Blumea 54, 2009: 162–165
www.ingentaconnect.com/content/nhn/blumea

RESEARCH ARTICLE

doi:10.3767/000651909X475950

Georeferencing specimens by combining digitized maps with SRTM digital elevation data and satellite images: a Bornean case study

N. Raes[1], J.B. Mols[1], L.P.M. Willemse[1], E.F. Smets[1,2]

Key words

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

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 2 577 additional collections from Indonesian Borneo, belonging to 1 744 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.

Published on 30 October 2009

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., La Tierra gazetteer, or BioGeomancer). 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 a bias in 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. 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 phytogeographical patterns of Borneo derived from species distribution models (Raes & Ter Steege 2007, Raes et al. 2009), 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. 1). Malaysia and Brunei have a long history of botanical collecting and local biodiversity studies (Proctor et al. 1983, 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 minimum. 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

[1] National Herbarium of the Netherlands, Leiden University branch, P.O. Box 9514, 2300 RA Leiden, The Netherlands;
corresponding author e-mail: raes@nhn.leidenuniv.nl.
[2] Laboratory of Plant Systematics, Institute of Botany and Microbiology, Kasteelpark Arenberg 31, B-3001 Leuven, Belgium.

© 2009 Nationaal Herbarium Nederland

You are free to share - to copy, distribute and transmit the work, under the following conditions:
Attribution: You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
Non-commercial: You may not use this work for commercial purposes.
No derivative works: You may not alter, transform, or build upon this work.
For any reuse or distribution, you must make clear to others the license terms of this work, which can be found at http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode. Any of the above conditions can be waived if you get permission from the copyright holder. Nothing in this license impairs or restricts the author's moral rights.

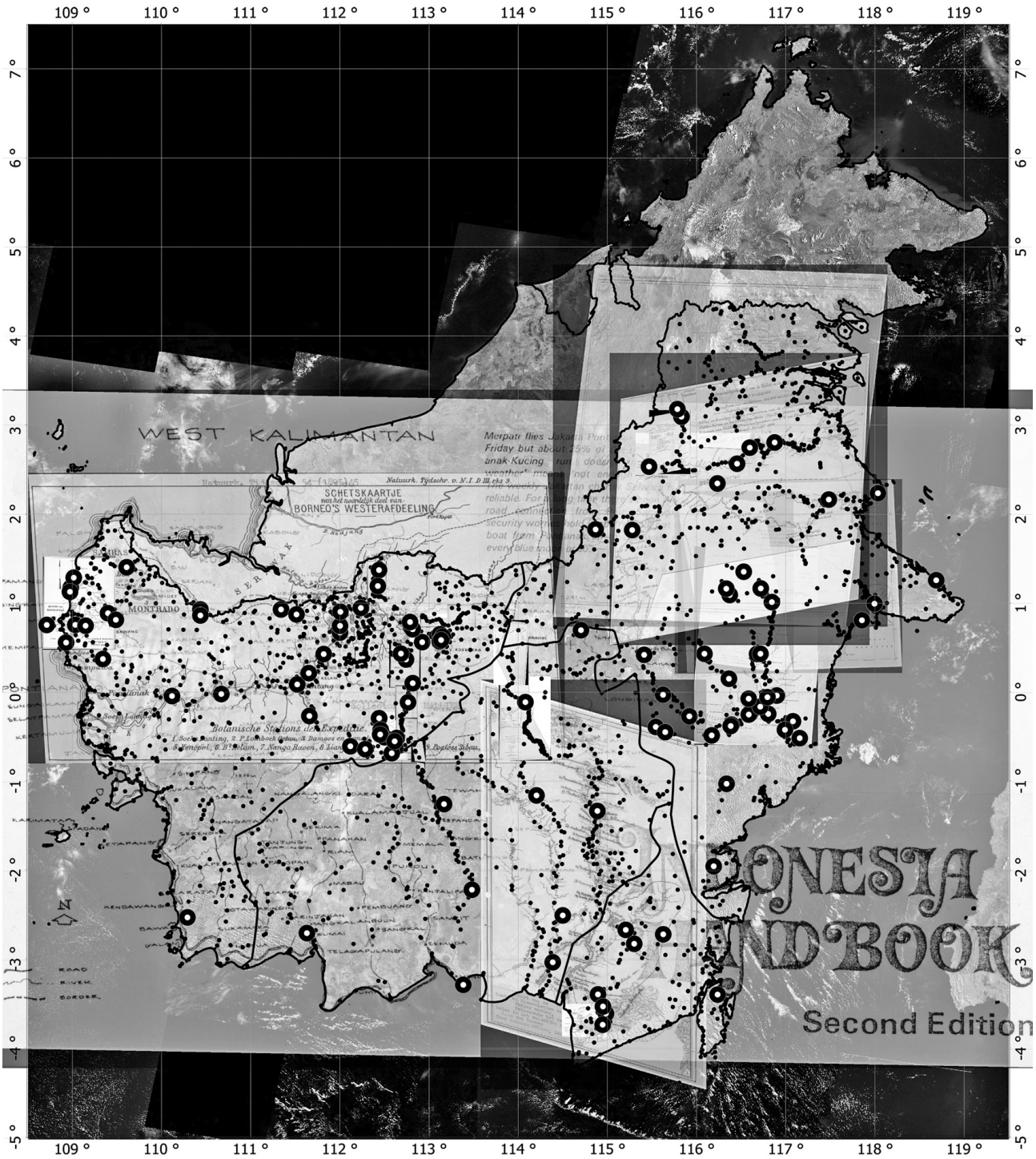

Fig. 1 Landsat 7 image of Borneo (geographic projection) superimposed with a selection of georegistered digitized maps. Black dots indicate georeferenced localities retrieved from the maps. White dots indicate georeferenced localities where actual collections were made which otherwise could not have been georeferenced.

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 1). These maps were used to 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 1). Secondly, we downloaded the SRTM 90 m resolution digital elevation data, and the 28.5 m resolution Landsat 7 (c. 2 000) images of Borneo. The 100 m resolution JERS-1 SAR 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.

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 90 m 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. 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 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 Fig. 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 3 269 unique localities from the digitized maps listed in Table 1, and merged these with the

Table 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 (Map IXb, c)	
Topographical and geological sketch-map of the Samba River (Map 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

Borneo gazetteer of the NHN database. These localities are represented by black and white dots in Fig. 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, 2 577 (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 Fig. 1. While this is only 4.1 % from the total of 3 269 georeferenced unique localities, the additionally 2 577 georeferenced collections represented 1 744 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 collection makes a valuable contribution, and reduces 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.

Acknowledgements We like to thank Ben Kieft for digitizing all the maps, Bart Meganck for usefull suggestions to improve an earlier version of the manuscript, and we acknowledge the NWO-Groot 'Building databases for life'-project (Grant# 175.010.2003.010) for funding the digitization of all Borneo collections used for this research.

REFERENCES

Aiba S-I, Kitayama K, Repin R. 2002. Species composition and species-area relationships of trees in nine permanent plots in altitudinal sequences on different geological substrates of Mount Kinabalu. Sabah Parks Nature Journal 5: 7–69.

Ashton PS. 2005. Lambir's forest: The world's most diverse known tree assemblage? In: DW Roubik, Sakai S, Hamid Karim AA (eds), Pollination ecology and the rain forest: Sarawak studies: 191–216. Springer, New York.

Ashton PS, Hall P. 1992. Comparisons of structure among mixed dipterocarp forests of north-western Borneo. The Journal of Ecology 80: 459–481.

Buys DW, Witkamp H, Endert FH, Siebers HC, Bosch DFK. 1927. Midden-Oost-Borneo expeditie 1925. Kolff & Co., Weltevreden.

Dalton B. 1978. Indonesia handbook, second edition. Moon Publications, London.

Dunselman J. 1939. Aantekeningen over zandpadangs in Borneo's Westerafdeeling. De Tropische Natuur 28: 71– 82.

Endert FH. 1933. De proefbaanmetingen in Sangkoelirang (Z.O. Borneo). Tectona 26: 50–92.

Graham CH, Ferrier S, Huettman F, Moritz C, Peterson AT. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in Ecology, Evolution 19: 497–503.

Hallier H. 1895. Rapport over de botanische tochten in Borneo's Westerafdeling gedurende de Borneo-expeditie 1893–94. Natuurkundig Tijdschrift 54: 406–449.

Hortal J, Lobo JM, Jiménez-Valverde A. 2007. Limitations of biodiversity databases: Case study on seed-plant diversity in Tenerife, Canary Islands. Conservation Biology 21: 853–863.

Kennedy M. 2000. Understanding map projections. Environmental Systems Research Institute, Inc., Redlands, USA.

Moerman DE, Estabrook GF. 2006. The botanist effect: counties with maximal species richness tend to be home to universities and botanists. Journal of Biogeography 33: 1969–1974.

Molengraaff GAF. 1900. Borneo-expeditie – Geologische verkenningstochten in Centraal-Borneo (1893–94) – Atlas 22 bladen. Brill & Gerlings, Leiden, Amsterdam.

Müller S. 1857. Reizen en onderzoekingen in den Indische Archipel gedaan op last der Nederlandsche Indische regering tusschen de jaren 1828 en 1836. Frederik Muller, Amsterdam.

Nieuwenhuis AW. 1910. Tochten naar Boven-Boeloengan en de Apo-Kajan. Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap 27: 263–306.

Parnell JAN, Simpson DA, Moat J, Kirkup DW, Chantaranothai P, Boyce PC, Bygrave P, Dransfield S, Jebb MHP, Macklin J, Meade C, Middleton DJ, Muasya AM, Prajaksood A, Pendry CA, Pooma R, Suddee S, Wilkin P. 2003. Plant collecting spread and densities: their potential impact on biogeographical studies in Thailand. Journal of Biogeography 30: 193–209.

Peterson AT. 2006. Uses and requirements of ecological niche models and related distributional models. Biodiversity Informatics 3: 59–72.

Potts MD, Ashton PS, Kaufman LS, Plotkin JB. 2002. Habitat patterns in tropical rain forests: a comparison of 105 plots in northwest Borneo. Ecology 83: 2782–2797.

Proctor J, Anderson JM, Chai P, Vallack HW. 1983. Ecological studies in four contrasting lowland rain forests in Gunung Mulu National Park, Sarawak: I. Forest environment, structure and floristics. The Journal of Ecology 71: 237–260.

Proctor J, Lee YF, Langley AM, Munro WRC, Nelson T. 1988. Ecological studies on Gunung Silam, a small ultrabasic mountain in Sabah, Malaysia. I. Environment, forest structure and floristics. The Journal of Ecology 76: 320–340.

Raes N, Roos MC, Slik JWF, Van Loon EE, Ter Steege H. 2009. Botanical richness and endemicity patterns of Borneo derived from species distribution models. Ecography 32: 180–192.

Raes N, Ter Steege H. 2007. A null-model for significance testing of presence-only species distribution models. Ecography 30: 727–736.

Reddy S, Davalos LM. 2003. Geographical sampling bias and its implications for conservation priorities in Africa. Journal of Biogeography 30: 1719–1727.

Slik JWF, Poulsen AD, Ashton PS, Cannon CH, Eichhorn KAO, Kartawinata K, Lanniari I, Nagamasu H, Nakagawa M, Van Nieuwstadt MGL, Payne J, Saridan A, Sidiyasa K, Verburg RW, Webb CO, Wilkie P. 2003. A floristic analysis of the lowland dipterocarp forests of Borneo. Journal of Biogeography 30: 1517–1531.

Stolk JJ. 1907. Opsporing van den zwervenden stam der Penjaboeng-Poenan's op de waterscheiding der Barito met de Mahakan en Kapoeas (Midden-Borneo), in Oct. 1905. Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap 24: 1–27.

Van Balen S. 1996. The ornithological importance of the Danau Sentarum Wildlife reserve in West Kalimantan. Sarawak Museum Journal 50: 123–144.

Van Diessen JR, Van den Brink P, Bek W. 2004. Grote Atlas van Nederlands Oost-Indië / Comprehensive atlas of the Netherlands East Indies. Asia Maior Publishers / KNAG, second extended edition, Zierikzee.

Van Walcheren EWF. 1907. Eene reis naar de bovenstreken van Boeloengan (Midden-Borneo) (12 Nov. 1905 – 11 April 1906). Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap 24: 755–844.

Whittaker RJ, Araújo MB, Jepson P, Ladle RJ, Watson JEM, Willis KJ. 2005. Conservation biogeography: assessment and prospect. Diversity and Distributions 11: 3–23.

Winkler H. 1927. Die planzendecke Süd-Ost Borneos? Institüt für algemeine Botanik, Hamburg.

Witkamp H. 1932. Langs de Mahakam. Tijdschrift van het Koninklijke Nederlandse Aardrijkskundige Genootschap 49: 30–65.

Websites:

La Tierra gazetteer: http://www.tutiempo.net/Tierra/

BioGeomancer: http://www.biogeomancer.org/

SRTM digital elevation data: http://srtm.csi.cgiar.org/

Landsat 7 images: https://zulu.ssc.nasa.gov/mrsid/

JERS-1 SAR radar satellite images: http://www.eorc.jaxa.jp/JERS-1/GFMP/#SEA2/