APPENDIX D
Mapping Forest Biomass with Radar Remote Sensing – Chapter 5 Training Module

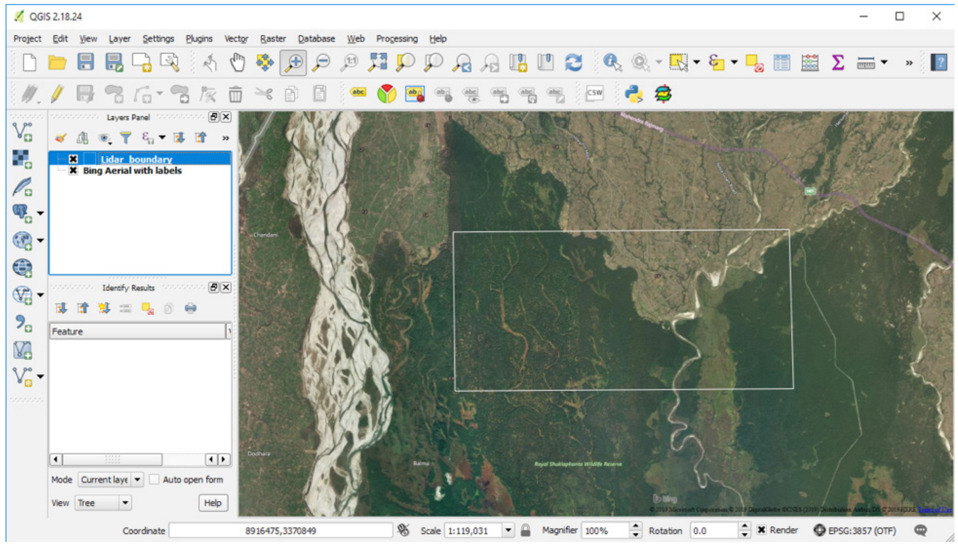


Figure 1.1 *QGIS interface displaying the study area in western Nepal. The white line represents the lidar boundary.*

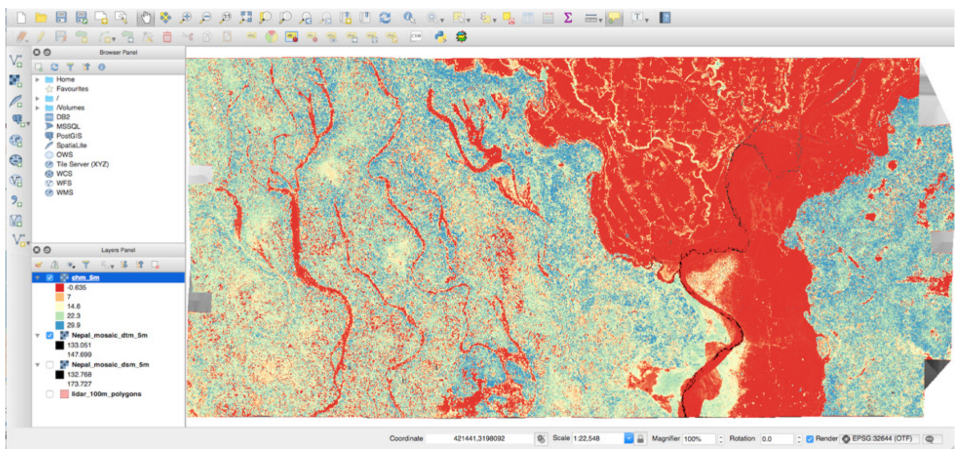


Figure 1.2 *Canopy Height Model created using the lidar-derived DSM and DTM. The image shows the Spectral color band stretched using min/max values.*

1 SOFTWARE AND DATA SOURCES

Software:

- QGIS, Microsoft Excel

Remote Sensing Data:

- ALOS 2 / PALSAR -2 RTC annual mosaic product
- Lidar data that overlaps with some of the ALOS 2 / PALSAR -2 data
- Forest inventory plot data that overlaps with some of the lidar data.

Note: All of the data to run the tutorial is included in the data.zip file for this chapter hosted on the SAR Handbook website. If you would like to do the same processing on your own data, you will need a SAR RTC product (steps for downloading from ASF or JAXA can be found in the tutorial for Chapter 6), lidar data covering part of your area of interest, and forest inventory data that overlaps with the lidar data.

2 AIRBORNE LIDAR AND INVENTORY PLOT DATA

Download the Data.zip file for Chapter 5 from the SAR Handbook website and unzip it in the desired location. The zipped file consists of three folders: Lidar, Ground_Plots, and ALOS with example data from Nepal.

Step 1: Open QGIS and add the shape file D:\NepalData\Lidar\Lidar_boundary. Add Google Aerial with Labels as the background (Web > OpenLayers plugin > Google Maps). Examine the location of the Lidar_boundary file to understand the study site location and the landscape across Western Nepal.

Step 2: Add the LIDAR DTM (Lidar > dtm_5m.tif) and DSM (Lidar > dsm_5m.tif) files to QGIS (Layer > add Layer > add Raster Layer). The images are provided as geotiffs with 5 meter spatial resolution.

Step 3: In this step, we will produce a Canopy Height Model (CHM) by subtracting the DTM from DSM. Using the raster calculator tool in QGIS (Raster > Raster Calculator), enter the following equation:

"Dsm _ 5m" – "Dtm _ 5m"

Since will be using this file multiple times, you may want to create a results folder and save the CHM to this new folder under the file name chm_5m.tif.

Step 4: Double click the CHM image name (chm_5m) in the layers panel of QGIS and explore the Layer Properties, including projection, display, and other image characteristics. The projection is in UTM Zone 44 N, Datum: WGS-84.

Step 5: From here you can also apply a color scale to the CHM that will highlight short to tall forests in the study region. With Layer Properties still open, go to Style > Render type > Singleband pseudocolor > Load min/max values > Min/max > Load > Color > Spectral (or any other color scale you like) > Apply. Your result should look similar to **Figure 1.2**.

Step 6: Use the Profile Tool to explore height distribution in the data. The Profile Tool is a Plugin and needs to be installed by going to Plugins > Manage and Install Plugins > Profile Tool > Install Plugin. After installation, the Profile icon  will appear in the toolbar. Click on the Profile icon > Select the CHM image in the Layers Panel > Click Add Layer > Draw a line at any place over the image (double click to end the line). Depending on where you draw your line, your result should look something like **Figure 1.3**.

Note: This image shows a typical example of a CHM profile which can be achieved drawing a line. The profile image can be saved for future use. Also be careful in interpreting your profile chart. It will start wherever you draw your line and move in whatever direction you ended your line. If you draw your line from east to west (instead of west to east), your profile will start in the east and move toward the west. In Figure 1.3, the line was drawn from east to west; therefore, the right part of the profile starts in the east and moves westward.

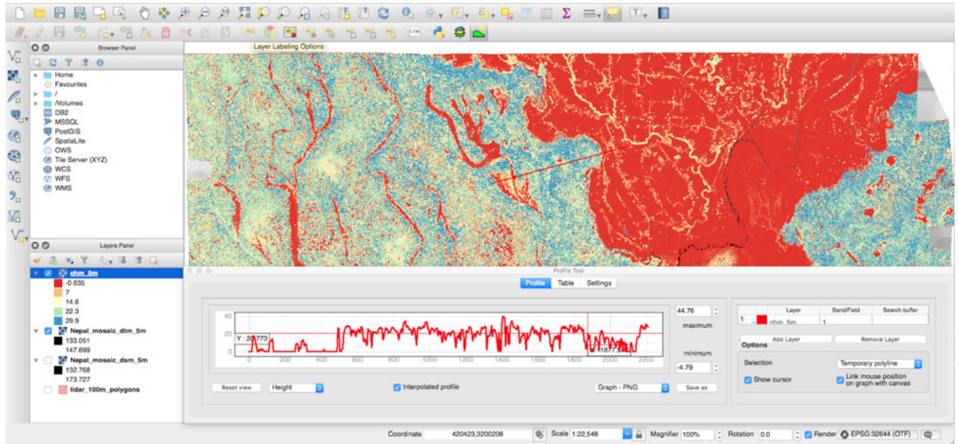


Figure 1.3 *Results of the profile tool showing canopy height variation across the red line drawn in the center of the scene.*

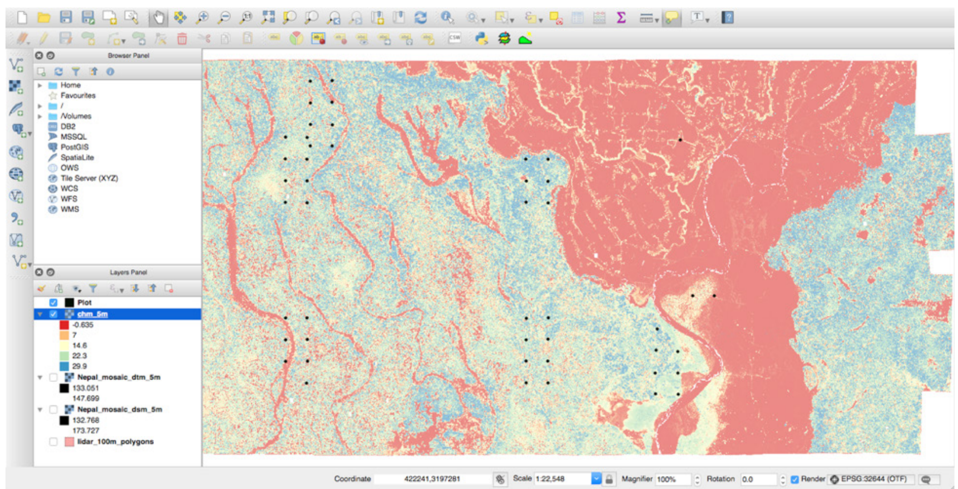


Figure 1.4 *Distribution of field plots (denoted by black dots) across the study area.*

Step 7: Play around with drawing different lines across different parts of the scenes. Where do you see the highest canopy height? The lowest? How does this pattern change across the landscape?

3 LIDAR BIOMASS MODEL DEVELOPMENT

Step 1: Add the shapefile containing the ground plot data (Ground_Plots > plot.shp) to the QGIS (Layer > Add Layer > Add Vector Layer). There are 47 small plots available for the study area. Each plot represents a 20 m radius ground footprint.

Step 2: In this step we use the Zonal Statistics Tool to extract the lidar-derived mean canopy height from each plot. Go to Processing > Tools > Search for Zonal Statistics. Once you open up the Zonal Statistics Tool, set your raster layer to chm_5m.tif and your vector layer to plot.shp and run the tool.

Step 3: A new shapefile (named Zonal Statistics) will be added to your Layers Panel. You can right click the new layer and select Open Attribute Table to view the data associated with each plot. If you scroll all the way to the right of the attribute table, you will see the zonal statistics you just calculated. Most important for the next steps is the _mean column, which contains the average canopy height for each plot.

Step 4: Next, you want to create an Excel spreadsheet from the Zonal Statistics attribute table. One method you can use is to to install the XYTools plugin (Plugins > Manage and Install Plugins > search for XYTools > Install Plugin. Next, make sure the Zonal Statistics layer is highlighted in your Layers Panel. Go to Vector > XYTools > Save attribute table as Excel file. Check the following fields: Object ID, AGB, _std, and _mean. Alternatively, you can right click the Zonal statistics layer in the Layers Panel and select Save As to save the data as a csv file,

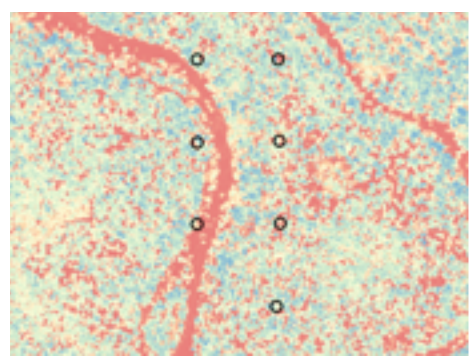
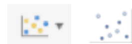


Figure 1.5 *A zoomed view of one set of field plots. Later in the exercise, we will compute the average of the pixel values that fall within the field plot set using the zonal histogram tool.*

OBJECTID	_mean	_std	AGB
1	7.85	7.49	97.77
2	16.91	10.45	290.54
3	18.54	9.26	263.21
4	18.09	8.51	163.86
5	9.73	9.12	172.78
6	14.20	8.93	210.74
7	10.09	11.72	68.58
8	19.36	5.97	176.68
9	18.65	4.99	277.04
10	16.30	7.34	146.01
11	15.67	10.48	268.26
12	14.44	11.25	250.54
13	23.47	4.73	693.69
14	12.61	10.09	103.73
15	20.79	5.25	325.56
16	12.06	9.66	109.76
17	19.56	7.90	810.71
18	23.26	3.70	687.63
19	19.08	6.65	293.98
20	17.72	7.76	399.16
21	20.73	4.74	258.01
22	23.15	7.44	591.53
23	24.49	2.96	579.44
24	24.98	5.60	530.78
25	18.08	5.29	344.34
26	25.93	5.96	657.66
27	21.93	6.20	764.19
28	21.94	5.97	441.79
29	25.34	6.43	582.33
30	16.16	8.53	238.40
31	23.44	5.13	674.51
32	10.39	7.11	196.14
33	20.86	8.31	249.41
34	17.41	9.66	209.03
35	17.85	6.91	243.84
36	11.23	11.06	257.55
37	29.32	4.51	858.85
38	20.25	5.90	139.54
39	22.96	5.01	412.32
40	21.82	5.74	667.84
41	21.78	4.40	287.50
42	19.47	1.95	233.49
43	19.46	4.95	330.82
44	20.05	1.84	395.18
45	3.88	3.39	63.51
46	8.17	4.75	138.14
47	3.11	2.41	80.94

Table 1 *Plot values for _mean, _std, and AGB.*

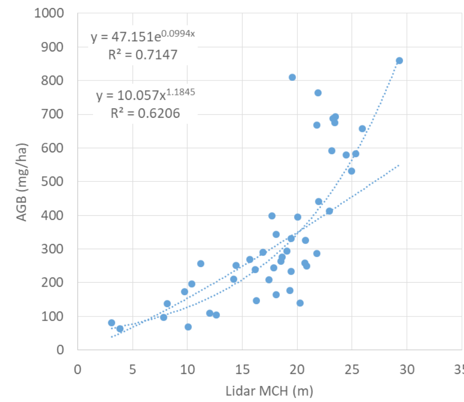
which can be opened in Excel. The values in your table should look similar to those in **Table 1**.

Step 5: Open your Zonal Statistics data in Excel. Here we want to plot the aboveground biomass (AGB) of the plots with respect to the MCH (_mean) values. You can do this by selecting the data in the AGB and _mean columns and creating a scatterplot using the scatterplot tool (Insert > Scatterplot )

Step 6: Now we want to fit the best model to present the data. You can do this by right clicking one of the points in the scatterplot and selecting add trendline. From the format trendline pane, you can evaluate various trendline options for the best fit and display the equations that go with each trendline. In this case, the best model is a power-law. However, as the plots are small and the sensitivity of height to capture the high biomass values of small plots may saturate, use other functions such as exponential (as shown in the inset graph on this page). During the training given at SERVIR-HKH, there was a consensus among the participants with local knowledge for limiting the maximum biomass in the study site to 1000 Mg/ha.

Step 7: Here we create a CHM image of 40 m pixel resolution by performing an 8x8 resampling of the chm_ 5m image (There are many ways to do this in QGIS, but one is to go to Processing > Toolbox > SAGA > Raster tools > Resampling, select chm_5m as your Grid, leave upscaling and downscaling method as Nearest Neighbor, make the Cellsize 40, and run the tool). Save this file as chm_40m. This process will produce a 40 m average MCH (mean top canopy height) image.

Step 8: Now we are going to create an AGB map for the entire area covered by LiDAR data by using the relationship we identified between the plot AGB and the LiDAR canopy height. Using Raster Calculator (Raster > Raster Calculator) apply the best fit equation you derived in step 6: 47.151 * exp(0.0994*MCH) to the 40 m resolution LiDAR MCH data to develop AGB map from the Lidar image as shown below. Save the result as LiDAR_agb_40m.tif

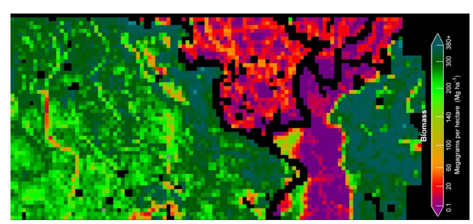


a. Note: MCH is the resampled lidar map created in step 6 (chm_40m).

b. Note: The Raster Calculator tool in QGIS does not have an "exp" function; therefore, you can replace the equation with: 47.151 * (2.718282 ^ (0.0994 * MCH)).

Step 9: The output of step 8 will be a AGB map that covers the same area as your LiDAR data. Now we want to resample the LiDAR AGB map from 40 m to100 m (1-ha) spatial resolution (There are many ways to do this in QGIS, but one is to go to Processing > Toolbox > SAGA > Raster tools > Resampling, select LiDAR_agb_40m as your Grid, leave upscaling and downscaling method as Nearest Neighbor, make the cellsize 100, and run the tool). Save the output as LiDAR_agb_100m.

Step 10: Display the final image and apply a color scale as part of QGIS color ranges from dark red to green from low to high biomass (below).



Step 11: Refer to Chapter 5 for forming the LiDAR biomass models and the uncertainty depending on the plot size and LiDAR pixel size. In this exercise, a simple method was used to develop the model. The sources of uncertainty and the quantification and propagation of errors are discussed in more detail in the chapter.

4 ALOS BIOMASS MAPPING (NEPAL)

In this exercise, the LiDAR estimated biomass map will be used to train the SAR ALOS image to develop AGB map from the ALOS data for the larger study area.

4.1 Radar Processing

Step 1: In this exercise, we use the LiDAR derived biomass map (lidar_agb_100m) as the reference data to develop a model for radar estimation of biomass. The inventory plot data are small and are not suitable to extract data from ALOS PALSAR data. Therefore, we use the 1-ha resolution LiDAR based AGB map for both developing model and testing the results.

Step 2: For this exercise, the ALOS 2/PALSAR 2 annual mosaic for 2015 is provided in the folder Data.zip folder (ALOS). Data are in HH and HV polarizations. For practice, you may want to download the ALOS PALSAR data for 2015, 2016, 2017 directly from the JAXA website. See steps for downloading in the Chapter 6 Training Appendix.

a. Note: If you download the imagery directly from JAXA, the website provides data in a grid of 1-degree tiles; you will need to select N29E080 for this site.

b. Note: If you compare the images for 2015 – 2017, you will notice some temporal variability in the backscatter due to variations in environmental conditions such as soil moisture or phenology. In this exercise, we will use data for 2015.

Step 3: Open the N29E080_15_sl_HV_F02DAR and the N29E080_15_sl_HH_F02DAR files (Data > ALOS) in QGIS. Note that these ALOS RTC annual mosaics are ready to use at source. Radiometric terrain correction and precise geometric corrections have already been performed.

Step 4: The backscatter data in the ALOS RTC annual mosaic comes as a digital number (DN) and needs to be converted to gamma naught dB for analysis. To convert the DN to dB values apply the following equation using Raster Calculator (Raster > Raster Calculator):

$$\text{Gamma\_dB} = 10 * \log_{10} [(DN)^2] – 83.0$$

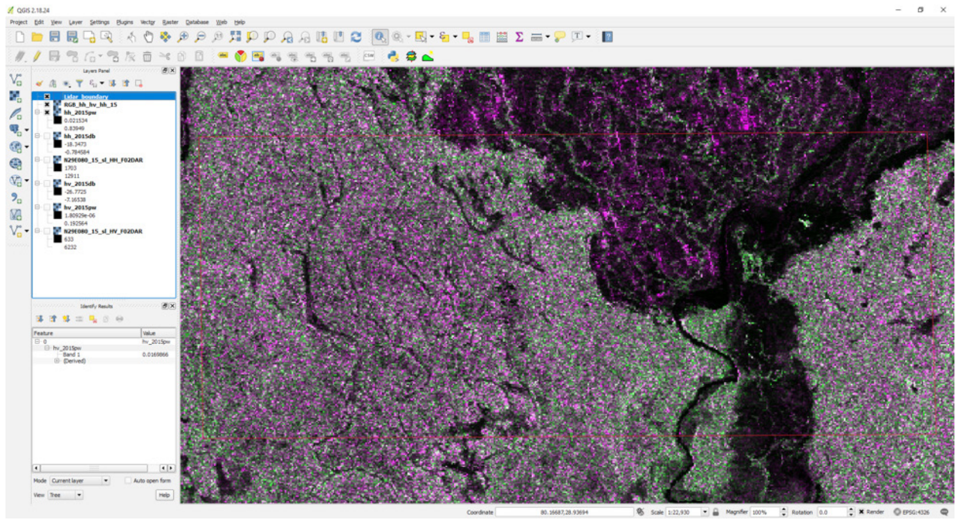


Figure 1.6 *SAR RGB image (HH, HV, HH) derived from ALOS PALSAR data.*

Next, convert the dB values to power backscatter by applying the following equation using Raster Calculator (Raster > Raster Calculator):

Gamma_pw = 10^(0.1*Gamma_dB)

You will need to do this step for both the HV (N29E080_15_sl_HV_F02DAR) and HH (N29E080_15_sl_HH_F02DAR) images.

Save the Gamma_pw result as gamma_pw_HV (or gamma_pw_HH for the HH polarized data). At the end of this step, you should have created two new files, gamma_pw_HV and gamma_pw_HH.

Step 5: Now we will create an RGB image for visualizing the backscatter power in color. You may consider a three-band composite, where R: HH (gamma_pw_HH), G: HV (gamma_pw_HV), B: HH (gamma_pw_HH). You could also use a ratio of HV/HH as the blue band (Calculate the HV/HH ratio using Raster Calculator). To create a multiband image, go to Raster > Miscellaneous > Merge > Edit > Type. Another option is to copy and paste following gdal command into the Edit box. Note that your data folders may be different:

gdal_merge.bat -ul_lr 80.0 29.0 81.0 28.0 -separate -of GTiff -o D:/Data/Results/RGB_hh_hv_hh_15.tif D:\Data\Results\hh_2015pw.tif D:\Data\Results\hv_2015pw.tif D:\Data\Results\hh_2015pw.tif

Since the HV polarization is most sensitive to forest structure, areas that have high backscatter in HV (showing up as green in **Figure 1.6** are likely to have higher AGB values as well.

4.2 Radar Biomass Model

Next, we produce samples from Lidar data to compare with radar measurements and develop a best-fit model. Refer to Chapter 5 for more detail on how to choose the appropriate LiDAR samples and issues related to the differences in date and the changes that occur between radar and LiDAR data. Any changes of landscape can easily introduce large discrepancies between LiDAR derived biomass and radar backscatter measurements.

Step 1: Open the resampled LiDAR biomass map at 1-ha (lidar_agb_100m). We can create a random or systematic sample dataset. To facilitate a systematic sample, we created a shapefile with horizontal and vertical polygons (lidar_100m_polygons.shp) which we used to create a systematic sample of points (lidar_systematic_sampling.shp). Open "lidar_100m_polygons.shp" (Data > Lidar) in QGIS.

Step 2: Use the lidar_systematic_sampling points to extract all of the 1-ha values from the LiDAR biomass map(lidar_agb_100m) (Processing > Toolbox > SAGA > Vector > Raster > Add Raster values to points) and save the data as a shapefile with the file name lss_100m_agb. This sampling strategy produces 1822 sample points that we will use to develop the model and quantify the uncertainty.

Step 3: Since the LiDAR based biomass data represent plot size of 1-ha, now we are going to resample the power backscatter radar data into 1-ha (100m) pixel size to match. In QGIS, resample the gamma_pw_HV and gamma_pw_HH files to 100m (Processing > Toolbox > SAGA > Raster tools > Resampling). Save these files as gamma_pw_HV_100m and gamma_pw_HH_100m.

Step 4: Next, we extract the resampled HH and HV power backscatters (gamma_pw_HV_100m and gamma_pw_HH_100m) from the radar images to the sampling points (lss_100m_agb) (Processing > Toolbox > SAGA > Vector > Raster > Add Raster values to points). Save this result as lss_100m_agb_sar. When you save the shapefile, it will also create a .dbf file, which easily can be opened in Excel.

Step 5: In Excel, open the lss_100m_agb_sar.dbf file. Before doing any analysis, we need to clean our data by removing all rows that have lidar AGB values that are negative or equal to zero (likely representing water pixels or erroneous data) from the spreadsheet. We also need to eliminate any rows where the lidar AGB value has missing data or no data (NAN, -9999, or 9999). After all cleaning (removing rows where AGB is zero or bad points), you should have a spreadsheet with 1649 data points.

Step 6: Now create two scatterplots, one that shows HH vs. AGB and a second that shows HV vs. AGB. You can use the same methods described in Section 3, step 5. Note that you should see a large spread of values, in part due to differences between the lidar and ALOS PALSAR acquisition times. Since the data collection did not occur on the same date, there may be some land cover or soil moisture change that could cause error in your model. Additional issues could include georeferencing discrepancies, topographical effects, errors due to speckle, and potential incidence angle variations. Be sure to consider these limitations as you work to improve your model and interpret your results.

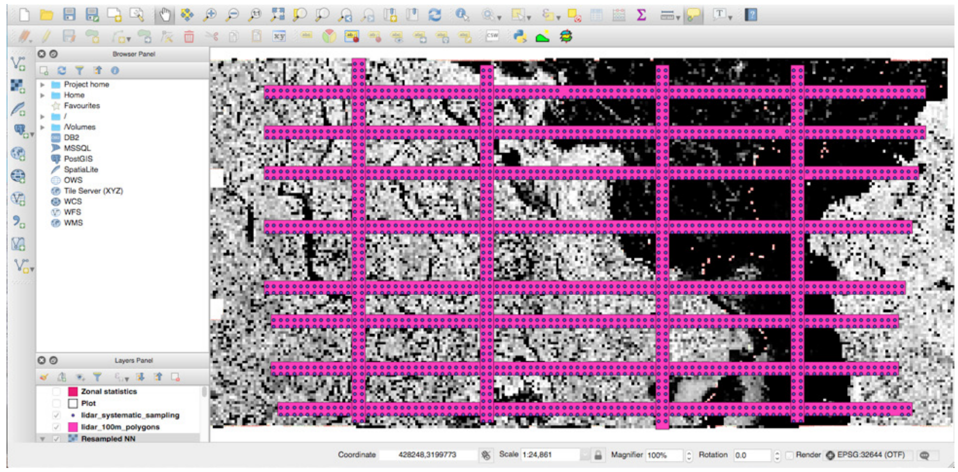


Figure 1.7 *Horizontal and vertical polygons (pink) used to create the systematic sample points (dark blue) as inputs to generate the AGB model are displayed.*

Step 7: Fit a logarithmic or a power-law to the both HH and HV SAR data to see the strength or weakness of the relationship between radar backscatter and lidar derived biomass (See section 3, step 6).

4.3 Radar Biomass Mapping

Step 1: In the previous section, we looked at the relationship between AGB and HH and HV backscatter. In this section, we focus on the relationship between AGB and HV backscatter only, as HV polarization has the strongest sensitivity to biomass. However, other radar polarization measurements and model fits are discussed in the text of Chapter 5.

Step 2: In section 4.2, step 7, we generated a best-fit model based on a power-law: AGB = 57696*(HV^2.0042). We will use this equation to model biomass from backscatter. Using Raster Calculator (Raster > Raster Calculator), apply this equation to the HV backscatter image at 100 m spatial resolution (gamma_pw_HV_100m). Save this file as HV_biomass_100m. Note that the equation is developed from 1-ha (100m) LiDAR derived map and should only be applied at the same resolution radar image. One cannot apply this equation to any other resolution (smaller or larger) radar image without introducing additional uncertainty.

Step 3: Evaluate the saturation in this model. Although the fit shows no saturation, however, the data shows

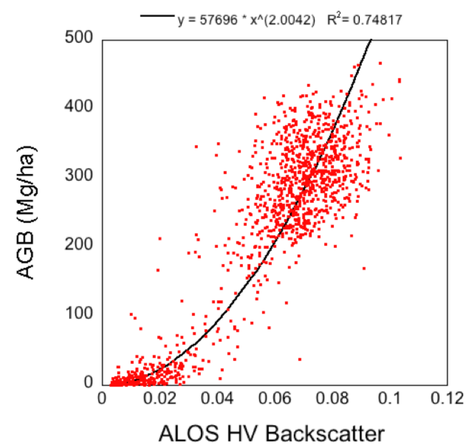


Figure 1.8 *An example of a best fit model between AGB and HV backscatter.*

that HV backscatter has almost no sensitivity to biomass above 200 Mg/ha for these forests (See Saatchi et al. 2011, or Saatchi et al., 2007 for other alternative equations and saturation of the radar data).

Step 4: Display the map of forest biomass from the 2015 ALOS PALSAR 2 image (HV_biomass_100m) using a color range to show the variation of biomass across the image. Step 5: Read values of the biomass from the image and visually compare it to the reference biomass map derived from the lidar image (lidar_agb_100m).

Step 6: Mask out all pixels above 200 Mg/ha to show that the map has large uncertainty over areas of above 200 Mg/ha and cannot be trusted (Raster > Raster

Calculator > "HB_agb_100m >=200" to create mask). Although you can leave the map untouched by explaining the fact that the map has large uncertainty in areas where AGB is > 200 Mg/ha.

Step 7: Clip the radar AGB map (HV_agb_100m) to the same extent as the LiDAR AGB map (lidar_agb_100m) using Raster > Extraction > Clipper in QGIS. Save this file as HV_agb_100m_clp.

Step 8: Here we are going to calculate the percent difference between the clipped radar-derived AGB map (HB_agb_100m_clp) and the lidar AGB map (lidar_agb_100m) using the raster calculator (Raster > Raster Calculator). Note that the lidar and backscatter AGB maps should be the same size for this step. Use the following equation in Raster Calculator to calculate the percent difference in AGB:

$$\text{Diff} = 100^* (b1-b2)/b1$$

Where b1 is the LiDAR map (lidar_agb_100m) and b2 is the radar map (HV_agb_100m_clp).

Step 9: Display the difference map in percentage and provide a color range to show the range of values and include the color range on the side for presentation of the results (**Figure 1.10**).

4.4 Improving the AGB Map

Here, we improve the radar biomass model and AGB mapping by using multi-temporal radar imagery. In an ideal scenario, ALOS PALSAR data from different seasons and over time from the same or multiple years can be downloaded and used to reduce the effects of soil moisture and phenology and improve biomass mapping.

Step 1: Download the ALOS PALSAR mosaic data for 2015, 2016, and 2017 from JAXA website (You can use the data included in Data > ALOS or see the Chapter 6 training appendix for information on downloading ALOS PALSAR data from the JAXA website).

Step 2: Using the steps provided previously, for each year, calculate the gamma power of the HV backscatter (Section 4.1, step 4) and resample each image to 100m (Section 4.2, step 3). At the end of this step, you should have three HV backscatter images resampled to 100m: one for each year.

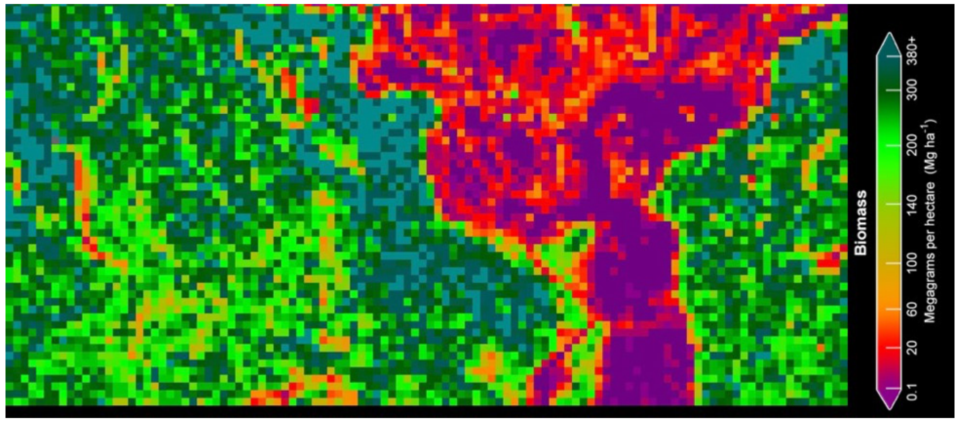


Figure 1.9 *Example results showing variation in biomass derived from HV backscatter.*

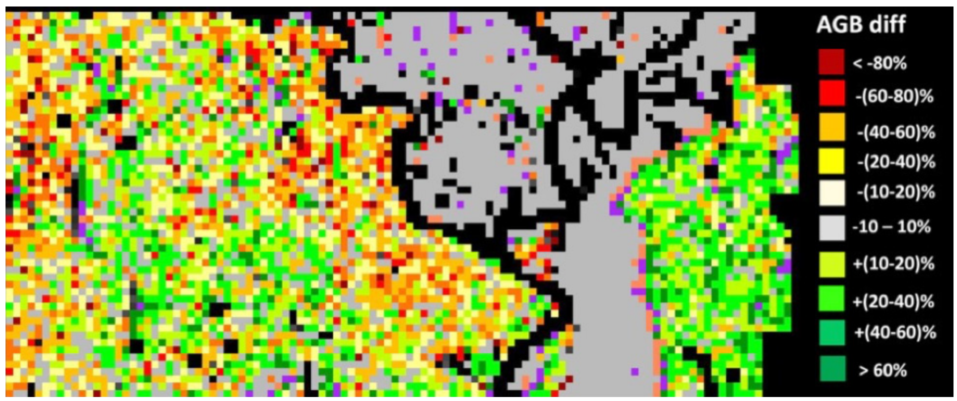


Figure 1.10 *An example of an AGB percent difference map between lidar-derived AGB and backscatter (HV)-derived AGB.*

Step 3: Use the lidar systematic sample points (lss_100m_agb) to extract the HV backscatter values for each year (See Section 4.2, step 4). Save the output as lss_100m_agb_sar15_17.

Step 4: Open the lss_100m_agb_sar15_17.dbf in Excel. Remember to clean the data as described in Section 4.2 step #5). Next, create a new column where you average the backscatter values from 2015, 2016, and 2017 to create a mean backscatter value in the spreadsheet.

Step 5: Create a scatterplot with the lidar-derived AGB and the three year HV mean in the spreadsheet. Develop a new model using the power-law function for simplicity.

Step 6: Before we apply the model we created in step 5, first we need to average the three backscatter images to create one single image of HV. Remember to use your gamma power images that have been resampled to 100m. You can use Raster Calculator to average the three images in QGIS.

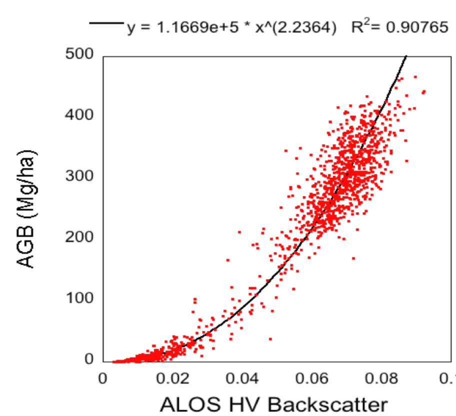


Figure 1.11 *An example of a best-fit model between AGB and a 3-year average of HV backscatter.*

Step 7: Based on your results from step 5, apply the best-fit equation AGB= 116690*(HV^2.2364) to the averaged HV image using Raster Calculator. Display the results with an appropriate color scale.

Step 8: Develop a percent difference map between LiDAR AGB and the new HV AGB map and color the range of biomass difference in percentage and display it (See Section 4.3 steps 7 and 8).

Step 9: Compare the new percent difference map with the earlier version derived just from the 2015 backscatter. Where do you notice differences in the overall negative and positive percent differences?

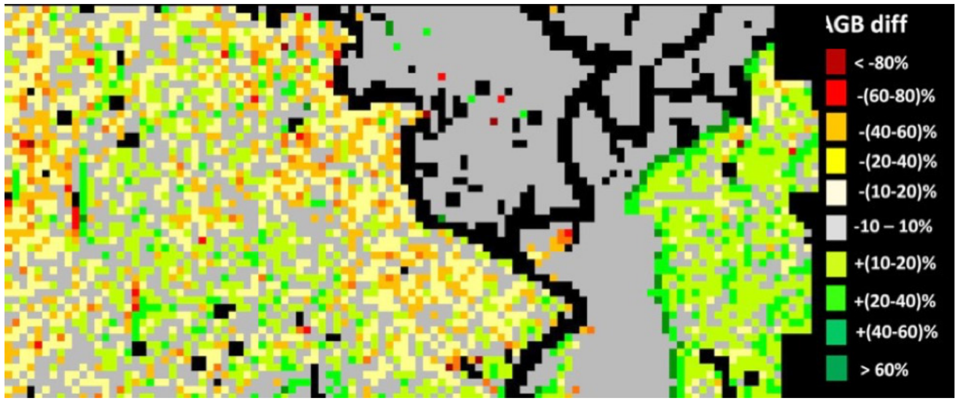


Figure 1.12 *Example of an AGB percent difference map between lidar AGB and a three-year average of HV*

4.5 Evaluating Uncertainty in the AGB Map

By assuming that we have several sources of errors that introduce uncertainty in the pixel level estimation of biomass, we can calculate the total uncertainty associated with estimating AGB at the pixel level by assuming all errors were independent and random, by using:

$$\varepsilon_{AGB} = \sqrt{\varepsilon^2_{measure} + \varepsilon^2_{model} + \varepsilon^2_{sampling} + \varepsilon^2_{prediction}}$$

where each of the terms are the relative errors at that pixel scale. Detailed description of error analysis and uncertainty assessment of the map are given in the Chapter 5. Here we examine three steps for evaluating the uncertainty of the map.

Step 1: For pixel level prediction, use model fit parameter uncertainty to simulate several biomass maps by bootstrapping the coefficients using the range of parameter uncertainty.

Step 2: Generate several maps (100 if the image is small as in the Nepal case) or about 20-30 if the image is large.

Step 3: Calculate the mean and variance of the bootstrapping approach and show the variance as a new map.