

GNIS: Geographic Nomenclature for the Efficacy of Geospatial Data

D.N.D. Hettiarachchi

Survey Department,

No. 150, Bernard Soyza Mawatha, Narahenpita, Colombo 05, Sri Lanka,

hettiarachchidnd@gmail.com

Keywords: GNIS, Geographical names, geospatial, standardization

Abstract:

Many sectors related with geospatial technologies, such as Remote Sensing, Photogrammetry, Geographic Information Systems, Global Navigation Satellite System, data modelling, algorithms, etc., mainly provide geospatial data as well as results generated through various analysing methods to the end users. Generally, a coordinate system is assigned to these data for positioning, scaling, etc. However, end users, including normal public, need positioning or tagging of geospatial data with Geographical Names for easy referring. Generally, geographical names available in other data sources such as topographic maps, administrative reports, etc., are used with the geospatial data sets without any validation. Although geographical names play a significant role in applications of geospatial data, limited attention has been paid to maintain and use standardized geographical names. Sri Lanka has initialized establishment of Geographic Names Information System (GNIS) to address this issue. In this context, the Government of Sri Lanka has appointed a Committee for Standardization of Geographical Names (CSGN).

Sri Lanka consists of 330 Divisional Secretariat divisions as administrative areas. CSGN selected Grama Niladhari Divisions (GND) in “Beruwala” Divisional Secretariat (DS) division in Western Province as the pilot area. Guidelines and principles defined by the CSGN for collection and standardization of geographical names were tested in the pilot area. Results revealed that

- a. about 42% of the geographical names, currently use by the public, are not recorded in any of the prevailing data sources including topographical maps, administrative reports, etc.
- b. about 20% of the geographical names, available in the prevailing data sources, are not in use, incorrect, having incorrect positions or misspellings, etc.

Therefore, a geographic nomenclature is vital for producing intended results from geospatial data. Once GNIS is finalized for the entire Sri Lanka, it will be a free online data source for all interested sectors.

1. Introduction

The Geographical name is the name applied to a feature on Earth (UNGEEN, 2002). Geographical features are mainly natural and man-made features of Earth or any other planet. Geospatial data are directly connected and fitting with geographical features. Therefore, geographical names are used with almost all type of geospatial data for easy referring and understanding. Use of accurate geographical names will enhance the validity of geospatial data. Unfortunately, most of the geospatial data users and producers don't pay attention to check accuracy level of geographical names before using them. The objective of this study is to identify accuracy level of geographical names available in various data sources which are used by geospatial data users and producers in Sri Lanka, the introduction

of geographical names standardization procedures, and discuss requirement of GNIS which will function as Geographic Nomenclature for the efficacy of Geospatial Data.

2. Methodology

United Nations has introduced a program for Standardization of geographical names. The word standardization, as applied to geographical names/ toponyms, is defined by the Group of Experts as:

- a. The establishment, by an appropriate authority, of a specific set of standards or norms, for example, for the uniform rendering of toponyms;
- b. Rendering an item such as a toponym in accordance with such norms (UNGEGN, 2006)

Standardized geographical names can be considered as the most accurate and officially acceptable names for all kind of applications. For checking the accuracy of existing geographical names which are available in various data sources such as topographical maps, administrative reports, etc., can be compared with standardized geographical names. Completeness and accuracy checking of prevailing data sources were done as follows.

- i. What percentage of geographical names, currently use by the public, are recorded in the prevailing data sources such as topographical maps, administrative reports, etc., (completeness of prevailing data sources).
- ii. What percentage of geographical names, available in the prevailing data sources, are not in use, incorrect, having incorrect positions or misspellings, etc., (accuracy of prevailing data sources).

A suitable system, GNIS, will be identified for the dissemination of standardized geographical names for free use by all interested parties. Sri Lanka consists with 330 Divisional Secretariat Divisions. Beruwala Divisional secretariat division located in the Western Province of Sri Lanka was selected for this study.

2.1 Authority for Standardization of Geographical Names

The Cabinet of Ministers of Sri Lanka government has appointed a committee for standardization of geographical names (CSGN) which functions as appropriate authority. CSGN consists of 16 members representing different subject areas such as mapping, archaeology, linguistics, geography, history, cultural, postal, home affairs, agrarian, land, and ocean geography.

2.2 Standardization Procedure

Establishing guiding principles, policies and procedures are the most important formalized courses of action in the standardization process. Principles constitute the fundamental doctrines used for guiding the national standardization, encompassing general adherence to local usage, use of a particular script, and areas of responsibility. Policies are rules covering specific details devised to deal with problems and the means of implementing standardization. Formalized policies may include rules covering name changes, treatment of derogatory names commemorating living persons, name duplication, and the use of minority-language names. Principles are seldom changed but policies may be changed over time to meet new situations or to address unforeseen events (UNGEGN, 2006). Standardization process identified by the CSGN is shown in the figure-1 below.

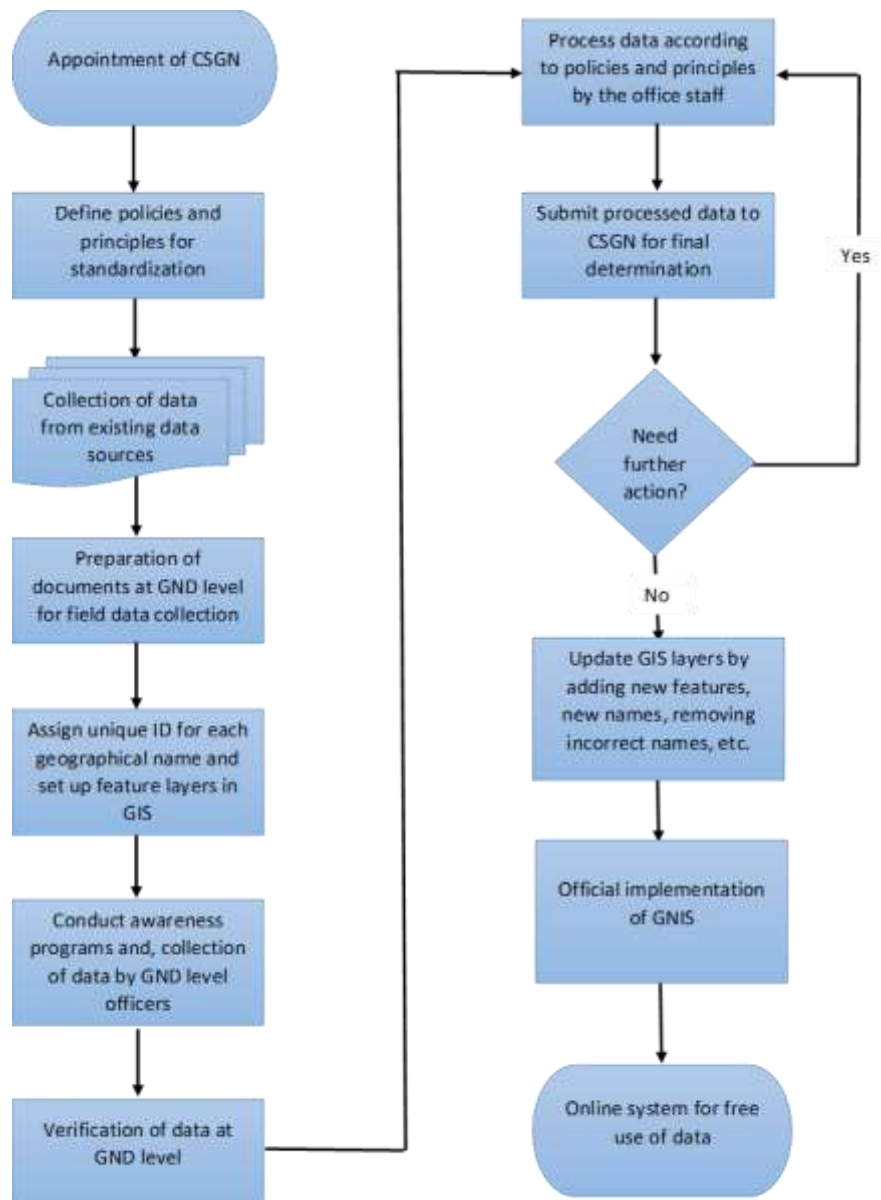


Figure-1: Geographical names standardization process

2.3 Data Collection

According to standardization process identified by the CSGN, it was decided to define geographical feature groups for easy management and processing of geographical names as there are unlimited geographical features on Earth. CSGN identified following groups which are considered vital in many areas such as GIS, mapping, administration, communication, industry, commerce, legal, science, education, internet, etc.

- a. Place names (town, village, administrative divisions, etc.)
- b. Archeological and religious places (temple, Mosque, kovil, archeological site, sculpture, etc.)
- c. Water sources and related places (stream, river, tank, lagoon, waterfall, harbor, etc.)
- d. Roads and Junctions
- e. Natural places (Mountain, valley, plane, cave, etc.)
- f. Important features other than above

For a given feature, the authority should decide most appropriate name from among a number of variant names. However, a single feature may have more than one standardized name (UNGE, 2002). Therefore, the collection of geographical feature names from the field is one of the important events in the standardization process to identify current names used by the public, officials, etc. CSGN designed instruction leaflet, forms and maps for field data collection after studying all levels of data collection, data analysing, data storage, etc. In addition, CSGN decided to collect names available in other data sources. Map making, land surveying, and human history have very close association for a long period. The resemblance to this, Sri Lanka also has a long history way back to 1690 in mapping and surveying activities. These activities were restructured with the establishment of the Survey Department as first Department in Sri Lanka on 2nd August 1800. Therefore, Survey Department has a wide range of collection of maps and statutory survey plans since 1800. Out of them, “one inch to one mile”, “1:50,000” and “1:10,000” scale topographic map series and large scale statutory survey plans contains geographical names which are valuable data sources that can support standardization process of geographical names (Hettiarachchi, 2016). Other organizations such as Divisional secretariats, Archaeological department, Postal Department, Elections Department, etc., are also having huge collection of geographical names related with their information.

Total land area of Sri Lanka is 65,525 km², and it is covered with 14,022 Grama Niladhari Divisions (GND). As GND is the smallest administrative area, it was decided to collect Geographical Names at GND level with the support of officers in charge of GNDs. Awareness program for the GND level officers are conducted at Divisional level, and required documents (forms, maps, etc.,) are provided to them for data collection. Data collected by GND level officers are checked and cross-examined with already available data sets and details are more clarified with the support of them. All collected data are marked on the maps of GND for easy referencing, consistency and maintenance of a GIS.

2.4 Data Processing and Standardization

Office staff of Geographical Names branch of Survey Department follows the guidelines and principles defined by the CSGN. Data collected through GND officers and other means are used for processing. All standardized data are submitted to the CSGN for their views and approval. All problematic names are also submitted to the CSGN for their final determination. GIS compatible data layers are updated using the information marked on maps and comparing them with satellite images. Road names, are checked with local authorities as local authorities have a mandate for naming roads. Standardized geographical names are shown in GIS data layers. Figure-2 shows part of the Beruwala Divisional Secretariat area.



Figure-2: Standardized geographical names in GIS layers

3 Results and Discussion

Results were analysed as follows after completion of standardization process of Beruwala Divisional Secretariat area.

Completeness of prevailing data sources:

| | |
|---|-----------|
| Number of geographical names finalized after standardization process | 1516 |
| Number of new geographical names (not available in any of the prevailing data sources) identified after standardization process | 639 (42%) |

Accuracy of prevailing data sources:

| | |
|--|-----------|
| Number of geographical names available in the prevailing data sources (Topographical maps, official records, etc.) | 1105 |
| Number of incorrect (not in use, having incorrect positions or misspellings, etc.) geographical names available in the prevailing data sources | 228 (20%) |

Results were compared with statistics of another Divisional Secretariat area (Kalutara) and found that completeness and accuracy were much more likely.

4 Conclusion

All types of spatial data acquiring technologies such as satellite remote sensing, aerial photography, LiDAR surveys, etc., generate valuable and up to date geospatial data sets which can be utilized in many applications by various type professionals, planners as well as normal public. However, to give a real meaning for those data sets, they should be tagged with location-based information, mainly with geographical names. Therefore, maintenance of standardized, officially accepted, a complete and accurate database of geographical names is indispensable. Studies reveal that geographical names available in prevailing data sources are not verified and maintained a desired level of quality. Therefore, it is important to maintain a database to provide standardized geographical names for interested parties. Users should be able to access GNIS through web-based facilities for making online queries based on feature type, location, etc. It also helps to search for historical names, variant names, and preserving and promoting a nations toponymic heritage etc. GNIS will help for collecting geographical names from authorized officials and the general public. Such a GNIS will support the efficient and effective use of geospatial data in various applications.

References:

- Hettiarachchi, D.N.D., 2016, Information Bulletin No. 50, United Nations Statistics Division, United Nations Publication, New York, 47 p.
- UNGEGN, 2002, Glossary of Terms for the Standardization of Geographical Names, United Nations Publication, New York, 18, 19 p.
- UNGEGN, 2006, Manual for the national standardization of geographical names, United Nations Publication, New York, 9-10, 33 p.