recent activity, for earthquakes greater than magnitude 6. Wayne Thatcher and Thomas Hanks of the USGS at Menlo Park and James Hillman of Caltech searched the seismological records and concluded that a small section of the San Jacinto fault near Anza is locked in place and apparently has not ruptured since 1890. While most researchers would agree that the likelihood of a moderate earthquake is much greater in such a situation, no one can say when it might happen. To further complicate matters, the San Jacinto fault is on what is currently taken as the southern boundary of the enigmatic southern California uplift.

Despite the frustrations of those trying to predict earthquakes, researchers generally believe that prediction is a realistic goal. The Japanese and Alaskan earthquakes have reinforced the belief that at least some damaging quakes have precursory phenomena. But many believe that it will be at least 10 years, perhaps longer, before widespread, reliable prediction of major earthquakes is achieved.

—RICHARD A. KERR

Geodesy: Dealing with an Enormous Computer Task

In 1974, the National Geodetic Survey (NGS), which is part of the National Oceanographic and Atmospheric Administration (NOAA), embarked on a project of unprecedented magnitude. The project is to readjust the North American Datum—a network of reference points whose longitudes, latitudes, and, in some cases, altitudes must be known to an accuracy of within a few centimeters. Now halfway through their readjustment project, NGS investigators are confident that they can complete the job by 1983, as scheduled. They point out, however, that this project is one of the largest computer tasks ever attempted (about 700 hours of computer time are allotted).

The sheer size of the problem of computing the coordinates of these reference points has forced NGS researchers to deal with problems that arise, as they are fond of putting it, "from pushing the state of the art to its limits." For example, a key step in the problem is to solve about 2.5 million nonlinear equations in 400,000 unknowns. For smaller problems, this task is textbook simple. However, according to Charles Schwarz of NGS, the NGS problem is the largest system of nonlinear equations whose solution has been attempted. The size of this problem forced the investigators to deal with, and protect themselves against, the possibility that the computed solutions would be worthless, or even meaningless.

The North American Datum consists of about 200,000 points described by their longitudes and latitudes and about 500,000 points described by their longitudes, latitudes, and altitudes. This network is necessary to regional planners, engineers, and surveyors, who need accurate reference points when they make maps and specify property boundary lines, to navigators, and to geophysicists who study the tectonic movements of the earth. Coordinates of these points must be known with great accuracy because detailed maps of regions, such as cities and counties, are made by measuring distances and directions to various sites from the datum points. Any errors in the coordinates of the datum points will lead to even greater errors in the coordinates of the intermediate points whose coordinates are determined from them. In fact, it is a basic principle of engineering that the accuracy of the coordinates of reference points be at least an order of magnitude greater than the accuracy of the coordinates of such intermediate points.

The extensive network constituting the North American Datum has been built up gradually since 1807. New points were continually added and their positions determined in reference to previously existing ones. The datum was last readjusted in 1927, when the coordinates of all points then in the network were calculated with reference to a triangulation station on Meades Ranch in Kansas.

Since 1927, more than 100,000 points have been added to the network, and the network has been extended from the continental United States to include Alaska (Fig. 1). The coordinates of these added points were calculated with reference to the 1927 points, which resulted in the propagation of errors in the positions of the 1927 points to errors in the positions of the added points. These errors were, in places, as large as 10 m. Moreover, the surface of the earth itself moved since 1927. In some areas, tectonic movements as large as 5 cm per year have been observed.

It became increasingly obvious that the current datum is no longer adequate for today's uses. In 1971, a National Academy of Sciences committee recommended that the datum be completely readjusted. The new datum will span all of the North American continent. The governments of Canada, Mexico, the republics of Central America, and Denmark (which administers Greenland) have decided to add their own networks to that of the United States and have agreed to support and cooperate with NGS to make the resulting North American Datum consistent.

Since techniques for measuring positions of points relative to each other improved tremendously over the years, NGS investigators believe they will have to do a minimum of remeasuring when they readjust the longitudes and latitudes of the datum points. Instead, they will concentrate on solving simultaneously a set of equations relating the position of every point of the datum to every other point and tying the whole network to a few reference points. They plan to make use of all previous measurements of the positions of the points in the network to get the best fit to the available data. They break their task into two parts. First is a monumental problem in data management and handling; next are the actual computations of the readjustment.

Data management and handling is a problem in part because NGS had not previously computerized its files. Robert Hanson of NGS says this project is, in effect, "dragging NGS into the 20th century." The position of each point in the network was determined several times over the years, but all this data exist on paper and has to be put into a form that can be read by a computer. After 4 years of work, the NGS investigators are only about halfway through entering their data on the computer cards and validating the data. Moreover, this data base must be validated to rule out errors in observations and in data entries into the computer.

To validate the data, NGS researchers are working with small subnetworks of 50 to 200 points and solving for the positions of these points relative to each other. Schwarz reports that, as might be expected, a number of errors were detected. The magnitude of this task is reflected by the enormous number of measurements that must be dealt with. John Bossler, director of the readjustment project, notes that the horizontal control points—200,000 points whose longitudes and latitudes, but not altitudes, are to be determined—are described by some 2 to 3 million observations.

The mathematical portion of the read-

0036-8075/78/0428-0421$00.50/0

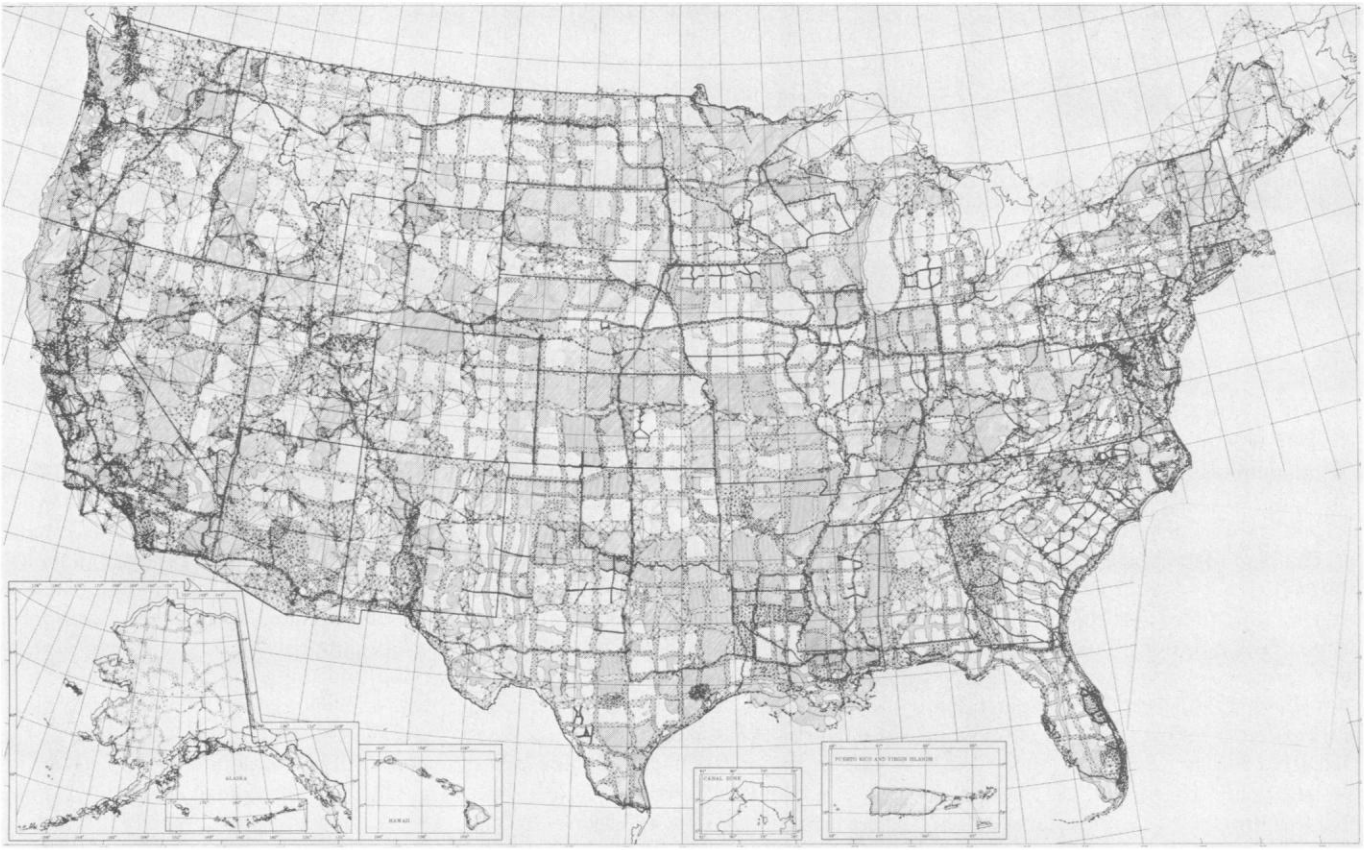

Fig. 1. Network of horizontal control points used for mapping the United States and possessions, showing the status as of July 1975. [Source: National Geodetic Survey]

justment is, in theory, straightforward. Thus far, NGS investigators are concentrating on the horizontal control points. In order to come up with corrections for the longitudes and latitudes of these points, they must solve simultaneously about 2.5 million nonlinear equations (one for each position measurement of the control points) in 400,000 unknowns (two for each control point—representing corrections for longitude and latitude). These equations are nonlinear because they contain functions, such as sines and cosines, necessary for describing the direction of one point relative to another.

The NGS investigators will first linearize their equations by means of a Taylor series expansion. They plan to solve the linearized equations and to use the solutions to improve their linear approximations to the original equations. Then the whole process of solving the linearized equations will be repeated. According to Schwarz, NGS investigators hope to get away with only one iteration, but if the solutions are not sufficiently accurate, they may have to iterate again.

In order to solve the linearized set of equations, NGS researchers must put these equations into a workable form. Since there are so many more equations than unknowns in the readjustment, the system of equations is said to be overdetermined, meaning that it may have no unique solution. The NGS investigators are using a well-known method, a least-squares adjustment, to obtain an answer that gives the best fit to the data. When they apply this method, they obtain a set of 400,000 equations in 400,000 unknowns. These equations should be solvable.

The actual task of solving the 400,000 equations is not so easy. First, there is the problem that the system of equations is too large to fit into the memory of a computer. Second, even if the problem is broken up into more manageable chunks, the answer that comes out of the computer may not be trustworthy. This is because computers can only store numbers of a fixed number of significant digits. Repeated arithmetic operations can cause significant digits to be lost. This loss of digits is especially serious when numbers of disparate sizes are added or subtracted. The NGS problem, as might be expected, requires an enormous number of arithmetic operations. For example, in one step of the problem the matrix resulting from the least-squares adjustment is solved by means of an algorithm known as a Cholesky reduction. This process alone requires about 10^{11} multiplications. Moreover, the numbers in the NGS problem are of disparate sizes.

The NGS researchers plan to break their problem into manageable pieces by a method known as Helmert blocking. This method, which was devised nearly a century ago, provides a way of ordering the points in the North American Datum. Groups of adjacent points are treated as a block. The matrix equation associated with each block is partially solved, and the blocks are mathematically pieced together. Schwarz points out that this process is algebraically equivalent to solving the whole system of equations simultaneously.

The NGS investigators worked with the Helmert blocking scheme so that it takes full advantage of the many zeros that occur in the NGS matrix, thus minimizing the number of necessary computations. The NGS matrix has a large number of zeros that arise because the position of each point of the datum is only measured with respect to a few adjacent points, not with respect to all other points in the network. The zero entries represent all the possible measurements between distant points that were not made.

Peter Meisl of the Technical University of Graz, in Austria, recently spent 10

(Continued on page 466)

RESEARCH NEWS

(Continued from page 422)

months at NGS analyzing the magnitude of the problem of round-off errors when the readjustment equations are solved by Helmert blocking. He concluded, on the basis of a statistical analysis of the problem, that round-off errors should not cause the computed solution to have fewer significant digits than are thought to be needed. (The NGS investigators want four or five significant digits. Meisl says they are certain to get two to three and are very likely to get four to five.) One reason for Meisl's optimism is that the zeros of the matrix facilitate the computations.

If the solutions to the readjustment equations are incorrect after all, will the NGS investigators recognize it? According to Schwarz, they will. He says they will look at the magnitudes of the position corrections specified by the solutions. If the solutions are correct, the magnitude of the corrections to particular points should be consistent with the magnitudes of corrections for nearby points. The solutions can also be analyzed to determine the estimated errors made in measuring the positions of points. The NGS investigators have some idea of what those errors should be, so they can decide whether the computer solutions make sense. In general, Schwarz says, the corrections should be on the order of 2 to 4 m.

Considering the magnitude of the readjustment project, it might be expected that NGS would be a large and well-funded agency. But NGS has a staff of just 300 people and an annual budget of just $10 million. Moreover, only about one-third of its staff is working on the readjustment of the horizontal control points. (Many of those not working on this readjustment project are surveyors who participate in projects, partially funded by state governments, to determine coordinates of points in between the control points of the datum.) Of the $10 million allocated each year to NGS, about $1 million goes toward paying for computer services. The NGS has no computer of its own but must buy computer time from NOAA.

Despite the small size and limited budget of NGS, its researchers are a proud and confident group who clearly view the task of adjusting the North American Datum as within their means. Hanson notes that the datum will undoubtedly have to be adjusted again, perhaps in 50 to 100 years. But next time, with the knowledge gained from this attempt, the task should be much easier.

—GINA BARI KOLATA