Possibilities and challenges in the application of multi-temporal airborne lidar data sets for the monitoring of archaeological landscapes

Ralf HESSE

State Office for Cultural Heritage Baden-Württemberg, Germany

Abstract: High-resolution digital elevation models (usually based on airborne laser scanning) have been applied for archaeological research for more than ten years. In some regions, repeated coverage is becoming available, resulting in opportunities for the detection of changes which have occurred in-between the different surveys. However, while DTM change detection is in principle very simple, the practical application faces a number of challenges. These challenges include spatial resolution, horizontal and vertical accuracy as well as impacts of vegetation cover and data processing (e.g. strip adjustment and vegetation filtering). In addition to these challenges, the issue of comparing DTMs with DSMs arises when lidar-derived DTMs are supplemented with lower-cost and often more easily acquired photogrammetric DSMs. As a result, the seemingly straightforward approach to monitoring archaeological landscapes by analysing multi-temporal elevation data sets is limited with respect to the detectability of relief changes and the achievable accuracy of the quantification of such changes. Because of the ongoing developments in terms of spatial resolution and accuracy, it is usually the (older) baseline data set which limits the applicability and informative value of change detection approaches. Therefore, it is expected that large area monitoring schemes based on airborne lidar will only become operational once repeated coverage by high-quality surveys becomes available. However, results achieved in small test areas in Baden-Württemberg are promising despite the mentioned challenges.

Keywords: monitoring, change detection, airborne lidar, ALS

Introduction

Given the large number of archaeological sites in many regions and the generally growing pressures from land use and development, monitoring their condition and detecting damages or threats can be an overwhelming task. While scant financial and staff resources are dedicated explicitly to the monitoring of archaeological sites and landscapes, "to monitor the condition of cultural monuments" (§1 DSchG Baden-Württemberg) is nevertheless commonly a legal obligation for cultural heritage institutions. This central central heritage management task calls for the recording, understanding and quantification of the impacts of natural and anthropogenic processes on archaeological heritage, because the detection of threats is indispensable for the development of successful conservation strategies. In particular where large (and/or poorly accessible) areas are concerned, remote sensing techniques are advantageous (e.g. HESSE 2015).

In recent years, airborne lidar (or ALS) data have become an important data source for archaeological research and cultural heritage management (e.g. OPITZ & COWLEY 2013). Aided by the increasing availability of such data and novel visualisation techniques, airborne lidar analysis has become common practice, in particular in the field of archaeological prospection (DONEUS & BRIESE 2011; BENNETT et al. 2013; KOKALJ et al. 2013; CRUTCHLEY 2013; GOJDA 2014). Even early on, the foreseeable availability of repeated coverage spawned the idea that multitemporal ALS data sets could be used to detect and quantify topographic changes and could hence allow a monitoring of archaeological sites and landscapes. Today, repeated ALS coverage is indeed becoming available, and opportunities for testing the potential for change detection arise. However, while change detection using multiple DTMs is in principle very simple (subtracting one DTM from the other), its practical application faces several challenges which limit the applicability of the approach. This paper presents challenges and limitations encountered in a test study as well as approaches towards solutions of these issues.

Example data sets

The primary requirement for the applicability of change detection using ALS-based DTMs is the availability of multi-temporal coverage. In Baden-Württemberg, one of the largest federal states of Germany, a state-wide ALS data set of 35,751 km2 (based on surveys in the years 2000-2005) is available and is being used for large area archaeological prospection (HESSE 2013). Additional data sets are available from surveys commissioned by the State Office for Cultural Heritage. Together with the state-wide data set these result in dual coverage of 183 km2 (with the first data set based on surveys in the years 2001 and 2002 and the second data set based on surveys in the years 2007-2009 and 2014) and triple coverage of 1.7 km2 (based on surveys in the years 2002, 2003 and 2008 for the first, second and third data set, respectively).

The data sets used in the present study pertain to the environment of the monastery Maulbronn, a UNESCO World Heritage. For this area a subset of the state-wide ALS data based on a survey in 2002 as well as a new data set based on a survey in March 2014 (commissioned by the State Office for Cultural Heritage) are available. The survey was primarily commissioned to detect and map the extensive traces of the medieval water system of the Cistercian monastery. The area has more than 50% forest cover. Archaeological relief features include the medieval water system, field systems of unknown age as well as early modern fortifications.

The second requirement for DTM change detection is that the two (or more) data sets have to have identical spatial resolution. The example data sets differ strongly in terms of ALS point cloud density. While the 2002 data set has approximately one ground point per square metre, the 2014 data set has approximately 25 ground points per square metre. As rasterized DTMs, this amounts to a spatial resolution of one metre and 20 centimetres, respectively. While it may be tempting to resample the lower-resolution data set to the resolution of the higher-resolution data set, this would be futile in terms of the detectability of spatially smaller changes. This is due to the limitation stated by the Nyquist sampling theorem (FERREIRA & HIGGINS 2011) that the highest recordable frequency component (i.e., smallest detail) is half the sampling rate (i.e., spatial resolution of the point cloud). Hence the resolution of the coarser (baseline) data set has to be used.

ALS accuracy and geo-referencing / co-registration issues

Assuming that both DTMs are error-free, a simplistic differencing of the two DTMs can be attempted. However, the resulting difference map shows apparent changes which would suggest a landscape-scale transformation of the topography. DTM errors thus have to be identified and, if possible, removed. The provider of the 2014 data set supplied the following information: GPS position error $\sigma < 0.15$ m, ALS point position error < 0.5 m. Assuming the same error margins for both data sets and taking into account error propagation, the combined ALS point position error is <0.71 m. Errors of up to 0.71 m would clearly render such data sets useless for most if not all archaeological monitoring purposes. However, this is only valid if all error is assumed to be random (not systematic). As no information is available regarding the nature of error (random vs. systematic), the goal is to remove the systematic error component in an attempt to reduce total error.

A conspicuous detail in fig. 1 is that – if the difference map is taken at face value and interpreted as a change map – the terrain appears to have been generally lowered on northern slopes and raised on southern slopes, and that both the lowering and the raising are stronger for steeper slopes. This is best explained by a north-south shift of the more recent DTM against the baseline DTM. This can be partially explained by the problem that different software treats the same data differently. For example, if the same point cloud is rasterized by different software (e.g. ENVI 4.6 vs. Global Mapper 10) the two resulting DTMs may be shifted against one another (fig. 2). In this case, it has not been resolved which of the two programmes is in error and whether the error occurs during rasterization or in saving the resulting file.

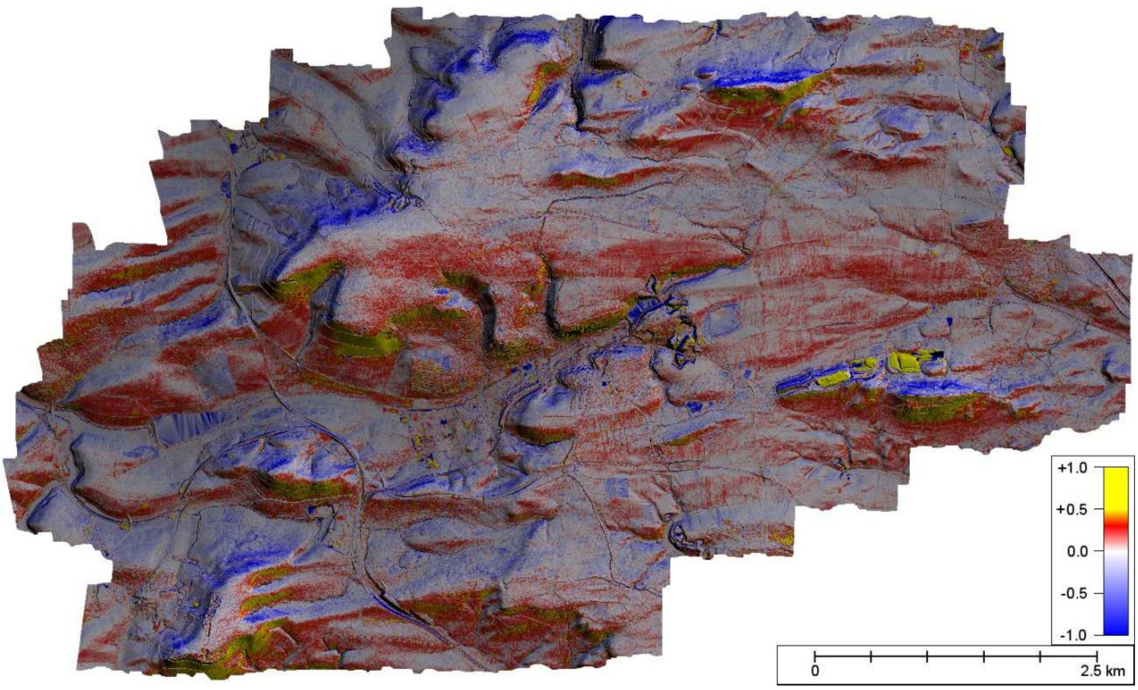

Fig. 1 – Colour-coded simplistic difference map draped over shaded relief image of the study area. Lidar data: LGL/LAD Baden-Württemberg.

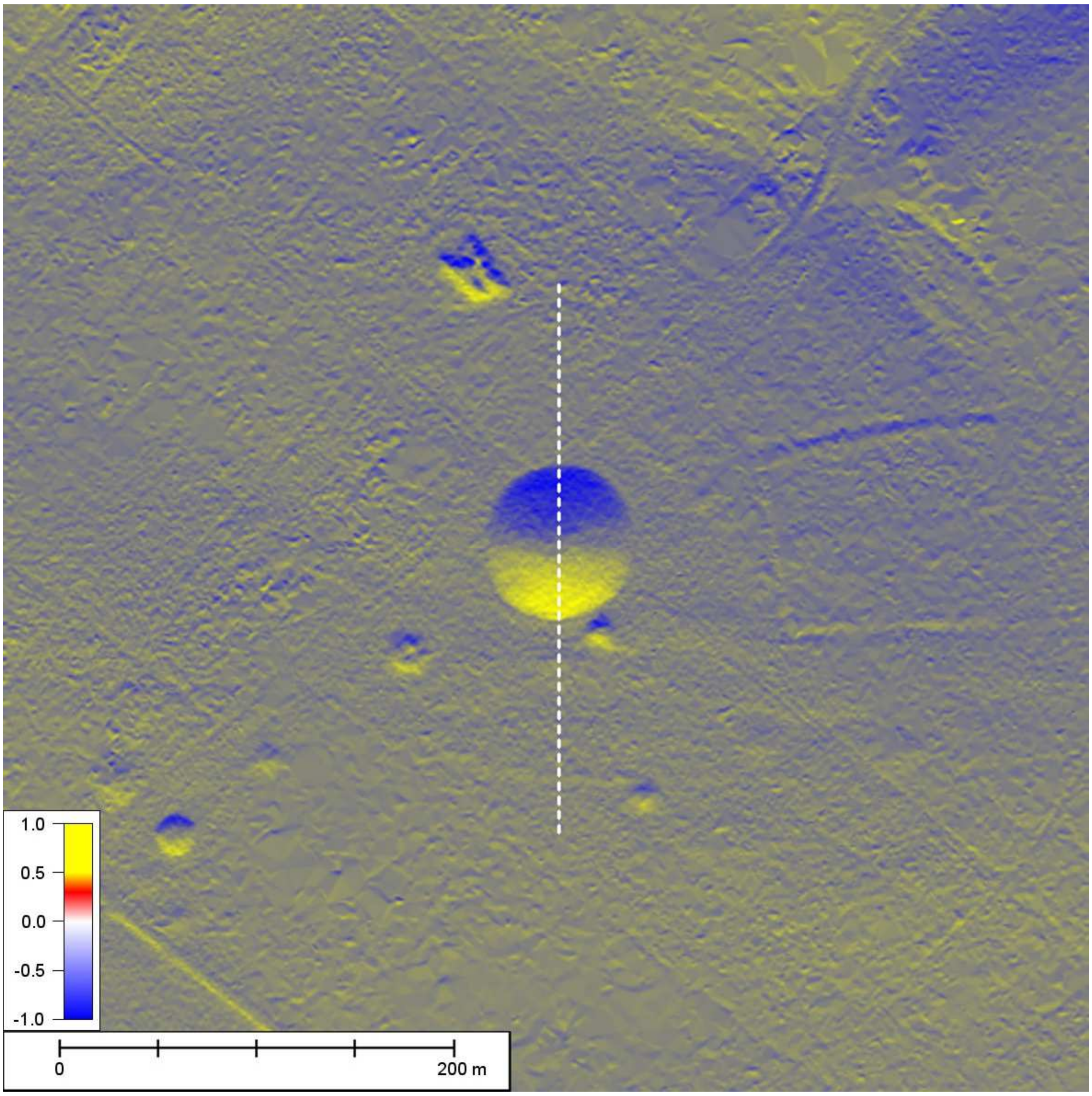

Fig. 2 – Colour-coded difference map of two DTMs created with different software from the same ALS point cloud. Lidar data: LGL/LAD Baden-Württemberg.

Dealing with systematic error

A straightforward approach for dealing with systematic error (accuracy and geo-referencing issues) is to find best fit co-registration between the two data sets. This can be done by shifting the data sets against each to find the X and Y offsets for which the fit between the data sets is optimized (fig. 3). Criteria for best fit co-registration can, for example, be minimized root mean square error (RMSE) or maximized correlation coefficient. In the present study, X and Y shifts were found to be sufficient to correct systematic error. In other cases, a rotation transformation may be perhaps necessary.

```
for all X offsets in the range -3 m to +3 m
    for all Y offsets in the range -3 m to +3 m
        shift data set 2 against data set 1 by X and Y offset
        compute RMSE
        if RMSE < minRMSE then
            minRMSE = RMSE
            remember X and Y offsets
apply remembered X and Y offsets to data set 2
```

Fig. 3 – Pseudocode for finding the best fit co-registration.

The described approach for finding the best fit co-registration may be affected by actual topographic changes. If actual topographic changes are known or assumed to be small and not biased towards certain slope aspects, their influence on the determination of best-fit co-registration is negligible. In areas with widespread and/or aspect-biased topographic changes, best fit co-registration has to be determined for unchanged references areas. This may pose the challenge of defining such unchanged reference areas before being able to perform the actual change detection. A further problem may be that the best fit co-registration offset may be spatially variable rather than constant over the area under investigation. Once the best fit co-registration offset has been determined, the spatial shift between the data sets can be corrected. In addition, a Z direction shift based on the average vertical difference between the two data sets can be applied. In the present case, a shift by -0.51 m and -1.77 m had to be applied in the X and Y direction, respectively (fig. 4).

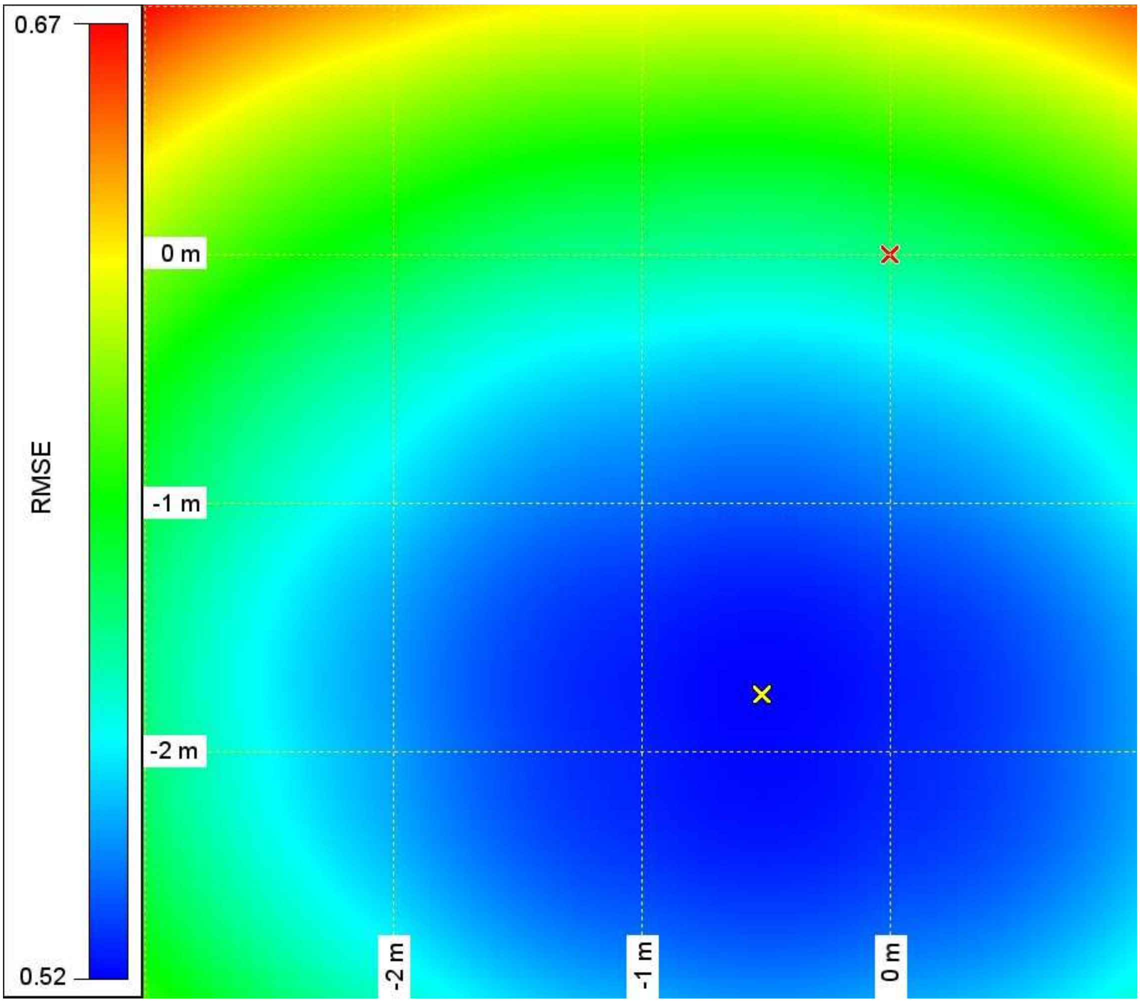

Fig. 4 – Dependence of RMSE calculated from two DTMs shifted against each other in X and Y direction. Position of lowest RMSE (yellow X in blue area) indicates best fit co-registration between data sets.

Difference map = change map?

After correcting for systematic error the difference map (fig. 5), there are no aspect-related apparent topographic changes which were conspicuous in the simplistic DTM differencing. The difference map histogram (fig. 6) shows a marked improvement from σ = 0.40 m to σ = 0.17 m (which may at least partially be due to actual topographic changes). This is a significant improvement in comparison to the theoretical error of up to 0.71 m.

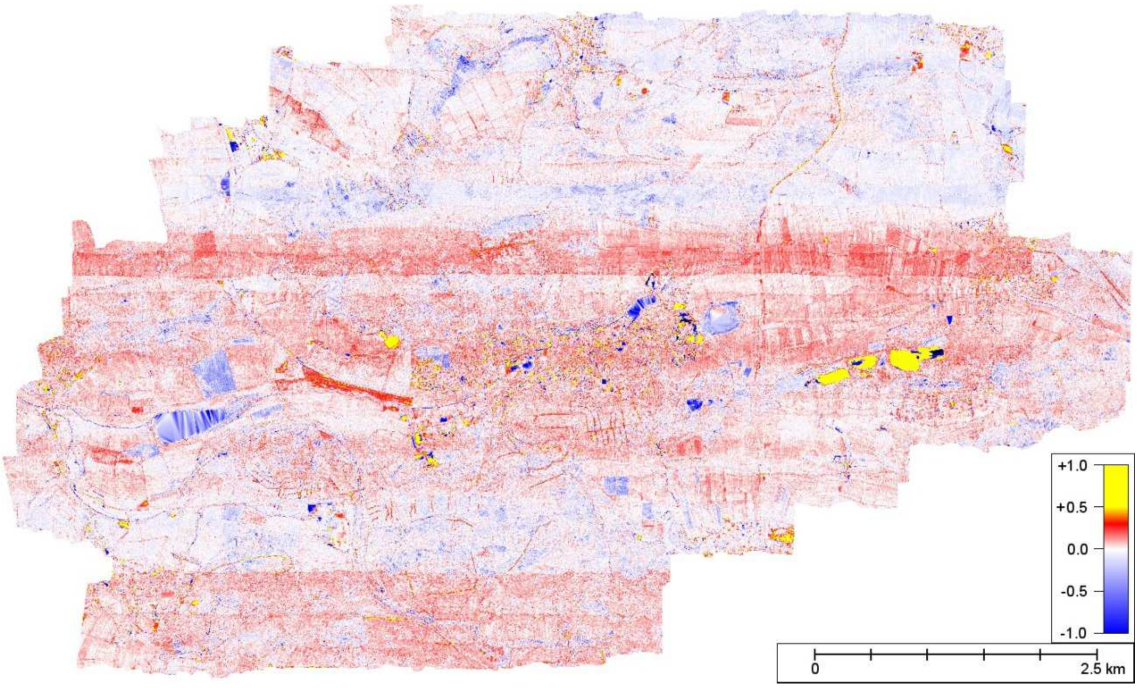

Fig. 5 – Colour-coded difference map of the study area after correcting systematic error by applying X and Y shifts based on minimized RMSE. Lidar data: LGL/LAD Baden-Württemberg.

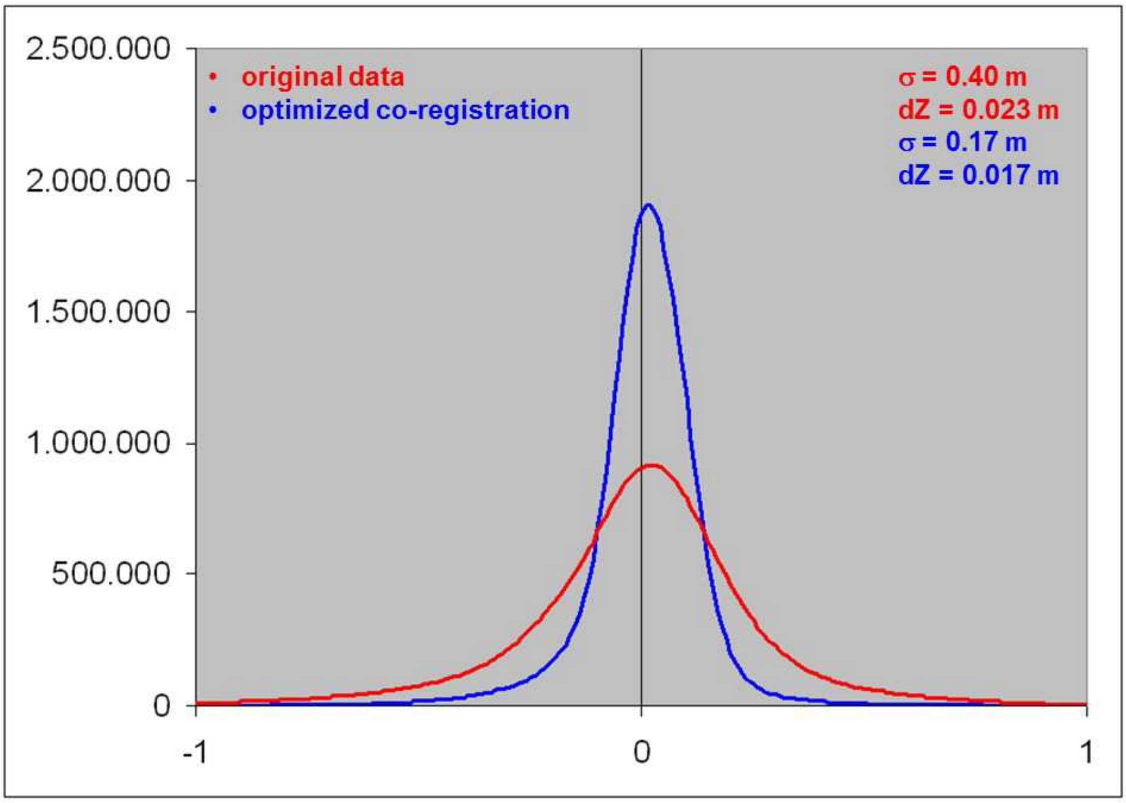

Fig. 6 – Difference map histograms without (red) and with (blue) optimized co-registration.

The question now is whether this difference map can be interpreted as a change map. In the difference map of the entire study area (fig. 5), east-west oriented bands of overall higher and lower DTM difference are noticeable. These correspond to individual flight paths and are thus artefacts related to strip adjustment. In a detail image (fig. 7), both the strip adjustment artefacts (here as positive difference map values) and rough-looking patches with negative values are visible. These patches of apparent relief lowering are due to improved penetration of low vegetation by the 2014 lidar survey as compared to the 2002 survey. Differences in vegetation penetration and filtering between surveys as well as differences in actual vegetation cover impenetrable by both surveys thus can have a marked effect on apparent topography and apparent topographic changes. In this case study, both the strip adjustment artefacts and the vegetation-related artefacts are mostly on the order of 0.1 to 0.2 m.

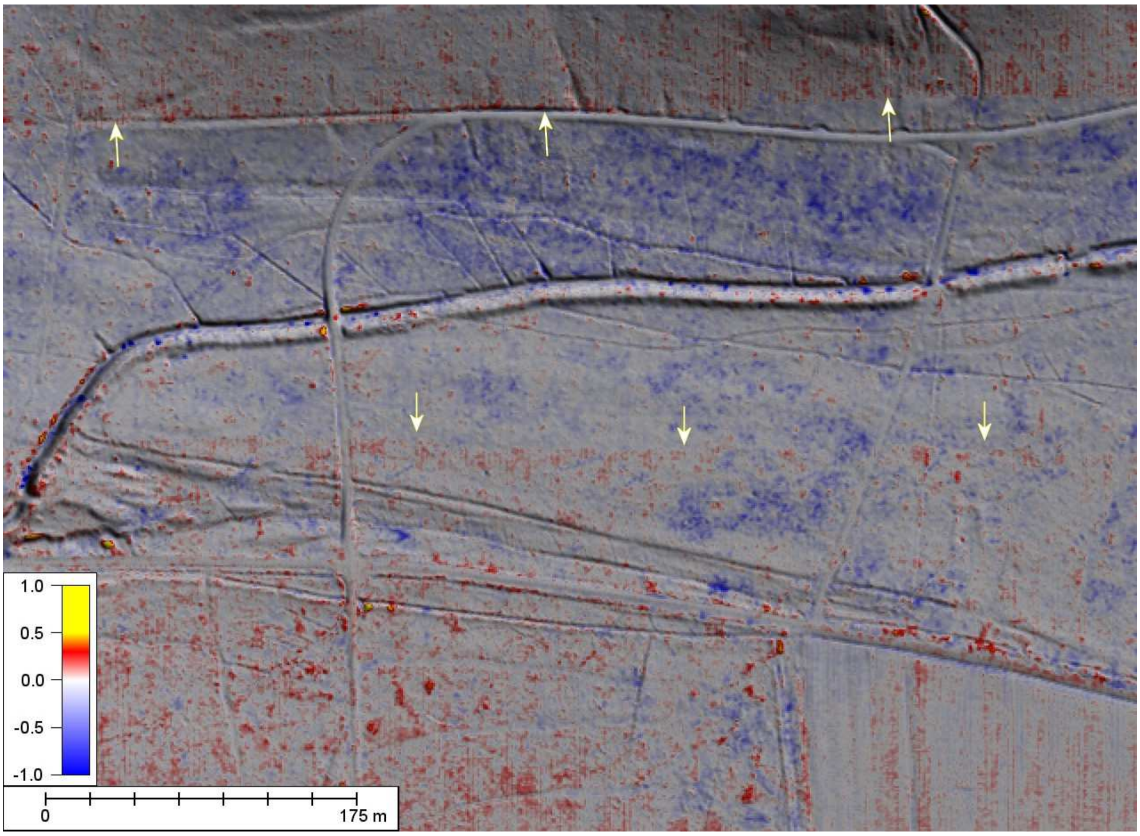

Fig. 7 – Detail of the colour-coded difference map after correcting systematic error, draped over shaded relief image of the 2014 data set. Arrows point to edges of flight paths visible as strip adjustment artefacts; blue patches (negative values) are due to better penetration of low vegetation by the second survey. Lidar data: LGL/LAD Baden-Württemberg.

There are, however, features in the difference map which can be interpreted as actual topographic changes. Fig. 8, for example, shows raised forestry lanes with heights between 0.2 and 0.6 m. However, without field investigation, it is not possible to state whether these features are earthen or gravel dams or rather piled up branches. The topographic changes shown in fig. 9 have been verified in the field. The ten metre wide and one metre high bank of material deposited alongside a forestry road also covers a portion of a newly discovered earthwork interpreted as an Iron Age rectangular enclosure (*Viereckschanze*).

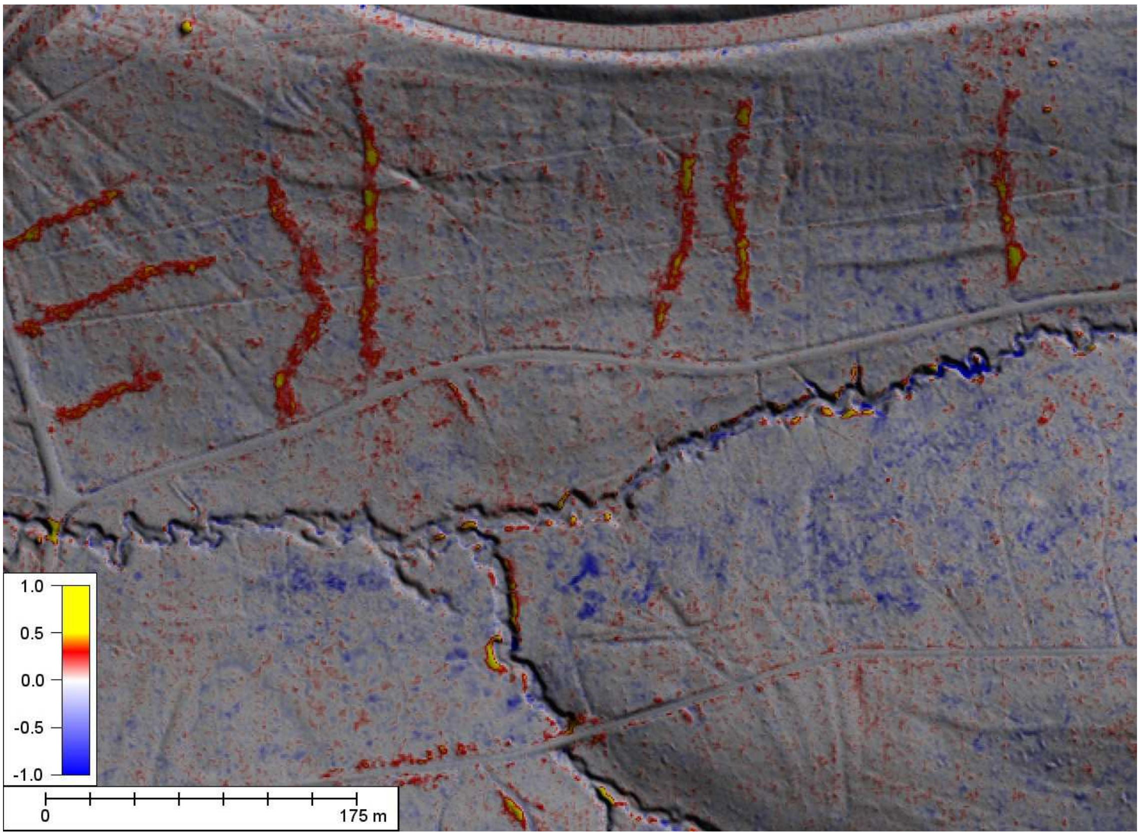

Fig. 8 - Raised forestry lanes detected by DTM differencing cut across historic water channels and lynchets. Lidar data: LGL/LAD Baden-Württemberg.

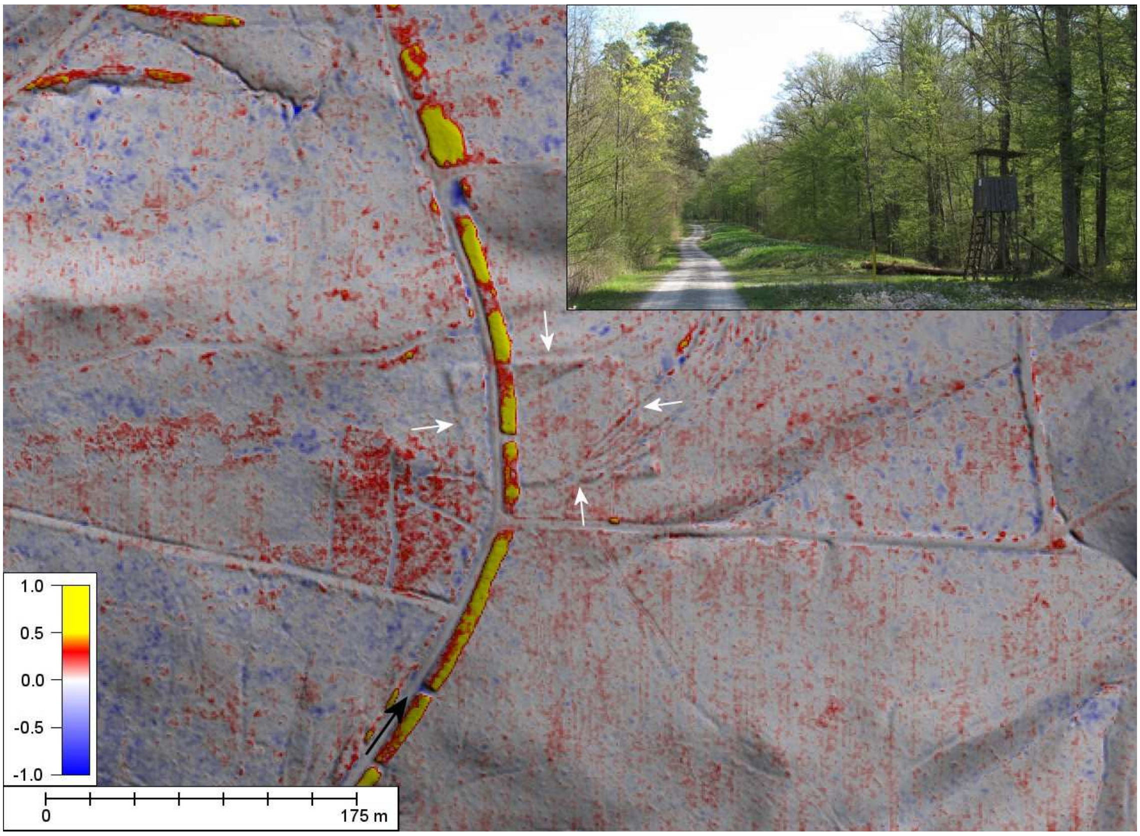

Fig. 9 – Ten metre wide bank deposited along forestry road between 2007 and 2014. Black arrow indicates viewing direction of photograph (inset); white arrows point out a newly discovered earthwork. Lidar data: LGL/LAD Baden-Württemberg.

SfM vs. ALS for change detection

The sole reliance of this data source for DTM change detection is usually not satisfactory. Where repeated ALS coverage is becoming available, this is presently only the case for small areas. Commissioning additional ALS surveys is still a very expensive endeavour, making high temporal coverage impossible. Perhaps even more importantly, if the earliest available ALS data set (commonly from the first decade of the 21st century or even younger) is used as baseline, it is impossible to detect and quantify topographic changes before that time. A possible solution to these problems is the creation of photogrammetric DSMs by Structure-from-Motion (SfM). The main advantages are that they can be created more easily and rapidly and at much lower cost than ALS data and that SfM-based DSMs can be created from legacy aerial photographs originally taken for aerial archaeology or other purposes. While such approaches produce promising results (RISBØL et al. 2015), creating photogrammetric DSMs suitable for integration in DTM change detection is commonly hampered by poor ground control, too few photographs and small area coverage. Furthermore, they cannot be applied in forests, and the general problem of comparing DTMs with DSMs in areas with vegetation cover has to be considered.

Conclusions

Change detection based on ALS-derived DTMs poses challenges resulting from ALS point cloud accuracy (including errors related to measurement and strip adjust), geo-referencing and co-registration issues as well as issues related to vegetation cover, penetration and filtering. ALS-based change detection can be feasible provided that systematic errors are removed. However, its applicability faces several limitations, including the poor detectability of small relief changes (in the present study, noise and artefacts are commonly on the order of 0.1 to 0.2 m, rendering the detection of lower relief changes all but impossible) and the achievable accuracy of change quantification. A fully automated change detection does not appear feasible with the data sets similar to those used in this study. Given the nature and magnitude of errors, visual interpretation of the difference map is and will likely remain necessary for change detection and monitoring based on DTM data sets of comparable accuracy. With improvements in spatial resolution and accuracy of ALS data, large area archaeological monitoring can be expected to become feasible in the future. However, it is usually the older (baseline) data sets which limits spatial resolution and accuracy of change detection. Therefore, repeated coverage by high quality lidar surveys will be required.

References

BENNETT, R., WELHAM, K., HILL, R.A., FORD, A., (2013). Using lidar as part of a multi-sensor approach to archaeological survey and interpretation. In: Opitz, R., Cowley, D.C. (eds.) (2013). Interpreting archaeological topography. 3D data, visualisation and observation. Occasional Publication of the Aerial Archaeology Research Group No. 5. Oxbow: Oxford. pp. 197-205.

CRUTCHLEY, S., 2013. Using lidar data – drawing on 10 years' experience at English Heritage. In: Opitz, R., Cowley, D.C. (eds.) (2013). Interpreting archaeological topography. 3D data, visualisation and observation. Occasional Publication of the Aerial Archaeology Research Group No. 5. Oxbow: Oxford. pp. 136-145.

DONEUS, M., BRIESE, C., (2011). Airborne laser scanning in forested areas – potential and limitations of an archaeological prospection technique. In: Cowley, D.C. (ed.), Remote sensing for archaeological heritage management. Proceedings of the 11th EAC Heritage Management Symposium, Reykjavík, Iceland, 25-27 March 2010. EAC Occasional Paper No. 5, Occasional Publication of the Aerial Archaeology Research Group No. 3, pp. 59-76.

FERREIRA, P.J.S.G., HIGGINS, R. (2011). The establishment of sampling as a scientific principle - a striking case of multiple discovery. Notices of the American Mathematical Society 58(10): 1446-1450.

GOJDA, M., (2014). Testing the potential of airborne lidar scanning in archaeological landscapes of Bohemia: strategy, achievements and cost-effectiveness. In: Kammermans, H., Gojda, M., Posluschny, A.G. (eds.), A sense of teh past. Studies in current archaeological applications of remote sensing and non-invasive prospection methods. BAR Internation Series, Vol. 2588. pp. 83-91.

HESSE, R. (2013). The changing picture of archaeological landscpes: lidar prospection over very large areas as part of a cultural heritage stategy. In: Opitz, R., Cowley, D.C. (eds.) 2013. Interpreting archaeological topography. 3D data, visualisation and observation. Occasional Publication of the Aerial Archaeology Research Group No. 5. Oxbow: Oxford. pp. 136-145. pp. 171-183.

HESSE, R. (2015). Combining Structure-from-Motion with high and intermediate resolution satellite images to document threats to archaeological heritage in arid environments. Journal of Cultural Heritage 16(2): 192-201.

KOKALJ, Ž., ZAKŽEK, K., OŠTIR, K. (2013). Visualizations of lidar derived relief models. In: Opitz, R., Cowley, D.C. (eds.) 2013. Interpreting archaeological topography. 3D data, visualisation and observation. Occasional Publication of the Aerial Archaeology Research Group No. 5. Oxbow: Oxford. pp. 100-114.

OPITZ, R., COWLEY, D.C. (eds.) (2013). Interpreting archaeological topography. 3D data, visualisation and observation. Occasional Publication of the Aerial Archaeology Research Group No. 5. Oxbow: Oxford.

RISBØL, O., BRIESE, C., DONEUS, M., NESBAKKEN, A. (2015). Monitoring cultural heritage by comparing DEMs derived from historical aerial photographs and airborne laser scanning. Journal of Cultural Heritage 16: 202-209.

Imprint:

Proceedings of the 20th International Conference on Cultural Heritage and New Technologies 2015 (CHNT 20, 2015)

Vienna 2016

http://www.chnt.at/proceedings-chnt-20/

ISBN 978-3-200-04698-6

Editor/Publisher**:** Museen der Stadt Wien – Stadtarchäologie

Editorial Team: Wolfgang Börner, Susanne Uhlirz