2020

Contents

Section 1

1.1 GIS Data Collection Standards

Spatial data collected by internal staff and external vendors/contractors with the purpose of being used in the enterprise geographic information system (**GIS**) must adhere to the NC State University Office of the University Architect's (**OUA**) Geospatial Data Standards. However, the following standards are not to be used in lieu of minimum requirements set by the NC Board of Examiners for Engineers & Land Surveyors (**NCBELS**) when data collection activities fall within the definition of the *"practice of land surveying"* as defined by GS89C-3.7.

The use of the Global Positioning System (**GPS**)/Global Navigation Satellite Systems (**GNSS**) satellites orbiting above the Earth and the land-based Terrestrial Positioning Systems (**TPS**) for accurately locating real world features and attributes have become widely accepted methods for collecting GIS data. GPS/GNSS and TPS receivers allow users to collect the locations in real world coordinates. For the purposes of this document, GNSS and GPS should be considered to mean the same thing. Please see expanded information describing these collection methods in the Appendix.

GIS Data Collection Standards does not define threshold accuracy values or minimum accuracy required for a given feature class. That information will be determined by the GIS Working Group on a project by project basis. If another collection method besides GPS/GNSS or TPS is used, please contact OUA: group-oua-gis@ncsu.edu

1.2 GPS/GNSS Data Collection and Delivery Standards

Any GPS receiver used to collect data for OUA must be of Mapping Grade quality or better, and must:

- Routinely achieve 1 meter or better horizontal accuracy, using either real time or post processed differential corrections.
- Operate in a 3D mode, where the receiver requires signals from a minimum of five satellites to determine "fixed" position.
- When mapping real world features, the receiver must be able to collect and store multiple positions; the minimum number depending on the quality required for that particular feature.
- Must have enough internal storage capacity for a typical day's worth of data collection without the need to transfer to a PC.

- Configurable for user settings like geometric dilution of precision (**GDOP**), positional dilution of precision (**PDOP**), elevation mask, and logging rate.
- Be able to utilize existing GIS data for base map, and export in a compatible format for enterprise GIS from unit or desktop software (geodatabase or shapefile).

1.3 Field Data Collection Parameters

Field staff must have a thorough understanding of GPS basic concepts, and receiver operation. How a receiver's critical parameter settings affect data collection must be very well understood. The staff must also have familiarity with the types of features that are to be located, and must be able to recognize/interpret features in the field. To achieve target accuracy all collected GPS data must be differentially corrected, either in real time or in a post process step.

OUA requires checking into at least one "known" point before beginning data collection, and after. This field procedure helps ensure the correct reference base station is being used, and that obstructions and satellite positions are optimal for accuracy.

Position Mode: All position fixes must be determined with 5 or more satellites. Manual 3D or overdetermined 3D (5 satellites minimum) modes are acceptable. 2D fixes (using only 3 satellites) are not acceptable. 3D positions generated from 2D fixes supplemented with user entered elevations are also not acceptable.

The following other minimum settings or conditions must also be maintained for all collected locations:

- Elevation Mask: 15 degrees above the horizon.
- PDOP Mask: Max PDOP = 3-4
- Minimum Positions: Set to achieve minimum level of accuracy accepted (1 meter), unless granted a different requirement by OUA.
- Logging Intervals (Epochs): Intervals for features will be at least 3-5 seconds, with line and area feature accuracies depending on the velocity at which the receiver is traveling and the nature of the feature and obstructions present.
- 3D Coordinate Quality: This parameter setting will be set to allow the logging of coordinate values along with position fixes for data collected.

1.4 Data Dictionaries

A data dictionary is a list of attributes to be collected, the field characteristics for each attribute and acceptable values (if appropriate). If a data dictionary exists for a feature, OUA requires that it be used for data collection. This will ensure that data collected will be compatible with existing data in the enterprise GIS. These may be requested from OUA by contacting group-oua-gis@ncsu.edu .

1.5 Processing of GPS Field Data

All GPS data collected for OUA must either be captured with the RTK method using NCGS RTN Real Time Kinematic or the NC GNSS Real Time Network, or post processed using desktop software before the data can be used in the enterprise GIS. The GNSS processing software must be able to download GNSS data files from the GNSS receiver, and perform differential corrections. In addition, it must allow exporting the corrected data to an OUA compatible format (ESRI File Geodatabase, or Shapefile), in the appropriate state plane coordinate system. Refer to the section below for specific information on the required coordinate system.

1.6 Reference Networks and GPS Base Stations for Post Processing

For Reference Network Corrections, several resources exist in North Carolina such as the NC GNSS Real Time Network (account with NCGS required).

For post processed differential corrections, several resources exist for GPS base station data in North Carolina, such as National Geodetic Survey's Online Positioning User Service (OPUS).

1.7 Output to GIS

OUA uses the ESRI software platform and requires GIS data submittals to be in Esri File Geodatabase (.gdb), or Shapefile format (.shp). The GNSS software must allow for exporting to one of these formats.

In addition to feature coordinate and field entered attribute data, the GNSS software used should be capable of generating attribute information for exported features. This information will include data about the quality of the GNSS position fixes that were used to generate the features. At a minimum, the following attributes must be produced for exported features, and if it is not automatically generated by the GNSS software it must be manually entered:

Point Features	**Line and Polygon Features**
Maximum PDOP	Maximum PDOP
Receiver type	Receiver type
Date of collection	Date of collection
Time of collection	Time of collection
Data file name	Data file name

Total positions	Total positions
Positional Accuracy 2D or 3D	Worst horizontal precision
Orthometric Height	Average vertical precision
Vertical Accuracy	Worst vertical precision
Antenna Height	Antenna Height

1.8 Elevation Data

If the collection of accurate elevation data is required as part of a project, it will be referenced to the North American Vertical Datum of 1988 (NAVD 88). Elevations must be generated as orthometric heights (relative to mean sea level) determined using GEOID12 model (Continental US).

Data must be collected with a Survey grade GNSS receiver when elevational or vertical data is required as part of a project. To achieve reliable vertical accuracies, regardless of the type of GNSS receiver being used, the antenna must be mounted on a pole or similar device which allows the user to measure, maintain, and record a consistent antenna height.

1.9 Naming Convention

It is important that GNSS data and attribute field names delivered to OUA follow the same structure already created for existing data (unless it's a new feature class). Not all of the following rules may apply to all datasets, but it's important to understand the geospatial naming convention already in place to resolve any conflict that might arise on import of delivered data into the enterprise GIS or added.

- No leading numbers or special characters.
- Do not include spaces, dashes, underscores or other special characters
- Do not use prefix or suffix for data type
- Do not use geometry type as suffix
- Avoid using reserved words
- Limit names to 10 characters or less (Shapefile only)
- Always provide alias names for fields

1.10 Deliverables

If GNSS data collection work is being performed by an outside agency, final deliverables shall meet the minimum requirements as set forth by NCBELS, including:

- All GNSS field data files, both uncorrected and corrected versions, must be submitted. If field data was collected in real time differential mode, then there will not be uncorrected files, and only the real time corrected files are necessary. If edits are made to corrected files (i.e., fixes deleted or offset), copies of both edited and unedited are to be submitted.
- All GNSS to GIS export files, using North Carolina State Plane Coordinates (**NCSPC**), in the NAD 83 (2011) horizontal geodetic datum, in US survey feet units, in Esri File Geodatabase, or Shapefile.
- All GNSS processing log files pertaining to post process differential correction and GIS export (if produced by the GNSS processing software).
- GPS Data dictionary files, defined for project attribute storage.
- Metadata (ESRI item description as a minimum; FGDC; ISO)

1.11 Project Report

The contractor or vendor must submit a project report that includes the following information.

- An introduction describing the project. Include the project name, the names of NC State University departments involved/supported, the purpose and goals of the project, the project's study area, and data collection (including accuracy) requirements.*
- Reference Network or GPS base stations used for the project. If local base stations (stations other than Reference Network) were used, the setup procedure must be described in detail, along with the operation, collection parameter settings, and what steps were used to establish the reference position.

All of the files for the Deliverables and the Project Report should be in a compressed format and be organized into a logical directory structure. For example, the files could be organized by date of data collection, and then into subdirectories for Data and Export. Uncorrected, corrected field data files, post process differential correction log files, and data dictionary files would reside in the Data subdirectory. GIS export files and associated export log files would reside in the Export subdirectory.

* An agency may be excused from including report items indicated with an asterisk with consent of OUA prior to delivery.

Section 2

2.1 Geospatial Data Standards

There are multiple types (and quality) of data sets created and used by OUA within the enterprise GIS. To facilitate sharing, integration, and compatibility, OUA requires all data generated for, and managed by the enterprise to adhere to basic standards as follows. All geospatial data must meet or exceed the documented standards regardless of scale. Testing against existing base data combined with statistics (about how and what data was collected) ultimately determines the accuracy of the data in question. This process will confirm relative accuracy of new or unintegrated geospatial data, and how compatible it is with the enterprise system.

2.2 Spatial Reference

Digital data provided to or produced for OUA needs to be in the North American Datum 1983/2011 (NAD83/2011) horizontal geodetic datum and referenced in North Carolina State Plane Coordinates (NCSPC); and in the North American Vertical Datum of 1988 (NAVD 88). The NCSPC is the official survey base for the State of North Carolina.

The specifics of the referencing system requirements are as follows:

- Projection: Lambert Conformal Conic
- Geographic Coordinate System: GCS_NAD_1983_2011
- FIPS Zone: 3200
- False Easting: 2000000.0 (as per ESRI)
- False Northing: 0.0 (as per ESRI)
- Central Meridian: -79.0
- Scale Factor:
- Latitude of Origin: 33.75
- Linear Unit: US Foot
- Angular Unit: Degree
- Horizontal Datum: North American Datum of 2011
- Vertical Datum: NAVD 88
- Spheroid: GRS1980
- Semi Major Axis: 6378137.0
- Semi Minor Axis: 6356752.314140356
- Inverse Flattening: 298.257222101

OUA requires that all coordinate values be reported in units of US survey feet, with the units clearly defined in the attached metadata. This requirement applies to all ground survey data as well, which must be submitted with equivalent NCSPC values for all points, if the points were not originally captured in NCSPC.

2.3 Data Creation and Collection

There are many techniques that can be used to create geospatial data, which in turn can be submitted to OUA. The two most common are:

A). Compiling data from base sources like digital imagery, using interactive editing, digitizing overlays which are then scanned and georeferenced, vector based analysis, classification, etc.

B). Data captured from the field using measurement technology instruments like TPS, GPS/GNSS, Digital Levels, and Lidar.

Base Sources:

- Any recent digital aerial photography at large or small scale must meet or exceed the NSSDA (Standard for Spatial Data Accuracy) testing methodology. Accuracies for historical digital imagery (prior to 2002) can be reported using the NMAS (National Map Accuracy Standards).
- Non-digital (hard copy) base maps: Since some data needed to represent historical conditions may have to be generated from non-digital sources, those sources must meet a NMAS threshold value for the appropriate scale.
- User geo-referenced digital imagery: Geo-referencing is the process of defining a coordinate system and a projection for an undefined data source, such as a historic map or image. In those cases where the data submitted to OUA was generated from a source geo-referenced by the data provider, the source material should be identified and its accuracy characteristics described, along with a full description of the geo-referencing process used by the data provider. Control point files with a report about accuracy, errors, and procedures are also to be provided.
- Vector data sets: Data submitted to OUA may be based on existing vector data sets. In these cases, a full description of the accuracy of the base vector data sets (including the accuracy of the source layer) used to create the data needs to be reported.
- In cases where data was created from base sources not referenced to NCSPC, all data will be projected to NCSPC before submittal to OUA.

2.4 Accuracy Requirements

Data submitted to OUA must be accompanied by a full description of the processing used to create said data, and of its accuracy. This description should include information on both the positional accuracy (horizontal and vertical), and the accuracy of any attribute information captured.

Base information used to create geospatial data must at least meet the NMAS threshold accuracy standards as listed in the following table:

Table 2.4 National Map Accuracy Standards

Scale	NMAS Accuracy (feet)	NSSDA Accuracy (feet)	NMAS Accuracy (meters)	NSSDA Accuracy (meters)
Large scale	1/30 inch (map)			
1:1,200	3.3	3.8	1.0	1.2
1:2,400	6.7	7.7	2.0	2.3
1:6,000	16.7	19	5.1	5.8
1:12,000	33.3	38	10.1	12
Small scale	1/50 inch (map)			
1:24,000	40	46	12.2	14
1:63,360	106	120	32.3	37

2.5 Supported Data Formats

OUA's enterprise system is built using ESRI software and there are several acceptable formats for data to be submitted. These formats are listed below in order of preference, with any specific requirements applicable to each:

1). **File Geodatabase**:
Data developed and submitted to OUA shall be in a compatible version of ESRI's file geodatabase (10.5.1 version as of 1/31/2020, contact group-oua-gis@ncsu.edu for verification). All data submitted must be topologically correct.

Geodatabases will adhere to at least the following standards:

- All feature classes included in the geodatabase will exist in one or more feature data sets
- The XY coordinate system for all feature datasets and feature classes will be in NCSPC
- Z coordinate system will be NAVD_1988
- The XY tolerance will be at least 0.01 ft. (A closer tolerance may be used where the accuracy of the data supports it, such as that collected with a survey grade GPS)
- The XY resolution will be at least 0.01 ft.
- Topologies will be created for all feature datasets and feature classes and all topologies will be checked before being submitted to OUA. Topologies may not have errors.
- For topologies that involve more than one layer, the most accurate layer will be given the highest rank

The minimum topology rules are:

- Features will not be duplicated
- Coincident boundaries will be corrected within a feature dataset (features that share boundaries with features in other feature classes in the dataset)
- Linear features will not overlap; i.e., all line intersections will require a node
- Linear features will maintain correct arc directionality for any data set with flow directions.
- Linear features will not have pseudo-nodes unless they are required to maintain a change in arc attribution
- Polygons must close
- Polygons will have no overshoots or dangles
- Polygons will not overlap
- Polygons sharing edges will not have gaps

Topologies should be submitted as part of the geodatabase delivered to OUA so that completeness and accuracy can be easily verified. In some cases, OUA GIS may develop project specific geodatabase templates for data submittal consistency. Users should investigate with the department whether or not geodatabase templates have been created and posted for download before developing a geodatabase for a project specific data submittal.

2). **Shapefiles:**

All shapefile data sets must include at a minimum the following files:

- .shp (the file that stores the geometry)
- .shx (the file that stores the feature geometry index)
- .dbf (the file that stores the feature attribute information)
- .prj (the file that stores the coordinate information)

When applicable, the following files should also be submitted:

- .sbx, .sbn (the files that store the spatial index of the features)
- .ain, .aih (the files that store the attribute index of active fields in the attribute table)

Shapefiles must be created so that the following basic topology rules are not violated:

- Features will not be duplicated
- Linear features will not overlap; i.e., all line intersections will require a node
- Linear features will maintain correct arc directionality for any data set with flow directions
- Linear features will not have pseudo-nodes unless they are required to maintain a change in arc attribution

- Polygons must close
- Polygons will have no overshoots or dangles
- Polygons will not overlap or self-intersect
- Polygons sharing edges will not have gaps
- Polygons will have one and only one label point

3). **Raster Data**

Raster data sets that encompass imagery or elevation data are less likely to be submitted to OUA than vector data (points, lines, polygons). However, in those cases where submissions of raster are required, the following data formats will be accepted.

- ArcInfo grids (integer or floating point)
- Triangular Irregular Networks (TINs)
- MrSid (version 2, 3 or 4) with world file
- Tiff, Geotiff, with world file
- Jpeg, jpeg2000 with world file
- ERDAS Imagine
- ASCII

All raster data sets must contain projection information, and projections will be defined with the following parameters:

- Projection: North Carolina State Plane
- Units: Feet
- Datum: NAD83

All raster data submitted to OUA must meet minimum positional accuracy standards as described in Section 2.4.

4). **CAD**

While there are many different CAD file formats, there are two major formats which will be accepted by OUA:

- AutoCAD DWG (AutoCAD 2017, Format Version 2013 and higher)
- AutoCAD DXF (AutoCAD 2017, Format Version 2013 and higher)

All CAD data submitted will comply with the following minimum specifications:

For all formats, the CAD files reference is NCSPC US Feet, NAD83. Unreferenced files will not be accepted. All data will be exported using a 16 decimal places option, so that double precision accuracy will be maintained. In addition, a text file listing individual layer names and descriptions shall be submitted with each CAD data set.

- All CAD files regardless of software version used to create the files will follow the structure outlined in the NC State mapping drawing template, which includes NC State standard layers, linetypes, and block symbols. The current version can be downloaded at: http://www.ncsu.edu/facilities/files/NCSU_CIV-SRV_TEMPLATE.dwg
- Annotation for each layer shall be placed in separate annotation layers

At least two separate digital files are required for each submission:

- One file will be a digital file in format containing the full survey drawing. This drawing must be created at its real North Carolina State Plane Coordinates, NAD83, NORTH position and the view shall be un-rotated from the coordinate system so that the NCSPC NORTH points are vertical in the screen.
- One file will contain all objects on the correct layer. Meaning, multiple features will be included on the same layer (ie all manhole points on manhole layer, all property lines on property line layer).
- Specific requirements may be requested on a project by project basis.

5). Other Data formats

Other data formats may be accepted by OUA based on project requirements. It is strongly suggested that those submitting data first check with OUA (group-oua-gis@ncsu.edu) to verify acceptable formats.

2.6 GIS Data Matrix

Internally, existing geospatial data editing will utilize a GIS Data Matrix document for rules/relationships within the enterprise system. Please contact OUA (group-oua-gis@ncsu.edu) to access and coordinate with this important document. For new features created or submittals, the following dataset and attribute naming conventions are applicable:

Dataset naming conventions

- Dataset names will contain only alphanumeric characters (i.e. letters, numbers)

- Dataset names will start with a letter
- Dataset names will be entirely in lowercase
- No spaces, dashes, or special characters other than an underscore will be used
- Dataset names will be 10 characters or less; common abbreviations should be used where applicable

Attribute field naming conventions

- Attribute field names will contain only alphanumeric characters (letters and numbers) and underscores
- Attribute field must start with a letter
- No spaces, dashes or special characters other than an underscore will be used
- Attribute field names will be 10 characters or less to avoid data conversion issues with truncation

2.7 Metadata for GIS and GPS Data

GIS and GPS data must conform to the Esri Item Description model at a minimum. This includes the following:

- Thumbnail snapshot of data
- Summary explaining general purpose of data
- Description which provides a detailed explanation of the purpose, data collection methods, expected maintenance schedule if any, expected accuracy and other pertinent information.
- Credits indicating what organization, department and persons created and/or maintain the data
- Use limitations or disclaimers
- Spatial extent

FGDC and ISO standards are also accepted as long as they also contain the required information listed at a minimum.

Additionally, avoid using fields in the database to store metadata about the feature class, e.g. a Date Loaded field. Such information is required in the metadata and therefore superfluous in the attribute table. Only exceptions are to either capture row specific metadata, e.g., Modified By, where each record may have different values, or where the origin source of the individual features may vary and need to be tracked. In the second case, the metadata should also document the fact that there are multiple data origins.

Appendix

Please contact OUA (group-oua-gis@ncsu.edu) for a complete list of information sources consulted and used in the creation of this document.

GPS/GNSS Concepts

GPS is the United States constellation of satellites, which was the first to become usable. Since the GPS constellation was deployed, several other satellite constellations have also been deployed including the Russian Glonass, European Union Galileo and Chinese Beidou. Many newer model receivers can be configured to receive signals from one or more of these constellations. This can increase accuracy and reduce the time needed to occupy a location. All of these constellations are now referred to collectively as Global Navigation Satellite Systems or GNSS.

The vertical accuracy of any GNSS receiver is typically equal to one and a half to two times its horizontal accuracy. To achieve reliable vertical accuracies, regardless of the type of GNSS receiver being used, the antenna must be mounted on a pole or similar device which allows the user to measure, maintain, and record a consistent antenna height.

To ensure that the appropriate type of GNSS receiver is matched to the mapping application, an understanding of receiver capabilities and limitations is required. For most applications, there are three types of GNSS receivers:

- Recreational Grade **(Low Accuracy)**

These units have a typical average horizontal accuracy of fifteen to sixty feet, with limited to no information about the quality of position displayed or collected on the unit. They normally do not have the ability to "post-process" field data that is collected or utilize real time corrections for improving positional accuracy. Recreational GNSS receivers can only be used to navigate to a general area. Smartphones and tablets are a good example of recreational GNSS.

- Mapping/GIS Grade **(Medium Accuracy)**

These units have a typical average horizontal accuracy that ranges from four inches to fifteen feet. They also have the ability to log raw GNSS data (allowing desktop software to “Post Process” and improve positional accuracy dramatically). This category of GNSS receiver also has the ability to communicate with a reference network (Real Time Corrections), store attributes of features, use a data dictionary, and transfer data from the GNSS device to a PC in a GIS compatible format. Units such as a Trimble Geo or Leica Zeno 20 are good examples of this type of GNSS receiver.

There are also several external smart antennas on the market which can be paired with other devices such as tablets or smartphones. This will increase the accuracy of those units turning them into a GIS or higher grade receiver.

- Survey Grade **(High Accuracy)**

These units can achieve sub-centimeter level accuracy, and are traditionally used by land surveyors for projects like boundary, topographic, or geodetic surveys. They have the ability to both post process data collected at a later date, or connect to a Reference Network for real time corrections.