UAV Remote Sensing Data Handling: A Transition from Testing to Long-Term Data Acquisition for Intensive Forest Monitoring

STUART KRAUSE[1,2,3], MAXIMILIAN STRER[1], JAN-PETER MUND[2] & TANJA GM SANDERS[1]

Abstract: Unmanned Aerial Vehicles (UAVs) and their application in intensive forest monitoring is opening up new possibilities for forest research. Recent trials have shown promising results for the extraction of tree parameters and the quantification of phenological observations. Long-term data acquisition strategies require the handling of large amounts of data as well as increased computing performance requirements. This study explores the challenges in transitioning from the testing phase of UAV-based remote sensing technology to a long-term data acquisition scheme capable of handling Big Data by means of robust archiving and processing workflows as well as the incorporation of cloud computing.

1 Introduction

For the past two years, the use of Unmanned Aerial Vehicles (UAVs) as a nearfield remote sensing platform were tested at the Britz forest research station in Brandenburg, Germany. The possibility to acquire high temporal, spatial and spectral resolution datasets by means of self-determined guidelines rather than adapting to secondary sources is of interest (TORRESAN et al. 2016). During this testing phase, selected datasets were acquired, processed and analyzed on local computers and in-house servers. The next stage in the progression of this rapidly developing technology is now however, the establishment of long-term data collection schemes which require the capacity to handle the storage and processing of "Big Data" (LI et al. 2016).

Recent studies at the Britz research station showed that UAV mounted multispectral sensors can obtain accurate geometric features such as tree heights and crown diameter and promising results have been demonstrated in the quantification of phenological phases. Remote sensing datasets acquired in the field for experimental studies were successfully processed and analyzed on localized Linux and Windows workstations. The transition to a more permanent and long-term data collection and processing scheme does however pose the challenge in terms of data storage and processing capabilities.

The concept of Big Data has been in discussion for a number of years now and it is estimated that the volume of global scientific data doubles every two to three years (GUO et al. 2014). Due to an increase in the interest of remote sensing data for environmental research, the issue of acquiring, storing, analysing, visualizing and sharing such data is becoming a challenge (LIU 2015) and solutions are required.

[1] Thünen Institute of Forest Ecosystems, Alfred-Möller-Str. 1, Haus 41/42, D-16225 Eberswalde, E-Mail: [stuart.krause, maximilian.strer, tanja.sanders]@thuenen.de
[2] Eberswalde University for Sustainable Development, Alfred-Möller-Str. 1, D-16225 Eberswalde, E-Mail: Jan-Peter.Mund@hnee.de
[3] University of Bonn, Meckenheimer Allee 166, 53115 Bonn, Germany

Big Data is not only attributable to the volume and velocity of data as in storage and computing capabilities, however also the variety and complexity of remote sensing data (MA et al. 2015). In terms of long-term UAV image data collection by a research institution, the amount of data would not necessarily be deemed as Big Data especially when compared to a typical satellite ground station which could be receiving up to a terabyte of data per day. Relatively speaking however, a low altitude (flying height < 100 m) UAV employed by a forest research institution during phenological phases can acquire up to several terabytes of raw data per year. The storage of such data as well as photogrammetric products could prove challenging for institutions not yet prepared for long-term acquisition schemes which could potentially carry on for decades. Furthermore, the processing, analysis and visualization of such data requires costly high-performance computing hardware.

Cloud computing, also known as "Serverless Computing" (STIGLER 2018), allows the implementation of on-demand self-service computing resources without having to invest in expensive localized high performance hardware. Users can access an online control panel and configure "instances" in the cloud depending on current computing requirements. An instance can be created in a "pay-as-you-go" manner where users are charged only for the time an instance is running. Popular cloud computing services such Amazon Web Services (AWS), Microsoft Azure and Google Cloud can run Linux and Windows multiple instances configured to a user's processing requirements.

In this study, the concepts of Big Data and cloud computing relative to UAV multispectral data acquisition for long-term intensive forest monitoring will be explored. Hereby, proposed protocols for the archiving and processing of acquired UAV-based remote sensing datasets as well as resulting photogrammetric products will be presented.

1.1 Study Site

The Britz research station is an intensive forest monitoring site located near the town of Eberswalde, Germany (Fig. 1) and under administration of the Thünen-Institute for Forest Ecosystems. Since the mid-1970s the site has been implemented for the continuous acquisition of weather and forest hydrology data, as well as individual tree parameters in for example incremental tree growth by dendrometers and tree health with sap-flow sensors. Recently, the site is also being used as a testing ground for the development of UAV-based remote sensing methods to modernize and digitize typical national forest inventories and intensive forest monitoring (i.e. Level II and ICP Forests). Currently, the extraction of individual tree attributes in terms of tree height and crown diameter as well as the quantification of phenological observations provided promising results.

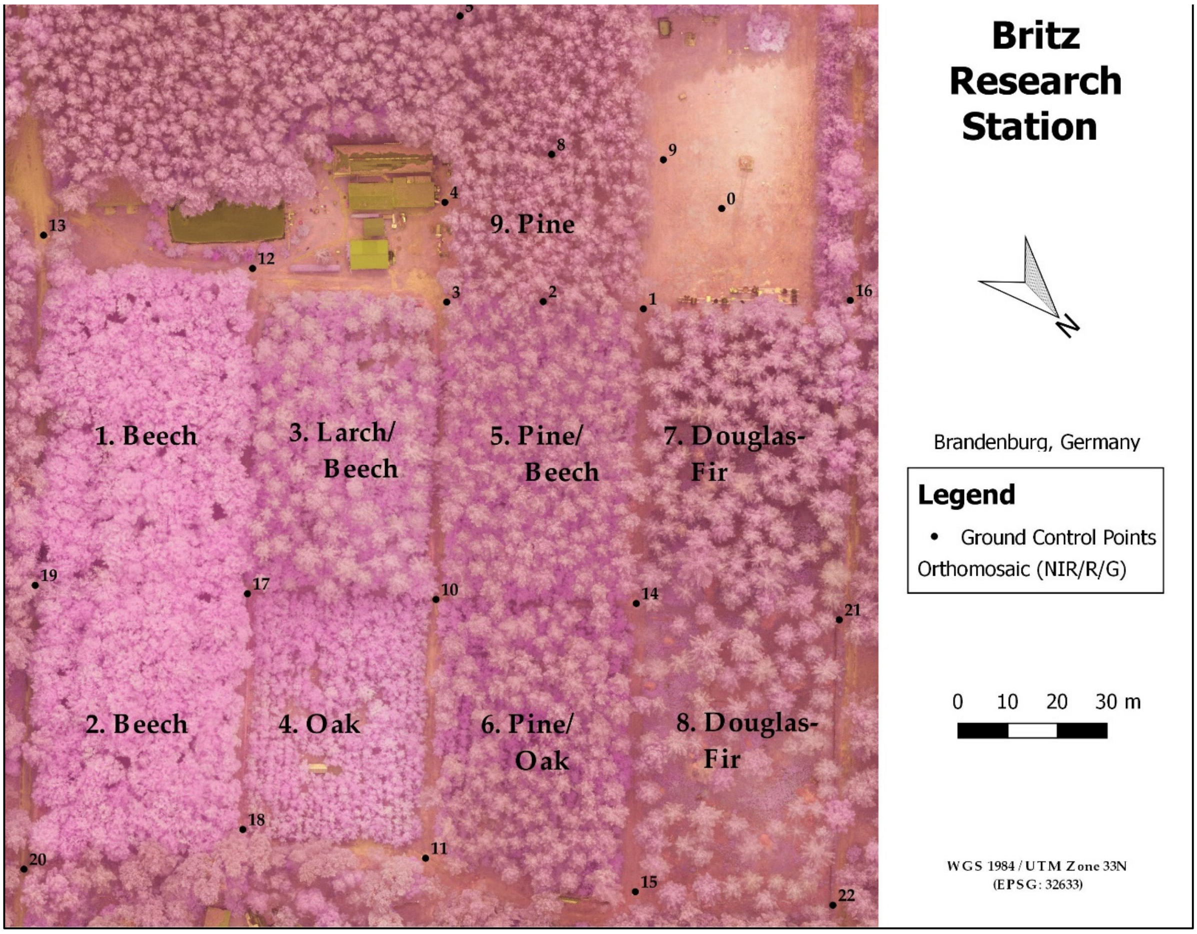

Fig. 1: Britz Research Station in Brandenburg, Germany. Nine tree stands totaling an area of approximately 4 ha.

1.2 High Resolution Datasets

The benefits of employing UAVs and their ability to be mounted with various lightweight active and passive sensors at a high temporal, spatial and angular resolution is of interest for forestry research. Cost-effective consumer-grade optical sensors mounted on UAV Photogrammetry (UAVP) platforms can acquire multi- and hyperspectral imagery for the purpose of creating very dense photogrammetric point clouds as well as very high resolution (VHR) Orthomosaics with Structure of Motion (SfM) software (Fig. 2).

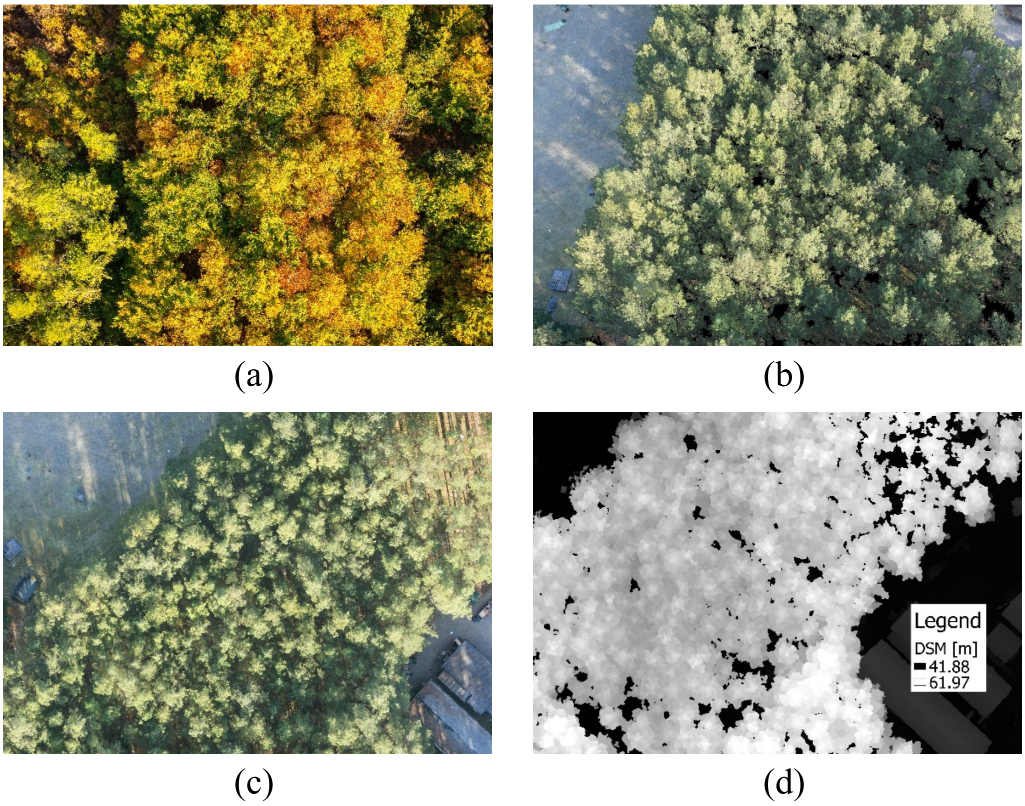

Fig. 2: (a) Aerial Image (b) Point Cloud (c) Orthomosaic (d) Digital Surface Model

For the purpose of assessing changes in forests over time, there is an interest in the capability to acquire data at a high temporal resolution. UAVs can be flown daily depending on weather conditions whereas open access passive satellite-based sensors such as Landsat-8 and Sentinel-2ab have revisit times of 16 and 5 days respectively. UAVs will be limited to a much smaller area (< 5 ha), however can acquire VHR (< 2 cm) imagery which will not be affected by cloud cover in terms of occlusion of the earth's surface. An upscaling of UAV datasets for inventory plots could cover large areas with the incorporation of multispectral satellite data and terrestrial measurements and observations.

In terms of spatial resolution, UAVP can provide a ground sampling distance (GSD) of typically 0.08 to 0.1 meters depending on flying height and the camera's intrinsic parameters (Tab. 1). At such high resolutions, not just individual trees can be detected, however also the apical meristems and individual leaves. Furthermore, high angular resolution refers to the possibility to capture imagery at various angles as opposed to only at nadir (90°). Low oblique imagery can be implemented to parts of trees normally subjected to occlusion from nadir in order to create more complete 3D information, as well as account for the ground area at the base of tree stems. Additionally, terrestrial imagery taken horizontally from the ground floor can account for areas below the forest canopy and aid in the extraction of tree stem diameter. Terrestrial imagery can also be fused with aerial nadir and oblique imagery for the purpose of creating highly dense point clouds depicting an entire forest stand. Table 1 shows the parameters for the Sony A7r RGB and Near-Infrared (NIR) cameras typically used at the Britz research station.

Tab. 1: Parameters for the Sony A7r camera.

Sensor	Sensor Type	Focal Length	Sensor Size	Bands	GSD at 80 m
Sony A7r	Full Format	35.3 mm	7360 x 4912 pixels	RGB/NIR	0.01 m

1.3 Data Acquisition and Processing

During the testing phase, various datasets were acquired from six flight campaigns in early 2018 for the purpose of creating 3D models of forest stands comprised of imagery acquired at the nadir, oblique and horizontal angles. The flight campaigns covered a Scots Pine stand and a European Beech stand of an area of 0.4 ha and 0.5 ha respectively (Fig. 1). During the phenological phases of 2018, weekly flight campaigns capturing multispectral imagery were carried out for the entire Britz research station covering nine tree stands totaling an area of approximately 4 ha. Flight plans were calculated to capture overlapping multispectral imagery for all plots within two missions.
After each flight mission, RAW imagery is copied to a portable hard drive alongside GPX logfiles and flight mission metadata. Later in the office, the data acquired in the field is archived on a local server awaiting further processing. Processing begins with the direct georeferencing of RAW imagery with the GPX logfile. To follow, georeferenced RAW imagery are corrected for vignetting if necessary, and histograms are matched for exposure and converted to 8-bit or 16-bit Tagged Image File Format (TIFF). Images are then loaded into the photogrammetry software and GNSS RTK Ground Control Points (GCPs) are marked within imagery. The SfM photogrammetry software, once set up, processes for the most part semi-automatically depending on the software package. Photogrammetric products such as point clouds, orthomosaics and Digital Surface Models (DSMs) are then further processed for analysis (i.e. Canopy Height Models).

Tab. 2: Data size distribution of a typical nadir flight mission with reference to the processing stage.

Sensor	Format	RAW (Image)	RAW Count	GeoTIFF	GeoTIFF (Selected)	Point Cloud	Orthomosaic
Sony A7r	ARW	35 MB	350	300 MB	300	2 GB	2 GB

2 Proposed Protocols

In order to plan for future flight campaigns and long-term UAV data acquisition, protocols were developed based on the example of a weekly flight campaign for phenological observations carried out with an Octocopter equipped with RGB and NIR sensors. Ideally each flight mission is radiometrically calibrated with a calibration target and flown in winds under 5 m/s to limit tree movement and aircraft instability. Each mission is required to be accompanied with metadata which is updated during the processing chain:

- Average windspeed throughout mission
- Flying height, estimated side- and frontal-overlap, focal length, sensor size, and GSD.
- After SfM processing: sensor intrinsic parameters, camera positions, actual GSD, RMSE, point cloud density, image list, and number of GCP markings.

Table 3 displays a description of the proposed processing levels alongside estimated processing times. RAW images are also converted to Adobe Digital Negative (DNG) format for the purpose of insuring prolonged support. Depending on storage capabilities, Level 2a-c can be deleted when no longer required. Important to note is that the Level 2a processing stage is only carried out in the case that GCPs and/or radiometric calibration panels were not available.

Tab. 3: Proposed processing levels for UAV acquired multispectral imagery (adapted from BIESEMANS & EVERAERTS 2006).

Product Level	Description	Est. Processing Time
Level 0	RAW imagery is recorded directly from the sensor. Imagery is stored on a portable hard drive alongside GPX log file. All images are included, including those acquired before and after the automatic flight grid.	Directly after flight mission (10 minutes)
Level 1a	RAW data converted to DNG and archived on local server alongside GPX log file and metadata containing flight plan parameters including wind speed and quality rating.	Upon returning to the office (60 minutes)
Level 1b	Copies of original images are georeferenced with the log file, selected for specific plot coverage and uploaded to the S3 (AWS) cloud storage. Plot selection added to ID in database as well as folder name.	30 minutes plus upload time
Level 1c	Georeferenced RAW imagery is copied to an EC2 instance (AWS) and corrected for Vignetting and exposure histograms are matched. Images are converted to 16-bit GeoTIFF and RAW imagery is deleted from the cloud (S3).	60 minutes (mostly automated)
Level 2a	Images are processed with SfM software without GCPs or radiometric calibration. Outputs are point clouds and orthomosaics.	2 hours
Level 2b	Images are processed with SfM software including GCPs.	2 – 3 hours
Level 2c	Images are processed with SfM software with GCPs and radiometric calibration.	3 – 3.5 hours
Level 3	Point clouds (LAS) and orthomosaics (GeoTIF) are clipped to a selected forest stand, downloaded from cloud and stored on a local server. EC2 instance(s) is (are) terminated.	30 minutes
Level 4a	Canopy Height Model is calculated.	< 10 minutes
Level 4b	Canopy Height Model and indices are calculated.	< 20 minutes
Level 5	Product levels 2a-c, and 3 are made available to the public (Levels 4ab optional).	To be determined

3 Discussion and Outlook

The linking of the archived imagery to metadata is essential for the selection of homogeneous datasets suitable for time-series analysis as well as linkage to satellite imagery for resampling and upscaling. The use of RAW imagery is important for the flexibility in image processing for scientific applications (VERHOEVEN 2010) however for camera specific RAW formats (i.e. Sony ARW) it is unknown whether such formats will be supported in the future. The decision to archive original datasets in DNG format in addition to the Sony ARW format, is based on the concept of “future proofing” (KODDE 2014) with the assumption that Adobe and other software packages will provide DNG support well into the future.

The 16-bit TIFFs created during the processing stages of Level 2a-c can be deleted in order to minimize storage space if necessary. In the event of retroactive processing, TIFFs can be recreated according to defined protocols, where the image list and exposure settings will require to be accessed in the metadata in order to insure identical processing parameters.

With Level 5, the goal would be to provide open-access data to the public in the form of orthoimagery and point clouds for research and educational purposes, and should be in accordance with the European Spatial Data Infrastructure (INSPIRE) initiative (SEIFERT 2011). This could be accomplished by a Web Mapping Service where users search and visualize datasets in a web browser and specific datasets can be downloaded directly or ordered via an Email request.

4 Conclusion

Remote sensing with UAVs for the purpose of intensive forest monitoring can accumulate large amounts of data that require adequate storage capacity and high-performance computing. The possibility to link acquired UAV remote sensing data to terrestrial measurements and observations as well as satellite data is of interest. With an understanding of the concept of Big Data in terms of the volume and complexity of data acquired through UAV remote sensing, robust processing and archiving protocols can be created in order to insure the future integrity and utility of such data. Furthermore, the use of cloud computing can provide high-performance computing capabilities eliminating the necessity to invest in costly hardware.

5 References

BIESEMANS, J. & EVERAERTS, J., 2006: Image processing workflow for the PEGASUS HALE UAV payload. 2nd International Workshop "The Future of Remote Sensing", ISPRS Inter-Commission Working Group I/V Autonomous Navigation.

GUO, H., WANG, L., CHEN, F. & LIANG, D., 2014: Scientific big data and Digital Earth. Chinese Science Bulletin, **59**(35), 5066-5073.

KODDE, M., 2014: Future-proof Data Storage. https://www.gim-international.com/content/article/future-proof-data-storage, accessed 10.12.2018.

LI, S., DRAGICEVIC, S., CASTRO, F. A., SESTER, M., WINTER, S., COLTEKIN, A., PETTIT, C., JIANG, B., HAWORTH, J., STEIN, A. & CHENG, T., 2016: Geospatial big data handling theory and methods: A review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing, **115**, 119-133.

LIU, P., 2015: A survey of remote-sensing big data. Frontiers in Environmental Science, **3**.

MA, Y., WU, H., WANG, L., HUANG, B., RANJAN, R., ZOMAYA, A. & JIE, W., 2015: Remote sensing big data computing: Challenges and opportunities. Future Generation Computer Systems, **51**, 47-60.

SEIFERT, M., 2011: INSPIRE Data Specification for harmonised Orthoimageries. Photogrammetrie - Fernerkundung - Geoinformation, **2011**(6), 447-455.

STIGLER, M., 2018: Beginning Serverless Computing - Developing with Amazon Web Services, Microsoft Azure, and Google Cloud. 1st edition, Apress, 216 p.

TORRESAN, C., BERTON, A., CAROTENUTO, F., DI GENNARO, S. F., GIOLI, B., MATESE, A., MIGLIETTA, F., VAGNOLI, C., ZALDEI, A. & WALLACE, L., 2016: Forestry applications of UAVs in Europe: a review. International Journal of Remote Sensing, **38**(8-10), 2029-2036.

VERHOEVEN, G. J. J., 2010: It's all about the format – unleashing the power of RAW aerial photography. International Journal of Remote Sensing, **31**(8), 2009-2042.