Borneo : a quantitative analysis of botanical richness, endemicity and floristic regions based on herbarium records

Raes, N.

CHAPTER 2

Georeferencing specimens by combining digitized maps with SRTM digital elevation data and satellite images

Niels Raes, Johan B. Mols, Luc P.M. Willemse and Erik F. Smets 2009.
Proceedings of the 7th International Flora Malesiana Symposium.
Blumea 54(1)

Abstract

For numerous scientific purposes collection records need to be georeferenced. Although the geographic coordinates of many of the collection localities are available in gazetteers, especially collections from tropical areas of the world are still not georeferenced. In an attempt to georeference these localities for Indonesian Borneo we used digitized old maps which were georegistered with SRTM digital elevation data, and Landsat 7- and JERS-1 SAR radar satellite images. This enabled us to georeference 2577 additional collections from Indonesian Borneo, belonging to 1744 taxa, which were collected at 134 previously not georeferenced localities. This applied methodology enables researchers to georeference their historical collections for biodiversity, biogeographical, and global climate change impact studies.

Keywords
georeferencing; georegistration; historical map; Landsat; JERS-1 SAR; SRTM digital elevation data

Introduction

One of the most important aspects of digitized herbarium- and natural history museum records in order to be used for i.e. biodiversity assessments, predicting the effects of -habitat loss, -potential for species' invasions, and -climate change effects (Graham *et al.*, 2004; Peterson, 2006), is that they need to be accurately georeferenced. Most collections made during the last two decades have coordinates taken with GPS equipment. The older collections, and notably those made in the 19th and early 20th century, often have only named collection localities. In order to make these older collections useful for floristic- and biogeographical research, the collection localities need to be georeferenced with the aid of a printed-, or one of the many online gazetteers (i.e. Alexandria Digital Library Gazetteer[1], La Tierra gazetteer[2], or BioGeomancer[3]). This works fine as long as the localities refer to rivers, mountains, villages etc. in western countries.

For many localities, such as small settlements, creeks, and hills in remote tropical areas, however, the coordinates have either never been assessed, or have not been made available in a gazetteer. For the purposes mentioned above, especially the collections made in remote areas can be very important, since these areas have often been visited only once by a collecting expedition. Complicating matters even further is the fact that the named localities on the labels of the collections gathered during the 19th, and early 20th century expeditions, regularly refer to vernacular names used by local guides at that time. Frequently these localities are currently known under a different name, which makes it impossible to find them in a gazetteer. Furthermore, these remote areas are likely to suffer most from the 'Wallacean Shortfall' (Whittaker *et al.*, 2005), a phenomenon that certain geographical regions are far less sampled than others, resulting in biased collection densities (Parnell *et al.*, 2003; Reddy & Davalos, 2003; Moerman & Estabrook, 2006; Hortal *et al.*, 2007). To reduce the impact of the 'Wallacean Shortfall' to a minimum, it is important to georeference as many collections as possible from these already under-collected areas. Fortunately, during the early expeditions often maps were made that indicate the collection localities and their corresponding names used at the time.

1 http://middleware.alexandria.ucsb.edu/client/gaz/adl/index.jsp
2 http://www.tutiempo.net/Tierra/
3 http://www.biogeomancer.org

These maps are generally stored in the very same institutions that harbour the collections.

Instead of trying to calculate the coordinates of collection localities with a ruler, based on map coordinates printed in the margins, we aimed at geographically positioning digitized expedition maps by matching them with SRTM digital elevation data and high-resolution satellite images in a geographic information system (GIS), through a process known as georegistration.

Methods

This study is part of the assessment of the botanical diversity, -endemicity, and floristic regions of Borneo derived from species distribution models (Raes & ter Steege, 2007 - Chapter 3), hence this island was used as the model. The northern and western parts of Borneo belong to the countries Malaysia and Brunei and cover 27.5 % of the total area; the remainder – the Kalimantan provinces – belong to the country of Indonesia (Fig. 2.1). Malaysia and Brunei have a long history of botanical collecting, and local biodiversity studies (Proctor *et al.*, 1983; Proctor *et al.*, 1988; Ashton & Hall, 1992; Aiba *et al.*, 2002; Potts *et al.*, 2002; Slik *et al.*, 2003; Ashton, 2005). Therefore, many collection localities of these countries have been georeferenced, and are available in a printed-, or online gazetteer. From the total of 166,757 digitized collections of Borneo present in the database of the National Herbarium of the Netherlands (NHN), 69.6 % was collected in Malaysia and Brunei. This makes it even more important to include as many georeferenced collections from the Indonesian Kalimantan provinces as possible, in order to reduce the effects of the 'Wallacean Shortfall' to a inimum. Especially for Indonesian Borneo – with its extensive network of rivers and creeks running between mountains and hills, with many small settlements along their banks – localities often have only local names which were never georeferenced. Fortunately, there exist a reasonable amount of detailed and published expedition maps from the 19th, and early 20th century (Table 2.1). These maps were used the retrieve the coordinates for as many collection localities of the Kalimantan provinces as possible.

Georegistration of digitized maps and georeferencing collection localities

The first step in the georegistration process is to digitize all available maps at high resolution (Table 2.1). Secondly, we downloaded the SRTM 90m resolution digital elevation data[1], and the 28.5m resolution Landsat 7 (circa 2000)[2] images of Borneo. The 100m resolution JERS-1 SAR[3] radar satellite images were obtained from DVD-ROM (free of charge). These data were imported in a geographical information system, Manifold GIS (Manifold Net Ltd.), and projected to a geographic projection (Kennedy, 2000).

Thirdly, the digitized maps were georegistered in Manifold GIS. Georegistration is the process of adjusting an image (the digitized maps) to the geographic location of a 'known good' reference image (the geographically projected satellite images, and SRTM digital elevation data). The georegistration process starts with the identification of one 'known good' reference feature, i.e., a major city, main river, mountain top, or an extrusion of the coast line with a (online) gazetteer. This reference point is marked on both the satellite image and the digitized map, based on the coordinates retrieved from the gazetteer. This gives an indication about the geographical position of the map, and the area it covers.

1 http://srtm.csi.cgiar.org/
2 https://zulu.ssc.nasa.gov/mrsid/
3 http://www.eorc.jaxa.jp/JERS-1/GFMP/#SEA2/

Table 2.1. List of georegistered digitized maps and their references.

Source		Reference
Geological explorations in Central Borneo		Molengraaff, 1900
	Topographic map of the north-eastern part of West Kalimantan (Map I)	
	Geological map of western Central Kalimantan and part of South Kalimantan (Map II)	
	Geological sketch-map of a part of the Kapoewas-river basin and the great lakes (Map III)	
	The Soengai Embaloeh (Map V)	
	The Soengai Mandai (Map VI)	
	The Upper Kapoewas (Map VIIa)	
	The Upper-Kapoewas, the Boengan, the Boelit and the track from the Boelit-river across the waterparting to the Mahakam-Basin in East-Borneo (Maps VIIb,c)	
	The Seberoewang and the Embahoe (Map VIIIa)	
	The Seberoewang (Map VIIIb)	
	From the Boenoet, the Sebilit and the Tebaoeng across the Madi-Plateau to the Melawi-Valley, the Lekawai and the Schwaner-Mountains (Map IXa)	
	The Boenoet (Maps IXb,c)	
	Topographical and geological sketch-map of the Samba River (Maps Xa-e)	
Comprehensive atlas of the Netherlands East Indies		van Diessen et al., 2004
	West Kalimantan pp. 350-351	
	Central and West Kalimantan pp. 352-353	
	East Kalimantan pp. 360-361	
	South and East Kalimantan pp. 362-363	
Miscellaneous		
	Banjermasing, Martapoera and part of the Lawoet areas 1845	Müller, 1857
	Kaart van de kust- en binnenlanden van Banjermasing	
	West Kalimantan	Hallier, 1895
	Sketch-map of the upper Barito (Boesang and Bakaäng) at the watersheds of the Barito-Mahakam, the Mahakam-Kapoeas and the Kapoeas-Barito.	Stolk, 1907
	Sketch-map of the Kajan, Bahau and Poedjoengan	van Walcheren, 1907
	Sketch-map of the Boeloengan and the Apo-Kajan	Nieuwenhuis, 1910
	Expeditie N.O. Borneo 1925. Reisroute v/d botanist F.H. Endert	Buys et al., 1927
	Midden-Oost-Borneo-Expeditie 1925; Endert F.H.	
	Overzicht van de tot dusverre verkregen topografische resultaten	
	Map I. Travels in the Serawai area	Winkler, 1927
	Map II. Travels in the upper Kapuas area	
	Along the Mahakam	Witkamp, 1932
	Sankoelirang	Endert, 1933
	Reede van Singkawang	Dunselman, 1939
	West Kalimantan	Dalton, 1978
	Mahakam river	
	Danau Sentarum Nature Reserve- West Kalimantan	van Balen, 1996
	Sketch-map of central East Borneo	Unknown

Next, as many reference points that were indicated on the digitized map (i.e. villages, river bends, tributaries, hill- and mountain tops), and also are recognizable on the satellite image, were marked on both the digitized map and the satellite image. Most frequently we used the Landsat 7 images, because these have the highest resolution and the most detail. However, when a location on the digitized map was obscured by a cloud cover on the Landsat 7 satellite image, we switched to the JERS-1 SAR radar satellite image, which penetrates through the cloud cover. To identify mountain tops we used the SRTM 90m resolution digital elevation data.

Finally, the digitized map is superimposed on the satellite image based on the reference points on the satellite image, thus is georegistered (Fig. 2.1). We allowed a certain degree of transformation of the digitized maps during the georegistration process to correct for differences in map projections, i.e. the way the round earth is flattened (Kennedy, 2000), and to overcome geographical measurement errors. Remind that most of the digitized maps are originally more than a century old, and that the equipment used at the time was not as accurate as the GPS equipment used today. To georeference the remainder of the localities that were not used as reference points, we superimposed the digitized and georegistered map (set as transparent) on the satellite images. By adding the remaining localities as points on a new data layer in the GIS, we were able to retrieve their coordinates, and thereby georeferenced them. This process was repeated for all available digitized maps at the NHN-Leiden University branch (Table 2.1). The named localities with their corresponding georeferenced coordinates were exported to a spreadsheet file and merged in the Borneo gazetteer of the NHN database.

Results and discussion

In total we used 34 digitized maps. From the selection of maps shown in Figure 2.1 it is clear that they differ greatly in the area they cover, and thereby in their amount of detail. The extent to which the maps are presented as diamond shapes, instead of rectangles, indicates the accuracy of the original maps, the differences in map projections, and the degree of transformation required to match the digitized maps with the satellite images. It should be kept in mind, however, that these maps, in many cases were developed based on compass readings. Nonetheless, they were often very accurate, and allowed us to georeference many map features. It is often argued that rivers are unreliable reference points, because they change their course during time. Our georegistration experiences confirm this fact, nevertheless the ancient river bends were in many occasions clearly visible as oxbow lakes, which were regularly used as reference points in the georegistration process.
In total we georeferenced 3269 unique localities from the digitized maps listed in Table 2.1, and merged these with the Borneo gazetteer of the NHN database. These localities are represented by black and white dots in Figure 2.1. From the 50,067 (30.1%) digitized collections from Indonesian Borneo stored at the NHN, we were able to georeference 40,646 records (81%) using various sources. Of these 40,646 records, 2577 (6.34 %) were georeferenced with localities retrieved from the digitized maps. These records could be attributed to 134 unique named localities and are represented as white dots in Figure 2.1. While this is only 4.10% from the total of 3269 georeferenced unique localities, the additionally 2577 georeferenced collections represented 1744 unique taxa. Although this percentage is lower than we initially had anticipated, considering the much lower collection density of the Indonesian part of Borneo, any additionally georeferenced collections make a valuable contribution, and reduce the impact of the 'Wallacean Shortfall' to a minimum.

At the same time the additions to the Borneo gazetteer can be used by other researchers enabling them to georeference the records of their taxa of interest. The methodology of georegistration allows researchers to assign accurate coordinates to their specimens based on historical maps, while at the same time illustrating the importance of historical maps for current research themes.

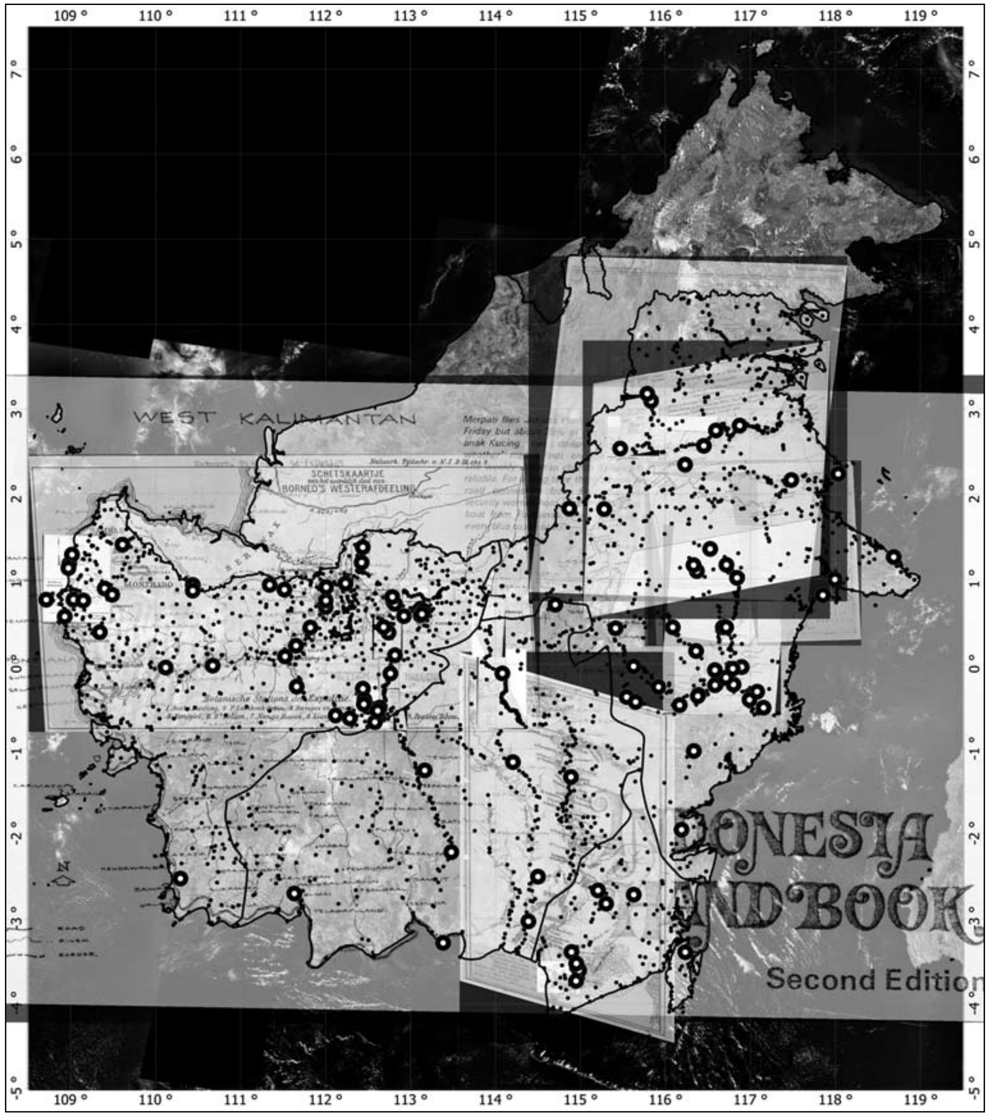

Figure 2.1. Landsat 7 image of Borneo (geographic projection) superimposed with a selection of georegistered digitized maps. Black dots indicate georeferenced localities retrieved from the maps. White dots indicate georeferenced localities where actual collections were made which otherwise could not have been georeferenced.

Acknowledgements - We like to thank Ben Kieft for digitizing all the maps, Bart Meganck for usefull comments to an earlier version of the manuscript, and the staff of the National Herbarium of the Netherlands for gathering all the maps from old publications and books which were used for this paper.