GPS for Environmental Applications: Accuracy and Precision of Locational Data

Peter August, Joanne Michaud, Gharles Labash, and Ghristopher 9mith

Abstract

Environmental scientists are beginning to use the Global Positioning System (GPS) for in situ determination of the location of point and line features. The accuracy of data collected by inexpensive cPs unifs can be quite variable. In this study, we estimated the accuracy and precision of a simple GPS receiver and community base station system. We found that, under ideal conditions, 95 percent of the locations we derived were within 73 m of true without differential correction and within 6 m of true with correction. Taking the average of repetitive fixes at a single location increases accuracy and precision, especially if more than 50 sequential fixes are used. There is little correlation of positional accuracies obtained at different stations or between raw and differentially corrected data. There can be measurable day-to-day variation in accuracy that may not be related to PDOP (positional dilution of precision) conditions.

lntroduction

Global Positioning System (crs) technology is rapidly becoming a common method for in-the-field collection of environmental data (Leick, 1987; Long et al., 1991; Slonecker and Carter, 1990; August, in press; August et al., in press). Under ideal circumstances, cPs can provide reliable and accurate position and elevation information. Data collection procedures at this time are, however, not yet turnkey in ease and simplicity (Wells et al., 1986). A simple GPS mission can require significant planning, preparation, and data manipulation to ensure the highest quality results (Puterski et al., 1990; Adkins and Lyon, 1991). There are many sources of error that can degrade the quality of cps-derived positional data. These include obstructions on the horizon, interference of satellite signals by forest canopy, atmospheric disturbances, poor satellite geometry, Selective Availability (sa), and reflection (multi-pathing) of satellite signals (Hurn, 1989; Puterski et al., 1990; Wilkie, 1989). Through careful mission planning and post-processing of field data, it is possible to minimize these degrading effects. It is important that new users of GPS be made aware of the fundamental accuracy and precision of cps-derived data and the sources of error that can degrade positional accuracy.

The purpose of this study is to assess the quality of horizontal position data obtained from an inexpensive cPS receiver system under excellent field conditions. Specifically, we wanted to determine (1) the effectiveness of differential correction (Hurn, 1989) in removing the distortion caused primarily by Selective Availability (SA), (2) how data averaging improved the quality of position estimation, and (3) the observed levels of precision within and between data collec-

tion sessions. Our objective was to determine levels of accuracy of cPs data under near ideal field conditions. We did this by comparing the actual position of two first-order (Antenucci et al., 1991) geodetic control points with positions estimated from 300 replicate GPS fixes obtained during 20 different data collection periods. This study was not designed to identify or measure different sources of error that degrade cps data. Moreover, we did not evaluate the accuracy of elevations provided by crs. Although cPS receivers are capable of providing excellent data on vertical positioning, extreme care must be taken in choosing the appropriate reference datum (geoidal or ellipsoidal) for the application (Wells et al., 1986).

Methods

Hardware and Software

Field data were collected with a three-channel Trimble Pathfinder Basic portabie receiver (firmware version 3.14, Trimble Navigation Company, Sunnyvale, California). Data were logged in the WGS-84 datum with a firmware-imposed PDOP (positional dilution of precision) mask of 4.0, a signal-tonoise ratio threshold of 5, and an elevation mask of 11 degrees. Data used in differential correction were obtained from a Trimble Community Base Station system with a Pathfinder Professional receiver ffirmware version 1.10) located at the Environmental Data Center, University of Rhode Island. Both the base station and rover unit were set to log data in manual gD mode using signals from four satellites. The coordinates for the base station antenna were obtained by triangulation from two first-order National Geodetic Survey control points located within 2 km of the base station (41° 29' 20.17" N latitude; 71° 31' 39.76" W longitude). Three survey-quality Trimble 4000 series GPS receivers were simultaneously operated at the base station location and the two control points for one continuous hour to gather the data for the triangulation. We confirmed the accuracy of the base station antenna coordinate by re-calculating the location using conventional traverse methods. The coordinates obtained from the two independent surveys were within 9.24 cm of each other. The horizon around the base station antenna was essentially unobstructed in all directions (Figure 1).

Data logging at the base station was done using Trimble PFCBS software (version 1.37). Mission planning and waypoint downloading were performed using Trimble PFBASIC software (version 2.0-15). Differential correction was accomplished with Trimble PFINDER software (version 1.46) using

0099-1112/94/6001-41\$03.00/0 @1994 American Society for Photogrammetry and Remote Sensine

Photogrammetric Engineering & Remote Sensing, Vol. 60, No, 1, lanuary 1994, pp. a1-45.

Environmental Data Center, Department of Natural Resources Science, University of Rhode Island, Kingston, RI 02881.

PEER-REVIEWED ARTICLE

the measurement space method (a pseudo-range correction process). The PLOTSSF software (Lawrence Fisher, In Litt) described in Puterski et al. (1990) was used to view graphically the cPS data. ASCII files for statistical analyses were created with PFBASIC software. All Trimble and PLOTSSF software was run from an 80386-25 MHz microcomputer. Data were analyzed with PC-SAS software (SAS Institute, Carey, North Carolina) on an 80386-16 MHz microcomputer.

Data Collection Protocol

All data were collected during conditions when pDoP was below 4,0, as predicted by mission planning projections, The accuracy of mission planning projections for optimal data collection times used in this study was tested on four different occasions by comparing real-time PDOP values reported by the community base station system with those projected by mission planning for the same one-minute interval. The mean deviation between the two was Iess than 0.1 PDOP (SD $= 0.069, n = 157$

For a given sample, the community base station was initialized and positions were continuously logged to a data file at the rate of approximately one fix per second. For every sample session, we collected data at two monumented geodetic control points, both of which meet FGCC (Federal Geodetic Control Committee) standards for first-order accuracy (Antenucci et al., 1991). University station is located 0.51 km

east of the base station and Pleasont station is located 2.0 km northwest of the base station. Both sites have relatively uncluttered horizons (Figure 1). While the base station was operating, we collected at each location 300 sequential fixes continuously logged at a rate of approximately one per second into the waypoint buffer of the rover unit. Satellite data reception for the rover was facilitated by an external antenna mounted atop a 2-m staff on a tripod positioned directly above the control point monument. Data were collected for both locations within a 30-minute period for each sample. Data logging at the base station was stopped after the field data had been collected by the rover unit.

The fixes for *Pleasant* and *University* stations were downloaded into the base station microcomputer, differentially corrected and, along with the uncorrected (raw) coordinates, stored as ASCII files. The PROJECT utility from the GIS software package ARC/INFO (Environmental Systems Research Institute, Redlands, California) was used to convert the data from decimal degrees (provided by the PFBASIC program) to Universal Transverse Mercator (UTM) metre units (Zone 19). The final ASCII file contained the date of collection, a code for station name (University or Pleasant), the replicate fix number (1 to 300), the raw coordinates for the fix expressed in UTM metres, and the differentially corrected coordinates in UTM metres for the same fix. A session's data were disregarded if any of the 300 points were unable to be differentially corrected. All data were collected over 13 different days from 20 March to 23 April 1992. The final dataset consists of 20 sample sessions each with 300 replicate fixes obtained from the two first-order control stations.

Data Analysis Protocol

The fundamental unit of our analyses is the distance between the known location for Pleasant or University station and the computed location determined from the cps rover unit. Exact locations for the monumented control points were originally reported in the NAD-as reference system. The GPS locations we measured were based on the WGS-aa reference system. The differences between NAD-83 and WGS-84 reference systems are slight (Schwarz, 1989; Brown, 1992) and do not require standardization for analyses such as we present.

For this study, accuracy is represented by the distance between the known and GPS-derived positions for each station. As the calculated location approaches the known position, accuracy is increased. Precision is represented by the variation among repeated measurements of accuracy and is reported as the standard deviation of the mean distance from true for replicated samples. We also report our results using percentile summary statistics. The 50th percentile, or median value, is the distance from true that encompasses 50 percent of all the fixes in a sample. This is also known as the Circular Error Probability (CEP). We also provide distance measurements for the 90th, 95th, and 99th percentiles.

An important component of the analysis was to determine how accuracy was improved by averaging sequential fixes during a continuous data recording session. To estimate the effects of averaging, we calculated the mean UTM X value and mean UTM Yvalue for 5, 10, 20, 50,100, 200, and 300 successive fixes. For each average UTM X and average UTM Y , we calculated the distance from true. Equalitv of variance was tested with the Variance Ratio Test (Zar, 1974). Tests of the equality of means were performed using the Student's T statistic or analysis of variance (ANOVA) (Zar, 1974).

Systematic directional bias was estimated in two ways. For a sample of fixes, we calculated the mean deviation in the X (east-west) axis and the mean deviation in the Y(north-

PEER - REVIEWED ARTICLE

TABLE 1. RAW (NOT DIFFERENTIALLY CORRECTED) GPS DATA COLLECTED FROM UNIVERSITY AND PLEASANT GEODETIC CONTROL SITES. THE NUMBER OF SEQUENTIAL FIXES THAT WERE AVERAGED ARE VARIED FROM NO AVERAGING AT ALL (ONE REPLICATE PER FIX) TO 300 REPLICATES PER FIX, THE MEAN DISTANCE IS THE DIFFERENCE (IN METRES) BETWEEN THE GPS-DERIVED LOCATION AND THE TRUE POSITION.

TABLE 2. DIFFERENTIALLY CORRECTED GPS DATA FROM UNIVERSITY AND PLEASANT GEODETIC CONTROL SITES. THE NUMBER OF SEQUENTIAL FIXES THAT WERE AVERAGED ARE VARIED FROM NO AVERAGING AT ALL (ONE REPLICATE PER FIX) TO 300 REPLICATES PER FIX. THE MEAN DISTANCE IS THE DIFFERENCE (IN METRES) BETWEEN THE GPS-DERIVED LOCATION AND THE TRUE POSITION.

south) axis. We tested the hypothesis that the mean deviation in each axis was equal to 0 using a T statistic. If the null hypothesis was rejected, the mean was not equal to 0 and directional bias existed; eastward if mean X was positive, northward if mean Y was positive. We also assessed directional bias by using standard methods of analysis of circular data (Batschelet, 1981). For each fix, we calculated the angle of deviation from true. For a sample of angles, we used the Rayleigh statistic to test the null hypothesis that the angles were uniformly distributed in all directions and were not clustered in a single direction. If the null hypothesis was rejected, the sample of angular deviations was not uniformly distributed and a directional bias was present.

Results and Discussion

The accuracy and precision of raw data not subjected to differential correction for Pleasant and University stations are reported in Table 1. The average distance from true for a single locational fix was 32 m for University station and 39 m for Pleasant station with 95 percent of all fixes falling within 67 m (University station) and 93 m (Pleasant station). These levels of accuracy are within the limits of error imposed by SA which was enabled during our study. The Department of Defense states that 95 percent of GPS fixes will be within 100 m when SA is operational (Georgiadou and Doucet, 1990). When 300 sequential fixes are averaged, the mean distance

PEER - REVIEWED ARTICLE

Figure 2. Representative array of 300 sequential fixes obtained at the P/easant station on 1 April 1992 (A) without differential correction, and (B) after differential correction. The three circles have radii of 2 m, 5 m, and 10 m. The station is located in the center of the inner circle. The arching pattern in the raw data is due primarily to the distortion caused by Selective Availability.

quential fixes.

TABLE 3. PEARSON PRODUCT-MOMENT CORRELATIONS (r) FOR RAW AND DIFFERENTIALLY CORRECTED GPS DATA. THE 300 REPUCATE AVERAGE PER FIX DATA WERE USED IN THIS ANALYSIS. NONE OF THE R VALUES SIGNIFICANTLY DIFFERED FROM $0(n=20)$.

from true is reduced to 25 m (University station) and 27 m (Pleasant station) with 95 percent of all data falling within 44 m (University station) and 73 m (Pleasant station).

Differential correction markedly improved the accuracy and precision of cps data (Table 2, Figure 2). The average distance from true for single fixes was 5 m (Pleasant station) to 7 m (University station) and 95 percent of all single fixes were within 10 and 15 m (Pleasant and University stations, respectively). When the 300 replicates per data recording session were averaged, the mean distance dropped to less

44

than 3 m for both stations; 95 percent of the fixes fell within 4 m for Pleasant station and 6 m for University station.

Averaging of sequential fixes improved the accuracy of position estimation (Table 2). Averages of 50 to 300 sequential fixes produced notable improvement in accuracy (Figure 3); however, averaging fewer than 20 replicate fixes only slightly improved the accuracy. The precision of corrected fixes is significantly improved by averaging, The variance of the mean of the 300 replicates per fix data was significantly less than the variance of the single replicate per fix data (Variance Ratio Test; $F_{Pleast} = 16$, $p < 0.001$, $F_{University}$ 5.9, $p < 0.001$).

Accuracies obtained for the two stations were similar, The mean deviation for the 300-fix distance (with differential correction) did not differ significantly between the two stations (T = 1.45, $p > 0.05$), although the data obtained for University station were significantly more variable than data obtained at Pleasant station (F = 4.1, p < 0.005). Interestingly, raw data (not differentially corrected) obtained for Pleasant station were significantly more variable than raw data from the University station ($\tilde{F} = 3.2, p < 0.01$). It is unclear why differential correction improved the precision of Pleasant station data more than the University station data. There were no significant correlations among uncorrected and corrected data (300 sequential fix average dataset) within or between stations (Table 3). These results indicate that it is difficult to predict the accuracy of differentially corrected fixes based on the accuracy of the same data prior to the correction process. Accuracy at one location does not appear to be a good predictor of accuracy obtained at another site.

Consistent directional biases were slight, or non-existent, at the University and Pleasant stations. The mean deviations in the east-west axis (300 replicate average data with differential correction) were -0.5 m for *Pleasant* and -1.13 m for University. The mean deviation for Pleasant station did not differ from 0 (T = 0.17, $p > 0.05$) but the deviation at University station in the east-west axis was significantly less than 0 (T = 1.13, $p < 0.05$) and this indicates a trend toward a western bias. The mean deviation in the north-south axis was 0.18 m at Pleasant and -0.5 m at University. Neither deviations in this axis differed significantly from 0 ($T_{Plexant}$ $= 0.6, p > 0.5;$ T_{University} $= 0.3, p > 0.5$. The mean angles of deviation for both stations are given in Table 4. In no case did the distribution of angles differ from uniform,

There were significant differences in the accuracy of GPS fixes among days. Two-way ANovA indicated a significant sampling day effect at both *Pleasant* ($F = 2.0$, $p < 0.05$) and University (F = 3.1, p < 0.001) stations when the 50 replicate average per fix data were used. Average sample (i.e., the mean of the six 50 replicate averages per sample day) accuracies ranged from 1.4 to 5.4 m at Pleasant station and 1.7 to 7.4 m at University station. The 50 replicate average data were chosen for this analysis because of the need to have replication within sampling days ($n = 6$ /sample day) and this level of averaging produced a reasonably good estimate of true location (Figure 2).

Conclusions

Our results indicate that a relatively inexpensive three-channel GPS receiver, coupled with the capability to apply differential correction to field data, can produce accurate and precise locational data at medium scales (1:12,000 to 7124,000) of resolution. Under good conditions (low PDOP and few obstructions to satellite signals), differentially corrected GPS positions will be within 6 m of true location. Averaging replicate fixes significantly improves the accuracy. We found that using fewer than 20 replicates per fix only slightly improved accuracy, whereas using 50 replicates or

TABLE 4. DIRECTIONAL BIAS IN RAW (UNCORRECTED) AND DIFFERENTIALLY CORRECTED GPS DATA FOR THE PLEASANT AND UNIVERSITY STATIONS. THE AVERAGE OF ALL 300 SEQUENTIAL REPLICATES PER SAMPLE WAS USED PER FIX. THE MEAN ANGLE IS THE AVERAGE DIRECTION OF THE CALCULATED LOCATION RELATIVE TO THE ACTUAL POSITION, r IS THE LENGTH OF THE MEAN VECTOR AND IS A MEASURE OF THE VARIATION AROUND THE MEAN ANGLE. R IS RAYLEIGH'S R STATISTIC WHICH TESTS THE NULL HYPOTHESIS THAT THE SAMPLE OF ANGLES IS UNIFORMALLY DISTRIBUTED AROUND A CIRCLE. P IS THE PROBABILITY THAT THE DATA ARE UNIFORMALLY DISTRIBUTED AROUND A CIRCLE.

more markedly improved both accuracy and precision. The cost of replication is slight as most GPS receiving units have data logging capabilities. Replication simply means spending a few more minutes at the data collection site to acquire additional fixes. Differential correction significantly improves the accuracy of GPS data. There can be significant day to day variation in accuracy at any given location. Accuracies obtained at one field site are not necessarily correlated with those obtained at other locations within the same time period. The results presented here should be viewed as near optimum for the kind of equipment we used. Data obtained under less than favorable conditions (e.g., poor satellite geometry or obstructions between the receiver and the satellite constellation) will likely be less accurate than those presented here.

Acknowledgments

We are grateful to Earl Patric and the College of Resource Development at the University of Rhode Island for making funds available to purchase the GPS equipment used here. Don White and Spencer Drake provided generous amounts of time and expertise in the GPS survey of the base station antenna. John Stachelhaus and the Rhode Island Department of Transportation Survey Section conducted the second survey of the base station antenna. Larry Fisher, Thad Mauney, Mike Charpentier, and Art Lange provided invaluable technical advice during these experiments. The EPA Region I office provided training opportunities and workshops on using GPS and these were very helpful in designing this study. Courtney Conway, Dan Civco, Frank Golet, Carol Baker, and John Stachelhaus critically evaluated the paper and their comments are most appreciated. This is Contribution Number 2781 of the Rhode Island Agricultural Experiment Station. Funding was provided, in part, by the Rhode Island Cooperative Extension Service.

References

- Adkins, K. F., and J. G. Lyon, 1991. Use of Aerial Photographs to Identify Suitable GPS Survey Stations, Photogrammetric Engineering & Remote Sensing, Vol. 57, No. 7, pp. 933-936.
- Antenucci, J.C., K. Brown, P.L. Croswell, M. J. Kevany, and H. Archer, 1991. Geographic Information Systems: A Guide to the Technology, Van Nostrand Reinhold, New York, 301 p.
- August, P. V., in press. Applications of GIS in Mammalogy: Building a Database, Applications of GIS in Mammalogy (S. McLaren and J. Braun, editors), University of Oklahoma Press.
- August, P.V., C. Baker, C. LaBash, and C. Smith, in press. Geographic Information Systems for the Storage and Analysis of Biodiversity Data, Methods in the Assessment of Mammalian Biodiversity (D. Wilson and R. Rudran, editors), Smithsonian Institution Press, Washington D.C.
- Batschelet, E., 1981. Circular Statistics in Biology, Academic Press, New York, 371 p.
- Brown, A., 1992. The GPS Coordinate System Explained, GIS World, Vol. 5, No. 2, pp. 70-71.
- Georgiadou, Y., and K. D. Doucet, 1990. The Issue of Selective Availability, GPS World, Vol. 1, No. 5, pp. 53-56.
- Hurn, J., 1989. GPS: A Guide to the Next Utility, Trimble Navigation Ltd., Sunnyvale, California, 76 p.
- Leick, A., 1987. GIS Point Referencing by Satellite and Gravity, International Geographic Information System (IGIS) Symposium: The Research Agenda (R.T. Aangeenbrug and Y.M. Schiffman, editors), Association of American Geographers and the National Aeronautics and Space Administration, Vol. 2, pp. 305-317.
- Long, D. S., S. D. DeGloria, and J. M. Galbraith, 1991. Use of Global Positioning System in Soil Survey, Journal of Soil and Water Conservation, Vol. 46, No. 4, pp. 292-297.
- Puterski, R., J. A. Carter, M. J. Hewitt III, H. F. Stone, L. T. Fisher, and E. T. Slonecker, 1990. Global Positioning Systems Technology and Its Application in Environmental Programs, GIS Technical Memorandum 3, Environmental Monitoring Systems Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Las Vegas, Nevada.
- Schwartz, C. R., 1989. Relation of NAD 83 to WGS 84, North American Datum of 1983 (C. R. Schwartz, editor), NOAA Professional Paper NOS 2, pp. 249-252.
- Slonecker, E. T., and J. A. Carter, 1990. GIS Applications of Global Positioning System Technology, GPS World, Vol. 1, No. 3, pp. $50 - 55$
- Wells, D., N. Beck, D. Delikaraoglou, A. Kleusberg, E. Krakiwsky, G. Lachapalle, R. Langley, M. Nakiboglu, K. Schwarz, J. Tranquilla, and P. Vanicek, 1986. Guide to GPS Positioning, Canadian GPS Associates, Fredericton, New Brunswick, Canada.
- Wilkie, D. S., 1989. Performance of a Backpack GPS in a Tropical Rain Forest, Photogrammetric Engineering & Remote Sensing, Vol. 55, No. 12, pp. 1747-1749.
- Zar, J. H., 1974. Biostatistical Analysis, Prentice Hall, New Jersey, 620 p.

(Received 15 September 1992; accepted 30 December 1992; revised 14 January 1993)

Do you know someone who should be a member?

Pass this Journal and Pass the Word!