NEW PRSOPECTS FOR LAND AREA MONITORING IN INSULAR SOUTHEAST ASIA WITH MULTI-SOURCE 10-30 M RESOLUTION REMOTE SENSING DATA

Jukka Miettinen, Chenghua Shi and Soo Chin Liew

Centre for Remote Imaging, Sensing and Processing (CRISP), National University of Singapore (NUS), Blk S17 level 2, 10 Lower Kent Ridge Road, Singapore 119076
Email: crsjukk@nus.edu.sg, crssc@nus.edu.sg and scliew@nus.edu.sg

KEYWORDS: Google Earth Engine, Sentinel-1, Landsat, data fusion

ABSTRACT: Due to the difficult climate conditions in insular Southeast Asia, large scale land area monitoring efforts in the region have been largely limited to coarse resolution products. In this paper we highlight the recently greatly improved availability and processing capabilities of 10-30 m resolution data and show their prospects for land area monitoring in the region using an example of a regional land cover mapping approach taking advantage of both optical and radar satellite data. The mapping example demonstrates the feasibility of automated large scale 10-30 m resolution land cover mapping in the region. The resulting finely detailed map contains 11 land cover classes revealing a great deal of contextual information through visual evaluation. The map has around 76% overall accuracy, with class wise accuracies of 59-97%. The major limitations currently relate to the incapability of separating primary forests from other tree covered areas in an automated manner, as well as to the large variety of different landscapes (e.g. home gardens and tea plantations) all classified as shrubland. Overall, we see great potential in the recently improved availability and processing capabilities of 10-30 m satellite data for land cover mapping in insular Southeast Asian conditions. However, further development is needed to overcome some remaining crucial limitations.

1. INTRODUCTION

Due to the notoriously difficult atmospheric conditions in the per-humid insular Southeast Asia (Brunei, Indonesia, Malaysia, Singapore and Timor Leste; Figure 1), land area monitoring activities have been traditionally greatly hampered by chronically inadequate availability of cloud free satellite observations in optical datasets. This has been particularly the case for high resolution optical data (≤30 m spatial resolution). Most regional level land area monitoring efforts have therefore been based on coarse resolution satellite data. However, coarse resolution remote sensing datasets do not allow proper identification of the complex and often fragmented insular Southeast Asian landscapes. Furthermore, the proportion of fragmented landscapes is all the time increasing, thereby reducing the usability of coarse resolution land cover products in the region. On the other hand, radar data alone is not sufficient to produce the level of thematic detail in land cover mapping products that would be required by the user community.

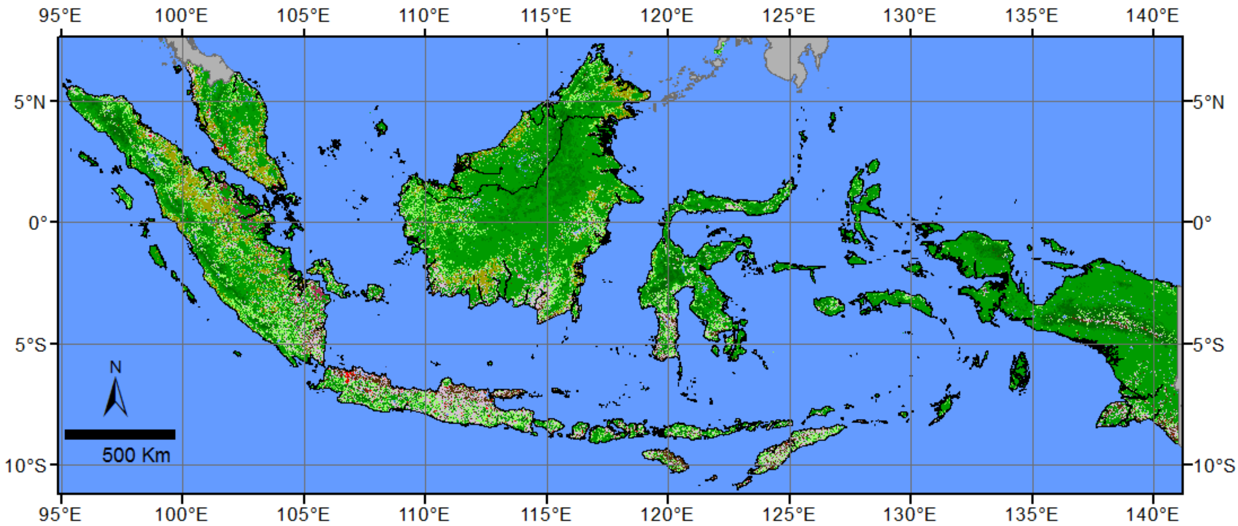

Figure 1. Insular Southeast Asia, as defined in this paper.

However, over the past few years, a dramatic shift in the availability of 10-30 m resolution optical satellite data has taken place. Combined with the existing Landsat 7 satellite, the successful launches of Landsat 8 and the two Sentinel-2 satellites have ushered in a new era for land area monitoring in the region. From late 2017 onwards, Landsat 7+8 and Sentinel 2A+2B satellites are providing 10-30 m resolution optical data in matching wavelength bands in frequency of nearly 120 observations pear year (i.e. on average around once in every three days). This unprecedented temporal frequency of observations will allow, even in the difficult atmospheric conditions of insular Southeast Asia, creation of full coverage cloud free composite images at least annually, potentially in as shorts as semi-annual intervals. These optical datasets can be used in combination with the Phased Array Synthetic Aperture Radar two (PALSAR-2) and Sentinel-1 radar data. For the first time in history, there is now an adequate volume of 10-30 m resolution satellite data available for land cover mapping in insular Southeast Asia.

In this paper we highlight the new prospects opened up by the recently greatly improved availability and processing capabilities of 10-30 m resolution data for land area monitoring in insular Southeast Asia. The dramatic rise in the amount of data calls for new data processing and classification approaches. Here we will 1) highlight the data availability and powerful processing capabilities of Google Earth Engine and 2) highlight the effectiveness of data fusion with a fully automated decision tree based classification approach for regional level mapping in insular Southeast Asian conditions.

2. MATERIALS AND METHODS

2.1 Google Earth Engine (GEE) – Application Programming Interface (API)

The Google Earth Engine (GEE) is a cloud-based platform for global satellite (and other GIS) data analysis. It provides a vast selection of publicly available remote sensing data stored in Google's servers and enables direct processing of these data in Google's computational infrastructure optimized for parallel processing of geospatial data. For more information please visit https://developers.google.com/earth-engine/. The GEE Application Programming Interface (API) is available in both JavaScript and Python. The API supports complex geospatial analyses including e.g. image processing, classification, change detection and vector-based extraction of image statistics. It provides a library of functions which may be applied to imagery for display and analysis. Through the API, users can create their own algorithms and recombine existing algorithms.

In the example presented in this paper we used the GEE JavaScript API to create all the remote sensing datasets (apart from PALSAR-2 data) used in the land cover classification. However, the PALSAR-2 data can also be uploaded to GEE in the users's workspace and utilized together with other datasets available in GEE. The classification decision tree was originally developed outside GEE but later translated into JavaScript, allowing the entire land cover mapping example to be run inside the GEE API.

2.2 Datasets

Altogether three different satellite datasets were used in the example land cover mapping presented in this paper: optical Landsat data as well as Sentinel-1 and ALOS-2 PALSAR-2 radar data. Both Landsat and Sentinel-1 data are available in GEE.

For the optical Landsat data, a yearly composite image was created as in Miettinen et al. (2016). Landsat 7 ETM+ (Enhanced Thematic Mapper) and Landsat 8 OLI (Operational Land Imager) Top of Atmosphere reflectance (TOA) data were combined selecting the median observation of all valid observations during the year 2015. An observation (pixel) was considered valid if the reflectance in blue band was less than 0.2 and an additional criterion was met. The additional criterion was dependent on vegetation greenness and was defined as: (1) If NDVI was less than 0.6, the ratio between red and the 2.1 μm band needed to be less than 1.0 or (2) If NDVI was greater than or equal to 0.6, the ratio between red and the 2.1μm band needed to be less than 2.5.

The Sentinel-1 dual-polarization Ground Resolution Detected (GRD) products in GEE have gone through pre-processing by Sentinel-1 toolbox including: 1) thermal noise removal, 2) radiometric calibration, 3) terrain correction using 30 m resolution Shuttle Radar Topographic Mission (SRTM) elevation data and 4) conversion to decibel values. In the example presented in this paper we created a composite image using the median value of all available dual polarization (VV and VH) data acquired on the Interferometric Wide Swath (IW) mode in 10 m spatial resolution between August and December 2015.

In addition, the insular Southeast Asian part of the ALOS-2 PALSAR-2 25 m resolution global mosaic product for the year 2015 provided by Japan Aerospace Exploration Agency (JAXA) at the ALOS-2 website (http://www.eorc.jaxa.jp/ALOS/en/palsar_fnf/fnf_index.htm) was used. This dataset is created by mosaicking images of HH and HV backscatter coefficients acquired by PALSAR-2 in Fine Beam Dual (FBD) mode. Orthorectification and topographic corrections have been applied (ALOS 2016).

Finally, the 30 m spatial resolution Shuttle Radar Topography Mission (SRTM) elevation data (Farr et al. 2007), two ocean and coastal area masks as well as two additional masks outlining the areas where 1) *Acacia* monocultures are mapped and 2) oil palm monocultures are not mapped were needed in the classification. For further details on these masks and on the data processing in general, please refer to Miettinen et al. (in review).

2.5 Decision tree classification

The decision tree classification used in the example presented in this paper was originally developed in Miettinen et al. (in review). The decision tree was specifically built for automated land cover classification approach that would produce annual land cover maps without the need for yearly sample data collection in the insular Southeast Asian conditions. In total the decision tree utilized 14 different variables from the satellite and auxiliary datasets and resulted in 11 land cover types (Table 1). Note that all of the classes describe purely land cover features, as opposed to land use. For further details of the development and structure of the classification decision tree, please see Miettinen et al. (in review). The map was produced in 30 m spatial resolution. The accuracy assessment of the land cover map was based on 1000 sample points visually interpreted in very high resolution satellite images available in Google Earth.

Table 1. Land cover classes produced by the decision tree classification (from Miettinen et al. in review).

Land cover type	Description
1. Water	Both inland and ocean water areas
2. Mangrove	Mangrove tree and shrub cover
3. Bare soil	Bare soil areas containing none or only senescent vegetation, typically including e.g. newly cleared areas, bare agricultural land and roads
4. Open	Areas with herbaceous vegetation, typically including e.g. agricultural land with annual crops, grasslands and areas (either rural or urban) with open (<25%) canopy woody vegetation
5. Shrub	Open to closed (25-100%) canopy woody shrub and young secondary regrowth, less than ~5 m. Note that this class is expected to also include broken canopy (25-60%) tree cover areas due to mixture of remotely sensed signals.
6. Lowland tree cover	Closed canopy (>60%) woody vegetation taller than ~5 m in elevation up to 750 m above sea level (asl)
7. Lower montane tree cover	Closed canopy (>60%) woody vegetation taller than ~5 m in elevation from 750 m to 1500 m asl
8. Upper montane tree cover	Closed canopy (>60%) woody vegetation taller than ~5 m in elevation above 1500 m asl
9. Oil palm	Closed canopy oil palm monoculture. Note that also other palm stands (e.g. coconut and sago) may be included in some areas.
10. *Acacia*	Closed canopy (>90%) *Acacia* spp. monoculture
11. Built-up	Areas dominated by buildings or other man-made structures

3. RESULTS

During the data processing and dataset preparation the GEE proved to be a truly powerful tool for creating composite images utilized in the land cover classification. Due to the yearly compositing period and high number of 10-30 m images needed to cover the entire insular Southeast Asia, in total thousands of images were used in the creation of the map. This would not have been possible without a centralized image database with integrated processing tools. The GEE was also found to be very convenient working environment for algorithm development since the effects of changes in various parameters can be seen within seconds by executing the modified code. Figure 2 provides an example view of the GEE working environment with a year 2015 Landsat 7+8 composite image visible. Note that GEE is not the only cloud service providing this type of service. For example the Australian Data Cube (http://www.datacube.org.au/) aims to provide similar data processing possibilities.

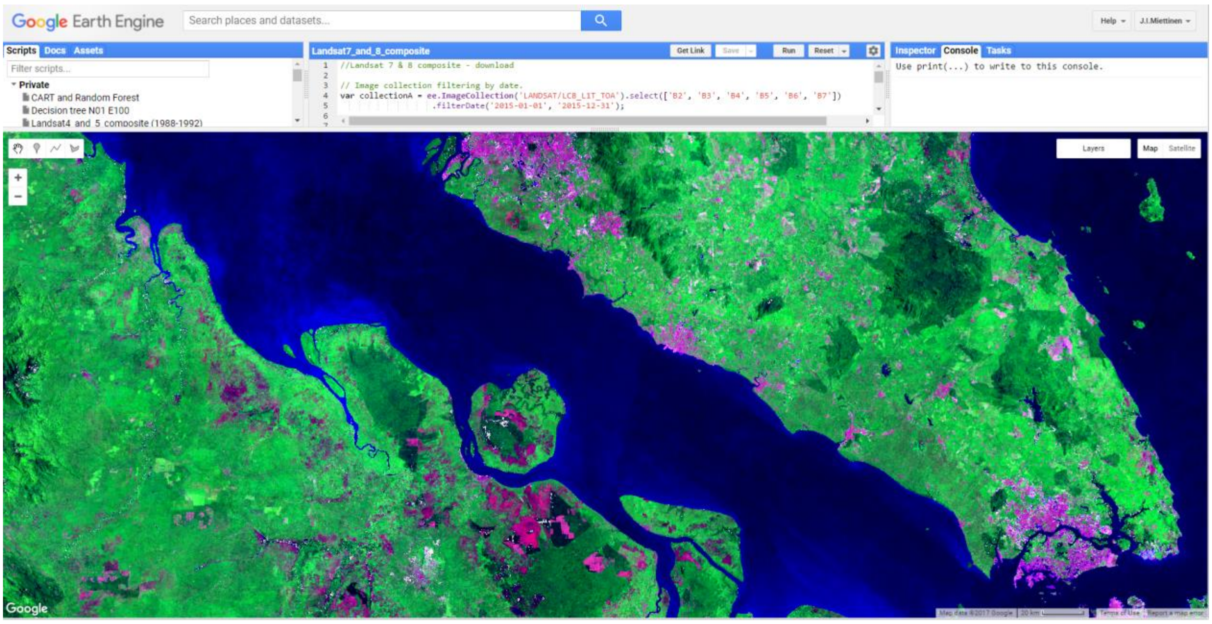

Figure 2. An example view of the GEE working environment with 2015 Landsat 7+8 composite image.

The fully automated decision tree classification resulted in a land cover map with unprecedented spatial and thematic detail in insular Southeast Asian regional level land cover mapping. The 30 m spatial resolution helps to deliver a great deal of contextual information of the area in visual map interpretation (Figure 3). The overall accuracy of the map was around 76%, with class wise accuracies varying between 55% and 97%. Particularly the high user's accuracy of 90% for closed canopy oil palm was considered encouraging. This may potentially enable efficient monitoring of the extent of closed canopy oil palm plantation in the region that produces the great majority of all palm oil in the world (SPOTT 2017). However, the noticeably lower producer's accuracy of 71% indicates underestimation of the extent of oil palm plantations. Further research is needed to understand the strengths and limitations of oil palm mapping with the approach demonstrated in this paper.

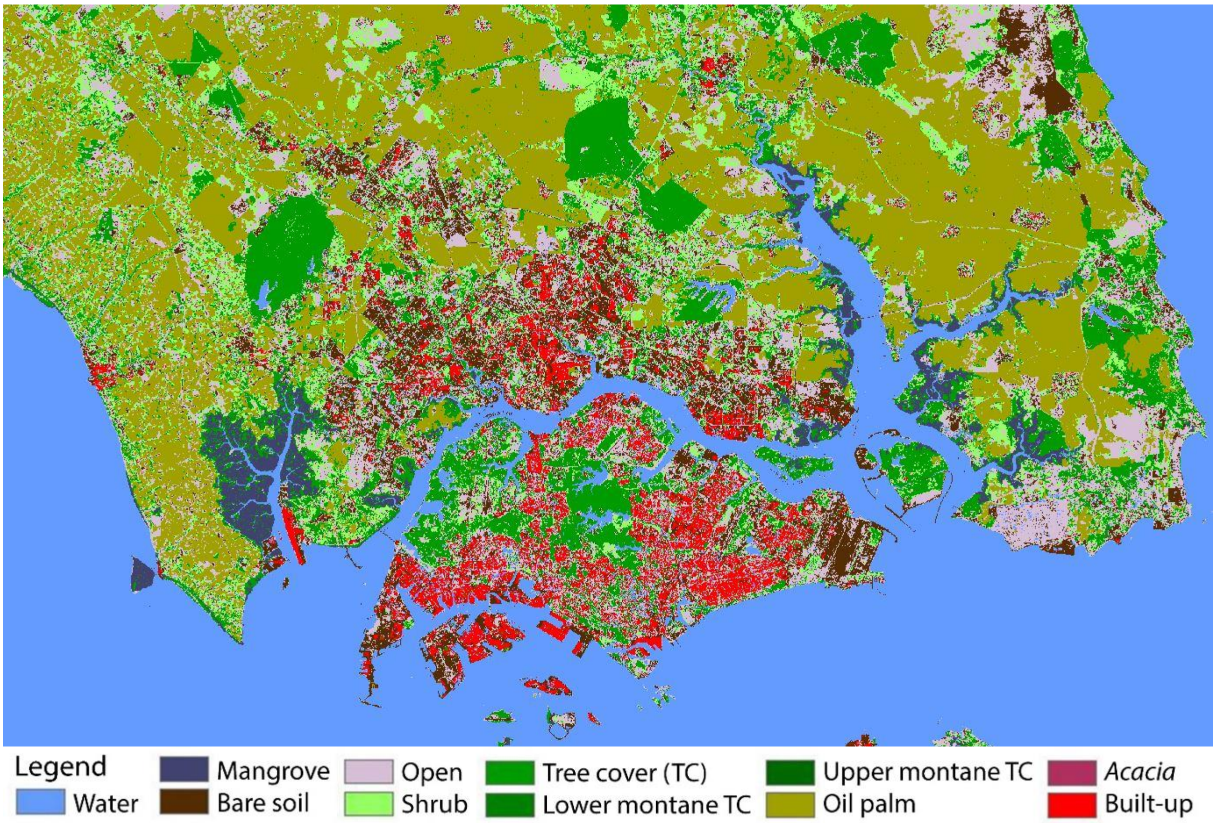

Figure 3. Example of the 30 m resolution land cover map around Singapore.

The main limitations of the current mapping approach are related to the identification of primary forest areas and the wide range of landscape features classified into the 'Shrub'-class. The first limitation is partially due to the fact that the decision tree classification only utilized band ratios which reduce the separability of old growth forests and regrowth. Currently, all closed canopy tree cover areas (apart from oil palm and *Acacia*) are classified into the tree cover classes (Figure 4). The problems with the 'Shrub'-class, on the other hand, are related to the great variation of land cover features within the region, many of which essentially exhibit a fine scale mixture of open soil and vegetation fractions, all classified into the 'Shrub'-class. These types of landscape features include e.g. true shrublands and young monoculture plantations as well as open canopy tree cover, home gardens, mosaic agricultural areas and tea plantations. For the usability of the map it would be essential to find ways to separate these land cover features in the future.

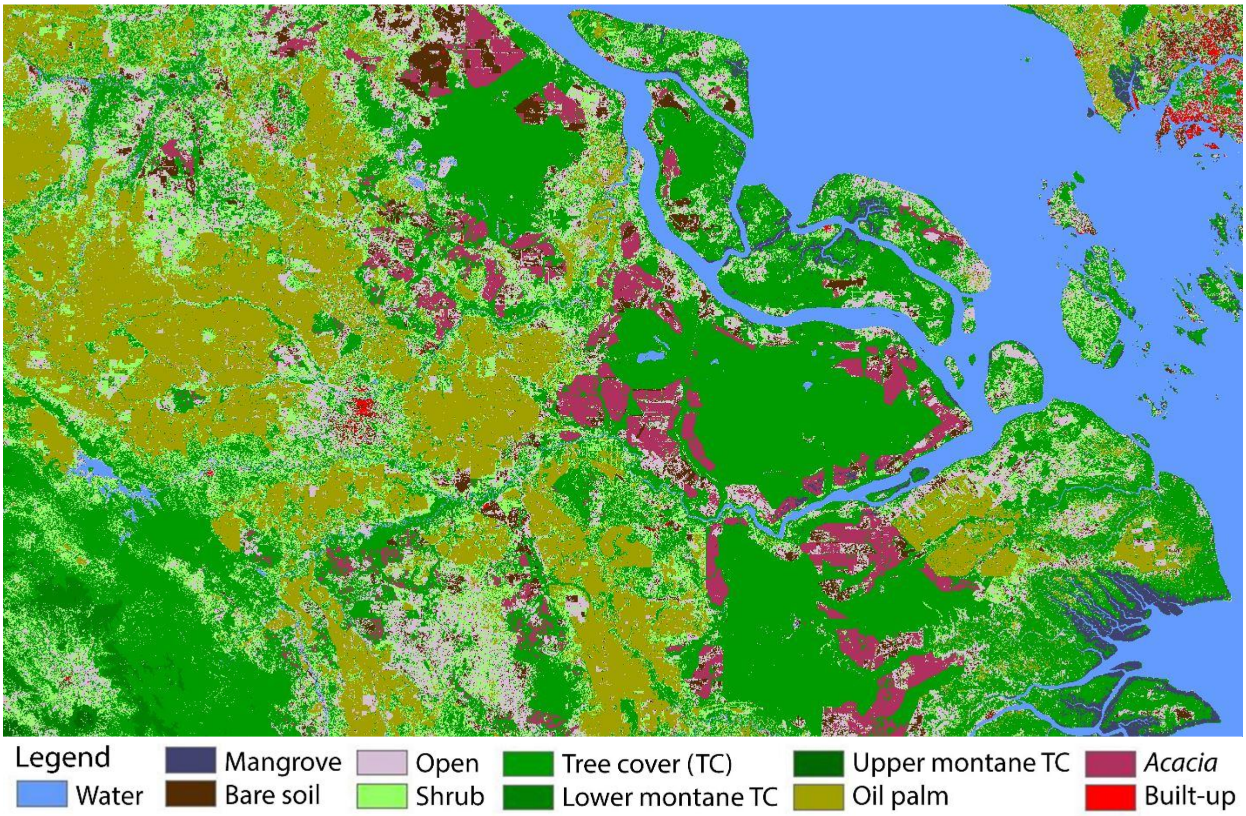

Figure 4. Example of the 30 m resolution land cover map in Riau, Sumatra, Indonesia.

4. DISCUSSION AND CONCLUSION

Regardless of the remaining shortcomings of the results, the land cover mapping example presented in this paper highlights the great potential of the recently expanded availability of 10-30 m resolution data for land cover mapping in insular Southeast Asia. Utilizing the datasets and processing power of Google Earth Engine, supported by ALOS-2/PALSAR-2 data, we were able to create high quality composite datasets suitable for yearly land cover mapping. Furthermore, the entire process can be run within Google Earth Engine, enabling output of land cover maps for any area of interest within the region in seconds, for any time period (depending on data availability naturally).

The classification approach is designed to be generally applicable, potentially with further improvements, also to the currently fast cumulating Sentinel-2 data. This may also help in solving the remaining limitations of the classification approach highlighted in this paper. For the users of land cover products in the region, it would be essential to be able to separate primary forest areas from other tree cover. Similarly, better identification of the different types of land cover features currently classified as 'Shrub' is needed to be able to take full advantage of the 10-30 m resolution land cover products. Further research along these lines is needed.

REFERENCES

ALOS 2016. Global 25 m resolution PALSAR-2/PALSAR mosaic and forest/non-forest map (FNF): Dataset description. Japan Aerospace Exploration Agency (JAXA) and Earth Observation Research Center (EORC), Japan.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D. and Alsdorf, D., 2007. The Shuttle radar topography mission. *Reviews of Geophysics*, 45: RG2004.

Miettinen, J., Shi, C. and Liew, S.C., 2016. Land cover distribution in the peatlands of Peninsular Malaysia, Sumatra and Borneo in 2015 with changes since 1990. *Global Ecology and Conservation*, 6, pp. 67-78.

Miettinen, J., Shi, C. and Liew, S.C. (in review). Towards automated 10-30 m resolution land cover mapping in insular Southeast Asia. *Geocarto International* (in review).

SPOTT 2017. Sustainable Palm Oil Transparency Toolkit (SPOTT). London, UK: Zoological Society of London. Available at: http://www.sustainablepalmoil.org/asia/, accessed 22 Jun 2017.