Assessment of Practical 3-D Geodetic Accuracy for Static GPS Surveying

Pasi HÄKLI, Hannu KOIVULA and Jyrki PUUPPONEN, Finland

Key words: Static GPS, GPS accuracy, observing session, baseline length

SUMMARY

Static GPS measurements have been used for precise surveying over two decades. Nevertheless, consistent information about relation between accuracy, baseline length and observing time has been missing. Surveyors have been dependent on the information from various sources (vendors of the GPS equipments, ambiguous guidelines by different companies and institutions, etc) as well as on their own experiences.

The outcome of this study gives a relation between the accuracy, baseline length and observing time for static GPS as an easily readable graph. The chart covers all the conventional baseline lengths and observation times as well as broadcast and precise ephemeredes.

This study is a part of an ongoing project that studies the quality of geodetic GPS at the Finnish Geodetic Institute. We used data that covers distances between 0.6 and 1,069 km and observing sessions between 10 min and 24 hours. Over 10,000 baselines were processed with broadcast and precise ephemeredes. The set of data used in the study is a random sample chosen from the data from several GPS campaigns. This way it was to give a realistic picture of accuracy by averaging e.g. the influence of atmosphere and satellite geometry. The accuracy is presented for individual baselines i.e. adjustments were not applied.

A surface was fitted over the rms values. Since the data was rather heterogeneous a series of fitting schemes were tested and the one with the best fit was chosen. The goodness of fit (R^2) for the best fit was 0.91 for broadcast and 0.87 for precise orbits. As a result we generated a graph that shows 1-5 cm regression lines of accuracy as a function of baseline length and observing time.

Assessment of Practical 3-D Geodetic Accuracy for Static GPS Surveying

Pasi HÄKLI, Hannu KOIVULA and Jyrki PUUPPONEN, Finland

1 INTRODUCTION

GPS has been used for surveying since mid 1980's. Surveyors have utilized different modes of the GPS as soon as they have become available. In the beginning of GPS era static measurements rapidly took over traditional surveying methods. Nevertheless exact information about accuracy, reliability and required observing times were missing. Surveyors were dependent on the information from various ambiguous sources (guidelines by GPS vendors, different institutions, etc) as well as on their own experiences. Even today surveying society may feel a lack of consistent and research-based knowledge of an optimum way of using GPS for practical surveying.

There have been several scientific studies on GPS accuracy. Beutler et al. 1989 gave an empirical formula for connecting baseline length with accuracy. Their intension was to give a rule for the best possible GPS accuracy of the time. Similar studies were conducted e.g. by Dong and Bock (1989), Davis et al. (1989) and Larson and Agnew (1991). They all gave GPS accuracy as a function of baseline length. Eckl et al. (2001) added observing time as one factor of GPS accuracy. They used GPS data with 4-24 hour sessions and baseline lengths between 26 and 300 km. Soler et al. (2006) expanded observing times to cover also shorter sessions (1-4 hours). Dogan (2007) studied also the accuracy of GPS with observing times between 4-24 hours. He followed the methodology of Eckl et al. (2001). All these studies are valuable and give information of the reachable accuracy when scientific GPS softwares are used. The aim of these studies has mainly been to demonstrate the highest reachable accuracy for special scientific purposes like deformation studies.

Our primary goal was to create a general guideline that surveyors might find useful for practical purposes. We are using commercial GPS software with default parameters, covering observing sessions between 10 minutes and 24 hours and baselines between 0.6 km and 1,069 km. Both broadcast and precise orbits are used with the same GPS data. In this paper we show a relation between accuracy, observing time and baseline length as a one easily readable graph.

This study is a part of an ongoing project that studies the quality of geodetic GPS. The goal of the project is to investigate both real-time and post-processing geodetic GPS in practice. Real-time applications, namely real-time kinematic (RTK) and Virtual Reference Station (VRS™) concept, were studied in 2003-2005 (Häkli, 2004, 2006 and 2007; Häkli and Koivula, 2006).

2 TEST METHODS

The goal of this study was to find a relation between observing time, baseline length and accuracy in static GPS surveying. Naturally attainable accuracy is not only a function of baseline length and observation time. It may also be influenced by sampling rate, satellite orbits, satellite geometry, atmosphere, carefulness of the observer, site-dependent effects like obstacles and multipath conditions (see e.g. Granström 2006), the equipments and post-processing software. In this study we tried to take into account as many of these factors as possible.

2.1 Test Field and Reference Coordinates

For the static test field permanent or semi-permanent sites were chosen (Figures 1-3). The test field consists of the Finnish permanent GPS network FinnRef®, Suurupi permanent GPS station in Estonia and concrete pillar stations from regular GPS research campaigns. These campaign points are concrete pillars with an antenna platform that has a 5/8-inch hole in the middle. The GPS antenna is mounted on top of this platform. Use of this kind of pillars and permanent GPS stations ensures that our results are free from centering and height reading errors. Also all the sites have a good visibility so that the observations were made in a good observing conditions.

Total of 31 stations or pillars were used in the test and a set of baselines between them was chosen. ITRF2000 coordinates of the points were determined with Bernese software using at least 24 hours of data. ITRF coordinates were derived from Metsähovi (METS) IGS station. The epoch of the coordinates is 2004.5 that is the mean epoch of the data used in the test. Standard errors are in the order of few millimeters. These coordinates were used as reference coordinates in the test.

2.2 Data and Baseline Processing

Static data of the test was to cover wide range of baseline lengths and observing times. Baseline lengths cover the distances between 0.6 and 1,069 km (figure 4). Observing sessions vary between 10 min and 24 hours (figure 5). These scales were considered to fulfill the requirements of conventional static GPS surveying. All the observations were made with choke ring antennas that were mounted either on steel masts or concrete pillars. The data is a random sample of chosen baselines from several GPS campaigns in 2003-2005. This way it was to give a realistic picture of accuracy by averaging e.g. the influence of atmosphere and satellite geometry.

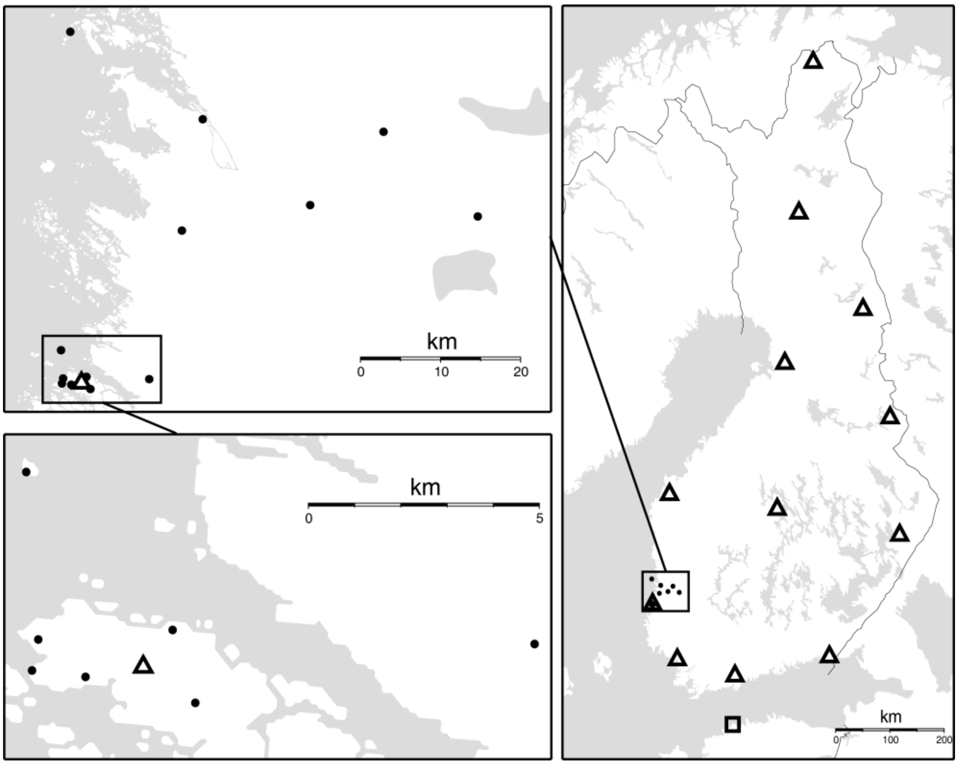

Figure 1. Test points used in the study. Triangles are the FinnRef® permanent GPS stations. Quadrangle is Suurupi permanent GPS station in Estonia. Circles are concrete pillars for research purposes.

Figure 2. A pillar of Romuvaara permanent GPS station. Three FinnRef® stations have this kind of concrete pillar as antenna platform. Same kind of pillars is used on the geodynamical test fields used in this study as well.

Figure 3. A mast of Tuorla permanent GPS station. This kind of a 2.5 m height steel grid mast is a typical antenna platform at FinnRef GPS sites.

The data was processed with commercial GPS software (Trimble Total Control) using standard processing parameters. These were mainly default options of the software and partly chosen to reflect the parameters that most of the surveyors use. However, instead of the default antenna models, NGS antenna models were used for the processing. Results are presented for individual baselines i.e. adjustments were not applied. In general, accuracy is expected to improve if proper network adjustment is utilized. A total of over 10,000 baselines were processed for the study, half of them with broadcast and half with precise ephemeredes. Figure 6 shows one example of the raw 3-D accuracy results when broadcast ephemeredes were used. It is clear how the results deviate more over longer baselines as expected.

All the processed data had an observing interval of 30 s since that is a nominal sampling rate of FinnRef® permanent GPS stations. In order to clarify if observing interval would have a significant role in this study, a separate test campaign was performed to study the influence of observing interval. A test was performed on another test field using tripods. Lengths of the 18 measured baselines varied between 1.8 and 32.5 km. Original observations were collected with 5 s sampling rate, but the same data was processed with 10 s 15 s and 30 s intervals as well. We did not find any significant improvement by using higher sampling rate (see figure 7). In all the cases we saw no or only a slight change in expectable accuracy when higher sampling interval was used. Mostly the slight improvement was seen on longer baselines with short observing sessions.

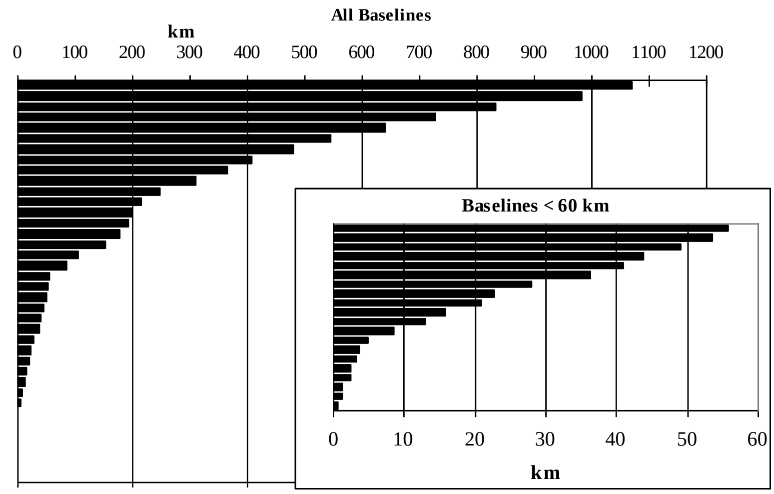

Figure 4. The baseline lengths cover the distances between 0.6 and 1,069 km. The graph behind show all the baseline lengths and the graph in front is zoomed to better show the lengths under 60 km.

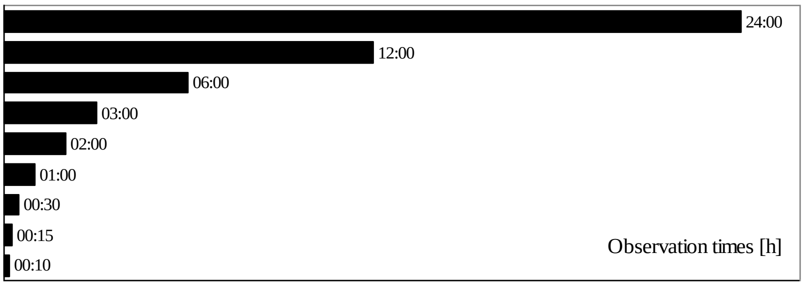

Figure 5. Observing sessions varied between 10 minutes and 24 hours.

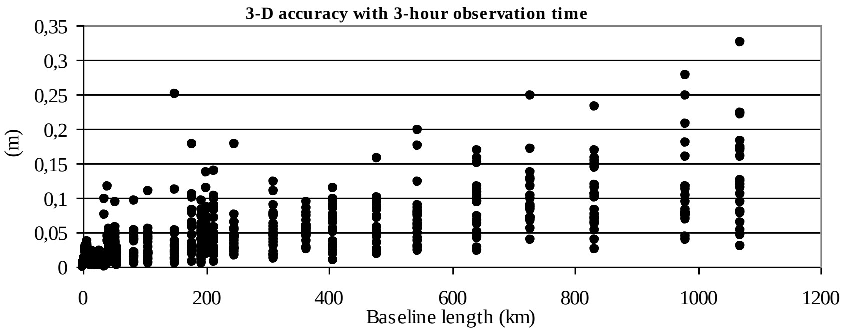

Figure 6. 3-D accuracy of different baseline lengths using broadcast ephemerides. Observing time is 3 hours.

2.3 Pre-processing of the Baseline Results

GPS results for the computed points are given in cartesian X, Y, Z coordinates. By comparing these to reference coordinates we get an error vector that consists of components ΔX, ΔY and ΔZ. However, we use the length of error vector (ΔS) as accuracy measure that can be considered one-dimensional. Hence accuracy is only one component in our 3-D case, baseline lengths and observing times being the others (Fig. 8). These form the components of one grid point where accuracy is the variable (observation). Each grid point consists of approximately 20 baseline results and the data sets used in surface fitting include 267 grid points for broadcast orbits and 275 grid points for precise orbits (5,062 and 5,178 baselines respectively).

In order to find the connection between baseline length, observing time and accuracy, a 3-D surface fit to the grid points is needed. Since the scales are wide both in baseline and session lengths, results are quite heterogeneous. Therefore pre-processing of the results is needed. Results were pre-processed first on individual observation level (1-D) and then group-wise (2-D).

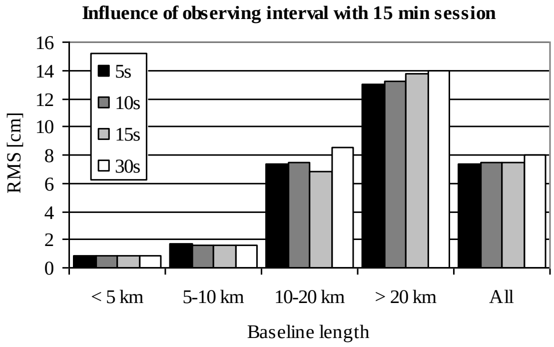

Figure 7. The influence of different observing intervals on static GPS results.

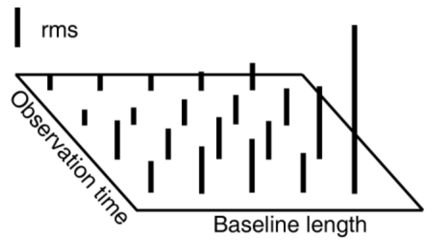

Figure 8. Data pre-processing at grid point level. Grid point components are baseline length and observation time and accuracy.

2.3.1 Individual Observations

The results were checked at individual observation level in order to get consistent (internally accurate) data sets. This is essential because each grid point consists of only 20 observations and in the presence of gross errors the accuracy is biased. Because of the same reason a rms (root mean square) was chosen as accuracy measure. For example 95^{th} percentile would have been more visual measure but also more unreliable because of the number of observations. An example of data set was shown in figure 6.

In the beginning a simple pre-elimination of gross errors was performed. Errors greater than 50 cm for individual baseline were considered to be gross errors and they were eliminated from the results. The limit is more or less arbitrarily chosen but one cannot consider errors larger than this as geodetic accuracy under any circumstances. However, an additional condition to the limit was set. If the error was smaller than 1 cm + 1 ppm, the result was not discarded. This avoids individual results being rejected offhand even if commonly used condition (rather conservative though) of GPS distance-dependency is fulfilled.

According to the normal distribution, all data points outside the 99.9% (approximately 3σ) deviate significantly from the rest of the data and therefore they can be considered as gross errors. This method was used after pre-elimination with a few iteration rounds. Approximately 1.5% of individual observations were eliminated during this stage. From the rest of the observations rms was calculated for each grid point.

2.3.2 Grid Points

At the second step of pre-processing data was checked on grid point level in order to get consistent data for the surface fit. This is arguable because of heterogeneous data, e.g. data includes large errors (rms) at long baselines and short observation times, which could not be considered as geodetic accuracy. If the data were not processed at all a high-order surface would be needed to model the data.

Grid points were checked series-wise so that observing times were kept constant and baseline length and accuracy were variables (Fig. 9). Each series were checked against the outliers and also possible truncation point was searched. The truncation point is expected in cases where observations are simply too few to accomplish the accurate baseline processing in GPS computation. After these points GPS solution gets radically weaker and accuracy is poor. This is the case especially for long baselines measured with short observing times.

For outlier detection we chose polynomial fit method. High-order polynomial describes the data well and only coarse errors fall outside data set. The selected method is very conservative and ensures minimum rejections. First the optimum degree of the polynomial was searched. The degree was chosen with standard errors of the fit. Then n-degree polynomial was fitted through the series and fit errors (e) were compared to three times standard deviation of the fit errors ($3\sigma_e$). If fit error exceeded the limit, the grid point was rejected. Only a few gross errors were found. In the case of 1-hour series (Fig. 9) the polynomial degree was n=11 and no outliers were found.

After the outlier detection a possible truncation point of the series was searched. The selected search method was differencing the consecutive grid points in the series. Squared differences show a possible step in the series as a peak. Peaks were also checked against 3σ rule described earlier. The method found some significant discontinuities in accuracy in the series of shorter observing sessions (which were also visually seen from the plots). The series were truncated from those points. The rms at these points was already outside the scale of the interest and therefore only bad data was rejected. For example the 10- and 15-minute series were truncated for both broadcast and precise orbit results at approx. 20 km baseline lengths and therefore following surface fit is not valid for longer baselines with these observing times. In 1-hour case (Fig. 9) results with broadcast orbits were truncated from 544 km.

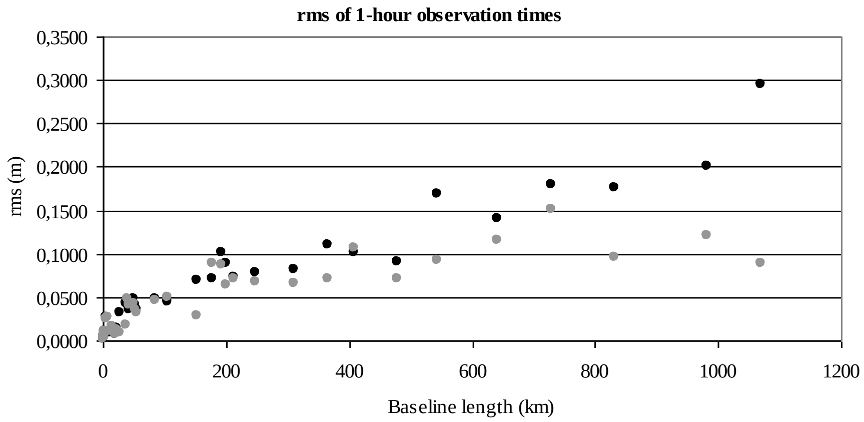

Figure 9. rms values for broadcast (black) and precise (grey) when 1-hour sessions were used.

2.4 Surface fi

2.4.1 Method

After the grid data was pre-processed we used a multiple regression to fit a surface to the grid points. Several different surfaces were tested. After some tests it became evident that high-order surface or data linearization is needed since the grid is not evenly spaced and data is varying. Since we wanted to keep the number of surface parameters low, data linearization and plane fit were chosen. Various linearization methods were tested and after comparisons we ended up using 10-base logarithm and square root for linearization. The final plane fit was performed with the equation:

Equation 1. Formula used in the surface fit.

$\log 10(z) = a + b \cdot \log 10(x) + c \cdot \sqrt{y}$, where

x	baseline length [km]
y	session duration [h]
z	rms of accuracy [m]
a, b, c	coefficients of surface fit

Plane was fitted iteratively with the least squares method. We computed the fit errors ($e_i = z_i - Z_i$) and if the grid point deviated more than 3□$_e$ it was discarded in iterative process. Iterations were continued until all outliers were rejected. The outliers lie mainly outside the area of interest i.e. rms of rejected data is high. The number grid points used in plane fit are 199 for broadcast orbit results and 239 for precise orbit results.

2.4.2 Goodness of fit

As a quality measure we use here R-squared (R^2) and rms of the fit. R-squared is the coefficient of the determination and tells how well the fitted surface can describe the data. rms instead shows the size of the remaining fit errors in the surface. However one must remember that rms gives the value that only roughly 68% of the measurements fall into (with an assumption of normal distributed data).

Rms of the final fit is 8.2 mm for broadcast orbit results and 10.9 mm for precise orbit results. Iterative process improved these values by a few millimeters, being in the beginning 13.6 mm and 13.0 mm, respectively. rms is fairly good if we take into account the heterogeneity of the original data. Also high R-squared value (0.907 for broadcast obits and 0.866 for precise orbits) proves that the model is suitable and describes the data successfully (Table 1).

The remaining fit errors at the grid points are drawn into the figure 10. We drew 1-5 cm regression lines to the figure. This range was thought to reflect geodetic accuracy however emphasizing to the lower end. The upper end may be considered in some less accuracy-demanding measurements but for accurate surveying one should consider to use lower end (< 2 cm). As the figure shows most of the largest fit errors lie outside the area of interest. The rms of fit errors for the accuracies < 5 cm is even better, 7.3 mm for broadcast orbit results and 8.7 mm for precise orbit results (Table 1).

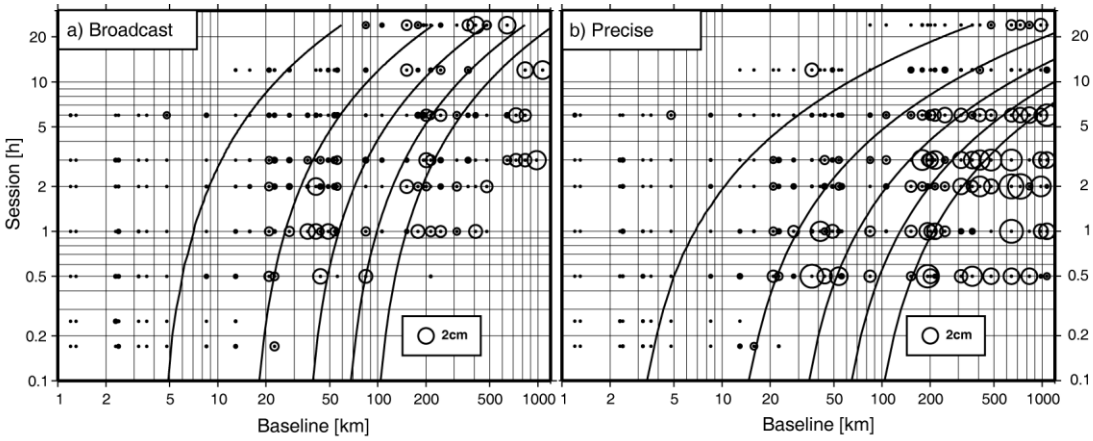

Figure 10. Grid points and remaining errors of the surface fit (circles) for broadcast orbits (a) and precise orbits (b). Figure shows also the regression lines for 1-5 cm accuracies.

Table 1. Some figures of the surface fit. R^2 and rms of fit prove that the surfaces represent data well. rms (area) is the rms of fit in the defined area (z < 5cm).

Value	Broadcast	Precise
R^2 [-]	0.907	0.866
rms [mm]	8.2	10.9
rms, area [mm]	7.3	8.7
Max error [mm]	23.6	32.4
Iterations	7	4
# of grid points in fit	199	239

3 RESULTS

3.1 Accuracy vs. Baseline Length vs. Observing Time

The graph in Figure 11 shows the connection between baseline length, observing time and rms accuracy. With the graph one can estimate one of above-mentioned variables if two of them are known. Both baseline length and observing time axes are in logarithmic scale. Regression lines show the rms accuracy of 3-D geodetic GPS for broadcast orbits (solid line) and precise orbits (dashed line). General trend is that longer the baselines are more one gain from using precise orbits. A grey shadowing indicates the area where the regression lines and the model are not valid. This is mainly due the fact that there were not enough data for reasonable GPS solution for that specific baseline length (e.g. truncated series) or that there were no grid points available.

It is relatively easy to use the graph. For example, if one knows the baseline length and desired accuracy, it is easy to read the recommended observing time. At the same time it may be decided if the precise orbits should be used or not. For example, one may estimate observing time for 20 km baseline and 1 cm rms to be a bit more than 4 hours for precise orbit solutions.

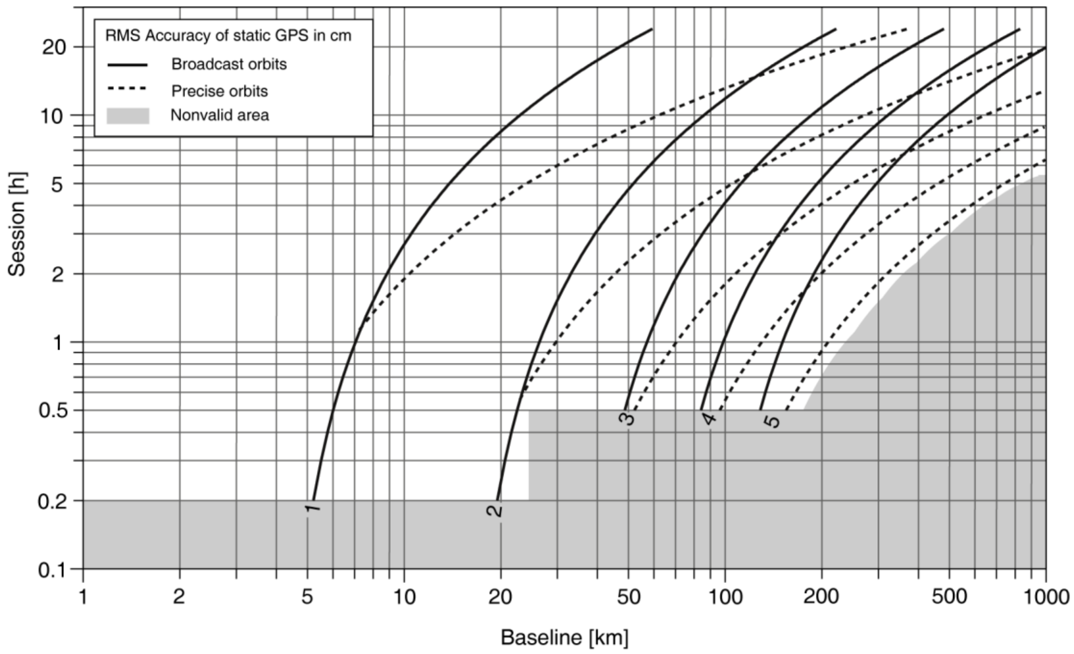

Figure 11. Static GPS results for individual baseline solutions with broadcast ephemerides. Dependency of baseline length, observation time and accuracy (rms). Regression lines are for accuracies of 1-5 cm with 1 cm intervals.

3.2 Remarks on Given Accuracies

The regression lines are generated from rms values that generally indicate that 68% of the results will be better than the accuracy given by the graph. If higher confidence level is required, one may get an approximate of 95% confidence level by multiplying rms with the factor 2. It is also worth noticing that the results show 3-D accuracy. In GPS the height component is typically 2-3 times worse than horizontal components. This means that if you are only interested in horizontal coordinates on a map projection, the graph gives pessimistic accuracy estimates. On the other hand you are on the safe side getting the desired accuracy with even higher probability.

While using the graph one must remember that all resulting values are only estimates (even if based on real measurements) and one’s own results may be strongly dependent on other factors. The sites used for the test have permanent installation with no possibility to centering or height measuring errors. The sites also locate in good observing environments. If your baselines are in an urban or forestry area or have lots of obstacles, the results may be worse than the graph indicates. There are differences in post-processing softwares as well. This graph was generated with using the default parameters recommended by the software we used. If you are experienced user and know what to do with all the parameters, you may reach higher accuracies that were shown here. Also, the graph gives the accuracies for independent baselines, so one may expect better results with proper network adjustment.

4 DISCUSSION

Our primary intention was to create a general guideline to help surveyors in planning and executing e.g. engineering surveys. Earlier studies have been mainly reaching the limits of scientific GPS processing. The main difference between our and earlier studies is that we computed our baselines with commercial software with default parameters. We also expanded observing sessions to very short time spans for static surveying and covered the conventional baseline lengths that surveyors need. Even though the differences in ideology, we made some comparisons to earlier studies in order to weight our results against them.

Beutler et al. (1989) developed a formula giving the optimal performance of GPS. They made an assumption that they have ideal processing software and all biases could be solved. They also used fiducial point concept for estimating orbits (since precise orbits were not available at that time). In our study 24-h results represent the best performance of GPS and thus these were compared to those of Beutler et al. (1989). Their results fall between our broadcast and precise orbit results. Our broadcast results are 1-6 mm less accurate and precise results 0-22 mm better than those of Beutler et al. for baselines up to 1,000 km. The difference depends on the baseline length being the largest at the longest baselines (1,000 km). Our results show the development occurred during the two decades. It is possible to achieve same accuracy (and even better) with commercial software than the “best possible” at that time because of e.g. drastic improvement of precise orbits, satellite constellation and more sophisticated GPS technology.

A more recent study by Eckl et al. (2001) studied GPS accuracy for baselines 26-300 km and estimated also influence of observing time span for sessions 4…24 hours. They used data from National CORS network and processed data with PAGES (scientific GPS-processing software) with precise IGS (final) orbits. They gave a simple formula to quantify the accuracy of GPS and did not find significant distance-dependency. Soler et al. (2006) tested the formula with observing sessions 1-4 hours. They concluded that equation cannot be extrapolated to sessions less than 3 hours. We found a slight distance-dependency especially on short observing sessions and if we compare our results to the time span of their equation (3…24 h), the accuracies converge at 80-100 km for 3-12 h and at 200-300 km for 24-hour sessions. However, the results are not fully comparable since Eckl et al. used precision (repeatability) as accuracy measure.

The study and comparison showed that it is difficult or even impossible to tell one truth of the GPS accuracy. Earlier accuracy studies have mainly been performed in connection to some scientific questions like crustal deformation analysis. These studies were limited to the most accurate scientific softwares and precise orbits only. However, e.g. continuous reference station networks offer a new kind of access to the reference frame for surveyors using rather commercial than scientific GPS softwares. Our study intends to fill that gap by offering a simple graph where one may find the connection between GPS accuracy, observing time and baseline length for both precise and broadcast orbits.

REFERENCES

Beutler, G., I. Bauersima, S. Botton, W. Gurtner, M. Rothacher and T. Schildknecht (1989): Accuracy and Biases in the Geodetic Application of the Global Positioning System, Manuscripta Geodaetica, 14, 28-35, 1989.

Davis, J.L., W.H. Prescott, J.L. Svarc, and K.J. Wendt (1989): Assessment of Global Positioning System measurements for studies of crustal deformation, Journal of Geophysical Research, 94, B10, 13,635–13,650, 1989.

Dogan, U. (2007): Accuracy analysis of relative positions of permanent GPS stations in the Marmara region, Turkey. Survey Review, 39, 304, pp. 156-165.

Dong, D. and Y. Bock (1989): Global Positioning System Network Analysis with Phase Ambiguity Resolution Applied to Crustal Deformation Studies in California. Journal of Geophysical Research, 94, B4, 3,949-3,966, 1989.

Eckl M.C., R.A. Snay, T. Soler, M.W. Cline and G.L. Mader (2001): Accuracy of GPS-derived relative positions as a function of interstation distance and observing-session duration. Journal of Geodesy 75: 633-640.

Granström, C. (2006): Site-Dependent Effects in High-Accuracy Applications of GNSS. Licentiate thesis. Chalmers University of Technology, Department of Radio and Space Science.

Häkli, P. (2004): Practical test on accuracy and usability of Virtual Reference Station method in Finland. FIG Working Week 2004, Conference proceedings. Athens, Greece, May 22-27, 2004. 16pp.

Häkli, P. (2006): Quality of Virtual Data Generated from the GNSS Reference Station Network. Conference proceedings. Shaping the Change, XXIII FIG Congress, Munich, Germany, October 8-13, 2006. 14pp.

Häkli, P. (2007): A case study on Virtual Reference Station Concept in Finland. Survey Review 39, 304 pp.145-155

Häkli, P., and H. Koivula (2006): Quality of Geodetic GPS. Conference Proceedings. NKG General Assembly, Copenhagen, Denmark, May 28 – June 2, 2006. 8pp.

Larson, K.M. and D.C. Agnew (1991): Application of the Global Positioning System to Crustal Deformation Measurement 1. Precision and Accuracy. Journal of Geophysical Research, 96, B10, 16,547-16,565, 1991.

Soler, T., P. Michalak, N.D. Weston, R.A. Snay and R.H. Foote (2006): Accuracy of OPUS solutions for 1- to 4-h observing sessions. GPS Solutions, 10(1):45-55. DOI 10.1007/s10291-005-0007-3.

CONTACTS

Pasi Häkli, Hannu Koivula, Jyrki Puupponen
Finnish Geodetic Institute
Department of Geodesy and Geodynamics
P.O.Box 15
FI-02431 Masala
FINLAND
Email: pasi.hakli@fgi.fi, hannu.koivula@fgi.fi, jyrki.puupponen@fgi.fi
Web site: http://www.fgi.fi/