SPECIES DISTRIBUTION MODELING TO AID REMOTE SENSING OF THE STARCH-RICH SAGO PALM IN THE PHILIPPINES

Meriam M. Santillan[1,2] and Jojene R. Santillan[1]
[1] Caraga Center for Remote Sensing and GIS, College of Engineering and Information Technology, Caraga State University, Ampayon, Butuan City, 8600, Philippines
[2] Division of Geodetic Engineering, College of Engineering and Information Technology, Caraga State University, Ampayon, Butuan City, 8600, Philippines
Email: meriam.makinano@gmail.com, santillan.jr2@gmail.com

KEY WORDS: Species distribution modeling, remote sensing, DOMAIN, Sago palm, Philippines

ABSTRACT: Mapping very specific vegetation using satellite images could be an arduous task especially if the coverage is very large. A good approach would be to avoid processing satellite images where the species is unlikely to be found and focusing only on areas that they are likely to exist. In this paper, we propose the use of species distribution modeling (SDM) as an important procedure prior to actual use of satellite images in mapping uncommon vegetation species such as the Sago palm. Here, SDM through the use of the DOMAIN model together with location data of Sago palms were employed to identify areas where the species are likely to occur. The likelihood of occurrence is computed based on the degree of similarity of bioclimatic conditions of a certain location to those of where Sago palms have been actually found. We used DIVA GIS to implement the DOMAIN model using values of 19 WorldClim bioclimatic variables at selected Sago palm locations. A total of 181 random locations were utilized in the DOMAIN analysis. Another set consisting of 376 randomly selected locations (independent from the DOMAIN set) were utilized to validate the potential species distribution maps (at varying degrees of similarity), specifically to check if how many Sago palm locations may be correctly predicted or omitted if the image analysis will only focus on areas identified by the DOMAIN model as bio-climatically similar. Results show that if the similarity index is to be used as basis to locate specific areas where remote sensing image analysis is to be employed, the accuracy to completely map actual Sago palm locations will decrease when a higher similarity index is selected. For classifications requiring at least 85% accuracy, a similarity index cut-off of 95% can be chosen. It was observed that the SDM approach was found to considerably reduce areal coverage of image analysis while maintaining acceptable accuracy of detecting actual locations of Sago palm. As the approach only targets locations where there is high likelihood of occurrence of a species, it reduces the number of images to analyze and can save time and resources.

1. INTRODUCTION

In this paper, we share our experiences in working out an approach to aid the mapping of an uncommon but highly valuable plant species, the Sago palm (*Metroxylon sagu* Rottb.), using remote sensing. The Sago palm, as illustrated in Figure 1, has a trunk which contains starch. It is considered to be the highest starch producer at 25 tons per hectare per year (Bujang 2008). It is now grown commercially in Malaysia, Indonesia and Papua New Guinea for production of Sago starch and/or conversion to animal food or fuel ethanol (McClatchey et al. 2006).

In the Philippines, the mass propagation and commercial utilization of Sago palm has gained interest from the government in order to develop and sustain a large-scale Sago starch industry that can supply raw materials for food and non-food uses, as well as for production of high value products such as ethanol and lactic acid. However, information on the present location and distribution of Sago palms is incomplete, and it cannot be ascertained whether there is enough supply of Sago logs to drive and sustain a large scale Sago starch industry. Mapping the location of existing Sago palms is necessary in order to determine current supply as well as to characterize its habitat. Once these characteristics have been identified, it is then possible to locate other areas that have the same habitat characteristics for Sago palms to grow. However, Sago palms are considered to be very uncommon in the Philippines. In fact, there were only very few locations in Visayas and Mindanao where Sago palms have been reported to exist, and in most of these locations, they occur as large clusters. Many areas remain to be explored. In this case, the use of satellite remote sensing becomes a convenient tool because of its wide coverage. But mapping very specific vegetation such as the Sago palm using satellite images could be an arduous task especially if the coverage is very large such as Visayas and Mindanao. If mapping is to be done using Landsat images, about 20 scenes are needed to be processed, with land area within the images totaling to 147,000 km^2. Satellite remote sensing has been known to be a useful platform for mapping land-cover in wide areas, but the process of applying it for Sago palm mapping maybe counter-productive considering that this vegetation species is not widely distributed as compared to other species (e.g., coconut). There might be a scenario when one or more images of an area have been acquired and processed but later on it was found that Sago palm cannot exist in this area because of unsuitable conditions. This leads to waste of time and resources. A better approach would be to avoid processing satellite images where Sago palms are unlikely to be found and focusing only on areas that Sago palms are likely to exist.

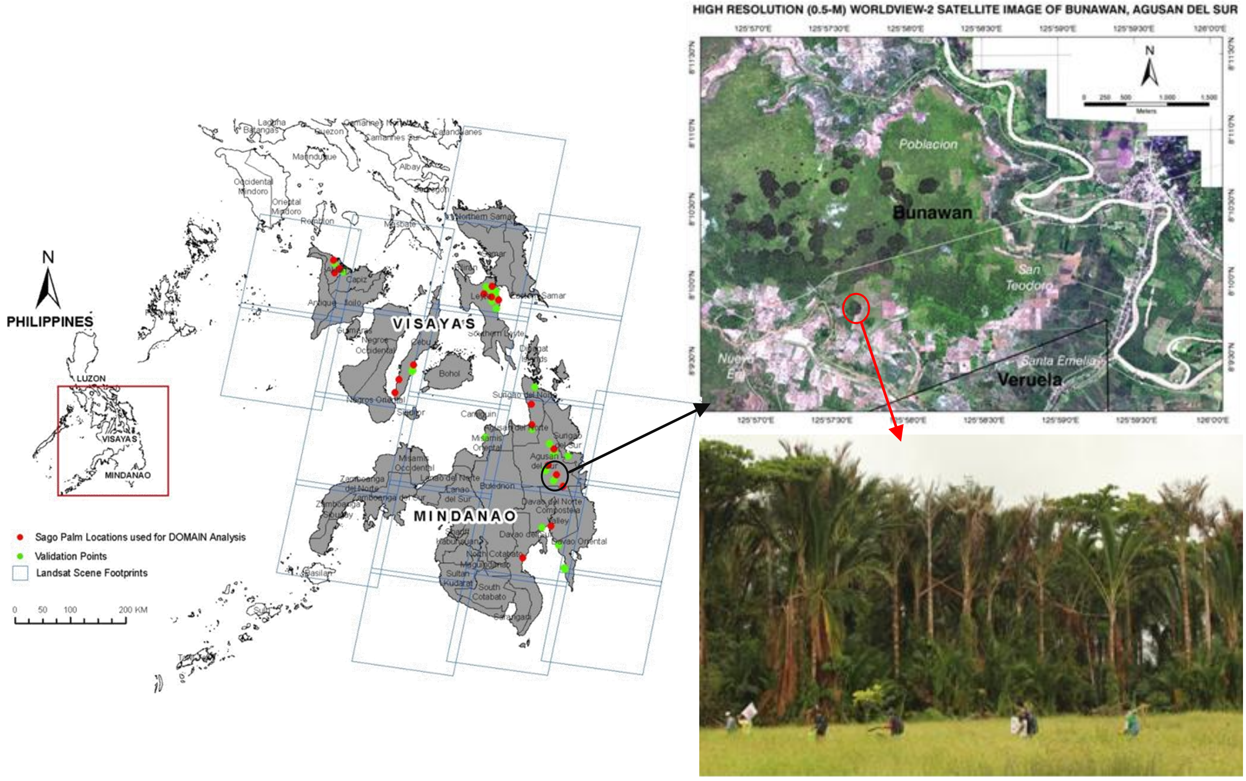

Figure 1. Map showing generalized locations of Sago palms in Visayas and Mindanao, Philippines, overlaid with Landsat scene footprints. Also included is a high resolution Worldview-2 image and photograph of a Sago palm stand.

2. OBJECTIVE

We propose the use of species distribution modeling (SDM) as an important procedure prior to actual use of satellite images in mapping uncommon vegetation species such as the Sago palm. Here, SDM together with location data of Sago palms are used to identify areas where the species are likely to occur. The likelihood of occurrence is computed based on the degree of similarity of bioclimatic conditions of a certain location to those of where Sago palms have been actually found. The importance of SDM is evaluated in terms of reducing the areal coverage to be subjected to image analysis, and the expected mapping accuracy when the approach is to be utilized.

3. MATERIALS AND METHODS

3.1 Sago Palm Data

To illustrate the importance of SDM, we used location data of Sago palms obtained from a series of field surveys conducted between February 2012 - May 2013 in several provinces in Visayas and Mindanao, and enhanced through interpretation of available high resolution Worldview-2 satellite images (UP TCAGP 2013). The data represents 557 unique Sago palm locations.

3.2 Species Distribution Modeling by Similarity of Bioclimatic Conditions

SDM is based upon the theory that the presence of a species in a location obeys three constraints: (i) the local environment allows the population to grow, (ii) the interactions with other local species (predation, competition, mutualism, etc.) allow the species to persist, and (iii) the location is actually accessible, given the dispersal abilities of the species (Hirzel and Le Lay 2008). These constraints determine the geographical distribution of the species. It follows that, theoretically, it is possible to reconstruct a realized ecological niche for a species from the environmental variables measured at the locations it occupies.

For this work, we implemented the SDM of Sago palms using the DOMAIN model (Carpenter et al. 1993). This model derives a point-to-point similarity metric to assign a classification value to a potential site based on its proximity in environmental space to the most similar occurrence. The Gower metric, $d_{AB} = \frac{1}{p}\sum_{k=1}^{p}\frac{|A_k - B_k|}{range(k)}$, which is the sum of the standardized distance between a known point (*A*) and candidate point (*B*) for each predictor variable (*k*), is used to quantify the similarity between two sites. The standardization is achieved using the predictor variable range at the presence sites to equalize the contribution from each predictor variable. Similarity is then calculated by subtracting the distance from 1 and multiplying by 100. The maximum similarity between a candidate point and the set of known occurrences is assigned to each grid cell.

We used DIVA-GIS version 7.5 (Hijmans et al 2012) to implement the DOMAIN model using values of 19 WorldClim bioclimatic variables at selected Sago palm locations. WorldClim (Hijmans et al. 2005), which can be obtained from http://www.worldclim.org/, is a set of 19 global climate layers (climate grids with resolution of ~900 x 900 m) depicting temperature and precipitation that can be used for mapping and spatial modeling in a GIS or with other computer programs (Hijmans et al. 2000). In DIVA GIS, the Domain procedure calculates the Gower distance statistic between each cell on the map and each Sago palm location, using the values of the 19 climate variables. The distance between point A and grid cell B for a single climate variable is calculated as the absolute difference in the values of that variable divided by the range of the variable across all points. The Gower distance is then the mean over all climate variables.

To make the analysis more realistic, we only used Sago palm location data in areas where it has been explicitly reported to exist, which are in Aklan, Cebu and Leyte (in the Visayas), and in the provinces of Agusan del Norte, Agusan del Sur, Davao del Norte and Davao del Sur (in Mindanao). The rationale behind this is that information about the location of Sago palms in other areas aside from those already reported is still unknown - that remote sensing is yet to be used to find them. And since the area covered is large, identifying locations with similar bioclimatic conditions is necessary to minimize the number of images and the effort needed in image analysis. A total of 181 location points (randomly selected and unique in 900 x 900m grid cells of the WorldClim layers) were utilized in the DOMAIN analysis. Another set consisting of 376 randomly selected locations (independent from the DOMAIN set) were utilized to validate the potential distribution maps (at varying degrees of similarity), specifically to check if how many Sago palm locations were correctly predicted or omitted if the image analysis will only focus on areas identified by the DOMAIN model as bio-climatically similar. The validation points are located in different provinces including those provinces where Sago palm location data has not been used in the DOMAIN analysis.

4. RESULTS AND DISCUSSIONS

The similarity index map derived using the DOMAIN model for Visayas and Mindanao is shown in Figure 2. The percentage values indicate the similarity in the bioclimatic conditions of a certain location to those in the actual locations of Sago palms. It can be noticed that the likelihood of finding a Sago palm decreases as the similarity index increases. It follows that as the similarity index increases, there will also be a decrease in area where Sago palm is likely to be present (Figure 3).

If the similarity index is to be used as basis to locate specific areas where remote sensing image analysis is to be employed, the accuracy to completely map actual Sago palm locations will decrease when a higher similarity index is selected. As shown in Figure 4, a similarity index cut-off of greater than or equal to 86% will still provide 100% correct prediction of Sago palm locations even though there is a significant decrease in land area. Further examination revealed that the accuracy decreases to 99% when the similarity index is greater than or equal to 87%, and this continues to decrease until the accuracy is only 50% when the similarity index is 100%. This pattern is expected as the area where Sago palm is likely to be present becomes more specific and the conditions now equate to the bioclimatic

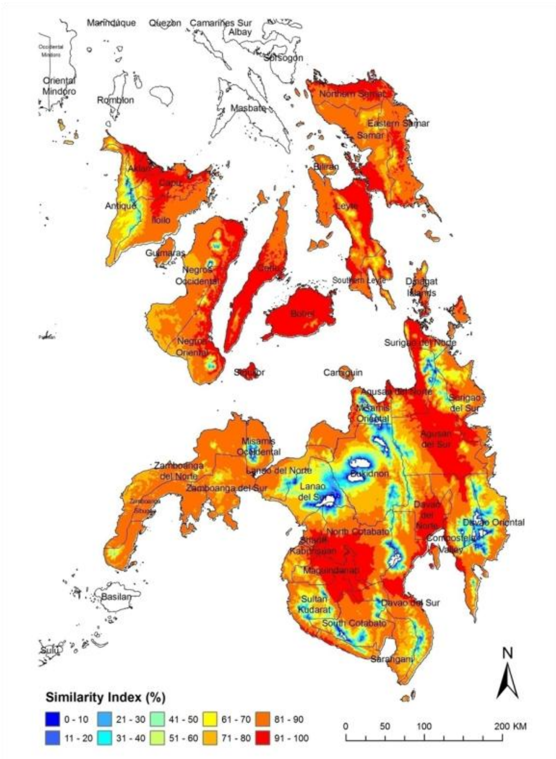

Figure 2. Similarity index map derived using the DOMAIN model which is indicative of the likelihood of occurrence of Sago palms.

conditions at the observed locations. For classifications requiring only a minimum of 85% accuracy, a similarity index cut-off of 95% can be chosen (i.e., to select areas with >95% index).

It can be observed that there will be a considerable decrease in land area when species distribution modeling using the DOMAIN model is employed prior to image analysis. The original area of 147,000 km^2 was greatly reduced to ~64,000 km^2 when only those areas with 86% similarity index are chosen (Figure 4). This translates to 56% reduction in coverage area for image analysis which is good in terms of resources and time management. A 95% similarity index cut-off which could provide at least 85% accuracy in detecting Sago palms will only require ~13,000 km^2 of land area to be analyzed.

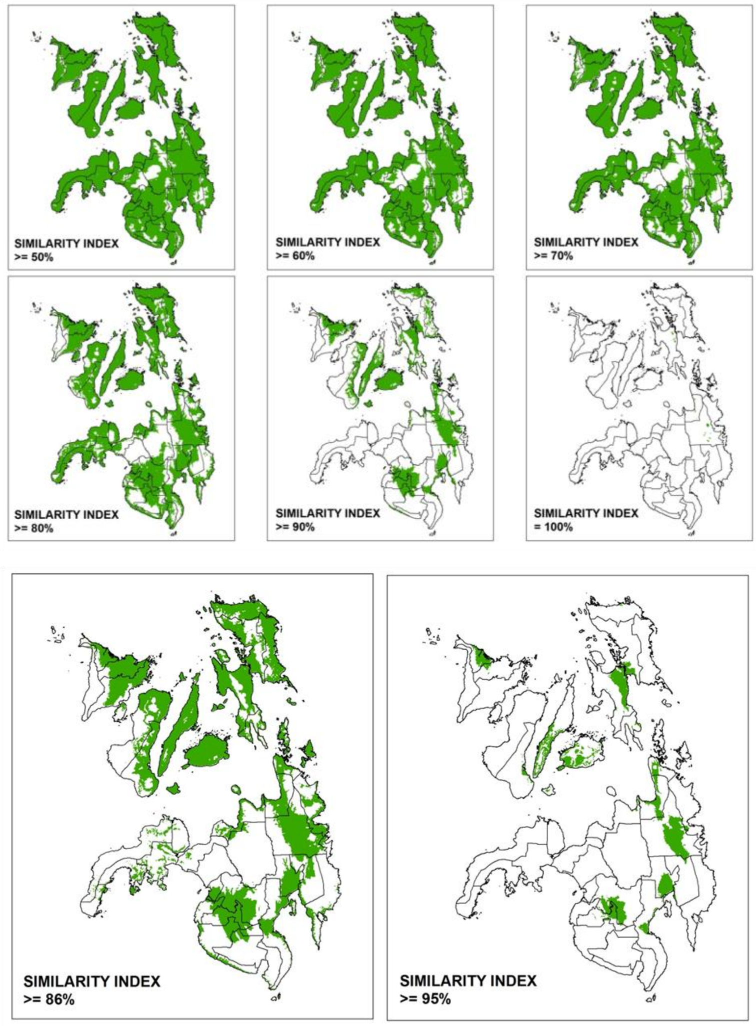

Figure 3. Series of maps showing locations at different similarity index cut-offs which are useful in selecting areas where images can be analyzed to detect Sago palms.

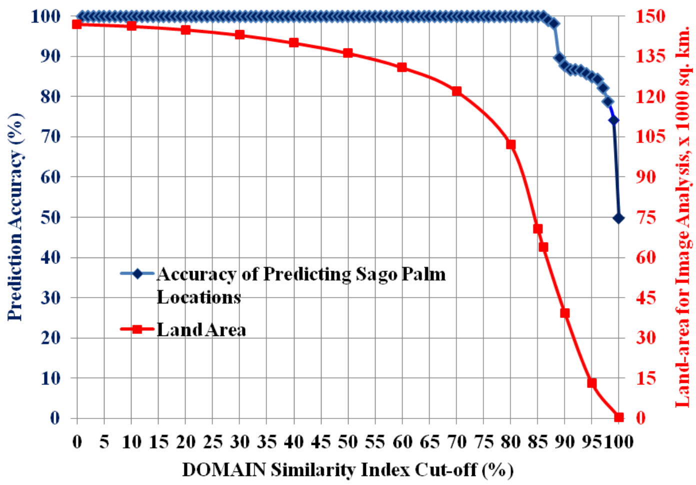

Figure 4. Graph showing accuracy of predicting Sago palm locations and land area for image analysis in various percentages of similarity indexes.

5. CONCLUSIONS

In this paper, we showed that employing SDM using the DOMAIN model can be a valuable procedure prior to employing remote sensing image analysis to map uncommon vegetation species such as the Sago palm. The SDM approach was found to considerably reduce areal coverage of image analysis while maintaining acceptable accuracy of detecting actual locations of Sago palm. As the approach only targets locations where there is high likelihood of occurrence of a species, it reduces the number of images to analyze and can save time and resources. Future studies are needed to verify the applicability of the approach as a preliminary procedure in remote sensing-based mapping of other uncommon vegetation species.

ACKNOWLEDGEMENTS

This paper is an output of Projects II.2 and II.3 under the “Sago Bioresource Assessment for Sustainable Industry Utilization Using Remote Sensing Geospatial Analyses and Suitability Relationships” research program. We are grateful to the Philippine Council for Industry, Energy and Emerging Technology Research and Development of the Department of Science and Technology (PCIEERD-DOST) for the financial support. Portions of this research was conducted in 2013 when the second author was still affiliated with the Research Laboratory for Applied Geodesy & Space Technology of the Training Center for Applied Geodesy and Photogrammetry & Department of Geodetic Engineering, University of the Philippines, Diliman, Quezon City.

REFERENCES

Bujang, K.B., 2008. Potentials of Bioenergy from the Sago Industries in Malaysia. In: Encyclopedia of Life Support Systems. Retrieved May 15, 2013, from http://www.eolss.net/sample-chapters/c17/E6-58-12-12.pdf.

Carpenter, G., Gillison, A.N., Winter, J., 1993. DOMAIN: a flexible modeling procedure for mapping potential distributions of plants and animals. Biodiversity and Conservation, 2(6), pp. 667-680.

Hijmans, R.J., 2012. DIVA-GIS Version 7.5. www.diva-gis.org, International Potato Center, Peru.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25(15), pp. 1965-1978.

Hijmans, R.J., Guarino, L., Cruz, M., Rojas, E., 2001. Computer tools for spatial analysis of plant genetic resources data. Plant Genetic Resources Newsletter, 127, pp. 15-19.

Hirzel, A.H., Le Lay, G., 2008. Habitat suitability modelling and niche theory. Journal of Applied Ecology, 45(5), pp. 1372-1381.

McClatchey, W., Manner, H.I., Elevitch, C.R., 2006, Metroxylon amicarum, M. paulcoxii, M. sagu, M. salomonense, M. vitiense, and M. warburgii (sago palm), In: Species Profiles for Pacific Island Agroforestry, edited by Elevitch, C.R., Permanent Agriculture Resources (PAR), Holualoa, Hawaii.

UP TCAGP, 2013. Terminal Report: Project II.3. Mapping Sago Habitats and Sago Suitable Sites using Optical and Radar Image Analysis and Suitability Relationships, Training Center for Applied Geodesy and Photogrammetry, University of the Philippines, Diliman, Quezon City.