Thailand National Biodiversity Database System with *webMathematica* and Google Earth

W. Katsarapong, W. Srisang, K. Jaroensutasinee, and M. Jaroensutasinee

***Abstract*—**National Biodiversity Database System (NBIDS) has been developed for collecting Thai biodiversity data. The goal of this project is to provide advanced tools for querying, analyzing, modeling, and visualizing patterns of species distribution for researchers and scientists. NBIDS data record two types of datasets: biodiversity data and environmental data. Biodiversity data are specie presence data and species status. The attributes of biodiversity data can be further classified into two groups: universal and project-specific attributes. Universal attributes are attributes that are common to all of the records, e.g. X/Y coordinates, year, and collector name. Project-specific attributes are attributes that are unique to one or a few projects, e.g., flowering stage. Environmental data include atmospheric data, hydrology data, soil data, and land cover data collecting by using GLOBE protocols. We have developed web-based tools for data entry. Google Earth KML and ArcGIS were used as tools for map visualization. *webMathematica* was used for simple data visualization and also for advanced data analysis and visualization, e.g., spatial interpolation, and statistical analysis. NBIDS will be used by park rangers at Khao Nan National Park, and researchers.

***Keywords*—**GLOBE protocol, Biodiversity, Database System, ArcGIS, Google Earth and webMathematica.

I. Introduction

BIODIVERSITY Database is database for collecting biodiversity data. Biodiversity data refers to scientific information, primarily about biological species and specimens. At the species level, such data would include the scientific names of the species and all of its synonyms; the common name(s) of the species; and other information about the species, such as a description of the species, its physiological properties, genetics, geographic distribution, phylogenetic relationships, role in the dynamics of ecosystem processes including cases of invasions, applications, and etc. Specimen-level data including samples for molecular analysis, would include the scientific name of the species to which the specimen belongs; information on where, when and by whom the specimen was collected; where the specimen is currently located; who identified it; what is the specimen number; and other associated information derived from the specimen (e.g., living culture, frozen tissues, photographs, parasites, and hosts) and any other related field notes written by the collector of the specimen.

Because of humanity's dependence on natural systems, information about biodiversity and ecology is vital to a wide range of scientific, educational, commercial, and governmental uses. Biodiversity and ecosystems are themselves interdependent. Ecosystems and the diversity of species they support underpin our lives and our economies in very real, though often underappreciated, ways. The living things with which we share the planet provide us with clean air, clean water, food, clothing, shelter, medicines, and aesthetic enjoyment. Yet, increasing human populations and their activities are disturbing species and their habitats, disrupting natural ecological processes, and even changing climate patterns on a global scale. These are greater stresses on the natural world than humanity has ever generated in the past. Since biodiversity is arguably the most precious resource on Earth, it is becoming more and more important that we actively conserve biodiversity and protect natural ecosystems in order to preserve the quality of human life. As human populations and their demands on the natural world grow, our accumulated knowledge about biodiversity and the environment will become ever more important in the effort to develop a sustainable world.

Recognition of this has led to the National Biological Information Infrastructure in the United States, to the Environmental Resources Information Network in Australia, and to a number of regional biodiversity information networks (NABIN, IABIN, EIOnet, and others). Indeed, the recommendation by an international working group established by the Global Science Forum (formerly Megascience Forum) of the Organization for Economic Cooperation and Development (OECD) that the nations of the world establish and maintain a Global Biodiversity Information Facility (GBIF), which is poised to become a reality in early 2001, is a direct outgrowth of both concern about the environment and the economy, and the acknowledgment that the complexity of biodiversity and

Manuscript received November 30, 2006. This work was supported in part by the Biodiversity Research and Training, and Complex System research unit, Institute of Research and Development, Walailak University.

W. Katsarapong is with School of Science, Walailak University, 222 Thaiburi, Thasala District, Nakhon Si Thammarat, 80161, Thailand (phone: +6675672005 ; fax: +6675672038; e-mail: manee601@hotmail.com).

W. Srisang is with School of Science, Walailak University, 222 Thaiburi, Thasala District, Nakhon Si Thammarat, 80161, Thailand (phone: +6675672005 ; fax: +6675672038; e-mail: wsrisang@gmail.com).

K. Jaroensutasinee is with School of Science, Walailak University, 222 Thaiburi, Thasala District, Nakhon Si Thammarat, 80161, Thailand (phone: +66 75672005; fax: +6675672004; e-mail: krisanadej@gmail.com).

M. Jaroensutasinee is with School of Science, Walailak University, 222 Thaiburi, Thasala District, Nakhon Si Thammarat, 80161, Thailand (phone: +6675672005; fax: +6675672004; e-mail: jmullica@wu.ac.th).

ecological datasets reflects the complexity of natural systems. It has become apparent that practitioners in the computer science and information technology fields must become as invigorated by and invested in the biodiversity and ecological information domain as are the biologists, who collect, generate, query, and interpret the data [1], [2].

Thailand National Biodiversity Database System (NBIDS) has been developed by Walailak University team and funded by Biodiversity Research Training Program (BRT). The goal of this project is to provide advanced tools for querying, analyzing, modeling, visualizing patterns of species distribution found in Thailand for researchers and scientists. This paper is discussing the first year prototype of NBIDS in term of database development, application of NBIDS, and the incorporation between Biodiversity database and The GLOBE program.

II. Materials and Methods

NBIDS is a web base system designed for four main features: database, data analysis tool, data visualization tools, and GIS tools. NBIDS database is developed using SQL technology. We have developed web-based tools for data entry and data access. NBIDS data record two types of datasets: biodiversity data and environmental data. Biodiversity data are species presence data and species status. The attributes of biodiversity data can be further classified into two groups: universal and project-specific attributes. Universal attributes are attributes that are common to all of the records, e.g. X/Y coordinates, year, and collector name. Project-specific attributes are attributes that are unique to one or a few projects, e.g., flowering stage. Environmental data include atmosphere, hydrology, soil, and land cover data using GLOBE protocols.

Data analysis tools for NBIDS are statistical analysis tools and computational modules for each research project. Examples of computation modules' outputs are biodiversity index, mosquito house index, coral data and fish morphometric data.

Data visualization tool is developed using *webMathematica* [3] technology. This tool is an interaction tool for visualizing graph by high performance computing power of *Mathematica* software.

Google Earth is well used for GIS visualization. Other information is added to the default Google Earth map. This information includes Khao Nan national park boundary, national park stations, landmark, administration map, transportation, LandSat images, and geo-computing data. Geo-computing data have been developed using ArcGIS program. These data include DEM, aspect, and flow direction.

NBIDS has five user types: system manager, project manager, researcher, senior scientist, and system administrator. Project manger is a principle investigator of each project. Project manager collects data in the field and inputs data on the website. Project manager views and makes changes his own data. System manager and senior scientist have accessed to and viewed data from all projects. However, only system manager and system administrator can make some changes and modification all data and system.

III. Results

Prototype NBIDS is now online at URL http://www.nbids.org since November 2005. NBIDS is tested with data collected from Khao Nan national park, Nakhon Si Thammarat, southern of Thailand.

A. Data

Now NBIDS contains 23 sub-projects, 5386 biodiversity data records; 182 species: 9 vascular plants species, 10 fish species, 91 coral species, 52 reptiles species, 15 butterfly species, and 5 mosquito species, and 213 Environmental data records from 8 permanent study sites and 1781 field study sites.

B. Web Tools

1. We can do data entry to NBIDS using a web-service (Fig. 1).

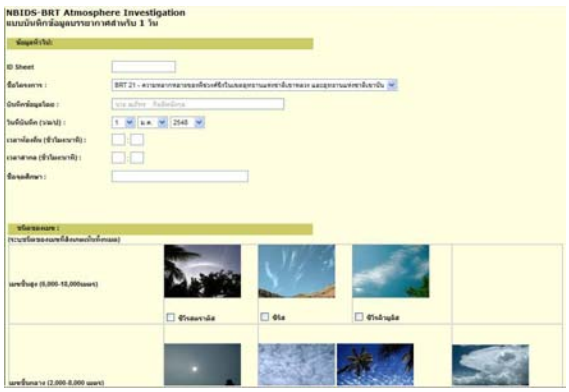

Fig. 1 Example of NBIDS data entry page

2. We can search information from NBIDS such as study sites, species name, common name, family, physical parameters and date (Fig. 2).

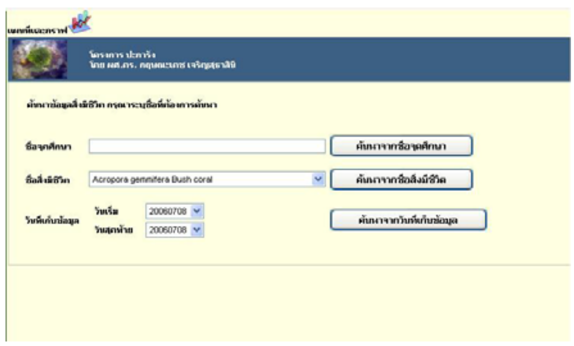

Fig. 2 Example of NBIDS search page

3. We can do data visualization in NBIDS using *webMathematica* (Fig. 3).

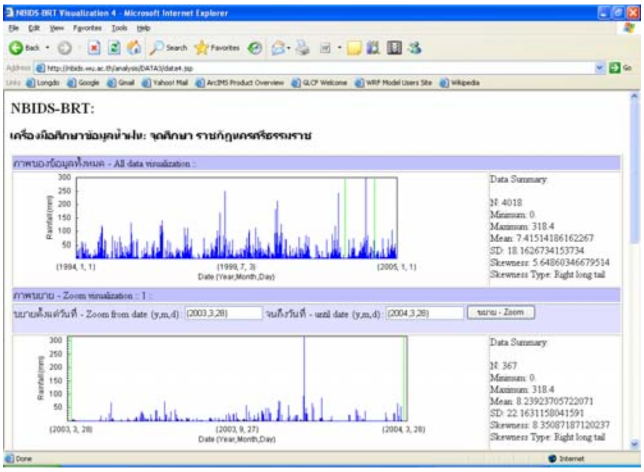

Fig. 3 NBIDS visualization tool, interactive graph and descriptive statistics using *webMathematica*

4. NBIDS shows locations of study sites on Google Earth (Fig. 4).

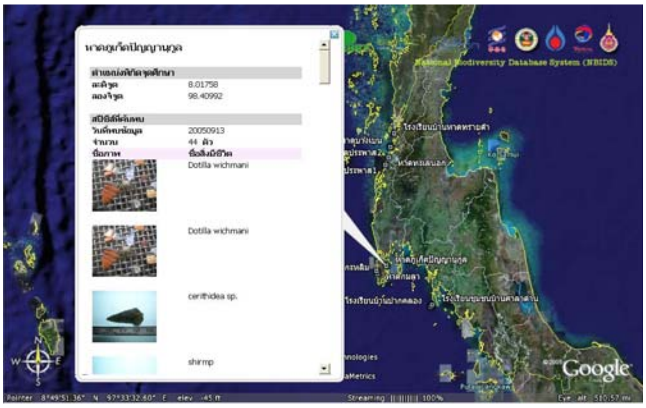

Fig. 4 NBIDS visualization tool, descriptive area and species data using Google Earth

IV. Discussions

GIS tools of NBIDS can help scientists and researchers to plan their research because the tool developed is compatible with Google Earth which is easy to use. This Google Earth can demonstrate maps, area boundary, transportation, and LandSat images (Fig. 5, and Fig. 6). With these pictures, NBIDS can help researchers to understand area, select their study sites effectively and plan their experiments appropriately. When scientists are doing their research, they can use this GIS tool for observing and constructing some relationship between geographical data, environmental data, and species presence data. Furthermore, scientists could model niche characterization and potential distribution of species using some mathematical and computational methods. Tools for a mathematical modeling have been planed to add in NBIDS in the near future. NBIDS is an effective tool for studying the relationship among species. All NBIDS data are stored in the same universal attributes that make these data comparable. For example, coordinates of species occurrence are collected in the same units that make the study possible and luminous. NBIDS data are stored in a security system. Only permitted users can access to their own data. However, when scientists need to compare the relation among species, permission for accessing data can be granted by the Principle Investigators of the projects.

Fig. 5 NBIDS GIS tools, Google Earth map with additional LandSat images and Khao Nan national park boundery showing in red color

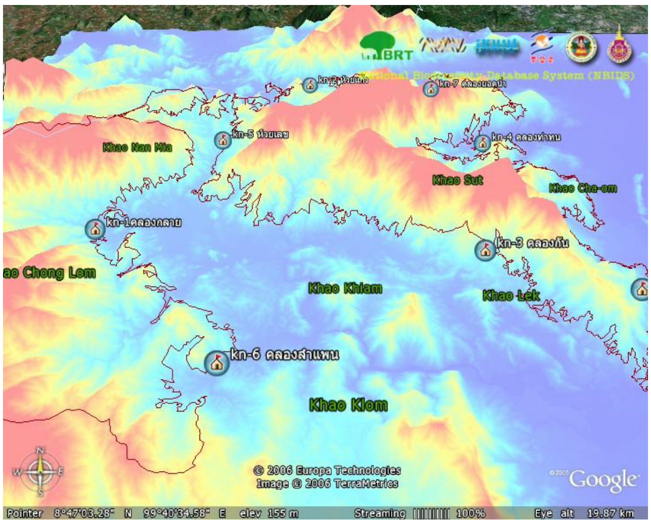

Fig. 6 NBIDS GIS tools, Google Earth map with additional DEM calculated using ArcGIS software, Khao Nan national park boundary and Khao Nan park stations showing in flags as park stations no. 1-8

V. Case Study

We selected one case study to demonstrate how NBIDS was used. The case study was "Seasonal Prevalence of *Aedes aegypti* and *Ae. albopictus* in Three Topographical Areas of Southern Thailand." This study investigated the seasonal prevalence of *Ae. aegypti* and *Ae. albopictus* larvae in three topographical areas (i.e. mangrove, rice paddy, and mountainous areas). Samples were collected from 300 households in dry and wet season in nine districts in Nakhon Si Thammarat province. *Ae. aegypti* and *Ae. albopictus* were found in 17 out of 26 types of water containers in mangrove, rice paddy, and mountainous areas. *Ae. aegypti* and *Ae. albopictus* laid eggs in different container types depending on topographical areas. *Ae. aegypti* larvae were found most in preserved areca jars in mangrove and mountainous areas and in banana trees in rice paddy areas. *Ae. albopictus* larvae were found most in preserved areca jars in mangrove areas, in plant plates in rice paddy areas, and in metal boxes in mountainous areas. All *Ae. albopictus* larval indices were higher than *Ae. aegypti* larval indices in all three topographical areas. HI and BI were different between mosquito species whereas not different in three topographical areas. Both of HI for *Ae. aegypti* and *Ae. albopictus* in all three topographical areas were greater than 10%, which indicated high risks of DHF transmission in these areas [4] (Fig. 7).

Fig. 7 NBIDS GIS tools, and Google Earth map with Seasonal Prevalence of *Aedes aegypti* and *Ae. albopictus* in Three Topographical Areas of Southern Thailand

Acknowledgments

This work was supported in part by PTT Pubic Company Limited, the TRF/BIOTEC Special Program for Biodiversity Research and Training grant BRT T_549002, Complex System research unit, the Institute of Research and Development, Walailak University, and DPST project (to W. Srisang), IPST, Thailand.

References

[1] R. M. Chefaoui, J. Hortal, and J. M. Lobo, "Potential distribution modelling, niche characterization and conservation status assessment using GIS tools: a case study of *Iberian copris* species," Biological Conservation, vol. 122, pp. 327-338, 2005.

[2] M. A. Lane, J. L. Edwards, and E. S. Nielsen, "The Challenge of Rapid Development, Large Databases and Complex Data," in *Proc. 26th Inter. Conf. Very Large Databases*, Cairo, Egypt, 2000.

[3] S. Wolfram, *The Mathematica Book* 5th ed. Wolfram Media, 2003.

[4] W. Preechaporn, M. Jaroensutasinee, and K. Jaroensutasinee, "Seasonal Prevalence of *Aedes aegypti* and *Ae. albopictus* in Three Topographical Areas of Southern Thailand". Enformatica Transactions on Engineering, Computing and Technology, vol. 19, 2007. to be published.