History of NAD27 And Subsequent ReadjustmentsIn Ontario

BY Z. A. JIWANI AND K. M. KELLY

Introduction

The stage was set for geodesists when Isaac Newton first announced in his Principia of 1637 that the figure of the earth most closely approximated an oblate spheroid - that is a surface which departs from an exactly spherical form by a flattening at the poles. Shortly after this deduction, geodesists packed their bags full of mathematical and instrumental paraphernalia and searched to see whether such a theory was valid. The first attempt at verification of Newton's hypothesis by actual field measurement was initiated by the French Academy of Sciences. They despatched two expeditions, one to Peru in 1735, and the other to Lapland in 1736. The task of the two expeditions was to determine the length of a degree of latitude near the equator and the arctic circle respectively. Sure enough, Newton was right - the crude results showed that the northern degree was the greater. Ever since, geodesists have laboriously toiled to determine a mathematical spheroidal surface that "best" fits the entire earth (but more often localized portions of it). Even to this day the search goes on.

The Figure of the Earth

THE IDEA of a geodetic datum will be better understood if we first define what is meant by the "figure of the earth". Indeed, all land and marine surveys are made upon the surface of the earth and we need to become familiar with the various properties and definitions of this surface.

The physical figure of the earth consists of that surface upon which we dwell and perform all our survey operations. This is the surface depicted on an ordinary topographical map. The geoidal figure of the earth is the actual shape of the earth as described by an equipotential surface, or a surface that is everywhere perpendicular to the direction of gravity. It is the surface which most closely coincides with the mean elevation of the oceans. Local topographical features, which in effect. amount to variations in the density of crustal material, produce local irregularities in the geoid.

Because the physical and geoidal surfaces are highly irregular, and surveys must be performed on these surfaces, it is necessary to relate our surveys to a regular, mathematical surface that is better adapted to mathematical computation, and that will provide a uniform system of reference. That mathematical surface which most closely approximates the geoid is an ellipsoid of revolution, also referred to as the spheroid or ellipsoid. Horrendously complex survey computations would be necessary if survey data was reduced to the geoid; however, the spheroid affords us with a surface upon which relatively simple (well maybe!), mathematically accurate survey computations can be made.

Before the advent of earth satellites and the ease and rapidity with which gravity surveys can be made today, geodesists had to determine a suitable spheroid (a suitable spheroid is one in which the deviations from the geoid are as small as possible) primarily from astronomical and triangulation data obtained from the measurements of meridian arcs. Gravity observations aid in accurately determining the shape of the geoid, and thus the general shape of the spheroid and its polar flattening, but do not provide much useful information or the actual dimensions of the spheroid.

The relationship between these three surfaces - the physical, the geoidal, and the spheroidal - are shown in Figure 1.

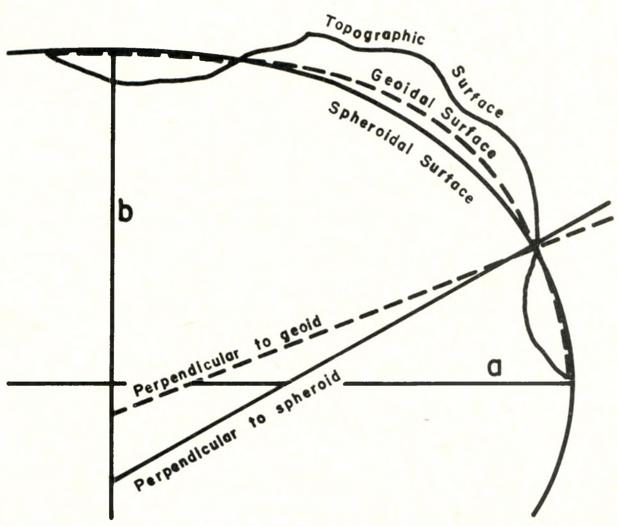

Fig. 1

A variety of spheroidal surfaces have been derived by geodesists in the past. The most important are those computed by *Bessel* in 1841, *Clarke* in 1866 and 1880 and *Hayford* in 1909. The *Hayford* spheroid was universally adopted in 1924 by the International Union of Geodesy and Geophysics as a basis for scientific study. Since then the IUGG has proposed that the New International 1980 spheroid be adopted for a new North American Datum to be established in 1983. (Now postponed to 1985 but still to be referred to as NAD83.) At present, however, the *Clarke* spheroid provides the basis for our current 1927 North American Datum. Hence the *Clarke* 1866 spheroid is, at present, the mathematical surface upon which all geodetic survey computations are made and their resulting positional values and relationships recorded.

Horizontal Geodetic Datum

Numerous definitions, some simple, some complex and some conflicting, exist for the term geodetic datum. From the standpoint of the surveyor, an ideal geodetic datum would be "... a system of stable survey monuments at some convenient spacing with unique and invariable coordinates appreciably more accurate than any survey work that they might be used to control" [*Jones*, 1973]. To date the best approximation of this ideal datum at least in Southern Ontario, has been the 1927 North American Datum (NAD27).

The National Oceanic and Atmospheric Administration 1973, in its publication, *The North American Datum*, defines this datum as follows: "The North American Datum (NAD) is the base of reference on the North American Continent to which all geodetic control surveys of the United States, Canada and Mexico are related. Collectively, these national network surveys are commonly referred to as the North American Datum. The geodetic datum, or starting point, for the North American Datum is a monumented point in Kansas called MEADES RANCH. The latitude and longitude for MEADES RANCH are referenced to the semi-axes of the *Clarke* Spheroid of 1866. The *Clarke* Spheroid of 1866 mathematically defines the spheroid which most closely fits the geoid (the actual shape of the earth) in North America."

The history of this datum, however, indicates that it has fallen far short of the ideal datum stated above. In order to realize just why NAD27 has proven inadequate we need to define the term "geodetic datum" more precisely as well as gain some insight as to how such a datum is established.

A geodetic datum is a mathematical means to adjust and record a network of control stations [*Fischer*, 1974]. It is the base of reference for the computation of horizontal control surveys in which the curvature of the earth is considered. The datum is basically composed of five quantities (more rigorous definitions define more quantities, i.e. *Bomford*, 1971 defines eight quantities): an adopted latitude and longitude of an initial or datum point, an adopted azimuth of a line from this point, and two constants (polar and equatorial radii) defining the reference spheroid. Note that the choice of these last two spheroidal parameters is essentially arbitrary, and hence any convenient spheroid that most closely approximates the geoid in a given area may be adopted. Indeed this closeness of fit between the geoid and spheroid constitutes one of the primary criteria for the mathematical computation of a reference spheroid for a particular portion of the earth's surface. However, back to datum definition.

Traditionally, the defining of a geodetic datum upon which all geodetic survey operations would be referenced and subsequently adjusted consisted of selecting a convenient point on the surface of the earth as a datum or origin point, adopting its astronomic coordinates as geodetic coordinates, observing an astronomic azimuth as an initial direction, choosing a convenient reference spheroid as a computational surface, and surveying to your hearts content!

The fact that the elevation of the datum point above the geoid is not specified in the above datum definition, implies that a convenient value of zero metres can be adopted for the geoid-spheroid separation at this point. Also, the adoption of astronomic coordinates as geodetic coordinates for the datum point, implies that the deflection of the vertical is zero at this point. Since, deflection of the vertical is simply the angular value of the deflection between the normal to the geoid and the normal to the reference spheroid (see Figure 1), then a zero deflection means that the geoid and spheroid are either tangent at the datum point (if the height above the geoid is zero) or parallel at the datum point (if the height above the geoid is not zero).

Although the position of the reference spheroid chosen for NAD27, (i.e. the *Clarke* 1866 spheroid) did not depart significantly from the geoid; it was assumed that the closeness of the two surfaces would ensure that errors caused by neglecting the geoid-spheroid separations in horizontal control computations would be negligible. It was the most logical assumption to make at that time when the shape of the geoid was more guessed at than known [*Hayford*, 1909].

In fact, when NAD27 was computed there were not sufficient astronomic or gravity measurements to determine geoid-spheroid separations, not only at the datum point but throughout the entire triangulation network used in those computations. In addition there were not sufficient astronomic observations to accurately determine the deflection of the vertical at the datum point. [*Vanicek and Wells*, 1974.] Hence, for all intents and purposes, zero values were assumed for both geoid-spheroid separation and deflection of the vertical at the datum point in the establishment of NAD27.

With the aid and abundance of more accurate gravity, and astro-geodetic deflection data and good geoid charts, these assumptions are no longer valid. As we will see later, it is largely on the basis of these assumptions inherent in NAD27, along with the increased accuracy of making geodetic observations, that the failure of this datum as a suitable reference system for geodetic surveying has become acute. Before we examine the effects of the NAD27 approximations on horizontal control networks in Canada, a brief historical background of NAD27 is in order.

History of North American Datum

Early triangulation in the United States consisted of pockets of control networks widely scattered throughout the continent. Each network had its own unique datum which was determined by astronomic observations for that particular area. As the triangulation expanded into other areas the detached portions grew until eventually they were linked to form a continuous chain of triangulation. Obviously, since each network was based on a different datum, discordant values for the positions of the junction points resulted. It became necessary to abandon each unique datum and produce one common datum to which all the triangulation could be referenced.

The first official datum in the United States was the New England Datum, adopted in 1879, and referenced to the *Clarke* Spheroid of 1866. This datum had its origin as triangulation station PRINCIPIO in Maryland and was based on adjustments of networks in the northern and eastern states. An adjustment in 1901, of the New England Datum, which by then extended into the south and west, resulted in a new datum designated the United States Standard Datum, and had as its origin triangulation station, MEADES RANCH, in Kansas.

As much of the triangulation of the Geodetic Survey of Canada was connected to that of the United States, it was formally agreed in 1913, by both Canada and Mexico, that these countries would base their triangulation networks on the same standard datum. Because of the international character of the datum, it was renamed the North American Datum. Stations are said to be on the North American Datum when they are connected with the station MEADES RANCH by continuous triangulation, or by precise traverse, and when all associated geodetic computations are subsequently reduced to the *Clarke* Spheroid of 1866 [*MacTavish*, 1952].

Serious problems of fitting new surveys into the existing network arose by the middle 1920's. As a result, a five year period from 1927 to 1932 was designated, and all available primary data (approximately 25,000 monumented control points) were readjusted into the system known as the 1927 North American Datum (NAD27). The arc along the 49th parallel in western Canada, area triangulation in Ontario and Southern Quebec, and a loop in New Brunswick were included in the adjustments for NAD27.

After adjustment of the eastern part was completed, a discrepancy of approximately 10 metres in latitude along the U.S. border in northern Michigan was detected. The U.S. portion of the network in Wisconsin and Michigan was subsequently readjusted to absorb this discrepancy. It should be noted that the 10 m discrepancy was not eliminated but simply distributed over a large portion of the local network. The result of this local adjustment was to produce a set of coordinates that are compatible with each other but not necessarily with triangulation networks very distant from the northern Michigan area.

As Canada built upon the NAD27 networks and spread northwards from MEADES RANCH, holding the NAD27 coordinates fixed, problems similar to those encountered in northern Michigan arose. Distortions of magnitudes much larger than those expected due to normal error propagation were encountered when attempts were made to fit new first-order networks with previously established stations. The primary cause of these distortions stemmed from the assumptions made in the definition of NAD27 and the process of constraining new measurements to fit old coordinates. The remedy was the infamous "regional readjustment".

Distortions in NAD27

As mentioned earlier certain approximations were made when NAD27 was originally defined. One of these was the adoption of zero metres for the value of the geoid-spheroid separation in North America. This approximation was based on the assumption that since the *Clarke* Spheroid of 1886 so closely fitted the geoid in North America, then errors caused by neglecting the geoid-spheroid separations in horizontal control survey computations would be negligible. However, the errors caused by disregarding geoid-spheroid separation in control survey computations performed on NAD27 have since been found to be quite significant and, in fact, are one of the primary causes for the relative scale distortions inherent in NAD27 networks and hence horizontal control networks in Canada.

Geoid-spheroid separations are important factors in the reduction of measured distances to the reference spheroid. In order to accurately reduce measured distances to the reference spheroid, the heights of the end points above the spheroid must be known. As shown in Figure 2, if orthometric or geoidal heights h_1, h_2 obtained from ordinary levelling are used in the reductions, and the geoid-spheroid separations N_1, N_2 are neglected then the measured distance ι, will not be reduced to the reference spheroid as S, but will be reduced to some other spheroidal surface as S_o. The heights that should be used for proper reduction of measured distances to the spheroid are $H_1 = h_1 + N_1$ and $H_2 = h_2 + N_2$. Hence, to correctly reduce measured distances to the reference spheroid, geoid-spheroid separation must be known.

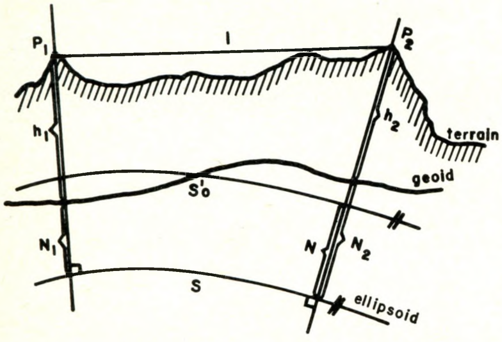

Figure 2 [Merry and Vanicek, 1973]

The approximate distance reduction used for the Canadian first-order horizontal control networks amounts to using the geoidal heights h_1, and h_2 instead of spheroidal heights H_1, H_2 [*Merry and Vanicek*, 1973]. The error in neglecting geoidal height in distance reduction is a systematic scale error whose sign rarely changes over large regions and therefore, the error will have a tendency to accumulate. Geoidal models of Canada indicate that geoid-spheroid separation over most of the continent is around $\pm$ 10m, thereby implying a scale error over all first-order triangulation of about 2 PPM. This scale error is significant when one considers the length of the primary triangulation network across Canada.

Another approximation made in defining NAD27 is that horizontal angles measured at a station are assumed to be on the reference spheroid - that is, no corrections for deflecting of the vertical, skew normals or normal section separations are applied. The most significant of these corrections is that of the deflection of the vertical. As shown in Figure 3, a theodolite set-up at station P, will be dislevelled

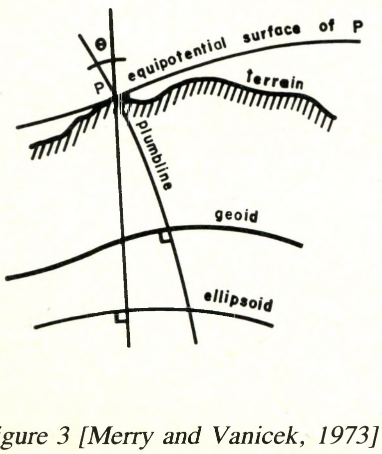

Figure 3 [Merry and Vanicek, 1973]

with respect to the reference spheroid by an amount equivalent to the deflection of the vertical, θ. If this is the case then directions and/or angles measured at station P cannot be considered to be on the spheroid. In order to reduce the observations to the spheroid the appropriate deflection corrections must be applied.

Experimentation by the Geodetic Survey of Canada has shown that even very small deflection corrections, as exhibited throughout Canada, can affect a large scale control survey. Although, the deflection correction at a single station may be small, it is capable of producing an appreciable regional distortion of a horizontal network of up to a few PPMs [*Merry and Vanicek*, 1973].

Hence with the increased accuracy of making geodetic observations, it is not surprising that control network distortions, caused by approximations in defining and computing NAD27 have affected even the work of the local surveyor. For this reason, regional readjustments have been undertaken in Canada and the U.S. in an attempt to minimize the effects of these distortions. Ontario is not without its share of such readjustments.

Regional Readjustment

Many of the distortions in control networks described earlier went undetected until they were finally isolated through the process of fitting and adjusting. Regional readjustments were then performed to distribute discrepancies over larger areas so that network distortion could be minimized. These regional adjustments are a patching process that does not eliminate errors, but rather distributes and minimizes their impact.

In Canada several regional adjustments have been undertaken in areas where distortions in NAD27 coordinates were unacceptable and to improve relative accuracies to more acceptable standards for local users. In this case a readjustment would sacrifice overall continuity for local network accuracy [*Jones*, 1973]. Such regional readjustments have been made in Prince Edward Island, Nova Scotia, southern Quebec, and in southern and northern Ontario.

In some cases, such as southern Quebec, the readjustments define an independent datum. This is effected by arbitrarily holding one station fixed as the origin at the NAD27 coordinates and introducing additional Laplace azimuth control into the adjustment [*Jones*, 1973]. However, the Ontario 1974 readjustment had many NAD27 values around the perimeter of the adjustment area held fixed and the remainder of the provincial adjustment was constrained by these fixed first-order stations. (See Figure 4.) Hence, the Ontario 1974 readjustment can be said to be a true regional readjustment with the resulting coordinate values referenced to NAD27; even though these values are often referred to as "1974 datum" values.

Figure 4
Shaded portions indicate areas held fixed in the 1974 readjustment

Based upon the technology and additional field measurements available at the time, the 1974 readjustment afforded a network of primary horizontal control stations with distortions sufficiently absorbed and minimized so as to produce for local users a set of coordinate values of much higher accuracy standards. Since then, however, modern satellite technology has made it possible to produce even more accurate coordinate values for these same primary triangulation stations.

Between 1974 and 1976 the Geodetic Survey of Canada performed two adjustments to test their adjustment techniques and to evaluate and compare the terrestrial triangulation and satellite Doppler networks in Canada. The first of these two adjustments has been labelled the MAY76 Test Adjustment according to its completion date [*Beattie*, et al, 1978]. In Ontario, coordinate values based on the MAY76 Test Adjustment are often referred to as MAY76 coordinate values. However, similar to the 1974 readjustment, the

MAY76 Test Adjustment does not define a new datum.

The test adjustment utilized a framework of triangulation stations extending across the continent whose coordinate values were determined by both conventional triangulation and satellite Doppler methods. The Doppler positions of these framework stations were transformed to the NAD27 system by applying appropriate datum shift parameters and they also provided the constraints for the adjustment. The adjustment was then performed on the *Clarke* Spheroid of 1866. Additional primary and some secondary control were then adjusted to the framework stations by holding the MAY76 adjustment coordinates fixed.

Hence coordinate values for control stations said to be on the MAY76 datum are in actual fact improved NAD27 coordinate values. The improved control station coordinates result from the fact that Doppler satellite observations, which provide absolute or geocentric positions, were included in the adjustment and served to remove some of the scale and orientation distortions inherent in the existing primary framework. Hence, the use of MAY76 Test Adjustment coordinate values may afford a better fit into the existing network of densification surveys performed between fixed control stations which have MAY76 Test Adjustment coordinates attached to them.

As a result two sets of coordinate values exist for many primary and lower-order control stations in Ontario. The two sets, however, despite their misleading labels of "1974 datum" and "MAY76 datum" respectively, are both based on the 1927 North American Datum. One may go only as far as to say that the NAD27 coordinate values produced by the above described adjustments are better than the original NAD27 values before these adjustments - that improvement being the removal of some of the scale and orientation distortions present prior to the respective adjustments.

With the availability of more accurate gravity, astro-geodetic and satellite data, geodesists have ample ammunition to attempt to redefine that spheroid which best fits the global geoid and hence determine a more refined model for the figure of the earth. At the same time, the numerous regional readjustments that have resulted in frequent coordinate changes has confirmed the obvious - that a comprehensive readjustment of all Canadian - indeed, all North American - horizontal control networks must be undertaken. The joint project of readjusting on a redefined datum has been confirmed and appropriately termed Redefinition and Readjustment of North American Geodetic Networks - NAD83. This will be the subject of the next column.

●

References

Babbage, G., *The Role of Geodetic Survey in Canadian Geodesy*. Paper presented at the CIS Centennial Meeting.

Beattie, D. S., Blais, J. A. R., Pinch, M. C., *Test Adjustments of the Canadian Primary Horizontal Network*. Proceedings of the Second International Symposium on Problems Related to the Redefinition of North American Geodetic Networks, Arlington, VA. April 24-28.

Bomford, G., 1971, *Geodesy. Third Edition*, Oxford University Press, Great Britain.

Clarke, D., 1951. *Plane and Geodetic Surveying, II: Higher Surveying, Fourth Edition*, Constable & Company Ltd., London.

Fischer, Irene, 1974. *The Role of the Geoid in Datum Definitions*. Proceedings of the International Symposium on Problems Related to the Redefinition North American Geodetic Networks, Fredericton, New Brunswick, May 20-25.

Jones, Harold E., September, 1973. *Geodetic Datums in Canada. THE CANADIAN SURVEYOR*, (27), 2.

MacTavish, W. H., October, 1952. *The 1927 North American Datum, THE CANADIAN SURVEYOR*, (11), 2.

McLellan, C. D., 1974. *Geodetic Networks in Canada*. Proceedings of the International Symposium on Problems Related to the Redefinitions of North American Geodetic Networks, Fredericton, New Brunswick, May 20-25.

Merry, C. L., Vanicek, P., March, 1973. *Horizontal Control and the Geoid in Canada. THE CANADIAN SURVEYOR*, (27) 1.

National Acadamy of Sciences, 1971. *North American Datum*.

National Oceanic and Atmospheric Administration. 1973. *The North American Datum*.

Vanicek, P., Wells, D. E., 1974. *Positioning of Geodetic Datums*. Proceedings of the International Symposium on Problems Related to the Redefinition of North American Geodetic Networks, Fredericton, New Brunswick, May 20-25.

Discussions with Mr. Paul Henderson of the Geodetic Survey of Canada.

Horizontal and vertical field control surveys have been recently completed or are in progress in the areas listed below. Some will be completed by the time this OLS Quarterly is received. Users and those in related fields may obtain more information by contacting the authors at the Ministry of Natural Resources, Surveys and Mapping Branch, Geographical Referencing Section, Toronto [Tel. (416) 965-4789].

CONTROL SURVEYS RECENTLY COMPLETED OR IN PROGRESS

AREA	TYPE OF SURVEY	TENTATIVE DATE OF COMPLETION
Alexandria, Town of	H/V	April, 1983
Almonte, Town of	H/V	November, 1982
Barry's Bay, Village of	H/V	April, 1983
Bath, Village of	H/V	April, 1983
Bruce Mines, Town of	H	Completed
Cache Bay, Town of	H/V	Completed
Campbellford, Town of	H	Completed
Casimir, Jennings and Appleby, Township of	H/V	Completed
Cardinal, Village of	H/V	January, 1983
Carleton Place, Town of	H/V	November, 1982
Chesterville, Village of	H/V	January, 1983
Cumberland, Township of	H/V	March, 1983
Eilber & Devitt, Township of	H/V	Completed
Gananoque, Town of	H	Completed
Hagar, Township of	H/V	March, 1983
Hastings, Village of	H/V	December, 1983
Hearst/Cochrane Block	H	December, 1983
Iroquois, Village of	H/V	January, 1983
Lansdowne, Township of	H	Completed
Marmora, Village of	H/V	Completed
Massey, Town of	H/V	March, 1983
Matilda, Township of	H/V	March, 1983
Norwood, Town of	H/V	December, 1983
Oso, Township of	H/V	February, 1983
Owen Sound/Orillia Block	H/V	October, 1983
Oxford County	H/V	Completed
Picton, Town of	H/V	January, 1983
Powassan, Town of	H/V	March, 1983
Rainy River and Southwest Kenora, Districts of	V	July, 1983
Spanish River, Township of	H/V	March, 1983
Terrace Bay, Township of	H/V	January, 1983
Thunder Bay/Schreiber Block	H	September, 1983
Trenton, City of	H/V	April, 1983
Tweed, Village of	H/V	Completed
White River, Township of	H/V	April, 1983
Winchester, Township of	H/V	January, 1983

Note: Tentative date of completion prior to September 1983 implies the survey may be completed but returns have not been submitted to the Ministry of Natural Resources for approval.