11.1 INTRODUCTION

The *general* objectives of the National Geodetic Satellite Program (NGSP) were, first, to get sufficiently improved positions for satellite-tracking stations that errors in connections between major datums could be materially reduced, and, second, to get a better determination of the Earth's gravitational field out to the 15th degree and order in the expansion in spherical harmonics. An evaluation of the requirements for such positions and for orbital prediction led to quantification of these objectives and to the setting of *specific*, numerical objectives. It was decided that global geodetic projects would require accuracies better than ±10m (standard deviation) in each coordinate in an Earth-center-of-mass, North-oriented system and better than ±3 mGal in the average value of gravity over 12°×12° regions. It was found that these two objectives made a third necessary—the quality of the data provided by the various tracking stations participating in the program would have to be determined. Preceding chapters have described how the NGSP set about achieving these objectives, and have given in detail the results of the program as they were determined separately by the participants.

An inspection of the results of the program shows that the *general* objectives have been met. The positions of enough stations on North American, European, South American, Tokyo, and Australian Geodetic datums have been determined to reduce the errors in ties between these datums by at least 50 percent. The number of terms in the series-expansion of the gravitational field has almost doubled. But instead of there being one set of coordinates and one gravity field, there are at least seven different major sets of coordinates and five different fields. Of course, if the various sets agree with one another to within the tolerances set by the specific objectives of the program, then the differences will be irrelevant from a practical standpoint (although they may be interesting from a scientific standpoint). But if the various sets do not so agree, then either the specific objectives of the NGSP have not been attained or a suitable set will have to be found to meet each of the objectives.

Unfortunately, the answers demanded by this assessment are not easy to obtain. In fact, a close examination not only of the various results, but also of the methods used in getting them, leads to the conclusion that the specific objectives of the NGSP were either too generally stated to allow one to tell whether they were met or were unobtainable. The existence of different results may indicate merely that the participants have gotten answers to different questions, all of which are contained within the original statement of the purposes of the program.

In this chapter, therefore, the results cited in chapters 2 through 9[1] will be examined to see if the objectives of the program, as set forth in the first paragraph, have actually been met. As will be seen, the answer is "yes" as far as the general objectives are concerned, and "almost" as far as the specific objectives are concerned. But it is not possible to select from the various sets one that probably meets the specific objectives, and it is not within the scope of the chapter to create a compromise that does. Analysis of methods and results shows that the standard deviations assigned to the results are indications of precision, not accuracy, and cannot be used to rank the various sets in order of accuracy. There is enough information available that at least a guess can be made about why the solutions differ, and the main thrust of this section will be toward exploring the extent and reasons for the

[1] Only the results of chapters 3, 5, 7, 8, and 9 will be examined in detail, since only these were produced specifically to satisfy the program's objectives.

discrepancies. The order of discussion will be the same as that of the objectives set by the NGSP: coordinates (sec. 11.3), gravitation (sec. 11.4), and evaluation of observational data (sec. 11.5). Because the validity of the results depends so much on the statistical methods used in getting them, section 11.2 reviews briefly the *statistical theory* involved.

11.2 THEORY

11.2.1 General

For many years the results of geodetic computations consisted only of angles, distances, and coordinates without any information on the reliabilities of these data. Although Gauss introduced the method of least squares in the 19th century, applied it to the adjustment of geodetic data, and explained its probabilistic implications, even today there still exist many large geodetic networks in which the coordinates of the control points are known but not their standard deviations. In such situations the reliability of the data is a matter of one's confidence in the organization or individual who produced them. There is no reliable quantitative evaluation possible, and one cannot make satisfactory numerical estimates of the accuracy of results computed with the use of such data.

The situation for the NGSP data is fortunately much better since standard deviations were calculated for most of the geodetic quantities. Furthermore, almost all the results given in the report have been evaluated by their authors by using two or three different methods rather than only one. The first and universal basis for evaluation is of course the standard deviation or the covariance. All results were obtained by means of the method of least squares, and the standard deviation and covariances of the results are contained in the matrix $\underline{\underline{\Sigma}}_X$, where the corrections $\underline{X}$ to the unknowns and the residuals $\underline{Y}$ are connected by the equation

$$\underline{X}=[\underline{\underline{A}}^T\underline{\underline{\Sigma}}_Y^{-1}\underline{\underline{A}}]^{-1}\,\underline{\underline{\Sigma}}_Y^{-1}\underline{\underline{A}}^T\underline{Y}$$

and

$$\underline{\underline{\Sigma}}_Y=\underline{\underline{A}}\,\underline{\underline{\Sigma}}_X\underline{\underline{A}}^T$$

connects the covariances of $\underline{X}$ with the covariances of $\underline{Y}$ through the matrix $\underline{\underline{A}}$ of coefficients. (See ch. 1 for more complete discussion, or see, e.g., Anderson, 1958.)

The covariances are useful principally in comparisons between results and as indicators of accuracy. As indicators of internal consistency the correlation coefficients are more suitable. Denoting the elements of $\underline{\underline{\Sigma}}_X$ by σ_{ij} and the correlation coefficients by ρ_{ij}, we have

$$\rho_{ij}\equiv\frac{\sqrt{\sigma_{ij}}}{\sqrt{\sigma_{ii}\,\sigma_{jj}}}$$

As a first approximation, the quantities $\sqrt{\sigma_{ii}}\equiv\sigma_i$ can be interpreted as the bounds between which there is a 67 percent probability that the true value of x_i lies. The ρ_{ij} indicates roughly the extent to which x_i and x_j vary together, a value of 0 indicating that they are independent and a value of 1 that x_i and x_j are functions of each other. But, almost without exception, interpretations of σ_{ij} and ρ_{ij} as anything more than the roughest indicators of where the truth lies can lead to great trouble. There are many reasons for this; the most important is the fact that the observation equations themselves are only guesses and, often enough, only rough guesses. Almost always there are present in the observations systematic effects that are not accounted for in the observation equations. So it is not at all unusual for two scientists working independently to come up with values of x_i which differ by three to four times the amounts of the σ_i's that they find. (Such anomalies are particularly noticeable when star catalogs are being compared, but can also be found in tables of coefficients C_n^m, S_n^m, λ, ϕ, h, and so on.) Perhaps the most common, dangerous, and unwarranted error found in scientific and engineering reports is the assumption that σ_{ii} is a correct estimate of error with respect to the true value of x_i, rather than being only a first, and often poor, approximation to the error.

The second basis for evaluation is comparison with the results given by other organizations. Such results are usually derived from different types of measurements or from different sets of the same type or just from more measurements. The closeness of agreement is considered a good, if not quantitative, indication of how good the results are. An outstanding example of this kind of evaluation is that used by the National Geodetic Survey (NGS) (ch. 7), in which NGS's results obtained by geometric means are compared with the Naval Weapons Laboratory (NWL) by analyzing orbital perturbations. This is a valid comparison because the results were obtained using completely different methods and using completely different sets of observations. On the other hand, to evaluate the results of the Smithsonian Astrophysical Observatory (SAO) (ch. 9) and of NASA/Goddard Space Flight Center (GSFC) (ch. 5) by intercomparison does not help much, since SAO and NASA used many of the same observations and used similar theories. Again, comparison of the results of Ohio State University (OSU) and of NGS does not help in their evaluation because NGS's data are a subset of OSU's.

Even when the values derived by different scientists agree, there is no guarantee that the values are correct. The agreement merely means that the scientists were working with similar sets of data and with similar theories. And conversely, the fact that the values disagree does not mean that only one can be correct. For example, one cannot compare NGS's values for points' locations directly with SAO's values or those of NASA/GSFC because the values are given in different coordinate systems, and the radius of the earth derived by NGS is not directly comparable with that derived by the Jet Propulsion Laboratory (JPL) from *GM* because the radii found by the two organizations refer to totally different concepts.

The third basis for evaluation is comparison of results with values whose accuracies are known. For instance, one can compare gravity computed from observations on satellites with gravity measured on the surface of the Earth; or one can compare coordinates and/or distances derived by satellite geodesy with corresponding values derived by surveying on the surface. Unfortunately, very few useful referents of this kind are available. For evaluating the accuracies of the NGSP's coordinates, we have the coordinates of stations as determined by conventional, first-order surveys. But the accuracies of such coordinates are, when known at all, known satisfactorily only within local datums and not with respect to a global system as is desirable for evaluations of NGSP's accuracies. A similar situation exists in evaluating NGSP's gravitational fields. Values of suitable accuracy are known for less than 25 percent of the Earth's surface. The regions in which accuracies are well enough known are fortunately globally distributed and connected by gravimetry. Nevertheless, lack of suitable data on the other 75 percent of the surface introduces undesirable uncertainties in evaluation of NGSP's gravitational field.

Some interesting tests of the ability to determine precision and accuracy were made at NWL by R. Anderle in 1972. The data from the Department of Defense (DOD) frequency-measuring equipment (ch. 3) were used. Precision was tested by using different sets of data in various combinations with different sets of gravitational constants. The accuracy was tested by comparing station locations found from satellite data with station locations given by the NGS geodimeter traverse in the United States.

11.2.2 Effects of Discarding Data

One interesting and important characteristic common to the reduction procedures of all participants has been to throw out data that differ from their expected values by more than a certain amount. This discarding is known by various names: filtering, preprocessing, data improvement, and so on. It is, of course, contrary to sound statistical principles if applied rigorously to data from a Gaussian distribution. All participants have assumed that the errors in the data have a Gaussian distribution. NGS's investigations have shown, at least, that this distribution applies approximately to its data (ch. 7).

The proper application of the rule for discarding data is to use it for identifying those values which differ greatly from the expected values. The background of a suspect value is investigated, and an explanation for its deviation is sought. If a valid reason can be found, the value is discarded. Such explanations as an error in copying or the known existence of a fault in the equipment provide adequate reasons. But if a valid reason cannot be found, the value should be retained regardless of how far it may be from the expected value. The assumption of a Gaussian distribution implies that values far from the expected values must be anticipated. Absence of such values would be as much of a reason for suspecting the data as their presence would be. So the discarding of values farther than a certain amount from the expected value is a direct violation of the assumption that a Gaussian distribution is present. The result of such discarding is to distort the distribution of values and to lower the root-mean-square error (rmse). If the distribution were Gaussian, the rmse would be a standard deviation and would have a probabilistic interpretation. Since, after the discarding, the set of values is no longer Gaussian, the rmse is no longer the same as the standard deviation. Nor is the weighted average any longer the best value. The problem of how to find the standard deviation from these processed data is not particularly difficult but has not been extensively studied. Grubbs (1950) and Remmer (1969) are good references for this matter.

It is easy to show that the true s.d., σ_t, of the truncated distribution is related to the putative σ by the relation

$$\sigma_t^2 = \sigma^2 (1-k)$$

where

$$k \equiv \frac{2u_0 e^{-u_0^2/2}}{1-2\Phi(-u_0)} \cdot \frac{1}{\sqrt{2\pi}}$$

$$\Phi(-u_0) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-u_0} e^{-u^2/2} du$$

and u_0 is the point of truncation. (The assumption is that the distribution is truncated at $u = \pm u_0$.)

If the rejection point is set at around 3σ, the rmse of the truncated set of data must be increased by 3 percent to get the standard deviation. If the cutoff is lower, the increase is greater. But all values used in the NGSP were so close to 3 that the increase is still less than 5 percent in all cases. Since the standard deviations themselves cannot be trusted to better than ± 10 percent at best, the effect of truncation would therefore seem to be negligible. In general, this may seem to be true. Unfortunately, some participants have discarded data in several cycles of processing. Expected values were compared with those found, data discarded, and new "expected" values computed on the basis of the abridged set. The new values applied again for still further discarding, and so on. Since the second set did not follow a Gaussian distribution, the effect on it is much more difficult to analyze. If the cycling is not continued too far (say, three times), we can assume that the effect of treating the distribution as if it were normal is insignificant. Then two discardings increase the σ by 6 percent and three cyclings increase it by about 10 percent. One difficulty with applying these numbers to the results cited in this volume is that those data finally used in getting the results have been put through such an involved process of sifting, checking, correcting, and discarding that keeping accurate track of the number of data discarded, their places in the distribution scheme, etc., is almost impossible. A safe rule would be to increase all standard deviations given in this book by at least 10 percent. This will be unfair to those organizations like OSU which discarded almost no data except those probably invalid. There are, however, so many other ways in which "improvements" are, often unintentionally, introduced into the reduction process that the 10 percent increase is much more likely to be conservative than radical.

Among the many complicating factors that made the computation of σ's unrealistic is the non-Gaussian character of almost all data gathered during the NGSP. For example, if the errors in the α and δ of a satellite had a Gaussian distribution, all values should be possible. But since the camera has a limited field of view and since the Earth in any case is not transparent, errors of more than 180 degrees are physically unlikely. The limits can, of course, be cut down to within a few minutes in most cases. The resolution of the equipment is another factor acting in the opposite direction. Many geodesists and mathematicians have looked into these problems. (See, e.g., Henriksen, 1967, for consideration of mathematical limitations and Stearn, 1964, and Boyarsky, 1965, for experimental considerations.) For these and other reasons, the σ's in this book are best considered as expressions of precision rather than accuracy.

11.2.3 Inner Constraints

The statistical procedures applied by OSU to obtain results cannot in all cases be considered mathematically identical to those used by the other participants. In particular, OSU has applied the method of "inner constraints" (ch. 8) in obtaining the origin of its coordinate system (but not in obtaining its orientation). Since the location of an origin is usually dictated by practical considerations rather than mathematical ones, the advantage gained by selecting an origin that leads to smaller σ's is debatable. But, because the method does produce lower standard deviations, its validity can be challenged. A careful analysis of the mathematics (Blaha, 1971) shows that the method is valid. It also shows, however, that the improvement in σ's is not obtained with respect to an arbitrary reference system but with respect to one defined by the data themselves. A geodetically useful frame of reference must be established with respect to physical objects (see discussion of datums in ch. 1). A system established by inner constraints is determined by the data themselves and has presumably less utility than a local datum or a datum with origin at the Earth's center of mass.

11.3 COORDINATES

The coordinates resulting from the NGSP are presented in chapters 3 and 5 through 9. Table 11.1 gives, for each point involved, the location of the tables containing the initial (local) coordinates and tables giving their final computed coordinates. (The stations themselves are listed, in order of increasing longitude, in ch. 1, table 1.27.) These coordinates should, if the mathematics is correct, be independent of the values initially assumed for them. Of course, the utility of the final computed values will depend, in very many applications, on the coordinates of each point as given originally in its local datum. This information is given for most of the points in chapters 3, 7, 8, and 9.

11.3.1 Evaluation Based on Data and Method

The coordinates given in this volume have been derived by one of two methods: either by using pure (or nearly pure) geometry or by using the theory of dynamics with or without geometry. Since the two methods are quite different, one would expect to get identical answers only if the data were the same and the theories were mathematically equivalent. Neither requirement has been met. The first requirement can be gotten around to some extent. Through the work done by NASA (ch. 5) on the third objective, and through internal evidence on the performance of the various instruments, different kinds of data from the same locations (including locations tied together by local survey) can be weighted to give an approximate equivalence. Enough positions have been occupied in common by different kinds of instrumentation that one may expect the lack of complete correspondence of sets of tracking stations to be a minor factor. These are, of course, guesses, and a rigorous investigation of the extent to which changes in data affect results has not been made. That a substantial part of the differences noted in results (fig. 11.1) is attributable to differences in data is certain; the extent is not certain. That an additional substantial part is caused by differences in method is probable. Locations determined by geometry must give the shape of the Earth. This is by definition. Locations determined by dynamics depend for their location on the orbits of the satellites used. These orbits do not depend on the shape of the Earth but are related to its figure, which depends on the gravitational field. The resulting locations should therefore also relate to the figure of the Earth. That is, if the gravitational field were known perfectly (along with the minor perturbing forces), then the orbits could be determined perfectly. The location would be determined to the accuracy allowed by the observations and would be in the same system of coordinates as the orbit. This system is, unfortunately, at present not absolute (i.e., geometrically related). One can therefore expect that the locations determined by dynamics will be related to the figure of the Earth because it is customary in this method to determine locations and figure simultaneously. The extent to which the locations are affected by the figure of the Earth will depend on (1) the accuracy of the observations, (2) the equations used for the orbit, and (3) the number and kinds of constants used for approximating the gravitational field.

The geometric theories used by the National Ocean Survey (NOS), NASA/GSFC, OSU, and SAO are mathematically equivalent except for OSU's use of inner constraints (ch. 8). Since the effect of using inner constraints is simply to translate the origin, all results should be the same if they are put into the same coordinate system and if the data are the same. The results of NASA/GSFC (ch. 5) and SAO (ch. 9) were obtained by using dynamics as well as geom-

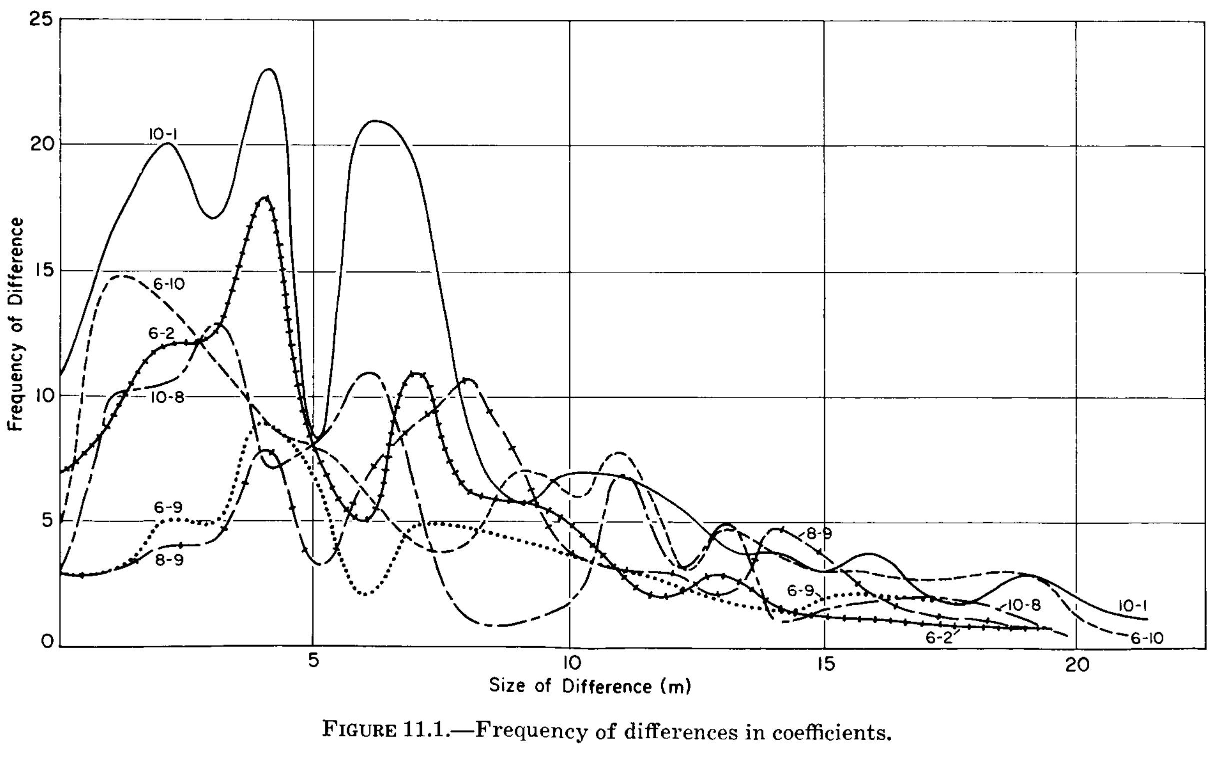

FIGURE 11.1.—Frequency of differences in coefficients.

etry, so only the results of NGS and OSU can be compared as geometric models. Comparison shows that the coordinates do not agree. This means that the differences, which are considerable (see fig. 11.1), must arise because the data are different and/or the applications of the theories are different. The former of these causes is certainly present. NGS used observations by one kind of camera (BC–4) on one satellite (PAGEOS), together with seven interstation distances determined by conventional survey (ch. 7). OSU used observations by more than four different kinds of cameras (BC–4, MOTS–40, PC–100, Baker-Nunn, and a few special types) and two kinds of radar (5-cm and SECOR). OSU used the same baselines as NGS but considerably different weights were used.

Furthermore, NGS used a slightly different set of BC–4 camera stations than did OSU, and OSU used observations on many different satellites. Therefore the differences between results are caused in large part by the considerably different data used. Theoretically, the results (WN14) of OSU should therefore be superior to those (WGN) of NGS. This is true, however, only if the additional observations used by OSU are properly weighted. But the only extensive series of experiments made to determine these weights—those made by NASA/GSFC (ch. 5) and Wallops Flight Center (WFC) (ch. 6)—were not completed in time to affect the reductions of OSU. The early, provisional results of these experiments are also, in some ways, difficult to interpret (sec. 11.4), and their use would therefore not have been advisable. The weights that were applied to the data were therefore derived from analyses of the data alone. Assignment of correct weights is not guaranteed, and the likelihood of erroneous weighting exists.

This is not the place for a detailed discussion of the weights to be assigned to the observations. Such a discussion is given by J. Berbert in chapter 5, and a discussion on Berbert's results is presented in section 11.4. The most detailed and extensive study of the errors present in a particular set of data is that of NGS on the errors in data from BC–4 cameras (ch. 7).

SAO's figure 9.13 in chapter 9 shows that, in the examples given, the axes of the error ellipses have the same orientation whether geometry or dynamics is used to find the direction. The ratios of the axes differ, however, and the centers are from 1σ to 3σ apart. The conclusion would be, if these figures are typical, that there are real differences between results obtained by geometry and results obtained by dynamics.

Anderle's use of a comparison between "dynamic geoid" and "geometric geoid" as a means of finding out how close NWL's coordinate system is to the Earth's center of mass is ingenious but inconclusive at present. The geometric geoid to which he refers is based on dynamics just as much as is the dynamic geoid, which is based on SAO's coefficients C_n^m and S_n^m, and the comparison is between geops both derived by dynamics. A further complication in this case is that the geoid based on SAO's coefficients is itself of unknown accuracy, as can be seen by comparing it with other geops.

As Anderle points out (ch. 3), the fact that station positions derived from Doppler data in tests in the United States agree with positions derived by conventional survey on the ground to within 1 to 3 meters does not mean that coordinates outside the United States are good to this accuracy. There is also the fact that the conventional survey itself is good only to 2 to 5 m overall. The accuracy of the global set of positions therefore may be better than 6 meters, but not, probably, as good as 1 to 3 meters.

11.3.2 Evaluation by Comparison of Results

Results cannot be evaluated on the basis of the results themselves. What is needed is an external set of standards with which comparison can be made. No such standard of unimpeachable accuracy exists. The closest we have to this is the set of seven baselines in North America, Europe, Australia, and Africa, which were used to insert lengths in to the WGN of NGS. Unfortunately (see next section), there is disagreement about the accuracies of these baselines to an extent that makes it unwise to depend on them for evaluation.

Since the σ's of the baselines cannot at present be relied on, the σ's attached to those coordinates which were derived by using the baselines cannot be relied on either. They must be considered measures of precision rather than accuracy. This view is supported by a study of the differences between corresponding coordinates in different models. Whatever inferences are drawn, the results about their accuracies must come from comparison between results. Since no one set of results can be chosen from the evidence as "best," the inferences can only be indicative, not final or absolute.

Figure 11.1 shows the frequency with which differences of 0 to 20 m (the largest) between corresponding sets in GEM 6 (1), NWL–9D (2), WGN (6), GSFC '73 (8), and SAO SE III (9) occur. Six curves, corresponding to differences (6) – (9), (6) – (2), (6) – (10), (10) – (1), (10) – (8), and (9) – (8), are shown. (Only five of these curves are independent, of course.) The coordinates were rotated, translated, and scaled into a common system before differences were taken. The transformations were not based on full sets of stations common to all participants, but enough common stations were used so that the difference between this system and that obtained with a full system is small. The differences given are probably within 2 m of the correct values.

Note that in general the differences vary from 0 to 21 m, the most frequent difference being 4 m. A closer look at the figure shows that WGN and NWL–9D show remarkably few large differences, whereas WGN and SAO SE III also agree more closely in this respect than do WGN and OSU's set. The average difference between WGN and NWL–DOD is 4.5 m, between WGN and OSU's set, 7.0 m, and between SAO's SE III and GSFC '73, 9.2 m. Other interesting deductions could be made, but it is obvious that even if we add to those differences the σ's for corresponding sets, the results will still be close to the $\pm$10–m limits, although they will not always be within those limits. But there are enough differences greater than 10 m present to make it doubtful that acceptance on the basis of average differences would be permissible—i.e., would ensure that whatever set was chosen met the requirements.

It must in any case be remembered that the differences are for coordinates in the same systems. As indicators of error, the differences in the systems themselves, as well as the differences in coordinates, should be considered. This consideration is taken up in section 11.3.2.5.

11.3.2.1 Comparison With External Standards: Baselines

It is interesting to compare the lengths of the baselines established for use in NGS's WGN with lengths computed by OSU and NWL from their results. Table 11.2 gives the identifying numbers of the stations terminating the lines, the datums governing the lines, the approximate lengths of the lines, their standard deviations as estimated from the survey, and the differences of values from these lengths. The discrepancies for all the baselines except those in North America and the baseline from Hohenpeissenberg to Catania are much greater in OSU's case than the originally estimated σ's should allow. The line from Tromsø to Hohenpeissenberg is also suspect in NWL's analysis. Only if we accept OSU's estimates of the σ's of the surveyed length, do all the differences fall below 3σ. Since the results obtained by NWL are quite independent of the lengths from traverse, the line from Tromsø to Hohenpeissenberg is suspect even though it has the lowest value of σ^2 of all the baselines.

The σ's quoted for the original surveys (column 3) are smaller (in absolute value) than the nature of the survey would lead one to expect. This is particularly true for the European baseline out of Tromsø and for the two Australian baselines. The line from Hohenpeissenberg to Catania has a relative error of 1.2×10^{6}, which is not unreasonable for a line going through the Alps. OSU's estimates (column 5) seem reasonable and make the results of both OSU and NWL reasonable. But, as was noted earlier, this apparent reasonableness of values is not evidence.

11.3.2.2 Comparison With External Standards: Distances Computed by Triangulation or Traverse

One useful, if not decisive, way of evaluating the coordinates given in this book is to compute from them the distances between various pairs of stations and to compare these distances with distances computed by using results of surveys carried out on the surface. Unfortunately, this method has been adopted only for one set of coordinates, that designated as GSFC '73 (ch. 5). The values given for GSFC '73 would indicate an agreement between satellite-derived and conventionally derived distances of, on the average, 5 m or so. But there are two reasons for being hesitant about accepting the 5-m value. First of all, the rms error in the distances computed from conventional survey is probably between 3 and 5 meters or more. A realistic estimate of the rms error in the satellite-derived distances would have to take this into account. But the second reason makes such an accounting difficult. The distances were apparently derived independently, but a glance at the geometry shows that they are not actually all independent. (The bar graphs shown are therefore misleading.) A number of different sets could be selected, each containing independent distances. The associated differences will differ from set to set; from the information now available there is no way of telling which set is the correct one. If, as seems reasonable from the evidence, the rms error of the distances calculated from the GSFC '73 data is assumed to be less than 10 m but greater than 5 m, the error in each coordinate would be between 3 m and 6 m. To these errors would, of course, have to be added the errors caused by errors in the coordinate system itself.

11.3.2.3 Comparison With External Standards: Miscellaneous

There exist a number of stations, not participants in the NGSP, whose distances from each other or from the Earth's axis of rotation have been computed by methods quite different from those used by the NGSP's participants. These distances have been used by NASA/GSFC and others for comparison with distances computed from NGSP's stations. The comparisons are given in chapters 4, 5, and 9. The comparisons are, unfortunately, not accompanied by an adequate error analysis. Although comparisons show agreement to within 5 m on the average, with excursions up to over 15 m in some cases, the lack of supporting information makes it impossible to infer from the comparisons anything about accuracy or precision. This is unfortunate, since results obtained by quite different methods are involved. One can say that the results do not contradict each other, but neither do they contradict an estimate that the NGSP's coordinates contain errors of over 10 m on the average.

11.3.2.4 Influence of the Reference System Used

Table 11.4 summarizes the differences between the coordinate systems (WGS's) used in this volume for satellite geodesy and the datums controlling the large horizontal networks. Table 11.5 summarizes the differences between the WGS's themselves. The data in table 11.4 are most useful from a geodetic and practical standpoint. They not only provide the necessary data for going from one datum or coordinate system to another, but also show clearly that the relationships are not well known, or at least not known to the degree of accuracy required by the NGSP. Of course, the systems of NGS and OSU are not strictly comparable either with each other or with the systems derived by dynamics. However, we can expect that the differences between local datums should be the same in either system. But the distances between origins of the Australian and

European datums are, e.g., 273 m in the system (WGN) of NGS and 289 m in the systems of WN14 (SU), or a difference of about 9 m in each component. These are purely geometric systems. SAO SE III gives a distance of about 360 m. The corresponding difference between EU50's and NAD 1927's origins in these systems is 49 m, and between EU50 and SAD 1969 is 37 m. On the other hand, the difference is only 8 m going from NAD 1927 to SAD 1969. These numbers lead to interesting speculations, but since no definite conclusions can be drawn, we will not go further.

The differences between local datums and global datums derived by using dynamics should be comparable, since the global systems not only have the same orientation but also, presumably, have the same origin, the Earth's center of mass. A glance at table 11.5 shows that coordinates of the origins are close together, but not as close together as the requirement for ± 10 m (sec. 11.1) would require. Coordinates of the center of the Australian Geodetic Datum 1965 differ by up to 18 m in X, 31 m in Y, and 35 m in Z. Even for NAD 1927, in which a large block of stations occurs, we have differences of 33 m, 41 m, and 14 m in the individual coordinates. One of the reasons for these differences is of course the very different number of stations used by the investigators in determining the constants involved. But it is not a major factor, as a comparison of the DOD/NWL column with the other will show. (NWL had the smallest number of stations per datum.)

The parameters in tables 11.4 and 11.5 are arranged as

X (m)	rotation about the X axis ($'' \times 10^2$)	
Y (m)	rotation about the Y axis ($'' \times 10^2$)	scale difference $\times 10^6$
Z (m)	rotation about the Z axis ($'' \times 10^2$)	

The comparison in table 11.5 is skimpy because lack of time made it impossible to compute the many relationships involved. Those interested and able can extend the comparisons by using table 11.4, taking the geodetic datums as intermediaries.

The outstanding characteristic of the values in table 11.5 is, first of all, the large values for X, Y, Z and, second, the large size of the rotation about the Z axis. The close agreement between GSFC '73 and SAO SE III undoubtedly results from commonality of data. The closeness of GEM 6 to GSFC '73 (except in longitude) is not explainable on this basis. In assessing the effect of rotations, note that the linear equivalent of angle x is approximately one-third the number given, multiplied by the cosine of the angular distance from the angle of rotation.

11.3.2.5 Discussion of Particular Sets

11.3.2.5.1 SECOR EQUATORIAL NETWORK

The Defense Mapping Agency/Topographic Command (DMA/TC) estimates (ch. 3) that the coordinates in the SECOR Equatorial Network (SEN) have standard deviations (in accuracy) of the order of ± 20 m. This is a large value and is not in accord either with the assessment from NASA's evaluations of SECOR (ch. 5 and sec. 11.5) or with OSU's results using data from SECOR (table 11.3). It results from comparisons of interstation distances computed from DMA/TC's results with distances obtained by conventional survey. It does accord with DMA/TC's own estimate of SECOR's accuracy. This indicates, if all tests are valid, that the data from SECOR can provide standard deviations better than ± 20 m if properly handled.

The reasons for SEN's failure to reach its potential strength are difficult to assess from the information available. An obvious partial explanation lies in the combination of weak geometry intrinsic to SECOR with the less than optimal geometry enforced on SEN by the distribution of occupiable sites. Although Blaha (1971) has pointed out that the configuration involved in determining, by geometric means alone, the location of a fourth point from three known points is such that small errors in distance measurements result in large errors in the coordinates of the fourth point if the four points are nearly coplanar, this conclusion does not apply to SEN. It holds only if more than 3 of the 12 coordinates are unknown. This consideration does affect OSU's procedure and results. (OSU—ch. 8—also used the data from the SEN.) However, it does not apply to the results from DMA/TC. DMA/TC used geometry only to obtain preliminary values for the coordinates; for the final coordinates, they used the short-arc method (ch. 1, ch. 3), in which a simple orbit is fitted to the observational data. The geometry could still be poor (DMA/TC gives no information on this point), but it is reasonable to suppose that a sufficient number of passes was observed at each station to give a good geometry. Certainly, the satellites GEOS–1 and GEOS–2, with their inclinations 59° and 106°, would provide good geometry for an "equatorial" network like SEN, and the SECOR series of satellites (ch. 3) had a good selection of inclinations and heights.

It is possible that the theory used by DMA/TC contains an error or was inadequate. No error is apparent in the theory given in chapter 3, although investigation of finer details might uncover one. The possibility of inadequacy is more likely. The original specifications on SECOR called for an rms error (in range) of ±1 m. GSFC's and OSU's evaluations of SECOR indicate that SECOR can give distances to better than ±5 m.

OSU, by requiring that the distances measured by SECOR agree with estimated distances to within a reasonable value (ch. 8), was able to keep the resulting rms error in SECOR-measured distances to well within GSFC's estimate. Since DMA/TC's results indicate that the standard deviations are 3 to 4 times larger than instrumental evaluation and OSU's results would show they should be, the conclusion is reached that the theory is inadequate rather than erroneous.

Errors too subtle to be evident from the theory may exist without being uncovered by OSU's analysis, particularly if they cause systematic errors of the same size as the deviation of OSU's results from, say, GEM 6 (ch. 5) or SE III (ch. 9). The refractive theory would be particularly suspect in this case.

Table 11.3 compares the coordinates in SEN and WN14. SEN has, according to chapter 3, the same origin and orientation as WGN. It is obvious that the heights are systematically off. Even when the difference in ellipsoids (about 12 m) is added, the difference still amounts to about 6 m on the average. This difference is almost certainly caused by the failure of DMA/TC to apply that method of correction used by OSU whereby the allowable error in SECOR's ranges was bracketed between assigned limits.

11.3.2.5.2 THE "COMBINED" SOLUTION OF NGS

A purely geometric solution has no relationship with the figure of the Earth that can be found from the data themselves. A full discussion of this point is presented in chapter 8, and OSU did not attempt to relate its system (WN14) to the figure of the Earth. For certain reasons connected with its intended use of WGN, however, NGS obtained a further set of coordinates, the "combined solution" (designated here as WGN–C) which *is* related to the figure of the Earth. This relationship was found by enforcing a certain amount of agreement between the coordinates in WGN–C and those in NWL–9D after appropriate transformations for scale, etc. (see ch. 7). The WGN–C therefore lies close to NWL–9D where the stations were colocated. Because of the way WGN–C was produced, it is not a purely geometric solution. Furthermore, its evaluation as an entity distinct from WGN and NWL–9D is easily subject to misinterpretation. WGN–C is therefore considered here as a compromise between WGN and NWL–9D. This does not imply that it is better or worse than either of the others, and present evidence is insufficient to prove from the characteristics of either of the others what its own accuracy may be.

11.3.2.5.3 COORDINATES OF STATIONS OF DMA/AC AND AFCRL

The coordinates given by DOD/AC for positions in South America (ch. 3) are tied to the NGS's WGN 1973 (ch. 7). The accuracy of these positions therefore depends strongly on the accuracy of WGN 1973. Separate evaluation does not seem warranted. The stations serve very well to strengthen the results for NAD 1927 and could be included, where σ's permit, in adjustment of triangulation in North and South America. The coordinates of the stations on Johnston Island and Bermuda are of course not connected to the others and should be judged individually. (See ch. 3.)

11.3.3 Summary

From the preceding evaluation, it appears that the general (geometric) objective of the NGSP has been attained, and handsomely. But if one adopts a no-compromise attitude towards the relation of the various results to each other and to the NGSP, one must conclude that the specific (numerically stated) objectives of the NGSP have not been conclusively attained. There are six major sets (NWL–9D, SEN, GEM 6, WGN, WN14 and SE III) of coordinates resulting from responsibilities assumed at the start of the NGSP and several sets (GSFC '73, DMA/AC, etc.) resulting from later involvement in the NGSP. The σ's found for the coordinates in any one set are for the most part better than ± 6 m or are less than 10 m in absolute value for the total error in location of a station. If we look at the differences between systems (tables 11.4 and 11.5) and between the coordinates themselves when they are transferred to a common system (fig. 11.1), we find that these differences are, when combined, larger for most stations than the tolerances allowed by the objectives. To adopt *one particular* set not only would result in losing a few or many stations, depending on the set selected, but also could not in any case be justified on the basis of objective evidence now available. The selection would have to include subjective judgments.

To adopt *average* values for the coordinates cannot be justified on any basis that is theoretically sound. Some major systems have in the past been defined in this way, but the justification lay not in the scientific evidence but in the political situation. There is no reason to believe that the differences between sets are random; there is even less reason to believe that they have a Gaussian distribution. This is strongly shown by OSU's plotting of the rotational relationships between systems (ch. 8). To arbitrarily select one particular system would make, in the present circumstances, much more sense than to compute a scientifically meaningless compromise. Such a set would have to be derived by weighting the individual values. The only nonsubjective weights available are those intrinsic ones calculated by the participants, and an analysis of the differences shows that at best one of the six sets of standard deviations can be correct.

The most reasonable recommendation is that a user adopt that set which (1) contains a set of stations suitable for his work (i.e., as regards number of points, proximity to user's areas of interest, etc.) and (2) was derived by geometric methods (NGS or OSU), if the user is concerned only with geometry, or by dynamics (NWL, NASA/GSFC, or SAO), if the gravitational field is also involved. The closer the user's situation is geodetically to that of one of the NGSP's participants, the better will be the user's results employing that participant's model.

In one sense, fortunately, we can say that the ± 10–m objective has been partially met as far as some individual stations are concerned. Our reason for claiming that we have fallen short of the goal is the existence of differences larger than the ± 10 m. But such differences do not exist for all the stations. There is a reasonable number of stations for which the differences in the coordinates lie well within the ± 10–m limits.

For such stations, we can certainly claim that the objective has been met. Of course, there is still the fact that coordinate systems themselves differ by considerable amounts. In the absence of more information, we cannot rigorously compute correct σ's for the transformations. We assume, therefore, that the transformation between systems (i.e., rotation of one system to the true system) contributes half the total error, the errors within a system contributing the other half. At the risk of oversimplifying the situation, we then assume that the objectives are met by those stations whose coordinates agree to within ± 7 m and whose intrinsic σ's (those calculated by the participant) also lie close to ± 7 m. This results in the list of stations presented in table 11.6, which can be considered a list of fundamental stations.

In discussing the reasons for the observed differences in results, mention was made of the differences that must be caused by use of dynamics rather than geometry. The participants have not analyzed the statistical implications of the theories in sufficient depth that numerical values for these differences can be calculated. A very rough guess at what the difference should be is 1 to 2 m. This is considerably less than the differences calculated for the two sets most closely associated, NGS's WGN and NWL's NWL–9D (ch. 3). The differences between the coordinates in WGN and NWL–9D are for the most part considerably larger than would be expected from the σ's computed for either set. The same holds true for differences between OSU's WN14 and NWL–9D. But the differences between OSU's and NGS's values are also too large. At present we can say only that the geophysically significant difference between results using geometry and results using dynamics is not separable from differences arising from other causes.

The estimate that the coordinates of the majority of the stations have standard deviations of the order of over ± 10 m when the uncertainties in the reference system are included is supported by the results of a computation by Marsh, Douglas, and Kloska (unpublished paper, 1973) of the coordinates of over 50 satellite tracking stations which participated in the tracking project ISAGEX in 1970–1971. Some of the data from this project were used in deriving GSFC '73 (ch. 5). Computation using a larger set of data from ISAGEX alone showed that the new set of coordinates differed from the previous set by a representative amount of 20 m.

The one definite and important conclusion we can draw is that the data accumulated in the NGSP have not been exploited to the extent either possible or desirable. The existence of σ's generally well below ± 10 m in each coordinate and the existence of unexplained discrepancies considerably larger than this shows this definitely. It shows, furthermore, that one of the reasons for insufficient exploitation is that the theories, detailed as they are, are still inadequate.

11.4 EVALUATION OF REPRESENTATIONS OF THE GRAVITATIONAL POTENTIAL

The three major specific objectives of the NGSP were stated in section 11.1, and the first of these, determination of the locations of tracking stations to ± 10 m in a center-of-mass system, was discussed in section 11.3. The second objective, determination of the average gravitational field to ± 3 mGal over 12×12 degree regions, will be discussed in this section. Because many users of the results are interested in actual values rather than average values, the results will be considered from several different aspects.

11.4.1 Evaluation by Comparison Between Models

As a result of the NGSP, two major determinations were made of the coefficients $\{C_n^m, S_n^m\}$ in the representation of the gravitational potential as a series of associated Legendre polynomials (ch. 1). These two resulting representations (or models) are those of GSFC (GEM-6) and SAO (SE III). They are discussed in chapters 5 and 9 and extend the representation to beyond $n=16$, $m=16$. These chapters give some indications of how good the representations are, and this information is used here in evaluating the models. (The models of Applied Physics Laboratory (APL), Koch and Wilet, and Rapp are also discussed in order to give a perspective to the two major representations.) An interesting and valuable approach is, however, to compare the representations with each other and with derivatives of the true potential. The latter comparison has also been made by Decker (1972) and is discussed later. The former comparison, while not as definitive, since it gives only relative values, does give an immediate indication of the variability.

Table 11. 7 gives the differences (and the percent differences) between the coefficients $\{C_n^m, S_n^m\}$ of SAO's SE III, of APL 5.0, of GEM 5 and GEM 6, and of Koch and Witte (1971), with respect to the corresponding coefficients of Rapp (ch. 3). The coefficients of Koch and Witte, APL 5.0, and GEM 5 are derived from tracking data alone. The coefficients of SAO SE III and GEM 6 are derived by using both mean gravity anomalies and tracking data. The set of coefficients used as referrent is that of Rapp (ch. 3), which is also based on mean gravity anomalies and on tracking data. Since Rapp's set used the coefficients in GEM 3 (ch. 5), one would expect GEM 5 and GEM 6 (which are descendants of GEM 3) to agree more closely with Rapp's set than any of the others do.

A perusal of table 11.7 shows that this expectation is indeed true. Similarly, as one would expect, SAO's set comes in third, and APL 5.0 and the set of Koch and Witte (1971) come in last—i.e., have the largest differences from Rapp's. If the contributions of the gravity anomalies were of major importance, one would also expect to find SE III agreeing more with GEM 6 than GEM 5 does. But this is not the case. GEM 5 and GEM 6 are in fact much closer together than are SE III and GEM 6. So we can conclude that the differences between sets depend so overridingly on the tracking data that the inclusion or omission of data on gravitation at the surface is unimportant.

But a glance at table 3.37 (ch. 3) shows that Rapp's set, GEM 3, GEM 4, and SE III gave essentially the same rms error in the orbit. Obviously, since the variation in coefficients is 600 percent from the reference set even for $n<8$, the sets determined by tracking data, with or without gravity data, must be considered as being representative more of the orbits used than of the Earth's gravitational field.

The percentage deviations [2] increase as n increases; there are values of over 3000 percent. Also interesting is the fact that the standard deviations for the zonal coefficients show differences an order of magnitude or more greater than the standard deviations given by SAO (ch. 9). The same pattern is almost certainly followed also by the tesseral coefficients.

One could take the data in the table and compute average deviations, rms deviations, and so on. But the usefulness of such numbers is uncertain because there is no evidence, external or internal, indicating how close any one of the five sets is to being correct. The sizes of the deviations cannot be trusted. R. L. Decker (1972) evaluated APL 5.0 against DMA/AC's accumulation of some 24 000 means of gravity anomalies over 1×1 degree "squares." The evaluation showed that APL 5.0 deviated more from DMA/AC's values than did any of the other models investigated. This agrees with the indications of the table that APL 5.0 and the models of Koch and Witte (1971) are least representative of the field. It also showed that the models of Lerch *et al.* (1972a,b,c), Gasposchkin and Lambeck (1970), and Rapp (1972a) agreed reasonably well with the mean values (to within ±13 mGal at worst and to within ±8 mGal at best). This again one would expect from the table. The sets used by Decker are of course not the same as those given in the table, with the exception of APL 5.0. But there is no evidence that the present sets are more than a moderate improvement over

[2] Care must be taken in considering these percentages, since for $n>$ about 12, the C_n^m, S_n^m are frequently very small.

Decker's sets. Trial computations indicate an agreement to within ±4–5 mGal, and this is supported by, e.g., figure 9.11 (ch. 9) and figures 5.47 and 5.48 (ch. 5). Consequently, one would not expect the relative standings (with respect to rms deviations or error) to be changed significantly.

The lack of agreement among sets and also between sets and the real Earth (as we know it) would seem to lie at least partly with the phase information in the coefficients. This is what the frequent serious imbalance in differences between C_n^m and S_n^m would indicate.

An obvious conclusion from the table is that representation of the gravitational potential by series of associated Legendre function is not a good idea for practical use, and is only of limited scientific value. This harsh judgment is supported not only by the variation in values shown in table 11.7, but by the following consideration. First, contributions from a particular term

$$\frac{k^2M}{a_0}P_n^m\ (\sin\psi)\ (C_n^m\cos m\lambda+S_n^m\sin m\lambda)$$

do not represent the contribution from a single region but from regions all over the Earth; i.e., the effect of a region is smeared out. Except in the case of harmonics of low degree, physical interpretation is therefore not rewarding. Second, to give average values of potential over regions of area A km^2 requires about $5.03\times10^6/A$ numbers. But to ensure that these are completely represented by spherical harmonics requires about as many coefficients to be specified. Thirdly, it is obvious that the gravitational field is not overly sensitive (as measured by allowed error) to large changes in the $\{C_n^m, S_n^m\}$. Almost the only advantage of using spherical harmonics is the ease with which algebra is carried out.

Rapp (1973) has compared the gravitational constants in GEM 6, SE III, and his own solution. He gives the comparisons of the rms averages when only those for the same $\{n,m\}$ are considered and also when all coefficients are considered. Table 11.8 is abstracted from Rapp's paper.

11.4.2 Evaluation by Comparison with Gravimetric Data

11.4.2.1 Evaluation on the Basis of NGSP's Stated Geodetic Objective

If we adhere strictly to the stated numerical terms of the NGSP's objective (sec. 11.1), we are concerned not with how well the various representations (models) can reproduce the gravitational field, but with how well they can reproduce averages over areas of a given size. There is a considerable difference between the two considerations. If we use as a measure of reproducibility the rms difference, then we want

$$E[(\bar{g}_{sn}-\bar{g}_{Tn})^2]$$

not

$$E[(g_{sn}-g_{Tn})^2]$$

We have, to work with,

$$(\bar{g}_{sn}),\ g_{Tnmi}$$

which are, respectively, the average value of g_{sn} over area n, and the measured values of gravity in the same area. Ideally,

$$\bar{g}_{nm}=\frac{1}{A_n}\int_{A_n} g_{Tno}\,dA_n$$

where $g_{Tno}(\lambda,\phi)$ is the true value of gravity at (λ,ϕ) in area n. g_{Tno} differs from the corresponding measured g_{Tnmi} (λ,ϕ) because of errors in measurement, reduction to the surface, etc. Let ϵ_{Tnmi} be the error in measurement. But $\bar{g}_{Tn}$ will differ from the average computed from g_{Tnmi} not only because of the ϵ_{Tnmi}, but because only a sample $\{g_{Tnmi}\}$ of g_{Tn} is taken. The error introduced into $\bar{g}_{Tn}$ because of improper sampling is denoted by $\bar{\epsilon}_{Tnr}$ and is called the error of representation. Then

$$\bar{g}_{Tn}=\frac{\sum_i (g_{Tnmi}-\epsilon_{Tnmi})}{I}-\epsilon_{Tnr}$$
$$\equiv\bar{g}_{Tnm}-\bar{\epsilon}_{Tnm}-\bar{\epsilon}_{Tnr}$$

If we define ϵ_{Tnri} by

$$\epsilon_{Tnri} \equiv g_{Tnmi} - \epsilon_{Tnmi} - \bar{g}_{Tn} \qquad (11.1)$$

we have quantities which indicate the variation of g_{Tno} over area n. The variation is not random. They and the average $\bar{\epsilon}_{Tnr}$ depend on how the sample is chosen. As I increases, $\bar{\epsilon}_{Tnr}$ will approach zero, but this does not guarantee that for a given I, the average of $\bar{\epsilon}_{Tnr}$ over many n will approach zero.

The behavior of ϵ_{Tnmi} and ϵ_{Tnm} is quite different. In general, ϵ_{Tnm} will not be zero, but will be small if care has been taken to calibrate the measuring equipment frequently. The $\{\epsilon_{Tnmi}\}$ are then, hopefully, random. $E[(\epsilon_{Tnmi})^2]$ will vary little within regions of considerable extent but can vary considerably from region to region. Typical values in oceanic areas are 70 to 20 mGal where measurements were made 10 years or more ago, and ± 2 to 5 mGal where recent measurements have been made. On land, the rms error is characteristically better than ± 0.1 mGal except in regions where heights are not well known. Now, using ν instead of n to prevent confusion with the degree n of spherical harmonics,

$$\begin{aligned} E[(\bar{g}_{s\nu} - \bar{g}_{T\nu})^2] = {} & E[(\bar{g}_{s\nu})^2] + E[(\bar{g}_{T\nu m})^2] \\ & + E[(\bar{\epsilon}_{T\nu r})^2] - 2E[\bar{g}_{s\nu}\bar{g}_{r\nu}] \\ & - 2E[\bar{g}_{s\nu}\,\bar{\epsilon}_{T\nu r}] \\ & - 2E[\bar{g}_{T\nu m}\,\bar{\epsilon}_{T\nu r}] \qquad (11.2)[3] \end{aligned}$$

Note that $E[(\bar{g}_{s\nu} - \bar{g}_{T\nu})^2]$ is not a measure of the agreement between two differently derived gravity fields, but a measure of agreement between averages. As the size of A increases, $\bar{g}_{s\nu}$ and $\bar{g}_{T\nu}$ become less representative of the field but agree more with each other. From equations (11.1) and (11.2),

$$\begin{aligned} & E[(\bar{g}_{s\nu} - \bar{g}_{T\nu})^2 - (\bar{g}_{s\nu} - \bar{g}_{T\nu m})^2] \\ & = E(\epsilon_{T\nu i})^2 - 2E[(\bar{g}_{s\nu} - g_{T\nu m})\,\epsilon_{T\nu r}] \end{aligned}$$

From the results of either SAO (SE III) or NASA (GEM 6), it is obvious that $E(\bar{\epsilon}_{T\nu 1})^2$ (which is roughly equivalent to $E(\epsilon_T)^2$ in their tables) is small in comparison with $E[(\bar{g}_{s\nu} - \bar{g}_{T\nu})^2]$ and $E[(\bar{g}_{s\nu} - \bar{g}_{T\nu m})^2]$. If it is also assumed that $E[(\bar{g}_{s\nu} - \bar{g}_{T\nu m})\bar{\epsilon}_{T\nu r}] \approx E(\bar{\epsilon}_{T\nu r})^2$, the tabulated quantity $E[(\bar{g}_{s\nu} - \bar{g}_{T\nu})^2]$ can be used as an estimate of difference between the two average gravitational fields.

Assume again that $E[(\bar{g}_{s\nu} - \bar{g}_{T\nu})^2]$ varies inversely as ν, and use the best estimates given. (For SE III this is given for n and m equal to 18 and for 5×5 degree regions containing 20 or more 1×1 degree regions for which average values are available.) For GEM 6 this is given for n and m equal to 16 and for the same kind of 5×5 degree regions. The value of $\sqrt{E[\bar{g}_{s\nu} - \bar{\bar{g}}_{s\nu})^2]}$ for 12×12 degree regions is then approximately ± 5 mGal for GEM 6 and ± 4 mGal for SE III. In view of the many assumptions made to arrive at these values, one must conclude that the values are decidedly on the optimistic side. But since optimism is, under the circumstances, at least as justifiable as pessimism, the easy corollary is that the results are close to what they should be. It is obvious, however, that we could have considerably more confidence in the results being within the desired limits if they were less than ± 3 mGal rather than greater.

This conclusion is deliberately made weak because of the many still unresolved problems involved. The information that relates to accuracy of the models presented by GSFC and SAO is contained in table 9.43 (ch. 9) and tables 5.60 and 5.61 (ch. 5). Both SAO and GSFC give values for all the quantities defined in Kaula (1966a) and repeated in chapter 9, although most of the quantities are significant only for—and defined for—the case where the representation of the

[3] Kaula (1966a) broke the quantity $\bar{g}_{sn}$ (actually $\overline{\Delta g}_{sn}$, but, as was remarked earlier, the distinction can be ignored in this discussion) into several components: $\bar{g}_{sno}$, the value of $\bar{g}_{sn}$ that would result if a complete and correct representation in terms of associated Legendre functions were available; ϵ_{gn1}, the error caused by errors in those coefficients $\{C_n^m, S_n^m\}$ explicitly present; and $\bar{\epsilon}_{gn2}$, the error caused by defining certain C_n^m, S_n^m to be zero, i.e., omitting certain terms. Such a breakup has no importance for the present evaluation, since only $\bar{g}_{sn}$ is relevant to the program's objectives. The values given by SAO and NASA for Kaula's components therefore need not be considered.

gravitational field is derived from observations only on satellites. The values given for the mean square error in average gravity anomaly are not worldwide in distribution—that is, do not cover the complete globe—but are restricted regions for which at least one average value over a $1° \times 1°$ figure is known in each $5° \times 5°$ figure for SAO and 10 values in each $5° \times 5°$ figure for GSFC. The result of such selection is to concentrate on regions in which gravity has been measured accurately. The greatest s.d. used is 5 mGal and is for $5° \times 5°$ figures in which at least $1° \times 1°$ figure has an average gravity anomaly associated with it. The values given are therefore much too low if we wish to evaluate the representation globally, and a global basis is, of course, the one that makes most sense geodetically and orbitally. The average values computed from the representations are the values of gravity computed at the center of each figure. This introduces another error, although a small one. Also, it must be kept clearly in mind that quantities in the tables are all given to 1 in the units place, implying that the various gravity and gravitational quantities are good to 0.5 mGal at least. This implication is simply not true. The most thorough analysis of the errors in average gravity anomalies is probably that of DMA/AC (Decker, 1972), and that analysis indicates that except in regions with a high density of recent gravimetric data (and such regions are not numerous), 1–mGal errors are exceptional.

11.4.2.2 Evaluation by Comparison With the Gravitational Field

In section 11.4.2.1, the quantities considered were average values over regions $12° \times 12°$ in area. For many applications, the value g_{Tn} (λ,ϕ) of gravity at locations $(\lambda,\phi;H \equiv 0)$ is more important than $\bar{g}_{Tn}$. But comparison of g_{sn} with g_{Tn} would serve no useful purpose, since $\bar{g}_{sn}$ is calculated from a series that has been truncated long before even the contributions of the gravimetric data to it can be reproduced. So it makes more sense to compare g_{sn} with averages of g_{Tn} over regions of 1×1 degrees, since the $\{C_n^m, S_n^m\}$ were derived from such averages. But if we do this, we find immediately that only in small regions can $E(g_{sm} - \bar{g}_{Tn})^2$ be expected to be small. SAO (ch. 9) compared g_{sn} with $\bar{g}_{Tn}$ along fixed latitudes and longitudes in North America and in the Indian Ocean. The rms difference was about ± 4 mGal. If R. L. Decker's (1972) evaluation of gravimetric data is anywhere near accurate, there are only limited regions in North America, Europe, and Africa where one can expect to find $[E(\bar{g}_{Tn})^2]^{1/2}$ to better than ± 5 mGal. The value is between ± 5 and ± 15 mGal for most of the land area and greater than ± 15 mGal for most of the oceanic areas where g has been measured at all. Part of each of these values results from measurement error and part from error of representation. It is not necessary to make the breakdown, however. It is enough to conclude that a representative value of $[E(\bar{g}_{sn} - \bar{g}_{Tn})^2]^{1/2}$ is at best ± 15 mGal globally and ± 4 to ± 5 mGal in regions where accurate data have been used.

The values computed by using orbital theory can be compared with an independent set derived by Hajela (1973) using gravimetric data. Hajela's values range up to 8 mGal. In the northern hemisphere, 3 mGal is representative of the rms differences; in the southern hemisphere, 6 mGal. The range of 3 to 4 mGal for the rms difference in average gravity anomaly from orbital theory plus gravimetric data is therefore reasonable on a global basis.

11.4.2.3 The Model of Rapp

In the previous discussion, R. Rapp's model (ch. 3) was not included because the model is based on an earlier version of GEM 3 and on gravimetric data which do not differ greatly from those used in GEM 6. The major differences, if there are any, between Rapp's model and GEM 6 must result from the different procedures used in adjusting the data. An analysis of the procedures shows that they cannot lead to results that differ by more than the uncertainties that already exist in the results. The evaluation of GEM 6 can therefore be taken as applicable to Rapp's model as far as attainment of NGSP's objective is concerned. The question of reliability of the $\{C_n^m, S_n^m\}$ is of course another question. This was covered previously.

Rapp has, using H. Moritz's "collocation method," combined the coefficients $\{C_n^m, S_n^m\}$ of GSFC's GEM 3 with the average gravity anomalies compiled by DMA/AC (ACIC). The results are close enough to those obtained by means of the usual method of least squares that any advantage in using a collocation method rather than the usual one is completely obscured by the uncertainties in the results.

Rapp's solutions are valuable, nevertheless, for their independence of method. Of course, since they use much the same data as the various GEM's, one would expect the various models to be close to each other. In the sense that they differ less from each other than they do from APL's and SAO's models, this expectation is true. To the extent that GEM 6 is based on more data than Rapp's coefficients are, one would expect GEM 6 to be superior to Rapp's coefficients. This expectation cannot be either proved or disproved on the basis of analyses made so far.

11.4.3 Evaluation by Comparison of Computed Orbits With Observations

The NGSP was intended to provide *geodetic* information. For this reason, the specification on gravitational errors was written in terms of a geodetically meaningful concept—the average value of gravity over an area of given size. This average, if known, together with a few other data, can be converted into approximations to the height of the geoid above a selected spheroid. The primary basis for evaluation of the gravitational part of the NGSP's results must therefore be in agreement with gravity.

As a secondary basis for evaluation of the gravitational part of the NGSP's results, *any* observable effect of the gravitational field may be used. The drawback to using such an effect is that it is contaminated by the presence of factors other than the gravitational field, so that these factors must be accounted for. The most readily available and observable effect of the gravitational field, as far as participants in the NGSP were concerned, was the orbital motion of spacecraft, with the observables being the directions or distances to the satellite at known times. The rate of change of distance between station and satellite can be computed from measurements of the Doppler shift in radio waves emitted by the satellite (ch. 2 and ch. 5). Since the relation between Doppler shift and component of velocity is simple, the component can be treated, to the accuracy we are concerned with, as if it were an observable. Hence there are three "observables" available for checking the accuracy of the gravitational field. They were used by NASA/GSFC and by SAO (which used only direction). The results are summarized in Table 11.9.

SAO has computed orbits for GEOS–A, GEOS–B, and D1D using SE III. The orbits gave residuals, over 2-day periods, whose rms values varied from 2 to 17 m, with almost 50 percent between 5 and 10 m. Since the satellites used were the same ones used in deriving the $\{C_n^m, S_n^m\}$, the results must be considered an *indication* of the accuracy of the gravitational field computed from the $\{C_n^m, S_n^m\}$. They do not, as shown earlier, tell anything about the accuracy of the coefficients, and, as SAO carefully points out, they result from errors in many quantities other than the coefficients.

SAO, in chapter 9, explains the difficulties that prevent satisfactory evaluation by orbital analysis. A very important difficulty not mentioned is that of obtaining independent data. It should be remembered that the orbits on which GSFC's and SAO's results depended were obtained by the adjustment of values of over 600 independent constants exclusive of the orbital elements, that many disagreeable observations were discarded, and the weighting of equations was not entirely objective. With this amount of freedom available for computing orbits, any residuals computed on the basis of observations already used in the adjustment cannot be expected to tell much about the accuracy of the constants. Even residuals from observations not so used must be suspect if they refer to the same satellites. For these reasons, the rms variations in residuals quoted by GSFC and SAO are unsatisfactory. GSFC, in chapter 5, gives an rms variation of 2''74 for orbits computed using GEM 6 and 2''37 for orbits computed using GEM 5. GSFC also used disstances measured by laser systems. The rms variation when using GEM 5 was 1.54 m; when using GEM 6 it was 1.64 m. These variations were for periods of 5^h rather than the 168^h within which camera data were used by GSFC and the 48^h used by SAO. The accumulation of errors in the orbits should not follow a three-dimensional, random walk, so the effects of discrepancies in periods cannot be reliably accounted for. Because of the short period covered by the arcs in GSFC's computations of residuals in range, the errors in the gravitational field cannot have much effect on the orbit or at least could be expected to be swamped by effects of errors in location of instrument, inadequacies in theory, etc.

For these reasons—those given by SAO and those given above—we must conclude that the evidence so far available for evaluating the models on the basis of orbits' accuracies is inadequate. Not only must we have completely independent data available, but there must be enough of these data that the errors in the gravitational field can be reliably separated from those in other constants. The tables provided in chapter 5 do show, however, in their comparison of variations, anomalous behavior from model to model, and further investigation to explain this behavior is urgently needed. (The tables give results using SE II. Variations have also been computed using SE III. The rms variations are slightly larger than those for SE II. However, because of the anomalous behavior mentioned, the increase may not be significant.)

11.4.4 Evaluation by Reference to the Geoid

The geoid is in theory derivable if gravity is known over it and if a connection between it and a suitable spheroid (center of mass at origin, etc.) can be established at one spot. The differences between a geoid calculated from one of NGSP's gravitational models and one calculated from astrogeodetic and/or gravimetric data could therefore be used as an indication of how good NGSP's representation is. But unless the geoid used as reference is considerably more accurate than the one to be evaluated, the comparison is not going to tell very much. There are unfortunately no geoids of this kind available. There are detailed representations over limited areas, such as North America (Fischer *et al.*, 1967), Europe (G. Bomford, 1972), and Australia (A. G. Bomford, 1971). But these are representations of geops (equipotential surfaces) which are either not part of the geoid or are connected to it by satellite-connected data. The differences therefore contain systematic errors which cannot themselves be computed.

Global gravimetric geoids are available—e.g., Uotila's geoid (1964). Unfortunately, these geoids are based on data which are a subset of the data used in producing GEM 6 and SE III. This makes their use as referents undesirable because the differences found could not be interpreted unambiguously.

The best standard of comparison, as far as independence of method of derivation is concerned, is an astrogeodetic geop, since it is least influenced by the *values* of gravity. There are objections to using such a geop—e.g., difficulty of connection to other geops, rapid rise of rms error with distance from datum point under certain conditions, and so on. However, Rapp (1973) has compared geoids computed from his own, GEM 6's, and SE III's sets of $\{C_n^m, S_n^m\}$ with astrogeodetic geops for North America and Australia. His comparison shows rms differences of

±2.0 m, ±2.2 m, and ±2.6 m between the astrogeodetic geoid in Australia and his revised model, GEM 6, and SE III. It shows corresponding values of ±4.4 m, ±3.9 m, and ±6.1 m with respect to the astrogeodetic geoid in North America. As remarked previously, geoid (or geop) is not at present a satisfactory concept for comparison because it is difficult to get definitive connections which are also completely independent of data from satellites but still accurate enough for the purpose. (Rapp's revised model uses GEM 5 rather than GEM 3 as support for the gravimetric data.)

11.4.5 Summary of Evaluation of Gravitational Field

The NGSP has produced two major representations of the gravitational field—GEM 6 and SE III—three, if GEM 5 is considered an independent representation rather than only a minor variant of GEM 6. These two models have been evaluated (or, more correctly, considered) in relation to each other and to models of lesser importance, in relation to areal averages of gravimetric measurements, and in relation to their effects on orbits calculated from them. The first and most important conclusion from these evaluations is that the models produced are considerable improvements over those available at the start of the NGSP and that the general objectives of the NGSP have, in this respect, been more than satisfactorily met.

But if the specific objective of the NGSP requiring a certain accuracy of the models is considered, we cannot say with certainty that the results are satisfactory. We do know that the individual coefficients differ from model to model by amounts which are much too large. These differences indicate that the results are less representative of the gravitational field than they are of the gravitational field plus the combination of observations and orbits involved. Part of the reason for the discrepancies must also be attributed to the fact that the procedures used in reduction are such that harmonic analysis does not correctly separate the effects. Consequently, the different models involved different numbers of terms; the effects of the residuals were distributed differently among the corresponding coefficients. This is a clear indication that the present method of representation is inefficient and inadequate.

We also know that while the average gravity anomalies computed from the models are close to those obtained from gravimetric data, the rms error is not sufficiently below the 3 mGal required for us to be certain that the objective has been met. In fact, the estimates available would indicate that the rms error is closer to 4 mGal than to 3 mGal, over a 12×12 degree square. (Note that most of the values given in tables 5.60 and 9.43 are not relevant to the basic objectives.)

Finally, we must conclude from a study of the residuals from observations on the satellites that the gravitational fields provide orbits which may be satisfactory considering the rms errors in the observations themselves, but that the data presented to support this conclusion are insufficient and inconsistent and have not been adequately analyzed. NASA/GSFC has, it must be said, been extremely conscientious in comparing computed distances, angles, velocities, etc., with observed quantities. But these comparisons have not been completely consistent as far as periods of time covered are concerned, and the information available for separating the contribution of the model from the contributions of other factors is inadequate. Much more work must be done to provide information for evaluating the models, and in any event the same criteria and methods of evaluation should be used by all participants, in particular by GSFC and SAO. No such common bases were used in the NGSP.

11.5 THE GEOID

Approximations to the geoid have been determined by four participants: APL (ch. 2), OSU (ch.3), NASA (ch. 5), and SAO (ch. 9). APL's geoid, being derived without reference to gravity on the surface, is useful primarily for evaluating the APL 5.0 potential. It cannot be considered a useful representation of the geoid, since it is not tied to surface measurements. As APL points out, its geoid does agree with Gaposchkin and Lambeck (1970) to within 10 m. This is a good approximation for a geoid derived entirely from one kind of tracking data. The agreement with SAO SE III is approximately the same.

There are many ways of comparing geoids. None in use at present is particularly convincing. The most common method is to take differences at equal intervals of longitude and latitude. Unfortunately, this will indicate that differences exist even though these differences may be caused only by one representation's being out of phase with the other. If the differencing is carried out at closely spaced intervals, contours of equal intervals can be drawn.

The diagram (fig. 5.38, ch. 5) comparing heights in GSFC '73 with heights from the geoid of Vincent and Marsh (ch. 5) shows that the surface determined from GSFC '73 is systematically lower than that of Vincent and Marsh.

11.6 EVALUATION OF THE PERFORMANCE OF SATELLITE-TRACKING SYSTEMS

One does not have to know how good a piece of equipment is in order to use the equipment and get useful service from it. This is especially true if, like satellite-tracking systems, the equipment is unique and known to represent merely one stage in the development of a rapidly advancing art. But one must have an estimate of the suitability of the equipment if one wants to make sure that the equipment develops and does not remain technologically retarded. Evaluation is therefore an essential part of the total knowledge about an instrument, as important as the manual of operation or the set of calibration constants.

Just what constitutes an evaluation depends partly on what one wants it for. There may be one evaluation of the performance of a tracking system if the data of the system are to be used only for orbit determination, another if the data are to be used to determine station coordinates, and a third if one is not sure of what the data will be used for. But what goes into an evaluation depends also on what one has available for making the evaluation. This is particularly true of satellite-tracking systems, where the system often includes not only the tracking station, but also the satellite, and where it is hard to find a standard against which to judge the system. For this reason, those evaluations that have appeared so far and which are reported in chapters 5 and 6 do not provide simple answers to the question of how "good" any system is, but say, "This system, if used for this purpose and compared with that system under these conditions, shows such-and-such differences."

It was pointed out earlier (sec. 11.1.3.1) that the characteristics of a tracking station are to some extent determined by the characteristics of the satellites with which they are used, or, what is almost the same thing, the characteristics of a tracking station determine what kind of satellite it can be used with. A fair comparison of the performance of one station with the performance of another therefore requires not only collocation of the stations, but also their use on the same satellite at nearly the same time. This is always difficult; it is often impossible. Even where near identity of measuring conditions could be found, there would be the fundamental difficulty that some of the stations measure angles, some distances, and some velocities. One cannot compare these data directly with each other; it is an apples-and-oranges situation. Any common standard of reference that is found may be so far removed from the basic data put out by the station, that the validity of the final comparison is hard to prove. For instance, laser DME can be compared with 6–GHz (5–cm) radar DME by locating the stations next to each other and then measuring distances to the same satellite at nearly the same time. Since the true satellite distances are not known and since refraction effects, and so on, are different for the two instruments, these factors must be accounted for in the comparison. But if one instrument measures distance and the other angle or velocity, it may be necessary to compare the ability of the instruments to give the correct satellite coordinates, which means bringing into the problem the orbit of the satellite. Then the comparison depends not only on the performance of the system *per se*, but also on the particular mathematics used in describing the orbit.

Since direct comparison of all tracking systems is impossible, comparison between systems of like natures must be considered the best that can be done. To go further, a common and valid standard or reference must be found. Only two standards need be considered: (1) the location of the satellite and (2) the location of the tracking instrument. Consider the several ways of getting a comparison on the basis of satellite locations.

(1) The data from each system are used to determine satellite locations in the manner best suited to that particular system: by simultaneous observations, by empirical curve fitting, or through orbital theory. Satellite locations for the same times are then compared. Insofar as a system is intended to be able to produce satellite locations, this is a fair method of comparison. It is not a valid method unless each system is used in its best geometric arrangement (for example, as a group of three SECOR stations arranged to form, with the satellite, a tetrahedron, or a set of two camera-type AME arranged to form, with the satellite, an equilateral triangle, and so on) and with suitable satellites. Assuming that this can be arranged, we then compare the systems on the basis of the standard deviations of those satellite locations determined by the data.

While this procedure is optimal in many respects, it is unsatisfactory in many others. It gives no answers to questions about the accuracies of the systems when used on satellites in general or about what the accuracies will be when the systems are forced to operate in geometries other than the best ones. Furthermore, there is no way of separating the effects of the theory on the performance from the effects of instrumental errors.

(2) The data from each system are all used in the same way to produce satellite locations—that is, by insertion of the data into the equations of the orbit, with the same orbital theory being used for all tracking. This method ensures that the systems alone are being compared so that any differences found do not arise from differences in theory. But now the hosts of error are attacking on the other flank. Every system has a best way of being used, and we are denigrating the performance of a system by forcing its data to conform to the same treatment as those of the others.

(3) The data from one particular system may be used in computing the orbit, and the values of the observables for each other system may then be computed by using that orbit. The "accuracy" of a system is then determined from formula (11.3), where y_{oi} are the computed values. This method is the one used by NASA/GSFC and NASA/Wallops Flight Center (WFC). It is excellent if the system used as standard is considerably more accurate than the other system involved and if performance under less than optimal conditions is wanted. For some of the equipment (such as FME) the conditions may indeed be much less than optimal. (For example, an insufficient number of stations may be used to determine the orbit, or the passes available may have poor geometry.) This method, therefore, can lead to misleading results.

The most important characteristic of a tracking system is its error—or at least its actual error σ_m compared with the error σ_r allowed for it. If the system measures values y_{oi} ($i=1$ to I) of an observable whose actual (true) values were y_{ti}, the measurement error of the system is defined to be

$$\sigma_m = \sqrt{\frac{\sum_i (y_{oi} - y_{ti})^2}{I}} \qquad (11.3)$$

The error performance or quality index of the system is the ratio

$$\rho = |\sigma_m/\sigma_r| \qquad (11.4)$$

where σ_r is the error allowed in or specified for the system.

A tracking system (ch. 1) measures one or more of the following quantities: phase angle, travel time of a pulse, frequency, or Cartesian coordinates of a point in a photographic image. Measuring these quantities can be considered equivalent to measuring the distance or direction of a satellite from the observer, or measuring that component of its velocity which is in the direction of the observer. But there is no way known at present of finding out the true values y_{ti} of these quantities. Equation (11.3) therefore cannot be used for evaluating the performance of a tracking system.

No completely satisfactory alternative has yet been found, or at least none has been used. The best alternative is to consider, as the most important characteristic, not the accuracy of the measurements of the system but the accuracy of the final results. With this criterion, the formulas for σ and ρ remain the same, but the definition of y_{oi} becomes different. Now y_{oi} is the value of a quantity for which accurate values are usually available.

The quantities $\{y_{oi}\}$ are now available, whereas they were not available if one used the first criterion. Furthermore, σ_r is now related directly to the user's needs rather than to the distance, direction, or velocity of the satellite, which are for most people only of minor interest.

One group has been engaged since the start of the NGSP in evaluating performance of the tracking systems used in the program. This group, under the direction of J. Berbert of NASA/GSFC, actually had two objectives: to evaluate the performance of the tracking systems and to calibrate those systems which were NASA's. Since the principal criterion used by the group for evaluation was the accuracies and precision of the systems, the procedures used in evaluating performance were in many cases the same as those used for calibration (except where calibration was done without satellites), and the results naturally also overlapped. But the objectives were not identical, and the results, while related, were not directly usable in either context. Berbert's group was more concerned with the problem of calibrating the instruments than with the problems of evaluating (comparing) them. As a consequence, the group's results, given principally for the calibration objective, do not convert readily into numbers that can be interpreted for evaluation (comparison). Table 11.10 gives values taken, with occasional slight changes, from the group's reports and intepreted as precision.

Note that the NASA group adopted the second approach discussed earlier. Instead of comparing measurements directly, the group assumed that the systems were intended to provide data for computation of orbits and compared measurements against quantities computed from the orbit. (The collocation tests did approximate direct comparison.) In the terminology of the group, the calibrations constants for a system are named: "zero-set bias coefficient" (symbol B_0) and "time bias coefficient" (symbol B_T), or "zero-set bias" and "timing bias." These constants appear in the linear equation

$$(y^i_{\text{obs}} - y^i_{\text{comp}}) = B_0 + \left(\frac{dy^i_{\text{comp}}}{dt}\right) B_T + \epsilon$$

where y^i_{obs} (y^i_{comp}) is the i^{th} measured (computed) value of the observable and ϵ is the measurement error. From the way in which B_0 and ϵ enter into the equation, the two quantities obviously cannot be separated unless some assumptions are made about the nature of B_0 and ϵ. The group assumed that B_0 was constant over long periods (one pass or longer), while ϵ varied randomly from measurement to measurement.

With respect to the data provided by the GEOCEIVERS (ch. 3), we should note that the stations involved in the test were located on or close to first-order control points. Many were on the precise traverse. If we accept the values derived for their coordinates, then the distances between stations are good to about 1 part in 10^6 (B. K. Meade, private conversation, 1973). The data from the stations will therefore be important in the new adjustment of the North American triangulation.

Conclusions on the Evaluation of the Tracking Instruments.—The work done by NASA/GSFC and NASA/WFC has provided a very large amount of potentially useful information on the errors associated with some instruments. The potential has not, unfortunately, been fully realized during the lifetime of the NGSP, and the natural process of evolution in instruments is fast rendering most of the information obsolescent. The work did make evident the capability of 5-cm radar for greater precision than was generally thought, and it did provide reasonable starting values for the σ's of observations with the instruments. It is obvious from the results that the σ of 5-cm radar data can be reduced still further. It is also obvious that a large systematic error must still be present in data from SECOR, although the error cannot be considered serious since (1) SECOR is no longer being used and (2) the data can be corrected by using OSU's method (ch. 8).

The evaluation done by NASA/GSFC is particularly valuable in showing how future work of this kind can be improved. The methods that were used extracted only part of the information present. Because of this and because the statistical methods used sometimes gave ambiguous results, there was—and still is—some disagreement over the interpretations of these results. The work done by Berbert *et al.* is therefore an excellent basis on which to build more powerful methods in future evolution of the performance of an instrument.

11.7 SUMMARY OF EVALUATION

The broad purposes of the NGSP, to get substantially improved coordinates for tracking stations, to get an improved model of the gravitation field, and to compare the performance of the tracking systems involved, can be said to have been well satisfied. There is no doubt that the six major datums on which most of the tracking stations were located (NAD 1927, SAD 1969, EU50, Adindan, and AGD 1965) have been related to one another to within ± 20 m. Solutions for individual sets of tracking stations (those contained in NWL-9D, GEM 6, WGN, WN14, and SE III) give standard deviations for the coordinates of important stations that are well below ± 5 m. For a selected number of stations, the rms error in the coordinates is probably less than ± 10 m regardless of the set from which they are taken.

As for the gravitational field, the coefficients for which values have been determined have been extended to beyond a full $n=16$, $m=16$, and in land areas at least, the average value of gravity over a 12×12 degree region can be computed to about ± 4 to ± 5 mGal.

In an extensive series of tests, comparisons have been made of the performance of BC-4, Baker-Nunn, and MOTS cameras, of MINITRACK, of SECOR, 5-cm radar, GRARR, and laser systems, and of the TRANET Doppler systems. By comparing observations against values computed from "standard" orbits, instead of against each other, the participants avoided the "apples-versus-oranges" difficulty.

The values derived for instruments' differences are probably better than $\pm 1''$ for the cameras, ± 1 m for the ranging instruments, and ± 1 cm/sec for the instruments measuring range rate. The values derived for biases in the data must be even better, since they are themselves averages and would be approximately as good as the standard deviations of the observations, divided by the square root of the number of observations.

As regards the *specific* objectives of the NGSP, the situation is less satisfactory. The first specific objective was to get the positions of enough stations on the major datums to ± 10 m (in each coordinate) in a geocentric system to enable these datums to be tied together to approximately the same accuracy and to get coordinates of other stations also to ± 10 m. It was implicit in the statement of the objective that positions of other points on these datums would then be fixed to the geocentric system either directly or, in most cases, through conventional survey on the local datums. The requirement may have been met, as was stated previously, for a selected set of stations and for six datums. It has not been met for all stations or for all the major datums. Furthermore, the assumption that it has been met for the six is a shaky one, since the assumption is first made that agreement of the participants on the values involved (to within the allowed uncertainty) is sufficient.

The discrepancies between coordinates of these and other tracking stations as determined by the various participants are so great, for the most part, that acceptance of one participant's set of coordinates will result in rejection of the coordinates of many stations determined by others. Every bit of evidence points to the conclusion that all the sets except perhaps one contain systematic and unexplained errors. The evidence seems to indicate that some of these errors lie in the definitions of the coordinate systems used, but that other error sources are also present. There is no doubt whatever that many of the sources could have been tracked down and the errors eliminated had the preliminary computations of results been more freely circulated and had the error analyses been carried out further. Certainly before a final set of coordinates is selected, the error analyses required to identify the sources of the errors will have to be made. Such analyses should be made in any case because so many geodetic and geophysical projects are being carried out using the methods developed by the NGSP. If nothing is done to clean up the work already done, these will carry within them the same errors that caused difficulties to the NGSP.

The second specific objective of the NGSP concerns the Earth's gravitational field. The evidence produced by the participants indicates that the NGSP has come close to achieving its objective of 3 mGal for $12^\circ \times 12^\circ$ quadrangles. Unfortunately, neither of the two major participants concerned with the gravitational field presented data relating directly to the NGSP's stated objective, and the data which were produced were, at best, inconclusive. The difficulties of evaluation were aggravated by lack of commonality between SAO and GSFC in test objects, at least as far as their use by SAO was concerned. GSFC did carry out parallel computations, using GSFC's and SAO's models, and these computations show a slight superiority of the GSFC model for computing orbits. But GSFC's data adduced for evaluation contain some anomalies that require explanation anyway, so nothing definite can be concluded.

As in the case of the conclusions about the geometric results, the gravitational results can be said to be capable of considerable improvement. Most of this improvement should come about by a definitive analysis of the errors. Considering the number of important and unanswered questions still existing in regard to the validity of the accuracies of the results, such an error analysis must be considered essential.

An inspection of the values of the coefficients $\{C_n^m, S_n^m\}$ shows, first, that the individual values for most of them have relative uncertainties of well over 50 percent, even as low as degree 6, while many of them have relative uncertainties of several thousand percent (based on differences from Rapp's model.) Representations of the gravitational field by series of spherical harmonics must therefore be considered completely unreliable. Since the various models are, however, well able to predict average gravity anomalies, as was mentioned previously, and to provide orbits that fit well to observed data, it must be concluded that much of the fault lies with the method of representation. Considered as a predictor of gravity anomalies, the various models are of course less successful than they are as predictors of average values, and an rms error of ± 15 mGal or poorer must be expected in all but a few land areas.

Evolution of the instrument performance, as contrasted to comparison of performance, was not an objective, but should be possible from the data accumulated and results obtained. There is not agreement among participants, or between participants and the editor, on the relation of these results to the precision and accuracy of the instruments. Since the objectives of the program have been met in a broad sense, if not in a narrow one, since many of the participants are satisfied with their own solutions, and since much of the equipment is by now obsolescent with no thought of future use, final evaluation of the instruments has no immediate importance. Final evaluation is important to the extent that without it we cannot claim that the results of the NGSP are either complete or completely understood.

Although this evaluation shows that only the general objectives of the NGSP have been reached, such a statement is of little value without a further statement of why the program did not reach its objectives and of what (if anything) should be done.

First, it is obvious that the specific objectives, although not reached, are yet within sight from the NGSP's terminus. The accuracy of coordinates is surely within 10 meters of where it should be, and the accuracy of the average anomalies within 2 to 3 mGal. Had the NGSP continued another two years, and had appropriate steps been taken, the accuracies could probably have been brought to the deserved values.

There seem to be only three important reasons why accuracies were not attained during the NGSP. One arises from a fundamental rule laid down at the very start of the program—that in order to ensure that the results should be independent and hence usable as checks against each other, the participants should work along independent lines. This rule, excellent in purpose and concept, was unfortunately adhered to with a fixity that preserved independence but prevented full cooperation in the tracking down of sources of disagreement. The second reason is that the error analyses carried out by the individual participants have not been of the depth and sophistication needed to completely support the results. The need for such depth and sophistication was of course not apparent until too late, because cooperation was not close enough to show that significant discrepancies were going to occur. And the third reason was, of course, that the discrepancies and the need for their explanation became obvious too late for the participants to take steps to do much about it. The NGSP ended at that point where the participants had just discovered the magnitudes of the discrepancies and realized the need for reducing them.

To what extent independence of operation prevented the discrepancies from being anticipated can only be guessed at. In any case, we know that the discrepancies exist, that they are larger than we would like, and that their causes are still uncertain. We also know that the methods that were used in analyzing the data can be further refined to allow deeper analyses of the data. Until such an analysis is carried out and the present discrepancies explained to everyone's satisfaction, processing of more data by present methods is not merely unnecessary, but is undesirable. Since present results of the NGSP are open to objections that prevent their being considered as meeting the program's objectives, since these same objections will affect acceptance of future results obtained by techniques similar to those used during the program, and since eliminating the objections will also provide the specific objectives for which the NGSP was designed, we must consider a deeper analysis of the data as necessary for satisfactory completion of the NGSP.

A final word as to the results of the NGSP: The judgment that the objectives of the NGSP have been only partly satisfied is true only with respect to the most stringent requirements that were imposed by NASA. If the more liberal and general requirements that were also put down by NASA—that the program lead to substantial improvement in the number and accuracy of geodetic locations and in the knowledge of the Earth's gravitational field—are considered, then the NGSP has more than adequately met these requirements. The number of control points that can serve for international connections has increased from approximately 20 to over 200, and the rms error has dropped from an estimated ± 50 to ± 100 m to an estimated ± 10 to ± 20 m.

APPENDIX

TABLE 11.1.—*Solutions for Coordinates of Stations*

No.	Model	Reference[a] Original coordinates	Final coordinates
1	GEM 5, GEM 6	8.3, 11.5	5.6
2	NWL-9D	3.3	3.5
3	DMA/AC	8.3, 11.5	3.5
4	NASA/WFC	8.3, 11.5	6.5
5	SEN (SECOR)	8.3	3.5
6	WGN (NGS)	7.3	7.5
7	AFCRL	7.3, 11.5	3.5
8	GSFC '73	8.3, 9.3, 9.5	5.6
9	SE III (SAO)	9.3	9.5
10	WN14 (OSU)	8.3	8.5
J	JPL	11.5	4.5

[a] References are to chapter and section.

TABLE 11.2.—*Comparison of Lengths of Chords as Determined by NGS, NWL, and OSU*

Stations at chords' ends	Datum	Length (m) of chord (original survey) Value	S.D.[a] Orig-inal[b]	S.D. NWL	S.D. OSU	S.D. NGS[b]	Difference from original survey (m) NGS	NWL	OSU
North America	NAD1927								
6002–6003		3 486 363.232	±3.53	---	3.49	1.75	−0.06	+ 2	2.7 ± 2.3
6003–6111		1 425 876.452	±1.59	---	1.59	0.72	+1.50	---	2.3 ± 1.4
Europe[a]	EU50								
6006–6065		2 457 765.810	±0.80	---	3.5	1.23	+0.10	+ 3	6.1 ± 2.0
6065–6016		1 194 793.601	±1.43	---	1.41	0.60	+0.42	− 1	−2.9 ± 1.3
Australia	AGD								
6023–6060		2 300 209.803	±0.88	---	4.60	1.15	−0.98	−11	5.9 ± 3.0
6060–6032		3 163 623.866	±0.98	---	∞	1.58	−2.76	−10	−4.5 ± 3.6
North Africa	Adindan								
6063–6064		3 485 550.755	±2.10	---	4.11	1.75	+2.60	− 1	10.6 ± 2.3

[a] The lengths and standard deviations given by R. Kube and K. Schnädelbach in an unpublished paper presented in 1973 at Athens are as follows:

6006–6065	2 457 765.44 ± 1.2 m
6065–6016	1 194 793.601 ± 0.9 m

For the second of these, the chord from Hohenpeissenberg to Catania, J. C. Gergen and B. K. Meade of the National Geodetic Survey, in an unpublished memorandum of 15 May 1973, give the same length but a standard deviation of ± 1.428 m.

[b] Taken from table 7.3, chapter 7.

TABLE 11.3.—*Comparison of Positions Computed by OSU and by DMA/TC*

Station	$\Delta\phi$	$\Delta\lambda$	Δh	Δh_{com}	Datum
5001	51	61	−8	7	NAD 1927
5648	52	8	−3	14	NAD 1927
5712	−7	−44	−8	9	SAD
5713	−38	53	−26	−9	Azores
5715	−92	3	−27	−10	Adindan
5717	−208	18	−29	−12	Adindan
5720	−261	33	−23	−6	Adindan
5721	−252	58	−35	−18	EU50
5722	+231	17	−16	1	Indian
5723	−164	−20	−19	−2	China
5726	−136	13	−18	−1	Luzon
5730	22	−14	−11	6	--------
5732	−55	29	34	−17	--------
5733	35	32	31	−14	Pac. Mid
5734	48	−91	−22	−5	NAD 1927
5736	122	−66	−12	5	Atl. Mid
5739	−38	53	−25	−8	Azores
5744	−120	89	−46	−29	EU50

Spheroids used:

	a	b	f
DMA	6378155 m	6356770.1 m	1/298.255
OSU	6378155 m	6356769.7 m	1/298.249

TABLE 11.4.—*Relationship Between Major Geodetic Datums and Systems Used for NGSP*

	Model[a]						
	DOD/NWL (2)	NGS/WGN (6)	GSFC '73 (8)	GEM 6 (1)	CNES	SAO/SE III (9)	OSU (10)
Adindan	−150 ----	−163 ----	------- -----	−147 -----	------- ----	------- ----	−184 ± 19 -----
	− 31 ---- ----	− 34 ---- ----	------- ----- ----	− 3 ----- ----	------- ---- ----	------- ---- ----	− 21 ± 11 ----- ----
	+199 ----	+207 ----	------- -----	211 -----	------- ----	------- ----	+200 ± 6 -----
ARC (Cape)	−120 ----	------- ----	------- -----	−126 -----	------- ----	------- ----	−152 ± 7 -----
	−128 ---- ----	------- ---- ----	------- ----- ----	−110 ----- ----	------- ---- ----	------- ---- ----	−126 ± 7 ----- ----
	−296 ----	------- ----	------- -----	−296 -----	------- ----	------- ----	−298 ± 10 -----
Australian	−125 ----	−124 ----	−137 34	−135 −100	------- ----	−117 ----	−118 −103
Geodetic	− 30 ---- ----	− 61 ---- ----	− 50 18 1.9	− 39 −120 2.4	------- ---- ----	− 39 ---- 2.44	− 41 − 99 1.2
(1965)	+148 ----	+145 ----	155 38	133 40	------- ----	+120 ----	+121 + 25
EU50	−729 ----	− 96 ----	−149 60	− 83 60	− 61 ----	− 87 ----	−134 41
	−105 ---- ----	− 79 ---- ----	−103 190 5.0	−116 40 −0.3	−128 ---- ----	−111 ---- ----	−153 − 27 7.2
	−121 ----	−126 ----	− 93 65	−120 −60	−150 ----	−134 ----	−145 51
Indian	+253 ----	------- ----	------- -----	------- -----	------- ----	------- ----	+165 ± 17 -----
	+291 ---- ----	------- ---- ----	------- ----- ----	------- ----- ----	------- ---- ----	------- ---- ----	+711 ± 10 ----- ----
	+359 ----	------- ----	------- -----	------- -----	------- ----	------- ----	+228 ± 11 -----
NAD 1927	− 29 ----	− 32 ----	− 43 −100	− 24 − 20	------- ----	− 31 ----	− 57 − 86
	+161 ---- ----	+121 ---- ----	162 − 20 0.9	+151 10 14	------- ---- ----	+155 ---- 2.4	+148 − 23 −0.8
	+183	+173 ----	179 − 5	+187 − 80	------- ----	+179 ----	+186 − 33
SAD 1969	− 77 ----	− 44 ----	− 44 74	− 63 60		− 73 ----	− 54 + 63
	0 ---- ----	2 ---- ----	+ 8 25 1.8	0 20 ----	------- ---- ----	− 3 ---- .86	− 30 17 −6.7
	− 43 ----	− 44 ----	− 46 28	− 32 0	------- ----	− 50 ----	− 43 + 12
Tokyo	−135 ----	------- ----	------- -----	−147 -----	------- ----	------- ----	−183 ± 10 -----
	+528 ---- ----	------- ---- ----	------- ----- ----	509 ----- ----	------- ---- ----	------- ---- ----	+506 ± 9 ----- ----
	+670 ----	------- ----	------- -----	686 -----	------- ----	------- ----	+686 ± 9 -----

[a]

ΔX (m)	Rx (in $'' \times 10^3$)	
ΔY (m)	Ry (in $'' \times 10^2$)	scale difference $\times 10^6$
ΔZ (m)	Rz (in $'' \times 10^2$)	

TABLE 11.5.—*Comparison of Systems Used for Satellite Geodesy in NGSP Systems*

	Model[a]											
	WGN[b] (6)			GSFC '73 (8)			GEM 6 (1)			WN14 (10)		
NWL-9D	21	−11		---			−6			16	29	
(2)	19	− 5	2.5	---			3	− 3	−0.3	10	71	−0.14
	−16	64		---			8	−59		−3	−15	
WGN	---			---			---			[b]− 1	− 8	
(6)	---			---			---			7	− 5	−2.3
	---			---			---			12	11	
GSFC '73	---			---			0.5	0		14	96	
(8)	---			---			0.6	4	0.4	13	−30	0.24
	---			---			2.1	35		− 2	19	
GEM 6	---			---			---			21	7	
(1)	---			---			---			11	11	0.4
	---			---			---			2	12	
SE III	18	−12		−1	−4		---			14	−17	
(9)	26	30	1.3	2	−3	−0.6	---			14	37	0.1
	−21	49		−9	8		---			−10	15	
GEM 4	---			---			0.5	0	0	15	93	
	---			---			−0.4	− 2		12	− 2	0.2
	---			---			−0.2	4		2	12	

[a]

ΔX (m)	Rx (in $'' \times 10^3$)	
ΔY (m)	Ry (in $'' \times 10^2$)	scale difference $\times 10^6$
ΔZ (m)	Rz (in $'' \times 10^2$)	

[b] Values specially computed by Computer Sciences Corporation.

TABLE 11.6.—*List of Stations With Acceptable Differences in Coordinates*

Geographic region	Models[a]
European Datum 1950	
St. Michael	8, 9, 10
Nice	8, 9, 10
Dionysos	8, 9, 10
San Fernando	8, 9, 10
North American Datum 1927	
Blossom Point	1, 8, 10
Ft. Myers	1, 8, 10
Beltsville	1, 6, 10
Columbia	1, 8, 10
San Juan	1, 8, 10
Denver	1, 8, 10
Jupiter	1, 8, 10
Mt. Hopkins	1, 8, 10
South American Datum 1960	
Paramaribo	1, 2, 6, 9, 10
ARC (Cape) Datum	
Johannesburg	
Australian Geodetic Datum 1965	
Thursday Island	1, 6, 9, 10
Culgoora	1, 2, 6, 9, 10
Caversham	1, 6, 9, 10
New Zealand Datum	
Invercargill	1, 2, 6, 9, 10
Miscellaneous (Minor) Datum	
Mahe	1, 2, 6, 10
Mauritius	1, 6, 10
Heard	1, 6, 10
Wilkes	1, 6, 10
Zamboango	1, 6, 9, 10
Christmas	1, 2, 6, 9, 10

[a] 1—GEM 6
2—NWL-9D
6—NGS/WGN
8—GSFC '73
9—SAO SE III
10—OSU's WN14

[a] The 7-m requirement is relaxed when only one coordinate is involved and the excess is less than 10 m.

TABLE 11.7.—*Summary of Differences in Coefficients Using Rapp's Model (Chapter 3) as Referent*

Type of comparison		Representation				
		GEM 5	GEM 6	SE III	APL 5.0	Kock and Witte
Terms included		12 × 12	16 × 16	16 × 16	12 × 12	4 × 4
		Differences[a]				
$n = 1 \sim 4$	$m = 0$	11/6	11/6	8/6	171/90	33/16
	$m = 1 \sim N$	90/22	22/10	229/80	171/41	435/190
$n = 5 \sim 8$	$m = 0$	18/13	19/13	12/8	320/203	---
	$m = 1 \sim N$	92/19	66/26	305/121	250/100	---
$n = 9 \sim 12$	$m = 0$	---	266/133	16/7	231/129	---
	$m = 1 \sim N$	---	97/30	297/80	268/99	---
$n = 13 \sim 16$	$m = 0$	---	29/6	24/16	---	---
	$m = 1 \sim N$	---	71/26	151/39	---	---
Percent differences, number						
Between 0–20		56	48	38	30	---
20–40		13	14	9	15	---
40–60		4	4	6	8	---
60–100		3	3	2	7	---
>100		--	3	17	17	---
Largest percent up to (8,8)						
$m = 0$		29 ($n = 8$)	29 ($n = 8$)	19 ($n = 8$)	516 ($n = 8$)	---
$m = 1 \sim N$		83 ($n = 7$, $m = 7$)	671 ($n = 7$, $m = 5$)	920 ($n = 8$, $m = 1$)	816 ($n = 8$, $m = 1$)	---
Differences × 10^8						

[a] Maximum difference/rms difference.

TABLE 11.8.—*Comparison of Gravitational Fields*

	Set						
	Rapp		GEM 6		SE III[a]		
	A	B	A	B	A	B	([b])
	--------	---	26 (49%)	---	58 (72%)	----	$\Delta\left\{\begin{matrix}C_n^m\\S_n^m\end{matrix}\right\}\times 10^9$
Rapp	--------	---	3	3.7	7.2	7.4	[b]ΔN (m)
	--------	---	5.3	7.7	10.0	11.1	[c]Δg (mGal)
	26 (49%)	---	--------	---	56 (80%)	---	$\Delta\left\{\begin{matrix}C_n^m\\S_n^m\end{matrix}\right\}\times 10^9$
GEM 6	3	3.7	--------	---	6.4	6.6	ΔN (m)
	5.3	7.7	--------	---	8.6	9.6	Δg (mGal)

[a] SE III, set B is taken only up to $n = 23$.
[b] $\Delta N \equiv$ difference in geoidal heights, rms value.
[c] Δg = difference in anomalies, rms value.

TABLE 11.9.—*Observational Errors Induced by Gravitational Representations (Average Value of RMS Error)*

Model	Distance	RMS error in direction, (7-day period)	Radial velocity (1-day period)
SAO SE III	4.9[a] (2^d)	----------	----------
GEM 5	1.54 m ($0^d.2$)	$\pm 2''.4$	±5.9 cm/sec
GEM 6	1.65 m ($0^d.2$)	$\pm 2''.7$	±5.5 cm/sec

[a] SAO (ch. 9) estimates that 2 to 3 m are contributed by errors in coordinates. This would still not make SAO's values for the contribution of the gravitational field consistent with GSFC's, which must also contain errors resulting from erroneous coordinates.

TABLE 11.10.—*Precision of Instruments Used for Satellite Tracking*

Instrument evaluated	Precision[a]			Instrument used as standard	Reference[c]
	Distance[b] (m)	Angle (″)	Velocity[b] (cm/sec)		
Camera	------------	±2 BC-4, BN ±1 MOTS, PC-1000	------------	------------	1967 (7)
		±2.8	------------	Camera	1968 (3)
		±3	------------	------------	1966
PRIME MINITRACK	------------	±50 m track ±100 m along track	------------	MINITRACK Camera	1966
	------------	±165 m	------------		1970 (5)
		±20″	------------		1967 (7)
GRARR	7 (l.p.) 12 (s.p.)	------------	------------	GRARR SECOR	1966 (8)
	±10.3 l.p. 7 s.p.	------------	------------	Camera	1969 (3)
	−2 to +2 (l.p.) ±3 to ±5 (s.p.)	------------	------------	Camera	1968 (4)
	±2.5 (l.p.)	------------	------------	Laser DME	1967 (1)
SECOR	1.2 − 6 (s.p.)	------------	------------	Camera	1966
	3.4 (l.p.) 1.7 (s.p.)	------------	------------	Laser DME	1968 (2)
	(−3 to +43) l.p. ±1 to ±6 s.p.	------------	------------	Camera	1968 (4)
	±10	------------	------------		1967 (7)
FPQ-6	5 (l.p.) 1 (s.p.)	------------	------------	Laser DME	1968 (2)
	5	------------	------------	Laser DME	1969 (9)
FPS-16	3 (l.p.) (s.p.)	------------	------------	Laser DME	1968 (2)
Doppler DME TRANET	------------	------------	4.5/3 cm/sec (l.p.) 5.4/4 cm/sec (s.p.)	Laser DME	1968 (2)
GRARR	------------	------------	5 cm/sec		1967 (7)

[a] At lower frequency/higher frequency.
[b] l.p. = long-period random error, s.p. = short-period random error.
[c] References:

(1) NASA Document X-514-67-447, 1967.
(2) J. Berbert and H. Parker, NASA Document X-514-68-458, 1968.
(3) J. Lerch *et al.*, NASA Technical Note TN-D-5036, 1969.
(4) J. Lerch *et al.*, NASA Document X-552-68-101, 1968.
(5) J. Marsh and C. Doll, NASA Technical Note TN-D-5337, 1970.
(6) R. Agreen and J. Marsh, NASA Document X-552-69-539, 1969.
(7) J. Berbert, NASA Document X-514-67-315, 1967.
(8) NASA Document X-514-66-513, 1966.
(9) Leital and Brocks, *C-Band Radar Range Measurements: An Assessment of Accuracy*, 1969.