Vertical Datum and NAVD 88

Petr Vanicek

ABSTRACT. The paper explains what a vertical datum is and how it can be realized in nature. It briefly discusses some of the available techniques for vertical datum realization. Five possible approaches for datum realization within the context of North American Vertical Datum of 1988 (NAVD 88) are shown and the adoption of the most rigorous one is advocated. Finally, complications caused by temporal variations of the vertical datum are pointed out.

What is a Vertical Datum?

A vertical datum is a coordinate surface to which heights, taken as vertical coordinates of points, are referred. Three different kinds of vertical datums are used in geodesy:

1. The geoid—a reference surface for orthometric and dynamic heights

2. The quasi geoid—a reference surface for normal heights

3. The reference ellipsoid (horizontal datum)—a reference surface for geodetic (geometric) heights

We shall assume that the definitions of all those terms are known to the reader and will dwell here only on some finer points that may not be generally appreciated. For definitions, the reader is referred to Vanicek and Krakiwsky (1986).

The reference ellipsoid is used in various geodetic computations concerning heights. It is nevertheless inconvenient for practical uses because it completely disregards physical reality. As an example, for the NAD 83 reference ellipsoid the geodetic heights of the shoreline along the eastern coast of North America vary from -36 m in Chesapeake Bay to -13 m in Nova Scotia, Canada. The shoreline of Sri Lanka has geodetic heights of the order of 100 m.

The quasi geoid is a close relative of the geoid and, as such, is equally as practical to use. It has not been explicitly used in North America; therefore, we shall leave it out of further discussion here.

The geoid is one of the horizontal level surfaces shaped by Earth's gravity field. It is thus intuitively understood by everyone and, not surprisingly, it is the near universal choice of a vertical datum. The only conceptual complication is that an infinite number of horizontal surfaces exist, and one has to identify the one that is to be chosen as the datum.

Petr Vanicek is a professor of geodesy in the Department of Surveying Engineering, University of .New Brunswick, P.O. Box 4400, Fredericton, New Brunswick, Canada E3B 5A3. He is also editor of Manuscripta Geodaetica and Studia Geodaetica et Geophysica.

There are two practical options for identifying the desired horizontal surface:

1. The abstract option, whereby one specifies a constant value of Earth's gravity potential, W = Wo = const., which defines the geoid as one horizontal surface
2. The geometrical option, whereby one requires that the chosen horizontal surface (the geoid) approximates in a specific way the mean sealevel surface

In the context of the application discussed here (i.e., the geoid as a vertical datum), the geometrical option is normally selected.

We note as an aside that the term NAVD 88 is a misnomer. More properly we should be speaking of "North American Leveling Networks Readjusted/Redefined in 1988."

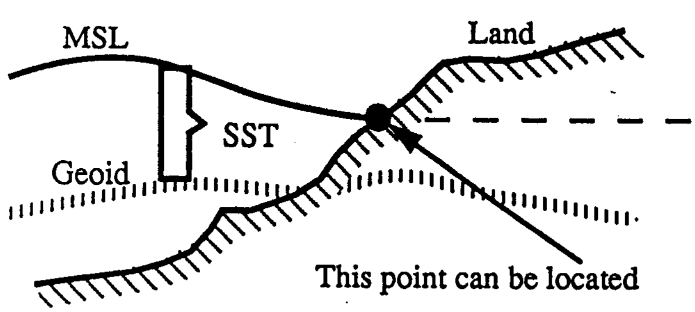
Realization of a Vertical Datum

While the concept of the geoid as a vertical datum is relatively clear, reality is much more problematic: The geoid does not exist in nature. We cannot simply start leveling from a point-of-zero orthometric height because such a point cannot be located in nature. Only approximate solutions to this problem exist. There are two such solutions available.

Direct Solution

A point of zero height (point on the datum) can be approximately located by observing sea-level variations by means of a tide gauge. Mean sea level (MSL) at a point is defined as a temporal average of sealevel observations for a specific epoch (Figure Ib). Spatially, MSL (for a specific epoch) is a surface (Figure la) that departs from the geoid by an amount called sea-surface topography (SST). We note the parallel between the definitions of SST and land topography as used in mapping.

SST is caused by sea dynamics and by prevailing meteorological phenomena (Montgomery 1937). Its magnitude is less than 2 m and when taken naturally, as the elevation of MSL above the geometrically de-



(a) spatial view

Figure 1. MSL and the geoid

Figure 1. MSL and the geoid.

fined geoid, it has generally a negative sign in polar regions and a positive sign in the tropical belt.

There are three different techniques available for the determination of SST:

1. Oceanic leveling (Forrester 1980; Levitus 1982)
2. Satellite altimetry combined with geoidal heights (Mather et al. 1976)
3. Local "zero frequency response analysis" for meteorological and other effects (Merry and Vanicek 1982)

The unpleasant fact is that none of these three techniques gives results with standard deviations smaller than 10 to 20 cm. Combinations of these techniques have been also considered by Cartwright (1984), but the potential accuracy improvement does not promise to be substantial. A slightly better accuracy can be obtained, however, for SST differences.

Indirect Solution

A point of zero height (point on the datum) also can be approximately located by computing the orthometric height H° of the terrain point directly above it (Figure 2). This orthometric height is computed from

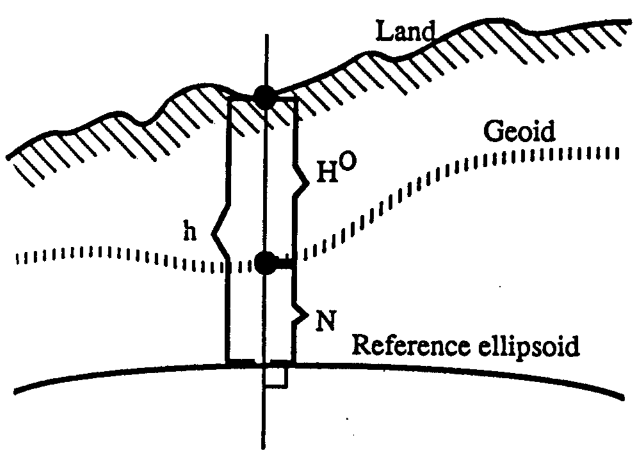
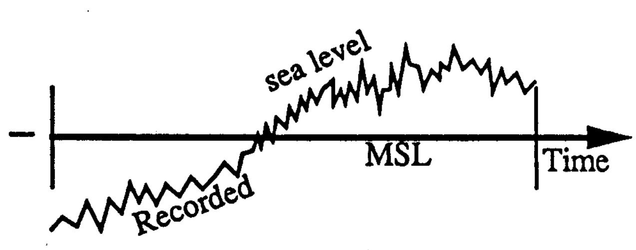


Figure 2. Indirect determination of a point on the datum.



(b) temporal view

the following expression:

$$H^{\circ} = h - N \tag{1}$$

where h is derived from the three-dimensional position (x, y, z) determined from space measurements, and N is the geoidal height.

While h can be established accurately, with a standard deviation considerably better than 5 cm from satellite laser ranging (SLR) (Committee on Geodesy 1985), N cannot. Even in areas well endowed with gravity and other kinds of data, the standard deviation of N is still only about 1 m (Vamcek et al. 1990).

It may be noted that the standard deviation of geoidal height differences AN between two points S km apart is typically less than 2S mm in nonmountainous areas of North America. This high relative accuracy, however, does not help solve the problems discussed here.

Vertical Datum for NAVD 88

Let us assume that the choice of vertical datum for the readjusted/redefined North American leveling networks (NALN) will be the geometrically denned geoid discussed previously. What practical choices are there for tying the new heights to this datum? It appears that there are five:

1. Use MSL at all reference tide gauges (i.e., those tide gauges that produce trustworthy data and are linked to the leveling network) as discrete datum points. This approach was used in the last adjustment of NALN in 1929 (Berry 1976). It tacitly assumes that SST at all reference tide gauges equals zero and that MSL therefore coincides with the geoid. Since this assumption is incorrect, this approach would introduce distortions amounting to many decimeters into the adjusted heights.
2. Initially hold an arbitrary height of one point of the network fixed in the adjustment while disregarding all the information provided by the tide gauges. Then shift the resulting heights by a constant value chosen so that the average

height of all reference tide gauges equals zero. This approach uses only the average of the reference tide gauges. The opportunity to use all valuable relative information contained in tide-gauge records for peripheral control of NALN is lost.

3. Estimate SST at reference tide gauges from realistic models and adjust NALN holding MSL-SST equal to zero for all reference tide gauges.
   This approach is clearly superior to options 1 and 2, but errors in estimated SST that may amount to a few decimeters—and to a much lesser degree also errors in computed MSL—would be forced into the network causing corresponding distortions of adjusted heights.
4. Do the same as in option 3, but instead of holding the estimated discrete datum points fixed at zero height, the points would be constrained by appropriate, realistic a priori weights. This approach utilizes correctly all information contained in reference tide-gauge records, prevents height distortions caused by errors in SST estimation and MSL determination, and checks the propagation of errors in NALN at the perimeter.
5. Do the same as in option 4, but add properly weighted differences of SLR station geodetic heights and high-precision geoidal heights. This approach would have all the good features of option 4 and would further homogenize the error distribution within the adjusted networks. This conclusion has been confirmed by simulation studies conducted by Tetreault (1987).

The last option is the most rigorous and should be used, therefore, in the readjustment of NALN. Mathematically, it is simple to implement. All that has to be done is to add to the usual system of observation equations (Vamcek and Krakiwsky 1986)

$$H_j - H_i = \Delta H_{ij}^{obs}, \text{ for all } i \text{ and } j$$

abbreviated as

$$AH = \Delta H^{obs} \tag{2}$$

another system of observation equations, namely, for all reference tide gauges $H_k = 0$ and for all SLR stations $H_l = H_l^{obs}$, abbreviated as

$$BH = H^{obs} \tag{3}$$

The least-squares solution $\hat{H}$ for the adjusted heights is then obtained from the following system of normal equations:

$$(A^T C_{\Delta H}^{-1} A + B^T C_H^{-1} B)\hat{H} = A^T C_{\Delta H}^{-1} \Delta H^{obs} + B^T C_H^{-1} H^{obs} \tag{4}$$

where C**AH** is the (diagonal) covariance matrix of observed height differences, and CH is the covariance matrix of tide gauge and SLR station heights. For the reason of theoretical rigor, geopotential numbers rather than observed heights should be used in the adjustment. It is known that observed height differences along a loop do not close theoretically to zero, while geopotential number differences with actual gravity variations accounted for do (Vamcek and Krakiwsky 1986). But this point has no direct bearing on the argument presented here.

In the options presented previously, we have not discussed techniques for obtaining MSL at reference tide gauges, or the epochs for which they should be computed. There is a reasonable agreement among geodesists how this should be done, and the interested reader may wish to read on this topic in Vanicek et al. (1984). It also should be mentioned that an improvement in MSL determination accuracy may be achieved by using differenced tide-gauge records. An exhaustive discussion of this subject can be found in Carrera et al. (1990).

SST at the reference tide gauges along the coast of North America should be studied, modeled, assessed, and taken into account in the forthcoming readjustment of NALN. It also would be beneficial to the overall effort if the geodetic height differences of SLR stations and their requisite precise geoidal height differences were also produced, assessed, and included in the readjustment. One suspects, however, that leaving the SLR stations out (i.e., using option 4 instead of 5) would not be as detrimental as using options 1 or 2 (i.e., leaving the tide-gauge records untapped). Just how far we can go towards the ideal solution will have to be decided by the U.S. National Geodetic Survey and the Geodetic Survey Division of Canada, the agencies responsible for carrying out the operation. Only these agencies are privy to all available information and are aware of existing temporal and financial constraints.

Time Variations of the Vertical Datum

In addition to the choices described previously, the operational agendes must come to grips with the following problem: All level surfaces of Earth's gravity field, including the geoid, change shape with time. This is in response to the continuously changing mass distribution within Earth. The most conspicuous and best-known example of these changes is the Earth body tide, responsible for the "breathing" of the geoid by about ±50 cm twice a day. Tidal variations, being either periodic or permanent in nature, do not present much of a problem. They can be averaged out in time and the permanent part accounted for.

Much more troublesome are secular variations caused by a variety of other phenomena. Of these, we should mention the global eustatic sea-level rise, estimated now to be about 1.5 mm/year. When the geoid is denned geometrically, this relentless rise of the vertical datum makes actual orthometric (and dynamic) heights all around the world shrink by 1.5 mm/year.

Paralleling global changes are well-documented cases of regional changes of the geoid such as those in the Hudson Bay region. Here, the geoid uplift reaches a maximum of about 1.5 mm/year in the center of Hudson Bay (Sjoberg et al. 1990), and while diminishing towards the perimeter, it affects an area several thousand kilometers across.

The temporal variations of the vertical datum take place in addition to real vertical crustal movements (Vanicek et al. 1987), which, of course, also affect heights. To our knowledge, no official policy has been adopted on how to deal with these changes. Some suggestions for a solution can be found in Carrera (1984).

ACKNOWLEDGMENTS

This paper, in an abbreviated form, was presented at the 1990 ACSM/ASPRS Annual Convention in Denver, Colorado. The manuscript was prepared while the author held the Alexander von Humboldt Distinguished Senior Scientist award at the Geodetic Institute of the University of Stuttgart, Federal Republic of Germany. The generous sponsorship of the Alexander von Humboldt foundation and the cordial cooperation of the host, E.W. Grafarend, are gratefully acknowledged. I also wish to thank C.R. Schwarz and two anonymous reviewers for their helpful comments on an earlier draft of this paper. W. Wells has done her usual flawless job editing and word processing several versions of this paper.

REFERENCES

Berry, R.M. 1976. "History of Geodetic Leveling in the United States." Surveying and Mapping, vol. 36, no. 2, pp. 137-153.

Carrera, G. 1984. "Heights on a Deforming Earth." Department of Surveying Engineering Technical Report No. 107, University of New Brunswick, Fredericton, New Brunswick, Canada.

Carrera, G., P. Vanicek, and M.R. Craymer. 1990. "The Map of Recent Vertical Crustal Movement in Canada." Department of Surveying Engineering Technical Report, University of New Brunswick, Fredericton, New Brunswick/ Canada.

Cartwright, D. 1984. Communication to Working Group on Vertical Datum of the United States. Washington, D.C.: NAS Committee on Geodesy.

Committee on Geodesy. 1985. Geodesy: A Look to the Future. Washington, D.C: National Academy Press.

Forrester, W.D. 1980. "Principles of Oceanographic Leveling." Proceedings of the Second International Symposium on Problems Related to the Redefinition of the North American Vertical Geodetic Network, Ottawa, Ontario/ Canada, pp. 125-132. Ottawa: Canadian Institute of Surveying.

Levitus, S. 1982. "Climatological Atlas of the World Oceans," NOAA Professional Paper No. 13. Washington, D.C.: U.S. Department of Commerce.

Mather, R.S., R. Coleman, and O.L. Colombo. 1976. "On the Recovery of Long Wave Features of the Sea-Surface Topography from Satellite Altimetry." University of New South Wales Unisurv G24, New South Wales, Australia, pp. 21-26.

Merry, C.L., and P. Vanicek. 1982. "Investigation of Local Variations of Sea-Surface Topography." Marine Geodesy, vol. 7, pp. 101-126.

Montgomery, R.B. 1937. "Fluctuations in Monthly Sea Level on Eastern U.S. Coast as Related to Dynamics of Western North Atlantic Ocean." Journal of Marine Research, vol. 1, no. 2, pp. 165-185.

Sjoberg, L.E., P. Vanicek/ and M. Kwimbere. 1990. "Estimates of Present Rates of Land and Geoid Uplift in Eastern North America." Manuscripta Geodaetica, vol. 15, no. 5, pp. 261-272.

Tetreault, P. 1987. "An Investigation of Possible Constraints for the Definition of the 1988 North American Height Network." Master of Applied Sciences thesis. Department of Civil Engineering, Survey Science, University of Toronto, Ontario.

Vanicek, P., and E.J. Krakiwsky. 1986. Geodesy: The Concepts, 2nd rev. ed. Amsterdam, The Netherlands: North Holland.

Vanicek/ P./ G.H. Carrera/ and M.R. Craymer. 1984. "Corrections for Systematic Errors in the Canadian Leveling Network." Survey Science Technical Report No. 10/ University of Toronto, Erindale Campus, Mississauga, Ontario.

Vanicek, P., P. Ong, and C. Zhang. 1990. "The New Canadian Gravimetric Geoid." Department of Surveying Engineering Technical Report, University of New Brunswick/ Fredericton/ New Brunswick, Canada.

Vanicek, P., P. Cross, J. Hannah, L. Hradiiek, R. Kelm, J. Makinen/ C.L. Merry/ L.E. Sjoberg/ R.R. Steeves/ and D.B. Zilkoski. 1987. "Four-Dimensional Geodetic Positioning: Report of the IAG SSG 4.96." Manuscripta Geodaetica, vol. 12, no. 3, pp. 147-222.