QUANTIFYING ABSOLUTE AND RELATIVE ERRORS OF HIGH RESOLUTION GEOSPATIAL DATA

Bishwa Acharya, President
Earth Mapping International
P O Box 491474, Lawrenceville, GA 30049
bishwa@earthmapping.com

John Rutan, Senior Associate
Earth Mapping International
Jrutan@mindspring.com

ABSTRACT

Absolute, intermediate, and relative errors are intrinsic and inherited values of any geospatial data that need to be redefined and quantified in the present context of ever-evolving sensors, platforms, and the Internet of Things (IoT). The absolute error being so small, many times it is assumed to be zero and thus not included in the error computation. Therefore, accuracies cited in the metadata may include only the relative accuracies of geospatial data. This paper will revisit the current practice of defining accuracy standards of geospatial data captured utilizing: Global Navigation Satellite Systems (GNSS), remote sensing satellite systems, airborne systems, and terrestrial digital sensors. Evolution of geospatial technology along with accuracy standards and specifications in the last twenty-seven years will be succinctly explained prior to proposing an improvement in defining and quantifying the absolute, intermediate and relative errors. The sources of absolute errors such as datums, map projections, GNSS, and conventional surveys will be analyzed and quantified. The most mistaken sources of errors: measurement units and truncation errors, geoid-ellipsoid separation for measurement of the vertical dimension and time measurements will be discussed. Optimal and cost effective approach for quantifying relative errors will be proposed. Limiting errors of sensors, platforms, and flying heights converted into ground sampling distance (GSD) will be explained. The absolute, intermediate, and relative errors will be merged to quantify the total error budgets and computation of root-mean-square errors. The geospatial error budgets for various scales of mapping will be suggested. A new standard for estimating error budgets based on GSD will be recommended.

Key Words: absolute error, intermediate error, relative error, GNSS, IoT

HISTORICAL BACKGROUND & INTRODUCTION

Understanding the Earth within and in relation to other planets and stars has fascinated humans since the beginning of the civilization. Understanding the necessities or practical aspects of navigation, retracing or relocating, warfare, and land ownerships (Earth surface) has been a human necessity since the beginning. The earliest history of map making or surveying which is a part of present-day geospatial technology goes back to about 3000 years ago, e.g., Babylonian clay tablets depicting the Earth. In Egypt, surveying was called “Stretching a Rope”, and a Surveyor was known as “Rope Stretcher”. Standards were established for measuring distance, angle, and height to create maps so that the surveyors (rope stretcher) could relocate the boundary to satisfy the owners of the parcels after flooding (Paulson 2005). Hand and pacing were also used as a basis for measuring length in the early civilization; ropes or chains were the main equipment to measure distance. We can safely assume that their intent was to be as accurate as possible.

In Vedic scripture, the universe is a cosmic egg like shape which expanded from a point called “Bindu”. It is like a living entity which expands and contracts. For many centuries, theological concepts biased the understanding of the Earth and the universe. Claudius Ptolemy’s (90-168 AD) contribution in measuring and projecting the Earth surface is considered as the start of the systematic mapping of the Earth. However, his understanding of the universe was based on the Earth at the center of the universe, i.e., geocentric concept. Nicholas Copernicus (1473-1543 AD) published the “Copernican Model”, which describes the heliocentric concept, i.e., the Earth revolves around the Sun in a circular orbit. This model became the basis for Johannes Kepler’s (1571-1630 AD) laws of planetary motion; the orbit of a planet is an ellipse and not a circle. Galileo Galilei (1564-1642) furthered the Copernican hypothesis and developed laws of dynamics. Building on Copernicus, Kepler and Galileo’s hypotheses, Isaac Newton (1642-1726 AD) discovered the three laws of motion. Albert Einstein’s (1879-1955) theory of general relativity, and his mass and energy equations (where mass and energy are a different manifestation of the same thing), and space-time or 4-D concepts are rooted in Newton’s law of universal gravitation.

The three scientists: Copernicus, Kepler, and Galilei provided a foundation for our knowledge to understand the universe, and Newton unified the concepts of all three scientists and furthered the understanding to the next level. The universal laws, discovered by Newton and the others, opened windows to the 20th-century space missions and technologies. Einstein’s theories brought us the space missions of the 21st century. With the advancement in rocketry, space travel to the moon could be as short as two days and Mars a few weeks. Our ability to view the Earth from space is possible due to the contribution by the above mentioned and many other scientists, engineers, planners, political leaders, and the supporting public.

On January 31, 1958, Explorer 1 launched and became the first U.S. satellite, using its single instrument to send back data about the radiation environment high above Earth's surface (JPL-NASA website). Since then, NASA has launched many missions including the remote sensing satellites (Landsat series: Landsat 1-8), NAVSTAR Global Positioning System (GPS: maintained and operated by US DOD), and the Gravity Recovery and Climate Experiment (GRACE) are directly related to the geospatial technology. Currently, NASA has twenty-five Earth Observing Systems (EOS), and the future plans are for a total of twenty-nine systems (NASA’s EOS website). There are several European, Asian and commercial satellites observing the Earth. The end products of present-day geospatial information systems are produced utilizing Science, Technology, Engineering, and Mathematics (STEM). The following STEM disciplines are directly related to geospatial information systems: astronomy, geodesy, photogrammetry, remote sensing, and computer science.

Global Navigation Satellite Systems (GNSS) include the following satellite systems: NAVSTAR GPS, Russian GLONASS, European Galileo, and Chinese Beidou. Modern GNSS receivers can collect data from all these four navigation satellites. The geodetic networks created using GNSS have become present-day backbones of all geospatial databases which are utilized for creating Comprehensive Geospatial Information Systems (CGIS), engineering applications, and research. The geodetic GNSS networks define the shape of the Earth for a given area and incorporate the geospatial data onto the defined physical surface of the Earth. The GNSS have become a primary tool to collect data for accuracy check and error analysis of any geospatial projects.

The current National Standards for Spatial Data Accuracy (NSSDA) and National Cartographic Standards for Spatial Accuracy (NCSSA) define the geospatial data collection and presentation techniques. The specifications of National Spatial Reference System (NSRS) is controlled by Federal Geodetic Control Committee (FGCC 1984); the positional accuracy of geodetic control points are reported as per the Federal Geographic Data Committee (FGDC 1998), and the ASPRS Accuracy Standards (1990) provides the accuracy standards for large-scale maps. All the above standards and specifications are the basis for updating the70-year old National Map Accuracy Standards (NMAS, US Bureau of the Budget, 1947).

Specifications and standards for data collection using newer sensors and technologies such as airborne and terrestrial Light Detection and Ranging (LiDAR), hyperspectral, RADAR, thermal sensors, etcetera, and specifications for spectral and spatial resolutions, flying-height-scales, image processing, conversion-compatibility, modeling-viewing, and data transfer need to be redefined in the present day technological situation. Issues such as metadata, database, data mining, quality control (QC), spectral resolution and band registration need more research especially for high-resolution data. Various aerial and terrestrial targets with large scale/high-resolution geospatial data collection should be designed to test the spectral resolution and spectral band registration, and should be included in the error analysis or computation of root mean square error (RMSE) of the observations.

METHODOLOGY

Geospatial observations have the limitation of measurement devices or biases that are considered as blunders, systematic errors, and random errors. The random errors are the part of this discussion. Measurements of time, distance, and angle (TDA) are used to create 4-D geospatial information systems. Various sensors placed on satellite, airborne and terrestrial platforms are used to measure TDA about the surface of the Earth. Astronomical observation systems collect data from extraterrestrial bodies such as stars or quasars to measure TDA. Optical and radio interferometry techniques are used in the astronomical observations and these are independent of Earth system, unlike the rest of the other TDA measurements, and thus we consider the random errors computed during the astronomical observations as absolute errors. The GNSS has two positioning techniques: absolute and interferometric or relative. Observations of both GNSS and Very Long Baseline Interferometry (VLBI) are influenced by the similar atmospheric parameters. The absolute GNSS positioning is not up to the accuracy required for high precision and resolution mapping, and therefore we are not considering the random errors from the absolute mode of GNSS observations as absolute errors. Any geospatial observations need a point of beginning. Horizontal and Vertical Datums are utilized as points of beginning for horizontal and vertical data observations. The random errors related to these two surfaces (Horizontal & Vertical Datums) can be considered as absolute errors. The errors generated during the data projection from the ellipsoidal to the spherical to the plane surface (map projections) can be minimized mathematically. Absolute gravity measurements are independent thus the random errors can be considered as absolute errors. Gravity values are critical for computation of orthometric heights during establishing benchmarks or geoid model computations which are most critical in establishing 3-D control points using GNSS technology.

Absolute Errors

Absolute errors are recorded or computed during observations of geospatial data that are independent, complete, and or of the highest accuracy achievable, **Table 1**. The following observations are considered for computation of absolute errors:

Observation Type	Basis and Error Ranges	Remarks
Horizontal Datum Vertical Datum Map Projection	WGS 84 (semi-major axis a= ± 100-cm)	Astronomical, gravity, VLBI and geodetic observations are used to derive semi-major axis and flattening.
Absolute Gravity Observations	± 10 μGal	Micro-g Lacoste Gravimeter A10 Specifications (Measurement Precision)
Astronomical Observations	± 0.1” H; 0.2” V (½ of 0.1 arc second is estimated)	Wild T4 Specifications (Measurement Precision)
Very Long Baseline Interferometry (VLBI)	± 5-cm	Baseline measurement is a relative system; however, it provides the highest accuracy independently
Time measurement	$1x10^{-9}$ to $1x10^{-12}$ Second (Nano - Pico Second)	GPS Clocks are the highly precise clocks which make the absolute positioning precise

Table 1: Sources of Absolute Errors

Relative Errors

Relative errors are computed when the data is compared with the known or given values from higher order surveys, **Table 2**. The following are the observations that are considered for relative errors:

Data Category	Sub-Category	Remarks
Geodetic	GNSS (Global Navigation Satellite Systems) Conventional Traverse Trilateration & triangulation	Millimeter to centimeter level errors Centimeter level errors
Topographic	Stereo photogrammetry LiDAR Ground survey Leveling	Flying height dependent ,, Instrument and distance dependent Millimeter to centimeter
Orthoimagery	Aerial triangulation Color Color infrared Hyperspectral Multispectral	Flying height dependent ,, ,, Sensor and flying height dependent ,,
Planimetric	Transportation Building Hydrographic Miscellaneous	Pixel size dependent, time tagged on the features ,, ,, ,,
Remote Sensing	Satellite Airborne	Sensor and resolution dependent Scale and pixel size dependent
Cadastral	Parcel based Numerical cadastre Torrens system	Based on base orthoimagery used Based on surveying techniques used
Land Use and Land Cover (LULC)	Level III Level IV	Basis: Anderson classifications Errors: based on orthoimagery used
Vegetation	Types	Basis: NDVI (Normalized Difference Vegetation Index)
Spatial Water	Area Quality	Based on orthoimagery used Based on orthoimagery classifications
Wetlands	Manmade types Natural types	Based on orthoimagery used Image classifications
Boundary	National State Local Zip codes	GNSS ,, ,, Digital maps
Marine and Coastal	Type	Image classifications
Geologic	Types	Based image classification and ground surveys
Geographic Names	Coverage and locations	Field verification, community services
Attribute and Non-Spatial	Locations	Based on data extraction and cartographic standards

Table 2: Sources of Relative Errors

GNSS Technology & Computation of Relative Errors

For the last quarter century, it is likely the GNSS technology established more geodetic networks and other control points than the entire history of geodetic control network establishment which started towards the end of the 17^{th} century. The first country to create the nationwide topographic mapping in the world was France. It was supervised by Giovanni Cassini (an Italian astronomer, mathematician, engineer, and astrologer) who was invited by King Louis IV to set up Paris astronomical observatory. Later, he took the leadership role for the topographic mapping with the immense power granted to him by the King to execute the project. The project was completed by his grandson Jacques Cassini. This project took more than a century and three generations to complete it (1670-1790). Gemma Frisius's technique of triangulation was utilized to establish ground control points. The technology was exported to England, India, and the USA in the 19^{th} century. It was the basis for establishing geodetic networks (first, second, third order) and then utilizing them to further establish topographic or cadastral mapping control and checkpoints. At present, the GNSS replaces triangulation, trilateration, and traverse surveys required for the geodetic networks, and control or checkpoints. The major sources of GNSS/GPS errors are as follows (Acharya and Popp,1994):

1. Orbit perturbation
2. Relativity
3. Tropospheric and stratospheric refraction
4. Multipath contamination
5. Ionospheric dispersion
6. Ground master control stations
7. Conversion of GPS height to orthometric height (Ellipsoidal to Geoidal height)
8. Conversion of Earth Centered Earth Fixed (ECEF) coordinate system

The sources two through five of the above-mentioned errors also imply to VLBI or astronomical observations. In the present-day geospatial data collection, GNSS plays a major role in achieving the desired accuracy and it is also used for checking and analyzing the geospatial measurement errors. The GNSS technology has increased the accuracy of geodetic control networks from 1:100,000 to 1:100,000,000, i.e., 1000 times. The first author supervised and tested the GNSS/GPS project for Tennessee Valley Authority (TVA) in 1988; the project was implemented following the draft FGCC standards which were developed to include specifications for the GNSS/GPS surveys (FGCC 1984 and FGCC 1988 Version 5 Accuracy Standards and Specifications for GPS Relative Positioning).

Very soon, there will be more than twenty GNSS satellites visible 24 hours a day for the most part of the Earth; with precise time measurement receivers, an absolute mode of GNSS positioning will meet many geospatial data capture requirements. As shown in **Figure 1**, pseudo range ($\rho_{1\text{-}23}$) measurement equations can be formed for 23 visible satellites as follows:

$$\rho_1 = \sqrt{\{(X_{s1} - X_p)^2 + (Y_{s1} - Y_p)^2 + (Z_{s1} - Z_p)^2\}} \quad (1)$$

$$\cdots\cdots\cdots$$

$$\rho_{23} = \sqrt{\{(X_{s23} - X_p)^2 + (Y_{s23} - Y_p)^2 + (Z_{s23} - Z_p)^2\}} \quad (2)$$

Also,

$$\rho_1 = c*(\Delta t)_1 \quad (3)$$

$$\cdots\cdots$$

$$\rho_{n=1\text{-}23} = c*(\Delta t)_{n=1\text{-}23} \quad (4)$$

Where,

$\rho_1 \ldots \rho_{23}$ are pseudo-ranges; c is the speed of light (299 792 458 m/s), and Δt is the time difference. Coordinates are in ECEF.

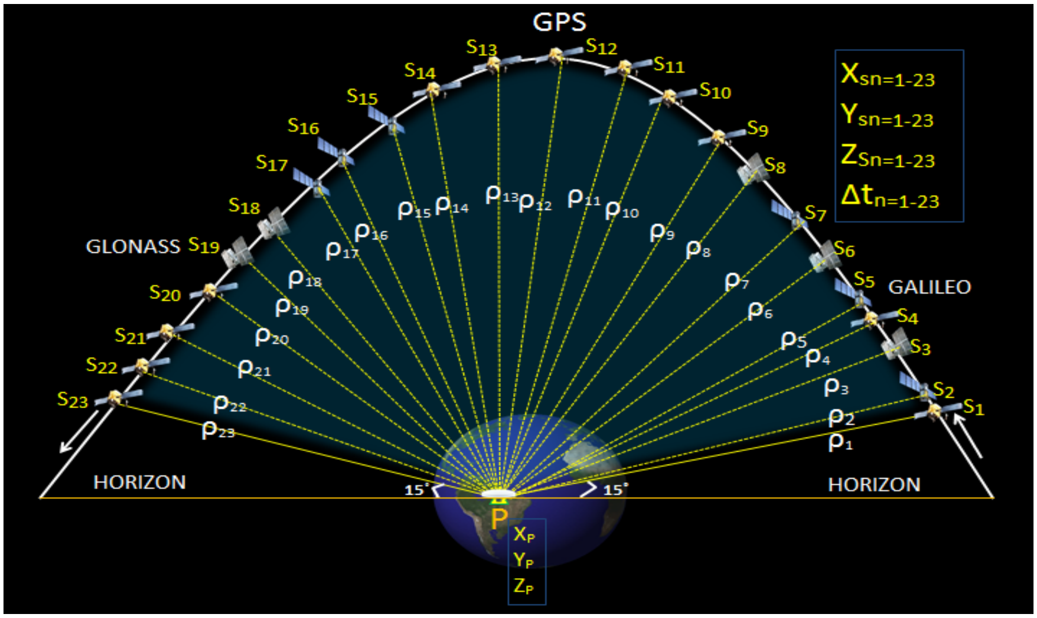

Figure 1: Conceptual GNSS satellite visibility and absolute positioning schema

Using the equation 3, one can estimate the relative errors of GNSS positioning in absolute observation modes, i.e., if the difference in time measurements between Satellite 1 and high precision GNSS ground receiver P is $10x10^{-9}$ (or 10 Nanoseconds), then the position measurement will be limited to 30-cm (299 792 458 * $10x10^{-9}$ = 0.299 meter or close to 30-cm). For many parts of the world, where there are no VLBI stations available, this absolute mode observation with high precision receivers can be used to establish an initial or datum point towards establishing a High Precision Reference Network using GNSS static or relative or carrier phase observations. The absolute mode would be more precise than astronomical observation with the theodolite T4 ($1/20^{th}$ of arcsecond = ~ 150-cm), and less precise than VLBI (5-cm) observation.

Existing Accuracy Standards

The first written accuracy standards: National Mapping Accuracy Standards (NMAS) was for analog mapping and based on map plotting limitations. The horizontal accuracy was divided into two scales:$1/30^{th}$ of an inch for 1:20,000 and larger; $1/50^{th}$ inch for 1:20,000 and smaller. Similarly, the vertical accuracy was defined simply as ½ of the contour interval plotted. The NMAS is based on 90% circular map error (Bureau of Budget, 1947). In March 1990, ASPRS Professional Practicing Division approved the "ASPRS Accuracy Standards for Large-Scale Maps" (ASPRS Annual Report 1990).

IoT and Future Accuracy Standards

Smart cities, smart electric grids, machine to machine communications, and communications among and between all objects located on the surface of the Earth will require geospatial information of the objects (objects located in the universe not included). The accuracy of geospatial information is critical to making the IoT work properly and safety of the systems in the future. Since the digital picture of any object is an accumulation of pixels covering the object; therefore, the accuracy suggested in this paper based on pixel resolution would include the accuracy requirements of any future IoT.

Calculation of Cumulative Root Mean Square Errors (RMSEs)

In the geospatial observations, the root-mean-square-error (RMSE) is calculated to measure the difference between observed values and the true or known or model predicted values. For large scale mapping it is critical to include the absolute, and intermediate or known higher order errors in the relative error calculations. It is proposed that RMSEs of absolute, intermediate or known, and relative or predicted value for horizontal or vertical data should be calculated as follows:

$$RMSE_A = SQRT.\ \{1/n \sum_{i=1}^{n} \left(X_{Obs,i} - X_{T,i}\right)^2\} \quad = \sigma_{A,} \qquad (5)$$

$$RMSE_K = SQRT.\{1/n \sum_{i=1}^{n} \left(X_{Obs,i} - X_{K,i}\right)^2\} \quad = \sigma_K \qquad (6)$$

$$RMSE_R = SQRT.\{1/n \sum_{i=1}^{n} \left(X_{Obs,i} - X_{R,i}\right)^2\} \quad = \sigma_R \qquad (7)$$

Where$_{,\ A,\ K}$, and $_R$ stand for absolute, known or intermediate, and relative or predicted errors.

EMI performed many countywide LiDAR and color orthoimagery projects in the USA. In all projects, EMI establishes GNSS/GPS networks first whose accuracy is dependent to VLBI or HARN control point accuracies, and then it establishes Real Time Kinematic (RTK) or lower order control or checkpoints. An example project data are drawn from Lincoln County Wisconsin, countywide LiDAR mapping project (all 137 checkpoint data not shown in **Table 3**). EMI established a network of 21 first order control points using VLBI or HARN control points from the National Geodetic Survey database, and then it established 137 checkpoints in different land use areas using RTK survey. In this example, it is assumed that the VLBI or HARN control network had an absolute error of 5 centimeters, the 21 points control network established by EMI had a relative error of 7.5 centimeters, and the checkpoints had 11.4 centimeters (0.38-ft) relative errors (Acharya, 2008), **Table 3**.

It is proposed to compute the cumulative or total RMSE of any large scale mapping or engineering projects as follows:

$$\text{Cumulative or Total } (RMSE_T) = SQRT.\ \{(\sigma^2_A + \sigma^2_K + \sigma^2_R)\} \qquad (8)$$

Cumulative RMSE for the Lincoln County project = SQRT $\{(5)^2 + (7.5)^2 + (11.4)^2\}$ = **14.53 cm (0.48 ft.), 0.48 instead of 0.38 Table 3** represents closer to the accurate value of the error computations. Similar procedures should be implied in computing the horizontal RMS Errors. It should be noticed that the absolute error (σ_A) includes the datum and map projection errors.

Project accuracy statistics at 100%				Total checkpoints = 137		
Std. Deviation = 0.38		Mean = 0.01		Coefficient of skew = -0.60		Median = 0.04
RMSE (ft.) = 0.38 (11.4 –cm)						
Land use category	River bank, trees, canopy	Urban, residential	Fully covered canopy-trees	Bare earth, short grass	High grass, shrubs, crops	
RMSE by category	0.19	0.39	0.35	0.36	0.54	
Ground truth points	Easting	Northing	Elevation		LiDAR Points	Δ Elevation
b1	301491.84	251518.24	1594.924		1595.14	-0.22
b2	301454.85	251487.81	1596.609		1596.96	-0.35
b3	301457.15	251465.05	1596.63		1597.29	-0.66

Table 3: Lincoln County, Wisconsin LiDAR quality assurance data

The error computations for determining the accuracy of the high resolution and precision digital data are very critical in the present day context. Major applications of geospatial data are listed in **Table 4**.

WHY NEED ACCURACY STANDARDS OR ERROR ANALYSIS?

Geospatial data are utilized for various applications (see **Table 4**). The accuracy can be optimized by minimizing the errors. The importance of accuracy of geospatial data is generally expressed as a function of the following elements:

$$\text{Accuracy} = CS + CLM + RR + PS \tag{9}$$

Where,

CS = Cost Saving
CLM = Confidence Level Measure
RR = Risk Reduction
PS = Project Success

Standards are the specifications to be followed during a project implementation to meet the accuracy defined or desired.

RECOMMENDATION AND CONCLUSION

Large scale high precision and resolution geospatial data need a newer approach of defining error budgets. Prior to observations and data extractions, computation of various a priori errors, and a posteriori errors for post data extraction should be computed to provide an appropriate confidence level to the data users. An approach based on ground sampling distance (GSD) or resolution of data captured using various sensors should be developed. The Table 4 below provides suggested accuracy standards based on pixel resolution of any objects on the surface of the Earth:

Platform	GSD (σ)	Standard Deviation		RMSE (Position)
		(XY)	Elevation (H)	
Satellite (high altitude)	>1-m	½ σ	½ σ	Base* + ½ σ
Airborne (low to mid altitude)	<1-m	½ σ	σ to ½ σ	Base + ½ σ
UAS (low altitude)	>1-cm	σ	σ	Base + σ
Terrestrial (surface)	>1-mm	Project Specific	Project Specific	Project Specific

Table 4: Suggested pixel based accuracy standards for all digital geospatial information + IoT

*Base = platform, sensor, and instrument limitation

Specifications and standards for error analysis of spatial and non-spatial data such as planimetric features and parcel data, thematic or land use land cover data, feature attribute data, and street names that are used for the applications given in **Table 5** should be developed. National grid schema for large scale digital geospatial data should be defined and used for the consistency of the data. Metadata with time tagging of all features, point, and base imagery data should be developed. It is anticipated that the EMI team will continue this critical research on geospatial accuracy.

A brief historical evolution of geospatial technology (mapping the Earth) is discussed. GNSS technology has revolutionized the geospatial technology in terms of high accuracy and resolution data generation and minimized the errors. A newer approach to error analysis by computing absolute, intermediate or known, and relative errors and combining them to compute the total RMSE of the geospatial products is presented. Importance and applications of the geospatial data in the 21st century has been briefly presented.

Application	Accuracy Requirement	Remarks
E911(National security, emergency preparedness, and management)	Very high	Public safety and security
Infrastructure development data (engineering)	Very high	Cost saving, efficiency, project success
Climate change mitigation and adaptation	Very high	Survival of humanity
Economic development data	High	Transparency, equitable development
Sociopolitical development data	High	Societal harmony

Table 5: Major applications of geospatial data

SELECTED REFERENCES

Acharya, B. et al. (2008). RTK and Static GPS Surveys for the Collection of LiDAR Ground Control and Check Points in Lincoln County, Wisconsin, a GPS Survey Report, p 96

Acharya, B. and Popp, R. (1994). Project Planning for a GPS Geodetic Control Network, *Surveying, and Land Information Systems*, v 54, n 2, Jun, 1994, p 69-75

ASPRS Annual Report (1990). ASPRS Accuracy Standards for Large-Scale Maps, p 1068-1070

Bureau of the Budget (1947), "United States National Map Accuracy Standards", U.S. Bureau of the Budget, June 17

Fan, D. and Zhou G. (2005). An Effective Mathematical Model for Precision Analysis of GPS Data Post-Processing, ASPRS 2005 Annual Conference "Geospatial Goes Global: From Your Neighborhood to the Whole Planet" March 7-11, Baltimore, Maryland

Paulson J. (2005). WSHS2.2 Surveying in Ancient Egypt: From Pharaohs to Geoinformatics, FIG Working Week 2005 and GSDI-8 Cairo, Egypt

Ranzan, C. (2013). The Physical Nature of Length Contraction, doi:10.5539/apr.v5n1p87

http://www.polaris.iastate.edu/EveningStar/Unit2/unit2_sub2.htm

http://www.jpl.nasa.gov/jplhistory/early/firstsatellites.php

http://landsat.gsfc.nasa.gov/?p=10208

http://www.wild-heerbrugg.com/shop/index.php?cPath=1_3_5_25_63