# Labeling roads in satellite imagery

Problems with the current data set of Borneo

The current data are the result of the work of volunteers who have manually labeled roads in four 50 by 50 km areas of Southeastern Borneo (see Fig. 1 below). This has been done in Google Earth (GE) with imagery spanning over a decade (earliest labels are from 2005). The data suffer from the following issues:

- road labels are often imprecise, sometimes off by dozens of meters, cutting corners, etc.
- labels are partly incomplete, likely reasons being manifold:
  - area has not been covered by users
  - incomplete labeling by user, possibly also as per instructions (which say that roads near settlements are not to be labeled)
  - as road construction is continuous process, road labeling must be based on satellite imagery from a limited time period (say, 6 months). in GE however, the user is presented with a mish-mash of imagery seemingly taken at different dates. There is little control over the date /recency of the satellite imagery displayed, and in some instances the exact date is not even displayed.
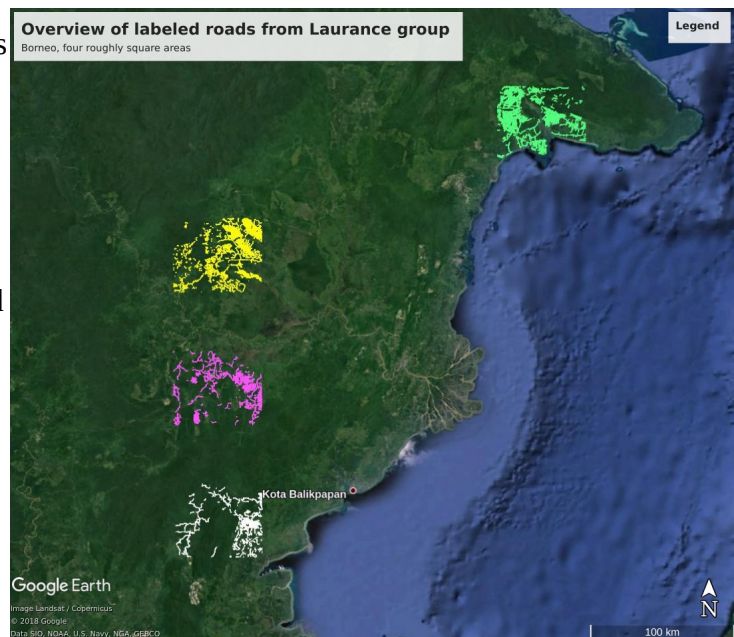
Labeling procedure

As most of the data is not of sufficient quality for our purposes, there is no compelling reason to follow the exact same mapping procedure as outlined in the manual provided by the Global Road Map Initiative (Sean Sloan, Laurance group). However, it does make sense to label roads in the same areas, not in the least as we may need the previous labels to determine a baseline.

Part 1: definitions and preparations

- decide on one of the four **areas of interest (AOI)** of Southeastern Borneo for mapping and make sure you have appropriate satellite imagery from a reasonably narrow and recent time interval (visible band, with little cloud cover)
- check whether the area has been worked on, and make sure that the resulting *.kml or *.kmz files are in the area's directory

- create or make sure you have a temporary directory with sufficient space into which GE can place temporary files
- as we will not be able to label all four AOIs in their entirety, we have to select subsets of these and ensure as best as we can that they are a representative cross-section, including paved and unpaved roads, roads through forests, roads through plantations, villages, and so on. So, it makes most sense to define several smaller, contiguous areas within each AIO. Let's term these **subAOI**. Currently, our agreement is that we label individual satellite images in their entirety, so a subAOI must correspond to the extent of a single satellite image or groups of adjacent images (no gap between them)

Part 2: import bounding box, previous labels and satellite image(s)

- start GE; on the lower left pane called Layers uncheck everything
- load from the AOI's data directory a file named boundingBox_xxxx.kml (where xxxx is the numerical code of the AOI) into GE (File-Open). You should see a 50 by 50 km square outlining the current AOI.
- the next step depends on whether previous labels for the AOI exist (look for files named e.g. RoadLabels_RDT_3093_a)
  - if road labels exist already: file-open <pick file>. Do this for each of the labeled subAOIs. **This is important so we don't label roads multiple times! For the same reason, it is important to save road labels in the AOI's directory and with the naming conventions as given below.** You should now see the subAOI as a 'map' with a corresponding symbol (round, blue stripes) and a subdirectory of the same name (see Fig. 2 below) as well as a polygon named 'boundingPoly....'.
  - if labels for the AOI do not exist, or if you want to start a new subAOI, create a new directory in the Temporary Places folder of the Places panel. Naming convention: RoadLabels_RDT_<numerical area code>_<alphabetical letter for subAOI>, e.g. RoadLabels_RDT_3093_a for subAOI 'a' in AOI 3093. Don't worry that you don't see the round map symbol as in Fig. 2, this will appear as soon as you save the directory (typically after having labeled some roads) and load it again in the next GE session.
- load a suitable satellite image: File-Open-<select image>
- GE will tell you that 'The imported image is larger than the maximum size supported by the hardware' (??) and ask you what to do. Choose 'Create Super Overlay' and in the next dialogue choose above-mentioned temporary directory. GE will take some time to load the data and in doing so create hundreds to thousands smaller sub-images in the temp directory. **Don't choose any of the other options** as this will result in images with lower resolution

Part 3: map roads

- see Figure 2 below to follow along:
- in the Places panel top left, make sure the directory (RoadLabels_RDT_...) or any of the paths in it are highlighted so the road labels are placed in that directory. **Again, if that directory is not highlighted, the new 'path' (road label) you are about to create will be placed anywhere in Temporary Places and very likely be lost** (unless you realized your mistake and move it into the directory right after finishing the path)
- start mapping: in the upper tools bar, click on 'Add path'. Up opens a window, into which you enter the label's name according to the following convention: <date of the underlying satellite image as yyyymmdd>_<initials>_<road code: 1 for paved, 2 for unpaved, 3 for road-like structures, e.g rivers>. Example: 20180803_hh_2. Yes, we may have many paths with identical names. If a road path spans images taken on more than one date, use the date of the most recent image.
- occasionally we may realize that we mislabeled a river for a road, or that we are constantly confused by its existence. In such cases it does make a lot of sense to keep or generate a label so we don't risk labeling it (again)

- set color: cyan for unpaved roads (by far the majority), white for paved roads, blue for rivers and red for other road-like structures (there are railways, power lines etc.)
- merrily click along, following the road as precisely as possible (this usually requires altitudes between 0.5-2 km, and frequent zooming in and out). When done with the current path, click OK in the window that popped up. Hints:
  - it's no problem if the road is not mapped along its entire length
  - while we're in mapping mode:
    - we can zoom in and out with the mouse wheel
    - the pop-up window can be shifted aside with the mouse
    - the whole map can be shifted with the arrow keys
    - the last click can be deleted via right mouse click
  - if you're not entirely happy with your path, you can correct it in many ways afterwards: right click on the path in the Places pane, properties. You can then
    - move individual points with the mouse (as soon as the mouse pointer turns into a pointed finger)
    - delete individual points via right mouse click
    - continue the path - click on the last point of the existing path and continue by extending it
- when done labeling, or in between, if you want to make sure that your last 20 minutes' worth of clicking don't get lost, right click the label directory/map and save it as a kmz file (Save Places as...). Save it in the directory in which all data from the Laurance group reside, not with the satellite images. Naming convention as explained above: RoadLabels_RDT_<numerical area code>_<alphabetical letter for subAOI>, e.g. RoadLabels_RDT_3093_a for subAOI 'a' in AOI 3093
- when really done labeling, create a bounding polygon around the area of labeled roads or extend the existing one (similar to paths, we can move individual points of a polygon by right click-properties). Inside the polygon, no unlabeled road must exist! Save the file again.
- It's a good idea to make a copy of the file and attach the present day's data. The files are very lightweight, so it won't hurt.
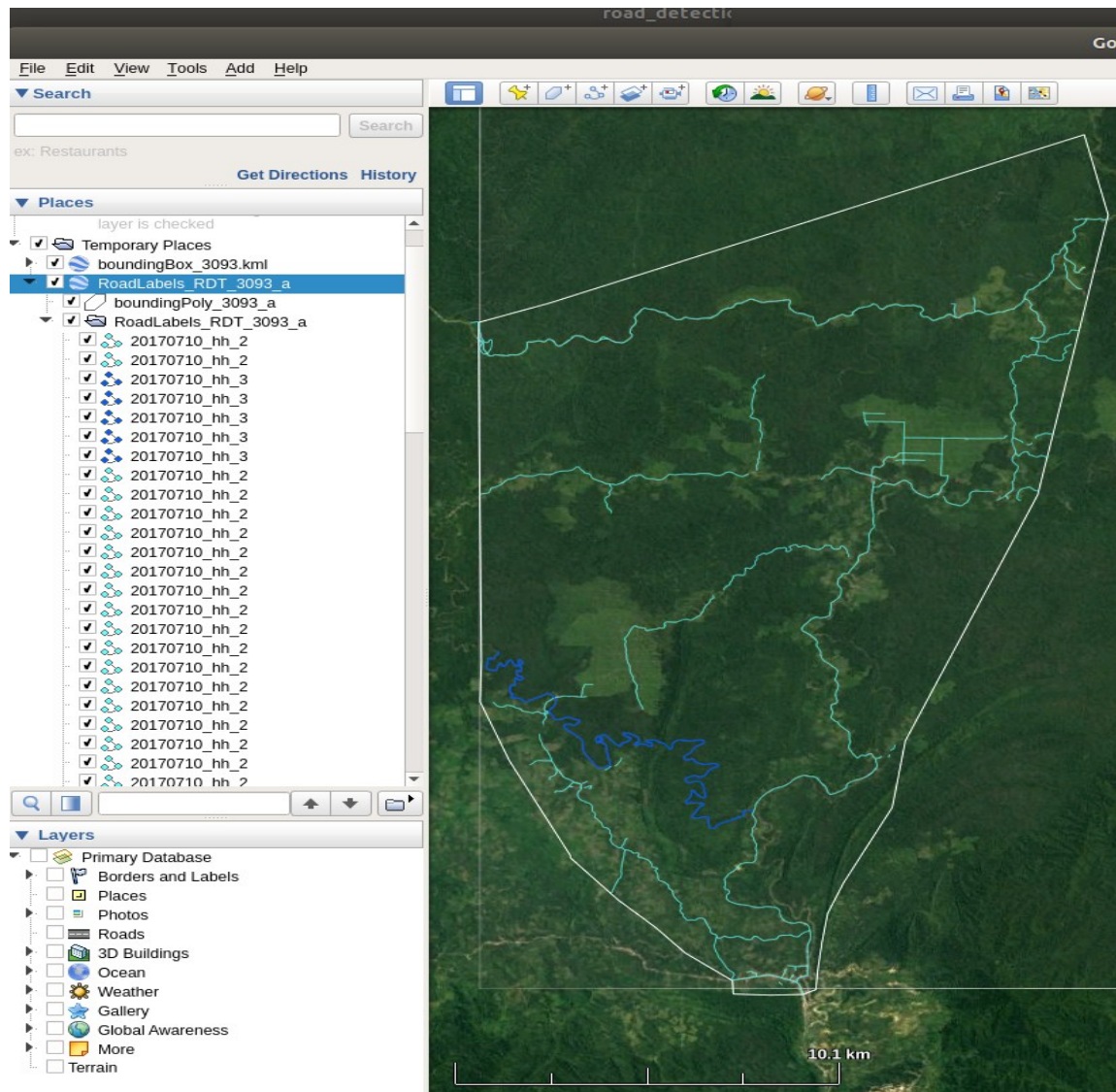
Fig. 2: Exemplary outcome of a happy clicking period. In the satellite image, cyan are the manually labeled (all unpaved) roads, blue is an inadvertently mislabeled river, white is the bounding polygon of subAOI 'a', and faint gray is the bounding box of the underlying AOI in Borneo (3093 in this case). In the Places pane left, all labeled road segments ('paths') are in directory named RoadLabels_RDT_3093_a (as described in the main text). The directory and the bounding polygon (boundingPoly_3093_a) are part of a GE 'map' named like the directory it contains, RoadLabels_RDT_3093_a. This map must be saved as a kml or kmz file.