Abstract

This work is based on a review of current scientific literature describing problems of different kinds of using GIS technologies in plant genetic resources (PGR) research. The first part of this work is focused on introduction to basic principles of using GIS and the most used computer tools for PGR management, such as DIVA-GIS and FloraMap. There is described the process of creating of maps and analyses from different kinds of germplasm data. The second part describing using GIS in concretes case of study, the biggest attention was given to problems of PGR conservation and use.

Key words: GIS, plant genetic resources, germplasm, conservation,

Abstrakt

Tato práce je přehledem současné vědecké literatury, jež se věnuje možnostem nejrůznějšího využití GIS technologii při hledání rostlinných genetických zdrojů. První část práce je zaměřena na základní principy použití GIS technologií, pozornost je též věnována nejpoužívanějším počítačovým programům; DIVA-GIS a FloraMap. Na základě dat popisujících genetický materiál, je možno vytvářet mapy nebo tyto data dále různým způsobem analyzovat. Základy těchto postupů jsou též nastíněny. Druhá část této práce se zabývá použitím GIS technologií v konkrétních odborných studiích, největší pozornost je věnována péči o zachování rostlinných genetických zdrojů a jejich následnému využití.

.

Klíčová slova: GIS, rostlinné genetické zdroje, zdroj genetického materiálu, péče o zachování,

Foreword

Plants provide the raw materials we rely upon for food, fiber, medicine and industrial products. Quite significant number of plants is threatened; mainly by human activities, for example by environmental pollution, deforestation, cattle breeding, spread of intensive agriculture etc. On the other hand, humans also provide protection of plant genetic resources, including the protection of wild species, plant populations and traditional crop varieties where they have evolved (*in situ* conservation), and collections of germplasm in genebanks and botanical garden (*ex situ* conservation). Plant germplasm is a natural resource of genetic materials that is conserved in the form of seed or live plants, representing current, obsolete and primitive crop varieties, wild and weedy relatives of crop species. Germplasm of wild species are collected or donated from around the world, and this material is conserved, but also available to a broad scientific community for basic research and development into crop cultivars. Information associated with germplasm accessions is as important as the actual plant material held in gene banks. This information includes various types of data, such as passport data, describing the identity; and characterization and evaluation data, describing germplasm. Data associated with germplasm are geo-referenced, and therefore can be analyzed in conjunction with other location data, from whatever additional source. For example climatic, vegetation and soil data from the study region can be used to estimate the distribution and phenology of the taxon concerned.

In this work, is described how spatial analysis of the geo-referenced data generated by the process of conservation and use genetic resources, using geographic information system (GIS), can feedback to enhance and facilitate the process, and indeed add value to the germplasm collections. This work starts with first part, which is about introduction to GIS technology and the second part is focused on the number of key stages of the genetic resources management process.

.

1. Introduction

1.1 Geographic Information System

Geographic information systems (GIS) were created on the basis of connection of traditional disciplines like cartography and geography, and modern disciplines like electronic computing and remote sensing. A GIS can be simply defined as 'a tool for managing information of any kind according to where it is located' (Treweek, 1999). It is a database management system, which simulate spatial data in graphics form, i.e. maps, and related, logically attached, non-spatial, attribute data, i.e. the labels and descriptions of the different areas within a map. Different kinds of information are geographically tagged ("georeferenced") and displayed together on a single map. By the combination of two or more data sets it is possible to view relationships between selected social or biophysical factors crops, livestock, forests, soils, waterways, climate, topography, population distribution, income patterns, education levels, roads, land tenure, administrative boundaries, and so on. In shortcut GIS is a tool for analysis and reanalysis, making and remaking maps, mixing and matching different kinds of information depending on what is of interest.

The main elements of a GIS are *data sources, data input, verification and editing, data storage, retrieval and management, data manipulation and analysis, and output.*

DATA SOURCES

There are two types of data, spatial and attribute, which are input into a GIS. Both types of data are encoded into a GIS database formats. Cornerstone of successful GIS implementation is establishment and maintenance of a robust spatial database. The most common general sources for spatial data are: *analogue maps, aerial photographs, sensing images, point data samples from surveys, existing digital data files*

Some regional or global scale datasets are available from organizations such as the Food and Agriculture Organization of the United Nations (FAO), the United Nations Environment Programme's Global Resource Information Database (UNEP/GRID), the International Soil Reference and Information Centre (ISRIC), the World Conservation Monitoring Centre (WCMC), the US Geological Survey and the international agricultural research institutes

(IARCs) of the Consultative Group on International Agricultural Research (CGIAR).Attribute data has an even wider variety of data sources. Any textual or tabular data than can be referenced to a geographic feature, e.g. a point, line, or area, can be input into a GIS.

DATA INPUT

There are four basic methods for inputting spatial data into GIS, which can be used singly or in combination: *manual digitizing, automatic scanning, coordinate geometry, using of existing digital data.*

Majority of GIS spatial data entry are done by manual digitizing, with using of *digitizer*. Which is an electronic gadget comprises a table upon which the map or drawing is placed. A hand-held magnetic pen traces spatial features. The coordinates of selected point are sent to the computer and stored. One of the most desirable functions of GIS is the ability of data transforming during digitizing process. Digitizing can be done in different modes, such as a *point mode* (single points are recorded one at a time), a *stream mode* (a point is collected on regular intervals of time or distance, e.g. every 3 meters) and a *spaghetti mode*s (lines are digitized by indicating a start point and an end point) etc.

The biggest advantage of current scanning technologies is the ability of capturing spatial features from map at a rapid rate of speed. There are also a lot of disadvantages such as high acquisition costs, problems with text and symbol recognition, need of a substantial amount of manual editing to create a clean data layer etc. In shortcut, scanners work best when the information on a map is kept very clean, very simple and without graphic symbology.

Coordinate geometry is very costly and labour intensive. This input technique is used for creating very precise cartographic definitions of property, and accordingly is more fitting for land records management at the cadastral or municipal scale and it is seldom used for natural resource applications in GIS.

The technique of using of existing digital data is becoming more and more popular because spatial data, like digital maps or regional and global scale datasets etc., can be easily found in a wide range of government and private sources. It is also possible to used data from CAD system. Because of wide variety of existing data formats, was developed data exchange/conversion software, which is easily available on the market.

The most time consuming operations are input, cleaning, linking and verification of the data, they take 60 to 80% of time, which is needed to complete the project. During the encoding

of spatial and non-spatial data can arise many kinds of errors so it is necessary to make data editing and verification.

DATA STORAGE

There is difference between storage of vector and rastr file. Vector files store geographic data as points, lines or polygons. Polygons represent areas of different sizes and shapes where a particular attribute is equal throughout. The raster data model has become the primary spatial data source for analytical modeling with GIS. The raster data model is well suited to the quantitative analysis of numerous data layers. To facilitate these raster modeling techniques most GIS software employs a separate module specifically for cell processing. Each individual cell is assigned a value for the variable being studied. Examples of vector data include maps of roads, or administrative zones. Typical raster datasets include satellite images, elevation (so called Digital Elevation Models) and interpolated climate data. Each type of data has advantages and disadvantages. However, modern GIS packages can handle both types of data, and analyses can be made across the two data types.

There are two basic types of storage, lokal (desktop) and linked (server). Local storage is used for storing of individual files, like Microsoft Office files (**.doc**, **.xls**, etc). This is not very useful for multiuser applicaton or enterprises. Linked or server storage is used for data storing in centralized DBMS (Database Management System). The advantage is that users can connect to link whatever data they need, so in this way it supports multi-user work.

DATA MANIPULATION AND ANALYSIS

The main criteria using to define a GIS, is its capability to transform and integrate spatial data. The spatial processing system (to manipulate the 'where') and database management system (for the 'what') of a GIS allow the user to bring together diverse datasets, make them compatible, and combine and analyse them. A distinction has been made in describing the analytical capabilities of GIS between deterministic and statistical methods.
Deterministic functionalities provided by GIS include network analysis, three-dimensional modelling and projection algebra. Spatial analysis includes the tabulation or mapping of basic summary statistics for data in areas of interest. Relevant tools include query facilities,

Boolean operations on attributes, map overlay facilities and buffer creation. More complex statistical approaches to the investigation, comparison and modelling of geographical patterns, includes spatial correlation, pattern and trend analysis, and interpolation techniques such as kriging.

DATA OUTPUT

There are several ways of outputting data from GIS. Displaying data on the screen is the most common way of outputting data from a GIS. Printing is probably the second most common way of data output. Printed maps allow such manipulations as selecting areas or layers for output, and changing scale and colour. The outputs of GIS analyses can also include tables, graphs and animations.

There are 2 most important computer tools for spatial analysis of plant genetic resources data, DIVA-GIS and Flora-Map.

1.1.1. DIVA-GIS

DIVA-GIS was created for scientists who cannot afford generic commercial GIS software, or do not have the time to learn how to use it, and for others who require a GIS that is specifically designed for genetic resources work. The DIVA-GIS software allows analysis of genebank and herbarium databases to elucidate genetic, ecological and geographic patterns in the distribution of crops and wild species. Analytical functions implemented in DIVA include mapping of richness and diversity, distribution of useful traits and location of areas with complementary diversity. DIVA helps improve data quality by assigning coordinates, using a large digital gazetteer. DIVA can also extract climate data for all terrestrial locations, which can be used to describe the environment of collection sites.

DIVA-GIS version 1.4, was developed at the International Potato Center (CIP) in collaboration with the International Plant Genetic Resources Institute (IPGRI), and with support from the System-wide Genetic Resources Program (SGRP). It is available free of charge from the CIP website (http:// gis.cip.cgiar.org). Base-map data (e.g. administrative boundaries altitude) for use with DIVA are also provided for use with DIVA via the Internet.

DIVA is allowed to import genebank databases including passport, characterization and evaluation data, using latitude and longitude fields. Sometimes are latitude and longitude unknown. In this case DIVA can assign the most likely coordinates only from locality information.(such as department, province and place name).

DIVA DESKTOP

Screen is made up of a map and its allied legend. A map is created using themes, what are geo-referenced databases. Each theme on the map is listed in the legend. With the themes such as altitude, national boundaries, main rivers and capital cities, can be made up map of the world. To manage the content of the map three menus are available: *file* (file and project managment, exporting data and maps, printing), *theme* (inspecting and managing individual themes), *view* (managing the map, e.g. zooming in and out)

There are also three additional menus, such as analysis, tools and help. The functions in these menus are discussed in more detail below.

DIVA TOOLS

Lack of coordinate data in genebank or data inaccuracy, seriously complicates spatial analysis of genebank data and make the results unreliable. The task of improving the data quality can be easier made by DIVA. DIVA can assign coordinates to accessions that have a locality description but no coordinates, and can help verify the accuracy of accessions that do have coordinates. The best system for the digital notation of geographic coordinates is decimal degrees. Whereas the sexagemal system contains numbers, symbols and letters (e.g., 12°34?12??S) the decimal system only has a number, and no letters, with the sign indicating the hemisphere (+ = N or W, - = S or E) (e.g., -12.5700). Decimal degrees should be stored with 4 or 5 decimals. At the equator, one unit of the fourth decimal (0.0001 degrees) equals about 10 meters (less at other latitudes; not affected by longitude). That should be precise enough, although in some countries people can use differential GPS (Global Positioning System) during collecting and have precision at the meter level. In those cases, 5-digits would be better.(Hijmans, 2001)

There is often lack of coordinte data in genebank databases, especially in older collections. For example, only 9% the accessions of six major genebanks of the United States Department of Agriculture genebanks have coordinates. That means that there is opportunity for

assigning coordinates to at least another 41% of the accessions (most accessions with coordinate data will also have a locality description).There can be used searching for the locality names on maps or in gazetteer. This is a list of names of geographic features and the coordinates of their locations. Digital gazetteers can facilitate searching.

DIVA is available to assign the coordinates to localities. The user defines an input file that should ideally have the following fields indicating where the accession was found: country, first and second administrative subdivision and up to two locality names. For both locality names one can indicate the distance and direction from the collecting site to that locality. Then is used a digital gazetteer (the database of foreign geographic feature names from the U.S. National Imagery and Mapping Agency) for searching the locality and assigning its coordinates to the accession.

When assignment of coordinates to all records is complete, data are imported to create a shapefile. This kind of file format is used to display data on the map. All the user has to do is specify a .dbf file, and the two numerical fields in that file that contain the altitude and longitude data.

When is shapefile made from a genebank database, it is sometimes full of errors, like spots on the wrong continent or falling in lakes. These questionable locations are easy to mark and correct. However, there will also be other errors than cannot be spotted so easilly. By simultaneously checking the accessions database and an administrative boundaries database, a new (temporary) database is created. For each accession, this new database contains the location names according to the genebank database *and* according to the administrative boundaries database. These names should be the same, and any mismatches probably reflect errors (or name or boundary changes!). By extract tool are assinged environmental data points. This can be useful because traits can be related to ecological conditions at the places the collections occurred. Currently, monthly mean data for minimum and maximum temperature and precipitation are included, but in future versions of DIVA more environmental variables will be available.

DATA ANALYSIS WITH GRIDS

Most of the various DIVA analyses are based on grids. By these grids is world divided into equal sized changeable cells. Potential calculation is then executed on each of the cells. Equal-sized cells are compared more objectively than areas such as countries or administrative region. This is one of the reason why is using of grids more and more popular.

The number of observations in each grid cell can be determined. Irrelevant points in the shape file may be excluded by (de)selecting on the basis of their value for a specific field in the database that describes the points.

There are three methods how can be the numbers of observation determinate into grid cell: The'Circular Neighborhood', 'Inverse Distance-Weighted' and 'Simple' methods. By the circular neighborhood option is calculation made for a circle with its center in the middle of a grid cell and specified radius. By using IDW method are inverse distance-weighted values assigned to the four nearest grid cells. By the 'simple'method are points selected into the grid cell they fall in. The weakness of this method is that a point that is on the border between grid cells is randomly assigned to one grid cell only. The nearness of the point to other grid cells is irrespected.

The function number of different values calculates the number of distinct classes of a certain variable that occurs in each grid cell. It means that input database can be consisted of the locations where different wild species are observed. The database field that indicates the species names can be selected and the number of different species per grid cell will be produced. Number of diversity indices is also calculated for each grid cell. From input database is selected a field for which a diversity index is calculated.

The complementary site selection procedure try to recognize sets of grid cells that are complementary to each other, it means to capture a maximum amount of diversity in as few cells as possible. This procedure is based on the algorithm described by Rebelo (1994). While the first cell choice is easy – the cell with highest species richness – the selection of the following cells depends on the previously chosen cell. This is because the species in the cell with the second highest number of species may also be present in the first cell. This is a non-linear optimization problem. Rebelo (1994) developed a procedure that calculates an approximate optimal solution, and this has been implemented in DIVA.

This procedure is used again and again. The value of each grid is calculated repetitively, based on the observations in that cell, and in relation to the observations in the cells already selected. If there are two or more cells with the same value, one is selected at random. Therefore, this procedure can lead to slightly different results in every time.

Statistics for each grid can be calculated, when the numerical variable is selected from the accession database. With the *Distribution Statistics* function, statistics for each unique value (class) of a multi-state variable can be calculated, for example, for each species in a database of wild relatives in a given genepool.

POINT-BASED DATA ANALYSIS

Point-based approaches, such as the Spatial Intra-specific Diversity (SID) software which was described by Nelson (1997), was implemented in DIVA like an alternative to the use of grids. Diversity indices are calculated based on all observations lying within a user-defined circle about each point. The results are assigned to the location of the central observation and output to a database. The results can then be mapped again in DIVA.

FREELY AVAILABLE DIVA-GIS DATABASES

Country-level database is free to download from http://gis.cip.cgiar.org. These databases can be completed by genetic resources data that are being mapped and analyzed. Also shapefiles with data on administrative boundaries, country boundaries and first and second level administrative subdivisions for most countries are available via internet. There are grids for all countries, including data such as altitude, land cover and population density. And many more public domain databases are freely on internet.
DIVA is probably most useful for analysis of distribution data covering larger areas, such as can be typically obtained from genebanks.

1.1.2. FloraMap

FloraMap (Jones and Gladkov, 1999) is another GIS application. This one was developed by CIAT primeraly for the prediction of the distribution of organisms in the wild, when little or nothing is known of the physiology of the species involved, so that analytical or symulation models cannot be used. This is important to plant genetic resources (PGR) workers because data on the geographic distribution of species are often scant and it would clearly be useful to be able to use what little is known to identify areas where a species has not been previously recorded, but where is might still be expected to occur. FloraMap assumes that the climate at the points of observation and/or collection of species (herbarium specimens, germplasm accessions etc.) is representative of the environmental range of the organism. The climate at these points is used as a calibration set to compute a climate probability model, which is then used to assess the likelihood of other sites being climatically suitable for the species.

THE BASE CLIMATE DATA

In FloraMap analyses are used climate data from 10-minute grid (similar to 18km at the equator). The grids were educed by interpolation from thousands of meteorological stations. Thirty-six climate variables are used: 12 monthly averages for temperature, rainfall, and diurnal temperature range. Temperature is standardised with elevation using NOAA TGP-006 digital elevation model (DEM) and a lapse rate model. Rainfall and diurnal temperature range are independent of elevation.

 A 12-point Fourier transformation is applied for adaptation geographic differences in the timing of major season. It rotates the climate data to a common time frame to make them directly comparable.The climate of a location in the Northern Hemisphere may be very similar to that of a location in the Southern Hemisphere but with difference in the beginning of the rainy season. The phase angle of the first frequency for the rainfall and temperature records are combined as a function of latitude and the resulting angle is substracted from each climate record.

Climate grids for Latin America and the Caribbean, Africa, Asia, Europe and USA are available for using in FloraMap.

COMPUTING THE CLIMATE PROBABILITY MODEL

The 36 climate variables described above are first extracted for the pixel in which each accession in the input file is located, and principal components analysis (PCA) is carried out on the resulting data. PCA recognizes new variables or dimensions (principal components) related to the original variables. The PCA is performed on the variance-covariance matrix since the Fourier analysis has already transformed the variables to comparable scales. FloraMap then fits a multivariete normal distribution to the principal components scores of the accessions (the user specifies how many dimensions to include)

MAPPING AND EXPORTING THE PROBABILITY SURFACE

For each pixel in the selected continent is calculated probability of belonging to the multivariate normal disribution. The final result is a probability surface for the continent(s) under consideration. These probability surfaces are ESRI® Shapefiles, with each pixel defined as a

square. They are available for other GIS applications, for carrying out further analyses and producing individually crafted output maps.

FLORAMAP DESKTOP

Map Window is used for displaying map and for managing the map layers, including probability surface and accession points. Study areas can be specified by roads, rivers, cities and political boundaries of Latin America, Africa and Asia. All these are displayed as a background to the probability layer. The principal components window is used for checking the principal components analysis. Choice of components and map probability limits are under user's control, by using convenient sliders or by entering the value desired. A scatter diagram indicates the distribution of the accession points in any combination of the principal component dimension. The cluster analysis tool provides access to seven powerful techniques to investigate the possibility of clustering. After find of cluster, the user can point and select a cluster and calculate the probability model for those accessions only, repeating the procedure for each cluster. The climate diagram can be displayed in Cartesian or polar coordinates and can show the data either by real date or in rotated form.

DIFFERENT USES OF FLORAMAP

The primary use of FloraMap is predicting the distribution of plants and other organism in the wild from very limited distribution data. Comparing such predicted distribution with localities where accession data were collected can be useful to germplasm collectors trying to identify gaps in existing collections. For example Jones at al. (1997) had some success with wild *Phasoleus* in South America.

The software's analytical facilities with data, which are used in prediction of distribution, also allow the 'retro-classification' of accessions. This process was coined by Steiner and Greene (1996), and it is the process of obtaining data by overlaying the locations of the sites on different thematic base maps of the appropriate scale and reading off the map attributes. Steiner and Green (1996) used in retro-classsification the Lotus collection of the US National Plant Germplasm System (NPGS)

The most important applications are: c*hecking data quality, predicting climatic adaptation in other areas, identifying groups of accessions with distinct climatic adaptations (ecotypes), comparing climatic adaptation*

For identifying errors in data is used the PCA scatter diagram. Any accessions that are obviously outliers in the PCA diagram should be checked for any inaccuracies in the latitude and longitude information.

The model describing the climatic adaptation of the species is usually applied only to the putative region of its distribution in the wild. With FloraMap is possible to apply the model to other place in the earth. For example, the distribution of species that are grown as forages and are native to Africa can be used to predict where they might be expected to yield a good response on introduction into South America or Southeast Asia. Libreros et al. (2000) presented a study of adaptation of South American fruit species in Mexico and Jones et al. (2000) a study of cross-continental climatic adaptation in *Desmodium*.

The clustering tool allows the user to investigate the existence of groups of accessions with distinct climatic adaptation within a collection. When such group is defined, individual probability of distribution has to be fit to the accessions in each group in order for the model to work correctly. The PGR workers exploit the possibility of groups association with genetic differences of significance for genetic resources conservation or use. Climatic groupings could be used in developing core collections. For PGR workers is very useful to compare the climate adaptation of selected groups of accessions. FloraMap was used in conjunction with data on collection of wild *Arachis* species to define and compare the climatic adaptation of different species. Users interested in adaptation to particular climatic conditions would then know which of the species are most suitable for their purposes.

FLORAMAP LIMITS

Though climate is widely considered to be one of the primary determinants of plant distribution, it is not the only one. FloraMap takes no account of factors such as soil, dispersal mechanism, natural vegetation, human intervention, physical barriers etc. The probability surfaces can be imported into other GIS, where they can by overlying with other geo-referenced information to refine the analysis.

FloraMap does not provide a once-and-for-all solution. As more distribution data become available, the results of the analysis can change significantly, in particular if the original accessions were in some way a climatically biased sample of the real distribution of the species

2. Aim of the thesis

These days it is more and more common to use computers for the simplification of work, or to achieve precision results. The aim of this work serves to act as a review of the potential of using of geographic information system (GIS) in plant genetic resources research. One of the biggest advantages of these technologies is both the possibilities of cooperation with other scientists through already generated databases and the possibility of perform all parts of research single-handed. This work provides some useful information describing the whole process of using the GIS whereby it is possible to choose the most suitable technique for everyone who is interested in plant genetic resources research.

3. Methodology

This work is based on a review of current scientific literature describing problems of using GIS technologies in plant genetic resources (PGR) research. Electronic databases of scientific literature, such as ISI Web of knowledge, Springer link and Science were used to create a collection of 26 scientific articles. These articles were used as resources of information for creating summary of advantages, disadvantages and possibilities of using GIS in PGR research. The first part is focused on introduction to basic principles of using GIS and the most used computer tools, such as DIVA-GIS and FloraMap. The second part describing using GIS in concretes cases of study.

In individual searching were used these keywords: GIS, DIVA-GIS, FloraMap, plant genetic resources, germplasm, in situ; ex situ conservation, biodiversity.

4. Results and discussion

4.1 Use of GIS in plant genetic resources conservation

In the process of using GIS in plant genetic resources conservation are most important these componenets: *Ecogeographic surveying, Field exploration, Design, management and monitoring of 'in situ' reserves, Germplasm evaluation, Use of genetic resources.*

4.1.1 Ecogeographic surveying

Herbarium specimens, germplasm accession passport data, field notes, maps and relevant literature, these are the sources of information needed for ecogeographic surveying. This information describe ethnobotany, ecological adaptation, geographic distribution and genetic diversity of a plant group. Along with information describing climate, geography, ecology and the human settings of study regions is possible to plan date, place and kind of collecting germplasm and choose the best place to establish genetic reserves and how to monitor and manage them. On the earth are places with different biodiversity so it is logical that some places are of higher priority for collection or conservation than others. Ecogeographic surveying is focused on identification these geographics areas: areas likely to contain germplasm of interest, areas with high environmentally, taxonomically or genetically diversity, areas which are complementary to each other, areas with danger of genetic erosion and areas which are still under-represented or even missing in conservation efforts.

AREAS WITH ASSUMED OCCURRENCE OF GERMPLASM OF INTEREST

GIS technologies might be very useful in searching of areas where a species has not been previously recorded but where they probably occur. First, they explore the areas where a species has been found, then estimate the conditions for occurrence. By using software (e.g. FloraMap) are extrapolated and displayed areas with the different level of similarity with areas of collection. Another possibility of application of GIS is objective prediction of the full distribution of species from incomplete point distribution maps, based on its ecological preferences. Most often these methods have been used at the species level, but it is possible to utilize them in the level of

botanical varieties, groups of similar accession (multivariate analysis based on e.g. morphological characterization) or farmer-recognized landraces. A species being chosen for conservation may be related with specific habitat, vegetation, land use and landform. In locating these features may be used combination of GIS technologies such as satellite and radar imagery. Is also possible to spot the isolated areas of cultivation, but only when these areas are larger than the resolution of the remote sensing system. Getting and analyzing satellite imagery could be costly and time-consuming work however some useful software is available online on the Internet. One of them was developed for FAO Global Information and Early Warning System, it is called WINDISP and is using for display and analysis of satellite images, maps and associated databases. Also another data derived from satellite sources such as some ecosystems and land use data are available to download from the Internet. Sometimes materials with very specific characteristics need to be conserved. GIS can derive areas where material with required adaptation might be expected to occur, this process is based on agricultural censuses, specialized surveys and crop experts. Geostatistical methods of interpolation can describe the spatial pattern of variation in genetic, morphological and agronomic traits among population. Analysis of genetic boundaries can show the areas with prevalent specific traits and identify areas that are relatively homogenous but different from each other for the characters being study.

A. Jarvis et al. (2005) describe the using of GIS for predicting occurrence of germplasm of interest in study about natural occurrence of *Capsicum flexuosum* Sendtn. in Paraguay. They selected preferred areas for germplasm collecting by model based on combination of a prediction of potential distribution on the basis of climatic adaptation determined from previous collecting sites; a map of preferred habitat of the species which is a forest margins; and a map of areas accessible by road.

The first step was collating the database of herbarium specimens and germplasm accession from the US National Plant Germplasm System, and also the determination of geographic coordinates for each observation. For predicting the distribution of *C. flexuosum* in Paraguay was used FloraMap together with climatic data from observations from 130 meteorological stations in Paraguay, for example monthly averages for temperatures, rainfall and diurnal temperature range. The result of this analysis is a map of potential climatic envelope where an organism could exist, but it does not account for factors such as edaphic preferences, dispersal mechanism, anthropogenic impacts or ecological habitat preferences. *(Figure 1.)*
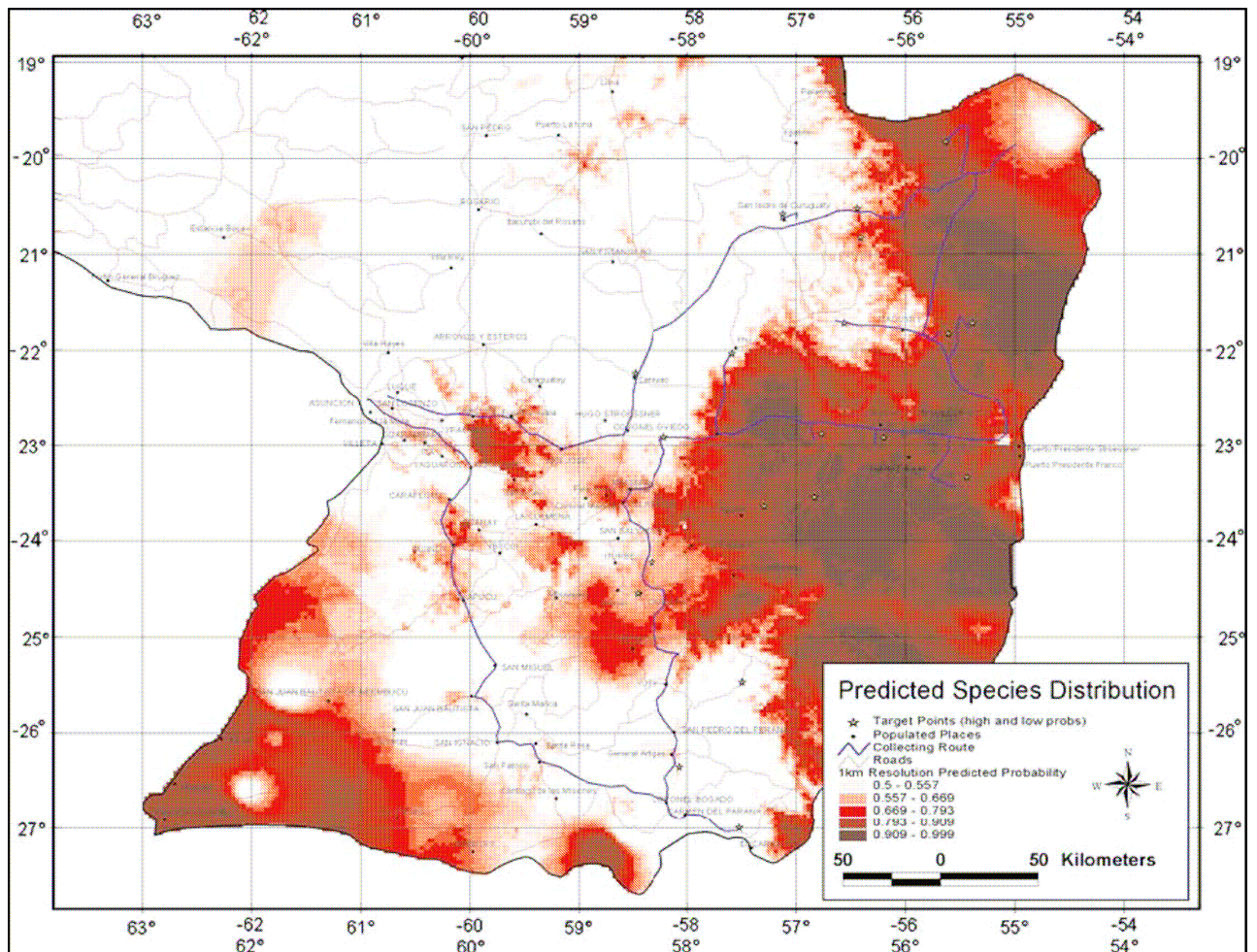
Figure 1. Example of projection of species distribution.

(Capsicum flexuosum in Paraguay, Jarvis, 2005)

The second step was identifying areas of potentially suitable habitat. For location of forests in Paraguay was used International Geosphere-Biosphere Program (IGBP) Land Cover Classification together with using of data set from United Sates Geological Survey (USGS) Global Land Cover Characteristics project, which is freely available on the internet. Five forest land-covers such as deciduous broadleaf forest, deciduous needleleaf forest, evergreen broadleaf forest, evergreen needleleaf forest and mixed forest were used to create a single coverage of forests in Paraguay.

(Figure 2.)

The third step was to locate roads and a 4-km buffer zone region by using the Digital Chart of the World. This step selected areas where collection is deemed feasible.
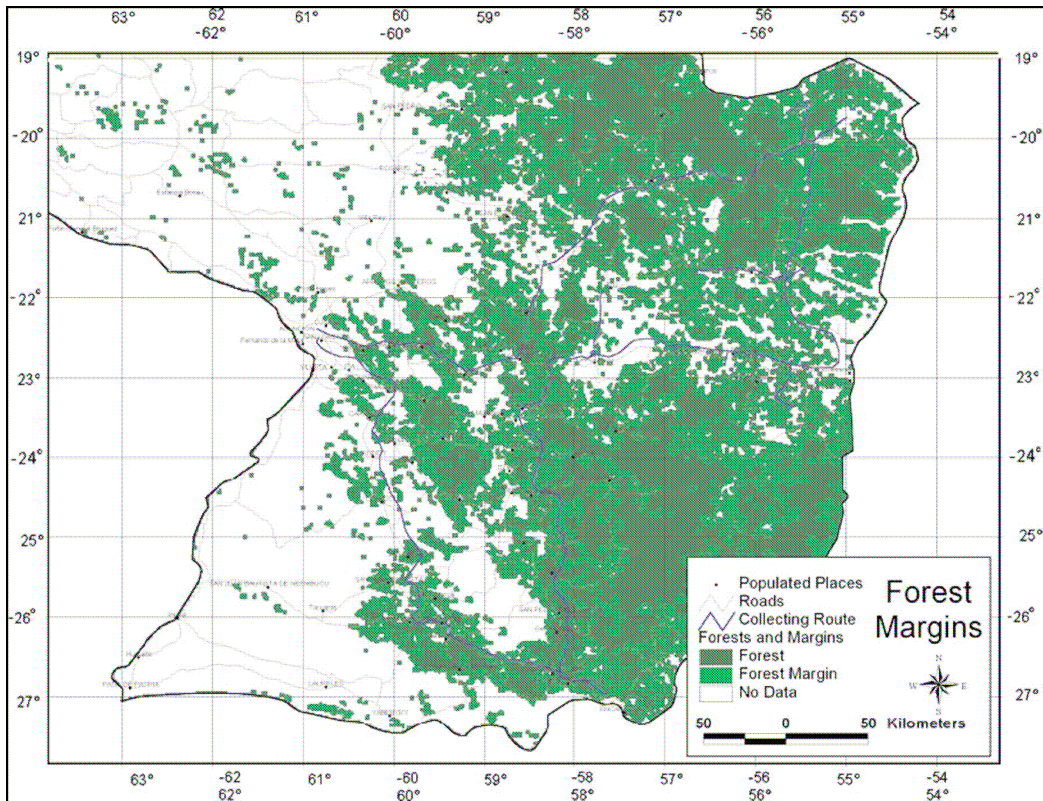
(Figure 3.)

Figure 2. Example of projection of forests and forest margins distribution.
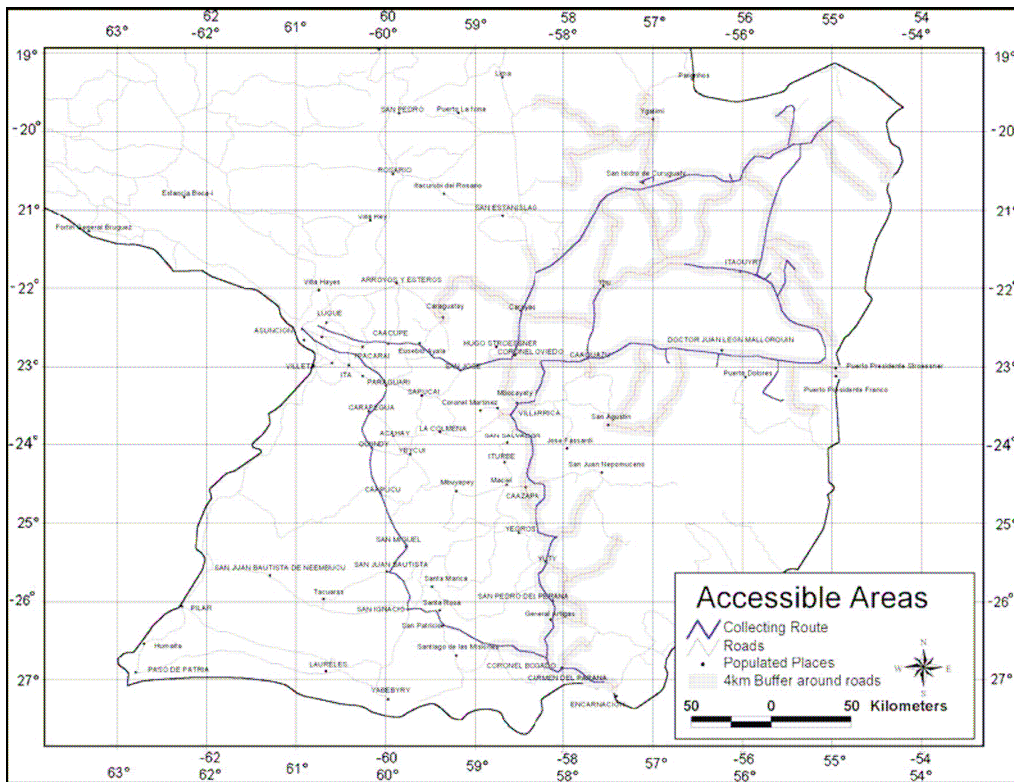
(southern Paraguay, Jarvis, 2005)

Figure 3. Example of projection of areas defined as being accessible to plant collectors.

(*Capsicum flexuosum* in southern Paraguay, Jarvis, 2005)

The result of the modelling refers an area of 6460 square km in south-eastern Paraguay where *Capsicum flexuosum* Sendtn. is predicted to occur. *(Figure 4.)*

In time of 15-22 March 2002, 20 sites were visited and explored for *Capsicum flexuosum,* and five new natural population of *C. flexuosum* was found. Germplasm samples from these natural populations have been deposited in US National Plant Germplasm System (NPGS) for long-term *ex situ* conservation on behalf of the Paraguayan national authorities. Of these one accession (WWQC 146) has been increased and assigned a Plant Introduction number (PI 631154) in NPGS Germplasm Resources Information Network (GRIN). The *in situ* conservation status of *C. flexuosum* in Paraguay is threatened by habitat destruction.
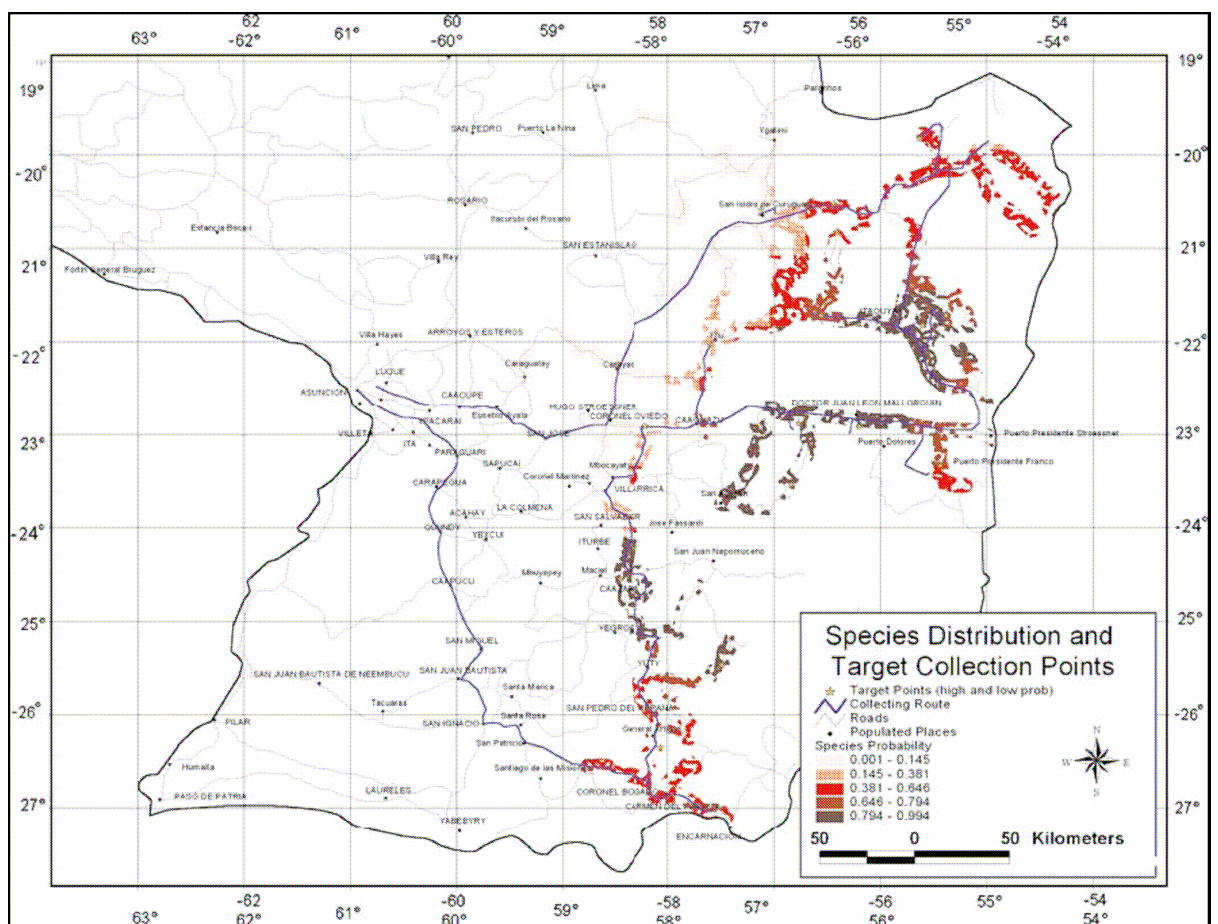

Figure 4. Example of projection of priority areas for collecting.

(*Capsicum fexuosum* in southern Paraguay, Jarvis, 2005)

AREAS WITH HIGH DIVERSITY

Some sites on the earth are more interesting in taxonomic or genetic diversity than others. Diversity is one of most important factor in genetic resources work. Study of diversity starts by separating the target area into a number of smaller zones, for each of which a measurement of diversity can be calculated. Ideally areas of equal shape and size, such as grid cells are the best. Sometimes is used presence or absence of species to investigate patterns in the taxonomic diversity in different grid cells. For calculation of morphological diversity of cultivated species of interest are using diversity measurements based on morphological characters or molecular markers. In this process can be very useful DIVA-GIS software, which calculates diversity indices for all the cells in user-defined grid given latitude, longitude and characterization data for a set of accessions and maps the results.

Another method for mapping diversity is 'point-based' method approach used by Spatial Intraspecific Diversity (SID). SID is able to read coordinates of each accession point after that is drawn a circle of user-defined radius around it. Accession points lying within this circle are used to calculate diversity within the defined radius, using the Shannon-Weaver or Simpson index. The result is allocated to the spatial location of observation. The advantage of this method is easy interpolation and modeling of the point-based values diversity.

Various studies used different environmental parameters as a substitute for species diversity. One example of GIS application in these problems is calculating with topographical data for predicting of the richness of the rare species. This is being applied at the intraspecific level in a study of relationship between environmental and human diversity and genetic diversity in cultivated groundnut in Ecuador and Guatemala. Even IPGRI and CIAT were participated on this study.

Use GIS in these problems was described by X. Scheldeman et al. (2007) in study of Distribution, diversity and environmental adaptation of highland papayas (*Vasconcellea* spp.) in tropical and subtropical America. The genus *Vasconcellea* is the largest within the family Caricaceae, holding 21 species, often collectively called highland papayas because many of them occur at higher altitudes. Five of the 21 species
(V. horovitziana, V. omnilingua, V. palandensis, V. pulchra, V. sprucei) are included in the IUCN Red List of Threatened Species and require special monitoring attention for their conservation. The process of measurement of diversity and distribution of *Vasconcellea* species was composed from several steps. Firstly, the number of observations was tabulated per species and per country. Then was selected an indicator of abundance or rarity of a particular species

which was the area of occupancy, defined by International Union for Conservation of Nature (IUCN). The area of occupancy was calculated by superposing a

2 x 2 km grid over the study area, followed by determining the number of grid cells occupied by each species (by DIVA-GIS) and by converting these to an effective areas by multiplying the number of occupied cells by 4 square km. Third step was to estimate the extant of the distribution area for each species by determining the average distance between all possible pairs of collection points of each species. The last step was to map observed species richness by using the point-to-grid richness analyses tool in DIVA-GIS, using a 1 x 1° grid (111 x 111 km at the equator). The circular neighborhood option, with a 2° diameter was applied to eliminate border effects due to the assignation of the grid origin.

The result of this study was a more detailed view of the geographic distribution of each *Vasconcellea* species. Areas with the highest richness for *Vasconcellea* are in Ecuador, there were found 16 of the 21 species, in the second place with 9 species are Colombia and Peru. In all other countries, five or less species have been collected. The Andes of northwestern South America represent the center of diversity of *Vasconcellea. (Figure 5.)*
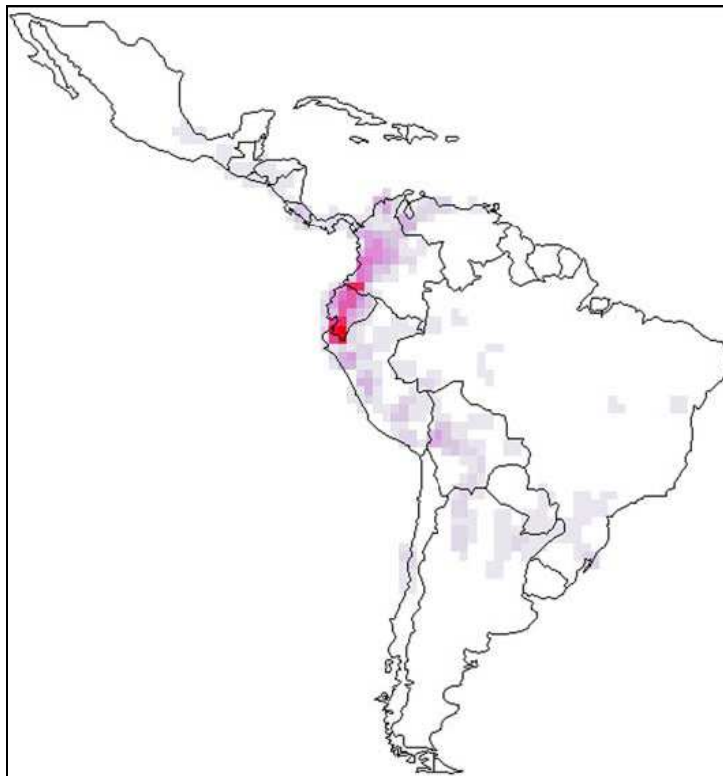

Figure 5. Example of projection of species richness. The darker areas indicate the highest
potential richness (*Vasconcellea* species in Latin America, maximum value: 11 species,
Scheldeman, 2007)

COMPLEMENTARY AREAS

Two areas may have equal richness or diversity of taxa or morphotypes, but the ones in one square may be similar to each other, while in other may be more different. Other things being equal, the second area would be the higher conservation priority. One way of optimal targeting of conservation effort is using multivariate statistics to classify or ordinate spatial units according to the species found there, on the basis of characterization or evaluation data, or in terms of environmental conditions. Another approach is method of iterative procedure based on choosing the smallest number of spatial units such that each species is placed in at least one unit. This method with complementary DIVA-GIS analysis can also be very helpful in problems of selecting potential biodiversity conservation areas.

UNDER-CONSERVED AREAS

GIS with its ability of superimposition of digital maps is used to identify individual species, species-rich areas, or vegetation types that are under-represented or even not represented in existing databases of protected areas or areas where germplasm collection has already been adequately carried out. Global datasets of protected areas are available from the World Conservation Monitoring Centre (WCMC). Identifying ecogeographical gaps in existing ex situ germplasm collections is more and more important. For example the coarse-grid collecting of the past 25 years is replaced with a more targeted, fine-grind approach

Checking and assigning of genebanks coordinates was discussed by Hijmans (1999). In this study was used data set from a case study on wild potatoes from Bolivia. This database compose of records from 18 expeditions from a period of more than 40years so it is not surprise that was found that more than 50% of the accessions had error. The process of checking coordinates has 3 parts; finding errors, determining the cause of error and assigning coordinates. By GIS can be detected three kinds of errors such as accessions in impossible places, e.g. oceans; accessions in unlikely places, e.g. unlikely altitude; accessions in the wrong place according to passport data. The first two kinds of errors are caused by wrong coordinates and can be detected by visual inspection of data. This method is using data on wild potato species of Bolivia from the Intergenebank Potato Database. Three suspicious isolated locations are branded with question mark. After clicking on the point, the Identify results window is shown. This window include information about the point such as locality description which indicates that the point should be

located in Las Delicias, in the Department of Potosí, although on the map the point is located in the Department of Santa Cruz. (Figure 6.) The coordinates in the database, and thus the location on the map, are therefore likely to be wrong. The third type of error is typically due to wrong co-ordinates and/or wrong names and can be detected using *overlay analysis* methods.
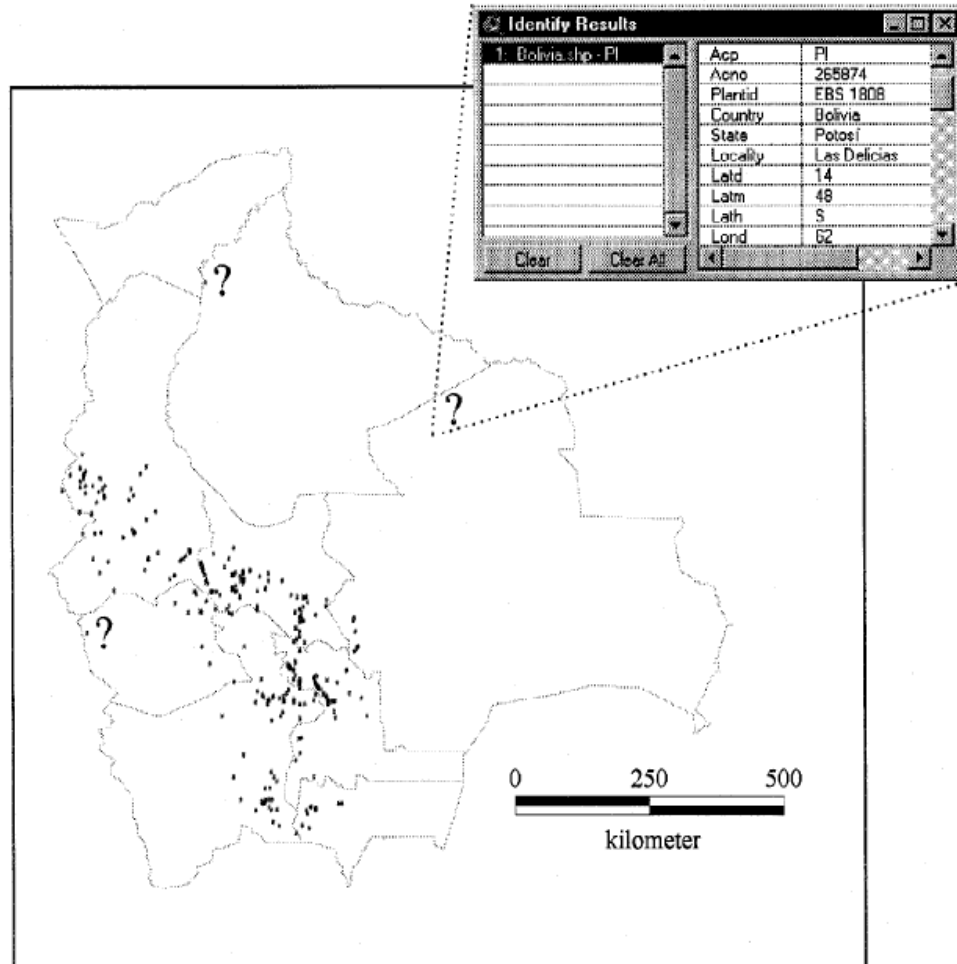

Figure 6. Example of projection of sites, where germplasm was collected (black dots).
(wild potatoes in Bolivia, Hijmans, 1999)

It is more difficult to determine causes of errors than to spot them. The causes of errors include incorrect reading of maps, sometimes caused by duplicate locations names and confusion about the coordinate system. The most common cause of errors in database are perhaps topographical mistakes. If a name in the genebank database does not correspond with the name in the administrative boundaries database, this could be due to wrong coordinates or to wrong names in the genebank database. Another error can be made in assigning and digitizing coordinates, such as switching of latitude and longitude, typing the wrong hemisphere or typing two digits in a number in the wrong order. If an error is due to a wrong name in one of the

databases, this is often caused by difference in spelling or by typographical mistakes. These errors are easy to trace and correct by inspecting alphabetical lists of the names in databases.

The different kind of errors and the difficulties in detecting and correcting them, highlight the importance of more precise bookkeeping by germplasm collectors and curators. One of the main reasons is that checking and improving the coordinates of germplasm database is tedious and time consuming.

.

ENDANGERED AREAS

Many areas in the world are threatened of genetic erosion, so it is logical that these areas are high priorities for conservation. Remote sensing can be used to document changes in the extent and characteristics of forests, wetlands, deserts, different land use types, cropping, etc. Remote sensing imagery together with agricultural censuses, news reports and fieldwork can be used as sources of information for predicting genetic erosion based on specific indicators that are associated with genetic erosion, for example desertification, soil erosion, extension of agriculture into increasingly marginal areas, etc. These indicators can be assessed by observables, for example decreasing development of vegetation (measured by NDVI), increasing outmigration and the appearance of new field in marginal areas, all of which can be managed by GIS.

These problems were described by L. Willemen et al. (2007), between 2001-2002, 285 farmers in 50 communities in the Ucayali region were interviewed to carry out a field survey. The survey included questions on the socio-economic situation of the farmers, their farming practice, cultural background, and current agriculture trends. Then were farmers asked about the number of distinct cassava varieties grown at the time of survey and about the observation of losses of varieties in the recent past. All survey results were summarized per community. Coordinates were surveyed using GPS, other spatial data were extracted from different digital sources, for example thematic maps of the region with the main roads, rivers, urban areas, and geomorphology layers was provided by The Instituto de Investigaciones de la Amazonía Peruana (IIAP), and the Centro Internacional de Agricultura Tropical (CIAP) provided deforestation map. In total, 295 cassava samples, the representative for the morphological diversity at community level, were collected and then used for *ex situ* conservation in Pucallpa and deposited in Peruvian national genebank.

The genetic erosion could not been measured directly on the field because of a lack of baseline and time series data. The presence or absence of genetic erosion was assessed by asking farmers whether they had noticed any disappearance of varieties in the last 5-10 years, and then

were defined a genetic erosion index for each community. DIVA-GIS was used for spatial data processing. Results showed correlation between genetic erosion and possession of livestock. By new settlers has been land in Ucayali region turned into pasture. Communities where a high number of farmers buy a planting material on the market also reported loss of varieties. Position of community is another factor involving the genetic erosion index is. Communities closer to the main road to Lima have clearly noticed more variety loss than settlements further away. Road to Lima is an important link to markets, so farmers in these areas are pressured to adopt commercial agricultural practices.

4.1.2 Fieldwork

Fieldwork is practical part of ecogeographic study. GIS technology can be useful in this part of genetic resources conservation through the development of field aids and providing information on the planning optimal time of field visits.

For filedwork are most important maps showing factors (specific or combination of them) of interest. Unluckily sometimes these maps are not available or the available maps are unwieldy to use. GIS can be used to create maps showing for example moisture, temperatures or ecogeographic gradients in the areas of interest.

Another problem could be timing of field visit. Because of difficult timing of field exploration simultaneously with appropriate stage of vegetation is often used satellite imagery. The advantage of satellite imagery is in providing data on the state of vegetation with relatively short time thereby collectors could be more precise in timing field visit. Data can be analyzed to provide surveillance of the state of crops and vegetation over the growing season. With satellite imagery is also possible to evaluate the stability of phenology in areas across the year. Areas with most predictable phenology were regarded the highest priority for field visit, because of the greater likelihood of finding seed of species of interest for collection.

4.1.3. Design, management and monitoring of *in situ* reserve

In designing of protected areas is applied the spatial analysis in GIS to questions of optimal size and shape or zonation. This type of application present the challenge of integrating data on the human population, such as demography, socio-economy or culture, with data on the biophysical environment and on other target taxon For developing a zoning plan for a reserve, buffer zones or core can be used information on species richness, distribution and abundance of

endangered species, disturbance and distribution of timber resources within a forest etc. The beginning is the definition of criteria being connected to ecological and socio-economic sustainability, this provides that the limits of the park will respect the natural features and that the size of protected areas will be enough for living requirements of the human population surrounding the buffer area. By using satellite imagery, biodiversity surveys, population studies and combinations of map are created GIS analysis to develop a buffer area balancing human and wildlife requirement.

Because of rapid rate of environmental and socio-economic changes, such as environmental pollution or global climate changes etc., is important to calculate with possible future scenarios during the reserve designing. GIS technologies allow to select and to manage reserve on the foundations of both current climatic and vegetations templates and different supposition of future changes.

At the moment that *in situ* reserve has been established is time to implement some form of management and to monitor its results on frequency, plenitude, demography and genetic diversity of species of interest. GIS may be used in defining and implementing the management plan.

A. Jarvins et al. (2002) described a geographical approach to identifying biodiversity hotspots and prioritizing regions for conservation, using peanuts wild relatives as a case study. The closest relatives of the peanut include 26 described wild species of the genus *Arachis*, section *Arachis*, all endemic to south-central South America (Bolivia, Argentina, Brazil, Paraguay an Uruguay). In this study were used 397 geo-referenced germplasm accession and herbarium specimens. The germplasm accession data were obtained from SINGER and GRIN. For predicting the distribution of *Arachis* by climatic models was used FloraMap. The lack of accessions (less than five) in eight species meant that it was not possible to carry out a sensible principal components analysis (PCA), so these species were omitted from the analysis. Some of these species have been identified as possible progenitors of the cultigen. So it is underlining the need for further collecting and conservation efforts.

The result is the map of two main high diversity areas, one in the southern part of the Department of Santa Cruz and the other in north-western Paraguay. This semiarid region contains natural savanna where populations of wild *Arachis* are most often found. Germplasm collection locations were overlaid on this map and show that no geo-referenced accessions exist from the region of interest in Paraguay and in southern part of the Department of Santa Cruz. These places were also considered as a possible place of origin for the cultigen, and thus the best place for *in situ* conservation.

4.1.4. Germplasm evaluation

The ability to find best areas for each accession is a key step in both germplasm multiplication and evaluation. The regeneration area must afford suitable conditions for any requisite triggers for the different stages of plant development, such as occurrence of pollinators providing flowering and seed production or lack of pest and diseases. GIS is able to identify potential areas for regeneration by overlapping climate and distribution maps for pests, diseases, pollinators and wild relatives. The environment of the areas of regeneration is compared with the environment of the areas of collection. There is very often used FloraMap. Areas of collection are carefully investigated for determination of the unique requirements of individual accession. This information is also used in germplasm evaluation by definition both ecological optimal and stress locations. The other application of GIS is in targeting and predicting the potential impacts of promising germplasm.

4.1.5. Genetic resources

The objective of conserving germplasm is for it to be used. The lack of data on accession and the large size of collection can significant difficulty using of conserved materials. GIS can identify interesting germplasm and show diversity structure within the collections. GIS may also help with improving the quality and quantity of the locality data associated with collection and completing passport data. By using of published or even online maps and collectors' notebooks can be determinate the latitude and longitude only by locality name. When localities are marked on the paper maps, digitizing of maps is used for estimating latitude and longitude.

Also available passport data can be inaccurate. Rapid data visualization by GIS can show obvious errors in latitude and longitude data resulting like bad position of accession.
Another data checking is doing by comparing locality descriptions in passport data with GIS datases of administrative boundaries. Retroclassification of the collection areas is focusing the attention on the most promising material for their specific purposes. Combination of germplasm location and soil maps is possible to identify candidates from plant collection what can be adapted on e.g. alkaline soils. Because of expensive or time-cosuming measurement of some traits is often done only a 'core' collection, which represented the genetic spectrum in the whole collection. The materials describing the core collection can be applied in the identification of the rest of the collection. Agroclimatical conditions at the collection area are one of the most often

using criteria in process of classifications. For example regional history of crop, length of growing season, photoperiod, soil type and moisture regime are used together with passport data to characterizations growth habit. GIS is also helpful in predicting the likely impact of the use of germplasm.

R. J. Hijmans et al. (2003) described the extent to which taxonomic, geographic and ecological factors can predict the presence of frost tolerance in wild potatoes. They used data collected in 12 countries in the Americas, covering most of the distribution area of wild potatoes. The coordinates of accessions were checked and then 1646-attested accessions from 87 species were used. Mean frost tolerance was calculated over species and series. By using DIVA-GIS was created a map of the mean and maximum observed frost tolerance. The mean minimum and maximum temperature during the year, and during the estimated growing season were calculated for each accession. Results show a significant association between species and frost tolerance, 5 species have a frost tolerance score of more 90% and 9 species have a score of 82% or higher. Highest frost tolerance was for *Solanum acaule ,* which is also one of the most common and widespread wild potato species; *S. albicans,* and for *S. commersonii.* Also significant association between species was proved, the highest frost tolerance have series *Acaulia, Commersoniana,* and *Megistacroloba.* Areas with high mean level of frost tolerance observations constitute area from central to south Peru, and a small part of adjacent northern Bolivia; and area stretching south and east from Paraguay into adjacent Argentina, Brazil and Uruguay, and small area in central Chile. (Figure 7). Species and series effects on frost tolerance are much stronger than effect of temperature. The map of areas with mean minimum temperature below 3°C within the area where wild potatoes occur is not a very good predictor of the level of frost tolerance in wild potatoes from these areas. (Figure 8) The map correctly identifies the Central Andes and parts of Chile, but misses the lowland area of Argentina and Uruguay, with *S. commersonii;* and wrongly predicts the southwestern USA as an area with frost tolerant wild potatoes. All these information could be used in searching the most fitted species for each area, and with it can save time and money in using of genetic sources.
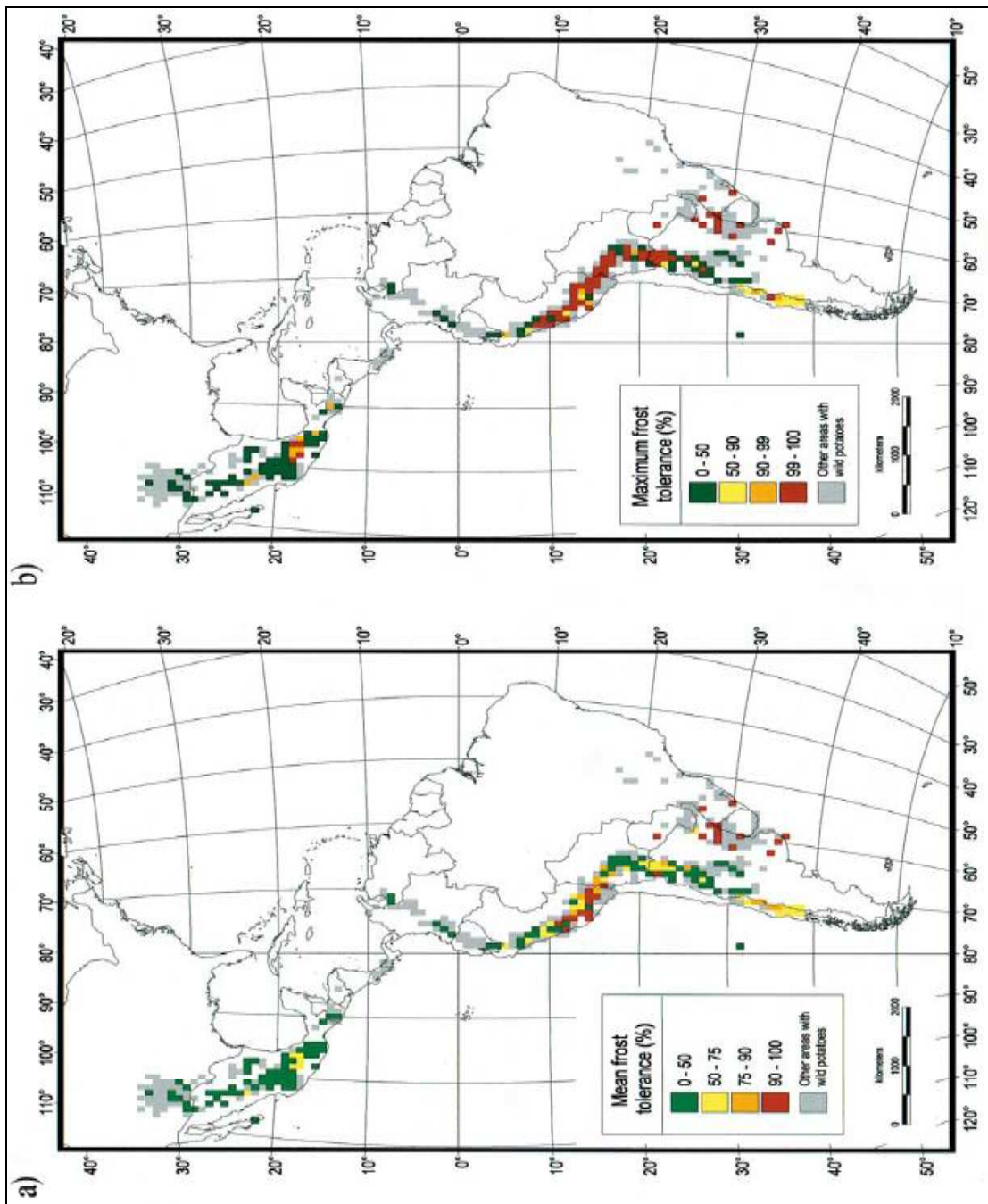
Figure 7.: Example of projection of the mean (A) and maximum (B) observed frost tolerance, percentage non-damaged tissue (wild potatoes, Hijmans, 2003)
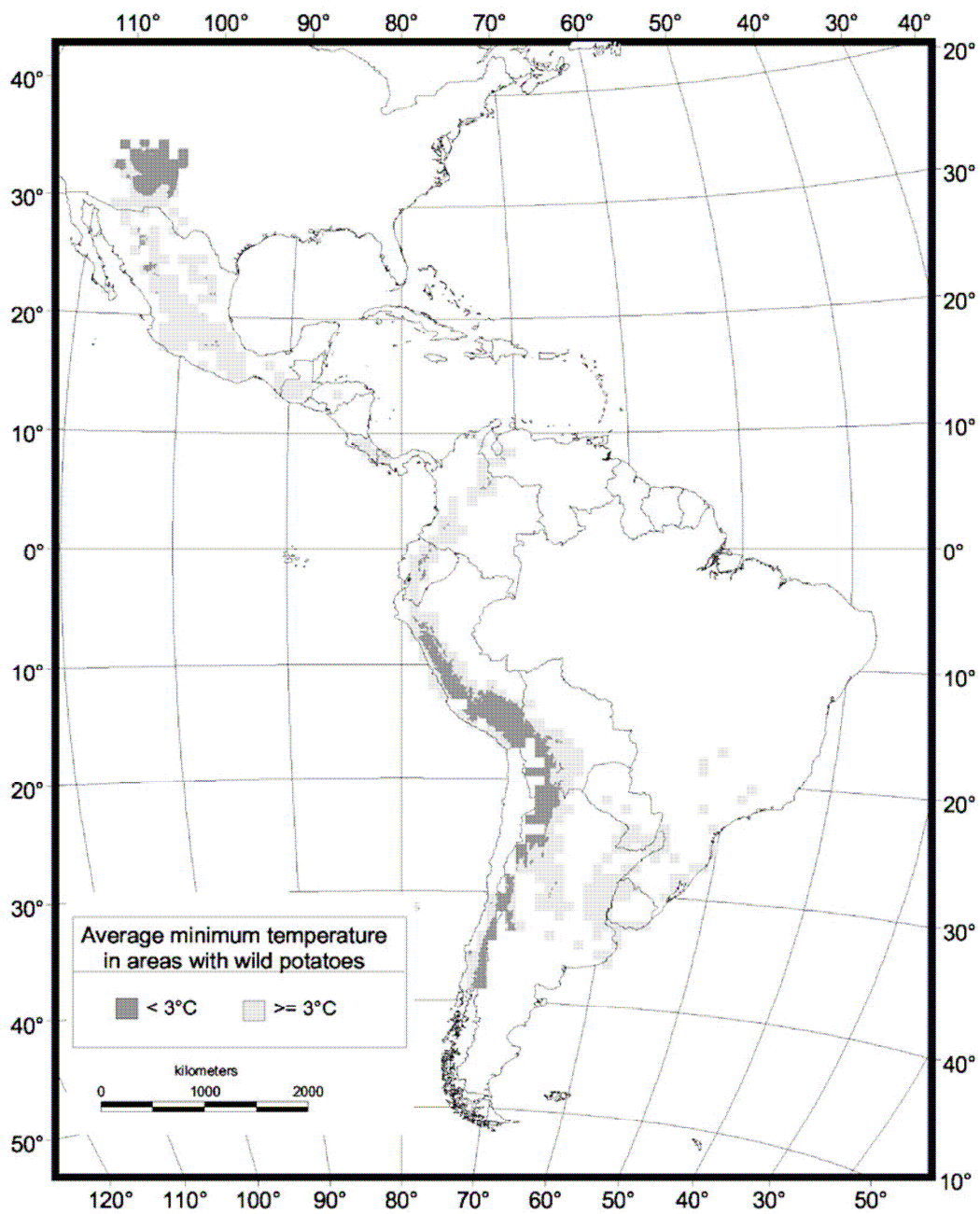
Figure 8: Example of projection of minimum temperature.
(areas where wild potatoes have been observed, Hijmans, 2003)

5. Conclusions

The use of GIS opens the way for new approaches in plant genetic resources management. GIS should significantly enhance the cost-effectiveness of conservation efforts, and facilitate use of germplasm by breeders and others. However, GIS technology has not been used by plant genetic resources conservation program, in the extent of its potential. The reason is quite high price of GIS hardware, software and data, and also its difficult obtainability and complex using. Another complication could be weighty resources limitation, especially in developing countries.

It is true that the current revolution in GIS technology is putting both the data and analytical tools within the reach of many, but it does not mean that the plant genetic resources conservations community will simply wait for the technology to come to its aid. On the contrary the community needs to take positive action itself at the international, regional and national levels. Another important step that could help wider adoption of GIS by plant genetic resources programs are web-based tools, where maps and accession points can be manipulated on-line with easy-to-use, browser-based and graphical interfaces. But the key role of process of GIS adoption plays access to data. GIS afford an ideal environment within which is possible to manage and analyze data from number of different sources. Genebanks should create linkage with institutions, which already have GIS capacity, such as national statistical and natural resources institute, mapping and survey departments and meteorological services. With available data and tools for their managing, is possible for everybody to use GIS in many cases of plant genetic resources study.

6. References:

DIVA-GIS tutorial, Available online at http://www.diva-gis.org/docs/DIVA-5_Tutorial.pdf

Draper, D.; Rosselló-Graell, A.; Garcia, C.; Gomes, C. T.; Sérgio, C.; 2003. Application of GIS in plant conservation programs in Portugal. *Biological Conservation*, 113: 337-349

Eade, J. D. O.; Moran, D.; 1996. Spetial economic valuation: Benefits trensfer using geographical information systems. *Journal of Environmental Management*, 48: 97-110

Ferguson, M.E.; Jarvis, A.; Stalker, H. T.; Williams, D. E.; Guarino, L.; Valls, J. F. M.; Pittman, R. N.; Simpson, Ch. E.; Bramel, P. J.; 2005. Biogeography of wild *Arachis (Leguminosae)*: distribution and environmental characterization. *Biodiversity and Conservation,* 14: 1777- 1798

Ganeshaiah, K. N.; Barve, N.; Nath, N.; Chandrashekara , K.; Swamy, M.; Shaanker, R. U.; 2003. Predicting the potential geographical distribution of the sugarcane woolly aphid using GARP and DIVA-GIS. *Current Science*, 85(11): 1526-1528

Greene, S. L.; Minoura, T.; Steiner, J. J.; Pentacost, G.; 2007. WebGMRS: Prototype software for web-based mapping of biological collections. *Biodiversity Conservation*, 16: 2611-2625

Guarino, L,; Jarvis, A.; Hijmans, R. J.; Maxted, N.; 2002. Geographic Information Systems (GIS) and the conservation and use of plant genetic resources. *Managing Plant Genetic Diversity, Rome: International Plant Genetic Resources Institute (IPGRI), pp.* 387-404

Guinard, O.; Arnaud, E.; Sharrock, S.; Preliminary analysis of *Musa* germplasm information system data for southeast Asia using the geographical information system software DIVA-GIS. *Advancing Banana and Plantain R&D in Asia and the Pacific*.11:.201-208

Guralnick, R.; Neufeld, D.; 2005. Challenges building online GIS services to support global biodiversity mapping and analysis: Lessons from the mountain and plains database and informatics project. *Biodiversity Informatics*, 2: 56-69

Hijmans, R. J.; Guarino, L.; Bussink, C.; Mathur, P.; Cruz, M.; Barrentes, I.; Rojas, E.; 2004. DIVA-GIS Version 4, A geographic Information system for the analysis of biodiversity data, Manual, Available online at http://www.cipotato.org/diva/docs/DIVA-GIS2b.pdf

Hijmans, R. J.; Guarino, L.; Cruz, M.; Rojas, E.; 2001. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*, 127: 15 - 19

Hijmans, R. J.; Jacobs, M.; Bamberg, J. B.; Spooner, D. M.; 2003. Frost tolerance in wild potato species: Assessing the predictivity of taxonomic, geographic, and ecological factors. *Euphytica* 130: 47-59

Hijmans, R. J.; Schreuder, M.; De la Cruz, J.; Guarino, L.; 1999. Using GIS to check co-ordinates of genebank accessions. *Genetic Resources and Crop Evolution*, 46: 291-296

Hijmans, R. J.; Spooner, D. M.; 2001. Geographic distribution of wild potato species, *American Journal of Botany,* 88(11): 2101-2112

Iverson, L. R.; Prasad A.; 1998. Estimating Regional Plant Biodiversity with GIS modelling. *Diversity and Distributions*, 2: 49-61

Jarvis, A.; Ferguson, M. E.; Williams, D. E.; Guarino, L.; Jones, P. G.; Stalker, H. T.; Valls, J. F. M.; Pittman, R. N.; Simpson, CH. E.; Bramel, P.; 2003. Biogeography of wild *Arachis:* Assessing conservation status and setting future priorities. *Crop Science,* 43(3): 1100-1108

Jarvis, A.; Guarino, L.; Williams, D.; Williams, K.; Vargas, I.; Hyman, G.; 2002. Spatial analysis of wild peanut distributions and the implications for plant genetic resources conservation. *Plant Genetic Resources Newsletter*, 131: 28-34

Jarvis, A.; Williams, K.; Williams, D.; Guarino, L.; Caballero, P. J.; Mottram, G.; 2005. Use of GIS for optimazing a collecting mission for a rare wild pepper (*Capsicum flexuosum* Sendtn.) in Paraguay. *Genetic Resources and Crop Evolution*, 52: 671-682

Jones, P. G.; Beebe, S. E.; Tohme, J.; Galwey, N. W.; 1997. The use of geographical information systems in biodiversity exploration and conservation. *Biodiversity and Conservation,* 6: 947-958

Jones, P. G.; Guarino, L.; Jarvis, A.; 2002. Computer tools for spatial analysis of plant genetic resources data: 2. FloraMap. *Plant Genetic Resources Newsletter*, 130: 1-6

Kuncoro, S. A.; van Noordwijk, M.; Martini, E.; Saipothong, P.; Areskoug, V.; Putra, A. E.; O'Connor, T.; 2006. Rapid Agrobiodiversity Appraisal (RABA) in the context of environmental service rewards: Protocols for data collection and case studies in rubber agroforests in Bungo district, Jambi. *Indonesia and Fragmented Forest in North Thailand*, pp. 51-58, ISBN: 979-3198-31-1

Miller, A. J.; Knouft, J. H.; 2006. GIS-based characterization of the geographic distributions of wild and cultivated populations of the Mesoamerican fruit tree *Spondias purpurea* (Anacardiaceae). *American Journal of Botany*, 93: 1757-1767

Quan, J.; Oudwater, N.; Pender, J.; Martin, A.; 2001. GIS and participatory approaches in natural resources research, Natural Resources Institut, the University of Greenwhich, England, ISBN: 0 85954 496 - 6

Scheldman, X.; Willemen, L.; Coppens d'Eeckenbrugge, G.; Romeijn-Peeters, E.; Restrepo, M. T.; Motoche, J. R.; Jimenéz, D.; Lobo, M.; Medina, C. I.; Reyes, C.; Rodríguez, D.; Ocampo, J. A.; Van Damme, P; Goetgebeur, P.; 2007. Distribution, diversity and environmental adaptation of highland papayas (*Vasconcellea* spp.) in tropical and subtropical America. *Biodiversity and Conservation*, 16: 1867-1884

Toivonen, T.; Kalliola, R.; Driesprong, A.; 2002. Western Amazonian GIS: Creation process of landsat TM image mosaic and internet map services, Department of geography, University of Turku, Finland, ISBN: 951-29-2300-9

Willemen, L.; Scheldeman, X.; Cabellos, V. S.; Salazar, S. R.; Gustino, L.; 2007. Spatial patterns of diversity and genetic erosion of traditional cassava (*Manihot esculenta* Crantz) in the Peruvian Amazon: An evaluation of socio-economic and environmental indicators. *Genetic Resources Crop Evolution,* 54:1599-1612