Datums - Map Coordinate Reference Frames
Part 2 – Datum Transformations

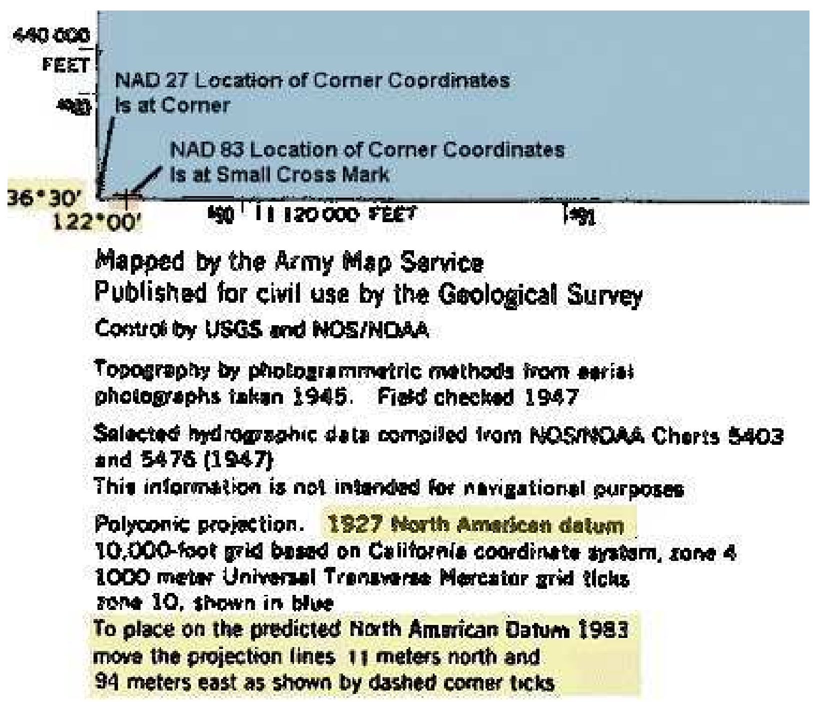
James R. Clynch
February 2006

VII. Datum Transformations

A. Basic Methods

Having data in one datum and needing the coordinates in other is a common occurrence. It happens routinely when you use a GPS receiver outside of North America. It may be hidden from the user, but a transformation must occur to display coordinates from a GPS receiver in any other datum than WGS84. Of course, map and chart makers need to often make these transformations. Re-surveying everything each time a datum change is needed is impractical.

Over any small area, say 100 km or so, the transformation will be just a constant shift in latitude, longitude and height. This is a practical statement, and as accuracy requirements increase, the area over which a simple offset can be used gets smaller. This is one common way that maps are "updated". An offset is just printed in the legend.



This is an example of a US Geological Survey topographic map that has been updated from NAD27 to NAD83 in this way. The basic map is the old NAD27 map. You must read the legend and make an adjustment if you want NAD83 (the same as WGS84) coordinates. The location of the corner coordinates is also show as a small cross in each corner. This is very useful in deciding whether to add or subtract the adjustments.

There still is the issue of how to compute these offsets for each map.

There are three common methods of making these transformations from one datum to another. In the science world, the transformation is often viewed from a vector perspective. The coordinates are transformed from Cartesian ECEF XYZ values of one datum to another. If latitude, longitude and height (LLH) are given or needed, the conversion to ECEF are done before the vector mathematics and then the new coordinates are converted back to LLH. Of course an ellipsoid definition (an "a" and "f") are needed for the LLH to/from ECEF. This method is usually called a **7-parameter transformation.** There are abridged versions of it with only **4 and 3 parameter transformations**.

A common method of directly transforming latitude, longitude, and height is the **Molodensky transformation**. This is a complex formula for the shift in latitude, longitude and height. It is complex because it is really a vector equation that is written out in its components. Also the values are scaled from lengths to arc-seconds of latitude and longitude. This method was very common before computers. It is still the most common method. There is a rudimentary method to deal with the distortion in datums. The number of parameters in this method is small.

A third method takes into account the distortions in the older transformations. A series of best-fit equations in the differences are generated. The equations are usually the same over a large area, such as the US, but there are different coefficients for small areas. This makes for a large database, but that is required if you wish an accurate transformation over a large area. This is often called the **Multiple Regression Method**.

B. Vector Method - The 7 Parameter Transform

The vector method is commonly used for the newer datums, or "frames". In this case it is assumed that there are negligible distortions and only some global changes are needed. The method deals in the earth centered, earth fixed, Cartesian coordinates (x,y,z).

It is assumed that there are three types of differences between the two frames:

a. The origin is different and a vector offset is given,
b. There is a rotation about each axis, and
c. There may be a scale change.

All these changes are assumed to be so small that many small parameter assumptions are valid. For example the sine of the rotations is replaced by the angle in radians and the cosine is replaced by 1. The order of the rotations is assumed un-important, something that is not true in general. The scale adjustment is also folded into the rotation matrix, also being order independent in this approximation. Usually the rotations are published in units of **milli-arc-seconds (mas)** or .001" . One mas is 5×10^{-9} radians. The **scale change** is also usually on the order of 10^{-9} .

The transformation equation commonly is found in two different notations. The older notations is:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{New} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{old} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} + \begin{bmatrix} \Delta s & \omega & -\psi \\ -\omega & \Delta s & \varepsilon \\ \psi & -\varepsilon & \Delta s \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{old}$$

Here:

$(\Delta x, \Delta y, \Delta z)$ is the shift in the origin,
Δs is the scale value,
ω is a rotation angle in radians about the z axis,
ψ is a rotation angle in radians about the y axis, and
ε is a rotation angle in radians about the x axis.

In some cases the vector on the right of the matrix is written as ($x-x_0$, $y-y_0$, $z-z_0$) where the zero values are the "primary point" of the original frame. In modern science work this is usually the center of the earth and thus omitted.

The alternate notation, common in the publications of information about the International Terrestrial Reference Frames (ITRF)'s is very similar in form, but uses different symbols.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_{New} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{old} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} + \begin{bmatrix} s & \omega_z & -\omega_y \\ -\omega_z & s & \omega_x \\ \omega_y & -\omega_x & s \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}_{old}$$

This general vector transformation is often called a "7 parameter transformation". It is usually used for transformation between datums with low distortion, such as the WGS's and ITRF's, which cover the world or a large area. For transformations over smaller areas there are "**4-parameter transformations**" that drop the rotations and the very common "**3-parameter transformations**" that just have a vector offset. Over a small area, most datum transformation can be adequately represented by the 3-parameter version.

C. Molodensky - The Historically Common Method

1. Vector Viewpoint of Molodensky Transformation

The basis of the Molodensky transformation is to assume that all differences are due to:

a. A shift in origin by a vector $\vec{\Delta}$ with components $(\Delta X, \Delta Y, \Delta Z)$,
b. A difference in ellipsoids of size, Δa and flattening Δf, and
c. All changes are handled with one term in the Taylor series.

There are two additional things that complicate the equations. First the shift effects are written out in components, not using vector notation. Second the results, which are initially a shift in east, north and up in distance units are converted to angles for latitude and longitude. This involves the **radius of curvature** in the two directions. (See separate section on geodetic coordinate conversions for details of $\mathbf{R_N}$ **and** $\mathbf{R_M}$**.**)

To begin the equations, note that the **unit vectors in the east, north, and up** directions are given by:

$$
\begin{aligned}
\hat{e}_E &= (-\sin\lambda, \cos\lambda, 0) \\
\hat{e}_N &= (-\sin\phi\cos\lambda, -\sin\phi\sin\lambda, \cos\phi) \\
\hat{e}_U &= (\cos\phi\cos\lambda, \cos\phi\sin\lambda, \sin\phi)
\end{aligned}
$$

The effects of the origin shift $\vec{\Delta}$ are easily obtained by taking the **dot product** of the unit vectors with $\vec{\Delta}$. The effects of the ellipsoid difference can be obtained with a few derivatives. This gives, in distance units for the shift in East, North, and Up:

$$
\begin{aligned}
\Delta E &= \vec{e}_E \bullet \vec{\Delta} \\
\Delta N &= \vec{e}_N \bullet \vec{\Delta} \quad + \quad \frac{\Delta a}{a} e^2 R_N \cos\phi\sin\phi \quad + \quad \Delta f \left[\frac{a}{b} R_M + \frac{b}{a} R_N \right] \cos\phi\sin\phi \\
\Delta H &= \hat{e}_U \bullet \vec{\Delta} \quad - \quad \Delta a \frac{a}{R_N} \quad + \quad \Delta f \frac{b}{a} R_N \sin^2\phi
\end{aligned}
$$

The equation for height is left as show above. The usual procedure for the East-Change is to convert it to arc-seconds of longitude and to convert North-Change to arc-seconds of latitude. This is done with the correct radii of curvature and the sine of one arc-second. The scaling lengths are:

$$
\begin{aligned}
L_E &= (R_N + H) \cos\phi \sin 1'' \\
L_N &= (R_M + H) \sin 1''
\end{aligned}
$$

The usual equations are found by dividing by these values:

$$\Delta\lambda'' = \frac{\Delta E}{L_E}$$

$$\Delta\phi'' = \frac{\Delta N}{L_N}$$

2. The Usual Statement of the Molodensky Transformation

Older textbooks and manuals give equations for the Molodensky transformation that would be used in a hand calculation. These give directly the shifts in latitude, longitude, and height. The angular values are in arcseconds. These formulas will be given here for completeness.

The standard form is often given as:

$$\Delta\phi'' = \left\{ -\Delta X \sin\phi\cos\lambda - \Delta Y \sin\phi\sin\lambda + \Delta Z\cos\phi + \frac{\Delta a}{a}(R_N e^2 \sin\phi\cos\phi) + \Delta f\left[\frac{a}{b}R_M + \frac{b}{a}R_N\right]\sin\phi\cos\phi \right\} \Big/ [(R_M + h)\sin 1'']$$

$$\Delta\lambda'' = \{-\Delta X \sin\lambda + \Delta Y\cos\lambda\} / [(R_N + h)\cos\phi\sin 1'']$$

$$\Delta h = \Delta X\cos\phi\cos\lambda + \Delta Y\cos\phi\sin\lambda + \Delta Z\sin\phi - \Delta a\frac{a}{R_N} + \Delta f\, b\frac{R_N}{a}\sin^2\phi$$

This is often shortened to the Abridged Molodensky Transform given by:

$$\Delta\phi'' = \{-\Delta X\sin\phi\cos\lambda - \Delta Y\sin\phi\sin\lambda + \Delta Z\cos\phi + (a\,\Delta f + \Delta a\, f)\sin 2\phi\} / [R_M \sin 1'']$$

$$\Delta\lambda'' = \{-\Delta X\sin\lambda + \Delta Y\cos\lambda\} / [R_N \cos\phi\sin 1'']$$

$$\Delta h = \Delta X\cos\phi\cos\lambda + \Delta Y\cos\phi\sin\lambda + \Delta Z\sin\phi - \Delta a + (a\Delta f + f\Delta a)\sin^2\phi$$

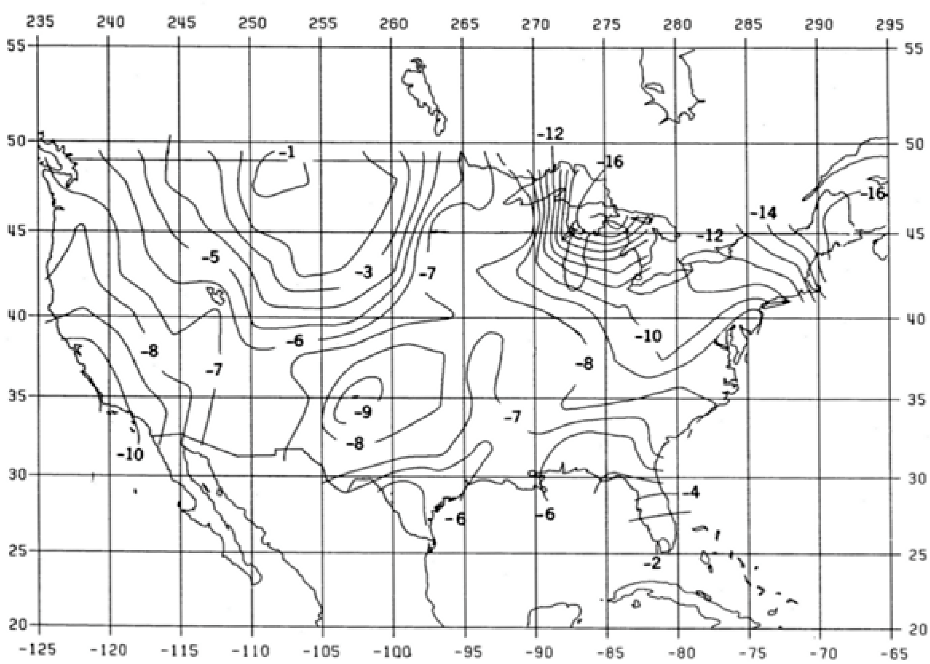
The abridged form is found by dropping any terms that are second order in small parameters (f, e, etc.). The addition of h to the radii is ignored, as the two are usually different by a factor of 1000 or more. For this reason, it is not critical if the height used is **orthometric** (H) or **geodetic** (h) where added to the R's. (See section on heights and **geoid** for details.) It will make only a small difference in an already small correction.

3. Modern Uses of Molodensky Transformation

The implementation of the Molodensky transform in imbedded in many geodetic programs. When WGS84 was initially published, a set of Molodensky transformation parameters was published between WGS84 and about 50 other datums. These were incorporated into most GPS receivers of the time, including all US Defense Department receivers. That is still the case.

The problem with the Molodensky transformation is the limited amount of data that it uses. There are only 5 numbers, three in $\Delta X, \Delta Y$, ΔZ and two in Δa and Δf. For small areas this is fine. But for larger areas such as the US with NAD27 and Europe with EU50, significant errors (10's of meters) resulted from the use of a single set of parameters. The original datums were distorted and needed more data to provide good datum shifts over the area of use. The ad hoc solution was to allow $\Delta X, \Delta Y$, ΔZ to be functions of position. These were then free parameters that were determined when the new datum was generated. An example of the resulting "**datum shift parameter ΔX** " is shown below. There are similar maps for ΔY **and ΔZ**. In fact NIMA has published maps of $\Delta X, \Delta Y$, and ΔZ for all major large area datums it uses.

The map for the "best" ΔX value is shown below. There are similar maps for ΔY and ΔZ. The ΔX values ranges from -16 to -1 m, the ΔY values from 172 to 187 m and ΔZ from 157 to 165m. There is about a 10 to 15 m variation in each axis. Of course the best-fit ellipsoids have only one offset, but in order to get better results from the Molodensky transformation, the value of $\vec{\Delta}$ is allowed to vary.

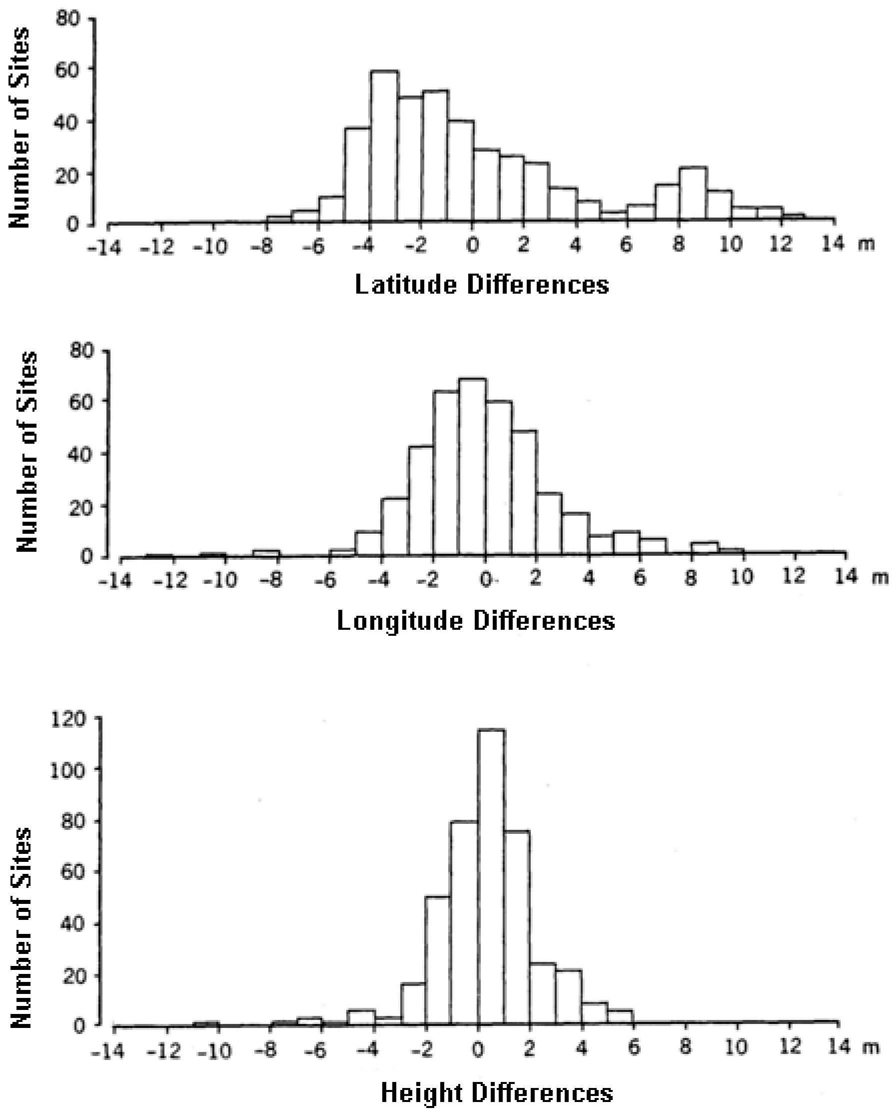


ΔX Component - meters
Datum Shift NAD 27 to WGS 84

These maps are seldom used in practice. The computer programs that do datum transformations use only a few Δ's. For the continental US it is common to use just two sets, on for east of the Mississippi river and one if west of it. This is true of the conversion routines in many GPS receivers as well as the DoD supplied PC conversion programs MADTRAN and **GEOTRANS**.

In the original unclassified report on the WGS84 datum, a table of about 50 sets of Δ's was given. There were only three sets for the continental US, the entire US, the US east of the Mississippi, and the US west of the Mississippi. This table has been incorporated into GPS many receivers and many non-scientific computer programs.

When the WGS84 datum was developed, there were many points accurately surveyed with satellite methods that were already known in NAD27 coordinates. The "datum shift parameters" for the whole US were tested with this set of points. The following is from the DMA WGS84 Report supplement.



Measured Differences
Transformed NAD27 vs. Satellite Measurements
Continent Average ΔX, ΔY, ΔZ

It is clear that for accuracies of 20 meters, this is adequate. However at 5 meters it is not. In particular the latitude errors are not random and have significant systematic character.

D. Local Fit Equations - Multiple Regression Method

The **multiple regression equations (MRE)** are ad hoc equations that provide for the shift in latitude and longitude as a function of position. They take the form:

$$\Delta\phi'' = A_0 + A_{1,0}U + A_{0,1}V + A_{2,0}U^2 + A_{1,1}UV + A_{0,2}V^2 + \ldots$$

and a similar equation for $\Delta\lambda''$ using another set of coefficients $B_{i,j}$. All the information is in the coefficients. The values of the independent variables, U and V are scaled latitude and longitude,

$$U = k\,(\phi - \phi_m)$$
$$V = k\,(\lambda - \lambda_m)$$

with k being a constant and (ϕ_m, λ_m) being a point near the middle of the area of validity.

NIMA published MRE shifts for the NAD27 to WGS84 valid for the entire US. These are shown in the figure below.

Multiple Regression Equations (MREs)
for Transforming
North American Datum 1927 (NAS) to WGS 84

Area of Applicability : **USA (Continental contiguous land areas only; excluding Alaska and Islands)**

MRE coefficients for ϕ and λ are :

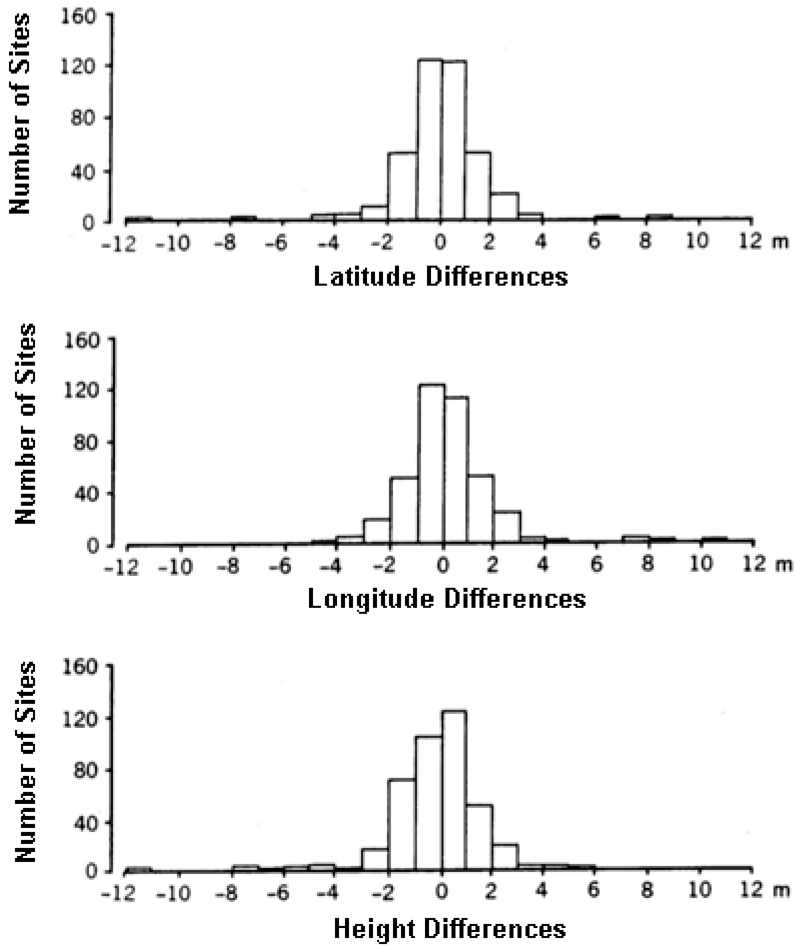
$$\begin{aligned}\Delta\phi'' = {} & 0.16984 - 0.76173\,U + 0.09585\,V + 1.09919\,U^2 - 4.57801\,U^3 - 1.13239\,U^2V \\ & + 0.49831\,V^3 - 0.98399\,U^3V + 0.12415\,UV^3 + 0.11450\,V^4 + 27.05396\,U^5 \\ & + 2.03449\,U^4V + 0.73357\,U^2V^3 - 0.37548\,V^5 - 0.14197\,V^6 - 59.96555\,U^7 \\ & + 0.07439\,V^7 - 4.76082\,U^8 + 0.03385\,V^8 + 49.04320\,U^9 - 1.30575\,U^6V^3 \\ & - 0.07653\,U^3V^9 + 0.08646\,U^4V^9\end{aligned}$$

$$\begin{aligned}\Delta\lambda'' = {} & -0.88437 + 2.05061\,V + 0.26361\,U^2 - 0.76804\,UV + 0.13374\,V^2 - 1.31974\,U^3 \\ & - 0.52162\,U^2V - 1.05853\,UV^2 - 0.49211\,U^2V^2 + 2.17204\,UV^3 - 0.06004\,V^4 \\ & + 0.30139\,U^4V + 1.88585\,UV^4 - 0.81162\,UV^5 - 0.05183\,V^6 - 0.96723\,UV^6 \\ & - 0.12948\,U^3V^5 + 3.41827\,U^9 - 0.44507\,U^8V + 0.18882\,UV^8 - 0.01444\,V^9 \\ & + 0.04794\,UV^9 - 0.59013\,U^9V^3\end{aligned}$$

Where : $U = K\,(\phi - 37°)$; $V = K\,(\lambda + 95°)$; $K = 0.05235988$

NOTE : Input ϕ as (-) from 90°S to 0°N in degrees.

While there are about two dozen coefficients in these expressions, these only do a fair job in representing the shifts. A set of a few hundred geodetic marks with NAD27 coordinates and accurate satellite positions was used to test this shift. The results were good but there were areas with significant distortions. A distribution of the results was computed.



Measured Differences
Transformed NAD27 vs. Satellite Measurments
All Continent Multiple Regression Equations

In order to do a better job, more data is needed. The approach used by the NGS for North America and a few other nations is to take the massive data sets used to define the newer, better datum and solve for simple fit equations in terms of difference of latitude and longitude from some reference points. And then have many of these reference points and apply the fit from it only in a small area. This involves a large data set. The NGS computer program NADCON uses this technique. It reads a database of about a megabyte. (Of course "large database" is a relative scale. Today this is not considered very large.)