Combined Analysis of Real-Time Kinematic GPS Equipment and Its Users for Height Determination

By W. E. Featherstone[1] and M. P. Stewart[2]

Abstract: Ellipsoidal and Australian Height Datum heights have been determined for a 60-point control network from five base stations by three separate users using three different makes of real-time kinematic (RTK) global positioning system (GPS) equipment. This was to determine whether these RTK GPS surveys (i.e. equipment and users combined) can meet vertical accuracy specifications for contract work with Main Roads Western Australia. A comparison of the differences between the RTK GPS ellipsoidal heights and the control indicates 51 mm (95% confidence) accuracy, accounting for errors in the control network. Likewise, the difference between the RTK GPS–derived Australian Height Datum heights and the control indicates 53 mm (95% confidence) accuracy. However, there was an increasing number of failed RTK GPS position solutions with increasing baseline length, which had to be omitted from these statistics.

INTRODUCTION

Real-time kinematic (RTK) global positioning system (GPS) surveying offers an efficient means of providing near-instantaneous positions. However, as with all other GPS measurements, it is inherently less precise and accurate for the determination of ellipsoidal heights, and there is the need to transform these heights to the local vertical datum. The short occupation periods and radiation technique usually used in RTK GPS surveying compound the former deficiency, yet RTK GPS has been widely adopted as an engineering surveying tool. Moreover, there has been little independent, objective testing of RTK GPS positioning precision, accuracy, and reliability (Edwards et al. 1999). Instead, RTK GPS users have tended to rely on the internal precision estimates provided by the software and manufacturers' brochures. Therefore, question arises as to the inherent precision, accuracy, and reliability of RTK GPS for ellipsoidal height determination, together with the precision and accuracy of heights when transformed to the local vertical datum.

During 1997, an RTK GPS test network was established by Main Roads Western Australia (MRWA), the Western Australian Department of Land Administration (DOLA) and Curtin University of Technology. The 60 ground marks comprising this network were subsequently coordinated by two existing MRWA survey contractors and one GPS vendor (collectively termed contractors throughout) using three different brands of RTK GPS equipment. To introduce an element of independence, no contractor was supplied with accurate positional information about the 60 points. For the sake of confidentiality, the contactors will be identified by A, which used Ashtech GPS receivers, T, which used Trimble GPS receivers, and L, which used Leica GPS

[1]Assoc. Prof., Dept. of Spatial Sci., Curtin Univ. of Technol., GPO Box U1987, Perth, WA 6845, Australia. E-mail: featherstone_we@cc.curtin.edu.au

[2]Assoc. Prof., Dept. of Spatial Sci., Curtin Univ. of Technol., GPO Box U1987, Perth, WA 6845, Australia. E-mail: stewart@vesta.curtin.edu.au

Note. Discussion open until October 1, 2001. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on December 29, 1999; revised July 18, 2000. This paper is part of the ***Journal of Surveying Engineering***, Vol. 127, No. 2, May, 2001. ©ASCE, ISSN 0733-9453/01/0002-0031–0051/ $8.00 + $.50 per page. Paper No. 22146.

receivers. These contractors were requested to conduct the RTK GPS surveys in a ''production mode'' so as to replicate the procedures that are used during contract surveys for MRWA. The contractors' surveys were conducted with respect to five base stations, situated within (Base 0) and approximately 1 km (Base 1), 2 km (Base 2), 5 km (Base 5), and 10 km (Base 10) from the test network. When considered together, this provides a sufficiently large sample (900 station occupations) to estimate accuracy confidence limits for the RTK GPS ellipsoidal heights and heights transformed to the Australian Height Datum (AHD).

The principal objective of these tests was to determine whether RTK GPS surveys can satisfy MRWA's vertical accuracy specifications for surface definition, which are 40 mm for sealed roads, 80 mm for formed surfaces, and 250 mm for natural surfaces (MRWA 1997). These specifications are assumed to be representative at the 95% confidence level and relate to the AHD since this is an engineering application. The need for the tests described in this paper has arisen because many of the contractors to MRWA have varying GPS-geodetic experience, use different RTK GPS hardware and software, and there is a dearth of any objective, independent assessment of RTK GPS systems. It is important to acknowledge that, as there was no control over the data collection and reduction strategies adopted by each contractor, these tests only allow for a *combined* estimation of the performance of the RTK GPS equipment and contractor. The contractors' RTK GPS results will be compared statistically with WGS84 ellipsoidal and AHD heights of the 60-point test network, which was achieved by treating the 900-point data set as a whole, as well as by dividing it among subsets based on each contractor and the base station used.

PROBLEMS WITH HEIGHT DETERMINATION USING RTK GPS

System-Related Errors

Incorrect Integer Ambiguity Resolution

The set of ambiguous integer numbers of cycles between the antennae and the GPS satellite constellation is estimated statistically by the software. All GPS data processing applies statistical confidence tests to determine whether the correct set of integer ambiguities has been estimated. However, because of the sometimes-invalid assumptions made about the stochastic nature of the ambiguity resolution process, the software's statistical tests can resolve ambiguities incorrectly. For instance, the RTK GPS software may be statistically satisfied with the ambiguity set that it has estimated, and inform the user so, but this may be incorrect (a suspected example of this will be shown later). Incorrect ambiguity resolution is more likely to occur in RTK GPS surveys because of the relatively short occupation times used, which prevents sufficient redundancy of GPS observations. The short occupation periods also make the integer ambiguity estimation process highly subject to localized and time-dependent biases, such as multipath (described later).

Optimistic Internal Precision Estimates

The error estimate of the positions given by GPS software in a single radiation represents only the internal precision of the computed solution and not the absolute precision or accuracy. This is demonstrated by the general need to scale GPS-derived variance-covariance matrices in GPS network adjustments. The internal RTK GPS positional error estimate is derived using least-squares theory, which must assume an *a priori* stochastic model. However, the stochastic models used in RTK GPS data processing generally rely on internally estimated noise values and thus cannot model systematic errors. The result is that the precision estimates made by RTK GPS software can be too optimistic by approximately one order of magnitude in relation to true absolute precision and accuracy.

Dilution of Precision and Satellite Availability

The geometry of GPS positioning is intrinsically weaker in elevation because the satellites used are situated only above the antennas. In terms of RTK GPS, the very short data spans used can compromise positional accuracy in areas of poor sky visibility. Most RTK GPS equipment requires a minimum of six satellites with a low position dilution of precision (PDOP), typically less than 6, for each initialization. Even with complete sky visibility, GPS offers only a global coverage of four or more satellites in view with a PDOP of less than or equal to 6 at 99.90% of the time (U.S. Departments of Defense and Transportation 1997). Therefore, even in ''perfect'' conditions, RTK GPS positioning cannot be expected to operate all the time, and in areas of restricted satellite availability, the periods when it should not be used become prolonged.

GPS Baseline Length

RTK GPS position solutions tend to degrade as the baseline length increases. This is predominantly due to atmospheric refraction effects, which become decorrelated and thus no longer cancel through differencing algorithms over longer distances. For these longer baselines, the RTK GPS software may find it more difficult to reliably resolve ambiguities, which can usually be remedied by increasing the observation time. Moreover, the increased noise decreases the reliability of the integer-fixed RTK GPS solution and thus increases the possibility of incorrect ambiguity resolution. It should also be mentioned that the radio link between the base and roving receivers affects RTK GPS. However, this should not affect the accuracy of the final coordinates, rather the ability of the roving receiver to compute a position at all.

Multipath and Electrical Interference

Multipath occurs when GPS signals are reflected from nearby objects before reaching the antenna. Electrical interference occurs when secondary sources or other transmitters and receivers distort the reception of the GPS signals or affect the receiver's circuitry. These are particularly problematic for RTK GPS because they act as a bias during the short occupations used, or can prevent satellite tracking completely. The only solutions at present are to select sites well away from potentially reflective areas, use ground planes or choke rings attached to the antennas, and apply digital filtering techniques. The selection of survey sites becomes more difficult in the case of electrical interference. For example, a 785 MHz television transmission source may affect a GPS receiver 15 km away (Finn 1997).

Antenna Phase Center Offsets

When surveying with GPS, the position is actually computed at the phase center of the antenna, then the software makes corrections to the survey mark. This phase center does not necessarily coincide with the geometrical center of the antenna, nor is it fixed in position with time. Therefore, it is preferable to use the same make and model of antenna from the same manufacturer so that these effects cancel to a large extent.

User-Related Errors

RTK GPS Surveying by Radiation

Positioning by a single radiation is an extremely undesirable approach to surveying, simply because any error can remain undetected for the radiated position and any points subsequently determined from this point. Terrestrial surveying techniques have always relied on redundancy, careful observation and data reduction; GPS surveying should be no exception. Therefore, despite the fact that RTK GPS surveying directly provides coordinates, the baselines should still be observed as a part of a geodetic-type network so as to aid in the identification of error. Accordingly, the accuracy of a RTK GPS position can only be estimated from a least-squares adjustment of a redundant set of observations to that point. One of the first published examples of the deficiency of RTK GPS radiations was given by Wylde and Featherstone (1995), where RTK GPS radiated positions were shown to differ from their independently determined positions by up to 3 m, which occurred even though the proprietary software indicated that the position was acceptable. However, it is acknowledged that RTK GPS hardware and software have advanced since 1994, when Wylde and Featherstone's tests were conducted.

Vertical Coordinate Transformations

The vertical coordinate transformation applied by propriety RTK GPS software in real time generally relies on precomputed geoid models. The geoid height used at the RTK GPS base station is not applicable to the roving stations, especially over longer baseline lengths, because the geoid-ellipsoid separation is a function of position. In addition, there is often an inconsistency between the local vertical datum and gravimetric geoid models (Featherstone 2000). Therefore, additional modeling of the difference between the geoid model and local vertical datum is sometimes required (described later). However, this usually has to be conducted postmission, thus defeating the objective of real-time height determination on the local vertical datum.

Bad Survey Practice

There are several errors, both system- and user-related, that downgrade the height computed using RTK GPS surveying. These errors also depend on the model of RTK GPS equipment used, but can be summarized as

- Incorrect antenna height defined for base and roving antennas
- Roving antenna not held steadily or vertically above the ground mark
- Incorrect coordinates defined for base-station antenna
- Insufficient time allowed for ambiguity initialization or reinitialization
- Disregard of system quality indicators, such as PDOP and the number of satellites available
- The use of single station occupations by only one radiation

The practice of positioning known points during a RTK GPS survey to provide a check must still be classed as bad survey practice. This is because it gives no error assessment for those positions observed in between checkpoints. The site-dependent nature of some RTK GPS errors precludes the assumption that if a RTK GPS position is accurate at a single point, it must be accurate for all points in the survey. Nevertheless, this approach is still preferable to single RTK GPS radiations, though not ideal.

ESTABLISHMENT OF RTK GPS TEST NETWORK

WGS84 Ellipsoidal Height Determination

The RTK GPS test network was established jointly by MRWA, the Western Australian DOLA, and Curtin University of Technology on and around the University's Bentley campus (Fig. 1). A combination of GPS- and terrestrial-geodetic methods was used to coordinate the 60 points with both WGS84 ellipsoidal and AHD height, respectively. A further five existing control points were included as the RTK GPS base stations. These were situated within (Base 0) and approximately 1 km (Base 1), 2 km (Base 2), 5 km (Base 5), and 10 km (Base 10) from the test network, so as to allow analyses with increasing baseline length. These base stations were deliberately chosen to run in an approximate east-west direction. This is because the geoid-WGS84-ellipsoid gradient is extremely steep in this direction (Featherstone 2000), thus providing a challenging vertical coordinate transformation for the contractors. If the geoid modeling techniques utilized apply here, it is reasonable to assume that they will apply in regions where the geoid gradient is not so steep.

The test network was installed in two phases. First, a geodetic network was observed by DOLA using a combination of terrestrial and GPS methods. The GPS measurements were taken using Ashtech Z12 dual-frequency GPS receivers and each baseline was observed in static mode for a minimum of 45 min. This network encompassed all existing geodetic control on and around the Bentley Campus, the five base stations, and 3 points from the 60-point test network. Secondly, MRWA observed the 60 ground marks within the framework of the DOLA network using fast-static GPS, optical leveling and traditional horizontal survey measurements. The rapid-static GPS surveys used five Ashtech Z12 dual-frequency GPS receivers. Two receivers were situated on nearby base stations selected from the DOLA control survey, and the other three were shifted, ''leap-frog'' fashion, along the series of 60 points. Each rapid-static GPS session was nominally of 15 min in duration, with a mini-

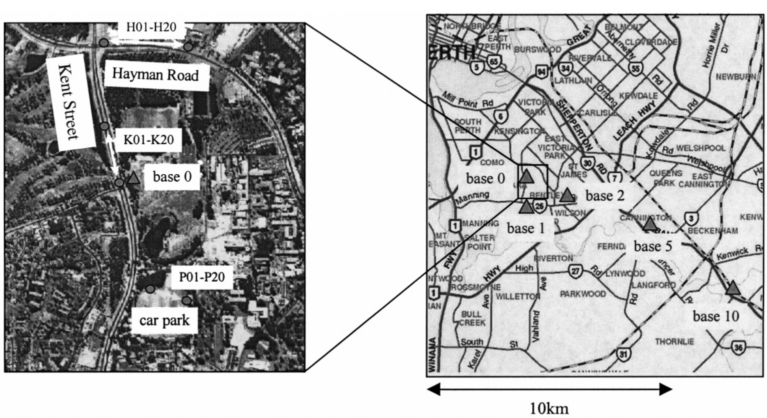

FIG. 1. Location of the Curtin RTK GPS Test Network Showing 20-Point Subsets along Hayman Road, Kent Street, and Car Park (Left) and Five Base Stations (Right)

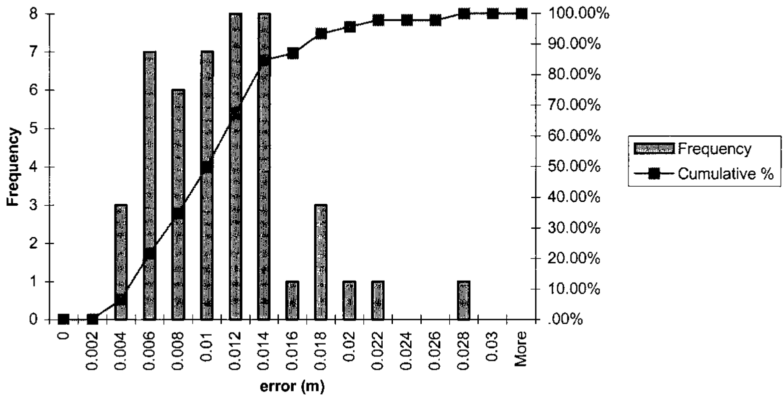

FIG. 2. Frequency Distribution and Accumulation of WGS84 Ellipsoidal Height Errors (95% Confidence) for 60-Point Curtin RTK GPS Test Network

mum of 30 min of data observed over longer baselines. The GPS results were also checked against independent solutions from terrestrial observations.

The test network was divided among three subsets comprising 20 points each. These were situationed linearly along Hayman Road (H0 to H20) and Kent Street (K01 to K20), and clustered in and around the south oval temporary car park on the Bentley Campus (P01 to P20) (see Fig. 1). This spatial distribution was designed to make the RTK GPS tests replicate, as closely as possible, typical MRWA contract surveys, with linear survey features to be positioned under varying degrees of GPS satellite visibility and site-dependent environmental factors (e.g. tree cover and multipath).

Three points, observed as part of the DOLA control survey, were held fixed during adjustment of the 20-point subsets. Therefore, for the subsequent analyses, the errors assigned to these points are simply taken from the adjustment of the DOLA control network. The 95% confidence error distribution for the WGS84 ellipsoidal heights in the 60-point test network is shown in Fig. 2. The largest error in ellipsoidal height is 27 mm, while the median is 11 mm. Note, however, that the distribution of errors in Fig. 2 is not normal. It can be seen that approximately 85% of control points have 95% confidence height errors of less than 15 mm. Therefore, this value has been adopted as a global estimate for the vertical accuracy of the control network. As the majority of points have errors of less than 15 mm, this estimate is probably too conservative and the median value of 11 mm may be more realistic. However, assessing the overall accuracy of a network such as this is difficult because mean and standard deviation computations cannot be applied to this nonnormal distribution. Therefore, it is preferable to use a slightly pessimistic estimate of the global vertical error, though further justification for this choice will be provided later.

AHD Height Determination

The AHD heights of the 60-point test network were established with respect to an existing AHD benchmark. The relative heights were established using class A leveling techniques with a digital barcode level and invar staves, and with a 4 mm per sqrt (km) allowable misclose (ICSM 1996). Accuracy estimates for geodetic leveling are difficult to determine from a prescribed survey tolerance. For instance, a field check of misclose will always be subject to the uncertainty of two equal and opposite errors being committed during the traverse. However, using the given allowable misclose over the mean distance of 50 m between test points implies an estimate of accuracy of 2 mm (95% confidence) in the AHD height at each point. It is acknowledged that the relative accuracy between adjacent leveled points is likely to be much less than this, because the class A leveling tolerance is not particularly suited to the very short distances between the points in this study.

An additional consideration is that RTK GPS does not measure the height difference between adjacent points in the test network, as is the case with the spirit leveling. Rather, RTK GPS measures the height difference between each point in the test network and the base station. Applying the class A tolerance over the maximum baseline length of 10 km, and acknowledging the aforementioned limitations, suggests an error of approximately 12 mm (95% confidence). Clearly, a compromise needs to be reached, while recalling that the objective of this project was to assess the performance of RTK GPS surveying as an alternative to spirit leveling for highway engineering surveys. Theoretically, the errors attributed to the control AHD heights should be applied over the GPS baseline lengths during tests of the RTK GPS performance. However, for the purpose of this project, this estimate of 2 mm (95% confidence) yields a realistic, if not completely rigorous, estimate of the error for AHD control heights in the test network.

ANALYSES OF RTK GPS–DERIVED WGS84 ELLIPSOIDAL HEIGHTS

Contractor-by-Contractor Analysis of Ellipsoidal Heights

Table 1 shows the mean, standard deviation, and number of differences observed between the WGS84 ellipsoidal heights from each RTK GPS test data set (i.e. for each contractor and for each base station) and the control WGS84 ellipsoidal heights. Importantly, the computed statistics do not take into account points where the contractor provided no position, which will be referred to as ‘‘null solutions’’ throughout. The statistics in Table 1 also assume that each RTK GPS position is independent of every other. This is not usually the case because points observed at adjacent time periods will be highly correlated due to the geometry of the satellite constellation and the RTK GPS processing procedure which must, in order to operate successfully, rely on data acquired previously. Neither are the position solutions among contractors necessarily independent because environmental considerations for GPS at each site, such as multipath, may similarly contaminate solutions taken at different times using different equipment.

In Table 1, differences can be seen among the means and standard deviations for each contractor and for different baseline lengths. Assuming that each data set is independent and a sample of a normally distributed population, some of these differences are statistically significant. When considering all points observed by contractors A and T (Table 1), both show significant mean offsets of 9 mm and −19 mm, respectively, from the control (Fig. 3). These offsets could be simply artifacts of the data sample or represent systematic RTK GPS system or user errors. While the results summarized in Table 1 can give an indication of the errors for each RTK GPS instrument and contractor with varying baseline length, it would be unwise to draw any firm conclusions from it. This is because the basic assumptions required for conventional statistical testing are unlikely to be valid for the smaller samples used.

Table 1 also indicates a slight overall increase in the scatter of the ellipsoidal height differences as the baseline length increases. For instance, the

TABLE 1. Mean, Standard Deviation, and Proportion of Observed Differences between WGS84 Ellipsoidal Heights for Each Base Station and Each Contractor

Base station	Contractor A (m)	Number of points	Contractor T (m)	Number of points	Contractor L (m)	Number of points
Base 0	0.003 ± 0.015	57 of 60	-0.024 ± 0.018	60 of 60	-0.006 ± 0.018	61 of 61
Base 1	0.014 ± 0.026	58 of 60	-0.013 ± 0.013	60 of 60	-0.002 ± 0.016	61 of 61
Base 2	0.011 ± 0.026	59 of 60	-0.022 ± 0.039	56 of 60	0.004 ± 0.029	61 of 61
Base 5	0.008 ± 0.034	58 of 60	-0.017 ± 0.026	55 of 60	0.015 ± 0.018	61 of 61
Base 10	-0.019 ± 0.026	59 of 60	-0.023 ± 0.030	20 of 60	0.025 ± 0.019	37 of 61
All, excluding Base 10	0.009 ± 0.026	232 of 240	-0.019 ± 0.025	231 of 240	-0.003 ± 0.022	244 of 244

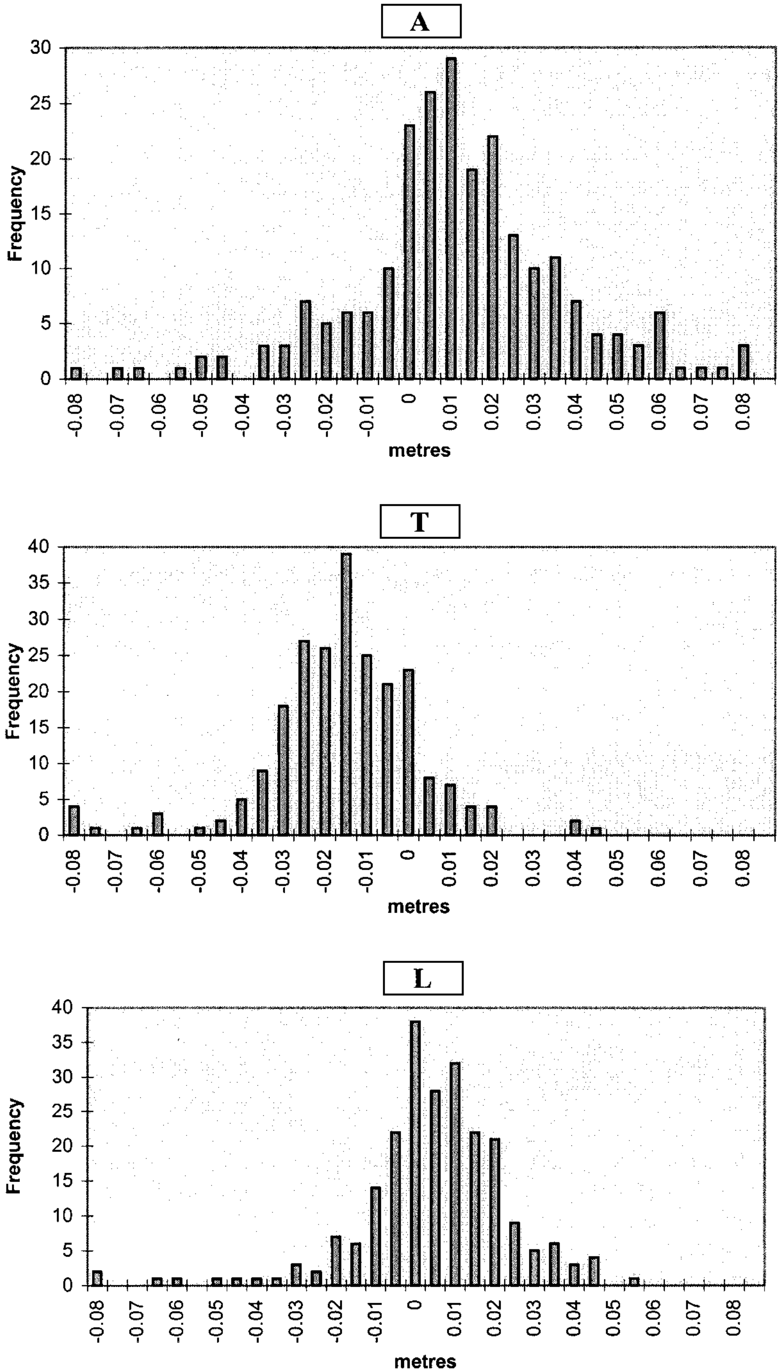

FIG. 3. Frequency Distribution of All Ellipsoidal Height Differences for Each Contractor's Data Combined for All Base Stations

Base 0 data sets have smaller standard deviations than the Base 2 and Base 5 data sets for most contractors. It may therefore be concluded (subject to the aforementioned restrictions) that a small amount baseline-length dependency is present up to distances of 10 km. Importantly, there is an overall increase in the proportion of null solutions with increasing baseline length. Therefore, as well as a degradation in the accuracy of RTK GPS ellipsoidal heights with increasing baseline length, there is a decrease in the reliability of the positioning with increasing baseline length.

Contractor L provided ellipsoidal heights at 61 points because it occupied an additional DOLA geodetic control point during its surveys. This additional point has been retained in this analysis so as to increase the sample size. This contractor also reoccupied the same test points by as many as five times for each base station. It must be acknowledged that this mode of data collection did allow for the *possibility* of preanalysis for each point by this contractor, where any points not agreeing with the remainder could simply be deleted, unbeknown to the writers. However, if this is the mode in which contractor L routinely conducts contract surveys for MRWA, it is commendable since a basic quality assurance procedure is applied. The possibility of regrouping and preanalyzing the data also applies to the other two contractors, because the same points were occupied separately from five base stations. Due to the absence of any explanation for the null solutions for some points, especially when they were coordinated by at least one other contractor, this cannot be ruled out as a plausible explanation for the missing data. These are purely speculations, but it is important to acknowledge the possibility for the sake of rigor. Therefore, care must be exercised if concluding that these null solutions are due only to RTK GPS system failure.

Analyses of All Data

Fig. 4 shows a histogram of the frequency distribution of the ellipsoidal height differences for all contractors' solutions measured from Base 0, Base 1, Base 2, and Base 5. The observations made from Base 10 have been omitted from this combined analysis because of the relatively large number of null solutions (Table 1). As the sample size increases, the effect of local correlations (for example between adjacent points observed a few minutes

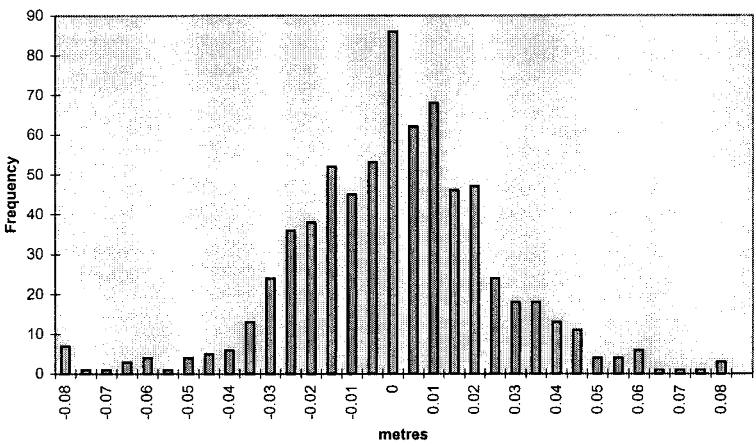

FIG. 4. Frequency Distribution of WGS84 Ellipsoidal Height Differences for All Contractors' Data (Excluding Base 10)

TABLE 2. Mean, Standard Deviation, and Proportion of Observed Differences in WGS84 Ellipsoidal Heights on Combining All Contractors' Solutions for Different Base Stations, Excluding Base 10

Base station	Number of solutions	Mean (m)	Standard deviation (m)
Base 0	178 of 180	0.009	±0.020
Base 1	179 of 180	−0.001	±0.022
Base 2	176 of 180	−0.002	±0.035
Base 5	174 of 180	−0.003	±0.029
All	707 of 720	−0.002	±0.027

apart by each contractor) is assumed to decrease, thus causing the frequency distribution to look more normal in shape. This permits the use of means and standard deviations in Table 2, which shows these statistics and the number of solutions grouped by base station for all contractors' data (excluding Base 10).

From Table 2, the distribution of ellipsoidal height differences for all solutions has a mean of −2 mm and a standard deviation of ±27 mm, which has been calculated from 707 of 724 possible position solutions (i.e. 17 null solutions). This suggests that, for this sample, the RTK GPS ellipsoidal heights do not differ statistically significantly from the control ellipsoidal heights. Also from Table 2, confidence limits can be derived, giving the expected range of errors for RTK GPS over baselines of up to 5 km in length. A single RTK GPS observation of ellipsoidal height falls within 53 mm of the control value 95% of the time. However, this estimate assumes no errors in the control network. Propagation of the 15 mm (95% confidence) error estimated for the control network into this confidence limit (assuming independence) reduces the value quoted to 51 mm (95% confidence). Recall, however, that this confidence limit does not take into account the fact that only 97.7% of the available points had positions supplied by the contractors.

Ellipsoidal Height Repeatability

Repeatability is defined and used here as the agreement of the RTK GPS ellipsoidal heights among contractors, rather than with the 60-point control network. Repeatability gives an estimate of the overall precision of the RTK GPS systems, and also allows a check for systematic errors in the control network. The repeatability has been computed by taking the mean of the three contractors' solutions for each point and computing the associated standard deviation (repeatability) for each base station. No repeatabilities were computed for Base 10 as insufficient position solutions were common among the contractors' data. The frequency distributions of these repeatabilities were not normal in nature, so Table 3 gives the median repeatability and the

TABLE 3. Median WGS84 Ellipsoidal Height Repeatability of All Contractors' Data for Different Baseline Lengths and the Proportion of Points with Repeatabilities Outside the Ranges ±20, ±40, and ±60 mm

Base station	Median (m)	>|±20 mm|	>|±40 mm|	>|±60 mm|
Base 0	0.019	38%	5%	0%
Base 1	0.021	52%	2%	0%
Base 2	0.018	46%	7%	3%
Base 5	0.025	67%	21%	8%

proportion of points that have repeatabilities outside the ranges of ±20, ±40, and ±60 mm.

From Table 3, an overall trend emerges indicating that, as distance from the base station increases, the percentage of points showing repeatabilities beyond ±20, ±40, or ±60 mm increases. These results support the conclusion drawn from Table 1 regarding RTK GPS solution degradation with baseline length. However, this inference is also subject to the limitation of an incomplete data sample. Since the repeatability estimate is independent of the control, the similar trends between Tables 2 and 3 indicate that no systematic error is present in the control network. Also, by comparing Tables 2 and 3, the vertical error in the control can be estimated to be approximately 15 mm (95% confidence), which is commensurate with the earlier accuracy estimation.

Analysis of Ellipsoidal Heights with Varying PDOP

It is generally assumed that the error in an RTK GPS solution will be directly proportional to the PDOP at the time the observations were made. Therefore, the differences between the ellipsoidal heights from the RTK GPS solutions and the control network versus PDOP are shown in Fig. 5. Only the Base 0 solutions have been considered so as to eliminate any potential effect due to the baseline-length dependency indicated earlier. All Base 0 solutions from each contractor are assumed to have no biases among them (Table 1) and are thus combined into a single sample.

Fig. 5 shows that no particular correlation between PDOP and the magnitude of the ellipsoidal height difference is observed for this data set. The largest PDOPs encountered were 6.85 and 7.37, which yielded a height difference from the control of −0.016 m and −0.039 m, respectively. Both these high PDOPs were part of the contractor T data set, which indicates that a different PDOP threshold was set in comparison with the other two contractors, both of whom never reported PDOPs of greater than 5.5 over their

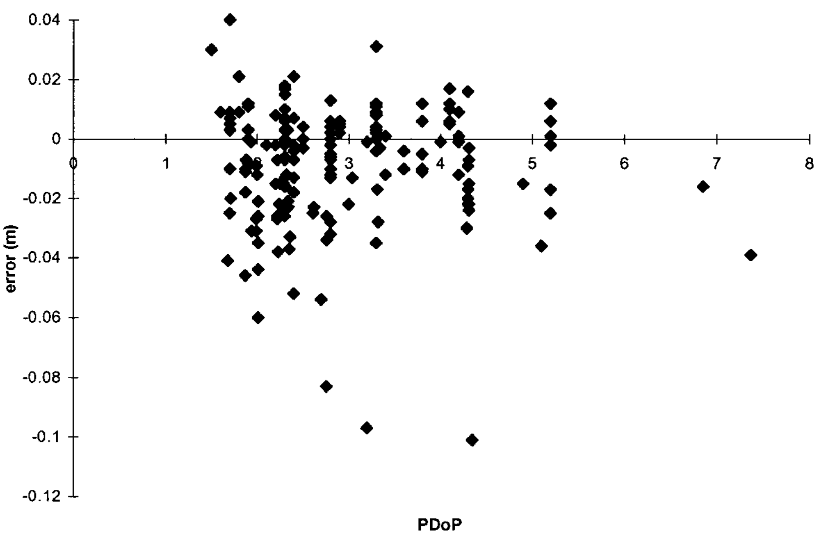

FIG. 5. WGS84 Ellipsoidal Height Differences versus PDOP for All Contractors' Data from Base 0

TABLE 4. Ellipsoidal Height Differences for Contractor T with Varying Number of Satellites and PDOP

Test station	Base station	Number of satellites	PDOP	Ellipsoidal height difference (m)
K20	Base 1	5	10.8	−0.013
P16	Base 5	4	13.6	−0.250
P17	Base 5	4	5.1	0.016
P18	Base 5	4	14.4	−0.181
P19	Base 5	4	15.3	−0.493
P20	Base 5	4	3.42	−0.892

entire data sets. From Fig. 5, PDOP does not appear to form the dominant source of error in RTK GPS positioning over the test network, provided that the PDOP is maintained below a value of 7. Indeed, in some cases, a high PDOP does not necessarily give a poor result; neither does low PDOP guarantee a good result. However, this must be qualified by the possible presence of other factors that affect the RTK GPS position solutions over the test network (notably multipath).

Table 4 illustrates the effect of poor PDOP on selected solutions from contractor T. The case of point P20 in the Base 5 data set is particularly instructive. This point was observed after points P18 and P19 had been observed with a high PDOP and only four satellites. The sequential increase in ellipsoidal height difference from P18 to P19 to P20 is indicative of incorrect ambiguity resolution. Although a low PDOP is observed at P20, the ellipsoidal height difference is larger than could be expected. This indicates that the RTK GPS software has fixed the wrong set of ambiguities at an earlier point, probably due to the high PDOP and low number of satellites. However, this assumes that a reinitialization occurred before observing point P18, though contractor T recorded no information to be able to verify this.

This type of problem can largely be eradicated using good survey practice. Generally, no RTK GPS solution should be accepted with a PDOP greater than 7 or with less than six visible satellites. Also, collecting continuous RTK GPS data for, say, around 10 min after reinitialization may have allowed the GPS software's internal statistical tests to identify the incorrect ambiguity resolution. However, these criteria provide no guarantee of the positional accuracy because of the presence of other system- and user-related errors.

ANALYSIS OF RTK GPS–DERIVED AHD HEIGHTS

Vertical Coordinate Transformation Techniques

WGS84 ellipsoidal heights can be transformed to the local vertical datum, in this case the AHD, in three principal ways (Featherstone et al. 1998)

1. The gravimetric method, where the geoid-WGS84-ellipsoid separation is interpolated from a precomputed grid of geoid heights then algebraically subtracted from the WGS84 ellipsoidal height. The best geoid model available in Australia at the time of the RTK GPS tests (1997) was called AUSGeoid93 (Steed and Holtznagel 1994).
2. The geometric method, where discrete geoid-WGS84-ellipsoid separations are computed from GPS measurements at existing vertical benchmarks, interpolated to other points, then algebraically subtracted from

the WGS84 ellipsoidal height. However, the geometrically derived geoid-WGS84-ellipsoid separation is limited to the combined accuracy of the GPS and AHD heights and the interpolation approach used.

3. The gravimetric-geometric method, where GPS measurements at existing benchmarks are essentially used to warp the gravimetric geoid model to fit the local vertical datum (Featherstone 2000). This makes the GPS-derived AHD heights compatible with existing benchmarks. This often is preferable because the geometrically derived geoid heights can account for any biases in the gravimetric geoid with respect to the benchmarks over each particular survey data. However, the combined gravimetric-geometric method is subject to the same precautions and limitations mentioned for the geometric method.

AHD Height Determination by Each Contractor

The coordinate transformation strategy described by contractor A was vague. It was applied using GPSTopo software via the selection of an ''anchor point.'' Given the limited amount of information and knowing that each contractor was supplied with the AHD heights of three points in the test network, it is likely that the geometrical method was used. However, it is unknown whether this software applies either a constant geoid-ellipsoid separation to the GPS ellipsoidal heights or interpolation of the geometrical geoid heights using a plane or any other surface.

The coordinate transformation strategy described by contractor T was supplied in relatively more detail. The heights provided by the RTK GPS system were termed ''pseudo-AHD'' heights by this contractor. It is assumed that this is because the AHD height of the base antenna was input to the base receiver. As such, the GPS-derived height of the roving antenna was determined by applying the change in WGS84 ellipsoidal height. If this was the case, then there is some question as to the accuracy of the GPS-derived ellipsoidal height differences. This is due to the incorrect WGS84 ''coordinate seeding'' from the base station. As a rule of thumb, every 10 m of error in the base station coordinates can introduce approximately 1 mm/km uncertainty in the GPS baseline vectors. As the geoid-WGS84-ellipsoid separation is approximately -32 m in the survey area, an error of approximately 3 mm/km may occur. This may provide an explanation for the bias observed in contractor T's data with respect to the control (Table 1). On completion of the surveys, contractor T downloaded the GPS-derived pseudo-AHD heights into Trimmap software. At this stage, the AUSGeoid93 gravimetric geoid heights were applied to these data. It is not clear what account was made for the application of the pseudo-AHD height during this post processing. Presumably, it was simply added or subtracted as a bias for each particular base station. Given these assumptions, the geoid modeling approach taken by contractor T was a variant of the gravimetric-geometric method.

The coordinate transformation strategy employed by contractor L was described in reasonable detail. The WGS84 coordinates were transformed to the Australian Geodetic Datum 1984 using a seven-parameter transformation. This, though strictly a horizontal coordinate transformation, yields ellipsoidal heights with respect to the Australian National Spheroid (ANS). Contractor L used these ANS ellipsoidal heights to determine AHD heights by applying the geoid-ANS-ellipsoid separation. A single value of this separation was determined for each of the three 20-station test subareas (Fig. 1) by subtacting the AHD height of a nearby point whose height was supplied. Therefore, the geoid modeling approach taken by contractor L can be classified as the geometrical method. Undoubtedly, the use of a constant value for each test subarea is the simplest to apply, but the proximity of each test subarea and the points for which the AHD height was known did allow the use of a plane surface for interpolation. It is also important to note that the transformation of WGS84 to ANS ellipsoidal height precludes the simple use of AUSGeoid93, which is referred to WGS84. Of course, AUSGeoid93 can be transformed to the ANS using the same seven-parameter transformation, but this introduces an unnecessary additional stage of calculations.

Contractor-by-Contractor Analyses of RTK GPS–Derived AHD Heights

Table 5 shows the mean, standard deviation, and number of the differences observed between the RTK GPS–derived AHD heights for test data set (i.e., computed by each contractor and for each base station) and the control AHD heights. These statistics could only be computed for those points provided by the contractors. The null solutions and the four points provided by contractor T with a PDOP greater than 10 were omitted from the analysis. The additional test point occupied by contractor L has also been omitted, because the AHD height of this station was not determined with class A geodetic leveling by MRWA. Table 4 indicates that significant biases remain between the RTK GPS–derived AHD heights and the test network. In some cases, such as for contractor T, the geoid modeling procedure reduces the biases observed for the GPS ellipsoidal height data (Tables 5 and 1). In other cases, particularly for contractor L over short baselines, larger biases are introduced. These biases are most probably due to the disparate effects of different geoid modeling approaches taken by each contractor.

The next analysis determined whether there was any baseline-length dependency in the differences between the RTK GPS–derived and control AHD heights. As seen earlier, a baseline-length dependency of the ellipsoidal heights over distances up to 10 km appears to exist. An additional baseline-length dependency is expected to have been introduced into the AHD heights because of the steep east-west geoid gradient over the test area. Therefore, this analysis will show, among other things, whether each contractor's application of the vertical coordinate transformation is correct. In Figs. 6–8, the AHD height difference is plotted against the distance supplied by each contractor. The distances appear to differ among Figs. 6–8 for the Base 10 observations, but this is due only to the fewer number of points positioned from this base station by the contractors. Figs. 6–8 also show that the AHD height differences for each test cluster are scattered about a different mean than that of an adjacent test cluster. It is most likely that this is due to incorrect modeling of the geoid gradient between each cluster within the test network (Fig. 1) for a single base station and between each base station.

Analysis of All RTK GPS–Derived AHD Heights

The analysis of all contractor's data for each base station is shown in Table 6. As seen for the ellipsoidal heights (Table 2), treating all the data as a single sample tends to average out the biases seen among the contractors and base stations in Table 5 and Figs. 6–8. From Table 6, the 95% confidence limit for all the RTK GPS–derived AHD heights is 53 mm with respect to the control. When accounting for the error in the control AHD heights of 2 mm (95% confidence), also gives a 95% confidence of 53 mm. This is because the estimated error in the control is insignificant in comparison to the observed height differences. However, these estimates are based on only the

TABLE 5. Mean, Standard Deviation, and Proportion of Observed Differences between Optically Leveled and GPS-Derived AHD Heights for Each Base Station and Each Contractor

Base station	Contractor A (m)	Number of points	Contractor T (m)	Number of points	Contractor L (m)	Number of points
Base 0	0.006 ± 0.015	57 of 60	-0.017 ± 0.018	59 of 60	-0.025 ± 0.026	60 of 60
Base 1	0.017 ± 0.025	58 of 60	-0.006 ± 0.013	60 of 60	-0.020 ± 0.024	60 of 60
Base 2	0.014 ± 0.027	59 of 60	-0.010 ± 0.013	56 of 60	-0.016 ± 0.034	60 of 60
Base 5	0.009 ± 0.034	58 of 60	-0.009 ± 0.025	55 of 60	-0.003 ± 0.020	60 of 60
Base 10	-0.018 ± 0.027	59 of 60	-0.017 ± 0.031	20 of 60	-0.003 ± 0.026	36 of 60
All, excluding Base 10	0.012 ± 0.026	232 of 240	-0.011 ± 0.018	234 of 240	-0.016 ± 0.027	240 of 240

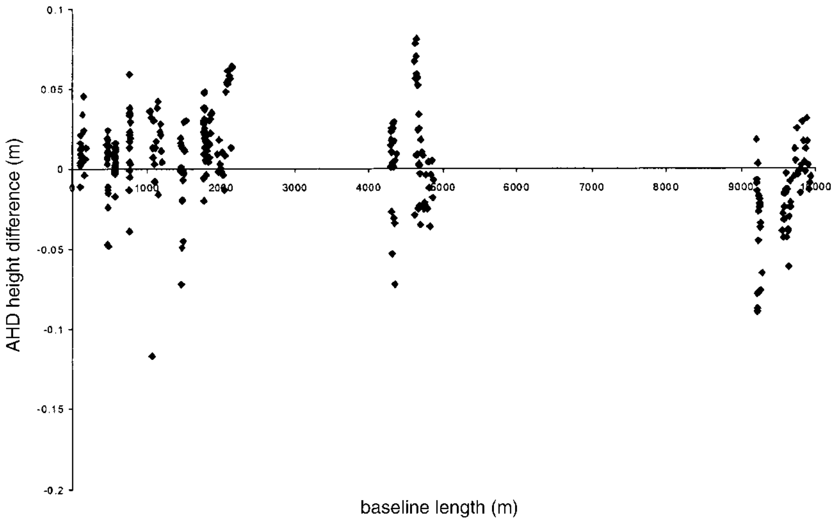

FIG. 6. Differences between Optically Leveled and GPS-Derived AHD Heights versus Baseline Length for Contractor A

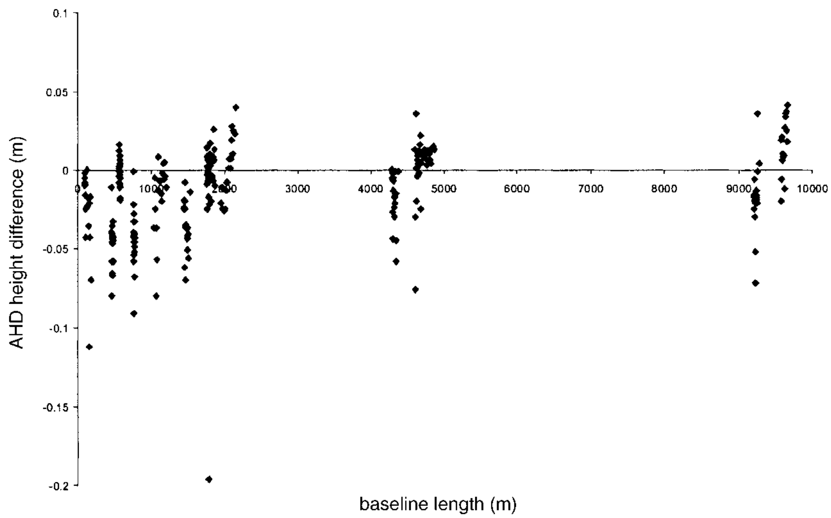

FIG. 7. Differences between Optically Leveled and GPS-Derived AHD Heights versus Baseline Length for Contractor T

97.7% of points supplied by the contractors and without those observed with a PDOP of greater than 10. When using all data in this way, it indicates that the RTK GPS–derived AHD heights collected by the contractors can meet MRWA's specifications for formed and natural surfaces, but not for sealed surfaces. Again, this conclusion deserves qualification because of the spatial and temporal nature of the accuracy of RTK GPS positioning. While every attempt was made to replicate contract surveys for MRWA, the conclusions

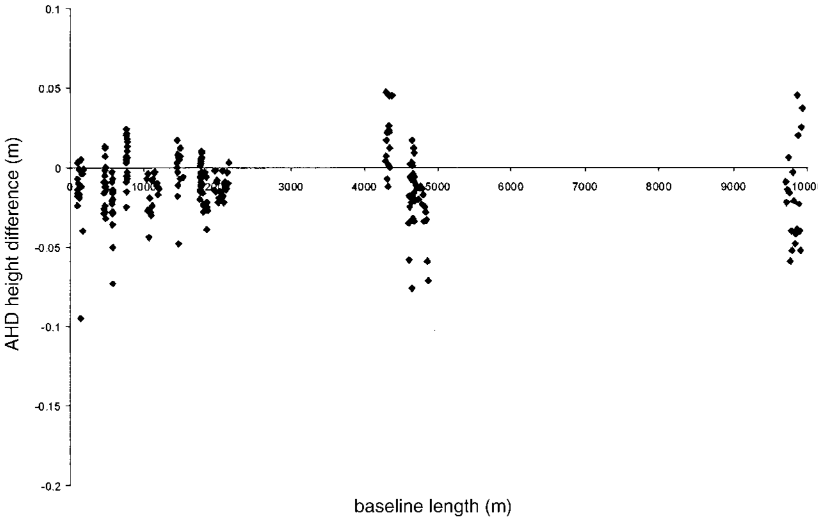

FIG. 8. Differences between Optically Leveled and GPS-Derived AHD Heights versus Baseline Length for Contractor L

TABLE 6. Statistics of Differences between Optically Leveled and GPS-Derived AHD Heights on Combining All Contractor's Data for Different Base Stations, Excluding Base 10

Base station	Number of points	Mean (m)	Standard deviation (m)
Base 0	176 of 180	−0.012	±0.024
Base 1	178 of 180	−0.003	±0.027
Base 2	175 of 180	−0.004	±0.029
Base 5	173 of 180	−0.001	±0.028
All	702 of 720	−0.005	±0.027

really only apply to the environment of the test network and at the times that the surveys were conducted.

Fig. 9 shows the AHD height differences plotted against the distance for all contractors' data. As with Figs. 6–8, Fig. 9 shows that the differences for any one base station are scattered about different means (Table 6). However, these mean differences are slightly smaller for shorter baselines, which most likely is due to incorrect modeling of the geoid gradient between each base station. Recall that this was expected to be the case, given the deliberate east-west orientation of the test network.

CONCLUSIONS AND RECOMMENDATIONS

A 60-point test network, established on Curtin University of Technology's Bentley Campus in Perth, Western Australia, has been used to evaluate the accuracy of three different brands of RTK GPS equipment as used by three different contractors. Therefore, these evaluations assess the *combined* performance of the RTK GPS equipment as used by each contractor, since they are inextricably linked. No evidence of any vertical bias in the test network

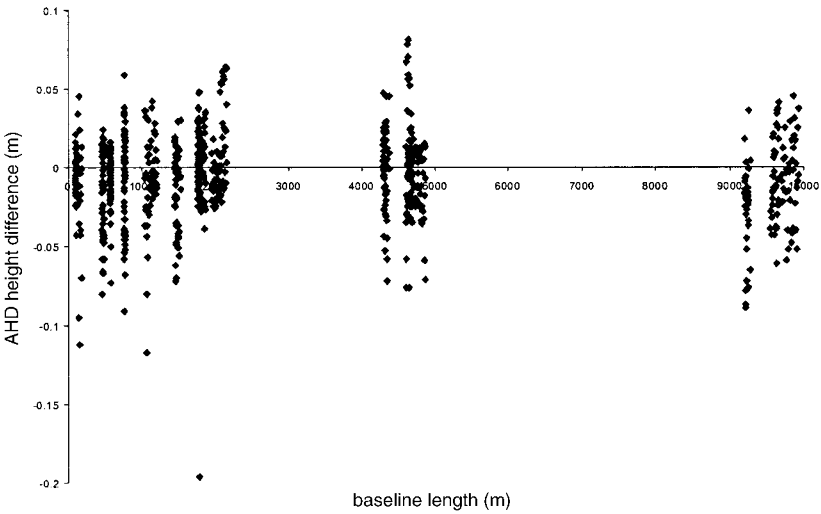

FIG. 9. Differences between Optically Leveled and GPS-Derived AHD Heights versus Baseline Length for All Contractors' Data

was detected, and the estimated error in the control WGS84 ellipsoidal heights of 15 mm (95% confidence) appears to be realistic.

The contractors used RTK GPS surveying with respect to five base stations to position these 60 points in ellipsoidal and AHD height. From the analysis of the data, it can be inferred that

- RTK GPS–derived AHD heights do not meet MRWA's specifications for sealed roads, but do meet the specifications for formed and natural surfaces.
- The accuracy estimate (accounting for errors in the control data) for RTK GPS ellipsoidal height recovery, derived using all contractor's data for baselines less than 5 km, is 51 mm (95% confidence).
- The accuracy estimate (accounting for errors in the control data) for RTK GPS–derived AHD height recovery, derived using all contractor's data for baselines less than 5 km, is 53 mm (95% confidence)
- There is degradation with increasing baseline length, both in terms of an increase in the number of null solutions and an increase in the magnitude of the height differences.

Importantly, the above confidence limits do not account for the null solutions, and in the case of the AHD heights, also omit positions determined with a PDOP greater than 10. Therefore, the confidence regions given represent 95% of the acceptable solutions, rather than 95% of all points occupied by the contractors. The overall reliability rate of the three RTK GPS surveys was 97.7% for baselines less than 5 km long.

The RTK GPS ellipsoidal heights appear to contain some system- and environment-dependent biases, which vary among contractors and base stations. For instance, statistically significant offsets from the control of approximately 9 mm for contractor A and −19 mm for contractor T were observed for baselines less than 5 km long. The magnitude and direction of the biases observed in the ellipsoidal heights subsequently change for the RTK GPS−derived AHD heights, which is due to the vertical coordinate transformation techniques utilized by the contractors. Overall, it appears that RTK GPS survey practice is variable among the contractors in terms of quality, and the application of geoid corrections to GPS data may not be well understood.

Finally, quality assurance remains an important issue for RTK GPS surveying. It must be reiterated that single radiations are insufficient to provide any form of quality control, even if some known points are positioned during a RTK GPS survey. RTK GPS simply cannot be used to the same level of accuracy at any time and in any environment. Generally, six GPS satellites with a PDOP value of less than 7 are the minimum configuration under which a RTK GPS user should consider operation. Care must also be taken during the reinitialization process as wrong ambiguity resolution can lead to severe errors. A suggested field procedure is to collect continuous RTK GPS data for approximately 10 min after each initialization, so as to increase redundancy and allow the software's statistical testing to verify the correct ambiguity set.

Any user of RTK GPS surveying must make specific choices regarding operation of the system. As in this study, variables such as maximum PDOP, minimum number of satellites, occupation times, modes of independent quality control and vertical coordinate transformations vary substantially among users. In the preceding analysis, no attempt was made to discern between different field procedures or to standardize observation techniques. Rather, contractors were instructed to ''do their own thing.'' Therefore, the results presented are only a broad indication of the capabilities of RTK GPS as applied in a real world situation. This may be in contrast with the observation environments and tests from which manufacturers' specifications are derived. It is thus recommended that further work be undertaken to firmly establish independent, objective estimates of precision, accuracy, and reliability for RTK GPS surveying in both horizontal and vertical directions. It would seem appropriate that such estimates should operate under an internationally agreed set of guidelines that standardize RTK GPS usage. This is considered an urgent need, given that RTK GPS surveying has established itself as a routine technique, seemingly due only to efficiency considerations and based on internal precision estimates and manufacturers' brochures.

ACKNOWLEDGMENTS

We would like to thank the following individuals and organizations: The Australian Research Council for supporting research in GPS height determination through grant A39938040; Main Roads Western Australia, particularly Hennie Kuenen, Allan Gregory, John Baker, and Geoff Xanthis, for providing the resources, data, and support with which to conduct this research; the Western Australian Department of Land Administration for assisting with the establishment of the RTK GPS Test Network; the Australian Surveying and Land Information Group for providing the AUSGeoid93 geoid model; and the contractors for providing their RTK GPS data.

REFERENCES

Edwards, S. J., Cross, P. A., Barnes, J. B., and Betaille, B. (1999). ''A methodology for benchmarking real-time kinematic GPS.'' *Survey Rev.*, 35(273), 163–174.

Featherstone, W. E. (2000). ''Refinement of gravimetric geoid using GPS and leveling data.'' *J. Surv. Engrg.*, ASCE, 126(2), 27–56.

Featherstone, W. E., Dentith, M. C., and Kirby, J. F. (1998). ''Strategies for the accurate determination of orthometric heights from GPS.'' *Survey Rev.*, 34(267), 278–296.

Finn, A. (1997). ‘‘GPS interference.’’ *Australian Meeting of the Civ. GPS Interface Com.*, Canberra, Australia.

Inter-governmental Committee on Surveying and Mapping (ICSM). (1996). *Standards and practices for control surveys*, Version 1.3. Canberra, Australia.

Main Roads Western Australia (MRWA). (1997). ‘‘Digital ground survey specification.’’ *Document 2300/05/09*. Issue 1, Perth, Australia.

Steed, J., and Holtznagel, S. (1994). ‘‘AHD heights from GPS using AUSGeoid93.’’ *The Australian Surveyor*, 39(1), 21–27.

U.S. Departments of Defense and Transportation. (1997). *1996 Federal Radionavigation Plan*, National Technical Information Service, Springfield, Va.

Wylde, G. P., and Featherstone, W. E. (1995). ‘‘An evaluation of some stop-and-go kinematic GPS survey options.’’ *The Australian Surveyor*, 40(3), 202–212.