[DM-05-052] Horizontal (Geometric) Datums

Abstract

A horizontal (geometric) datum provides accurate coordinates (e.g., latitude and longitude) for points on Earth's surface. Historically, surveyors developed a datum using optically sighted instruments to manually place intervisible survey marks in the ground. This survey work incorporated geometric principles of baselines, distances, and azimuths through the process of triangulation to attach a coordinate value to each survey mark. Triangulation produced a geodetic network of interconnected survey marks that realized the datum (i.e., connecting the geometry of the network to Earth's physical surface). For local surveys, these datums provided reasonable positional accuracies on the order of meters. Importantly, once placed in the ground, these survey marks were passive; a new survey was needed to determine any positional changes (e.g., due to plate motion) and to update the attached coordinate values. Starting in the 1950s, due to the implementation of active control, space-based satellite geodesy changed how geodetic networks were realized. Here, "active" implies that a survey mark's coordinates are updated in near real-time through, for example, artificial satellites such as GNSS. Increasingly, GNSS and satellite geodesy is paving the way for a modernized geometric datum that is global in scope and capable of providing positional accuracies at the millimeter level.

Keywords: coordinate systems, datums, horizontal datums, spatial reference systems

Author & citation

Kessler, F. (2022). Horizontal (Geometric) Datums. The Geographic Information Science & Technology Body of Knowledge (2nd Quarter 2022 Edition). John P. Wilson (Ed.).
DOI: 10.22224/gistbok/2022.2.6.

This Topic is also available in the following editions:

DiBiase, D., DeMers, M., Johnson, A., Kemp, K., Luck, A. T., Plewe, B., and Wentz, E. (2006). Vertical datums. The Geographic Information Science & Technology Body of Knowledge. Washington, DC: Association of American Geographers.

Explanation

1. Definitions
2. Geodesy and Datums
3. Earth's Shape and Size
4. Classic, Transitional, and Modern Geometric Datums Used by the United States
5. Modern 3-D Geometric Terrestrial Reference Systems

Definitions

Earth centered-Earth fixed Cartesian coordinate system: a Cartesian spatial reference system representing locations on or near Earth's surface as X, Y, and Z measurements from its center of mass or geocenter.

Geocentric: referring to Earth’s center as in its center of mass.

Geodetic control: a set of control stations established by geodetic methods. The data of geodetic control consist first of the distances, directions, and angles between control stations. These are converted to geodetic coordinates and azimuths. The latter, in turn, may be converted into other kinds of coordinates such as plane coordinates in a State Plane Coordinate System.

Geodetic datum : a set of constants specifying the coordinate system used for geodetic control (i.e., calculating coordinates of points on the Earth).

Geodetic network: a network whose points are survey stations or gravity stations.

Geometric geodesy: a branch of geodesy that combines geometric datums/reference frames, latitude, longitude, ellipsoid height, and state plane coordinates.

Horizontal datum: a geodetic datum specifies the coordinate system in which horizontal survey marks are located.

Reference ellipsoid: an ellipsoid of specified dimensions and associated with a terrestrial reference system or a geodetic datum.

Survey mark: a dot, the intersection of a pair of crossed lines, or any other physical point corresponding to a point in a survey. The physical point to which distances, elevations, heights or other coordinates refer.

Terrestrial Reference Frame: a realization of a terrestrial reference system.

Terrestrial Reference System: a theoretical spatial reference system that co-rotates with Earth through space and time.

Triangulation: a survey method in which the points whose locations are to be determined, together with a suitable number (at least two) of points of known location, are connected in such a way as to form the vertices of a network of triangles. The angles in the network are measured and the lengths of the sides are either measured or calculated from known points and lengths.

Many of the terms used in the first part of this chapter are discussed in related GIS&T BoK entries on Earth’s Shape, Sea Level, the Geoid, Geographic Coordinate Systems (forthcoming), and Vertical (Geopotential) Datums, whose content presents the foundatonal knowledge on which this entry on Horizontal Datums builds.

2. Geodesy and Datums

Friedrich Helmert (1880) provided an early definition of geodesy as a field focused on the science of measuring Earth’s size and shape. He divided geodesy into four disciplines, two of which are horizontal and vertical control: horizontal refers to a coordinate system (e.g., latitude and longitude) used to define spatial positions while vertical deals with heights on Earth’s surface. Geodetic control establishes the orientation, scale, and accuracy of the system in which those coordinate values and heights are established, monitored, and updated, creating a geodetic network. A series of survey marks for which accurate coordinate values and heights are known provides public access to the network. Surveyors, for example, access a geodetic network via survey marks to complete their projects (e.g., constructing a tunnel) using accurate coordinate positions and heights. Historically, a geodetic network comprises separate but integral horizontal and vertical geodetic datums. A geodetic datum is any quantity or set of such quantities that may serve as a referent or basis for calculation of other quantities as applied to Earth's surface (NGS, 2001). Increasingly, space-borne technologies facilitate a unified datum in which horizontal and vertical control are combined to form a single terrestrial reference frame.

3. Earth's Shape and Size

For centuries, people believed that Earth was spherical. For example, Eratosthenes of Cyrênê (276-195 BCE) assumed a spherical shape based on direct observation (e.g., observing eclipses). While a spherical assumption can be used to carry out a systematic measurement of Earth's circumference, and while such an assumption satisfies medium and small-scale thematic mapping endeavors, high accuracy survey work or precise large-scale mapping efforts require more detailed knowledge of Earth's exact shape. A spherical Earth assumption ignores the subtleness by which its true shape undulates due to, for example, variations in the gravity field and crustal motions that are important to defining modern horizontal datums.

3.1. The Geoid

Earth's shape is controlled, in part, by its gravity field, which makes it more complex than a sphere. To illustrate, Figure 1 shows point P resting on Earth's surface with a unit mass a near the surface. Gravity acts to attract P by the unit mass toward Earth's center via centripetal force. This force is shown by vector F. As Earth spins, a centrifugal force results (indicated by the vector C), in pulling P away from Earth's surface. Combining centripetal and centrifugal forces results in the gravity force (indicated by the vector G). Mok and Chao (2012) describe that the magnitude of G represents the potential energy of P. Since P can exist at any point on Earth's surface (e.g., P'), and since the mass of its crust varies (e.g., at mass b) and its surface is dynamic, the potential of G varies across Earth's surface (e.g., P' has a different C). Measuring G across Earth's surface creates a potential or geopotential surface. This geopotential surface defines the **geoid**, a more accurate representation of Earth's undulating shape. The geoid's primary role is to provide a zero-height surface. See entries on Earth's Shape, Sea Level, and the Geoid and Vertical (Geopotential) Datums for more discussion about Earth's geopotential surface, the geoid, and vertical (geopotential) datums.

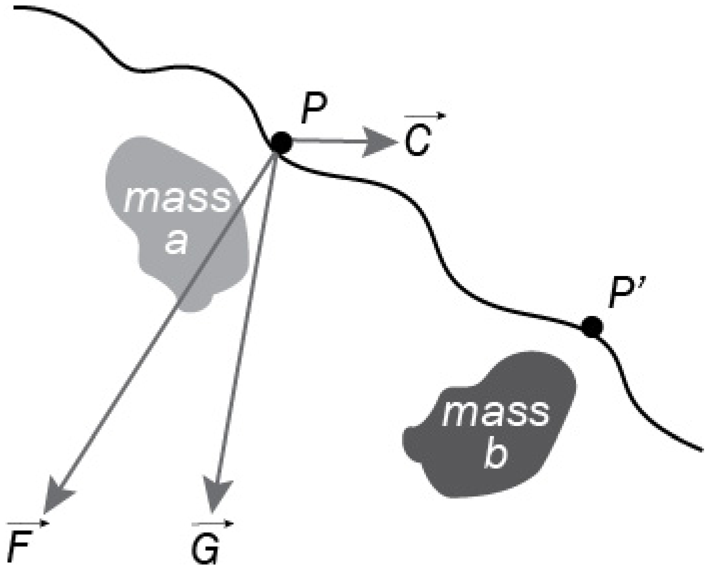


Figure 1. A simplification of the gravitational attraction F on P, the centrifugal force C pulling at P, and the resulting force of gravity G. Source: author. Image after Mok and Chao (2012).

3.2. Ellipsoids

The idea illustrated in Figure 1 supports Newton's theory of universal gravitation, put forth in his Principia Mathematica, and is useful in defining Earth's shape as an **oblate ellipsoid**. Newton asserted that a large body rotating about its central axis at high speeds would, experiencing centrifugal forces, bulge at the equator and compress at the poles. Modern measurements suggest the difference in radii between the plane in which the equator and prime meridian rest is approximately 21.384 km. This difference was determined by taking the difference in the radii values for the GRS80 reference ellipsoid.

3.2.1. Reference Ellipsoids

Mok and Chao (2012) relate that while the geoid is a more accurate representation of Earth's shape, its surface is not well suited for a horizontal coordinate system. Due to the geoid's undulating surface, attempting to establish a regular spatial coordinate system by applying simple mathematical rules is problematic. Instead, a **reference ellipsoid**, a solid body that is defined by simple mathematics, and which creates a smooth surface that closely approximates the geoid's shape, is used.

3.2.2. Reference Ellipsoid Parameters

A 2-D cross-section of an oblate ellipsoid reveals an ellipse (Figure 2A). To fit the reference ellipsoid to the geoid's size and shape requires defining the ellipsoid's parameters. Historically speaking, five parameters are used, including: semimajor axis (a), semiminor axis (b), flattening (f), first eccentricity (ε2), and second eccentricity (ε'2).

The **semimajor axis** (a) and the **semiminor axis** (b) is the distance from the ellipsoid's center to its perimeter along the equatorial plane and the distance from the ellipsoid's center to one of its poles along the polar or meridional plane, respectively (Figure 2A).

Meyer (2010) elaborates that the size of the ellipse is set by the value of a and the shape of the ellipse is set by some value of eccentricity, either the value of b and f, ε2, or ε'2. Values of a and b are used to compute values of f, ε2, and ε'2, which, individually, can be used to describe the noncircularity of an ellipsoid. The literature often interchanges e and ε as symbols representing eccentricity. Figure 2B illustrates values of b and f as they define the noncircularity of two ellipsoids. Note that the **flattening** (f) parameter is sometimes reported as a reciprocal, also known as **inverse flattening** (1/f). Meyer (2010) and Lu et al. (2012) provide an extensive discussion of the mathematics of the ellipse computations.

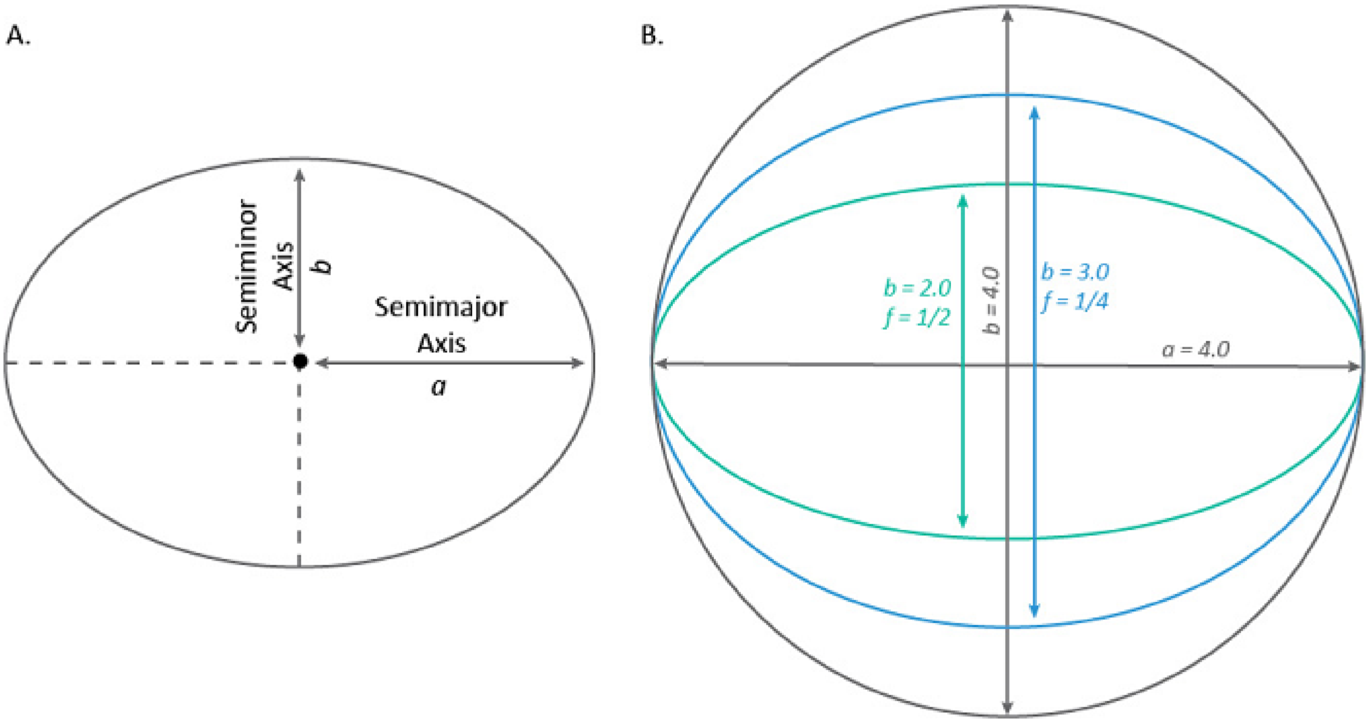


Figure 2. A 2-D cross-section of an oblate ellipsoid illustrating the parameters a and b (A) and the noncircularity values of b and f (B). Source: author.

Rotating an ellipse about its semi-minor axis b, keeping the value of a constant, results in a solid figure known as a **bi-axial ellipsoid of revolution**. Since the equatorial plane can also be treated as noncircular, Torge and Muller (2012) discuss the parameter c which rests in the equatorial plane orthogonal to b and results in a **tri-axial ellipsoid of revolution**. See entries on Earth's Shape, Sea Level, and the Geoid and Geographic Coordinate Systems (forthcoming) for further discussion of reference ellipsoids.

3.2.3. Locally or Globally Fitted Reference Ellipsoids

Depending on values chosen for a and f, for example, reference ellipsoids can be developed that fit specific regions of the geoid’s surface or its entirety. Figure 3A illustrates an ellipsoid that minimizes the separation between the ellipsoid and geoid surfaces for a localized region at M’ (e.g., Australia), but a greater separation is seen in other places (e.g., M). Figure 3B shows an ellipsoid that minimizes the overall separation (e.g., the separation seen at M and M’ are more balanced).

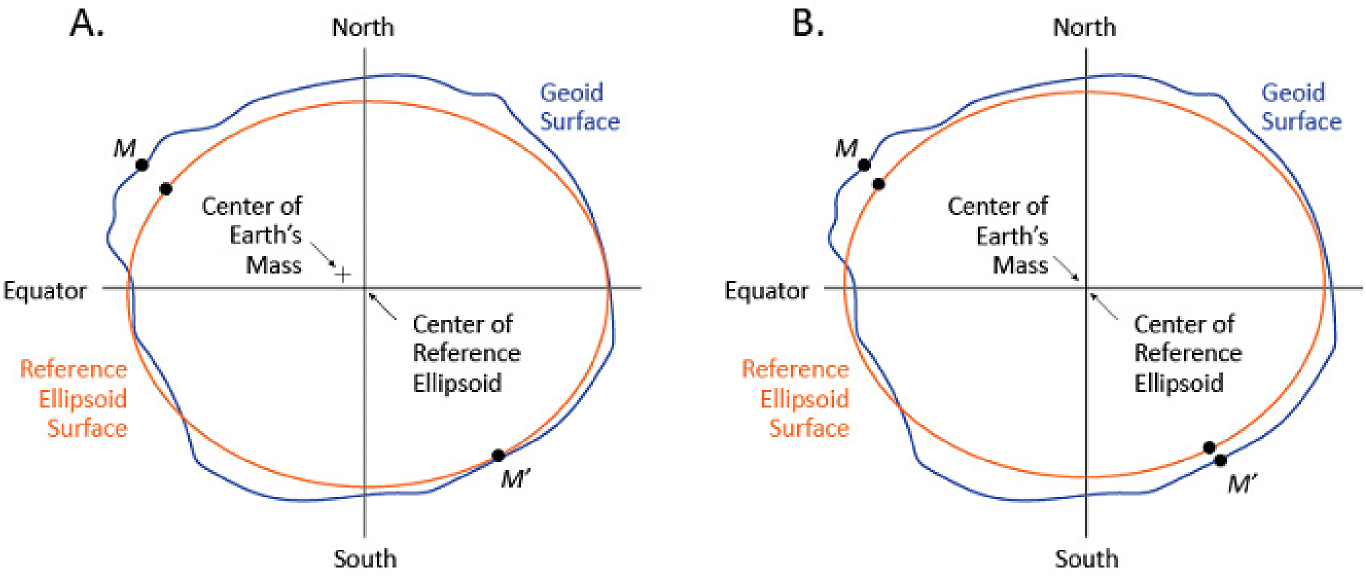


Figure 3. A highly exaggerated representation of a reference ellipsoid designed to 'fit' Australia (A) and the entire geoid (B). Source: author.

3.2.4. Common Reference Ellipsoids

Since the early 1800s, dozens of reference ellipsoids have been developed. Table 1 and Table 2 presents the parameters for selected locally and globally fitted reference ellipsoids, respectively. Values of a and b are reported in meters while values of 1/f are unitless.

Table 1. Common Local Reference Ellipsoids*

Reference Ellipsoid Name	Date of Publication	a (m)	b (m)	1/f	Intended Geographic Area
Airy	1830	6,377,563.396	6,356,256.909	299.3249646	Great Britain
Bessel	1841	6,377,397.155	6,356,078.962	299.1528128	Europe
Clarke	1866	6,378,206.4	6,356,583.8	294.9786982	North America
Clarke	1880	6,378,249.17	6356514.9	293.465	France and much of Africa
Everest	1830	6,377,276.345	6,356,075.4131	300.8017	India
Hayford / International	1909	6,378,388.0	6,356,911.9	297	Various Countries
International	1924	6,378,388.0	6,356,911.9	297	Europe
Krassovsky	1940	6,378,245.0	6,356,863.018	298.3	Soviet Union

* Values for a and f taken from the European Petroleum Survey Group's (EPSG) Geodetic Parameter Database (epsg.org/home.html). Some values of b were calculated using b = a • (1-f).

Table 2. Common Global Reference Ellipsoids*

Reference Ellipsoid Name	Date of Publication	A	B	1/f	Intended Geography Area
GRS67	1967	6,378,160.0	6,356,774.161	298.247167427	World
GRS80	1980	6,378,137.0	6,356,752.3141	298.257222101	World
WGS66	1966	6,378,145.0	6,356,760.0	298.25	World
WGS72	1972	6,378,135.0	6,356,750.52	298.26	World
WGS84	1984	6,378,137.0	6,356,774.7191	298.257223563	World
Fischer	1960	6,378,166.0	6,356,784.283666	298.3	The Mercury Project

* Values for a and f taken from the EPSG’s Geodetic Parameter Database (epsg.org/home.html). Some values of b were calculated using b = a • (1-f).

3.2.5. Earth Centered-Earth Fixed Coordinate System

A 3-D coordinate system can be described that locates points on Earth’s surface. The origin of this system, containing three orthogonal coordinate axes (X, Y, and Z), is located at Earth's center of solid mass called the geocenter (Figure 4). Each axis rests in its own plane that passes through the geocenter. Values expressed in X, Y, and Z constitute a geocentric Earth Centered-Earth Fixed (ECEF) Cartesian coordinate system and are designed for globally fitted reference ellipsoids. This ECEF coordinate system allows geodetic coordinates of latitude (φ) and longitude (λ), and ellipsoidal height (h), to be computed from X, Y, and Z (see Iliffe and Lott, 2000). Bowring (1976) presents suitable formulas for the inverse operation (i.e., using φ, λ, and h to compute X, Y, and Z).

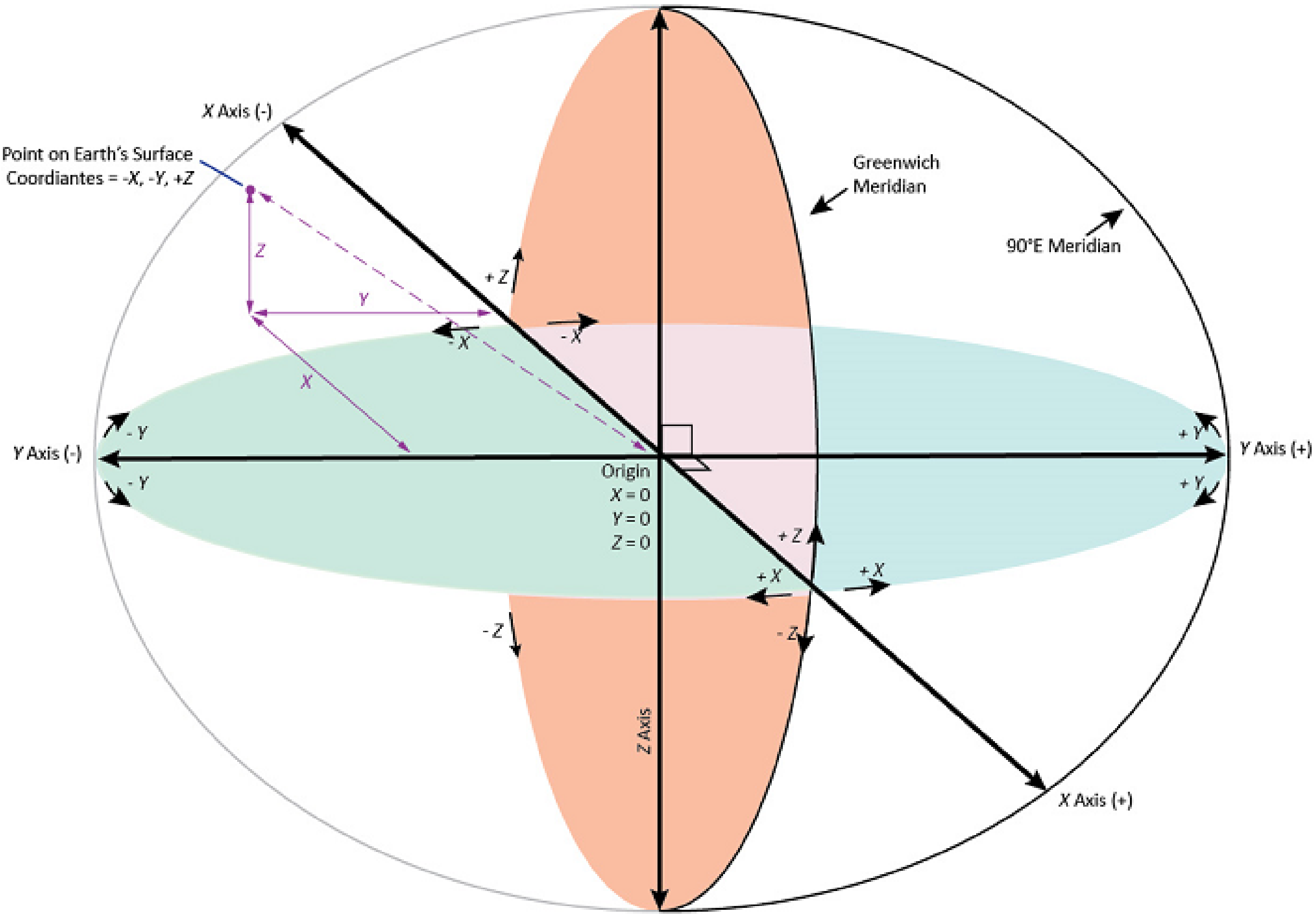


Figure 4. The X, Y, and Z axes of an ECEF Cartesian coordinate system. Source: author.

In an ECEF Cartesian coordinate system, the plane of the X-axis coincides with the internationally recognized plane of the conventional equator and is orthogonal to a chosen zero-meridian (today, Greenwich, England). This meridian is named the International Reference Meridian and is maintained by the International Earth Rotation and Reference Systems Service (IERS). Extending orthogonally from the geocenter and resting in the same equatorial plane is the Y-axis. Both X- and Y-axes rotate with the Earth about the Z-axis. The Z-axis is aligned with the North Pole (Earth's axis of rotation) which is the International Reference Pole.

4. Classical, Transitional, and Modern Geometric Datums Used by the United States

The mission of NOAA's National Geodetic Survey (NGS) is to define, maintain, and provide access to a unified National Spatial Reference System (NSRS). The NSRS is the official reference system for accurate latitude, longitude, height, scale, gravity, and orientation throughout the United States and its territories. It provides foundational geodetic control for the nation's transportation, mapping, and charting infrastructure, and it supports a multitude of scientific and engineering applications (NOAA, 2021). Increasingly, the NSRS relies on space-based technologies such as GNSS to provide active geodetic control to provide mm accuracy in defining the NSRS. Smith et al. (2015) report that monitoring global-based phenomenon such as sea level rise, subsidence, shoreline boundaries, and the motion of the continental plates requires a consistent coordinate system with cm or mm accuracies. Modernizing and providing access to the NSRS is an ongoing NGS project that has seen adjustments to horizontal datums developed for the United States and driving the development of the new North American Terrestrial Reference Frame 2022.

4.1. Classically Derived 2-D Geometric Datums

To build a locally fitted classical datum, a geodetic survey uses principles of geometry and geometric quantities such as distances and azimuths to create a horizontal (geometric) datum (Ewing and Mitchell, 1970). As Hooijberg (2008) states, a classically derived geodetic datum is based on the results of the geodetic survey that builds the geodetic network and possesses at least five parameters: 1) latitude of the POB, 2) longitude of the POB, 3) azimuth from the POB to another control point, 4) the value of a for a given reference ellipsoid, and 5) value of 1/f for a given reference ellipsoid.

Gossett (1959) explains that geodetic survey work begins with the “initial point” or “point of beginning (POB).” The POB is a passive physical marker (a survey mark) that can be permanently set in the ground with a carefully derived latitude and longitude through astronomical observation. From the POB, another intervisible survey mark is set in the ground at some azimuth and distance, creating a baseline. (This intervisible spacing is necessary due to limitations of the optical survey instrumentation available at the time. To extend the limited, ground-level intervisibility, steel Bilby towers (observation platforms) were constructed (Young, 1974).) From this baseline, a third control point is set, connecting the previous two survey marks to form a triangle. To ensure the integrity of the triangle net and to minimize measuring long distances of a given baseline, angles are recorded and checked for mathematical consistency (e.g., using the Law of Sines). This triangulation survey process was repeated many times over, creating corridors of quadrilaterals whose baseline nodes were survey marks. These quadrilaterals covered thousands of miles and provided information on Earth’s shape across a sizable portion of its surface (Figure 5). The result of this triangulation process constituted a geodetic network of survey marks that offered a surveyor access to a geometric datum.

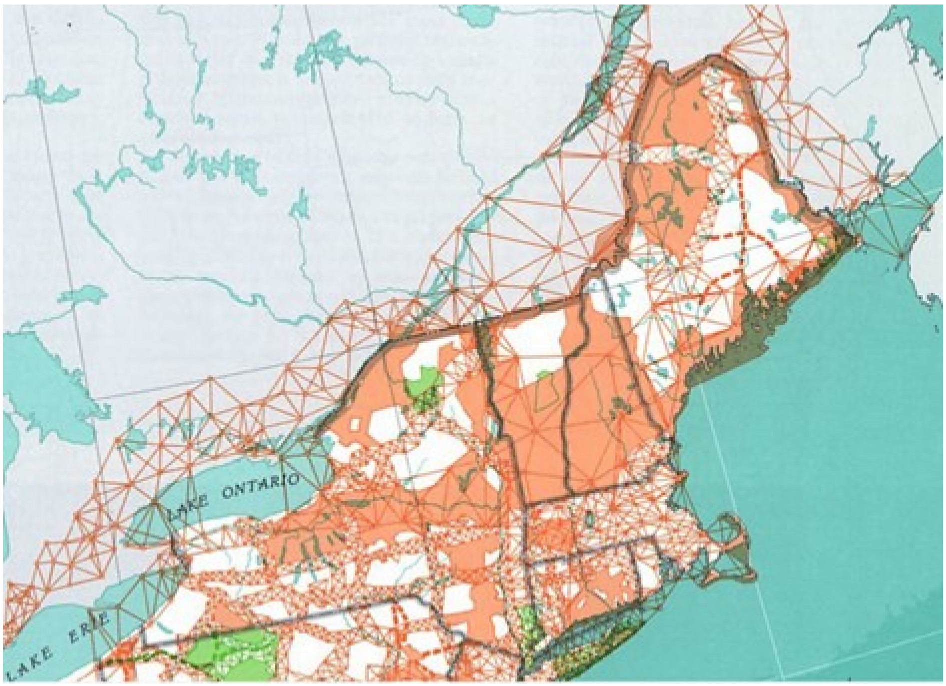


Figure 5. The quadrilaterals that resulted from the triangulation used to build NAD27. Source: The United States National Atlas, 1970.

It is easy to see how small errors in a triangulation process can propagate throughout the larger geodetic network. Accounting for this propagation through a least-squares adjustment can take place once the network is completed. Least squares is a statistical method that uses weighted equations derived from the coordinates of control points. The equations aim to minimizes the sum of the residuals in the measurement data where a residual is the amount of correction applied to a measurement for that measurement to fit into a best-fit solution. The least squares method also provides a means to statistically assess the results of the best-fit solution (Ghilani and Wolf, 2008). Although an adjustment can be applied to a geodetic network to balance out the errors, the errors are still there, nonetheless.

4.1.1. North American Datum 1927 (NAD27)

Triangulation was used in the United States and other countries to establish horizontal datums (e.g., see Wolffhardt, 2017 discussing the great survey of India). Dracup (1994) detailed the development of classically derived horizontal datums in the United States. The New England Datum 1879 was the first datum. The Clarke 1866 was the selected reference ellipsoid and the POB was named Principio, in Maryland. The choice of this POB was deliberate, as this location was approximately central to the geographic area under consideration. The triangulation used to develop this datum was extended south and west to “tie in” to other completed surveys along the Pacific and Gulf coasts, and in 1901, this extended network was named The United States Standard Datum. Meades Ranch, in Kansas, was then set as the POB, due to its centralized location in the coterminous United States (CONUS). In 1913, Canada and Mexico agreed to base their triangulation networks on The United States Standard Datum. After years of additional survey and adjustment work, the NAD27 was realized. This realization allowed public access to the datum so that, for a project, a surveyor could tap into the network through one of the existing survey marks.

Despite its completion, Schwarz (1989) discussed three limitations to NAD27. First, as new survey work was added to the existing geodetic network, there was no systematic, simultaneous adjustment across the entire network to ensure internal consistencies. Second, the network did not provide adequate density of control points. Third, the network did not systematically account for crustal motion. Collectively, the accuracy of the network on which NAD27 was built was found to possess error as much as one part in 15,000. As technology began made into geodetic surveying methods during the 1950s and 1960s, more favorable errors in the range of one to 50,000 or one in 100,000 became available.

Two other limitations of NAD27 are noteworthy but are understandable considering the time in which the datum was developed. First, the survey marks used to build the NAD27 network were passive. Once surveyed, the coordinates and heights were not updated unless a new geodetic survey was conducted. Second, the origin of the Clarke 1866 reference ellipsoid was assumed to be somewhere close to Earth’s center (Van Sickle (2017, 49) reports this distance separation to be 236m). At the time of NAD27’s development, however, technology did not exist to provide active, updated geodetic control, and knowing the exact location of this center was not of concern as surveys were usually localized in scope.

4.1.2. North American Datum 1983 (NAD83)

Realized in 1986, NAD83 was also classically derived, as it was built on NAD27’s geodetic network. However, two technologies changed the accuracy level of coordinate positions and distances used in the geodetic surveys that were utilized to develop NAD83: Satellite doppler (TRANSIT) and electronic distance measuring (EDM), both of which emerged during the 1950s. Between 1964 and 1996, seven United States Navy TRANSIT doppler satellites allowed X, Y, and Z to be simultaneously recorded at a single location. And by tracking the Doppler shifts in these satellite signals, any position on Earth's surface, with respect to its geocenter, could be determined with positional errors in the range of 5 to 1 m (Anderle, 1983). TRANSIT was used to establish the geocenter of the Geodetic Reference System reference ellipsoid (Hooijberg, 2008). EDM techniques allowed surveyors to accurately measure lengths of baselines with relative accuracies of one in 5,000 (Maling, 1996). Snay (2012) points out that EDM illuminated rather large distortions in the positional accuracies of NAD27’s geodetic network of control points – in some cases, greater than 1m. Collectively, TRANSIT and EDM provided quantitative evidence that NAD27 no longer met the positional accuracy that these technologies could provide. A new geometric datum was needed.

Schwarz (1989) explains five unique approaches that were implemented in developing NAD83. First, a system-wide, simultaneous adjustment of the NAD27 network incorporated both Doppler and EDM techniques during the triangulation process. Second, data collected during the adjustment of the NAD27 network was converted into machine-readable form so that computers could store and perform the complex calculations. Schwarz (1989) reports that this database was between 3-4 Gb. Third, a new reference ellipsoid (GRS80) was chosen to reflect a fundamental change that due to the still developing field of satellite geodesy cast NAD83 with a geocentric reference ellipsoid. Interestingly, there was ongoing debate on the merits of adopting a geocentric reference ellipsoid for NAD83. Chovitz (1989, 82) summarizes the argument by asking “which surveyor was to suffer the potential confusion caused by the switch to a new datum – the one utilizing a satellite survey system or the one attempting to observe and compute highly accurate conventional surveys?" Forward vision suggested that more of the former would exist as satellite technology progressed, and less of the latter. Fourth, to accommodate the enormity of a simultaneous least-squares adjustment of the NAD27 network required a different approach. Helmert blocking divided up the process into geographic "blocks" that required smaller computational tasks and resulted in large savings in computer storage and CPU requirements. Those individual blocks were later reassembled into a single solution (Pearson, 2005). Fifth, crustal motions and their impact on horizontal positions was modeled via Regional Deformation of the EArth Models - README (Snay et al., 1989). Figures 6 and 7 illustrate the horizontal change in longitude and latitude (in meters) between NAD27 and NAD83(1986).

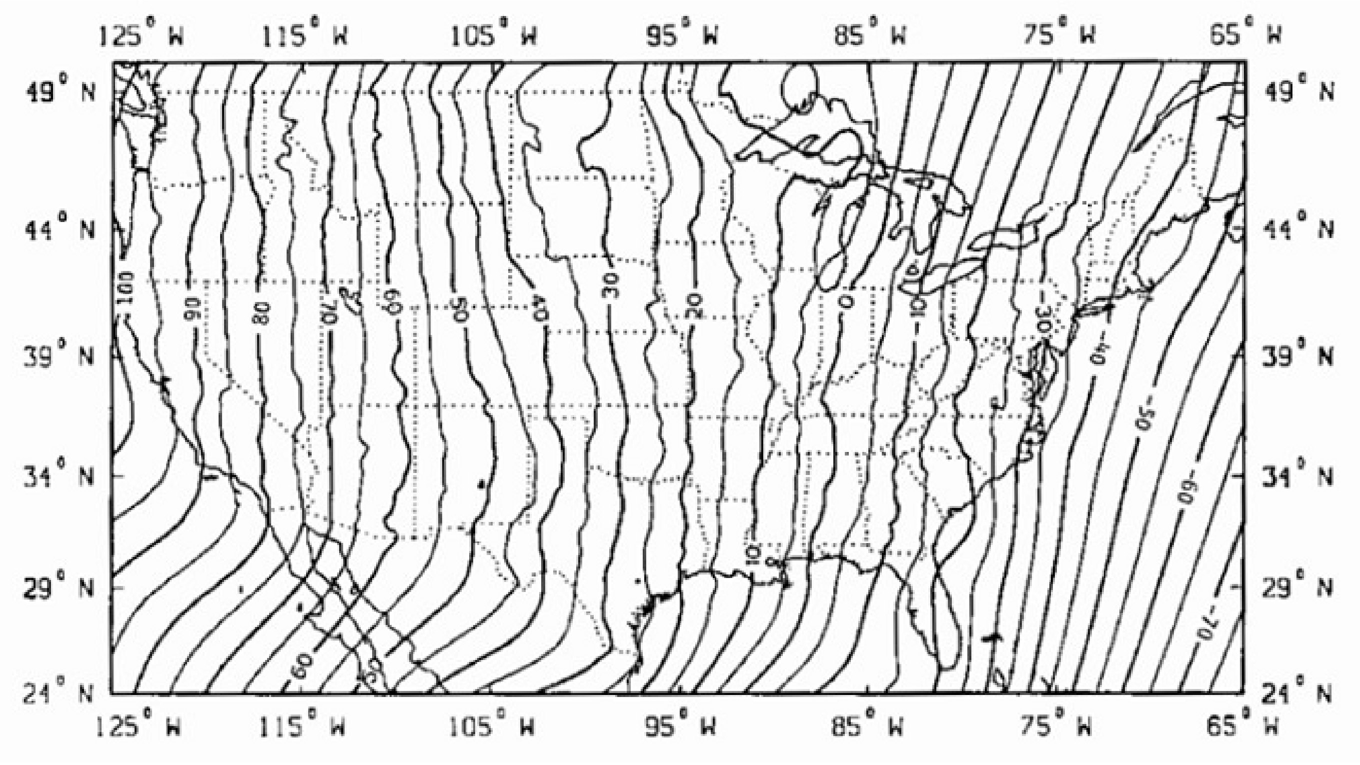


Figure 6. The horizontal change in longitude from NAD27 to NAD83 in meters. Source: [National Geodetic Survey](#).

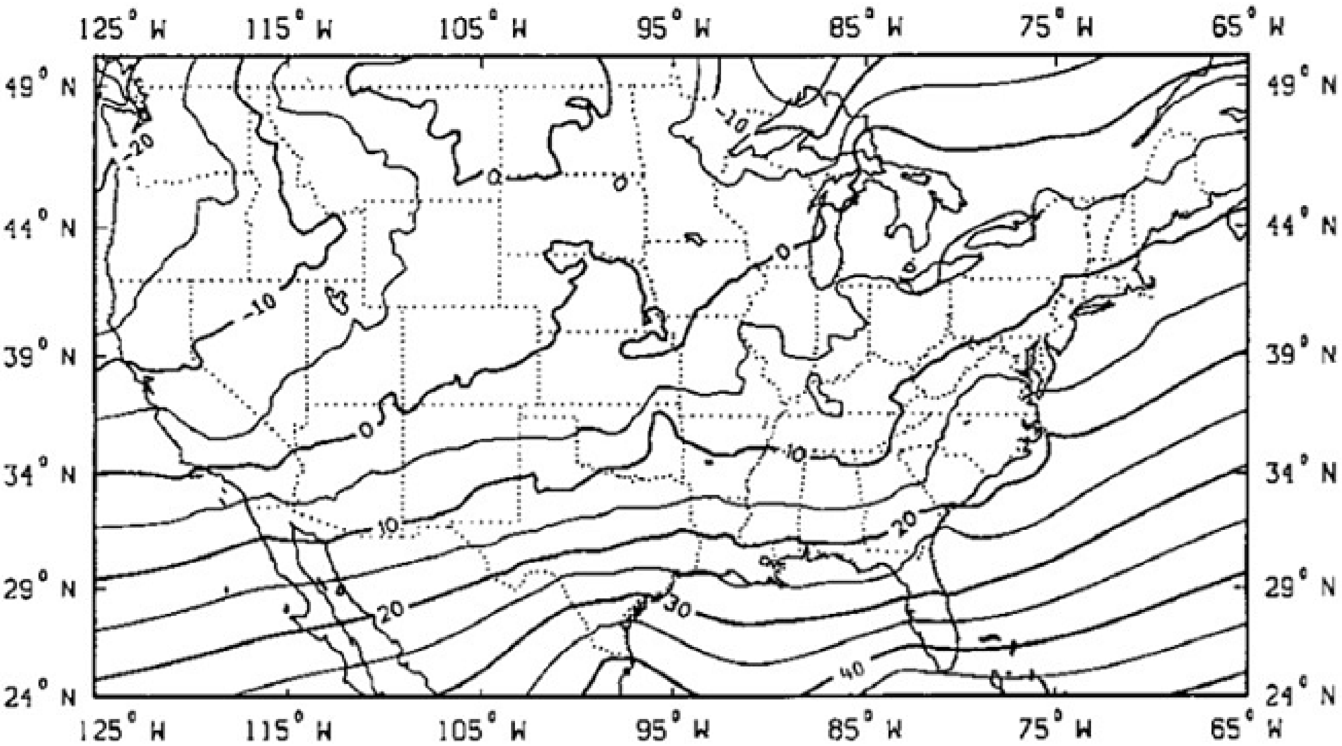


Figure 7. The horizontal change in latitude from NAD27 to NAD83 in meters. Source: National Geodetic Survey.

4.2. Transitional 3-D Datums: NAD83 Realizations

Since NAD27 and NAD83(1986) datums provided only latitude and longitude values to users, they are considered 2-D. If a surveyor desired to incorporate heights into their work, a separate 1-D vertical datum was available: NGVD29 or NAVD88. Discussed below, space-based technologies allowed the simultaneous integration of the horizontal coordinates and vertical heights into a modernized single 3-D datum that evolved through different realizations and adjustments to NAD83.

4.2.1. Artificial Satellite Technologies

Coupled with TRANSIT and EDM, other space-based technologies became available and were instrumental in developing a 3-D datum: global navigational satellite system (GNSS), very long baseline interferometry (VLBI), satellite laser ranging/lunar laser ranging (SLR)/(LLR), and DORIS (Doppler Orbitography and Radiopositioning Integrated by Satellite). GNSS refers to the general artificial satellite constellations that orbit Earth and provide accurate data that can be used to determine X, Y, and Z ECEF coordinates. Specific GNSS examples and their origins include GPS (United States), GLONASS (Russia), and Galileo (European Union). VLBI accurately measures distances (baselines) between radio signals received from intergalactic phenomenon (e.g., quasars) and is useful to determine Earth’s orientation and provide scale to a geodetic network. The most recent VLBI platform (VLBI 2020) expects to achieve accuracies of 1 mm in position and of 0.1 mm/year in crustal velocity (Schuh and Behrend, 2012). SLR/LLR is a global network of observation points that measures the travel time of ultrashort pulses of light emitted from ground stations to satellites equipped with retroreflectors (e.g., LAGEOS satellite). SLR/LLR provides instantaneous range measurements of millimeter-level precision that can be accumulated to provide accurate measurement of gravitational variations due to plate motions on a global scale as well as definition of a global reference frame (Sośnica, 2014). DORIS is a French satellite tracking system initially designed in the late 80s by Centre National d’Etudes Spatiales (CNES), the French Space Agency, as a dense global network of ground-based radio beacons that emits a signal at a known frequency that provides precise orbits to Low Earth Orbiting (LEO) satellites (Willis, et al., 2010).

4.2.2. Satellite Geodesy

Collectively, these space-based technologies ushered in a new field called satellite geodesy. The allied technologies allowed for the simultaneous determination of two or more non-intervisible triangulation stations through artificial satellites elevated sufficiently high above Earth’s surface that provided signals detectable over large distances. Snay (2012) reports that these technologies made it possible for accurate sub-centimeter positional measurements. Satellite geodesy also revealed that the geocenter adopted for NAD83(1986) and the original WGS84 realization, the GRS80 reference ellipsoid, was displaced by about 2.2 meters from Earth's true geocenter (Smith et al, 2015). Moreover, the scale of NAD83(1986) was approximately 0.0871 ppm shorter than the true definition of a meter (Snay and Soler, 2000). Increasingly then, satellite geodesy enabled geodetic surveys to produce positional accuracies that were higher than what NAD83(1986) provided. As a result, several adjustments to NAD83(1986) followed that modernized NAD83(1986) and improved its geodetic network accuracy.

4.2.3. NAD83(HARN)

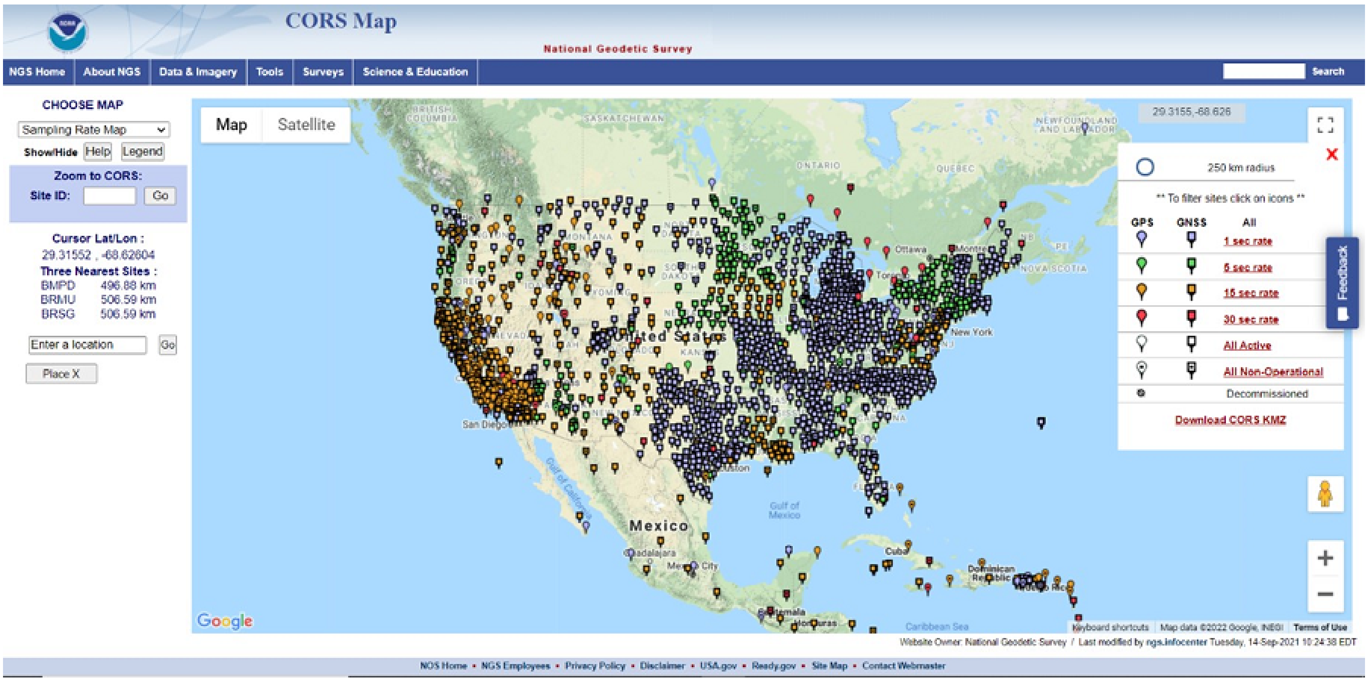
Beginning in 1989, and using GNSS, each state developed a statewide geodetic network called a high-accuracy reference network (HARN); HARNs were originally referred to as high-precision geodetic networks (HPGN). As development continued, each HARN network was tied into and adjusted with the surrounding states' networks (Strange and Love, 1991), creating 50 HARN geodetic networks. All HARNs were generally concluded by 1997. Snay (2012, 163) reports that “relative coordinate accuracies among HARN reference stations are better than 1 ppm.”

4.2.4. NAD83(CORS9X)

In 1994, the Continuously Operating Reference Station (CORS) was used to adjust NAD83’s geodetic network accuracies. A CORS relies on GNSS to provide “active” control consisting of permanently mounted GNSS antennas to fixed monuments (Figure 8). NGS collects, processes, archives, and publicly distributes positional and height data freely from the nationwide CORS network (NCN). There are numerous CORS stations (Figure 9) with varying density by state (i.e., more in tectonically active states like CA and fewer in more stable states like KS). Each CORS station provides coordinates and ellipsoid heights reported in X, Y, and Z and latitude and longitude, and ellipsoid height, based on GRS80 and ITRF (Stone, 2006). To ensure the real-time health of the CORS network and data, NGS participates in the International GNSS Service (IGS) that precisely computes daily orbits for the GPS constellation that serves the NCN (Dow et al., 2009). The CORS GNSS data and metadata augments a surveyor’s data collection, allowing them to directly tap into the NSRS. This CORS realization experienced three adjustments: NAD 83(CORS93), NAD 83(CORS94), and

NAD 83(CORS96).

Figure 8. A CORS antenna mounted to a post. Source: [National Geodetic Survey](#).



Sampling Rate Map

Zoom to CORS:
Site ID:

Cursor Lat/Lon :
29.31552 , -68.62604

Three Nearest Sites :

BMPD	496.88 km
BRMU	506.59 km
BRSG	506.59 km

29.3155,-68.626

○ 250 km radius

** To filter sites click on icons **

GPS	GNSS	All
		1 sec rate
		5 sec rate
		15 sec rate
		30 sec rate
		All Active
		All Non-Operational
●		Decommissioned

Download CORS KMZ

Website Owner: National Geodetic Survey / Last modified by ngs.infocenter Tuesday, 14-Sep-2021 10:24:38 EDT

Figure 9. The NCN showing the varying densities of CORS stations among the states, territories, and countries. Source: National Geodetic Survey.

In concert with the NCN, users have free access to the Online Positioning User Service (OPUS). With OPUS, users upload their GPS observations using a standard file format (e.g. RINEX) through the OPUS web interface. The uploaded data, processed through CORS, averages the baselines of three proximal CORS stations selected accordine to the user’s GPS data (Stone, 2006). Computed CORS coordinates and heights are returned to the user. Detailed metadata is also provided (e.g., overall RMSE, UTM and SPCS horizontal coordinates, and estimate of horizontal and vertical network accuracies).

Since its inception, each CORS collects GPS data 24-hrs a day, 7 days a week. Given the historical record of positional and height data at a CORS station, it is possible to compute and publish the station’s horizontal and vertical velocities. Usually, at least three years of GPS data is needed for a CORS velocity to be computed with a precision of approximately 1.0 mm/year (Snay, 2012, p. 167). Given sufficient spatial distribution of CORS velocity data, it is possible to describe plate motion, which allows CORS data to augment NAD83 into a 4-D datum (Snay, 2012). To make full use of a CORS data record, it is important to know the epoch. For example, NAD83(CORS96)’s epoch is January 1, 2002, which indicates the date when the published coordinates representing a consistent location of the CORS across space and time are valid — an important consideration in tectonically active areas (such as the western CONUS). NGS (2001) defines an epoch as a particular instant of time from which an event or a series of events is calculated. Epochs are expressed as a decimal year where a decimal year is equal to “year” plus “day-in-year/365”, or “366” for leap years (e.g., the epoch of May 5, 1991, the 128th day, is written as 1991.35). Knowing the epoch of coordinate values or heights is important when transforming coordinates from one epoch or datum to another. This transformation can be carried out using Horizontal Time Dependent Positioning (HTDP) software. Soler and Snay (2012) report that HTDP incorporates numerical models for both horizontal, crustal velocities throughout the United States and changes in the vertical due to earthquakes greater than 6.0 magnitude.

4.2.5. NAD83(NSRS2007)

Starting in 1997 and completed by 2005, NGS connected with each state to improve the HARN network through a readjustment to increase the density of HARN control stations. A goal of this process was to resolve inconsistencies between and accurately identify ellipsoid heights at HARN control stations via GNSS surveys and tie into the CORS network (Purcell and Potterfield, 2008). Once the resurvey was completed, NGS performed a simultaneous least-squares adjustment of all HARN data using leveling data from an additional 3,500 GPS surveys and produced consistent 3-D coordinates for approximately 70,000 passive reference stations (Snay, 2012). This realization was called NAD83(NSRS2007).

4.2.6. NAD83(2011)

In June 2012, NGS completed the National Adjustment of 2011 Project. This nationwide project involved an adjustment of NGS passive survey marks using GNSS. The adjustment used coordinates from the active NAD83(CORS96) network collected between January 1994 through April 2011. One notable change with this adjustment was the inclusion of separate plates as part of the unified NAD83 datum. CONUS states and territories were divided based on their tectonic plate and assigned individual datum tags and a single epoch: NAD83(2011), NAD83(PA11), and NAD83(MA11) refer to the North American, Pacific, and Mariana Plates, respectively. NAD83(2011) is relative to the North American Plate and defines coordinates for stations located in CONUS, Alaska and United States' territories in the Caribbean. NAD83(MA11) is associated with the Mariana Plate (east of the Philippine Plate) and defines coordinates in the Mariana Islands. NAD83(PA11) is assigned to the Pacific Plate and defines coordinates in Hawaii, American Samoa, the Marshall Islands and other territories found on this plate. The epoch for this realization was January 1, 2010 and was named NAD83(2011).

5. Modern 3-D Geometric Terrestrial Reference Systems

Today, a geometrical datum and the defining parameters are associated with a 3-D terrestrial reference system (TRS). A TRS is an abstract mathematical model that co-rotates with Earth and is described as ECEF where positions are expressed and have small temporal variations due to geophysical effects (plate motion, Earth tides, etc.). Snay and Soler (1999) describe four characteristics of a TRS. First, the X, Y, and Z axes of a 3-D coordinate system are oriented and located on or within Earth's surface. The origin of a 3-D coordinate system is located at Earth's geocenter of mass. Second, the adoption of the meter, the unit of length, should be the distance traveled by light in a vacuum during a time interval of exactly 1/299,792,458 seconds, which associates this distance to a theoretical rather than a physical measurement, which would contain uncertainties. Third, the shape and size of Earth's surface is approximated by a suitable reference ellipsoid. The center of the ellipsoid should be aligned to and positioned with the 3-D ECEF Cartesian coordinate system. Appropriate defining parameters (a, b, and f) are selected for a given reference ellipsoid. Fourth, a suitable model of Earth's gravity field is chosen (see DM-41 for further discussion on this point).

While a TRS is built on a conceptual model (e.g., the North American Datum), it is not physically attached to Earth's surface. A TRS is realized to a terrestrial reference frame (TRF) by a set of geodetic station coordinates and velocities on or near Earth's surface for a global geodetic network at a given epoch. Together, these station network coordinates and velocities constitute a TRF, which becomes the working model of the datum (e.g., NAD83(2011)).

5.1. World Geodetic System

In the late 1950s, the United States Department of Defense developed the World Geodetic System (WGS) to serve as a worldwide TRS. The first realization, WGS60, was based on the need for a unifying global TRS and was intended primarily for military applications (Snay, 2012). Satellite geodesy using Doppler (e.g., TRNSIT) and the collection of gravity data prompted WGS66 and WGS72 realizations (Slater and Malys, 1998). The geocentric reference ellipsoids for WGS72 and WGS84 were WGS72 and WGS84, respectively. The WGS84 originally used the GRS80 reference ellipsoid but has adopted slightly different ellipsoid parameters (see Table 2). Geocentric reference ellipsoids like GRS80 are more relevant in today's global mapping environment as GNSS satellites orbit around Earth's center of mass. Moreover, GNSS satellites return coordinates that are based upon global geocentric ellipsoids (e.g., WGS84), not local ones. WGS84 is the default TRS upon which

GNSS coordinates are delivered.

Slater and Malys (1998) report that the original realization of WGS84(1987) aligned with the Bureau International de l'Heure TRS84 (BTS84) but changed to GRS80 (which was first adopted with the NAD83(1986) realization). Developing WGS84 did not rely on any ground-based passive geodetic network. In 1994, WGS84(G703) was one of several terrestrial reference systems that were based completely on GNSS (GPS) observation data rather than Doppler data, and because of this, WGS84(G730) and other adjustments forward have agreed with ITRF91. Here, the "G" represents the fact that the realization was based on G (GPS), and 730 is the GPS week from 0 hours UTC on 2 January 1994. The following additional WGS84 "G" realizations (and the realization date) include G873 (1997.0), G1150 (2001), G1674 (2005.0), and G1762 (2005.0). These adjustments are important for high-precision calculations for satellite orbits but have little practical effect on typical high-accuracy survey applications (Van Sickle, 2017).

5.2. International Terrestrial Reference System

Earth's shape and size are dynamic. Thus, to be of use, a global TRF must also be dynamic. To meet this characteristic, the International Terrestrial Reference Frame (ITRF) has been realized. In 1988, under the auspices of the International Earth Rotation Service (IERS) the International Terrestrial Reference Frame (ITRF) was developed. ITRF is a geocentric datum, uses the GRS80 reference ellipsoid, and is considered the gold standard to which other TRFs are compared. The first realization was ITRF88 and was developed using a world-wide network of hundreds of control stations around the world whose positions and velocities were determined through satellite geodesy techniques such as GNSS (Figure 10). Since ITRF88, multiple IITRF realizations have been developed, the most recent being ITRF2014. The ITRF94, ITRF2000, and ITRF2008 realizations are aligned to WGS84(G873), WGS(G1150), and WGS84(G1674), respectively (Snay, 2012).

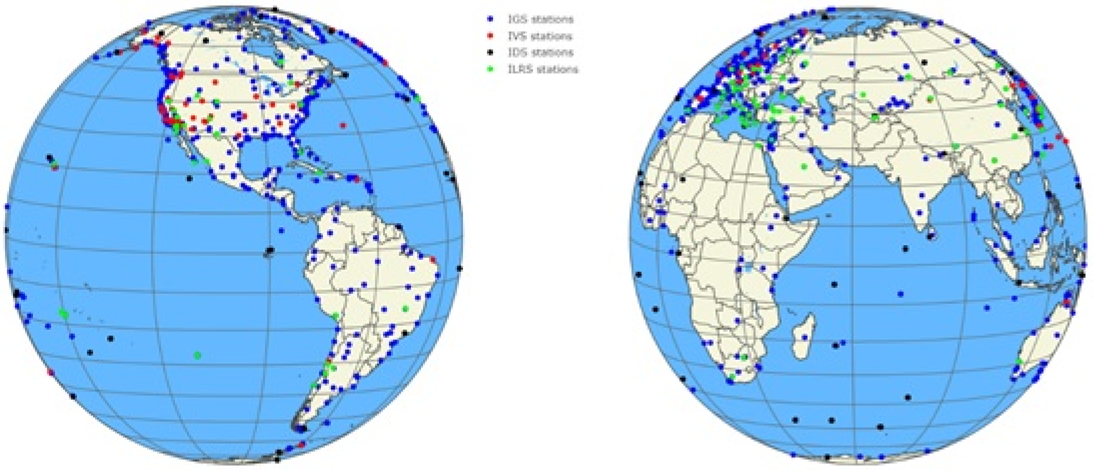


Figure 10. Maps showing the global distribution of ITRF control stations according to space-based measuring technology. Source: IERS

These ITRF adjustments used co-location, meaning that two or more stations have instruments that are operating and are surveyed in three dimensions (X, Y, and Z), using classical or GNSS geodesy, and that differential coordinates (DX, DY, DZ) are available

(Altamimi, et. al, 2009). Adjustments to ITRF consider plate motions by publishing velocities of the control stations by assuming that the “total angular momentum of Earth’s outer shell is zero…the total angular momentum on one plate is compensated by the angular momentum on the other plates (Snay and Soler, 2000, p. 2).” In other words, the absolute velocity is zero, and the reference frame and the plate move together. Points on the North American plate generally move horizontally at measurable rates. ITRF velocities for CONUS points on the North American plate range, for example, are from 10 – 20 mm/year but are larger for AK and HI (Snay and Soler, 2000, p. 2). In contrast, NAD83 asserts that the North American plate does not move, on average, relative to Earth’s interior. Using this assertion, points on NAD83 are “plate-fixed” and generally have no horizontal velocity relative to NAD83 (expect for locations near plate margins like California).

5.3. A 4-D North American Terrestrial Reference System

Throughout its history, NGS has provided horizontal control through the NSRS to its users through passive survey marks. Access to NAD27 and NAD83(1986) exemplified this approach for geodetic control, provided accuracies in the 1 to 10 m range, and was sufficient for most users of the NSRS (Smith et al., 2015). Using satellite geodesy, adjustments to NAD83 ushered in accuracies in the sub-cm level. Without an accurate and perpetually updated 3-D crustal motion model, there will never be complete coordination between the active and passive control of NAD 83. Deriving a new 4-D TRF model (e.g., a location’s latitude, longitude, ellipsoid height, velocity) was a driving force for the development of the North American Terrestrial Reference Frame 2022 (NATRF2022). Three general characteristics define the uniqueness of NATRF2022: Prioritizing the CORS network, plate rotation and the Euler Pole, and time dependency and velocity model.

5.3.1. Prioritizing the CORS Network

Under NATRF2022, the NCN will comprise separate but integral hierarchies. The first hierarchy comprises foundational NGS-owned or -operated stations. These foundational sites are equipped, monumented, and distributed by NGS to ensure the NSRS is accurately tied to the ITRF and accessible to GNSS users in the United States and its territories. The second hierarchy includes all of the sites remaining in the NCN. The location and maintenance of these sites falls upon each independent site operator. Not only does the CORS network publish ITRF coordinates, but they also supply coordinate positions and velocity data based on ITRF linking the NSRS with ITRF (NOAA, 2021).

5.3.2. Plate Rotation and Euler Pole

The idea of Earth’s crust being divided into moving and deforming plates (Figure 11) was proposed by Wegener, 1915. However, proof of the concept would not arrive until the 1970s when space-based technologies provided the needed evidence. Horizontal plate motion (relative to ITRF) can be modeled as a rotation about a geocentric axis passing through a point on Earth’s surface. Leonard Euler (“oiler”) quantified this idea through a Euler Pole Parameter (EPP) that includes latitude, longitude, and rotation in, for example, angular degrees per million years (NOAA, 2021). Determining the location of the Euler Pole can be quantified through years of GNSS observation data at control stations distributed across a plate. Figure 12 appears to show the North American plate to rotate counterclockwise about a Euler Pole whose center is in the Pacific Ocean near Ecuador. This horizontal motion is also called Eulerian motion (NOAA, 2021). Locations further from the

Euler Pole appear to be rotating faster, but all points are rotating at the same angular velocity.

If one removes the Eulerian motion shown in Figure 12, a residual motion (called non-Eulerian motion) can be seen. Figure 13 shows that these non-Eulerian motions vary across the eastern (A) and western (B) portions of CONUS. In Figure 13A, the motions generally appear to move northerly (although some randomness is visible) between 1mm and 3mm per year. This motion can be attributed to glacial isostatic adjustment (GIA) – the historic effect of retreating glaciers from this area. Figure 13B also illustrates a random motion in the interior portion of the western states with 10mm per year, but a much larger northwesterly direction is exhibited along the West Coast. This larger motion is attributed to the deformation between plates in this area and the North American Plate. Across CONUS, these non-Eulerian motions are not comparable, are time-dependent, and need to be modeled in some fashion.

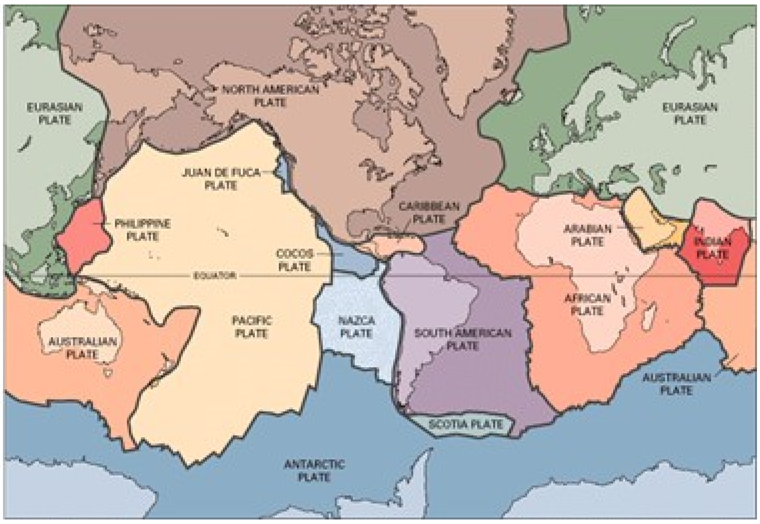


Figure 11. The major tectonic plates. Source: USGS.

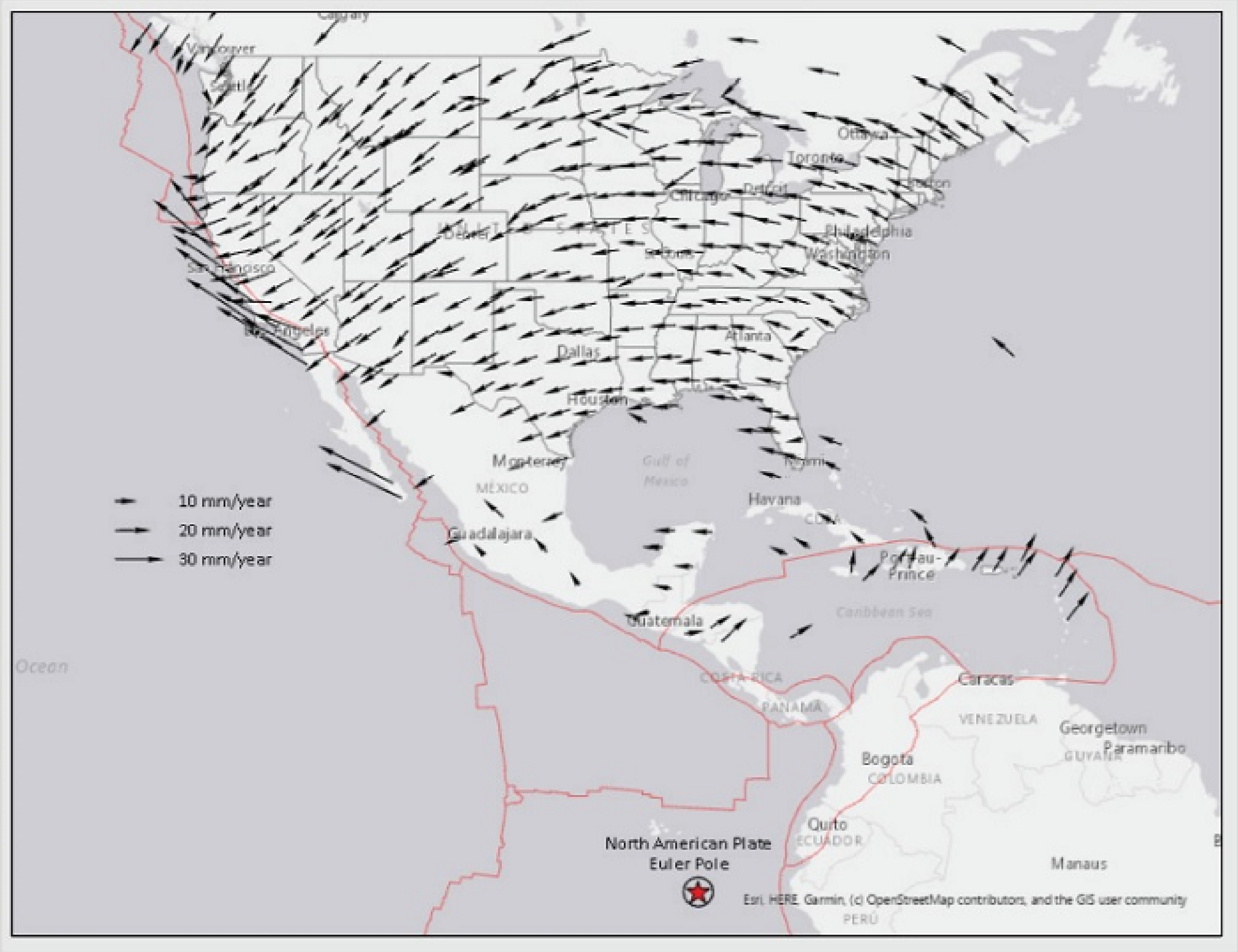


Figure 12. The Eulerian motion (rotation) of control stations about the Euler Pole for the North American Plate located near Ecuador. Source: NOAA, 2021.

Rather than combine CONUS with other states and territories into a single plate, NATRF2022 will contain four plate-fixed TRFs: NATRF2022, CATRF2022, PATRF2022, and MATRF2022. All TRFs will align to ITRF2020 at epoch 2020.00 (midnight, January 1st, 2020). This epoch defines the exact moment when coordinates from the four TRFs will agree with ITRF2020 (NOAA, 2021). Furthermore, each TRF rotates with the plates, is defined by separate EPP, and expresses coordinates in ITRF. By taking this approach, the Eulerian motion will be removed from each plate, leaving behind the non-Eulerian motion of a point’s coordinates. That non-Eulerian motion will become quantified as velocities within each TRF and will be modeled and related to ITRF2020, other TRFs, and across epochs.

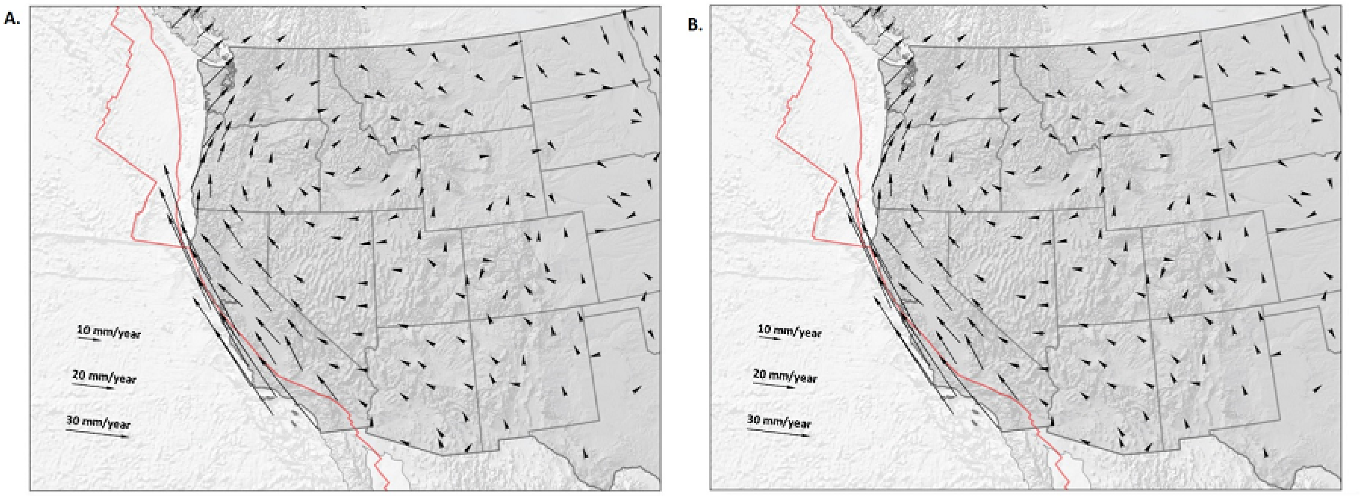


Figure 13. Non-Eulerian residual horizontal motion in the western half (A) and eastern half (B) of CONUS. Source: NOAA, 2021.

5.3.3. EPP and Velocity Model

A position cast on any of the four plate-fixed TRFs is a time-dependent Cartesian coordinate that can be transformed into one of the four TRFs or ITRF2020 with or without changing the epoch using a model. A plate rotation model for each TRF will be developed that relates a coordinate on a TRF to its corresponding ITRF2020 coordinate. This relationship will be defined by the EPPs about the three ITRF axes and quantified in a model called EPP2022 (NOAA, 2021). In areas of the continent where motion of the tectonic plate is fully characterized by plate rotation (e.g., Kansas), time-dependent coordinates will exhibit positional stability. All remaining velocities (including horizontal motions that are derived from adjacent tectonic plates, horizontal motions induced by GIA, other horizontal motions, and all vertical motions) will be captured by Intra-Frame Velocity Model (IFVM). The IFVM2022 will be useful to compare coordinates within one TRF, but at different epochs (NOAA, 2021). This is illustrated in Figure 14 where positional changes to a survey mark in North Dakota are expressed across the four plate-fixed TRFs and ITRF2022. The merging of the five lines to a common point in 2020 marks the epoch of 2020.00, where the positions aligned to TRFs and ITRF agree. Intuitively, the NATRF2022 trend line for this mark, located in the stable region of the North American Plate, illustrates horizontal stability, and the IFVM2022 is zero. In contrast, Figure 15 shows a mark located in the deforming Pacific Plate in California and shows an increase in the latitude separation from NATRF2022 stability based on the 2020.00 epoch. This separation will be modeled with IFVM2022.

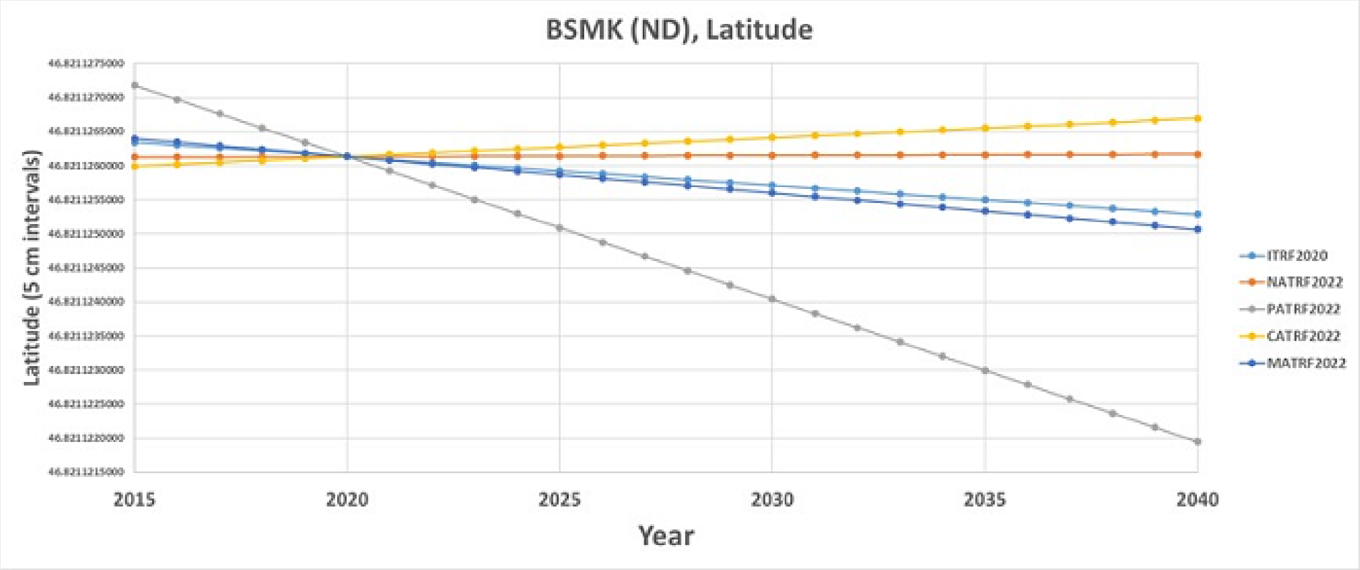


Figure 14. A graph showing the positional changes of a station in North Dakota across five TRFs from 2015 to 2040. Source: National Geodetic Survey.

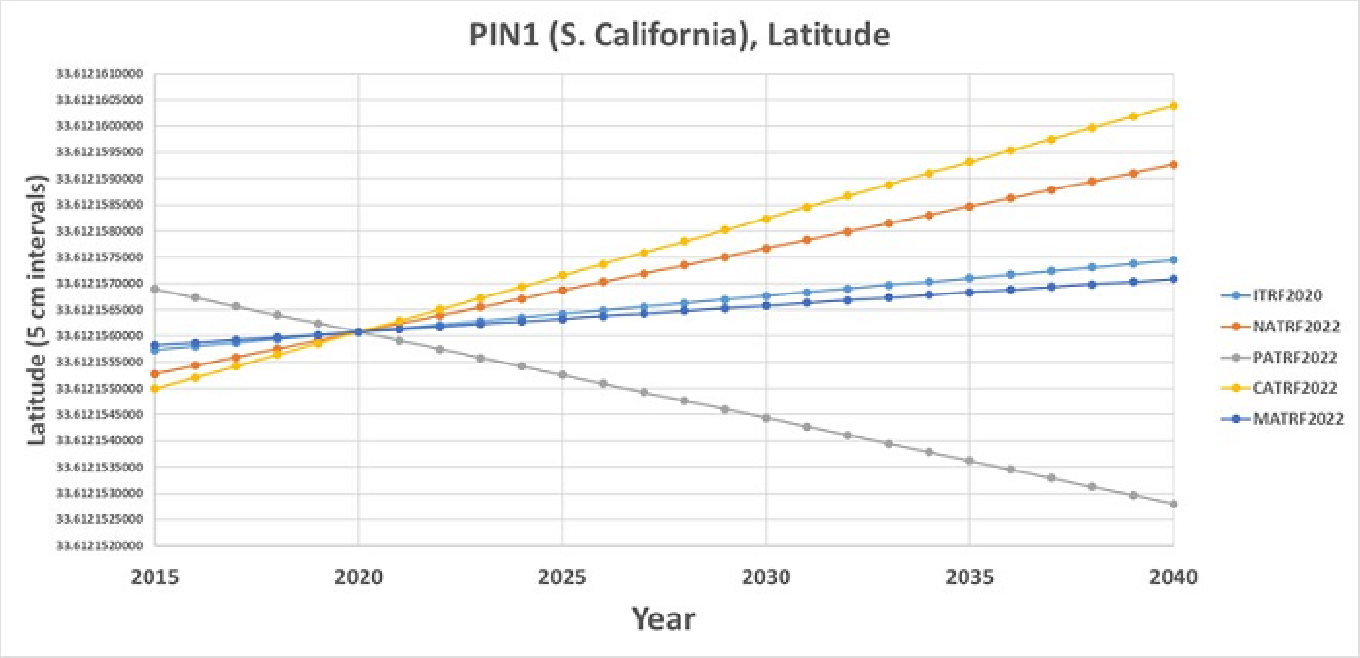


Figure 15. A graph showing the positional changes of a station in California across five TRFs from 2015 to 2040. Source: National Geodetic Survey.

References

Altamimi, Z., Angermann, D., Argus, D., Blewitt, G., Boucher, C., Chao, B., ... & Watkins, M. (2001). “The International Terrestrial Reference Frame.” International GPS Service for Geodynamics. 1999 IGS Annual Report. JPL, 400-978.

Anderle. R. (1983). Doppler Satellite Data Characteristics. Naval Surface Weapons Center, Dahlgren, VA.

Bowring, B. R. (1976). Transformation from spatial to geographical coordinates. Survey Review, XXIII (181), 323-327.

Chovitz, B. H. (1989). Datum Definition. In Schwarz, C. R. (editor), North American Datum of 1983. NOAA Professional Paper NOS 2, 81-85.

Dow, J., Neilan, R., & Rizos, C. (2009). “The International GNSS Service in a Changing Landscape of Global Navigation Satellite Systems.” Journal of Geodesy, 83(3), 191-198.

Dracup, J. (1994). Geodetic Surveys in the United States: The Beginning and the Next One Hundred Years, 1807-1940. US Department of Commerce, National Oceanic and Atmospheric Administration, National Ocean Service.

Ewing, C., and Mitchell, M. (1970). Introduction to Geodesy. American Elseiver Publishing Co., Inc. New York, NY.

Ghilani, C. D., & Wolf, P. R. (2012). Elementary Surveying: An Introduction to Geomatics. 14th Edition. Pearson Education. Upper Saddle River, NJ.

Gossett, F. (1959). Manual of Geodetic Triangulation (Vol. 4). US Government Printing Office. Washington, D.C.

Helmert, F. 1880 (1962). Die Mathematischen und Physikalischen Theorieen der Hoheren Geodasie, Part 1, 1884, Part 2. Minerva GmbH, Leipzig (Frankfurt / Main)

Hooijberg, M. (2008). Geometrical Geodesy: Using Information and Computer Technology. Springer, Heidelberg, Germany.

Iliffe, J.C. and Lott, R. (2000). Datums and Map Projections for Remote Sensing, GIS, and Surveying. 2nd Edition. Whittles Publishing, New York. 160p.

Lu, Z., Qu, Y., & Qiao, S. (2014). Geodesy. Springer: Berlin/Heidelberg, Germany.

Maling, D. H. (2013). Coordinate Systems and Map Projections. 1st Edition. Elmsford, NY: Pergamon Press.

Meyer, T. (2010). Introduction to Geometrical and Physical Geodesy: Foundations of Geomatics. Redlands, CA: Esri Press.

Mok, E. and Chao, J. (2001). Coordinate Systems and Datum. In: Chen, YQ., Lee, YC. (eds) Geographical Data Acquisition. Springer, Vienna.

Moritz, H. (1980). Geodetic Reference System 1980. Journal of Geodesy 74, 128–133.

National Geodetic Survey (NGS). (2001). Geodetic Glossary.

National Geodetic Survey (NGS). (2021). Blueprint for 2022, Part 2: Geopotential Coordinates. NOAA Technical Report NOS NGS 64. National Oceanic and Atmospheric Administration (NOAA).

Pearson, C. (2005). The National Spatial Reference System Readjustment of NAD 83. National Oceanic and Atmospheric Administration (NOAA).

Pursell, D., & Potterfield, M. (2008). NAD83 (NSRS2007) National Readjustment Final Report. NOAA Professional Paper NOS 60. National Geodetic Survey (NGS), National Oceanic and Atmospheric Administration (NOAA).

Schuh, H., & Behrend, D. (2012). VLBI: A Fascinating Technique for Geodesy and Astrometry. Journal of Geodynamics, 61, 68-80.

Schwarz, C. R. (Ed.). (1989). North American Datum of 1983 (NOAA Professional Paper No. 2). National Geodetic Survey, Charting and Geodetic Services, National Ocean Service.

Slater, J.A. and Malys, S. (1998). WGS 84 — Past, Present and Future. In: Brunner, F.K. (eds) Advances in Positioning and Reference Frames. International Association of Geodesy Symposia, vol 118. Springer, Berlin, Heidelberg

Smith, D. A., Roman, D. R., and Childers, V. A. (2015). Modernizing the Datums of the National Spatial Reference System. Marine Technology Society Journal. 49(2), 151-158.

Snay, R. A. (2012). Evolution of NAD 83 in the United States: Journey from 2D toward 4D. Journal of Surveying Engineering, 138(4), 161-171.

Snay, R. A., and Soler, T. (2000). Modern Terrestrial Reference Systems (Part 2): The Evolution of NAD 83. Professional Surveyor, 20(2), 1-2.

Snay, R., and Soler, T. (1999). Modern Terrestrial Reference Systems (Part 1). Professional Surveyor, 19(10), 32-33.

Snay, R., Cline, M., and Timmerman, E. (1989). “Crustal Motion Model.” In Schwarz, C. R. (Ed.) North American Datum of 1983. NOAA Professional Paper NOS 2, 141-152.

Sośnica, K. (2014). Determination of Precise Satellite Orbits and Geodetic Parameters Using Satellite Laser Ranging. Astronomical Institute, University of Bern, Switzerland.

Stone, W. (2006, April). The Evolution of the National Geodetic Survey's Continuously Operating Reference Station Network and Online Positioning User Service. In Proceedings of IEEE/ION PLANS 2006 (pp. 653-663). National Geodetic Survey (NGS), National Oceanic and Atmospheric Administration (NOAA).

Strange, W., and Love, J. (1991). “High Accuracy Reference Networks: A National Perspective.” Proceedings of the American Society of Civil Engineers Specialty Conference on Transportation Applications of GPS Positioning Strategy. ASCE, Sacramento, CA

Torge, W. and Muller, J. (2012). Geodesy (4th ed.). Berlin, Germany: De Gruyter.

Van Sickle, J. (2010). Basic GIS Coordinates. 3rd edition. CRC Press, Boca Raton, Florida.

Willis, P., Fagard, H., Ferrage, P., Lemoine, F. G., Noll, C. E., Noomen, R., ... & Valette, J. J. (2010). The International DORIS Service (IDS): Toward Maturity. Advances in Space Research, 45(12), 1408-1420.

Wolffhardt, T. (2017). Unearthing the Past to Forge the Future: Colin Mackenzie, the early colonial state, and the comprehensive survey of India. Book 4 of 6: Studies in British and Imperial History. Berghahn Books. Oxford, England.

Young, G. (1974). A mixed observational survey method. Bulletin Géodésique (1946-1975), 114(1), 349-363.