36 Geographic Information Systems (GIS) and the Conservation and Use of Plant Genetic Resources

L. Guarino,[1] A. Jarvis,[1] R.J. Hijmans[2] and N. Maxted[3]
[1]*International Plant Genetic Resources Institute, Regional Office for the Americas, IPGRI c/o CIAT, Cali, Colombia;* [2]*International Potato Center (CIP), Lima, Peru;* [3]*School of Biological Sciences, University of Birmingham, Birmingham, UK*

Introduction

The management of genetic resources is a complex, multi-faceted process. It involves a number of distinct stages, which are nevertheless linked and interrelated, from the selection of priority taxa, to the design and implementation of complementary conservation strategies and the development and exchange of the results of germplasm use (Maxted *et al.*, 1997a). These different components generate various types of data, including information on the identity (passport data) and characteristics of germplasm (characterization and evaluation data), which are crucial for the effectiveness of the process as a whole. In the end, these data refer to stands of wild or cultivated plants found growing in specific, known places. The location of these places on the surface of the earth is also included in the passport data of genebank and herbarium documentation. This means that the data associated with germplasm are 'geo-referenced', and therefore amenable to 'spatial analysis', or the description and modelling of patterns and relationships in geographical data (Bailey, 1994).

Each of the different components of the process of conservation and use of genetic resources not only generates, but also requires, data. For example, germplasm collection results in data on the distribution, the phenology, the ethnobotany and, once characterization and evaluation have been carried out, the genetic diversity of the target taxon. However, the collector will clearly benefit from considering – before venturing out into the field – any such data that may already exist in the literature and the documentation systems of genebanks and herbaria. Because they are geo-referenced, the data coming out of the genetic resources management process can be analysed not only on their own, but also in conjunction with other location data, from whatever additional source. Thus our collector, presented with scanty data on previous collections, could use climatic, vegetation and soil data from the study region to estimate the distribution and phenology of the taxon concerned.

In this chapter, we describe how spatial analysis of the geo-referenced data generated by the process of conservation and use of genetic resources, using geographic information systems (GIS), can feed back to enhance and facilitate the process, and indeed add value to the germplasm collections. We start with a brief introduction to GIS technology. We then examine a number of key stages of the genetic resources management process and discuss how GIS may be used to increase their efficiency and effectiveness. This chapter thus follows the structure of Guarino *et al.* (1999), while updating the information presented there with new examples and additional references (including relevant Internet resources).

Geographic Information Systems

A GIS may be formally defined as a database management system which can simultaneously handle spatial data in graphics form, i.e. maps, or the 'where', and related, logically attached, non-spatial, attribute data, i.e. the labels and descriptions of the different areas within a map, or the 'what'. Simpler definitions have been given, perhaps the most jargon-free being 'a tool for managing information of any kind according to where it is located' (Treweek, 1999). The main elements of a GIS are as follows (Guarino, 1995; Guarino *et al.*, 1999):

- data input, verification and editing;
- data storage, retrieval and management;
- data manipulation and analysis; and
- output.

Data input

Data can be entered into a GIS by digitizing paper maps and their associated attribute information *de novo* using a digitizing table or scanner, or by importing existing digital datasets, including remote sensing images. Genebank curators can import data from the database of their documentation system into a GIS, using the latitude and longitude fields in the passport data to provide the link to digitized maps and remote sensing images. 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). As an example, the spatial data holdings of the IARCs can be searched in the CGIAR Spatial Data Catalogue,[1] and a wide range of global datasets ranging from maps of terrestrial ecosystems to human population density are available at the UNEP/GRID web site[2] and from the Eros Data Centre of the US Geological Survey.[3]

[1] www.griada.no/cgiar.htmls.mdindex.htm
[2] www.grid2.cr.usgs.gov
[3] edcwww.cr.usgs.gov/earthshots/slow/

Data storage

There are two main types of GIS data: vector and raster. 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. In contrast, in raster (or grid) data, an area is divided into an array of regularly shaped cells. 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.

Data manipulation and analysis

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 (Bailey, 1994). 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, include spatial correlation, pattern and trend analysis, and interpolation techniques such as kriging (Bailey, 1994).

Output

GIS allows visualization of spatial data (maps) on a computer screen and the production of printed maps, allowing such manipulations as selecting areas or layers for output, and changing scale and colour. However, the outputs of GIS analyses are

not just maps, but can also include tables, graphs and animations. Three-dimensional visualization is one of the key topics in current GIS development, and is revolutionizing cartographic representation.

Use of GIS in Plant Genetic Resources Conservation and Use

In this section, we discuss how different types of spatial analysis can be applied to locality, characterization and evaluation data in a GIS environment to enhance the efficiency of genetic resources management. We concentrate on the following five components of the process:

1. Ecogeographic surveying;
2. Field exploration;
3. Design, management and monitoring of *in situ* reserves;
4. Germplasm evaluation; and
5. Use of genetic resources.

Ecogeographic surveying

Maxted *et al.* (1995) described the process of collating information on the taxonomy, genetic diversity, geographic distribution, ecological adaptation and ethnobotany of a plant group, as well as on the geography, ecology, climate and the human setting of study regions. The sources of the information for such 'ecogeographic surveys' will include herbarium specimens, germplasm accession passport data, experts, the formal and grey literature, field notes and maps.

It is only on the foundation of basic information such as (but not limited to) this that sensible conservation decisions can be made, for example regarding when, where and how to collect germplasm, and where genetic reserves might best be established and how they would need to be monitored and managed. This is because biodiversity is not evenly distributed over the surface of the earth. Some places will, therefore, be of a higher priority for collection or conservation than others. One of the main objectives of ecogeographic surveys is thus the identification of those geographic areas which are:

- likely to contain specific desired traits (adaptations), taxa or habitats of interest;
- highly diverse (whether environmentally, taxonomically or genetically);
- complementary to each other;
- currently missing or under-represented in conservation efforts;
- threatened with genetic erosion.

Areas likely to contain germplasm of interest

Although data on the geographic distribution of species are often scanty, even this limited information can be used to identify areas where a species has not been previously recorded but where it might still be expected to occur. GIS tools such as BIOCLIM[4] (Busby, 1991; see also GARP,[5] Genetic Algorithm for Rule-set Production, an extension of the BIOCLIM approach), DOMAIN[6] (Carpenter *et al.*, 1993) and FLORAMAP[7] (Jones and Gladkov, 1999; Jones *et al.*, 1997) attempt this type of extrapolation using climate data. They first estimate conditions at sites where a species has been recorded. Although the details of the methodologies differ, the software then derive a climatic 'envelope' for the set of collecting sites and display all other areas that have some level of similarity to the collecting sites. An example of the output provided by FLORAMAP is shown in Fig. 36.1. Afonin and Greene (1999) describe a similar methodology. Model-based approaches (e.g. Walker, 1990; Stockwell and Noble, 1992) have also been used to investigate the so-called 'potential' distribution of a species. An instructive, large-scale example of the application of GIS to the problem of 'objective prediction of the full distribution of a species from incomplete point distribution maps, based on its ecological preferences' is the work being carried out by the Royal Botanic Gardens, Kew,[8] in the context of Madagascar's Environmental Action Plan. These techniques have mainly been used at the species level. However, they can also be applied to infraspecific entities such as botanical varieties, groups of similar accessions (e.g. based on multivariate analysis of morphological or molecular characterization), or even farmer-recognized landraces.

A species being targeted for conservation may

---

[4] www.dino.wiz.uni-kassel.de/model_db/mdb/bioclim.html
[5] www.kaos.erin.gov.au/general/biodiv_model/ERIN/GARP/home.html
[6] www.cgiar.org/cifor/research/intro_d.html
[7] www.ciat.cgiar.org/floramap
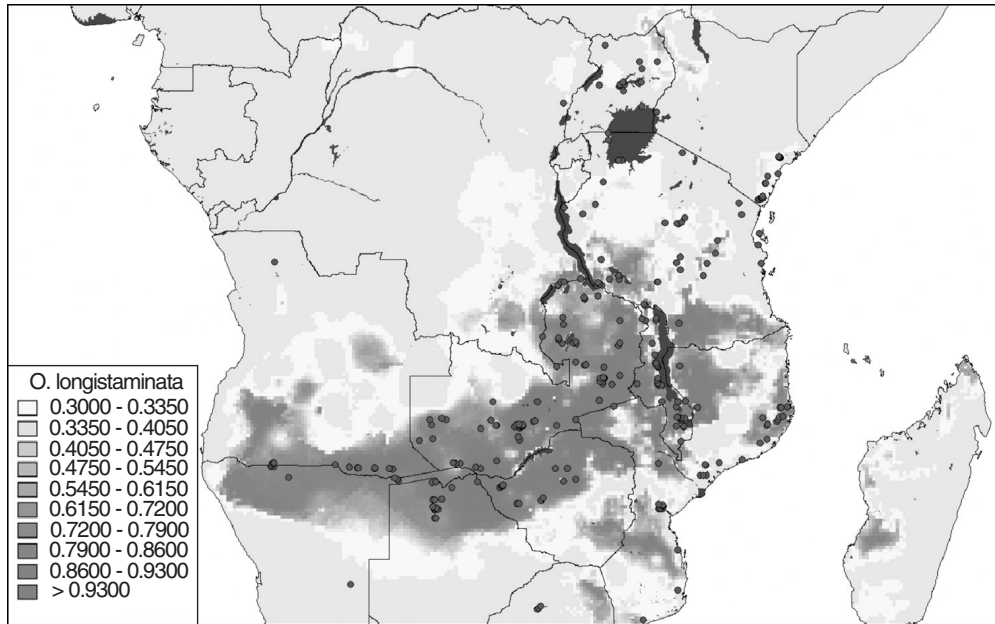[8] www.rbgkew.org.uk/herbarium/madagascar/plant_dis.html

Fig. 36.1. Results of FLORAMAP analysis of the distribution of *Oryza longistaminata* in southern Africa. The shading shows areas with increasingly high levels of climatic similarity with the sites where the species has been encountered (the dark dots, which represent germplasm accessions and herbarium records), and therefore in theory increasing probability of finding the species. The climate variables used were monthly rainfall totals, monthly average temperatures and monthly average diurnal temperature range. Map used courtesy of Kihika Kiambi (IPGRI). See text for more information on the FLORAMAP software.

be closely associated with a specific habitat, vegetation, land use or landform. Remote sensing imagery such as aerial photographs, satellite imagery and radar images, combined with other sources of data in a GIS environment, and properly ground-truthed, can be useful in locating these features. For example, Landsat™ satellite images were used by Veitch *et al.* (1995) to map heathland fragments in southern England, and by Afonin and Greene (1999) to locate herbaceous meadows in the western Caucasus. Isolated areas of cultivation can also be spotted, assuming they are larger than the resolution of the remote sensing system (about 30 m in the case of Landsat images, and 15 m for SPOT images), and in some cases even the crops being grown can be identified.

Obtaining, manipulating and analysing satellite imagery can be a complex, costly and time-consuming task, but some useful software tools are available. An example is WINDISP, which was originally developed for the FAO Global Information and Early Warning System (GIEWS), and is 'a public domain, easy to use software package for the display and analysis of satellite images, maps and associated databases'. It can be downloaded from the Internet.[9] Some data on ecosystems and land use derived from satellite sources are available preprocessed and free on the Internet.[10]

Sometimes material with very specific characteristics needs to be targeted for conservation, perhaps because a need for it has been expressed by the user community or it is specifically threatened. Here the problem is one of using environmental data to identify areas where material with the required adaptation

[9] www.fao.org/giews/english/windisp/windisp.htm
[10] For example see www.edcdaac.usgs.gov/glcc/glcc.html for global datasets derived from AVHRR 1 km remote sensing imagery.

might be expected to occur. GIS can be used to superimpose appropriate thematic coverage from different sources to identify such areas. Thus, for example, CIAT's *An Atlas of Cassava in Africa* (Carter *et al.*, 1992), an excellent hardcopy product of a GIS project, divides the cultivation area of cassava on the basis of production environments, based on agricultural censuses, specialized surveys and crop experts.

If characterization and evaluation of germplasm have already been carried out, geostatistical methods such as kriging and other forms of interpolation can be used to describe the spatial pattern of variation in genetic, morphological and agronomic traits (or combinations of traits) among populations. The effects of natural selection on a broad scale can then be explored by seeking correlations with environmental and other factors, and disentangled from the effects of isolation by distance and micro-environmental variation (e.g. Epperson, 1993; Monestiez *et al.*, 1994). Differential systematics can also be used (Kirkpatrick, 1974), involving the combination of different character contour maps into a single map of a systematic function, the ridges of which reveal areas where maximum change over distance occurs. These have been called genetic boundaries (see also Monmonier, 1973; Pigliucci and Barbujani, 1991). Such analyses can guide the user to areas where specific traits are prevalent, and also help the conservationist identify areas that are relatively homogeneous but different from each other for the characters being studied. This is discussed further in the section on complementarity.

High diversity areas

Some geographic areas show greater taxonomic or genetic diversity than others. Diversity is often the most important consideration at all stages of genetic resource work, yet remains a difficult parameter to map and analyse. Diversity studies usually begin by dividing the target area (or strata within the target area, e.g. climate zones) into a number of smaller zones, for each of which a measure of diversity can be calculated. Different geometric, political or socio-economic spatial units have been used (e.g. see references in Csuiti *et al.*, 1997), ideally areas of equal shape and size (to reduce the area effect on

diversity measures), such as square grid cells, are best. For example, Nabhan (1991) used presence/absence of species to investigate patterns in the taxonomic diversity of wild *Phaseolus* in different grid cells covering the Sierra Madre, Mexico. Measures of diversity based on morphological characters or molecular markers can also be used, as done by Pickersgill (1984) to calculate the morphological diversity of cultivated *Capsicum* spp. in different grid cells within Central and South America. Ferguson *et al.* (1998) calculated genetic diversity in wild *Lens* spp. using randomly amplified polymorphic DNA (RAPD) markers for different grid cells in the Mediterranean basin. To support this type of analysis by genebanks, IPGRI and CIP are collaborating in the development of software called DIVA,[11] which calculates diversity indices for all the cells in a user-defined grid given latitude, longitude and characterization data for a set of accessions, and maps the results.

A complementary method for mapping diversity is the 'point-based' approach used by the Spatial Intraspecific Diversity (SID) software described by Nelson *et al.* (1999).[12] SID reads the coordinates of each accession point in turn, and draws a circle of user-defined radius around it. All accession points lying within this circle are then used to calculate diversity within the defined radius, using either the Shannon-Weaver or Simpson index. The result is assigned to the spatial location of the observation for which the calculation was done. The advantage of this method lies in the ease with which point-based values of diversity can be interpolated and modelled. Figure 36.2 shows sample SID and DIVA outputs for the same dataset.

Various studies have investigated the use of different environmental parameters as surrogates for species diversity (Gaston, 1996; Faith and Walker, 1996). One example of the application of GIS to this kind of investigation is provided by Miller (1986), who showed that variation in elevation, calculated by GIS from topographical data for each of a large number of polygons in the southern Appalachian region, is a useful predictor of the richness of rare species in those polygons. This approach is being applied at the infraspecific level

[11] gis.cip.cgiar.org/gis/tools/diva.htm
[12] www.gis.ciat.cgiar.org/sid
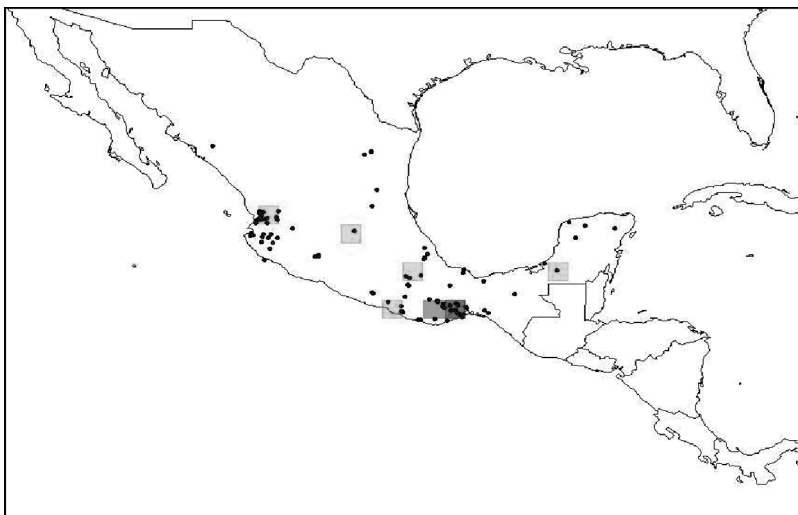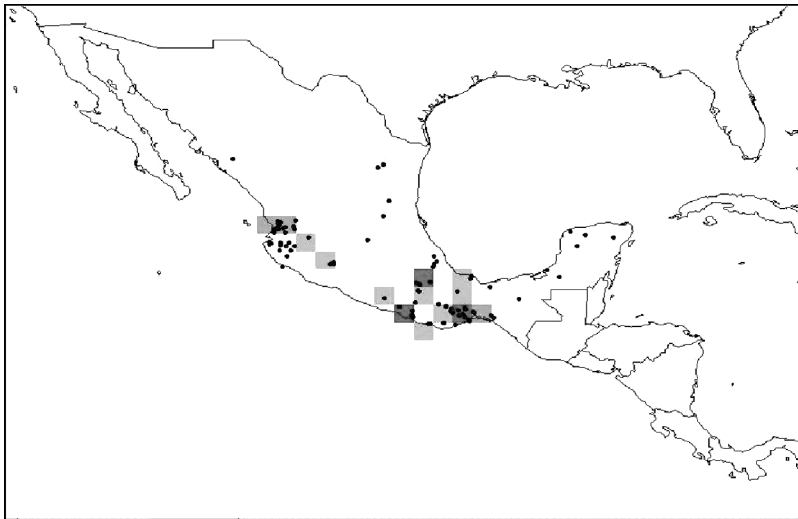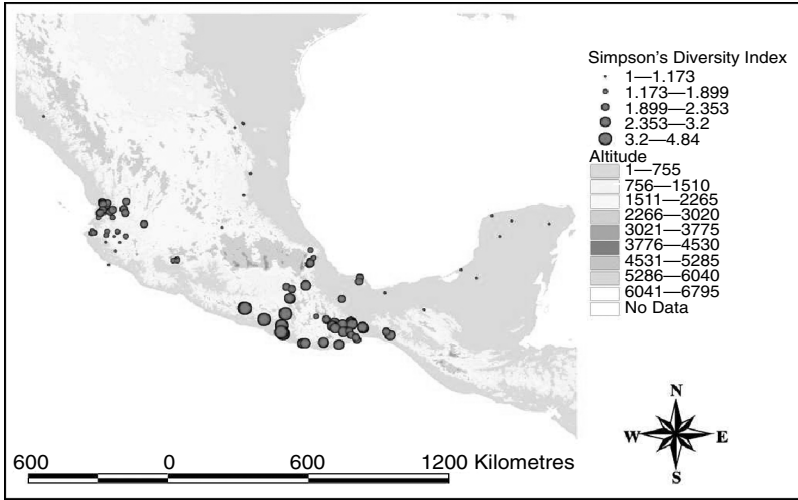
in a study of the relationship between environmental and human diversity and genetic diversity in cultivated groundnut in Ecuador and Guatemala. The study is a collaboration among the national programmes of the countries involved, USDA, IPGRI and CIAT (International Centre for Tropical Agriculture).

Complementary areas

Analyses of diversity can be refined further. One possible enhancement may be necessitated by the fact that two areas may have equal richness or diversity of taxa or morphotypes, but the ones in one square may be similar to each other (e.g. closely related), while those in the other may be more different. Other things being equal, the second area would be the higher conservation priority (Humphries *et al.*, 1995). The procedure described by Vane-Wright *et al.* (1991), and available in their WORLDMAP software,[13] allows the diversity measure to be weighted for the distinctness of taxonomic units, calculated from a phylogeny based on cladistic analysis. Another possible refinement might be to use the potential distribution of species, say as predicted by FLORAMAP, to calculate diversity values, to counteract the kinds of sampling biases described by Hijmans *et al.* (2000a).

It is not enough to simply identify areas that are highly diverse to maximize the amount of diversity protected for a given amount of effort, because all the higher diversity areas might actually contain the same diversity, as well as the same amount of diversity. One approach to the optimal targeting of conservation effort is to use 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 (Booth *et al.*, 1989; Pollak and Corbett, 1993; Corbett, 1998). Spatial units can then be targeted for conservation separately in a stratified manner from each distinct cluster of similar units.

More sophisticated methods are also available. Iterative procedures (e.g. Rebelo and Sigfried, 1992) can be used to choose the smallest number of spatial units such that each species, morphotype, etc. will be present in at least one (or two, three, etc.) unit(s) in the set. A recent study of 19 different techniques (Csuti *et al.*, 1997) found that this and various other heuristic techniques can all be very efficient at solving the problem of selecting potential biodiversity conservation sites. However, except when dealing with large, complicated datasets, the study recommends the use of a linear programming approach called a 'branch-and-bound' algorithm. The DIVA software tool supports complementarity analysis as described by Rebelo and Sigfried (1992), and is being used on a variety of datasets, notably the geographic distribution of wild potato species.

Under-conserved areas

The WORLDMAP software allows the user to select a grid square so that a subsequent run identifies those grid squares which are complementary to the selected one. If the selected grid square is the most diverse, the process approximates that of Rebelo and Sigfried (1992). On the other hand, if the selected grid square includes an existing protected area (a global dataset on protected areas is available from WCMC, the World Conservation Monitoring Centre[14]) the result is 'gap analysis.' The concept and some applications are discussed

---

[13] www.nhm.ac.uk/science/projects/worldmap
[14] www.wcmc.org.uk/cis

Fig. 36.2. (*opposite*) Comparison of diversity analysis using point-centred and grid-based approaches. The dataset consists of locality data for ten species of *Stylosanthes* in Mexico, collated from 119 herbarium specimens by Susana Gama of UNAM as part of a BADC-funded IPGRI research project. The first map shows the results of point-centred species-level diversity analysis using the SID software (Simpson's diversity index). The dots show the location of specimens. The size of the dots is proportional to the species diversity in a one-degree circle around each accession point. The second map shows the results of analysing the same dataset using the DIVA software. The dark grid cells are the most diverse (Shannon-Weaver diversity index), light grid cells the least diverse, and others intermediate. The third map shows the results of complementarity analysis as described by Rebelo and Sigfried (1992), and implemented by DIVA. The dark grid cell is the most diverse. The other grid cells are the additional ones that would have to be conserved to retain at least one population of each species. See text for more information on SID and DIVA.

in detail by Scott *et al.* (1993).[15] The process involves the use of 'digital map overlays in a GIS to identify individual species, species-rich areas, and vegetation types that are not represented or under-represented in existing biodiversity management areas.' These 'existing biodiversity management areas' could be protected areas, but also areas where germplasm collection has already been adequately carried out. Identifying ecogeographical gaps in existing *ex situ* germplasm collections in this way is increasingly important as the large-scale, coarse-grid collecting of the past 25 years is replaced with a more targeted, fine-grid approach.

Threatened areas

Areas under imminent threat of genetic erosion are clearly going to be high priorities for conservation (Guarino, 1999). But how can we predict where genetic erosion is going to take place before it actually happens? Remote sensing has been used to document and model changes in the extent and characteristics of forests (e.g. Skole and Tucker, 1993; Gastellu-Etchegorry *et al.*, 1993),[16] wetlands (e.g. Sebastini *et al.*, 1989), deserts,[17] different land use types (e.g. Zheng *et al.*, 1997) and cropping,[18] and must surely have a role to play in predicting genetic erosion.

Hutchinson and Weiss (1999) have recently described what that role might be. They present a model where different potential causes of genetic erosion are associated with specific indicators. They then list observable phenomena by which these indicators may be assessed, and finally the means by which information on the values of these observables might be obtained, including such diverse sources as remote sensing imagery, agricultural censuses, news reports and fieldwork. A similar conceptual framework has been successfully applied to early warning of crop failure and famine in Africa (for example by USAID's Famine Early Warning System, FEWS[19]). Thus, if environmental degradation is postulated as a possible cause of genetic erosion for a particular genepool, indicators might include desertification, soil erosion and the extension of agriculture into increasingly marginal areas. Observables might include decreasing peak NDVI (Normalized

Difference Vegetation Index, a measure of the development of vegetation) values over the course of years, increasing outmigration and the appearance of new fields in marginal areas, all of which can be managed in a GIS environment.

However, the link between these kinds of changes and genetic changes in specific crops or wild species of interest has still not been sufficiently investigated. We may have observables, but we are not sure in most cases whether they are in fact associated with genetic erosion, because this has rarely been measured. An exception is perhaps habitat fragmentation, because the link between this phenomenon and genetic diversity, particularly in tree species, has been the subject of a number of theoretical and practical investigations (e.g. Templeton *et al.*, 1990; Ledig, 1992). The use of remote sensing to monitor and predict genetic erosion in the same way as is being done for food security and famine remains a theoretical, though no less interesting, and tantalizing, possibility.

The somewhat different problem of identifying areas of past genetic erosion is exemplified by ongoing work by CIP and IPGRI on cultivated potatoes in Peru. In this study, isozyme data from the CIP *ex situ* potato collection are being used to estimate levels of genetic variation in each 20 km $\times$ 20 km^2 in a grid covering the whole potato-growing area of the country. Data on the extent of potato cultivation and the biophysical environment in each square are used as the independent variables in a multiple regression model, with genetic diversity as the dependent variable. Negative deviations from the model are then mapped to identify possible areas of genetic erosion, and correlations sought with a number of socio-economic variables (such as accessibility and population growth) to identify possible causative factors for any lower than expected level of genetic diversity.

Synthetic analyses: a case study

GIS tools can be used not only to map the different kinds of areas described above, but also, crucially, to show where they coincide. We have used a dataset on the distribution of 15 wild *Gossypium* species in Africa to illustrate a possible methodology for the

---

[15] See also www.gap.uidaho.edu/gap for an application of this methodology.
[16] edcwww.cr.usgs.gov/earthshots/slow/Rondonia/Rondonia
[17] www.medalus.leeds.ac.uk/medalus.html
[18] edcwww.cr.usgs.gov/earthshots/slow/Mozambique/Mozambique
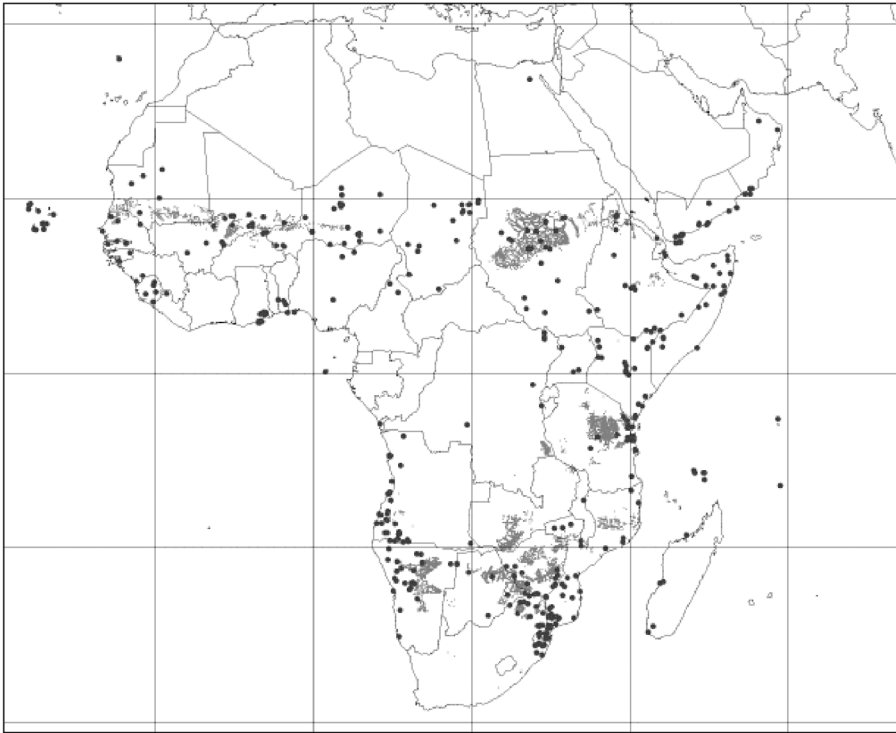[19] www.info.usaid.gov/fews

Fig. 36.3. High priority areas for conservation of wild *Gossypium* in Africa are shown in the shaded areas. Dots are locations of herbarium specimens. See text for details of the analysis.

integration of different datasets and analyses in a GIS environment in support of a large-scale ecogeographic study (Holubec, 1998). The methodology involves four separate analyses, the results of which are then combined to produce a single map of priority areas for conservation interventions.

1. Probability of high diversity. First, the SID software was used on the localities of 607 herbarium specimens to calculate species diversity within a 1-degree radius of each collecting point. The locality data of the top 20% diversity points were then imported into FLORAMAP. The FLORAMAP output was a probability surface displaying areas with climatic conditions most suitable to harbour a high diversity of *Gossypium* species.

2. Proximity to existing accession points. Next, a map was prepared which prioritizes areas most distant from existing collections, thus targeting geographical gaps in existing collections.

3. Proximity to roads. Third, an accessibility map was developed, with ease of access running from 1

(directly on a road) to 0 (20 km from the road). A CIAT dataset on road networks was used.

4. Genetic erosion risk assessment. Fourth, an index of genetic erosion risk was produced, using a variety of data sources. The first component of the index is severity of soil degradation, using a dataset from the International Soil Reference and Information Centre.[20] The threat of habitat loss was crudely modelled based on population growth. Those areas undergoing the highest rate of population growth were assumed to present the greatest risk of genetic erosion. This analysis was made using a dataset on population density[21] for 1970 and 1990. The final component uses an ILRI dataset on cattle density for the African continent (Kruska *et al.*, 1995). It is assumed that a high density of cattle produces a significant risk to wild species of cotton. These three components were put together with equal weighting to form a single index of genetic erosion risk.

Figure 36.3 shows the results of combining these four components into a single map, which highlights

[20] www.isric.nl/GLASOD.htm
[21] www.grid.cr.usgs.gov/clearinghouse/datalist.html

those areas most in need of conservation because of a combination of high diversity and high risk of genetic erosion. Such a methodology can be adapted to account for the specific objectives of a conservation programme. It could be refined further by, for example, overlaying the results of complementarity analysis or giving higher priority to specific taxa. It is currently being applied by IPGRI, USDA and national partners in support of a programme of exploration for wild *Arachis* spp. in Bolivia.

Fieldwork

The ecogeographic study is essentially desk-based. It needs to be refined and expanded, and its results implemented, through fieldwork. GIS technology can contribute to this phase of genetic resources conservation through:

- the development of field aids; and
- providing information on the optimal timing of field visits.

Field aids

Maps are crucial to fieldwork. Unfortunately, published maps showing a specific factor of interest to the conservationist (or combination of factors) may not be available. An example of the use of GIS to develop mapping products for use in the field is provided by Greene *et al.* (1999a). A number of moisture and temperature maps synthesizing some 60 monthly climate variables at 500 m resolution were prepared in support of a joint USA/Russia forage collecting programme in the Caucasus. Even when published maps are available they may be unwieldy to use: a single soil map of the study region in the Caucasus was synthesized from four different map series. These maps allowed the identification of ecogeographic gradients in the field, the slope of which influenced decisions about sampling frequency. They were also used to monitor, during the course of the fieldwork, which combinations of environmental conditions were being adequately sampled and which were not.

Timing

Rainfall can be unpredictable in both space and time, especially in the arid and semiarid tropics,

making it difficult sometimes to time field exploration work to coincide with the most appropriate stage of vegetation development. Satellite imagery can provide data on the state of vegetation with relatively short lag-times. Use of such data could allow collectors to be more precise in timing their visit, adding greatly to the cost-effectiveness of collecting trips. Meteosat and NOAA/AVHRR (National Oceanic and Atmospheric Administration Advanced Very High Resolution Radiometer) data on rainfall and the state of vegetation (as measured by NDVI), although fairly low in resolution, can be analysed to allow surveillance of the state of crops and vegetation over the growing season. Justice *et al.* (1987) describe the annual course of NDVI in a variety of East African vegetation types, and how this measure relates to the phenology of rainfall and plant growth. The USA/Russia forage collecting programme mentioned above used Landsat MSS (multi-spectral scanner) and TM (thematic mapper) imagery not only to locate the primary target of meadow patches, but also to evaluate the stability of their phenology across years (Afonin and Green, 1999). The patches with the most predictable phenology were considered the highest priority for field visits, because of the greater likelihood of finding seed of the target species for collection.

FAO's WINDISP software package has various features that are of relevance in this context, for example, tools to compare two images and analyse trends in a time-series of images. Much analysis may not be necessary, however, as there are sources of relevant pre-packaged data. It is possible for a collector in, say, Zimbabwe to determine the location of potential collection areas from the satellite imagery available on the Zimbabwe Meteorological Services Department's Vegetation Index Report web pages[22] or indeed FAO's GIEWS[23] and then visit the areas a matter of days later.

Design, management and monitoring of *in situ* reserves

The design of protected areas, including questions of optimal size and shape, zonation and networking (e.g. Given, 1994; Hawkes *et al.*, 1997) is a spatial problem, and spatial analysis in a GIS environment has been applied to them, although mainly in the

---

[22] www.weather.utande.co.zw/navigator/latest-vegetation.idc
[23] www.fao.org/giews/english/giewse.htm

context of ecosystem or animal conservation. Thus, Howard (1996) discusses how spatial information on species richness, distribution and abundance of an endangered species, disturbance and distribution of timber resources within a forest can be used to develop a zoning plan for a forest reserve, including different use areas, buffer zones and a core. This type of application presents the challenge of integrating demographic, socio-economic, cultural and other data on, and from, the human population with data on the biophysical environment and on the target taxon. For example, Fox *et al.* (1996) used GIS to map areas where the objectives of protecting red pandas and those of meeting the claims and grazing-land needs of communities living in the area either coincided or came into conflict. Harmsworth (1998) described an attempt to manage within a GIS information on the cultural values of different features of the landscape, flora and fauna, with a view to developing resource and environmental management plans more in tune with the requirements of local people.

A particularly comprehensive example of the application of GIS to the design of an *in situ* reserve is the work of Kremen *et al.* (1999) on the Masoala National Park in Madagascar. They defined a set of 13 design criteria relating to ecological and socio-economic sustainability, which, for example, included that special consideration should be given to rare and threatened habitats and species, that the limits of the park should follow natural features to make them easy to respect and that buffer areas should be large enough to meet the subsistence need of the human population surrounding the park. Information needs were then identified, combining maps, satellite imagery, biodiversity surveys and population studies. Data from these different sources were then analysed in a GIS environment to develop a park proposal balancing human and wildlife needs, including a core rainforest area, three satellite marine reserves and a large surrounding multiple-use zone. The proposal was discussed with local people and, once their approval was obtained, formed the basis of a national decree establishing the national park.

Given the rapid rate of environmental and socio-economic changes (including the threat of global climate change), it is important for reserve design to take into account possible future scenarios. GIS combined with computer modelling can be used to optimize reserve selection and management based not only on current climatic and vege-tation patterns, but also under different assumptions of change. Menon and Bawa (1997) used GIS to document land use change, deforestation and habitat fragmentation in the Western Ghats of India, and investigate the socio-economic drivers of these processes to develop predictive models. These were then used to estimate the current and future effectiveness of the protected area network for biodiversity conservation.

Once the *in situ* reserve has been established, some form of management intervention may need to be implemented, and its results on the frequency, abundance, demography and genetic diversity of target species monitored (Maxted *et al.*, 1997b). GIS can assist in defining and implementing the management plan. Although most examples of such applications relate to ecosystem management (e.g. Kessell, 1990), Liu *et al.* (1995) used a spatial population model in a GIS to simulate the potential effects of a range of forest management plans on the size of the population of a bird species. Working in a GIS environment can ease the application of such analytical techniques of plant demography as transition matrix sensitivity analysis (Silvertown and Lovett Doust, 1993), thus hopefully stimulating what is still their incomplete integration into the armoury of plant genetic resources conservation.

Germplasm regeneration and evaluation

A key step in both germplasm multiplication and evaluation is to decide what sites are best for which accessions. The environment of the regeneration site must not only provide suitable conditions for reliable flowering and seed production, but also any necessary triggers for the different stages of plant development (Sackville Hamilton and Chorlton, 1997). Unless controlled pollination is practised, pollinators must also be present. On the other hand, pests and diseases must not be prevalent and populations of species related to the one(s) being regenerated should not be found nearby. Climate data and distribution maps for pests, diseases, pollinators and wild relatives can be overlapped using GIS to identify potential sites for regeneration. If the choice of locations for regeneration is limited, which is usually the case, such data can be used to estimate the relative suitability of different available sites for the regeneration of different accessions, by comparison with the environment at their site of collection (e.g. using FLORAMAP). They can also

help to determine the unique needs of individual accessions so that growing procedures can be customized in sub-optimal sites.

As for germplasm evaluation, both ecologically optimal and stress locations are usually necessary. Classification of testing locations using GIS-derived climatic data (e.g. Pollak and Pham, 1989) can be used to determine the suitability of different sites for the evaluation of specific traits, or combinations of traits, of interest. If a site has already been chosen, GIS can be used to determine the extent of the region where similar conditions apply, and hence the relevance and potential impact of the evaluation work (Chapman and Barreto, 1996). CIMMYT's *Africa Maize Research Atlas* on CD-ROM (Hodson *et al.*, 1999) provides an interesting illustration of how GIS can support germplasm evaluation work. It provides a compendium of data layers and an easy-to-use GIS exploration tool, allowing researchers to 'characterize regions of interest and target and predict the potential impacts of promising germplasm'.

Use of genetic resources

The major aim of conserving germplasm is, of course, for it to be used. Adequate access policies, good linkages between genebanks and users and well-developed breeding and other use programmes are necessary conditions for genetic resources to be used to the full; but they are not sufficient. The lack of data on accessions and the large size of collections can still present significant bottlenecks to the use of conserved material. GIS can help to alleviate the problem in the first instance by improving the quality and quantity of the locality data associated with collections. Given adequate data on the origin of material, GIS can also help to identify particularly interesting germplasm and reveal the structure of diversity within collections. Finally, GIS can be used to assess the potential impact of the products of germplasm use.

Better locality data on accessions

Typically, a large proportion of germplasm accessions, especially older ones, lack latitude and longitude data, though elevation data are more common. GIS can help the collector and genebank curator in completing such passport data. Gazetteers, both published and on-line (e.g. the *Getty Thesaurus of Geographic Names*[24] and the *GEOnet Names Server*,[25] which provide the coordinates for about 3.5 million global geographic features, such as populated places), maps and collectors' notebooks can be used to determine latitude and longitude for accessions where only a locality name is available. With localities marked on paper maps, latitude and longitude may be estimated by using digitized maps of the area and employing the common facility of GIS software to return coordinates when a specific point is selected, say by moving the cursor to it.

Even when passport data are available, they are sometimes inaccurate. Rapid data visualization using a GIS can reveal obvious errors in latitude and longitude data resulting in accessions falling into the sea, a lake or the wrong country. Ecogeographic outliers will also be suspect. Chapman and Busby (1994) describe a method of spotting outliers which involves plotting the climate profile of each specimen and looking for entries that are out of step with the rest. Similarly, FLORAMAP can be used to identify accessions with climatically discordant localities and to carry out other data checking analyses (Jones and Gladkov, 1999). Data can be checked further by comparing locality descriptions in the passport data with GIS datasets of administrative boundaries. A methodology for such data-checking is presented by Hijmans *et al.* (1999).

Germplasm targeting

Next, once locality data are better and more complete, such data as altitude, major soil type, land use and vegetation, if missing from the field documentation, can be estimated by overlaying the locations of sites on different thematic base maps of the appropriate scale and reading off the map attributes. Interpolated climate surfaces can also be used to characterize conditions at each collection site given its latitude, longitude and altitude, for example using FLORAMAP. Steiner and Greene (1996) describe this process as 'retro-classification' of accessions and give an example of its usefulness using the *Lotus* collection of the US National Plant Germplasm System (NPGS) (see also Greene *et al.*, 1999b).

Retro-classification of the collection site can guide the use of the material by focusing the attention of users on the most promising material for their specific

[24] www.shiva.pub.getty.edu/tgn_browser/
[25] www.164.214.2.59/gns/html/index.html

purposes. For example, although they have been developed in the context of introduction of tree species, the climate matching programmes described by Booth (1990; 1999) for various countries are valuable tools for targeting the introduction of germplasm to specific areas for evaluation and use based on the conditions where it was collected. Similarly, Chapman and Taba (quoted by Chapman and Barreto, 1996) overlaid germplasm locations on soil maps to identify candidates from an extensive maize collection that might be adapted to alkaline soils. A geographical approach was also used by Beebe *et al.* (1997) to identify edaphic adaptation in beans as part of a breeding programme.

Structure of collections

'A core collection is a selected and limited set of accessions derived from an existing germplasm collection, chosen to represent the genetic spectrum in the whole collection and including as much as possible of its genetic diversity' (Brown, 1993). The core collection can be used to evaluate traits that are expensive or time-consuming to measure on the whole collection, for example complex yield and quality traits and general combining ability with local germplasm. These studies can then be used to identify material in the rest of the collection that might require further investigation.

Different workers have employed passport, characterization (taxonomic, morphological, molecular, etc.) or evaluation data, often in different, usually hierarchical, combinations, to classify accessions into groups and identify duplicates. However, the importance of location data is widely acknowledged and agroclimatic conditions at the collection site is often one of the more important types of criteria used in the classification process. The case of the *Phaseolus vulgaris* core collection described by Tohme *et al.* (1995) illustrates this – and the role of GIS – well. First, regions were prioritized based on the history of the crop. Then, interpolated surfaces for four parameters (length of growing season, photoperiod, soil type and moisture regime) were used to define 54 distinct environments, and each 10-minute grid cell was assigned to one of these classes. Passport data were then used to match each landrace accession to an environmental class. Finally, accessions in each environmental class were stratified according to characterization data (growth habit, grain colour and grain size) and selections made at random from within each stratum within each environmental class.

Environmental data on accessions derived from analysis using GIS can be used to investigate the structure of genetic diversity collections without going as far as developing core collections. Thus, analysis of various datasets using FLORAMAP, including *Passiflora* spp. in the Andes (Sergio Segura, personal communication, 2000), wild rice in southern Africa (Kihika Kiambi, personal communication, 2000) and *Stylosanthes* in South America (Jones and Gladkov, 1999), has revealed correspondences between the groupings of accessions by clustering using the climatic characteristics of collection sites and those derived from isozyme and DNA marker data. Indeed, Chapman and Barreto (1996) have suggested that as the linkage between genetic resources and environmental data is explored, 'geographic trait loci' may be identified, where particular parts of the genome are linked to adaptation to different conditions.

Assessing impact

GIS has a role to play not only in facilitating the use of germplasm, but also in predicting the likely impact of that use. Thus, a breeding programme could weigh the relative effects of achieving a range of different improvement objectives, and then concentrate on the ones likely to have the greatest impact with a particular priority target group. An example is provided by recent work at CIP on frost tolerance in potato (Hijmans *et al.*, 2000b). A crop growth simulation model was used in conjunction with high-resolution monthly climate surfaces for the Altiplano of southern Peru and northern Bolivia to predict the likely effect on yield in different areas within the region of increasing the frost tolerance of a standard variety by specific amounts. This type of analysis is strongly dependent on the availability of the kind of detailed, within-country, geo-referenced crop distribution data that are being assembled by CGIAR Centres (e.g. Huaccho and Hijmans, 1999; Hyman, 1999).

Prospects

The application of GIS to estimating the impact of improved germplasm opens the way for a fully integrated approach to the use of GIS in plant genetic resources management, in which GIS-based analysis guides and facilitates the process from start to finish and back again. First, the likely impact of a particular potential use product in a given area is assessed

using GIS; for example a more frost-tolerant potato variety. Next, retro-classification of past collection sites is carried out with GIS to identify germplasm in existing national and international collections which could be useful in developing the proposed product. In the potato case, interpolated surfaces for the occurrence or severity of frost in the Andes could be used. If sufficient suitable material is not available in genebanks, climatic data are used to identify areas where further germplasm collection should be done. This could be approached by super-imposing maps of frost severity and potato cultivation, or by running the collecting localities of frost-tolerant accessions through FLORAMAP. Once promising material has been obtained, GIS is used to identify appropriate evaluation sites. The best material coming out of the evaluation programme is then used to develop the improved product, the actual performance of which is finally fed back into the impact model to start the cycle all over again.

The potential is clearly there for GIS-based analysis of data on germplasm accessions to add greatly to the value of the data, and thus of the germplasm to which they pertain. It should significantly enhance the cost-effectiveness of conservation efforts, and facilitate use of germplasm by breeders and others. However, it must be admitted that GIS technology has not been taken up by plant genetic resources conservation programmes to the extent that one might have predicted on the basis of its potential. Part of the reason is that many such programmes, particularly in developing countries, have significant resource constraints, and GIS hardware, software and data are perceived as being expensive, difficult to obtain and complex to use.

The perceived 'barriers to entry' may still be too high, but the current revolution in GIS technology is putting both the data and the analytical tools within the reach of many. However, if the potential of GIS in the field of plant genetic resources conservation and use is to be fulfilled, the plant genetic resources conservation community needs to take positive action itself at the international, regional and national levels, and not simply wait for the technology to come to its aid. At the international level, the CGIAR Centres must show the same leadership in the application of GIS-based approaches to plant genetic resources conservation and use as they have been showing in their application to natural resources management. An important first step will

be the integration of more GIS capability into the International Crop Information System (ICIS) being developed by CIMMYT (Fox and Skovmand, 1996) and the on-line database of the System-wide Information Network for Genetic Resources (SINGER[26]). In fact, Web-based tools, where maps and accession points can be manipulated on-line with easy-to-use, browser-based, graphical interfaces will help wider adoption of GIS by plant genetic resources conservation programmes.

IPGRI has been collaborating with other CGIAR Centres, in particular CIAT and CIP, in implementing a three-pronged strategy of support to GIS adoption by national and regional programmes. This entails the following activities:

- development, testing and dissemination of a menu of basic methodologies;
- development, promotion and distribution of easy-to-use, inexpensive software tools (and data) specifically aimed at carrying out the analyses alluded to in the previous point; and
- awareness-building.

The approach that has been taken is to implement a series of case studies to develop methodologies, whenever possible based on the use of specialized software such as FLORAMAP, SID and DIVA. The hope is that the results of such case studies will be used to build awareness of the potential of GIS technology at both the technical and decision-making levels within the plant genetic resources conservation community.

At the regional level, plant genetic resources networks have a potentially pivotal role to play in the process of GIS adoption. If GIS technology is still not within the reach of many individual national programmes, it should be possible to find a mechanism whereby the members of a regional network can share the costs, as well as benefits, of GIS adoption. GIS provides an ideal environment within which to manage and analyse data from a number of countries in a common framework, and thus develop both regional and national conservation strategies based on a broad vision of the problems and opportunities.

At the national level, genebanks need to forge linkages with institutions which already have GIS capacity, for example meteorological services, the geography departments of universities, mapping and survey departments and national statistical and natural resources institutes. This will guarantee

---

[26] www.singer.cgiar.org/

access to data and expertise in the longer term and, therefore, sustainability. It will also eventually place genebank managers in a position to lobby decision-makers with tangible evidence of the gains in effectiveness that can accrue from the use of this technology, so that both national and international resources may be accessed for the eventual acquisition of in-house GIS capability.

It is certainly possible to visualize in a not too distant future GIS being used routinely and in an integrated fashion in support of plant genetic resources conservation and use as outlined above. However, before we are swept away by our enthusiasm, it is worth reminding ourselves of the fact that none of the analyses that have been described here would be possible without adequate data on germplasm, in particular accurate locality data in coordinate form. At the risk of seeming prosaic, the future of GIS adoption, with the improvements in efficiency of plant genetic resources programmes which it will bring, is likely to be more dependent on these programmes having solid documentation components, that is, effective data management systems and lots of data to manage, than on methodological or software innovations.

Acknowledgements

References

Afonin, A. and Greene, S.L. (1999) Germplasm collecting using modern geographic information technologies: directions explored by the N.I. Vavilov Institute of Plant Industry. In: Greene, S.L. and Guarino, L. (eds) *Linking Genetic Resources and Geography: Emerging Strategies for Conserving and Using Crop Biodiversity*. CSSA Special Publication No. 27. ASA and CSSA, Madison, Wisconsin, pp. 75–85.

Bailey, T.C. (1994) A review of statistical spatial analysis in geographical information systems. In: Fotheringham, S. and Rogerson, P. (eds) *Spatial Analysis and GIS.* Taylor & Francis, London, pp. 13–44.

Beebe, S., Lynch, J., Galwey, N., Tohme, J. and Ochoa, I. (1997) A geographical approach to identify phosphorus-efficient genotypes among landraces of common bean. *Euphytica* 95, 325–336.

Booth, T. (1990) Mapping regions climatically suitable for particular tree species at the global scale. *For. Ecol. Manag.* 36, 47–60.

Booth, T. (1999) Matching germplasm to geography: environmental analysis for plant introduction. In: Greene, S.L. and Guarino, L. (eds) *Linking Genetic Resources and Geography: Emerging Strategies for Conserving and Using Crop Biodiversity*. CSSA Special Publication No. 27. ASA and CSSA, Madison, Wisconsin, pp. 63–74.

Booth, T.H., Searle, S.D. and Boland, D.J. (1989) Bioclimatic analysis to assist provenance selection for trials. *New For.* 3, 225–234.

Brown, A.H.D. (1993) The core collection at the crossroads. In: Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. and Morales, E.A.V. (eds) *Core Collections of Plant Genetic Resources*. John Wiley & Sons, Chichester, UK, pp. 3–19.

Busby, J.R. (1991) BIOCLIM – a bioclimate prediction system. In: Margules, C.R. and Austin, M.P. (eds) *Nature Conservation: Cost Effective Biological Surveys and Data Analysis*. CSIRO, Melbourne, Australia, pp. 4–68.

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

Carter, S.E., Fresco, L.O., Jones, P.G. with Fairbairn, J.N. (1992) *An Atlas of Cassava in Africa: Historical, Agroecological and Demographic Aspects of Crop Distribution*. CIAT, Cali, Colombia.

Chapman, A.D. and Busby, J.R. (1994) Linking plant species information to continental biodiversity inventory, climate modeling and environmental monitoring. In: Miller, R.I. (ed.) *Mapping the Diversity of Nature*. Chapman and Hall, London, pp. 179–195.

Chapman, S.C. and Barreto, H.J. (1996) Using simulation models and spatial databases to improve the efficiency of plant breeding programs. In: Cooper, M. and Hammer, G.L. (eds) *Plant Adaptation and Crop Improvement*. CAB International, Wallingford, UK, pp. 563–587.

Chatelain, C., Gautier, L. and Spichiger, R. (1996) A recent history of forest fragmentation in south-western Ivory Coast. *Biodiversity and Conservation* 5, 37–53.

Corbett, J.D. (1998) Classifying maize production zones in Kenya through multivariate cluster analysis. In: Hassan, R.M. (ed) *Maize Technology Development and Transfer: A GIS Application for Research Planning in Kenya*. CAB International, Wallingford, UK, pp. 15–25.

Csuti, B., Polasky, S., Williams, P.H., Pressey, R.L., Camm, J.D., Kershaw, M., Kiester, A.R., Downs, B., Hamilton, R., Huso, M. and Sahr, K. (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. *Biological Conservation* 80, 83–97.

Epperson, B.K. (1993) Recent advances in correlation studies of spatial patterns of genetic variation. *Evolutionary Biology* 27, 95–155.

Faith, D.P. and Walker, P.A. (1996) Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity and Conservation* 5, 399–415.

Ferguson, M.E., Ford-Lloyd, B.V., Robertson, L.D., Maxted, N. and Newbury, H.J. (1998) Mapping the geographical distribution of genetic variation in the genus *Lens* for the enhanced conservation of plant genetic diversity. *Molecular Ecology* 7, 1743–1755.

Fox, J., Yonzon, P. and Podger, N. (1996) Mapping conflicts between biodiversity and human needs in Langtang National Park, Nepal. *Conservation Biology* 10, 562–569.

Fox, P.N. and Skovmand, B. (1996) The International Crop Information System (ICIS) – connects genebank to breeder to farmer's field. In: Cooper, M. and Hammer, G.L. (eds) *Plant Adaptation and Crop Improvement*. CAB International, Wallingford, UK, pp. 317–326.

Gastellu-Etchegorry, J.P., Estreguil, C., Mougin E. and Laumonier, Y. (1993) A GIS based methodology for small scale monitoring of tropical forests – a case study in Sumatra. *International Journal of Remote Sensing* 14, 2349–2368.

Gaston, K.J. (1996) Species richness: measure and measurement. In: Gaston, K.J. (ed.) *Biodiversity: A Biology of Numbers and Difference*. Blackwell Science, Oxford, UK, pp. 77–113.

Given, D.R. (1994) *Principles and Practice of Plant Conservation*. Chapman and Hall, London.

Greene, S.L., Hart, T. and Afonin, A. (1999a) Using geographic information to acquire wild crop germplasm: I. Map development and field use. *Crop Science* 39, 836–842.

Greene, S.L., Hart, T. and Afonin, A. (1999b) Using geographic information to acquire wild crop germplasm: II. Post collection analysis. *Crop Science* 39, 843–849.

Guarino, L. (1995) Geographic information systems and remote sensing for the plant germplasm collector. In: Guarino, L., Ramanatha Rao, V. and Reid, R. (eds) *Collecting Plant Genetic Diversity. Technical Guidelines*. CAB International, Wallingford, UK, pp. 315–328.

Guarino, L. (1999) Approaches to measuring genetic erosion. In: Serwinski, J. and Faberova, I. (eds) *Proceedings of the Technical Meeting on the Methodology of the FAO World Information and Early Warning System on Plant Genetic Resources, 21–23 June 1999, Prague, Czech Republic*. Research Institute of Crop Production, Prague, Czech Republic and FAO, Rome, Italy, pp. 26–28.

Guarino, L., Maxted, N. and Sawkins, M. (1999) Analysis of geo-referenced data and the conservation and use of plant genetic resources. In: Greene, S.L. and Guarino, L. (eds) *Linking Genetic Resources and Geography: Emerging Strategies for Conserving and Using Crop Biodiversity*. CSSA Special Publication No. 27. ASA and CSSA, Madison, Wisconsin, pp. 1–24.

Harmsworth, G. (1998) Indigenous values and GIS: a method and a framework. *Indigenous Knowledge and Development Monitor* 6, 3–7.

Hawkes, J.G., Maxted, N. and Zohary, D. (1997) Reserve design. In: Maxted, N., Forde-Lloyd, B.V. and Hawkes, J.G. (eds) *Plant Genetic Conservation: The* In Situ *Approach*. Chapman and Hall, London, pp. 132–143.

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

Hijmans, R.J., Garrett, K.A., Huaman, Z., Zhang, D.P., Schreuder, M. and Bonierbale, M. (2000a) Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. *Conservation Biology*.

Hijmans, R.J., Condori, B., Carillo, R. and Kropff, M.J. (2000b) Estimating the potential impact of frost tolerant potato cultivars in the Altiplano (Peru and Bolivia). *Proceedings of the Third International Symposium on System Approaches for Agricultural Development (SAAD3)*. CD-ROM publication. CIP, Lima, Peru.

Hodson, D.P., Rodriguez, A., White, J.W., Corbett, J.D., O'Brian, R.F. and Banziger, M. (1999) *Africa Maize Research Atlas* (v. 2.0). CIMMYT, Mexico DF, Mexico.

Holubec, V. (1998) Extinction threat of wild African *Gossypium* species in their center of diversity. In: Damania, A.B.,

Valkoun, J., Willcox, G. and Qualset, C.O. (eds) *The Origins of Agriculture and Crop Domestication*. ICARDA, Aleppo, Syria, pp. 286–290.

Howard, P.C. (1996) Guidelines for the selection of forest nature reserves, with special reference to Uganda. In: Bennun, L.A., Aman, R. and Crafter, S. (eds) *Conservation of Biodiversity in Africa. Local Initiatives and Institutional Roles.* Centre for Biodiversity, Nairobi, Kenya, pp. 245–262.

Huaccho, L. and Hijmans, R.J. (1999) A global geo-referenced database of potato production for 1995–1997 (GPOT97). *Production Systems and Natural Resource Management Department Working Paper* 1. CIP, Lima, Peru. www.cipotato.org/data/potato_atlas/gpot97.htm

Humphries, C.J., Williams, P.H. and Vane-Wright, R.I. (1995) Measuring biodiversity value for conservation. *Annual Review of Ecology and Systematics* 26, 93–111.

Hutchinson, C.F. and Weiss, E. (1999) In: Serwinski, J. and Faberova, I. (eds) *Proceedings of the Technical Meeting on the Methodology of the FAO World Information and Early Warning System on Plant Genetic Resources. 21–23 June 1999, Prague, Czech Republic*. Research Institute of Crop Production, Prague, Czech Republic and FAO, Rome, Italy, pp. 19–25.

Hyman, G. (1999) Crop distribution mapping: applications and techniques for broad-scale analysis of crop geography. In: Pande, S., Johansen, C., Lauren, J. and Bantilan, F.T. Jr (eds) *GIS Analysis of Cropping Systems. Proceedings of an International Workshop on Harmonization of Databases for GIS Analysis of Cropping Systems in the Asia Region.* Cornell University, Ithaca, USA and ICRISAT, Patencheru, India, pp. 91–96.

Jones, P.G. and Gladkov, A. (1999) *FloraMap: A Computer Tool for the Distribution of Plants and Other Organisms in the Wild*. CIAT, Cali, Colombia.

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

Justice, C.O., Holben, B.N. and Gwynne, M.D. (1987) Monitoring East African vegetation using AVHRR data. *International Journal of Remote Sensing* 7, 1453–1474.

Kessell, S.R. (1990) An Australian geographic information and modeling system for natural area management. *International Journal of GIS* 4, 333–362.

Kirkpatrick, J.B. (1974) The use of differential systematics in geographic research. *Area* 6, 52–53.

Kremen, C., Razafimahatratra, V., Guillery, R.P., Rakotomalala, J., Weiss, A. and Ratsisompatrarivo, J.-S., (1999) Designing the Masoala National Park in Madagascar based on biological and socio-economic data. *Conservation Biology* 13, 1055–1068.

Kruska, R.L., Perry, B.D. and Reid, R.S. (1995) Recent progress in the development of decision support systems for improved animal health. In: *AfricaGIS'95 Documents. Proceedings of the AfricaGIS'95 Conference, 6–10 March 1995, Abidjan, Ivory Coast*, pp. 524–538.

Ledig, F.T. (1992) Human impacts on genetic diversity in forest ecosystems. *Oikos* 63, 87–108.

Liu, J.G., Dunning, J.B. and Pulliam, H.R. (1995) Potential effects of a forest management plan on Bachman's sparrows (*Aimophila aestivalis*): linking a spatially explicit model with GIS. *Conservation Biology* 9, 62–75.

Maxted, N., van Slageren, M.W. and Rihan, J.R. (1995) Ecogeographic surveys. In: Guarino, L., Ramanatha Rao, V. and Reid, R. (eds) *Collecting Plant Genetic Diversity. Technical Guidelines*. CAB International, Wallingford, UK, pp. 255–285.

Maxted, N., Forde-Lloyd, B.V. and Hawkes, J.G. (1997a) Complementary conservation strategies. In: Maxted, N., Forde-Lloyd, B.V. and Hawkes, J.G. (eds) *Plant Genetic Conservation: The* In Situ *Approach*. Chapman and Hall, London, pp. 15–39.

Maxted, N., Guarino, L. and Dulloo, M.E. (1997b) Management and monitoring. In: Maxted, N., Forde-Lloyd, B.V. and Hawkes, J.G. (eds) *Plant Genetic Conservation: The* In Situ *Approach*. Chapman and Hall, London, pp. 144–159.

Menon, C. and Bawa, K.S. (1997) Applications of geographic information systems, remote sensing, and a landscape ecology approach to biodiversity conservation in the Western Ghats. *Current Science* 73, 134–145.

Miller, R.I. (1986) Predicting rare plant distribution patterns in the southern Appalachians of the south-eastern USA. *Journal of Biogeography* 13, 293–311.

Monestiez, P., Goulard, M. and Charmet, G. (1994) Geostatistics for spatial genetic structures: study of wild populations of perennial ryegrass. *Theoretical and Applied Genetics* 88, 33–41.

Monmonier, M. (1973) Maximum-difference barriers: an alternative numerical regionalization method. *Geogr. Anal.* 3, 245–261.

Nabhan, G.P. (1991) Wild *Phaseolus* ecogeography in the Sierra Madre Occidental, Mexico. *Systematic and Ecogeographic Studies of Crop Gene Pools* No. 5. IBPGR, Rome, Italy.

Nelson A., LeClerc, G. and Grum, M. (1997) The development of an integrated Tcl/Tk and C interface to determine, visualize and interrogate infraspecifc biodiversity. Internal document. GIS Laboratory, CIAT, Cali, Colombia.

Pickersgill, B. (1984) Migrations of chili peppers, *Capsicum* spp., in the Americas. In: Stone, D. (ed.) *Pre-Columbian Plant Migration. Papers of the Peabody Museum of Archaeology and Ethnology*, Vol. 76. Harvard University Press, Boston, Massachusetts, pp. 105–123.

Pigliucci, M. and Barbujani, G. (1991) Geographical pattern of gene frequencies in Italian populations of *Ornithogalum montanum* (Liliaceae). *Genetic Research* 58, 95–104.

Pollak, L.M. and Corbett, J.D. (1993) Using GIS datasets to classify maize-growing regions in Mexico and central America. *Agronomy Journal* 85, 1133–1138.

Pollak, L.M. and Pham, H.N. (1989) Classification of maize testing locations in sub-Saharan Africa by using agroclimatic data. *Maydica* 34, 1133–1138.

Rebelo, A.G. and Sigfried, W.R. (1992) Where should nature reserves be located in the Cape Floristic Region, South Africa? Models for the spatial configuration of a reserve network aimed at maximizing the protection of diversity. *Conservation Biology* 6, 243–252.

Sackville Hamilton, N.R. and Chorlton, K.H. (1997) *Regeneration of Accessions in Seed Collections: A Decision Guide*. Handbook for Genebanks No. 5. IPGRI, Rome, Italy.

Scott, J.M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., Dérchia, F., Edwards, T.C., Ulliman, J. and Wright, R.G. (1993) Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monographs* 123, 1–41.

Sebastini, M., Sambrano, A., Villamizar, A. and Villalba, C. (1989) Cumulative impact and sequential geographical analysis as tools for land use planning. A case study: Laguna La Reina, Miranda state, Venezuela. *Journal of Environmental Management* 29, 237–248.

Silvertown, J.W. and Lovett Doust, J. (1995) *Introduction to Plant Population Biology*. Blackwell Scientific Publications, Oxford, UK.

Skole, D. and Tucker, C. (1993) Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988. *Science* 260, 1905–1910.

Steiner, J.J. and Greene, S.L. (1996) Proposed ecological descriptors and their utility for plant germplasm collections. *Crop Science* 36, 439–451.

Stockwell, D.R.B. and Noble, I.R. (1992) Induction of sets of rules from animal distribution data: a robust and informative method of data analysis. *Math. Comput. Simul.* 33, 385–390.

Templeton, A.R., Shaw, K., Routman, E. and Davis, S.K. (1990) The genetic consequences of habitat fragmentation. *Ann. Missouri Bot. Gard.* 77, 13–27.

Tohme, J., Jones, P., Beebe, S. and Iwanaga, M. (1995) The combined use of agroecological and characterization data to establish the CIAT Phaseolus vulgaris core collection. In: Hodgkin, T., Brown, A.H.D., van Hintum, Th.J.L. and Morales, E.A.V. (eds) *Core Collections of Plant Genetic Resources*. John Wiley & Sons, Chichester, UK, pp. 95–107.

Treweek, J. (1999) *Ecological Impact Assessment*. Blackwell Science, Oxford, UK.

Vane-Wright, R.I., Humphries, C.J. and Williams, P.H. (1991) What to protect? Systematics and the agony of choice. *Biological Conservation* 55, 235–254.

Veitch, N., Web, N.R. and Wyatt, B.K. (1995) The application of geographic information systems and remotely sensed data to the conservation of heathland fragments. *Biological Conservation* 72, 91–97.

Walker, P.A. (1990) Modelling wildlife distributions using a geographic information system: kangaroos in relation to climate. *Journal of Biogeography* 17, 279–289.

Zheng, D., Wallin, D.O., Hao, Z. (1997) Rates and patterns of landscape change between 1972 and 1988 in the Changbai mountain area of China and North Korea. *Landscape Ecology* 12, 241–254.