The GPS Contribution to the Error Budget of Surface Elevations Derived From Airborne LIDAR
Matt A. King

Abstract—When using airborne LIDAR to produce digital elevation models, the Global Positioning System (GPS) positioning of the LIDAR instrument is often the limiting factor, with accuracies typically quoted as being 10–30 cm. However, a comprehensive analysis of the accuracy and precision of GPS positioning of aircraft over large temporal and spatial scales is lacking from the literature. Here, an assessment is made of the likely GPS contribution to the airborne LIDAR measurement error budget by analyzing more than 500 days of continuous GPS data over a range of baseline lengths (3–960 km) and elevation differences (400–2000 m). Height errors corresponding to the 95th percentile are < 0.15 m when using algorithms commonly applied in commercial software over 3-km baselines. These errors increase to 0.25 m at 45 km and > 0.5 m at 250 km. At aircraft altitudes, relative heights are shown to be potentially biased by additional errors approaching 0.2 m, partly due to unmodeled tropospheric zenith total delay (ZTD). The application of advanced algorithms, including parameterization of the residual ZTD, gives error budgets that are largely constant despite baseline length and elevation differences. In this case, height errors corresponding to the 95th percentile are < 0.22 m out to 960 km, and similar levels are shown for one randomly chosen day over a 2300-km baseline.

Index Terms—Aircraft navigation, error analysis, Global Positioning System (GPS), tropospheric propagation.

I. INTRODUCTION

SURFACE elevations derived using airborne LIDAR sensors controlled using positions from Global Positioning System (GPS) data are now used for a wide range of scientific and engineering applications [1]–[6]. These elevations have typical root-mean-square errors of 10–30 cm when compared to independent ground-truth data [1]–[3], [7]–[12]. The dominant component of the airborne LIDAR error budget is often the GPS positioning [3], [13]. However, LIDAR errors are typically examined over relatively small spatio-temporal scales without separating out the individual error components; a rigorous study of GPS-related error characteristics is currently missing from the literature. This is mainly due to a lack of long-running (ideally continuous) and large-scale data sets from which the error budget could be carefully examined (see [3] and [12] for two partial exceptions). This is, of course, a result of finite aircraft flight time, generally on the order of 2–6 h, and dedicated validation experiments require significant expense and effort to control the various error sources reliably [9].

When considering the GPS component of the error budget, an alternative approach is to use continuously operating ground-based GPS receivers that are spaced (in both horizontal and vertical directions) and analyzed in a way that is representative of typical flying distances and heights for LIDAR studies. The advantage of this approach is that an accurate “truth” value is available for the ground-based stations. Applying GPS data processing techniques identical to those applied to airborne data sets and then comparing to the truth values yield a robust assessment of the GPS contribution to the surface height error budget from airborne LIDAR data. The main difference is the dynamics of the aircraft. However, apart from turning maneuvers that may cause short loss of lock on low-elevation GPS satellites, the nature of LIDAR data collection demands stable flying patterns, and these differences are minimized.

In this paper, continuous GPS data spanning more than one year is processed with sites separated horizontally by between 15 and 960 km and vertically by between 400 and 2000 m. By using two different GPS software packages, kinematic GPS solutions are computed and their heights are compared to “known” values.

II. BACKGROUND

Precise and accurate positioning of aircraft using GPS was first performed in the late 1980s [14]. Kinematic positioning is fundamentally less precise than static positioning since station coordinates must be estimated each measurement epoch. That is, the redundancy is less in the least squares solutions. As a result, minimizing the number of other estimated parameters in the solution is highly desirable. For instance, real-valued ambiguity terms (N) must either be fixed to integer values (and then removed from the solution) or stabilized, such as that occurs within a Kalman filter. Incorrect ambiguity fixing results in a drift in the determined position which may exceed 0.5 m after as little as 15 min of flying [15]. The fundamental observable in most relative GPS processing software is the so-called double-difference (∆) carrier-phase observable, which may be written in units of distance as [16]

$$\Delta \nabla \Phi(t) = \Delta \nabla \rho(t) + \lambda \Delta \nabla N$$

$$+ \Delta \nabla I(t) + \Delta \nabla T(t) + \Delta \nabla \delta(t) + \varepsilon$$

(1)

where $\Phi$ is the measurement made at time $t$ by the GPS receiver of each observed satellite’s carrier phase, $\rho$ is the geometric distance between the antenna and the satellite, $\lambda$ is the wavelength, $I$ is the ionospheric phase advance along the
signal path, $\delta$ denotes multipath effects, and $\varepsilon$ denotes other small errors. Carrier-phase multipath is $\sim<0.05$ m [16], but this has been observed to propagate to aircraft height errors as large as 0.1 m with periods of $\sim\! 10$ s [17].

For precise and accurate GPS heighting, one important additional unknown is the tropospheric delay ($T$) at the elevation angle and azimuth of each satellite. The values for $T$ are normally related to a single zenith total delay (ZTD) term using appropriate mapping functions ($m$)

$$T = ZHD \cdot m_h(\theta) + ZWD \cdot m_w(\theta)$$

where ZHD and ZWD are the hydrostatic and wet components of ZTD, respectively, such that $ZTD = ZHD + ZWD$, and $m_h$ and $m_w$ are their respective mapping functions [18]. While not discussed further here, the mapping functions are not perfectly known and introduce some small error. Importantly for airborne LIDAR surveys, errors in accounting for ZTD are amplified by 2–3 times into station height errors [19], making them potentially a dominant source of error.

Unlike ionospheric effects, which can be negated through cancellation over short baselines or through forming an ionosphere-free linear combination of dual-frequency data [16], ZTDs must either be modeled or estimated. For a reference station and rover at approximately the same elevation, tropospheric effects are similar over distances of up to 10–20 km and almost fully cancel. Beyond this range and/or elevation difference, substantial relative ZTD errors may bias the derived aircraft elevations.

For baselines longer than 10–20 km, relative ZTD can increase to 0.2–0.3 m, although it is more typically 0.05–0.1 m. It is important to note that there are many applications where long-baseline LIDAR is significantly more cost-effective than the alternative of placing many local base stations [1], [20], and so, these effects must be considered.

However, ZTD is also important for aircraft only short (<10–20 km) distance from the GPS base station. This is because, for elevation differences of several hundreds of meters from the base station, as typically required in airborne surveys, relative ZTD can be up to 0.05–0.3 m even if the horizontal baseline is small [21]. This bias will vary according to flying height and also time (as the aircraft and weather systems move relative to one another, for instance). Dealing with relative ZTD is therefore critical to obtaining accurate and precise aircraft positioning using GPS at any distance from a base station.

The most convenient approach to account for relative ZTD is through application of a tropospheric model. However, while the ZHD may be modeled accurately using local pressure measurements at each receiver, ZWD models are generally less accurate [22], [23], leaving residual error as large as $\sim 1$–2 dm and consequent rover height errors 2–3 times this. This error will have a high spatio-temporal variability such that repeat LIDAR surveys may appear to show differences in ground elevation which could be erroneously interpreted as surface change (see, e.g., [3]).

An alternative approach, routinely used in geodetic-quality static processing packages, is to model the ZHD and then parameterize the residual, consisting mainly of ZWD plus ZHD model errors [24]. Due to the motion of the aircraft relative to the regional ZTD pattern, a new ZWD parameter must be estimated more frequently than for static applications—up to every measurement epoch. More sophisticated parameterization approaches are, however, possible [25]. For the resulting improved long-term accuracy, a necessary tradeoff of introducing these additional parameters is a decrease in epoch-to-epoch repeatability. Simplified estimation strategies, such as estimating scale factors to a ZTD model, are also possible [26], [27], but a constant scale factor may not be particularly valid for aircraft positioning (see also [28]).

By using parameterized ZTD, it is theoretically possible to achieve a similar level of accuracy and precision at any baseline length, even a few thousands of kilometers from the reference station [29].

In the next section, tests are described that were designed to determine typical GPS-related errors that could be expected from both conventional commercial software packages and from state-of-the-art packages incorporating residual ZTD estimation. For airborne studies, the GPS-related error budget is dominated first by ZTD and then carrier-phase multipath errors. These tests not only explicitly examine the effects of different approaches to dealing with ZTD but also implicitly include the effects of carrier-phase multipath and other smaller errors. Consequently, these tests provide bounds to the GPS-related error budget of airborne LIDAR data collection.

### III. DATA AND METHODOLOGY

#### A. Data

Two GPS data sets were assembled, and these are shown in Fig. 1. The United Kingdom data set consists of data from 13 continuous GPS receivers spaced at regular intervals from a reference receiver in London (LOND). This data set is designed to test the effect of increasing baseline length on kinematic GPS errors. The second data set consists of data from five continuous GPS receivers in the European Alps spaced at regular intervals in elevation, approximately equidistant from a reference receiver in Innsbruck (HFLK). This data set is designed to test the effect of different elevations on kinematic GPS errors. The elevation differences of the Alps network are reversed to those in airborne surveys, with the base station at elevation, but the conclusions will be unaffected. The baseline lengths and elevation differences are given in Table I.

The study period spanned 514 days over the period July 1, 2005 through November 27, 2006. There were few data outages and hardware changes during this period. Each of the sites recorded data at 30-s intervals. Typical airborne surveys are conducted using GPS data collected every 0.1–1.0 s. For low sampling rates, some GPS receivers use the carrier phase to smooth the pseudorange data, resulting in less noisy pseudorange for ambiguity fixing than would be possible when using data at higher data rates. This may mean that the tests described in the succeeding discussions will have higher rates of ambiguity fixing than would be possible with 1-s data. On the other hand, lower sampling rates also increase the likelihood of “false positives” in carrier-phase cycle-slip detection routines and hence increase the number of ambiguity parameters that need
Fig. 1. Location of sites in the (left) U.K. network and (right) Alps network. Symbols are shaded according to the site elevation.

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Length (km)</th>
<th>Height difference (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOND-BARK</td>
<td>2.9</td>
<td>38.1</td>
</tr>
<tr>
<td>LOND-NEOT</td>
<td>45.9</td>
<td>92.6</td>
</tr>
<tr>
<td>LOND-PETE</td>
<td>77.5</td>
<td>56.3</td>
</tr>
<tr>
<td>LOND-LINC</td>
<td>121.8</td>
<td>6.8</td>
</tr>
<tr>
<td>LOND-SCAR</td>
<td>195.8</td>
<td>37.4</td>
</tr>
<tr>
<td>LOND-NEWC</td>
<td>250.0</td>
<td>82.9</td>
</tr>
<tr>
<td>LOND-ABED</td>
<td>307.4</td>
<td>84.2</td>
</tr>
<tr>
<td>LOND-LERW</td>
<td>389.1</td>
<td>59.8</td>
</tr>
<tr>
<td>HFLK-ZOIF</td>
<td>625.2</td>
<td>61.3</td>
</tr>
<tr>
<td>HFLK-SBGZ</td>
<td>958.9</td>
<td>65.3</td>
</tr>
<tr>
<td>HFLK-ABED</td>
<td>147.2</td>
<td>-376.6</td>
</tr>
<tr>
<td>HFLK-SBGZ</td>
<td>140.8</td>
<td>-1060.6</td>
</tr>
<tr>
<td>HFLK-OBE2</td>
<td>86.3</td>
<td>-1742.7</td>
</tr>
<tr>
<td>HFLK-BZRG</td>
<td>90.6</td>
<td>-2055.0</td>
</tr>
</tbody>
</table>

TABLE I
BASELINE DETAILS FOR THE U.K. AND ALPS NETWORKS

to be fixed/stabilized. In practice, and with high-quality data, differences are often small, and the results of the tests described in the succeeding discussions may be regarded as being representative of the true error budget for airborne GPS positioning.

A truth data set was constructed by analyzing these same data using a conventional 24-h precise point positioning (PPP) approach [30]. In brief, site coordinates were estimated along with epoch-by-epoch tropospheric zenith delays and gradients and receiver clock terms. Nonfiducial satellite positions and clocks and Earth orientation parameters were held fixed to values provided by the Jet Propulsion Laboratory. The nonfiducial station coordinates were then transformed into the International Terrestrial Reference Frame 2000 (ITRF2000) [31]. The individual daily solutions were then combined in a least squares sense to estimate site-specific velocities and coordinates for the midepoch of the survey in the ITRF2000. Since the velocity gradients across the respective networks are small, these motions were ignored and only the station coordinates used in subsequent comparisons. None of the stations exhibited jumps or other anomalous motions that could affect the comparison.

B. Methodology

The data were processed in two different processing software packages. In contrast to the PPP approach used to determine the truth values, airborne data processing typically uses single or, more commonly, double differences [16] that result in positions relative to a base station. The two processing packages used here adopt such a relative processing strategy. The first package is a conventional commercial survey package, Trimble Total Control (TTC) v2.73 [32]. The algorithms present in TTC are proprietary but are believed to be representative of algorithms present in most off-the-shelf commercial survey packages, including those used for airborne surveys. Of particular relevance is the lack of parameterization of residual ZTD in kinematic applications. Instead, a modified version of the Hopfield [33] tropospheric model is here chosen [34] as found, for example in [35]. The model input values are defined by the Mass Spectrometer and Incoherent Scatter Extended Atmospheric Model 1990 (see http://nssdc.gsfc.nasa.gov/space/model/).

The second package is the Track v1.14 kinematic processing software developed as part of the GAMIT/GLOBK GPS processing suite [36], [37]. Track allows parameterization of residual ZTD at every measurement epoch within the context of a Kalman filter solution. The ZTD is mapped to the elevation.
angle of each satellite using the MTT mapping function [38]. Because of ZTD parameterization, Track represents a range of packages that are state-of-the-art in airborne GPS data analysis [21].

The data were processed in a way that best approximated real flight data processing. First, airborne surveys, run according to good practice, start GPS data collection at least 10 min prior to the flight and continue for 10 min after. This process assists in ambiguity fixing and hence can dramatically improve the determined aircraft positions. In our processing, the first and last 10 min of rover data were modified in such a way that only one rover position was effectively estimated during these windows. Second, the data were processed in each package to simulate flight durations of 3 h; a further set of solutions was performed in Track for 6-h sessions to test the effect of session length. The sessions arbitrarily straddle midday on each of the considered days. Third, for Track, the site motion constraints were set as for a real flight. This was 1000 m per 30-s epoch (equivalent to 120 km/h) for rover position and 1 mm/epoch for the ZTD estimates. These represent loose constraints. TTC does not allow site motion constraints to be entered, presumably since site motion is completely unconstrained in their solutions.

Reference station (LOND or HFLK) coordinates were fixed to the truth values. International GNSS Service final satellite orbits were held fixed in the solutions [39]. An elevation cutoff angle of 10° was used in all solutions. For consistency across various baselines, all baselines were processed using the ionosphere-free linear combination [16]. Corrections were made for antenna phase center variations. Solid Earth tides were modeled in Track, but their status is not known in TTC. Over baselines up to 100 km or so, the effect is negligible, but the longer baselines considered here may have residual solid Earth tides of several millimeters if these are not modeled. Ocean tide loading displacements were not modeled in either package, and for the longer U.K. baselines, this may also introduce diurnal and semidiurnal signals of several millimeters.

Data were rejected in a way that best replicated outlier rejection in airborne surveys, i.e., where no truth value is available. Epochs were rejected where elevation formal errors were greater than four times the median. In addition for the Track solutions only, epochs were rejected where ZTD estimates deviated by more than 0.2 m from their median value. TTC solutions were not possible for the four longest baselines in the U.K. network due to a hard-wired software limit, but the TTC solution quality over >100-km baselines is established by the available solutions. Furthermore, much of the data for site OBE2 would not process through either TTC or Track, and the OBE2 results should therefore be considered with caution.

The determined heights, relative to the reference station, were then differenced from the truth heights in order to compute a GPS-related height error for every 30-s epoch computed over the 514 days. These errors are discussed in the following section.

IV. RESULTS

The distributions of height errors are shown in Figs. 2 and 3, for the U.K. and Alps networks, respectively, in the form of box-and-whisker plots. The errors shown are dominated by ZTD-related errors and carrier-phase multipath, the latter being common to the TTC and Track solutions.

First, considering the U.K. network (Fig. 2), the TTC relative height solutions are very precise for the short 3-km baseline with negligible mean bias (Table II), with 95% of the data falling within 0.144 m (Table III). The Track solutions are much less precise (95% of data within ~0.25 m), partly due to frequent cycle slips in the BARK data being better handled by TTC than with the chosen Track settings. With increasing baseline length, the TTC solution precision decreases, although...
TABLE II
MEAN HEIGHT BIASES ACROSS ALL DAYS. THE HFLK–OBE2 BASELINE
SUFFERS FROM POOR DATA QUALITY

<table>
<thead>
<tr>
<th>Baseline</th>
<th>Track</th>
<th>TTC</th>
<th>Track</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>3h</td>
<td>3h</td>
<td>6h</td>
</tr>
<tr>
<td>LOND–BARK</td>
<td>-0.016</td>
<td>0.007</td>
<td>-0.010</td>
</tr>
<tr>
<td>LOND–STEV</td>
<td>-0.011</td>
<td>0.021</td>
<td>-0.008</td>
</tr>
<tr>
<td>LOND–NEOT</td>
<td>0.003</td>
<td>-0.016</td>
<td>0.003</td>
</tr>
<tr>
<td>LOND–PETE</td>
<td>-0.017</td>
<td>0.001</td>
<td>-0.013</td>
</tr>
<tr>
<td>LOND–LINC</td>
<td>-0.024</td>
<td>-0.018</td>
<td>-0.022</td>
</tr>
<tr>
<td>LOND–SWAN</td>
<td>-0.030</td>
<td>-0.012</td>
<td>-0.025</td>
</tr>
<tr>
<td>LOND–SCAR</td>
<td>-0.037</td>
<td>-</td>
<td>-0.033</td>
</tr>
<tr>
<td>LOND–NEWC</td>
<td>-0.035</td>
<td>-</td>
<td>-0.030</td>
</tr>
<tr>
<td>LOND–ABED</td>
<td>-0.043</td>
<td>-</td>
<td>-0.045</td>
</tr>
<tr>
<td>LOND–LERW</td>
<td>-0.173</td>
<td>-</td>
<td>-0.138</td>
</tr>
<tr>
<td>HFLK–ZOUF</td>
<td>-0.000</td>
<td>0.012</td>
<td>-0.003</td>
</tr>
<tr>
<td>HFLK–SBGZ</td>
<td>-0.026</td>
<td>0.000</td>
<td>-0.026</td>
</tr>
<tr>
<td>HFLK–OBE2</td>
<td>0.103</td>
<td>0.166</td>
<td>0.111</td>
</tr>
<tr>
<td>HFLK–BZRG</td>
<td>-0.006</td>
<td>0.067</td>
<td>-0.011</td>
</tr>
</tbody>
</table>

the mean solution bias remains small. In contrast, the Track solutions maintain a similar level of performance out to 389 km (LOND–NEWC), after which the solutions begin to become less precise. A slight decrease in precision with increasing baseline length is expected as the number of common satellites visible at each end of the baseline decreases. The 6-h Track solutions generally have smaller biases compared to the 3-h Track solutions, but the biases are systematically negative in both cases. These biases are generally larger than the TTC solutions. The Track LERW solution (959 km) is biased and has a large data range. The degradation of the Track solutions at long baselines is discussed further in the succeeding discussions.

Table III lists the height error value corresponding to the 95th percentile. Ninety-five percent of the epochs in the Track solutions have heights within ∼0.2 m of the truth value, at baseline lengths up to approximately 300 km. The 6-h solutions are more precise over the baselines > 300 km. The TTC solutions are more precise for the 3-km baseline, but for the other baselines, they are much less precise, with 95% of the epochs in TTC having heights within 0.25 m over 46 km and ∼0.5 m at 250 km.

From the U.K. network, then, it is possible to conclude that for very short baselines with small elevation differences, software that does not parameterize the ZTD is more precise than software that parameterizes the ZTD. Mean biases are small for these baselines where elevation differences are small, but this may not hold for sites at different elevations as explored with the Alps network.

Consider then the Alps network where baseline lengths are each close to 100 km, but baseline height differences range from 500 to 2500 m. This network more closely replicates the elevation difference for airborne surveys than the U.K. network. Fig. 3 shows that for ZOUF and SBGZ, the Track and TTC solutions are of comparable quality to the U.K. baselines of similar baseline lengths. As mentioned before, the OBE2 data are unfortunately of poor quality and should be ignored. Apart from this, the Track solutions have biases < 0.03 m, and 95% of the epochs are within 0.18 m (BZRG) or 0.08 m (SBGZ and ZOUF). Ninety-five percent of the TTC solutions are within ∼0.2–0.3 m of the height truth, with little bias at BZRG and SBGZ. At ZOUF, however, the mean bias in the TTC solution is 0.067 m.

The result for BZRG shows a distinct positive bias in the TTC solutions and much larger scatter than in the Track solutions for the same site. The HFLK–BZRG baseline has a 2055-m height difference (Table I). Closer examination of the BZRG height errors reveals a large temporal variation with approximately seasonal period (Fig. 4). In contrast, the Track solutions are essentially bias free and with small scatter. Indeed, the mean bias of the Track BZRG solutions is not substantially different from comparable-length baselines in the U.K. network which have small elevation differences.

The seasonal variations in the TTC relative heights for BZRG correlate well (R = 0.53) with the estimated ZTD from the Track solutions (Fig. 4), after smoothing using a 30-day running mean. The TTC solutions are up to 0.16 m from the true value,
and these biases persist for several days. The Track solutions, however, remain largely unbiased. Santerre [19] predicts that for every 0.10 m of ZTD error, station heights are biased by approximately 0.22 m. The Track-determined relative ZTD in Fig. 4 is shown scaled by a factor of two, and the agreement in amplitude strongly suggests a relationship between the TTC height errors and the unmodeled ZTD in these solutions.

Height biases that increase with aircraft elevation above the GPS reference station are to be expected in airborne GPS solutions that do not parameterize the ZTD [21]. However, examining the correlation of Track-derived ZTD with temporal fluctuations in TTC height solutions for the other two Alps baselines demonstrates that this relationship may not always be a simple one. For instance, while HFLK–ZOUF ($R = 0.56$) has a similar level of correlation to HFLK–BZRG, the HFLK–SBGZ baseline has a negligible correlation ($R = -0.09$). Nonzero mean ZTD for these two baselines is also not reflected in the mean bias of the TTC solutions. The actual height bias in TTC clearly depends on the actual residual ZTD left after applying the TTC tropospheric model; applying simple models is clearly insufficient when the highest precision and accuracy are required.
Fig. 5. Same as Fig. 2, but with days 180–310 in 2005 and 2006 removed.

V. DISCUSSION

Perhaps the majority of airborne LIDAR surveys are conducted several tens of kilometers from a base station. In this case, the results for LOND–BARK and LOND–STEV, together with the results of the Alps network, are the most representative. These suggest that biases of up to ~0.1–0.2 m may occur with solution scatter of similar magnitude, simply due to GPS analysis deficiencies.

Longer baseline solutions may also be desirable, and in many cases, these are required. While the Track solutions in Fig. 2 are generally precise, their quality is much reduced in the longest two baselines. Examining the coordinate time series revealed that the solutions, after outlier rejection, were degraded during days 190–320 in 2005 and 2006. This degradation was very clear on the longer baselines but also affected the shorter baselines, suggesting that the origin was poor LOND data quality at these times. The exact origin remains unclear but could be related to seasonally dependent local conditions reducing the LOND data quality (e.g., local obstructions or signal interference).

These epochs were removed, and Fig. 5 shows the revised version of Fig. 2. The values corresponding to the 95th percentiles are shown in the right-hand columns of Table III. Apart from BARK, the Track solutions show uniform improvement in precision. Ninety-five percent of solutions are now within 0.15 m or out to ~400 km (LOND–NEWC). Most dramatically, the two longest baselines now have relative height errors <0.22 m. The biases are also reduced and are no longer systematically negatively biased. To a lesser extent, the TTC solutions were also improved when the low-quality LOND data are removed. The BARK solutions in Track remain degraded since they are dominated by the large number of cycle slips.

Since poor base station data quality is not typical, Fig. 5 and the right-hand column of Table III may be regarded as a reliable representation of the baseline-length-dependent component of the GPS error budget for airborne surveys. The GPS-related error budget for airborne surveys may therefore be regarded as <0.25 m for baselines of many hundreds of kilometers, and <0.15 m for baselines ~400 km, when processed using algorithms similar to those in Track. For TTC-like solutions, the error budget quickly approaches 0.5 m. In solutions from both software packages, carrier-phase multipath will be an important, but identical, residual error source [17].

Multipath conditions will not be identical to those experienced on an aircraft. However, the U.K.-based sites we consider here are located on buildings (often on roofs) rather than on well-constructed geodetic monuments. Their near-field multipath field will therefore not be completely dissimilar to an aircraft. GPS antennas on aircraft have, of course, no far-field multipath source, although employed reference stations will do. Further work in understanding and mitigating aircraft multipath is required.

The error budget presented here may, however, be slightly conservative, for four reasons. First, when compared to 1 s or higher sampling rates typically used in flights, the 30-s GPS data sampling may be subject to a greater number of flagged cycle slips, thereby degrading the solutions. Second, in airborne surveys, the base station is normally located at, or near, the airport. Consequently, ambiguities can often be fixed to their (correct) integer values and removed from the least-squares solution prior to takeoff and following landing. The least squares solutions are strengthened as a result, and in-flight ambiguities should also be fixed more readily than the scenarios necessarily adopted in this study where the remote station begins and ends at some distance from the base station. That the 6-h Track solutions are generally closer to the true value than the 3-h Track solutions gives some support to this conclusion. Third, only a single base station was used—processing the
aircraft data against multiple base stations, as is allowed in some software (e.g., Track), will effectively average down some errors, and solutions will be less susceptible to cycle slips at a single station. Finally, the analysis in both TTC and Track has been done in an automated way—further improvements could be made through more careful data editing and model options available in both software packages.

An example of what is possible is shown in Fig. 6(a), where relative height errors are shown for a baseline approximately 2300 km in length in the high Arctic (Svalbard to Greenland). The data were selected at random and processed in Track (v1.16) using site constraints similar to those used for the U.K. and Alps networks, but the other processing options were tuned for these data. Two reference stations 30 km apart in Svalbard were employed. Both 1- and 30-s solutions were computed. Less than 50 epochs in either solution have been rejected as outliers. The 1-s solutions (95th percentile: 0.136 m) and 30-s solutions (0.191 m) are of comparable precision to the shorter baselines in Figs. 3 and 5. An uncorrected cycle slip may be degrading the solution from ∼22 h, and more careful analysis would further increase the precision. New well-controlled airborne LIDAR experiments are essential if our understanding of the spatio-temporal error budget is to be advanced further.

With increasing baseline length, the number of common satellites with which to form double differences is reduced. Solution precision is therefore affected. One commonly used alternative to long-baseline relative processing is kinematic PPP [8], [40], [41]. This approach has the advantage of not requiring a base station since precomputed satellite clock values are used and receiver clock errors are parameterized along with the station coordinates and, for the highest precision work, ZTD [30]. The increase in the number of parameters estimated is substantial, however. Accuracies are reported at the 1-dm level in height [8], [40], [42]. For completeness here, we show in Fig. 6(b) a kinematic PPP solution, generated using the GIPSY software [43], of the same KELY data set processed in Track. The kinematic PPP solutions show less low-frequency signal, but increased high-frequency signal. For this data set, the 95th percentile (0.114 m) is smaller than that of the relative solutions. Although the results of this brief comparison could not be extended to other data sets, it is likely that for very long baselines, kinematic PPP will produce the more precise time series, whereas relative processing should be more precise up to several hundreds of kilometers. The baseline length cutoff where kinematic PPP is preferred is not clear, however, and will certainly depend on base station data quality among other factors.

The relative heights obtained using TTC and Track contain fluctuations from the true value at many time scales. The relative height errors in Fig. 6(a) are representative of the variability for baselines of any length (although the magnitude will vary with baseline length; Fig. 5). In TTC-like solutions, where residual ZTD is not estimated, the temporal characteristics of height errors are comparable to those of the Track height errors, with occasional excursions from the true value lasting hours and longer. It should be noted, therefore, that many airborne LIDAR surveys cover the region of interest in much shorter periods of time—perhaps several tens of minutes. At these time scales, other error sources may dominate (e.g., GPS multipath), and any ZTD-related error in TTC-like solutions may appear like a near-constant bias. In these cases, and only over short (<10–20 km) baselines, it may prove to be more precise and accurate to remove the ZTD-related bias over a discrete region using contemporaneous ground-truth values (in the same reference frame as the aircraft positions), for example, using terrestrial laser scanning or kinematic GPS. This approach is currently widely adopted in the LIDAR community but has
the severe limitation of needing ground-truth data for every single survey; improved GPS positioning would clearly be preferential.

VI. CONCLUSION

While the U.K. data set has revealed that failure to parameterize ZTD will lead to seriously degraded GPS positioning over longer baselines (probably > 10–20 km), the Alps data set highlights the added complication of large height differences. A truer representation of a typical airborne LIDAR survey would have the height differences of the Alps network over < 20–30 km, although long-baseline surveys are also common and, indeed, logistically highly desirable. For short baselines, the GPS component of the LIDAR error budget is likely \( \sim 0.15 \) m, although because of the temporal correlation of these errors, individual surveys will occasionally have smaller errors. This value will be substantially affected by unmodeled relative ZTD, something that may be largely specific to the local climate, plus carrier-phase multipath and other smaller errors. Biases that vary from day to day and season to season may therefore be present in some LIDAR data sets, and these fluctuations may be erroneously interpreted as real elevation change. For longer baselines with no ZTD parameterization, the error budget is \( \sim 0.25 \) m. With high-quality data and careful data processing, this value may be extended, or even improved, up to baselines of several thousands of kilometers.

The precision of GPS solutions remains a major limiting factor in airborne LIDAR surveys, although the results also apply to GPS-controlled stereo photography or airborne gravimetry [44]. Of the various GPS-related limitations, accurately and precisely dealing with residual ZTD remains the most challenging, particularly since it is highly correlated with the heights of interest. Without ZTD parameterization, LIDAR surveys of the same feature on different days or months could yield height differences of \( \sim 0.1–0.3 \) m, simply due to unmodeled tropospheric variations. The size of the unmodeled troposphere in the Alps and U.K. is not as large as in some other locations, and hence, these biases may be understating the problem in these cases. Even with ZTD estimation, GPS heighting remains a limiting factor in airborne LIDAR surveys, and more sophisticated approaches and improved models must be developed for ZTD (see, e.g., [25]) and multipath; more observations from Galileo and/or GLONASS may also assist.

Finally, the drawback of the extra sophistication of software like Track is that it requires significantly greater levels of geodetic expertise, and this should be kept in mind when aiming for the highest precision and accuracy in LIDAR surveys.

ACKNOWLEDGMENT

The author would like to thank the MIT and JPL for making their software packages available. The work benefited greatly by the visit of O. Colombo to Newcastle in 2005. The author would also like to thank A. Luckman, N. Barrand, T. James, and T. Murray of Swansea University; A. Fox of British Antarctic Survey; I. Solovjanova of Russian Academy of Sciences; A. Adamek of the University of Silesia; and T. Abrahamsen of Store Norske for their assistance in collecting the Svalbard GPS data sets, and P. Clarke and the three anonymous reviewers, associate editor, and editor for their helpful comments on this paper. The GPS data used in this study were supplied by the NERC (U.K.) British Isles GPS Facility (BIGF), EUREF, and International GNSS Service archives.

REFERENCES


