Observation of a local gravity potential isosurface by airborne lidar of Lake Balaton, Hungary

A. Zlinszky1,2, G. Timár3, R. Weber1, B. Székely3,4, C. Briese1,5, C. Ressl1, and N. Pfeifer1

1Vienna University of Technology, Department of Geodesy and Geoinformation; Gußhausstraße 27–29, 1040 Vienna, Austria
2Balaton Limnological Institute, Centre for Ecological Research, Hungarian Academy of Sciences; Klebelsberg Kuno út 3, 8237 Tihany, Hungary
3Eötvös Loránd University, Institute of Geography and Earth Science, Department of Geophysics and Space Science; Pázmány Péter Sétány 1/C, 1117 Budapest, Hungary
4Interdisziplinäres Ökologisches Zentrum, TU Bergakademie Freiberg, Leipziger Str. 29, 09599 Freiberg, Germany
5Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology; Hohe Warte 38, 1190 Vienna, Austria

Correspondence to: A. Zlinszky (zlinszky.andras@okologia.mta.hu)

Received: 3 December 2013 – Published in Solid Earth Discuss.: 14 January 2014
Revised: 7 April 2014 – Accepted: 9 April 2014 – Published: 22 May 2014

Abstract. Airborne lidar is a remote sensing method commonly used for mapping surface topography in high resolution. A water surface in hydrostatic equilibrium theoretically represents a gravity potential isosurface. Here we compare lidar-based ellipsoidal water surface height measurements all around the shore of a major lake with a local high-resolution quasi-geoid model. The ellipsoidal heights of the 87 km² we sampled all around the shore of the 597 km² lake surface vary by 0.8 m and strong spatial correlation with the quasi-geoid undulation was calculated ($R^2 = 0.91$). After subtraction of the local geoid undulation from the measured ellipsoidal water surface heights, their variation was considerably reduced. Based on a network of water gauge measurements, dynamic water surface heights were also successfully corrected for. This demonstrates that the water surface heights of the lake were truly determined by the local gravity potential. We conclude that both the level of hydrostatic equilibrium of the lake and the accuracy of airborne lidar were sufficient for identifying the spatial variations of gravity potential.

1 Introduction

The aim of physical geodesy is the determination of level surfaces of the Earth’s gravity field (Hoffmann-Wellenhof and Moritz, 2005). Lakes are in theory affected by the variations of gravity, and the surface of any liquid at rest is part of a surface of equal gravitational potential (Brettenbauer and Weber, 2003; Merriman, 1881; Gomez et al., 2013). Variations in the ellipsoidal height of the standing water surface are therefore expected to correlate closely with variations in geoid undulation. Based on this assumption, mean water levels of lakes have been surveyed with GPS floats (Del Cogliano et al., 2007), water-level gauges and satellite altimetry (Cheng et al., 2008; Kingdon et al., 2008) in order to refine local geoid models. River water levels measured by satellite altimetry have been used as a reference for leveling gauge stations (Calmant et al., 2008). Low spatial resolution has always been a difficulty of such studies: radar satellite altimetry involves footprint sizes between 2–10 km (Connor et al., 2009). Satellite laser altimetry offers footprint sizes in range of 50–100 m, but track spacing remains in the range of several tens of kilometers (for temperate latitudes) which still does not allow for high-resolution mapping (Baghdadi et al., 2011). The limited range of geoid undulation values encountered within most study areas, together with the low spatial resolution of geoid models has also constrained assessment of the correlation between water surface height and geoid undulation, together with the effect of water currents and density differences on surface topography (Hipkin, 2000). A notable exception is the study of Borsa et al. (2008a), who survey a dry salt lake by high-resolution GPS topography mapping, and compare the results to local gravity measurements to prove that the salt flat resembles an equipotential
surface. While a vehicle-mounted GPS survey over the hard and quasi-stationary surface of a salt flat delivered height accuracies within 2.2 cm (Borsa et al., 2008b), a similar survey on a water surface encountered uncertainties in the range of 10–14 cm (Bouin et al., 2009) due to the superposition of waves and dynamic water surface height on patterns of geoid undulation.

We surveyed a lake where the surrounding gravity variations are well understood and have a wide range (> 1 m quasi-geoid height range), using airborne lidar for high-resolution mapping of the lake surface height. Airborne lidar is commonly used for mapping terrain topography (Wehr and Lohr, 1999), and stationary laser altimetry has been used for time series measurements of water surface height for radar calibration (Washburn et al., 2011). Airborne lidar has also been proven to deliver height measurements comparable with satellite altimetry over sea ice and open water (Connor et al., 2009). Recently, fine-scale height differences caused by internal standing waves in a coastal sea were also mapped by airborne lidar, proving that its accuracy and resolution are suitable for such studies (Magalhaes et al., 2013).

The geoid undulation in western Hungary increases near the axis of the Transdanubian Range, a series of hills of NE–SW orientation with elevations 600–700 m above sea level. Lake Balaton is located on the southeastern flank of this ridge; therefore the geoid undulation of its immediate neighborhood increases toward the northwest. This trend is explained by the isostatic unbalance of the Transdanubian Range which is a region of ongoing crustal uplift (Fodor et al., 2005; Timár et al., 2005; Sikhegyi, 2002). The center of this process is the axis of the hill chain. Repeated precise leveling has indicated maximum uplift rates of 1 mm year\(^{-1}\), with values between 0–0.2 mm year\(^{-1}\) in the forelands (Joó, 1992). Other methods have suggested slightly lower uplift rates, without disputing the general trend of uplift along the hill chain axis (Szanyi et al., 2009).

2 Objectives and hypotheses

We investigated two questions and hypotheses through this case study:

(i) To what extent does the lake resemble a level surface? Our hypothesis is that while many effects known to physical limnology influence the local height of a lake surface, observations can be timed to minimize these and the remaining effects are negligible or can be corrected.

### 1.1 Study area

We surveyed Lake Balaton in western Hungary, which is a shallow (average depth 3.3 m), large (597 km\(^2\)) and elongated lake of neotectonic origin (Sikhegyi, 2002) (Fig. 1).
(ii) Can lidar measure water surface elevation accurately enough for inferring variations in geoid undulation? Our hypothesis is that through strip adjustment based on terrestrial target surfaces, and correction of dynamic water surface height effects, the accuracy of the measurements can be sufficient to deliver data comparable with a geoid model and suitable for inclusion in the geoid modeling process.

3 Method

Lidar (also known as Airborne Laser Scanning, ALS) is a commonly used remote sensing technique capable of rapidly surveying a large number of points with elevations and horizontal positions accurate to a few centimeters (Wehr and Lohr, 1999). Lidar data points are collected with direct georeferencing, i.e., position and attitude of the scanning system are determined by GNSS (Global Navigation Satellite System) once every second and INS (Inertial Navigation System) data are the basis for the interpolation within this time interval, separately and independently for each laser pulse. GNSS allows determining the scanner position in geocentric Cartesian coordinates within the reference frame of the base station. These Cartesian coordinates are first converted to ellipsoidal latitude, longitude, and height (Hoffmann-Wellenhof and Moritz, 2005) and then projected to plane grid coordinates and ellipsoidal height. This process does not involve a geoid model, nor is it affected by gravity variations.

3.1 Airborne data and processing

We measured elevation of the lake water surface in 58 flight strips along the whole shoreline (Fig. 2a, b). Since the main goal of the survey was littoral vegetation mapping (Zlinszky et al., 2012), the flight pattern was optimized for continuous coverage of the coast and the coastal water surface. A Leica ALS50 laser scanner operating at 1064 nm (Zlinszky et al., 2012, 2011) was used with a strip-wise nominal ground point density of 1 point m\(^{-2}\) at the average flying height of 1400 m above ground level. The strip width varied between 600 and 1200 m and neighboring strips had an average overlap of 15%. The data were delivered in the global geodetic datum in UTM projection and this coordinate system was used throughout the study.

We used lidar strip adjustment (Filin, 2003; Kager, 2004; Ressl et al., 2011; Skaloud and Lichti, 2007) to improve the relative georeferencing of the flight strips by optimizing their relative alignment with respect to each other. In the first step, inclined planar surfaces (typically building roofs) were extracted automatically from the lidar points in each strip. During an iterative process, misalignment, lever arm, offset, scale, deflection angle, and individual global shifts of each strip were estimated in order to minimize the differences between corresponding planes in the overlapping parts of each pair of strips. The adjustment delivered optimal values for these parameters which we then used to transform all the lidar points of each strip to new locations. The shifts were determined independently from eventual water-level variations as they were calculated only from selected shore features. Typical shifts were less than 20 cm.

The OPALS software package (Pfeifer et al., 2014; Mandlburger et al., 2009) was used for interpolation and processing. Moving least-squares interpolation with a plane model was applied for creating a raster elevation model of the water surface with 1 m \(\times\) 1 m raster resolution. We used the lake outline and a vegetation map generated from the ALS data (Zlinszky et al., 2012) to remove the non-water areas. Visual quality control showed that cells with ellipsoidal heights lower than 149.5 m or higher than 152 m are mostly artefacts, mainly high points from vegetation, boats, etc. Therefore, cells with ellipsoidal elevations outside these limits were also excluded from further study (1 % of the lidar raster cells). As the height distribution contained some outliers (remaining non-water points), standard deviation would not have yielded a representative value. Therefore \(\sigma_{\text{MAD}}\) was used as robust estimator of the standard deviation. It is defined as \(\sigma_{\text{MAD}} = 1.426 \times \text{MAD}\), where MAD is the median of the absolute deviation to the median. For a normal distribution \(\sigma_{\text{MAD}}\) is equal to the standard deviation.

3.2 Correction of dynamic water-level effects

In order to correct for short-term local water-level variations of the lake during the survey, measurements of the water gauge network around the lake were investigated. These were collected at 9 stations (Fig. 1) in 15 m intervals, using floats connected to digital recorders (3 stations) or pressure sensors (6 stations). In order to remain independent from the datum surface of the water gauges (which is a geoid model itself), the local mean lake level (LMLL) during the 4 flight days was calculated separately for each gauge. A time series of water-level variations compared to the LMLL was calculated for each measuring station. This series was compared with GPS time recordings of each flight strip, and the ellipsoidal height of the elevation model pixels within each flight strip corrected with the difference between LMLL and local water level, measured exactly at the place and time of the strip. Correction values were between −6 and +2.5 cm, with a median of +0.16 cm and a \(\sigma_{\text{MAD}}\) of 1.31 cm. As a result, a water surface model was produced, with approximately 87 million data points characterized by horizontal coordinates and adjusted elevations above the WGS 84 ellipsoid.

3.3 Comparison with quasi-geoid model

As a basis for comparison with the local gravity potential isosurface, the HGE2000 Hungarian quasi-geoid model (Kenyeres, 1999; Kenyeres, 2000; Völgyesi et al., 2005) and the HGTUB2007 Hungarian quasi-geoid model (Tóth,
Figure 2. (a) Map of lidar-measured ellipsoidal water surface heights (inside flight strip outlines) and modeled gravity potential isosurface height (background raster) with the same elevation color scheme. Inset shows detailed height distribution of water surface measurements, histogram refers to lidar measurements (not to gravity potential isosurface model). (b) Map of lidar-derived normal water surface heights, obtained by correction with local deviation of LMLL and local quasi-geoid height. Histogram refers to water surface normal heights.
2009a, b) were considered, and HGTUB2007 was chosen due to its finer spatial resolution. We are aware that the quasi-geoid is not necessarily an equipotential surface, however, the Hungarian National Height System uses normal heights, and normal heights within Hungary deviate max. 2.8 cm from orthometric heights (Ádám, 1999), which is within lidar measurement accuracy. The calculation of HGTUB2007 is described in full detail in Tóth (2009b). Tóth (2009b) calculated the quasi-geoid from the following data sets: 6678 mean free-air gravity anomalies in 2′ × 3′ blocks based on more than 300,000 point gravity data; 276 vertical deflection components based on 138 astro-geodetic vertical deflections (both north and east components); 7452 surface gravity gradients (torsion balance stations) resampled from 27,005 measurement points; and 94 GPS leveling measurements of the Hungarian National GPS network (OGSH). The GPM98CR geopotential model was combined with the GRACE GGM02C model to the maximum degree and order 720 and used for reduction of the observations. RTM corrections were calculated based on the fixed mass model of SRTM3 heights. The residual gravity field of all observations was interpolated by least-squares collocation with the self-consistent logarithmic covariance model of Forsberg (1987), to a grid of 1.5′ × 1′ resolution. The estimated prediction errors of the model are below 2 cm inside Hungary (Tóth, 2009b), which also applies to the areas near the shore that we surveyed with lidar.

For comparison with the quasi-geoid model, the lidar-derived (and water-level corrected) water surface model had to be resampled to the same spatial resolution. In order to ensure that the remaining elevated non-water points in the data do not distort the heights, the 30th percentile of the water surface model cell heights within each cell of the quasi-geoid model was calculated and used for representing the water surface heights. As a further correction, cells that did not represent the water surface height because they were mainly over shore or wetland surfaces were removed, excluding 16 of the originally 207 data points, removing the imperfections of land and vegetation masking. The correlation of this low-resolution water surface height model with the quasi-geoid was tested by linear regression.

4 Results

The 87,077,358 water surface model raster cells had an ellipsoidal elevation range of 80 cm (Fig. 2a). The largest ellipsoidal elevations of the water surface are in the northwestern basin of the lake, decreasing gradually toward the southern shore with the lowest areas in the eastern corner (see supplementary data for high-resolution map). The ellipsoidal heights of the water surface in overlapping areas of strips surveyed on different days were similar within the accuracy range of the instrument (< 8 cm, Leica Geosystems AG, 2006). Water levels measured during the flight window showed some variation (explained in detail in Appendix A), with a total range of 20 cm for all stations. Water-level deviations from LMLL measured synchronously with the flight strips have a total range of ±6.0 cm for the 4 flight days (Fig. 4).

We created a simple gravity potential isosurface model by adding a constant (105.1 m, mean height of the lake surface above sea level) to the quasi-geoid height raster, considering the difference between orthometric and normal heights negligible for our study. Comparing this gravity potential isosurface model throughout the area of the lake with the measured ellipsoidal water surface elevations shows close correlations (Fig. 2a, see also supplementary material). The geoid undulation difference between the shores of the lake corresponds to the water surface elevation pattern in flight strips both parallel and perpendicular to the shore. The ellipsoidal
height range of the water surface heights was approximately 80 cm, with a dispersion (quantified by σMAD) of 20 cm (Fig. 2a). The water surface heights resampled to the quasi-geoid height of each cell: a linear regression with an $R^2$ value of 0.906 was calculated with a slope of 1.12, intercept of 99.95 m and σMAD of 5.17 cm. This agrees with our expectation that the pattern of water surface ellipsoidal heights is explained by the height variations of the gravity potential isosurface.

When the local quasi-geoid heights were resampled by bilinear interpolation to the spatial resolution of the water surface model ($1 \times 1$ m) and subtracted from the local water surface model elevations, the σMAD was reduced to 5.6 cm, and the elevation range of the water points to 30 cm. 78% of the points were within 15 cm and 36.1% of the 87 077 358 points within 5 cm (Fig. 2b). Hardly any flight strips show along-track differences, not even those spanning 15 cm of quasi-geoid height range along their length (across-track differences are mainly artefacts discussed in Appendix B). The geoid-corrected water surface heights are slightly lower than average in the SE corner of the lake and higher than average in the NW, a pattern that suggests that the height gradient of the lake may be even steeper than the gradient in the quasi-geoid model.

This is confirmed by the scatterplot of the measured ellipsoidal water surface heights and the corresponding geoid undulation (Fig. 3). A clear linear trend is visible, and the slope is very close to 1 (1.06), which agrees with our expectation that the geoid undulation controls water surface height variation in space. Nevertheless, water surface heights show a steeper profile than the geoid model, and fall clearly below the 1 : 1 line for the lowest quasi-geoid height values. Erroneously high or low artefact points are probably created by insufficiently removed non-water features and by waves.

5 Error budget

5.1 Relative georeferencing

For the lidar data of Lake Balaton no geometric ground control features were available, therefore the absolute georeferencing accuracy is a product of standard differential GNSS georeferencing (10 cm according to the flight operator), enhanced by the large number of measurements. However, the relative georeferencing of the strips (i.e., their mutual alignment) was improved by strip adjustment, as documented by σMAD of the differences in smooth areas of overlapping strip pairs. For all considered 305 pairwise strip differences (with around 106 000 000 difference values in total) σMAD improved from 12 cm (before) to 5 cm (after strip adjustment).

5.2 Dynamic water-level correction

Single water gauge measurements can be considered accurate within 1 cm, but the LMLL as a datum surface is slightly less accurate. Calculated as an average of the local water levels measured every 15 min during the 4 flight days, the errors of LMLL can be estimated from the range of local mean water levels calculated separately for each measurement day and each gauge. This is below 2 cm for all stations, and below 1 cm for 6 out of 9 stations.

A possible source of error in the height measurements was water as the target surface for lidar. Water is not a strong reflector at the wavelength used (1064 nm). The nature of reflection from the surface or penetration into the water column depends strongly on the look angle, therefore from different parts of the strips either no points were recorded, accurate measurements were made or (rarely) erroneously long or short ranges were measured. A detailed overview of the artefacts produced by water as a target surface and how this affected our measurements is provided in Appendix B.

While over oceans, dynamic water surface topography is known to include deviations from the gravity-driven equilibrium of more than 1 m (Seeber, 2003), this is much less prominent in a lake because of the shallow depth (3.3 m on

Figure 4. Water-level recordings with respect to LMLL of all 9 water-level gauges around Lake Balaton, during the 4 flight days. Triangles in the map depict the water gauges (colour coded to the water level graphs), rectangles crossing all graphs show actual flight times of lidar strips.
Table 1. Error budget, including standard deviation and median of each individual error source and the theoretical error propagation, compared with the dispersion found in the measurement data. Lines indicated in bold influence the final error budget. Lines not in bold show intermediate calculation steps, they only apply to the theoretical case without strip adjustment (Absolute point vertical accuracy) or to the individual strips affected by water artefacts. The total impact of water artefacts was calculated by averaging these values across the total data set.

<table>
<thead>
<tr>
<th>Estimated error budget, individual height sources in cm</th>
<th>Standard deviation</th>
<th>Median</th>
<th>Spatial distribution of error</th>
<th>Source of value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airborne lidar system</td>
<td>Absolute point vertical accuracy</td>
<td>8</td>
<td>0</td>
<td>systematic point error within each strip</td>
</tr>
<tr>
<td></td>
<td>Point accuracy after strip adjustment</td>
<td>5</td>
<td>0</td>
<td>mainly random except for strips with georeferencing artefacts</td>
</tr>
<tr>
<td>Water surface height effects</td>
<td>Waves (in 40 % of the strips, no waves in the rest)</td>
<td>9</td>
<td>0</td>
<td>periodic systematic</td>
</tr>
<tr>
<td></td>
<td>Total impact of waves on full data set</td>
<td>4</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Specular reflection (influencing ca. 10 % strip area in 30 % of the strips surveyed)</td>
<td>5</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total impact of specular reflection on full data set</td>
<td>0.15</td>
<td>0.45</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Smile artefact (influencing ca. 20 % strip area in 30 % of the strips)</td>
<td>2.5</td>
<td>−7</td>
<td>systematic</td>
</tr>
<tr>
<td></td>
<td>Total impact of smile artefact on full data set</td>
<td>0.15</td>
<td>−0.35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dynamic water topography</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimated error budget, total effect of height error sources on data [cm]</th>
<th>Standard deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total impact of water as target surface</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Total estimated height error budget</td>
<td>7.1</td>
<td>0</td>
</tr>
<tr>
<td>True total error budget (from measurement data, Fig 2b)</td>
<td>5.6</td>
<td>−2.2</td>
</tr>
</tbody>
</table>

average) and low current intensity at the time of flight. Nevertheless, a number of processes with proven effects in marine settings which might also potentially influence the height of the water surface in our case are discussed below:

5.3 Seiche and setup

Storms can produce disequilibrium of the lake surface, increasing the water level on the downwind side while decreasing it upwind (setup) or producing dynamic standing waves along or across the lake, with wavelengths of several kilometers (seiche). In case of Lake Balaton, this starts at sustained wind speeds above 5 m s\(^{-1}\) and such a displacement can reach 1 m in water level during storms winds of 20 m s\(^{-1}\) (Muszkalay, 1973; Somlyódy, 1983). According to local Meteorological Aerodrome Report (METAR) data, wind speeds never exceeded 5 m s\(^{-1}\) during the whole survey. Some setup and seiche effects can be observed in the water-level data (see Appendix A and Fig. 4), but these were corrected for as described in the methodology. The remaining uncertainty of the dynamic water topography correction is a consequence of using the difference of LMLL of the nearest gauge for each strip. In the worst possible case, this might produce an error in range of the difference in deviation of LMLL between two neighboring gauges. These differences had a standard deviation of 3 cm for the 4 flight days (Table 1).

5.4 Currents

On Lake Balaton, currents are driven by wind forcing and by seiche while thermal convection remains hardly observable (Muszkalay, 1973). The flow of water from the tributaries towards the outlet is also very weak, because significantly more water leaves the lake through surface evaporation than through the outflow. The patterns observed in ellipsoidal height do not match the main tributaries or the outlet of the lake, ruling out river plumes. Current flow can locally reach 1 m s\(^{-1}\) in cases of storms or strong seiche (Virág, 1998). While no direct current measurements have been carried out during our survey, the wind speeds were low (< 5 m s\(^{-1}\)) as documented by the METAR reports, and seiche displacements were within a few centimeters as detailed in Appendix A. Therefore we have a reason to believe currents were negligible.

5.5 Waves

Wind-induced surface waves truly affect local water-level height, at observation scales smaller than the wavelength. The wave lengths we observed on the lake are around 6 m or less, the vertical standard deviation of wave artefacts was 5 cm (Table 1), rare extreme wave heights encountered in the data were 30 cm from trough to crest. We believe this to be the most important source of height dispersion, however, it
does not contribute to the lake-scale pattern of ellipsoidal height that we derive our conclusions from, as wind-induced waves involve a periodic motion of water with equal amplitudes below and above the mean water level (Imboden, 2003).

5.6 Air pressure response

The rule of thumb for air pressure response is 1 cm decrease in water level for 1 mbar increase in air pressure (Ponte and Gaspar, 1999). However, this only applies to the open ocean and after a period of 1–2 weeks. In our case, since the flights took place within 10 days and since the lake was very shallow, most of the air pressure is directly forwarded to the lake bottom. Therefore we assume that air pressure changes could not have influenced the local lake level by more than a few millimeters during the measurement.

5.7 Lake Tides

Lake tides are known to have amplitudes of up to 10 cm in larger lakes (Trebitz, 2006). In case of Lake Balaton, the shallow depth and the relatively low water volume of the lake suggest this effect would be even smaller. During long-term investigation of water movements conducted in the 1970’s, no evidence of lake tides was observed (Muszkalay, 1973), therefore we do not expect these to have influenced lake level.

5.8 Density change

Due to its fresh water, shallow depth (3.3 m avg), limited inflow and small water volume (2 km$^3$) compared to its surface (600 km$^2$), Lake Balaton experiences very limited variations in water density. There is no thermocline, warming and cooling can only create maximum vertical temperature gradients of up to 1.1 $^\circ$C m$^{-2}$, and horizontal differences in surface temperature between basins have been documented to remain less than 3 $^\circ$C (Virág, 1998). The salinity of the water is also quite constant over its area and over time, so no major changes in density are expected. The height effect of local warming or cooling can dissipate relatively quickly on this scale, therefore no contribution of density change to lake topography can be expected.

5.9 Summary of error budget

The most important errors influencing our data are summarized in Table 1. In case of the single lake surface height cells in the raw data, the most important source of height error is the combination of GNSS positioning and lidar range measurement, causing errors around 8 cm (Leica Geosystems AG, 2006). With strip adjustment this quality can be assured and even improved, as in our case the differences between overlapping surface model heights (on land) showed a standard deviation of 5 cm. This is the sensor-based uncertainty in each of our water surface model cells (the water surface model is defined in Sect. 3.2).

The effects of water as a target surface are discussed in detail in Appendix B. Each lidar strip was inspected for such effects, and the typical amplitude and the size of the affected area noted. Waves were only encountered in part of the strips, and their amplitudes were low. Specular reflection influenced a very narrow part of some strips at nadir, creating systematic positive shifts in these data cells, but these have a very low proportion relative to the full data set. “Smile” artefacts (discussed in Appendix B again) were mutually exclusive with waves and specular errors, and produced a systematic negative shift at some strip edges. Dynamic water topography is expected to have produced errors within 3 cm standard deviation.

The total theoretical impact of water as a target surface can be calculated by summing the error distributions and calculating the median and the standard deviation of the result. No systematic shift in the data is expected here, since the systematic errors affecting the median are rare and low. The 5 cm standard deviation is mainly caused by waves and dynamic water topography. The total estimated height error budget additionally includes the post-adjustment relative georeferencing error.

The distribution of our water surface model heights corrected with the local quasi-geoid height (which also has a 2 cm standard deviation) shows a smaller deviation than theoretically expected from the inspected error sources. The −2 cm bias in absolute height (compared to the 105.19 official water level measured at Siófok during the flight days) may be caused by remaining absolute inaccuracies of the GNSS positioning (which is only partly corrected by strip adjustment) and uncertainties of water-level measurements. In the future, for such surveys, measuring of ground control points is highly advisable in order to avoid the eventual absolute error of GNSS.

If the data is to be used for geoid modeling, the error budget could be improved by two factors: automatic removal of specular and smile effects (see Sect. 6.2), and interpolation of water surface height into larger cells, removing periodic and some random error.

6 Discussion and outlook

Comparison of water surface ellipsoidal heights with the quasi-geoid model shows that these correlate very closely, with 90.1 % of the variations in water surface height explained by the quasi-geoid height variations. As far as the resolution of the geoid model allowed, the close correlation of the two data systems confirmed that standing water has a truly level surface. Variations in the ellipsoidal height of the lake water surface are mainly a product of the variations in local gravity potential represented here by the quasi-geoid height; the slight water-level changes induced by movement...
of water during the flight period were corrected for. One of the limitations of investigating this correlation is the resolution of the quasi-geoid model we used: similar to satellite altimetry, the resolution of the water surface is much higher than the resolution of the geoid model, therefore limiting comparison (Cheng et al., 2008). The high resolution of the lidar-derived water surface model shows short-wavelength patterns in height, and therefore potentially in geoid undulation, which are beyond the scale of the quasi-geoid model we used, and indeed most geoid models available. This implies that water surface ellipsoidal heights measured by lidar might be used in the future to refine local gravity variation models.

6.1 Comparison with other measurement methods

The theory of hydrostatic equilibrium being connected over large areas to the local gravity field has already been utilized for surveying the geoid over the oceans (such as ICESat and Envisat), but these measurements are affected by dynamic ocean topography (Hoffmann-Wellenhof and Moritz, 2005). In our case, the dense network of water-level gauges and the high measurement frequency allowed quantification of dynamic water height effects and correction of any measurement bias. This was possible since only water surfaces close to the shore were surveyed and because the flight took place under calm weather conditions. Our lidar measurement accuracy of 5 cm calculated from flat terrestrial surfaces after strip adjustment compares to the 4 cm accuracy obtained by Borsa et al. (2008b) through vehicle-mounted GPS surveys (over much larger area), and to the best accuracies measured from ship-mounted GPS under calm weather conditions (10 cm) (Bouin et al., 2009), but are slightly worse than the 2.7 cm error reached with a GPS catamaran in a marine setting by Bonnefond et al. (2003). Borsa et al. (2008a) find that 93% of the ellipsoidal height variations of a dry salt flat are explained by variations in geoid height, and come to the conclusion that the surface topography closely approximates the gravity equipotential surface. At the scale of the most detailed quasi-geoid model we could obtain, we found that 90.1% of the water surface topography variation is explained by variations in quasi-geoid height.

Point elevation measurements from GNSS buoys were already used as part of a leveling-based geoid survey (Gomez et al., 2013), and satellite gravimetry-based geoid change measurements have been used as a proxy for lake water levels (Awange et al., 2008; Calmant et al., 2008). However, active satellite altimeters were only applied in rare cases for mapping inland lakes as level surfaces, probably mainly because of their limited spatial resolution (Hoffmann-Wellenhof and Moritz, 2005; Cheng et al., 2008; Baghdadi et al., 2011).

Comparing satellite altimetry with terrestrial, airborne and satellite gravimetry over the Great Lakes, Kingdon et al. (2008) observe that lake surface altimetry follows short-wavelength gravity anomalies confirmed by ship- or airborne gravimetry better than a lower resolution model based on GRACE satellite data does. They also describe the absolute accuracy of the altimetry-derived quasi-geoid to be closer to the GRACE values than the high-resolution ship- or airborne measurements, which were more prone to systematic bias. This means that altimetry was more accurate than terrestrial gravimetry but delivers higher resolutions than satellite gravimetry. If the same principles apply to lidar, which is also based on altimetry, it can be expected that lidar-derived lake surface heights would deliver a valuable input to high-resolution geoid models. Compared to satellite altimetry, lidar is characterized by smaller footprints, even higher spatial resolutions and the possibility to survey large areas within a few hours. Similarly to satellite altimetry over oceans, the spatial resolution and area coverage of lidar could theoretically allow identification of dynamic water topography features of lakes such as plumes or eddies. Since our data only covered the waters nearest to the shore, and since the lake was near equilibrium, this could not be tested in our case.

Compared to GPS buoys which collect height data spread over longer periods to assess mean lake level, lidar delivers repeated measurements spread in space. This has comparable accuracy to the technique of Bonnefond et al. (2003) who used GPS receivers mounted on a ship to obtain area-covering sea level height measurements, but is more productive due to higher survey speeds and area coverage. However, lidar surveys provide a snapshot and are thus more easily distorted by the effects of water movement, requiring correction based on local water stages. The method we proposed only needs relative height changes at each gauge.

Compared to analysis of simultaneous water-level gauge readings such as Cheng et al. (2008), the advantage of lidar in our case is the higher number of data points, which deliver statistically stronger results while using coverage of the shore for accurate relative georeferencing. It is expected that in the future, lidar coverage of lake surfaces will further increase with the spread of bathymetric lidar (Mandlburger et al., 2011).

The typical methods for assessing geoid undulations over land are gravimetry and the surveying of leveled elevation benchmarks with GNSS (Seeber, 2003). While these surveys can deliver higher vertical accuracy due to static GNSS measurements, a drawback of GNSS leveling is the limited number of possible measurement points confining spatial resolution and noise removal. In our case, the lake itself serves as a leveling instrument providing a vast area where elevations relative to the geoid are shown to be constant. In combination with the high number of measurement points from lidar and the improved accuracy of strip adjustment, a comparable signal-to-noise ratio can be reached.
6.2 Outlook: including water surface lidar heights in geoid modeling

The full process of creating a new local quasi-geoid model including water surface height data from lidar and the evaluation of this quasi-geoid map would be outside the scope of the current paper. However, we believe that in the future, lidar-derived lake surface data can be used in a way similar to GPS-leveling data: lidar provides an ellipsoidal height while local water gauges provide (in the Hungarian case) the normal height of the same object, the water surface. In practice, the following steps (beyond those already described in the methods section) are recommended: any data points where elevation can be considered uncertain from visual inspection should be excluded (surface glint, smile effects, twist); lidar data points should be aggregated to larger raster cells (e.g. $100 \times 100$ m, 10,000 points), which also further reduces error; the leveling of the water gauges should be checked and any biases corrected; and for each lidar-derived raster cell, the local water stage should be interpolated from timed water-level records and proximity to gauges. After these steps, the quasi-geoid height of each raster cell can be calculated from the difference between ellipsoidal (lidar) and normal (gauge) water surface height and can be loaded to the least-squares collocation for the geoid model together with other data types. Lidar mapping of lake surface elevations can deliver information on the ellipsoidal height pattern of the water surface, and thus on the local gravity anomalies. These in turn can be used to collect information about the formation of the lake (Dietrich et al., 2013). Lidar surveying of lakes can be valuable for estimating the error budget of lower resolution regional geoid models and GPS-derived heights based on such models. Since many major European and North American lake shores have already been surveyed by lidar, there is a wealth of data available.

7 Conclusions

We conclude that a lake may reach a level of hydrostatic equilibrium that allows it to be used as a basis for mapping a gravity potential isosurface. This can be assessed on the basis of water gauge readings, which can also be used to correct the remaining lake-scale dynamic surface topography. The ellipsoidal height of the lake surface has been proven to closely follow the local geoid undulation, within the limits of instrument accuracy and quasi-geoid model resolution.

Our study also demonstrates that lidar measures water surface height accurately enough for inferring variations in local gravity potential. The strip adjustment procedure enhanced the georeferencing accuracy of the data, and the artefacts produced by the water surface only affected a small part of the data points collected. The high resolution of lidar data shows short-wavelength patterns in water surface height which are beyond the scale of the current geoid models. Airborne lidar of lake surfaces may therefore be introduced as a valuable new technique which should be further studied in support of physical geodesy.
Appendix A: Water levels of Lake Balaton during the studied period

Investigation of the water levels of the lake will be used to answer two questions:

(i) To what extent was the lake at hydrostatic equilibrium during the airborne survey?

(ii) How did variations in local lake water levels affect the ellipsoidal heights measured by lidar?

Hydrostatic equilibrium would mean that the lake was free from dynamic processes such as local currents, eddies and standing waves, and therefore the rules of hydrostasis would govern the pattern of its surface elevation. Whether the lake was in such a state can be determined by local water-level variations. The common datum of water gauges around the lake was determined by a geoid model, therefore it is subject to errors of the model and might not perfectly represent the true long-term surface of the lake in theoretical equilibrium. Therefore, we studied water-level changes independently for each gauging station: LMLLs for each station are expected to define this equilibrium across long-term measurements. During the four flight days, the daily LMLLs showed increases or decreases of a few centimeters, consequently it was assumed that slight changes of lake water volume were affecting these measurements (Fig. 4). In this case, measuring LMLL across longer periods of time might not necessarily have increased the accuracy of LMLL estimation.

While wind and currents of lakes are known to induce static and also periodic deviations from the equilibrium water level (setup and seiche), the wind speeds observed during the campaign were below the known threshold for such processes (Muszkalay, 1973). The flight data was collected during 21, 22, 23 and 26 August 2010. August 24 and 25 were left out because of cloudy weather. During Day 1, the highest water level, 122 cm above the station datum (8 cm above LMLL) was observed in Keszthely, where the water level slowly fluctuated between 112 and 122 cm with a period of ca. 5 h. While this periodic fluctuation with decreasing amplitude implies a standing wave (seiche), no water gauge showed the same movement in the opposite direction. It is already known that the Keszthely gauge typically records the largest dynamic water-level variations, since the small western basin of the lake is the most sensitive to the wind. This fluctuation was most probably caused by remnants of seiche from winds during the previous day (Fig. 4).

The lowest water level for Day 1, 103 cm (9 cm below LMLL) was observed in Tihanyrév in the Tihany straits and is an isolated low record both preceded and followed by continuous recordings around 111 cm (0.7 cm below LMLL). This pattern matches some measurements described by Muszkalay (1973) as a short-term standing wave formed within a harbor basin, which might well be the case since Tihanyrév has intensive ferryboat traffic. Neither of these deviations corresponded to the flight times of Day 1; in fact, during the afternoon flight the water level was within 2 cm of LMLL for all stations near the flight area of the day (Fig. 4).

During Day 2, the highest recorded water-level deviation was 5 cm above the LMLL, again an isolated measurement of 117 cm in Tihanyrév preceded and followed by values around 112 cm. The lowest level was 1.9 cm below LMLL, 110 cm measured at Siófok, preceded and followed by readings of 111 cm. During the time of the flight on that day, the deviation from LMLL was between −1.8 and +2.3 cm for the stations covered by the flight area (Fig. 4).

During Day 3, the highest differences of LMLL of +3.3 cm were measured over the course of an hour in the evening in Balatonfüzfő, not counting an isolated reading of +4.2 cm in the Tihanyrév harbour. The largest negative difference of LMLL was −2.9 cm in Siófok, sustained during several hours at night. The minimum of −2.7 cm compared to LMLL (111 cm above gauge datum) was continuously measured during the morning hours of the day in Keszthely. During the flight, deviations from LMLL were between −1.8 and +3.3 cm (Fig. 4).

During Day 4, the largest absolute differences with respect to LMLL were +10.3 and −7.7 cm, both observed in Keszthely. Here the water level rose from 114 cm (above the gauge datum) to 124 cm over 2 h and fell back to 112 over the next two hours. The largest negative deviation (−7.7 cm) happened during the flight in Keszthely, the largest positive difference in LMLL was +2.1 cm in Siófok (Fig. 4).

Apparently, even under calm conditions, standing waves and other dynamic processes can cause up to 10 cm differences compared to the average water level. The strongest fluctuations were observed in Keszthely and Balatonfüzfő, which are both near the ends of the lake and in narrow corners capable of producing water level changes higher than on the open lake surface due to a funnel effect. If this is the case and the lake level changes amplified by the funnel effect reach no more than 10 cm compared to the average water level, we can only expect minor dynamic lake topography along the more open shores. We thus conclude that while the lake was not at full hydrostatic equilibrium during the flight days, dynamic local water level changes were within ±5 cm for most measurement sites and within ±10 cm globally. Due to high frequency water-level measurements and some luck in the afternoon timing of most flights, the dynamic water-level variations actually affecting our measurements were even lower, typically within ±3 cm, in one isolated case reaching −7.7 cm. These variations were relatively well understood and corrected for by measuring deviations from LMLL at 9 stations with 15 min frequency.

Appendix B: Possible sensor artefacts due to water as a target surface

If the target surface of lidar is water, part of the incoming laser light energy passes through the surface, enters the water...
column and is absorbed, since water is a weak reflector at the near-infrared wavelength we used (1064 nm). Some of the pulse energy can nevertheless be scattered back to the sensor from below the water surface. The rest of the light is reflected from the surface, part of it with near-Lambertian characteristics, part of it specularly. In case of specular reflection, very high amounts of radiation are reflected back into the sensor at nadir, while off-nadir, the energy is reflected away from the sensor and none of it is received. In case of waves, every crest may have a surface inclined to produce a specular reflection towards the sensor. However, given a constant wave amplitude, chances for a reflection towards the sensor will decrease with increasing look angle.

The proportions of light absorbed, volume scattered, specularly reflected and scattered from the water surface depend on sensor wavelength, the optical characteristics of the water, the roughness and inclination of the water surface and the look angle of the outgoing pulse, discussed in detail in Guenther et al. (2000).

In case of very flat water, the pulse energy is reflected away from the sensor for most of the strip, with an insufficient return for triggering the detector and therefore no recorded point. At nadir and again over calm water, the pulse energy reflected specularly arrives back in the sensor, producing a point and also often a glint effect with very high echo intensity. This is known to result in slightly shorter range measurements (10–50 cm), as the high amount of incoming energy causes the system to detect a peak too early (range-walk effect). Obvious glint range-walk effects were encountered in less than 1% of the data (see histogram in Fig 2b), since for most of the areas where the water surface was completely flat, no points were registered as described above.

For most of the flight campaign, moderate waves were encountered. These were sufficient to produce non-specular reflection over most of the covered strip surface, but no data points were produced closer to the edges of the strips due to a combination of water absorbing the pulse energy and semi-specular reflection away from the sensor. The width of this “blank” edge again varies depending on local water surface characteristics. Closer to the strip centerline, but still nearer to the edge of the strip, slight decreases of the heights from the strip center to the strip border by about 5 cm (Fig. 2a, supplementary material) can be observed in some cases. This small systematic smile error may have two causes:

(i) The most likely reason for this observed smile error may lie in the interaction of the laser signal with the water surface as explained above. That is, perhaps in this part of the affected strips the pulse energy entered the water, and volume scattering of water contributed to the reflected energy besides surface reflection, and thus produced a measurement of erroneously long range, i.e., low water surface. The width of this “too-low” part of the strip is usually not more than 100 m.

(ii) With increasing look angle the portion of specular reflected energy away from the sensor increases, thus ever less energy is detected by the sensor. Since very high amounts of incoming energy (in the nadir) cause the system to detect a peak too early (resulting in too-short ranges), it seems logical that with decreasing amounts of incoming energy the system will detect a peak too late (resulting in too-long ranges; i.e., low water surface). The width of this too-low part of the strip is usually not more than 100 m. Having systematically inspected each flown strip for the “empty edge”, smile, specular center and wave effect, the following conclusions were made:

The width of the strip edges where no lidar echo from water was observed was typically 200 m on either side of the strip, and all lidar strips collected were affected by this. This varied by about 50 m depending on surface conditions, with narrower empty strips where waves were encountered, and exceptionally very broad empty strips where water was completely flat and specular reflection dominated. This suggests that the empty edge effect is caused by the dominance of specular reflection and absorbance above a certain angle of incidence between the pulse and the water surface.

Over completely flat water, in our case in the wind shadow of terrain (west of the Tihanyrév gauge and area around the Badacsony gauge), specular reflection dominated. Up to 15–20 cm overestimations of water surface height due to specular effects were observed in small parts of 20 of the 58 strips we used. All of these were in very calm water conditions. In some extreme cases (5 strips) the strip area would be completely empty except for the very narrow (5–20 m) centreline where points would have erroneous heights due to specular effects; implying that specular reflection dominated in these cases and data were only collected at nadir.

The artefact of erroneously low strip edges (smile) was encountered in 19 of the 58 strips, all of these in the eastern basins of the lake where the water is less turbid than in the west. Most strips are completely free from this artefact, suggesting no systematic internal errors of the scanning system were encountered. The affected strips were all on the northern shore in deep water and local wind shadow, again implying low turbidity. The smile effect typically occurred at scan angles near 20°, which follows our expectation that the entry of the laser pulse into the water column is limited to a certain angle of incidence (a fact well known in the bathymetric lidar domain).

Waves affected 23 of our 58 strips, nearly all collected during Day 3, some during Day 4 (Fig. 4). As expected, strips affected by waves were always free of smile and major specular effects. Wave amplitude was between 5 and 30 cm and wave length typically around 6 m between crests. Surface waves showed areas both above and below average water surface height.
A further, final cause of strip patterns may lie in remaining errors of the strip adjustment process. During surveying, a sensor problem was discovered to affect the data: the central position of the scanning mirror was determined with some uncertainty, resulting in a slight “twist” of the affected data strips along their longitudinal axis, with erroneously low cells on one edge and erroneously high on the other (+/− 5 cm). Strip adjustment corrected this effect as described in the text, but in some rare cases (definitely not more than 7 of our 58 strips), the low overlap with other strips means that some such error may have remained.

When all has been considered, these stripe artefacts seem to have a limited effect on the inferred lidar heights. As shown by the histogram of Fig. 2b, the distribution of the difference between ellipsoidal water heights and local quasi-geoid height remains rather narrow despite these. We believe a major cause for the local deviation from the mean is the effect of waves. The smile effect may be responsible for the slight skew of the distribution, with more low points than high, while specular effects fall mainly into the highest shown histogram bin (105.05 m) and have a very low proportion (0.2%). Twist effects also cause some broadening of the distribution, but their overall contribution is low since they were corrected successfully in the overwhelming majority of the strips.
The Supplement related to this article is available online at doi:10.5194/se-5-355-2014-supplement.

Acknowledgements. The airborne survey of the lake was sponsored by the European Community’s 7th Framework Programme (FP7/2008-2010) under EUFAR contract no. 2271; strip adjustment was funded by the GIONET FP7 project (PTTN-GA-2010-264509). A. Zlinszky and partly B. Székely and N. Pfeifer were funded by the Changehabbits2 FP7 project, C. Brieuse by the Ludwig Boltzmann Institute, B. Székely contributed as an Alexander von Humboldt Research Fellow. Contribution of G. Timár was partially supported by the Hungarian National Research Fund (OTKA) project NK83400. A. Zlinszky was partially supported by the project TÁMOP-4.2.2.A-11/1/KONV-2012-0064 “Regional effects of weather extremes resulting from climate change and potential mitigation measures in the coming decades”. The HGTTUB2007 geoid model was kindly provided by Gyula Tóth from the Department of Geodesy and Surveying of the Budapest University of Technology and Economics.

Edited by: A. Jordán

References

Mandlburger, G., Hauer, C., Höfle, B., Habersack, H., and Pfeifer, N.: Optimisation of LiDAR derived terrain models for river flow...