Victoria Scholl

**Useful sites:**

* <http://www.qgistutorials.com/en/index.html>
* Manual LiDAR editing. But I couldn’t get all the functionality working when using lastools with wine on mac ☹ <https://rapidlasso.com/2014/03/02/tutorial-manual-lidar-editing/>
* LASmoons (“lunar licenses for poor academics”) are complimentary licenses given to underfunded researchers for a certain number of moon cycles - <https://rapidlasso.com/lasmoons/> <https://rapidlasso.com/category/lasmoons/>
* Non-commercial software for viewing 3d things: <http://vterrain.org/Packages/NonCom/>
* University of California, Santa Barbara Dept. of Geography LAS reader and viewer code: <http://www.tpingel.org/code/lasread/lasread.html>
* Remote sensing regulations come under congressional scrutiny – laws that govern remotely sensed imagery. With the recent drone boom, the regulations are being reconsidered as they are outdated.
* LAS data classification codes: <http://desktop.arcgis.com/en/arcmap/10.3/manage-data/las-dataset/lidar-point-classification.htm>
* Swiss grid (CH1903) coordinate system is chosen so that all points in the country have positive coordinates, expressed in meters, and that easting is always greater than northing. It uses a fundamental point in Bern as the center of the projection, located at x = 600’000 E, y = 200’000 N. Reference frame is LV03 (Landesvermessung 1903) . <http://www.giangrandi.ch/soft/swissgrid/swissgrid.shtml>
* RSL Outing Excursion panning for gold albums: <https://photos.google.com/share/AF1QipOcaRucgdjCfOQ7wpEMGE7opciOrNUPAbr7yGp_xuHO99oDrMqVcdiQPiyYwmXHZg?key=cG1mQmZzWXlBRjY0NFlPdV8tS2l0bWIxSHpvU0x3>
* <https://photos.google.com/share/AF1QipP8sQPfneqiPY4e8oaR5XvTG33wiu5QRnxKUQHCl3jDqZQ3VlsXig7aULLrHMoV1g?key=WmhsNzhlbEJiSnlxUVAyR2Y1Q3h4cmRGaHhMdUNn>
* To turn off the screen but keep the computer awake: SHIFT + CTRL + POWER
* SwissPhoto - <http://www.bsf-swissphoto.com/>
* Gillian keeps a wiki page for notes and information. PhD coffee meeting notes: <http://gmilani-tux.client.geo.uzh.ch/mediawiki/index.php/Spectrolab_PhD_Coffee>
* Matthew Parkan’s useful LiDAR code: <http://mparkan.github.io/Digital-Forestry-Toolbox/>
* Map with aerial imagery for visual assessment of regions: map.geo.admin.ch
* RSL website calendar for conference info: <http://www.geo.uzh.ch/en/units/rsl/services/conference-calendar/>
* Geography Association at UZH: <http://www.geoteam.uzh.ch/>
* The Carnegie Airborne Observatory <https://cao.carnegiescience.edu/>
* FunDivEUROPE <http://www.fundiveurope.eu/>
* Collaborative RSL repository for sharing code: <https://rerun.geo.uzh.ch/login>
* ForestSAT 2016 program: <https://forestsat.blogspot.ch/p/program.html> a bridge between forest science, remote sensing, and geo-spatial applications. Held for the first time in latin America.
* WSL webmail: <https://webnotes.wsl.ch/>
* LiDAR metadata maintained by Matthew: <https://sites.google.com/site/lidaretforet/ressources/synoptique-des-donnees>
* WSL long term research sites: <http://www.wsl.ch/info/organisation/fpo/lwf/methodology/plots/index_EN>
* QGIS Tutorial to do someday: <https://www.lynda.com/QGIS-tutorials/Up-Running-QGIS/383524-2.html>
* Code by Devis and colleagues for automated feature selection with multispectral imagery: <http://remi.flamary.com/soft/soft-fl-rs-svm.html>
* Cool blog about computer vision concepts: <http://www.visiondummy.com/>

**Data**

* LiDARLabNAS 🡪 lidarlab (lidarlab, multispectral) 🡪 data

Laegern

Masoala Hall

Kanton Argau

Within each site, data varies but typically includes

* + LAS, DTM, MAT files for leaf on/off conditions.
  + RGBI Mosaic 16-bit tif
  + ORT directory.
  + CHM data product
  + Forest\_type map (immergrün/ evergreen / conifers, sommergrün / deciduous)
* Code:

sftp://vscholl@sftp.geo.uzh.ch

/group/lidarlab/asltools/

Unix Permissions notes

* 3 ownership relations: owner, group, all
* 3 access permissions: read (r), write (w), execute (x)
* A set bit has a letter, a clear bit has a dash ( - )
* chmod used to set or modify a file’s permissions
* *u* = user, *g* = group, *o* = others (not *u* or *g*), *a* = all
* *r* = 4, *w* = 2, *x* = 1. Sum of permissions used with chmod

Study Sites

* Main study site is Laegern
* Masoala Rainforest in the Zürich Zoo. It’s an indoor rainforest environment opened in 2993 for studying and practicing measurement techniques. More than 500 plant species grow there. <http://www.zoo.ch/en/visiting-zoo/areas/masoala-rainforest>
* Borneo – field campaign location during which a 1 hectare plot was imaged using terrestrial laser scanning. Reflective targets were placed along a center transect of the region, and a set of targets were moved along the path at each segment during scanning. These targets were used to register scans.

**Info / Meeting Notes**

09/19 Felix

We agree that it makes sense to move forward in the direction of applying tree detection methods to data across Switzerland, as described in the research proposal. One issue is that this data is rather heterogeneous, and methods have all only been applied to small, homogenous regions. Some areas have multiple time steps/series/dates of data collection while others have a single collect. We also do not have all of the data here to access at the moment.

There is data for the Laegern and Aargau cantons readily available - both LAS point cloud and DTM derived from it. Both leaf-on and leaf-off data is present for Aargau. However, the existing field data or ground truth is not exhaustive (it does not contain data for every single tree, so we can not assess the overall number of trees identified in this way). Fabian just manually created a morphological traits / tree crown map for Laegern, so that will be treated as ground truth for the ALS processing for that site.

There are two main ITD methods to use here; one that is successful for coniferous trees (Morsdorf et al. 2004, hybrid method that finds local minima as seed points) and the other for deciduous (Geodetic Voting by Parkan & Tuia 2015, graph based segmentation algorithm for leaf-off conditions). We want to use both methods and combine the results, since they have been validated in each respective forest type.

My first task is to classify coniferous from deciduous trees, so the proper tree detection technique may be applied to each tree type region (and then to possibly do some morphological processing to grow the region of interest for each tree type). For the Kanton Argau, Felix suggests utilizing leaf-off data since the conifers keep their leaves while deciduous trees lose them during this time. These results will be validated using multispectral imagery, and/or species mapping done manually (such as Fabian’s recent creation of a Laegern 2010 morphological traits table).

After the ALS data is read into MATLAB, its structure contains a variety of fields. The Return Number (also referred to as Echo) field is of interest. The numbers range from 11, 21, 22, 31, 32, 33, … 76, 77. The pattern is as follows: the first number indicates the total number of returns for that pulse. The second number indicates which of the returns the specific point is. For instance: 11 means a single return. 21 means the first of two returns. 2 means the second of two returns.

I must find all of the first returns in the point cloud based on the second of these two numbers. Felix suggested using the unique MATLAB function. Subtract the DTM from them, create a raster (raw2ras) and play with the resolution parameter. Is there a threshold we can use to identify coniferous from deciduous regions? Coniferous regions should have high values in the raster while deciduous should be low or noisy. Incorporation of intensity information may also be helpful.

9/22 Felix

Felix has reached out to Christian Ginzler at the WSL, and he says our plan is “Tip Top”! In German, that means it sounds good. He is on board with the plan to initially identify coniferous vs. deciduous areas of trees, then apply the appropriate processing to them. One day, we will hopefully organize a meeting with him and the others at WSL.

Canton Aargau has some products including forest type (sommergrün vs. immergrün). This was created by Reik based on phenology from multiple seasons, the difference between the leaf on/off data. The resulting map will be useful for me as truth (coniferous vs. deciduous), but this type of map is not available for many of the cantons (many only have one collection time of the year). There is also the MATLAB species data that Fabian created, which can be used to validate some tree classification results in the Laegern Canton. After the Aargau and Laegern cantons have been processed, others may be too. Then, individual tree detection can be applied. WSL will be able to then provide ground truth for individual trees within the cantons.

To view rasters in 2D, imagesc is often used. For a DTM, or other products that are generally smooth, surf may be used to view a 3D continuous surface. 3D surfaces are not generally constructed for looking at canopies due to pits and other variations (it will look noisy).

The masters students took a trip to SwissPhoto to learn about LiDAR work (apparently, a partner of the GIUZ).

9/26

We will be meeting with Christian Ginzler of the WSL on October 6.

Eidg. Forschungsanstalt für Wald, Schnee und Landschaft WSL

9/28

Spoke with Reik today during the coffee break to try to understand the data I’ve been looking at. (This includes LAS tiles for Kanton Aargau leaf on / leaf off, polygon data with the percentage of conifers and percentage of forest per polygon, raster statistics Matlab files containing canopy cover values). Some new information:

* The leaf on data was acquired in August, leaf off in March/April (so it’s not actually completely leaf off). In particular, beech trees maintain their leaves for most of the season, and only lose them once new leaves start to grow in. So the period of time where they do not have leaves is very small.
* The Aargau polygon data was created using inventory data, and it may not be very accurate. There exists ground truth with higher accuracy for Laegeren, although the forest type within that small study site is not necessarily representative of other forest types throughout all of the cantons.
* The canopy cover calculations involve relative differences between canopy cover of two seasons. This involves summing the number of points above a certain height threshold (for instance, 3m) and taking the ratio to the total number of points. However, this is sensitive to point density. The flight line overlap during collection varies between seasons, which results in different point density.
* The forest type map used a canopy cover difference threshold of > 0.35 to differentiate broadleaf from needle leaf types.
* He believes that crown delineation for broadleaf vegetation is truly impossible to do in a robust and accurate fashion. This is due to how the crowns grow and form together in dense forests. During leaf off periods, the stem-based detection methods may work well. But, during leaf-on periods, this is a problem.
* He suggests using the ORT imagery for visual inspection to find regions that are clearly broadleaf vs. needle leaf. Apparently, ORT imagery was collected simultaneously with the ALS data.
* Need to ask Felix what the goals are of the tree detection. Is it primarily stem location? Crown delineation? Neither, both?

Overall, Reik helped increase my awareness of the complexity and difficulty of using airborne laser scanning to measure vegetation structure. Many limitations exist, and the existing products / results should be approached with a critical mind. He encourages skepticism.

Gillian organized a PhD scientific coffee break, where current research topics, conference/journal dates and information, as well as computing resources were all discussed. Diego pointed me to ALS processing tools on Matthew Parkan’s github, which may come in handy. Many of the functions process the point cloud directly, although his watershed method processes the raster.

10/6 Meeting with Christian Ginzler and Felix at WSL Brimensdorf

Felix and I met with Christian at WSL (which stands for Eidg. Forschungsanstalt für Wald, Schnee und Landschaft) to discuss data availability and access. Our collaboration will be within the framework of the Swiss NFI (National Forest Inventory) / LFI (Schweizerischen Landesfortinventar), which has ground truth data available for validation of our methods. The maintain long-term plots and areas of study. Inventory data typically includes tree height, stem location, and species information. Uncertainties in these measurements varies.

ALS data exists for all Swiss cantons except for one (Lucerne. Apparently a large cost of 40-50 thousand CHF is being demanded, so who knows if/when that data will be acquired). Most of the data features leaf-off conditions. In addition, DTMs are available, and the data is already normalized. The cantons of Zürich and Zug (leaf on) data has a very high point density.

Other processing tools that exist to do individual tree detection include the [FINT](https://www.conftool.net/EARSeL-WS-Forest-2016-Krakow/index.php?page=browseSessions&print=head&form_session=39&metadata=show) (Find Individual Trees), which uses a raster-based (Normalized Surface Model) local maximum approach to identify dominant trees. Photogrammetry methods using passive imaging may be used to build up point clouds effectively.

The future of ALS forestry work includes using FPAs to detect LiDAR signals. Additionally, using UAS technology to save money and time (a 9 hectare area may be imaged in about 1 hour as opposed to an entire week using TLS). However, the systems are expensive (CHF 6000) and there are restrictions regarding altitude of flight and area of coverage.

We plan to meet regularly with Christian at the WSL or UZH every 5-6 weeks. In the beginning of next year (possibly January), I will plan to present my work and initial findings at WSL.

The first thing is seriously to differentiate between broad/needle leaf tree regions. Felix says that his Finnish colleagues believe strongly in using rasters of first-return data during the leaf-off season. We can test these methods using Aargau and also the Laegern site within it.

Christian gave me a forest type map of Aargau to be used as ground truth. (in TIF form, to be opened by qGIS).

17.10.16

Professor [David Coomes](http://www.plantsci.cam.ac.uk/directory/coomes-david), the head of the Forest Ecology and Conservation Group at the University of Cambridge spoke at a seminar at UZH today (part of the UZH URPP Global Change and Biodiversity Seminar Series). In his talk, *Ecology from 2000 metres up: what can airborne remote sensing tell us about life on earth*, he discussed some of the past and current motivations and methodologies for imaging vegetation with remote sensing technologies.

General science background: ALS allows for widespread, large-scale estimation of vegetation structure. Ecologists are concerned with monitoring species and changes in forests over time. A current global goal is more accurate estimation of carbon dioxide emission into the atmosphere due to deforestation, as well as carbon sequestration. Biomass can be estimated using metrics including TCH (top of the canopy height), BA (basal area), and wood density (WD, as a function of species). The transition from ABA (area-based approaches) to ITD (individual tree detection) methods and availability of high resolution LiDAR data has immensely increased the possible metrics, and accuracy, of remotely sensing vegetation.

[Greg Asner](http://globalecology.stanford.edu/labs/asnerlab/) (of [Stanford](https://cao.carnegiescience.edu/), famous within the ALS forestry community) is known for his LiDAR work in the Amazon. Asner argues that TCH is all that is needed for carbon estimates.

Coomes described the methods used by his group to estimate biomass, to identify individual coniferous trees. This is successful for the larger, bigger trees, and it has been shown that accurately measuring them is a decent proxy for biomass estimates. It has been shown that these trees are the predominant contributors of above ground biomass. But it yields a systematic underestimation of biomass. But it is hopeless for small, suppressed trees. Coomes claims that the next essential work in this field involves individual tree detection for suppressed trees in the lower strata of forests. This will lead to a more accurate assessment of basal area (which is apparently has the strongest relation to biomass, compared to other tree structure measurements).

The [SAFE](https://www.safeproject.net/) project is one of the largest ecological experiments in the world, studying the impact of humans on biodiversity and ecosystem function. The [FunDivEUROPE](http://www.fundiveurope.eu/) (fundamental diversity) project aims to promote understanding of tree species diversity.

Met [Florian Zellweger](http://www.wsl.ch/info/mitarbeitende/zellweger/index_EN), working at WSL, recently finished his PhD at ETH in collaboration with UZH. He recommends I also contact / look into the work being done by [Lars Waser](http://www.wsl.ch/info/mitarbeitende/waser/index_EN).

01.11.16

what MATLAB scripts do:

leafon\_leaoff\_pc\_stats.m

* Computes the boxplots (point-cloud based statistics) as a function of species (based on crown polygon information) for max, mean, median, std dev height

first\_last\_return\_heightdif.m

* Subtracts the last echo height from first echo height (per pulse) and creates boxplots for the average height difference per species (based on crown polygon information)
* Researchers at the Finnish Geodetic Institute have shown that this height difference is useful for classifying deciduous and coniferous tree areas.

14.11.2016 Felix Meeting

Regarding my processing (in place to classify tree type based on first and last echo height returns from LiDAR data only), there are still some questions related to the ideal raster size and morphological processing parameters. We need to talk with Christian to determine what is important in terms of the accuracy of deciduous vs. coniferous forest types (determining the accuracy of one type is inversely related to the accuracy of estimating the other). The forest type map that he gave us for the canton of Aargau was created based on ortho-rectified tif images. They may contain some differences/inaccuracies due to distortion introduced.

* How long will take to run the processing on the entire canton? Calculate the time for a single tile, then multiply. This is an important consideration for automating and expanding the method for a large area.
* Consider using MAT files instead of LAS, since the LAS reader is not efficient. Although there are not MAT files available for all cantons.
* What is the resolution of the map from Christian? 🡪 2m
* Compare the forest type products on a per-pixel basis. Subtract them for a difference.
* How much space do I have on my local machine? Will I need an external drive? 🡪 I have a 1TB hard drive with 860GB free currently.

17.11.2016

It seems that the forest classification based on LiDAR data alone may be more difficult than we thought. Classification data is available for the point cloud, but only ground and vegetation classes exist for the Aargau data. This means that areas with short vegetation or non-vegetation (such as buildings, not bare ground) will not be easy to identify.

My current processing overestimates the area covered by coniferous trees, compared to Christian’s method. I am curious, how accurate is his area map ? How often may this be generated? It would be great if a forest type map could be created using just LiDAR data, but my current method is not necessarily the most accurate, given the very small, limited area of ground truth (tree crowns within the Laegeren site).

We will meet with Christian next week to discuss how to move forward. I am excited to be implementing some individual tree detection methods. We need the code from Mathew Parkan for Geodesic voting.

21.11.2016

[Forest Facts and Figures in Switzerland](../papers/Forest+Report+2015.pdf)

From the WSL Forest Report 2015 and NFI/WSL 2011

Forest Area Statistics

* About 1/3 (31%) of Switzerland’s land surface is covered by forest
* Pure conifer forests account for 43% of forested area, broadleaf forests 25%
* 62% of forest area is covered with conifer forests, both pure (where 90% of the trees are conifers) and mixed (where 51 – 90% are conifers)
* 38% of forest area is covered by broadleaf
* Pure conifer forest has decreased by 8% since 1985, although it is still the most prevalent type of forest
* Pure broadleaf forest has increased by 5% since 1985, mainly at lower altitudes

Forest Functions

* Provide wood
* Store carbon dioxide
* Provide protection against avalanches and rock falls
* Clean drinking water
* Habitat
* Recreation

Forest cover & use

* Forests cover ¼ of the Swiss Plateau and ½ of the southern slopes of the Alps
* Most of the Swiss population live on the Swiss Plateau

Growing Stock

* Wood stock is the volume of stemwood from a particular forest area
* Most important capital for a forest owner, and accounts for the largest proportion of carbon stored in the forest

Inventory Practices

* Roughly 500 million trees grow in Swiss forests
* National Forest Inventory NFI uses a sampling method with 6,500 plots (each with an area of 2km2, 1.41x1.41km square) used to estimate inventory metrics for the entire country through extrapolation.
* 5 regions: Jura, Swiss Plateau, Pre-Alps, Alps, Southern Alps
* The height and diameter at breast height (DBH) of single trees within each sample plot are measured. On average, 13 trees are documented per plot, 84,500 trees total.

Species Diversity

25.11.2016

Christian from WSL met with Felix and I today at UZH. We discussed the initial results of my classification work and how to move forward.

The WSL forest type map was created using NDVI data during winter. (Areas with a high NDVI are coniferous.)

* How to classify/remove power lines from the ALS point cloud?
  + Standard deviation of intensity – distinctively high compared to forest.
  + Identify pixels containing power lines then remove points prior to segmentation
  + Research classification schemes
  + LAStools, cloudcompare
  + Test using Aargau
  + Zürich ALS PC data is fully classified using TerraScan software
* Local Maxima search using all remaining points, and implement a height threshold for local maxima (>3m is considered a tree in Switzerland)
* WSL forest type map was created in 2015 using data from 2012/2013. There has been a shift from coniferous-dominated to broadleaf forests for the last 15 years. Logging is going on as well, which could contribute to changes in forest cover. But 90% of the area should not change.
* Move forward with ALS-only forest classification.
  + Use ortho photos as ground truth. May be some spatial mismatch.
  + May need to train a neural network, consult Beni / Devis, or support vector machine
  + Tune thresholds of the algorithm
  + Incorporate number of single echos, intensity of single echos
* Data from Christian
  + Leaf-Off, Leaf-On Color Infrared Aargau imagery
  + Leaf-Off LiDAR CHM with Buildings classified
  + Leaf-Off LiDAR not normalized with Buildings classified
  + Aargau forest type mask for entire canton
* Two-stage segmentation
  + Local Maxima processing to do first segmentation, yields clusters of just conifer trees. Remove or classify these points.
  + Another segmentation round for broadleaf trees
* Use forest type map to assign areas for local maxima method identification (with a height threshold of 3m), then use all of the points for the segmentation of coniferous trees.
* CHM 🡪 local maxima 🡪 assign as coniferous or deciduous 🡪 PC-based clustering
* Output: labeled point cloud clusters, groups of points that belong together for each tree identification number, presumably single tree crowns.
* Methods exist to compute statistics for each cluster (center of gravity, max height)
* Borders
  + Segmentation becomes inefficient for large areas
  + Use small tiles with overlap
  + If a local maximum at tile border, make the hard cut based on its position and the buffer distance
  + Segtree function
* It would really be ideal to do this forest type classification with only ALS data. Although data fusion is great, the availability and collection conditions may not be ideal with spectral data.
* Validation: NFI number of stems per hectare and species per tree. We can compare the number of trees detected using the LiDAR data over the same 500 square meter plot.
* CHM created on sloped terrain will lead to over-estimated height. Need to correct for different height at different parts of the crown
  + Martin Isenberg commented on one of Greg Asner’s posts about the tallest tree in Borneo
  + Use Digital Surface model instead of CHM, no normalization error. Use 3m threshold for CHM.
  + Use normalized point cloud to identify buildings, remove these points from the non-normalized points.
  + Hossein tested single tree detection at Laegern but was not happy with the highly sloped regions. Crown geometry would be changed and crown segmentation may not work as well after normalization, so the local maxima processing should be done with non-normalized point cloud data.
  + Difference between normalized points (which look oblique) and non-normalized points at Laegern and the resulting tree segmentation?
* Single tree detection at such a large scale is very innovative! Not published or being done widely. Canton Aargau has \_\_\_ trees! … dominant trees? Or trees higher than 3m? Highest trees in the canton are located\_\_. If LFI data validates it? Number of trees estimated by the NFI. Suppressed and understory trees will be difficult to achieve. Hossein’s method iterates for suppressed trees but there are physical limitations (regarding how far the laser signal can reach into the canopy).
* Our next meeting: January

28.11.2016

* I must research other methods that classify forest type using only LiDAR data.
  + Many methods upon first look use linear discriminant analysis (LDA). Felix does not have experience with this, but it’s possibly implemented in MATLAB and I may also want to consult with Beni, Devis, Diego, etc.
  + Graph-based methods?
  + Normalized cut methods?
  + Connie Ko’s classification methods
* Research individual tree detection methods for deciduous species, ideally with published code that is available to us.
  + Matthew Parkan’s code is no longer available for use. Plus, it requires extremely high point density (20-40 pts per square meter). Although we have this for Laegern, this is not available for other areas and cantons.
* Ask Hossein for his code? His approach is iterative and good for multi-layered deciduous forests.
* Read the last review paper by Felix and Hossein, which covers many methods (IEEE)
* Laura Duncanson’s work in RSE 2012-13 may be applicable. Felix knows her and has given her data in the past, so hopefully she would be willing to share code with us.

Felix / Christian questions:

* Felix
  + Send me the lecture slides?
* Christian
  + When exactly was the data collected for the LeafOff\_CHM\_Buildings data?

6.12.2016

I am experimenting with some machine-learning based classification methods. As the threshold-based algorithm (using height difference) must be tuned manually, I am hoping that using an automated classifier may be more accurate (and less manual work / decision making) for the forest type classification method.

Random forest classification takes an input of samples each with a set of features. I am treating each pixel as a sample and currently using 6 features (median height difference, number of single echos, mean intensity of single echos, max height, mean height, and standard deviation of height within a raster cell). The true class labels were assigned based on the WSL classification image. I am using 3 classes: broadleaf, coniferous, and non-vegetation. Error is going to be introduced since the WSL map was created using data from 2013 and the LiDAR data is from 2014. By using the same number of samples per class, the accuracy of the random forest prediction looks more accurate.

Beni has suggested trying the automated feature selection code developed by Devis, although it was created for multispectral imagery, so its performance with LiDAR data is unknown. I may also apply a Markov random field to reduce some of the noise in the output.

Also, I want to assess the importance of the features I chose for my initial Random Forest classification input.

Literature – ALS forest type classification

* Brandtberg (2007, Sweden) Classifying individual tree species under leaf-off and leaf-on conditions using airborne lidar
  + Directed graph (digraph) of possible binary interactions of laser pulses with canopy, used for species classification

7.12.2016

* For the Random Forest machine learning approach, I am currently working to make the Markov random field code work to denoise the output classification map.
* The CSPottsPixel function computes pairwise interactions between a pixel and its four neighbors. The sigma\_rbf and lambda parameters impact the smoothness. The label cost can assign a penalty for certain classes to be next to one another (or in this case, it is more likely that the center pixel is surrounded by pixels of the same class).
* There are NaN values in feature 5 of the input (Xnew), have to figure out why!
* By evaluating the relative importance of the features, they can be modified/added/removed. Display them using bar(Mdl.OOBPermutedPredictorDeltaError). Look at the image processing machine learning slides from Beni for more ideas on complex features (such as using the pixel neigbors).
* Need to normalize the feature data so they all have the same range (0,1).

Christian stopped by and left a hard drive with additional LiDAR data (including leaf off, not normalized). He will leave the drive here. I need to return it at our next meeting in January.

I must work on the power line removal from LAS data. Also, look into Laura Duncanson’s method and see how complicated it is compared to Hossein’s method for ITD in deciduous areas. Christian wants to know it if would be possible to create more individual tree crown ground truth data (like the polygon map) for more regions in Aargau?

13.12.2016

Something that we must consider (also Danny and Fabian) is the increased point density in regions where flight lines overlap. Possibly using a normalization cell that extends beyond the data cell.

I corrected for the overlap point density using lastools’ lasoverage function, but then realized that the thresholds are no longer effective. And also, we could not process the entire canton without a LAStools license. So there needs to be a different way to do this in MATLAB.

22.12.2016

Referenced *Physically-based multi-temporal LiDAR traits for species discrimination in a temperate mixed forest* (2016) by Morsdorf, Bruggisser, et al for some specific info about the Kanton Aargau ALS data:

The Kanton Aargau data was collected in March/April (“leaf off”) and June/July (“leaf on”) of 2014 using the Riegl LMS-Q680i instrument flown at al altitude of 600 and 700m, respectively. The pulse length was approx. 4 ns and the echo density was approximately 15 and 30 points/m2, respectively.

Kanton Aargau vegetation includes the Laegeren site, approximately 15km northwest of Switzerland, on the border of the Swiss Plateau between the Jura and the Alps. Mixed beech forest with Norway spruce.

23.12.2016

Felix meeting –

Spending a lot of time on these pre-processing steps will hopefully save time later when we attempt to upscale the tree detection.

Laura Duncanson will be visiting in April. It may be a good opportunity to work with her, perhaps she will contribute her deciduous multi-layer tree detection code to us or become a collaborator.

* Forest Type classification
  + **Random forest** shows promise – keep working on it
  + **Noise removal** using MRF or even morphological processing
  + First do a **height threshold** to remove points (under 2m? 3m?) that are non-forest. Set no data value to 0? Then use the building classification data from Christian to remove the buildings. Should be left with just forested areas to do the coniferous vs. broadleaf classification.
    - Look at Reik’s 3D image – it shows that
* Copying over the 3D images of Aargau (one is a summertime RGB orthophoto overlaid onto the point cloud, the other illustrates height differences and standard deviation of features on the landscape) to the
* **Try to filter out the power lines by assessing points with high standard deviation!!!** 
  + Clearly visible in the 3D image rom Felix:

osgviewer /Volumes/FastDisk/RGB\_AG/AG3D.ive (in terminal)

* + Then, try the Finnish Geodesic Institute method only if that fails (due to the high computational cost)

5.01.2017

To copy files directly from laptop to mac, first get the IP address of the mac.

Type **ifconfig**in terminal on Mac for an address with this form: inet 130.60.16.182

On laptop, from terminal:

**ssh scholl@130.60.16.182**

Enter password. That’s it! To transfer files, use scp from terminal.

scp fileToCopy [scholl@130.60.16.182:/Path/to/file](mailto:scholl@130.60.16.182:/Path/to/file)

09.01.2017

For the RF classifier, I wanted to add vertical histogram data. 50 bins per pixel. Devis says, just add 50 more features ! To reduce the number, consider using PCA or taking the mode of the histogram.

Devis suggested that I sit in on his class in the spring semester (448) Thursdays 10:15-noon

Meeting with Felix

The power line problem seems to be more complex than Felix and I had anticipated.

I tried with my binning approach but it is difficult to separate the power line bins from the tops of trees. I used the standard deviation raster to locate pixels that potentially contain power lines. However, there are many forested areas (without power lines) with comparable standard deviation.

Some papers (3d classification of power-line scene from airborne laser Scanning data using random forests, by, using both voxel- and sphere-based volumetric approaches to create input for RF classifier) use methods such as the Hough transform to identify power lines based on their linear shape. However, the Hough transform requires binary 2D form and is best for finding connected, continuous lines (which are not present in the lidar and derived raster products).

But removing power line points is important to help improve the accuracy of the tree detection. I will try using a k-means clustering method with 3 classes (power line, trees, and ground) and setting specific seed points with locations. Perhaps then the power lines will be isolated?

AOI 2 (with the extremely high power lines) contains an interesting artifact as a result of the DTM normalization. The power lines get disconnected and indicate an inaccurately high estimate of the DTM in that location, so the trees are also higher than they should be in this ground-normalized portion. This is a reason to use the non-normalized ALS data for the tree detection. Felix can use this spot as a question in his LiDAR remote sensing class ☺

Swiss Photo was contracted to scan Zurich but a German company did the canton of Aargau.

I can use majority voting (basically, a filter to keep the value with the majority inside a 3x3 or 5x5 window) to reduce noise in the RF output for forest type classification.

Felix says to be mindful of the bin resolution and range when utilizing a vertical histogram as a classifier input. For instance, if the bin values are absolute (as opposed to being relative based on the local forest), then there may not be enough vertical resolution to distinguish structure in younger forests, in early stages of development. But a 1m resolution histogram for my forest type classification purposes is likely to be enough.

10.01.17

Kmeans is not good for separating classes that are not spherically separated in 3d space.

But it is for vertical assessment in a 2d pixel!

12.01.17

From WSL NFI flyer publication:

<https://www.lfi.ch/publikationen/publ/LFI_Flyer-en.pdf>

* NFI uses aerial photo interpretation to identify forest/non-forest regions, broadleaf/coniferous, trees/bushes or grass/buildings/roads/bodies of water/ snow and their height.
* There are about 500 million trees in Switzerlad
* 6500 sample plots with 1.4km grid mesh size.
* Each sample plot is visited (over a period of 9 years), many attributes are surveyed, an average of 12 trees are measured per sample plot.
* Forested area is growing by 0.3% every year