Diversity and ecology of the potato:

The use of spatial analysis in crop science

Promotoren: Prof. dr. ir. L.O. Fresco
Voormalig hoogleraar plantaardige productiesystemen met bijzondere aandacht voor de tropen, tevens persoonlijk hoogleraar externe betrekkingen

Prof. dr. M.J. Kropff
Hoogleraar in de Gewas- en onkruidecologie

Samenstelling promotiecommissie:

Dr. R.G. van den Berg, Wageningen University
Prof. dr. ir. A.K. Bregt, Wageningen University
Dr. ir. A.J. Haverkort, Plant Research International
Dr. T.S. Walker, International Potato Center

Diversity and ecology of the potato:

The use of spatial analysis in crop science

Robert J. Hijmans

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. ir. L. Speelman
in het openbaar te verdedigen
op 24 September 2002
des namiddags om 13.30 uur in de Aula

Hijmans, Robert J. (2002)

Diversity and ecology of the potato: the use of spatial analysis in crop science.
Ph.D. Thesis. Wageningen University.

ISBN: 90–5809–719–0

Key words: agro-ecological zoning, climate change, crop geography, crop science, diversity, frost, genebanks, impact assessment, GIS, late blight, potato, *Solanum* sect. *petota*, *Solanum tuberosum*, *Phytophthora infestans*, spatial analysis, taxonomy, spatial bias, wild potatoes.

Abstract

Spatial variation is a fundamental characteristic of agriculture, but crop scientists have largely ignored it, particularly at levels beyond the field scale. This thesis provides examples of analysis of spatial variation over larger areas. It contributes to our knowledge of the potato crop and its wild relatives, and to methodological progress in the use of geographic information in crop science. Part I deals with management of genetic resources. Methods that can be used to improve data quality in biological collection databases, and to assess the presence of spatial biases are discussed. The spatial distribution of wild potato species is analyzed, and the value of geographic, ecological, and taxonomic factors to predict the presence of frost tolerance in wild potatoes is studied. Part II deals with aspects of agro-ecological zoning approaches for research management. First, gobal potato distribution is described and analyzed. This is followed by a constraint-specific agro-ecological zoning study for potato late blight; a study on the potential impact of frost resistant potatoes; and an assessment of the effect of climate change on global potato production. In the final chapter, research needs and challenges to the further use of GIS in crop science are discussed.

Table of contents

Preface

Part I – Plant genetic resources

Part II – Agro-ecological zoning

Appendix: Color plates

Preface

The research reported in this thesis was carried out at the International Potato Center (CIP), in Lima, Peru, where I was asked to set up a geographic information systems laboratory. As I had always been fascinated by maps I did not hesitate to accept this task, even though at the time I did not have much experience in that field of work. It soon became clear that while the necessary technical skills were relatively easy to acquire, it was more challenging to come up with applications that went beyond producing pretty maps, and that could contribute to a research program which aims at improving food production in the developing world. This thesis is a reflection of my attempt to contribute to CIP's mission: people don't eat maps, but maps can support research and development for a world without hunger.

This thesis is also an important milestone on the hopefully long, but certainly winding, road of my scientific career. Many people have walked along this road with me, or supported me in other ways, and encouraged me to continue. I would like to thank all of you, colleagues, family and friends, for making this a pleasant and worthwhile journey, despite the occasional potholes. Your support is more valuable than a road map could ever be.

In particular, I would like to thank professors Louise Fresco and Martin Kropff, for their encouraging me to write this thesis. Both were also my teachers during my graduate studies and their skill and enthusiasm for agricultural sciences was a seed of inspiration without which I would not have opted for a career in science. I thank them for sharing their broad knowledge; for being responsive at a long-distance; and for their patience when I got caught up in other activities.

I am indebted to many of my colleagues at CIP. I particularly want to thank Tom Walker for his support and intellectual guidance, and Rebecca Nelson for inspiring discussions. While probably forgetting many others, I also want to thank the following colleagues at CIP: Meredith Bonierbale, Coen Bussink, Roger Cortbaoui, Jorge de la Cruz, Andre Devaux, Greg Forbes, Karen Garrett, Marc Ghislain, Michael Hermann, Luisa Huaccho, Henry Juarez, Maria Lozano, Sylvie Priou, Edwin Rojas, Alberto Salas, Monica Tuesta, Ivonne Valdizan, Zandra Vasquez, Hubert Zandstra, and Dapeng Zhang.

I am obliged to Luigi Guarino (International Plant Genetic Resources Institute) for introducing me to many aspects of plant genetic resources conservation and to David Spooner (USDA, University of Wisconsin) for helping me sharpen my thoughts and polish my skills. I have also worked with a number of colleagues from PROINPA, Bolivia. I thank them all, and particularly Enrique Carrasco and Bruno Condori, Willman Garcia, and Efrain Pozo.

I thank John Stares and Koos Stehouwer for editing.

I would like to acknowledge Jorge Andrade, Ramiro Carillo, Mirjam Jacobs, Rubí Raymundo, Marianne Schreuders, Tineke Troost, and Esther Wiegers, who worked with me as students and made significant contributions to my research program.

During my short visits to the Netherlands I have much appreciated the kind hospitality and support of the people in the Chairgroup of Crop and Weed Ecology of Wageningen University.

Very special thanks go to Mariette Malessy for her steady help in many aspects of the delivery of this thesis.

1. General Introduction

1.1 Introduction

This thesis consists of eight studies of the potato crop and its wild relatives. The first four studies (Part I) deal with the geographic distribution of wild potato species. The main question explored in these chapters is how spatial analysis can be used to improve conservation and use of wild potatoes and other crop genetic resources. In Part II, geographic and agro-ecological aspects of the production of the potato crop are analyzed. Quantitative methods that can be used to study spatial variation in crop production are assessed in four studies, with a view to improved research prioritization.

The aim of these eight studies is twofold: to increase our knowledge of the potato crop and its wild relatives and to contribute to methodological progress in the use of geographic information in crop science. Therefore, the intended audience of this thesis is anyone interested in the geography of the potato, as well as those who need to use analytical and quantitative methods to deal with spatial variability in agriculture.

In this introductory chapter, the contribution of geographers and crop scientists to spatial analysis in agriculture is briefly reviewed. This is followed by an introduction of the two main themes of this thesis: the management of plant genetic resources, and agro-ecological zoning for research management. Finally, an outline of this thesis is given.

1.2 Agricultural geography

The study of spatial variation is the essence of geography. Agriculture is a widespread and highly diverse economic activity and this explains why it has previously received much attention from geographers, perhaps more than any other subject (Kleinpenning, 1968; Gregor, 1970). However, the amount of attention devoted to this subject has declined dramatically over the last few decades, and the study of spatial patterns in agriculture currently receives little attention from geographers (Grigg, 1995).

Geography is the study of “where things are” and “why they are there”, and agricultural geography describes and analyzes spatial variation in agriculture. Early agricultural geographers focused on the role of environmental factors in explaining spatial patterns, but in the second half of the 20th century, other factors were studied in more detail, and environmental influences were neglected and left to the agricultural sciences (Grigg, 1995).

There have been numerous efforts to delimit agricultural geography from its related sciences, including the agricultural sciences (Gregor, 1970). Agricultural science is the study of the production of crops and livestock; it builds on more fundamental biological sciences (such as botany, genetics, physiology, and ecology) but also relates to the social sciences and economics. It is a technical science that is rooted in field and laboratory experimental work and which aims to solve practical production problems.

Agricultural geography, as well as biogeography, has contributed to the work of agricultural scientists. The body of literature associated with these two sciences is the basic stuff for background reading material and for introductory courses. However, only a limited amount of research has been carried out at the interface of agricultural science and geography. Despite the great geographic variation which occurs in agriculture, crop scientists rarely use spatial information and analysis to better address the problems they work on.

1.3 Geographic approaches in crop science

Perhaps the most obvious contribution made by geographers to crop science (or to crop geography) has been the elaboration of crop distribution maps, a deceptively simple endeavor. Outstanding examples include those by Bacon *et al.* (1948) for Europe and the Near East; the global crop distribution maps in the Oxford Economic Atlases (e.g., Economist Intelligence Unit and the Cartographic Department of the Clarendon Press, 1954); and Huke's (1982) rice production maps for Asia. These maps allow appraisal of spread and spatial variation in an area planted with different crops, and give a much more complete picture than the commonly used country level statistics. More complex studies include attempts to characterize agricultural systems rather than single crops (e.g., Duckham and Masefield, 1970; Kostrowicki 1974; Dixon *et al.*, 2001).

There are, of course, more notable examples of geographical approaches to crop science. Karl H.W. Klages, a professor of agronomy, argued for the incorporation of 'crop ecology' and 'ecological crop geography' into the agronomy curriculum (Klages, 1928). In his magnum opus, *Ecological Crop Geography*, Klages (1949) described the geographical distribution of crops, and the social and ecological factors that explain these patterns. The relationship between crop distribution and ecological factors has been described in great detail by Wilsie (1962) and there also have been many number of studies of the distribution of plant diseases (e.g., Weltzien, 1972; Zadoks and Rijsdijk, 1984).

Another important exception to the overall picture of dissociation of crop science and geography is the work by Nikolai I. Vavilov, arguably the foremost crop geographer of the 20th century. Vavilov led the first systematic and long-term genetic resource collecting effort. Based on his numerous and extensive expeditions, and subsequent field trials in Russia, he developed his theory of the centers of origin of cultivated plants (Vavilov, 1926). In later publications, he further improved and refined his theories (see Hawkes, 1999, for an overview). Although some of Vavilov's propositions are no longer accepted in their details, he advanced the field for many later studies of the origin and spread of crop plants, as De Candolle (1882), who first provided comprehensive studies of the origin of crops, had done before him. Later studies have often focused on the spread of crop plants, from their centers of origin to other regions (e.g., Bertin *et al.,* 1971; Simmonds, 1976; Sauer, 1993). Increasingly, such studies use biotechnological tools. For example, Heun *et al.* (1998) identified the site of einkorn wheat domestication using DNA fingerprinting; Hanotte *et al.* (2002) used microsatelittes to show independent domestication of cattle in Africa; and Zhang *et al.* (2002) used molecular markers in an attempt to trace the prehistoric introduction of sweetpotato from the Americas to Polynesia.

In summary, fundamental questions of crop geography are:

- Where did a crop originate, and where is it most diverse?
- Where is it currently grown and why?
- What are current levels of productivity and production constraints?

Today, with the advent of geographic information systems, these old questions can be addressed in many new ways.

1.4 Geographic information systems

The development of geographic information systems (GIS) software has spurred a renewal of the interaction between geography and other sciences, including crop science. GIS makes spatial analysis much more efficient, as huge amounts of data can be processed quickly, and accurately mapped, without the need for specialist cartographers. More importantly, GIS can make research more effective, as it allows new types of investigations. The capacity to combine many sources of geo-referenced information and for the quantitative treatment of these data (spatial modeling) allows complex analysis that was hardly possible previously.

In crop science, the use of GIS is mainly related to precision agriculture (White *et al.*, 2002). This reflects the typical scale of analysis used by crop scientists: the individual cultivated field. Nevertheless, studies of larger areas are also emerging. For example, Carter *et al.* (1992) used a geographic framework in which socio-economic and agro-ecological data described the spatial pattern of cassava production in Africa. Jones *et al.* (1997) used genebank databases in conjunction with climate surfaces to identify areas in South America where the wild common bean (*Phaseolus vulgaris)* might occur, but had not been recorded. There is also a growing body of literature in the field of agro-ecological zoning (e.g., FAO, 1978-81; Booth *et al.*, 1989; Duchateau *et al.*, 1997), and disease management (e.g., Nelson *et al.*, 1999).

This thesis is a contribution to GIS use in studies beyond the field level. It consists of a number of case studies exploring the interface of crop science and geography. Most of the studies in this thesis were carried out with the intention of providing information to allow more effective decision making in terms of research management at the International Potato Center (CIP). The objectives of CIP include the conservation of potato genetic resources, and improving potato production technology in developing countries.

1.5 Conservation and use of crop genetic resources

For millennia, farmers have selected those plants that best fit their needs, initially from wild populations, and later increasingly from cultivated populations (land races). From this process originated the diverse groups of crop genotypes that are both an important part of current agricultural technology and a resource for the development of new varieties. Use of genetic diversity allows us to meet the need for a continuous flow of new varieties that can increase yields, quality, and biotic and environmental tolerances in crops.

Unfortunately, the adoption of new “modern” varieties may lead to the loss of the very raw material on which they depend—i.e., genetic diversity. The replacement of a diverse array of land races by a limited number of new cultivars has led to a decrease in on-farm genetic diversity, or genetic erosion (Harlan, 1992). Man’s increasing influence on the biosphere has also led to a decrease in the genetic variability of the wild relatives of crop plants. Genetic erosion of crops, and of their wild relatives, decreases the potential for the adaptation of agriculture to changing circumstances, at either the farm level or through formal breeding programs. Although evidence shows that genetic erosion is not always as serious as is often suggested (Brush, 1995), it is important to locate, quantify, and monitor biological diversity, and actively conserve it where necessary.

This process entails many steps, from eco-geographic surveying before a collecting mission, to predicting the presence of useful traits in collections, to monitoring genetic erosion. For all of these steps, spatial variability needs to be considered. Although spatial analysis can clearly play a role in making plant genetic resource conservation and use more effective and efficient, there are only a few examples of studies where spatial analysis has been explicitly used (Guarino *et al.*, 2002). The first part of this thesis contributes to filling this important gap.

The studies in Part I are all on wild potato (*Solanum* sect. *Petota*), the wild relatives of the cultivated potato. According to the most recent comprehensive review, there are 199 wild potato species (Spooner and Hijmans, 2001). They occur over a large and ecologically diverse geographic area, from Colorado (USA) to Uruguay and Chile (Hawkes, 1990). They are used in breeding programs for disease resistance, environmental tolerance, and other agronomic traits of interest (Ross, 1986; Jansky, 2000). This combination of geographic, ecological, and taxonomic diversity, with their use in breeding programs, makes wild potatoes a model system for spatial analysis of the genetic diversity of a crop.

1.6 Crop production

Over the coming decades, demand for food is expected to increase sharply. Food production will have to rise substantially in order to avoid food scarcity, high food prices, and to alleviate hunger. Much of the land most suitable for agriculture is already in use, or is occupied by the last remaining patches of nature that should be conserved in their natural state. Therefore, further increases in food production should come, primarily, from increased productivity (production per unit land area, and per unit time). Increased productivity can partially be achieved by the use of more inputs, such as labor, water, fertilizers and pesticides. Yet there is widespread concern about the negative effects of intensive agricultural practices on biodiversity and human health, as well as about the sustainability of such agricultural production systems. Future agricultural production systems should be more productive, economically efficient, produce fewer unwanted side effects, and be ecologically sustainable. They should also be stable and resilient enough to cope with perceived future climatic instability, climate change, and changes in the biotic environment (e.g., pathogens). Hence, there is an obvious, but challenging, need for the development of improved agricultural technology.

Despite all these needs and concerns, support for international agricultural research is currently declining. This has sharpened the need to increase research efficiency, and to systematically evaluate the impacts of past research and the prospective benefits of proposed research (Alston et al., 1995; Wood and Pardey, 1997). However, impact evaluation and related activities are complicated by the site-specificity of agricultural production, and of the adoption and effects of agricultural technologies. A variety may be very useful in one location but not be adapted to another, because of differences in, for example, the weather, soil types, cropping systems, pests and diseases, farmer assets, and market demand.

Land evaluation (FAO, 1976), agro-ecological zoning (FAO, 1978-81), and related approaches may be used to stratify the world into homogenous agricultural research and development domains. Such approaches have been used for various purposes, including setting priorities for research (Gryseels *et al*., 1992); exploring crop growth potential (Van Lanen *et al*., 1991); studying the incidence of a livestock disease (Duchateau *et al.,* 1997); and predicting areas suitable for specific tree species (Booth *et al*., 1989). However, a number of conceptual and methodological problems associated with the delimitation of agro-ecological zones remain (Wood and Pardey, 1997).

The studies in the second part of this thesis contribute to an improved use of agro-ecological zoning by refining existing, and by developing new methods. These methods are quantitative and specific, have a relatively high resolution, and take the current area of crop distribution into account. All studies are on the potato crop. Potato is a globally important food crop. It is also one of the most widespread of crops, and is therefore a suitable subject for these studies.

Potato can produce a large quantity of food in a short period of time. Under ideal conditions, a potato crop can yield more than 15 t/ha dry matter (Stol *et al*., 1991), much higher than for cereals, and estimated average dry matter yield is currently at 10 t/ha in New Zealand and at 9 t/ha in the Netherlands (FAO, 2002). However, potato yields are often severely constrained by abiotic stress and by pests and diseases. This fact is reflected in the chapters of this thesis that deal with different production constraints, including frost and heat tolerance, and potato late blight.

1.7 Outline of this thesis

Part I

Genebank databases are the principal source of information on the distribution of crop genetic diversity. However, coordinate data in genebank databases are notorious for their scarcity and, in the case of those records that do have coordinate data, for occasionally being inaccurate. This makes spatial analysis of genebank data more complicated, and the results more uncertain. In Chapter 2, a protocol is described to check and improve the data quality of geographical coordinates in genebank databases.

Genebanks were not intended primarily as a source of geo-referenced information about a crop, but rather as a way to conserve as much diversity as possible. Therefore, the use of these data can be problematic and issues of information quality need to be considered. Due to spatial differences in recorder effort (Rich and Woodruff, 1992; Prendergast *et al.*, 1993b; Gaston, 1996), data on plant distributions are typically biased. In extreme cases, differences in the number of species between areas might reflect the amount of time spent there by recorders, and not the actual differences in distribution. In Chapter 3, indicators are proposed and used to assess the information quality of genebank data on wild potatoes in Bolivia.

In Chapter 4, a large database of collecting localities is analyzed to provide a comprehensive and quantitative description of the distribution of the 199 currently accepted wild potato species. Range sizes and abundance are discussed, and areas of high and complementary diversity are identified. This study aims to be a general reference for understanding the distribution of genetic diversity in this group of plants, which may be used for future collecting and conservation efforts.

Conservation and collection of wild potato species has been undertaken mainly on a taxonomic basis. For example, Spooner *et al.* (1999) justify their collecting expedition on taxonomic grounds. From a crop improvement perspective, however, the presence of useful agronomic traits, such as disease resistance, is more important then taxonomic diversity per se. Yet, the use of taxonomic diversity could be justified if it were a proxy for trait diversity, because there are many useful traits and often they cannot be observed when collecting. An alternative approach might entail the use of geographical and ecological information to stratify and sample wild potato populations (e.g., Peeters *et al.*, 1995). The relative merits of these different factors for the purposes of prediction have not previously been put to the test simultaneously—this is done in Chapter 5, using a case study on frost tolerance in wild potatoes.

Part II

The studies in the second part of this thesis contribute to improved use of agro-ecological zoning (AEZ) for research planning. An important limitation for many AEZ studies is that the results are difficult to interpret and use, because they are not related to the current area of the distribution of the crop. This may be partly due because of a lack of detailed crop distribution data. To fill this gap for potato, a global potato area distribution database was developed. Potato distribution is discussed in relation to other variables (latitude, altitude, population) in Chapter 6. The potato distribution data are used in subsequent chapters to interpret the results of AEZ studies.

Potato late blight is one of the most devastating crop diseases—as was dramatically illustrated by the Irish Famine of the 1840s. As the severity of late blight is highly weather dependent, it lends itself to agro-ecological zoning. In Chapter 7, global climate databases and disease forecast models are used to develop quantitative and constraint-specific AEZ for potato late blight. The main purpose of the study is to assist in targeting technology (resistant varieties and improved fungicide use) for late blight control in developing countries.

A new method of using AEZ approaches for estimating the potential impact of proposed research is described in Chapter 8. In this chapter, the potential impact of frost tolerant potato varieties on the Altiplano of Peru and Bolivia is quantified, using a modified potato growth simulation model and high-resolution interpolated climate data in a GIS. To avoid information loss, zoning, in the strict sense, is not used, as each small grid cell is treated individually (i.e., without aggregation prior to the calculations).

Finally, in Chapter 9, the effect of climate change on global potato production is quantified using current and predicted climate data and a crop growth simulation model. Whereas most studies of this type analyze a limited number of locations, this study is truly global. Particular emphasis is given to the potential to adapt to climate change by shifting the growing season or by using varieties with different maturity type. The goal of this study is to identify areas where the impact of climate change is likely to be strongly negative, and to identify the areas upon which research and development could focus through the development of heat-tolerant varieties.

Part I

Plant genetic resources

2. Improving coordinates in biological collection databases

Published as: Hijmans, R.J., M. Schreuder, J. de la Cruz, and L. Guarino, 1999. Using GIS to check co-ordinates of germplasm accessions. Genetic Resources and Crop Evolution 46: 291-296.

Abstract

The geographic coordinates of the locations where germplasm accessions have been collected are usually documented in genebank databases. However, the coordinate data are often incomplete and may contain errors. This paper describes procedures to check for errors, to determine the cause of these errors and to assign new coordinates, using Geographical Information Systems (GIS). These procedures can assist in improving the quality of genebank databases, and with that, increase the capability for analysis and use of crop genetic diversity.

2.1 Introduction

Information on the accessions (i.e., entries, genotypes) conserved in genebanks is usually documented in a database. The completeness and the quality of such databases are important determinants of the usefulness of the germplasm collection to which they refer. A genebank's database may include passport, characterisation and evaluation data. Passport data include a description of the location where the accessions were collected. The location is usually specified by the country and at least one administrative subdivision (e.g., state or department), and by a description of the locality where the accession was found. Often, the location is also specified with geographic coordinates (latitude and longitude).

When location data are in coordinate form, they can be used in a Geographical Information System (GIS). With a GIS, a number of analyses can be carried out that are of importance for managing and using the germplasm collection, and in planning further collecting. For example, Jones *et al.* (1997) used genebank databases together with climate surfaces to identify areas where the wild common bean *Phaseolus vulgaris* might appear, but had not actually been recorded. Other activities in which the analysis of geo-referenced genebank data can make a considerable contribution include the investigation of the taxonomic structure of collections (Jones *et al.,* 1997); the identification of areas of high diversity (Nabhan 1990; Frankel *et al.,* 1995; Hijmans 1997); the targeting of genetic resources for breeding programs (Nabhan 1995; Guarino *et al.,* 1998); the development of core collections (Guarino *et al.,* 1998); and the selection and design of sites for *in situ* conservation (Guarino *et al.,* 1998).

Unfortunately, in many cases the coordinates in the databases are (wholly or partly) missing, imprecise or wrong. As curators of genebanks strive to improve their databases, they face the task of completing and correcting the coordinate data (Hazekamp and Frese, 1992). The present paper intends to assist in that activity by describing how the coordinates of genebank accessions can be checked and improved, using GIS. We do not discuss the use of GIS in the further analysis of genebank data. The procedures described in this paper are, however, the first steps that are needed to undertake such analyses.

2.2 GIS

A GIS is a computer-based tool for managing geographical referenced databases and analysing spatial relationships. There are many different GIS software packages available. In this article we include, as an example, some commands from the widely used Arc/Info and ArcView software (trademarks of ESRI; http://www.esri.com).

To be able to use genebank data in a GIS, the data need to be in a format that the GIS supports. This is usually done by creating a text file with three numbers on each line: a unique identifier, longitude and latitude. Longitude and latitude should be in decimal degrees. In many genebank databases longitude and latitude are given in degrees (º), minutes ('), in some cases also in seconds (''), together with a hemisphere (North or South, and East or West). In a GIS, coordinates in decimal degrees are needed. Conversion to coordinates in decimal degrees is done with the following formula:

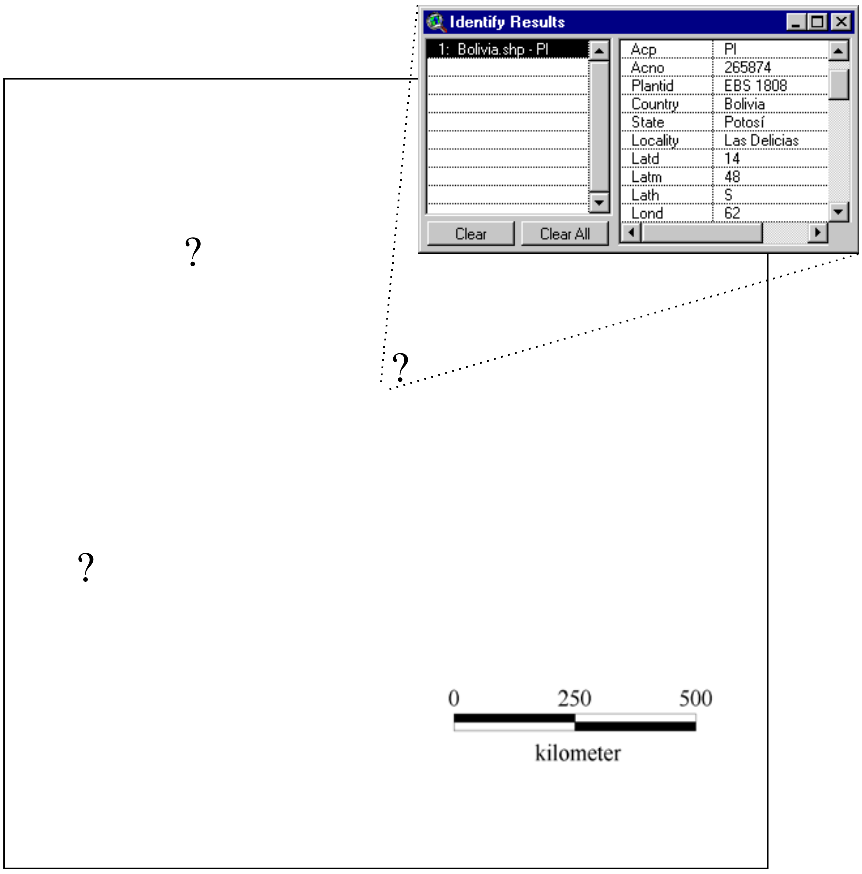
$$d^{o}\ m'\ s'' = h \times (d + m / 60 + s / 3600)$$

Where h = 1 for the Northern and Eastern hemispheres and -1 for the Southern and Western hemispheres (e.g., 30º 30' 0'' S = –30.5 and 30º 15' 55'' N = 30.265).

The unique identifier could be the collection number, but because this usually is a combination of alphanumeric and numeric characters, such as "SVGU 6505", this can easily lead to errors. It may, therefore, be better to use a unique *numeric* identifier that is related to the collection number. The text file is then imported into the GIS ("Generate" in Arc/Info). With a relational database operation, using the unique identifier, the points can be linked to the passport and other data from the original database.

2.3 Finding errors

Errors can be spotted by plotting the collection sites on a map with administrative boundaries. This can lead to the detection of *impossible* locations, that include, e.g., accessions located in an ocean or a lake, and *unlikely* locations, that are far away from all other accessions. In both cases, one should access the database to see if the suspicious locations are really wrong. Perhaps an accession in the ocean was actually found on a small island, that is not shown on the map, or, the accession may truly be isolated for other reasons, such as lack of collection efforts in that area.



Acp	PI
Acno	265874
Plantid	EBS 1808
Country	Bolivia
State	Potosí
Locality	Las Delicias
Latd	14
Latm	48
Lath	S
Lond	62

Figure 2.1 Bolivia, its departments, and locations where wild potatoes were collected (black dots). Probable errors are indicated with a question mark. The ArcView 'Identify Results' window is shown for one of the probable errors.

This visual inspection method is illustrated in Fig. 2.1, using data on wild potato species of Bolivia from the Intergenebank Potato Database (Huamán *et al*., 1996). Three suspiciously isolated locations are indicated with a question mark. For one of the locations, the "Identify Results" window is shown that appears after clicking on the point, using ArcView. On the map, the point is located in the Department of Santa Cruz, according to the coordinates in the genebank database. However, the locality description in the same database indicates that it should be located in "Las Delicias", in the Department of Potosí. The coordinates in the database, and thus the location on the map, are therefore likely to be wrong. The visual inspection method only works for the grossest of errors. These errors and other less conspicuous errors can be identified using the capabilities of a GIS to a greater extent.

By simultaneously querying the accessions database and the administrative boundaries database, a new database can be created ('overlay' analysis in GIS language; Arc/Info: "identity"). For each accession, the 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 reflect errors. This is illustrated in a simple example for an imaginary island that has three provinces, called A, B and C. Six accessions have been collected and stored in a genebank. The coordinates of the collection locations, according to the genebank database, have been plotted on a map of the provinces (Fig. 2.2). By querying the two databases, Table 2.1 is generated, pointing at accessions 2, 4 and 6 as possible errors.

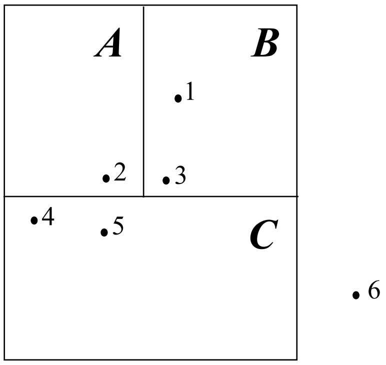


Figure 2.2 Imaginary island with three provinces, A, B, and C and the location of 6 germplasm accessions.

Table 2.1 The location of six germplasm accessions according to the genebank database and according to the administrative boundaries database. Discrepancies between the two databases (shaded entries) point at likely errors.

Genebank database		Administrative Boundaries
Accession	Province	Province
1	B	B
2	B	A
3	B	B
4	E	C
5	C	C
6	C	

2.4 Determining the cause of errors

It is easier to spot errors than to determine their causes. The causes of errors include incorrect reading of maps, sometimes caused by duplicate location names and confusion about the coordinate system. Typographical mistakes are perhaps the most common cause of errors in the database. Thus, 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.

In general, it is more likely that coordinates are wrong than that the name is wrong. Coordinates are often derived from the names in the first place. Moreover, because coordinates are, unlike geographical names, rather abstract entities to most people, it is more likely that errors be made in assigning and digitising them. Some common errors of this type are the switching of latitude and longitude, typing the wrong hemisphere or typing the last 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 differences in spelling or by typographical mistakes. These errors are easy to trace and correct by inspecting alphabetical lists of the names in the databases. Such errors are particularly common if names have been transliterated from other scripts (e.g., Arabic, Chinese or Cyrillic).

It is not always immediately obvious why names do not correspond. Such is the case, for example, when a collection is made near a border of an administrative unit. Apparent errors can be due to lack of precision in administrative boundaries maps, or because explorers did not exactly know in which administrative area they were when they made the collection. In such cases, the coordinates may be correct, and the discrepancies may be due to a wrong name. Another common cause of discrepancies in the records is the change in names of administrative units, or the creation of new ones.

When there is doubt about the location of an accession, it is useful to reconstruct the expedition's itinerary. As collection numbers are assigned sequentially, this may help to determine where an accession was collected. It is likely that accession *z* was collected in between the locations of *z-1* and *z+1*. This is not always the case, however, as collectors may travel up-and-down a road, so this rule should not be applied blindly.

One can also plot sub-groups of the database, and then compare the doubtful location of an accession in question with the location of the other accessions of the same taxon. For example, if all accessions of a species were found in the Bolivian Amazon, and a doubtful location is in the desert coast of Peru, the doubtful location is likely to be wrong. However, one should be very cautious when applying this procedure. To allow future interpretation of the genebank data, one should avoid downgrading the database by creating artificially reinforced spatial relationships. The exceptions to the general spatial patterns should not be changed/removed just because they are exceptions, but only when they are clearly wrong.

Additional variables, like altitude or vegetation, can also be used to verify the coordinate data. For example, when an overlay is made of the collection locations and an altitude map, a new database is generated, analogous to what we showed for administrative boundaries. If there is a big difference between the altitude according to the map and according to the collectors, the coordinates may be wrong. However, one should bear in mind that the altitude in the genebank might be wrong too, because of typographical errors or because altitude is often estimated without using proper instruments. The US Geological Survey has published a high resolution (approximately 1 km^2) altitude grid for the whole world that is available on CD-ROM and on the Internet (http://edcdaac.usgs.gov/gtopo30/gtopo30.html). This grid can be used to accurately verify the altitude of a given location. Problems with using vegetation maps to check coordinates include the fact that many genebank databases do not have much information on vegetation; that the use of different classifications systems complicate comparison; that patchy or mosaic-like vegetation may be missed on small scale maps; and that vegetation patterns change (Hall, 1994).

2.5 Assigning coordinates

If the coordinates of an accession are wrong, or absent, new coordinates need to be assigned where possible. Many old accessions generally have very little location data. In these cases, one cannot do much. If a record has a description of the locality where the accession was found, coordinates can be determined, using maps. The precision with which this can be done depends on the scale of the maps and on the locality description. Locality descriptions are sometimes very detailed: e.g., “Franz Tamayo, 10 m west (towards Pelechuco) of bridge crossing over Río Chullumuyo, on horse trail from Pelechuco to Mojos, ca 6 km east of Quiara, 1:250,000-scale map SD 19-10”; but others are very short, like “near Mojo” or “about 20 km from Cochabamba”. It is not clear what 20 km from Cochabamba means. In what direction? 20 km from the centre, or from the outskirts of town (and in what year)?

For reasons of precision, large scale (high resolution) maps, i.e., of 1:100,000 and larger, should be used where available. Especially in mountainous areas, it can be difficult to estimate distance on a map, because of the winding roads. In some cases, more precise descriptions can still be found in the field books or expedition reports of the explorers, rather than on the collecting forms and the databases that were derived from them.

Searching for names on maps can be time-consuming. Gazetteers, or lists of geographic names and their coordinates, make searching quicker and easier. There is no comprehensive world gazetteer yet available, but the Times' Atlas of the World (Times Books, 1988) contains an extensive one. The US Board of Geographical Names constructs the Official Standard Names Gazetteer that is available in country volumes. Another reference, containing more than 3 million names from all over the globe, is available on the World Wide Web (http://164.214.2.59/gns/html/index.html). Herbaria sometimes develop their own unpublished gazetteers, perhaps in a card catalogue or computer database. Details of the localities mentioned in standard Floras are sometimes published in separate volumes or appendices (Hall, 1994). Historical maps, atlases and gazetteers, and even travel books, can be useful sources of localities if names or boundaries have changed (Maxted *et al.*, 1995).

If new coordinates are assigned, the coordinate checking procedures described in the previous paragraphs should be applied again. Changes made in the database should be documented so that others may understand the reason for any change that was made. This would be especially useful in the event that any new error is introduced.

2.6 A case study

The importance of checking and assigning coordinates with GIS is illustrated by a data set from a case study on wild potatoes from Bolivia. Because the database consists of records from 18 expeditions from a period of more than 40 years, many errors could be expected. Applying the procedures described above, we found that more than 50% of the accessions had an error of one kind or another. By carefully studying the sources of the errors, and the location descriptions, most of the errors could be corrected (Table 2). Even recent data, collected with a Global Positioning System, contained some errors, due to typographical mistakes.

Table 2.2 Initial number of records, the number of errors per category (one accession may have more than one error), and the final number of accessions with acceptable coordinates, after applying our methods, for a database of wild potato germplasm from Bolivia.

Number of Records	**Number**	**Percentage**
Total	1420	100
Wrong names (province and/or department)	344	24
Wrong coordinates	202	14
Missing coordinates	483	34
Final number with coordinates (after corrections)	1039	73

2.7 Conclusions

We have described methods to verify coordinates of germplasm accessions and to assign new coordinates where they are absent or when errors are detected. These are important steps in improving the quality of a genebank database, and with that, the usefulness of the germplasm collection.

Using GIS, three kinds of errors can be detected:

- Accessions in impossible places, like oceans
- Accessions in unlikely places, e.g., widely separated from all others, or, at an unlikely altitude
- Accessions in the wrong place according to passport data.

The first two kinds of errors are due to wrong coordinates and can be detected by *visual inspection* of the data. The third type of error is typically due to wrong coordinates and/or wrong names and can be detected using *overlay analysis* methods.

The different kind of errors and the difficulties in detecting and correcting them, highlight the importance of more precise bookkeeping by germplasm collectors. A detailed and unequivocal locality description is crucial. The availability of Global Positioning System (GPS) greatly facilitates taking geographical coordinates.

However, in our case study, even the data of the accessions that were collected with a GPS had errors, both in the geographical names and in the coordinates, caused by typographical mistakes.

Checking and improving the coordinates of a germplasm database is tedious and time consuming. And even after applying the procedures described here, the database will likely still contain errors. However, given the dramatic increase of the data quality that is feasible, as shown by the data from the case study, it seems that the effort would be justified for many genebanks.

3. Assessing spatial bias in biological collection data

Published as: Hijmans, R.J., K.A. Garrett, Z. Huamán, D.P. Zhang, M. Schreuder and M. Bonierbale, 2000. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. Conservation Biology 14: 1755-1765.

Abstract

Genebank collection databases can be used for eco-geographical studies, under the assumption that the accessions are a geographically unbiased sample. We evaluated the representativeness of a collection of wild potatoes from Bolivia, and defined and assessed four types of bias: species, species-area, hotspot, and infrastructure bias. Species bias is the sampling of some species more often than others. Species-area bias is a sampling that is disproportionate to the total area in which a species is found. Hotspot bias is the disproportionate sampling of areas with high levels of diversity. Infrastructure bias is the disproportionate sampling of areas near roads and towns. Each of these biases is present in the Bolivian wild potato collection. The infrastructure bias was strong: 60% of all wild potato accessions were collected within 2 km of a road, as opposed to 22%, if collections had been made randomly. This analysis can serve as a guide for future collecting trips. It can also provide baseline information for the application of genebank data in GIS-based studies.

3.1 Introduction

The collection of plant genetic resources aims at providing access to the greatest possible amount of genetic variability in the species of interest and helps reveal the ecological and geographic distribution of plant species (Bennett, 1970). A genebank is a collection of a particular crop and its wild relatives and, ideally, includes at least one example of each alternative allele for each locus (Chapman, 1984). Since researchers can only hope to sample a fraction of the genetic variation that occurs in nature, it is important that this sample be as large as possible and contain the maximum amount of useful variation for both present and future use (Allard, 1970; Brown and Marshall, 1995).

The role of genebanks for the conservation and use of germplasm for crop improvement is well established. More recently, genebanks have been used as a source of data on the ecological and geographic distribution of species. For example, Jones *et al.* (1997) used genebank databases together with climate surfaces to identify areas where the wild common bean (*Phaseolus vulgaris*) might occur but had not actually been recorded. Analysis of geo-referenced genebank data also proves useful in identifying areas of high diversity (Nabhan, 1990; Frankel *et al.*, 1995); targeting genetic resources for breeding programs (Nabhan 1990; Guarino *et al.,* 1999); and selecting and designing sites for in situ conservation (Guarino *et al.*, 1999). However, the scientific value of these alternative uses of genebank data is questionable if the database of localities used for geographical analysis is small, skewed, or otherwise unrepresentative of the total natural distribution (Nabhan, 1990).

Diversity is not uniformly distributed in space or among taxonomic groups. Some genetic heterogeneity can be attributed to selection and gene flow and reflects eco-geographic adaptation, whereas other heterogeneity is due to factors such as founder effects and genetic drift (Guarino, 1995). Although ecological factors are a major determinant of genetic diversity (Bennett, 1970), the amount of genetic variation that can be attributed to eco-geographic factors rarely exceeds 50% (Chapman, 1989). As the distribution of diversity is not known prior to data analysis, successful collecting in terms of diversity may depend on the proper identification of populations closely adapted to specific environments and land-use patterns (Brown and Marshall, 1995). Therefore, it is often suggested that collectors take a stratified sample, dividing the population into as many ecological and geographical subpopulations as possible and drawing a sample from each one (Cochran, 1977). In doing so, it is important to include extremes of the range of a species (Allard, 1970). Although these subpopulations may not present great genotypic variation, they may harbor unique traits or taxa (Von Bothmer and Seberg, 1995). At present, the optimal scales for accurate sampling of and genetic diversity have not been determined. Brown and Marshall (1995) suggest sampling 50 plants from each of 50 sampling populations in each region, but they do not clearly define the size of those regions.

A genebank collection may deviate from the ideal sample for a number of reasons. Practical constraints include limits on accessibility, time, and funding. In addition, the collection may be skewed if the germplasm was collected with a specific objective such as rescuing material in risk of extinction, identifying a specific trait, or fulfilling a specific research need (Engels *et al.,* 1995).

Identification of optimal accession sample size for a representative genebank collection is an important issue, but will not be covered in this paper. Instead, we will analyze how genebank collections can systematically deviate from a representative sample in a relative sense, i.e., whether some species and/or areas have been sampled more intensively than others. Whether or not a sample may be considered representative depends partly on how representative sampling has been defined. Samples may be thought of as drawn from a pool of individual organisms or from a pool of geographic areas (Hayek and Buzas, 1997). If drawn from a pool of individual organisms, a random sample would tend to produce collections with proportionately more accessions from the more abundant species or common varieties. Whether or not that is desirable depends on the specific research goals. To assess whether a genebank collection is representative in the relative sense, we have defined four types of bias. Three of these biases are spatial: species-area, hotspot, and infrastructure bias. The fourth, species bias, is not. The presence of these four types of bias was studied for a collection of wild potatoes from Bolivia.

Species bias

Species bias refers to sampling more from one species than from another. Species bias might result from differences in the probability of finding one species relative to that of finding other species because of differences in abundance or from collectors' preferences for a particular species due to a specific expected trait or a specific research need. In terms of genetic resource applications, species bias could be justified if, for example, some species were considered more useful than others because they contained more genetic diversity or desirable agricultural traits.

Species-area bias

Species-area bias refers to over- or under-sampling a species in relation to the size of the area in which it occurs. Theoretically, in order to maximize genebank diversity the number of accessions collected per species should increase with the size of the area in which it is found, assuming a larger distribution area implies a proportionate increase in intraspecific diversity. This is true, in general, although biological diversity is more directly dependent on ecological diversity and environmental stress factors (Nevo, 1998). The increase in diversity with area tends to be logarithmic (cf. Rosenzweig, 1995; Magurran, 1988).

Therefore, a species-area bias exists if a collection contains too many accessions from some species, and hence too few from others, relative to the size of the areas in which the species appear. No species-area bias exists if each species is represented proportionately according to its geographic range. Unless all species distribution areas are the same size, there will be a species bias, a species-area bias, or both.

Hotspot bias

Hotspot bias refers to excessive or insufficient collection in certain geographic areas according to their level of genetic diversity. This is likely to occur if a genebank contains material from different expeditions. To avoid this bias, collection expeditions should cover complementary areas. Since collectors often seek particular species, they may to replicate the parts of previous expedition routes where those species have been found, and to ensure that they obtain a sufficient amount of germplasm. Although it is important to consult genebank and herbarium databases and the route maps of previous collectors to design a collecting strategy (Brown and Marshall, 1995; Nabhan, 1990), this might lead to redundancy if collectors only aim for areas where previous studies and collections indicate that high species diversity exists. Consequently, some areas may be explored more intensively than others, thus originating a hotspot bias. The collection would have too many accessions from certain highly diverse areas, in a relative sense, despite the potentially high level of species diversity.

Infrastructure bias

Infrastructure bias refers to oversampling near roads and towns. For reasons of efficiency, logistics, and convenience, collectors may tend to follow the roads that connect the main towns. Hermann (1988) showed that most Andean tuber crop collection sites in Ecuador were located along the Pan-American Highway and other major roads. Based on herbarium records, Von Bothmer and Seberg (1995) produced a distribution map for *Elymus cordilleranus* in South America. The map shows distinct aggregation around La Paz (Bolivia), Lima (Peru) and the Pan-American Highway in

Ecuador. This does not reflect species distribution very accurately (Von Bothmer and Seberg, 1995). Øllgaard (1995) analyzed geographic distribution of *Huperzia* based on collections considered representative and of sufficient quality for determining priorities in selecting biodiversity reserves. He found that these collections showed a strong tendency to be located near roads and special collection sites.

3.2 Methods

Wild potato genebank data

The wild tuber-bearing *Solanum* species (Solanaceae sect. *Petota*) are relatives of the cultivated potato (*Solanum tuberosum* and six other species [Hawkes, 1990]). They appear in the Americas from Colorado (USA) to Argentina and Chile and are most abundant in the Andes of Peru and Bolivia (Hawkes, 1990). Wild potatoes have been used in modern breeding programs to improve the cultivated potato since the early 1900s, when Mexican *S. demissum* was used to breed for resistance against *Phytophthora infestans*, the fungus that causes potato late blight (Ross, 1986). The present study was restricted to data for wild potatoes collected in Bolivia. An overview as well as in-depth information about wild potato species from Bolivia can be found in Hawkes and Hjerting (1989), Hawkes (1990), and Ochoa (1990). A recent wild-potato collecting expedition in Bolivia is reported by Spooner *et al.* (1994).

Our data source was the Intergenebank Potato Database (IPD). The IPD contains passport and evaluation data from the world's six main wild potato genebanks. Passport data include location variables such as the geographic coordinates and the locality, province, and department of origin for each accession. Each accession represents a number of potato seeds stored in one or more of the genebanks. The IPD currently has over 7128 records, of which 1440 are from Bolivia. The Bolivian data was drawn from 18 different expeditions conducted from 1953 to1994.

GIS

Each accession has passport data that include geographic coordinates. These were used to incorporate the IPD in a GIS, using IDRISI, Arc/Info, and ArcView software. This yielded a digital map of collection sites consisting of points with the IPD's records as attributes. The locations of these points were checked following procedures described by Hijmans *et al.* (1999). This included the production of an overlay of the collection site map and an administrative boundary map of Bolivian departments and provinces, to allow for comparison of the province and department names from both sources. If coordinates were wrong or missing, new coordinates were assigned when possible based on the IPD locality description. After applying these procedures, 1051 out of 1440 Bolivian accessions in the database had confirmed geographic coordinates (Fig. 3.1) and could thus be used in our study.

Species bias

There is an ongoing debate about species boundaries in wild potatoes (Spooner and Castillo 1997; Van den Berg *et al.,* 1998). We have treated all species, subspecies, and varieties included in the database as separate taxa, according to the identification by collectors and genebank taxonomists. Using this criterion, there are 45 wild potato taxa (hereafter referred to as species) collected in Bolivia, conserved in ex situ genebanks, and documented in the IPD. The number of accessions per species, A_S, was tabulated. The presence of a species bias was evaluated with a χ^2-test, comparing the observed number of accessions per species to the number per species that would be observed if accessions were equally allocated between all species.

Species-area bias

Bolivia was divided into 100 × 100 km grid cells (Fig. 3.1), and for each species, S, the number of grid cells in which it was collected, G_S, was determined. If there is no species-area bias, there should be a high degree of correlation between A_S and G_S. The relation between A_S and G_S was described using linear and logarithmic regression. A χ^2-test was used to test for systematic deviations from these two expected relationships, comparing the observed number of accessions per species to the number per species that would be observed if accessions were equally allocated to each species based on the number of grid cells in which it occurs.

Hotspot bias

To examine the presence of a hotspot bias, we divided Bolivia into 100 × 100 km grid cells (Fig. 3.1), as for the species-area bias. For each grid cell, G, we determined the number of accessions that were collected, A_G, the number of species that were collected, S_G, and the number of expeditions that collected a sample (accession), E_G. The relation between A_G, S_G, A_G/S_G and E_G was determined by linear regression. We define a hotspot bias to be present if there is a significant (and positive) correlation between A_G/S_G and E_G. The number of expeditions per area was determined based on published route maps (e.g., Spooner *et al*., 1994) or inferred by plotting collection sites on a road map. Large gaps between collection sites were not included, because the routes between them were not always clear and because they were likely to represent stretches where no active searching for wild potatoes took place.

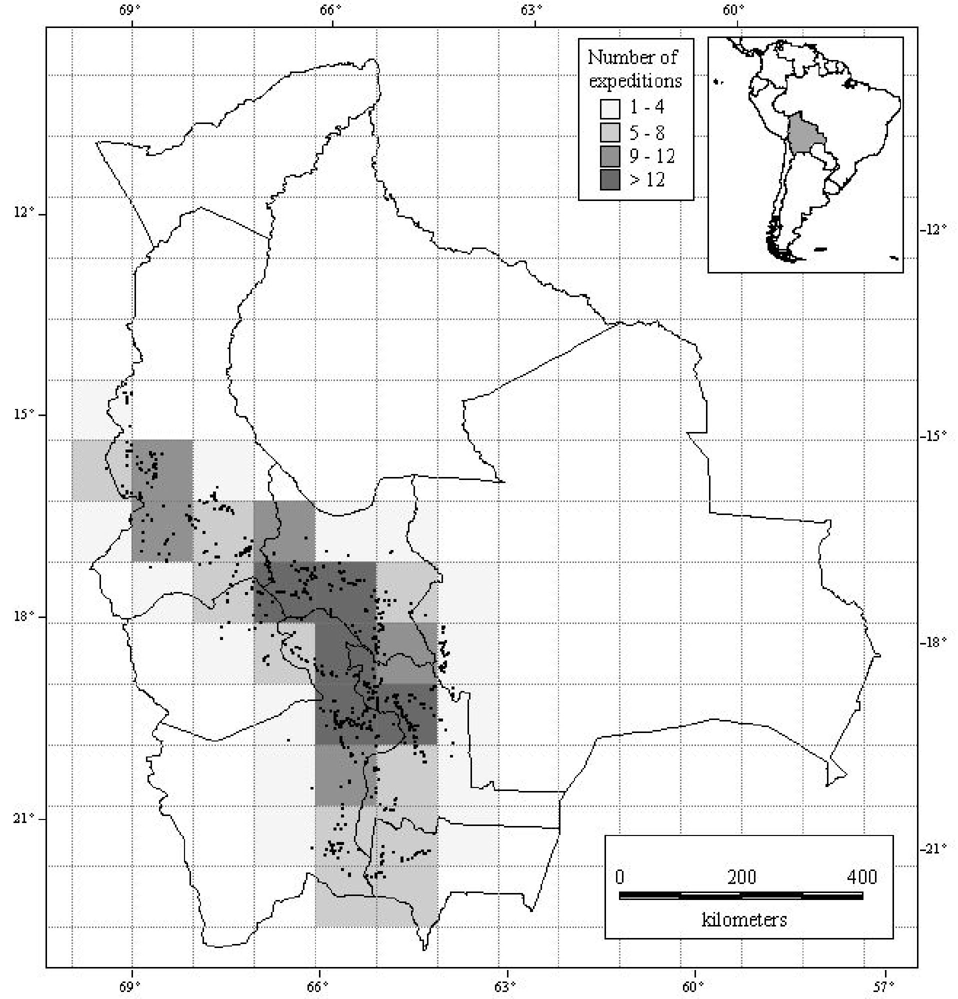


Figure 3.1 Bolivia, Departments (solid lines), locations where wild potatoes (*Solanum* spp.) were collected (dots, n = 1051), 100 × 100 km grid cells (dotted lines) and number of collecting expeditions per grid cell.

<u>Infrastructure bias</u>

To assess the presence of an infrastructure bias, Bolivia was divided into 1 × 1 km grid cells. For each grid cell, the distance to the nearest road, nearest main road, and nearest large city (i.e., the department capital) was determined. In turn, using the grid, the distance between each accession area and the main road was calculated. This yielded a specific distribution, which was then compared with the "random" distribution (i.e., the distribution of the distance between the [main] road or city over the grid cells where wild potatoes are know to occur), using a χ^2-test.

3.3 Results

Species-bias

According to our data, 45 wild potato species have been collected in Bolivia. The distribution of the number of accessions per species (A_S) is highly skewed (χ^2 test; $p < 0.01$; Fig. 3.2). *Solanum acaule* has been collected 166 times, while some other species appear only once. The number of accessions from *S. acaule* is more than double that from the second most collected species, *S. sparsipilum*. Together, these two species account for 23% of all accessions in the genebanks. Seven species yielded 52% of the accessions, and 24 species (almost half of the total collection) yielded 90% of the accessions.

The strong species bias evident in this case is clearly not desirable for maximizing the diversity of a genebank collection. However, this bias does not have a direct implication for geographical representativeness. *Solanum acaule* is widespread, and the high number of accessions from this species might be an accurate reflection of this.

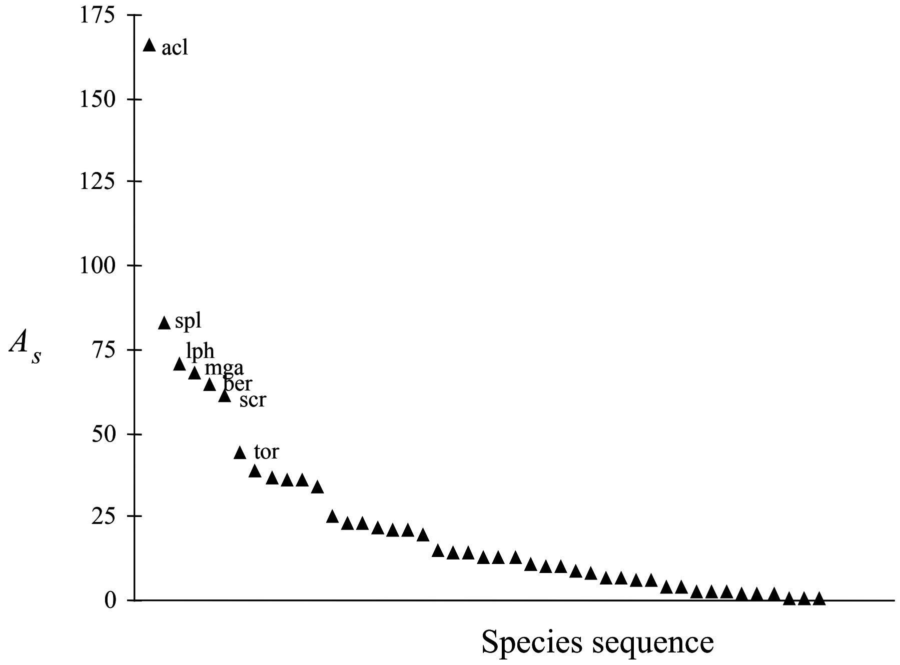


Figure 3.2 Number of Bolivian accessions per wild potato species (*Solanum* spp.), A_S, in the Intergenebank Potato Database (acl, *S. acaule*; spl, *S. sparsipilum*; lph, *S. lepthophyes*; mga, *S. megistacrolobum*; ber, *S. berthaultii*; scr, *S. sucrense*; tor, *S. toralapanum*).

Species-area bias

The number of accessions per species increases according to increases in the species' distribution area (Fig. 3.3). However, A_S increases disproportionally with G_S, and the observed relation between G_S and A_S is significantly different then the predicted linear and logarithmic relationships (χ^2 test; $p < 0.01$). *Solanum acaule* was strongly over-sampled. *Solanum sucrense*, *S. lepthophyes*, and *S. berthaultii* also were relatively over-sampled. However, not all species that were collected many times are necessarily over-sampled according to this criterion. The number of accessions from *S. sparsipilum* and *S. megistacrolobum*, for example, were about what would be expected on average given the area they occupy, despite their second and fourth rank on the abundance curve (Fig. 3.2). Other species such as *S. brevicaule* and *S. infundibuliforme* were relatively under-sampled.

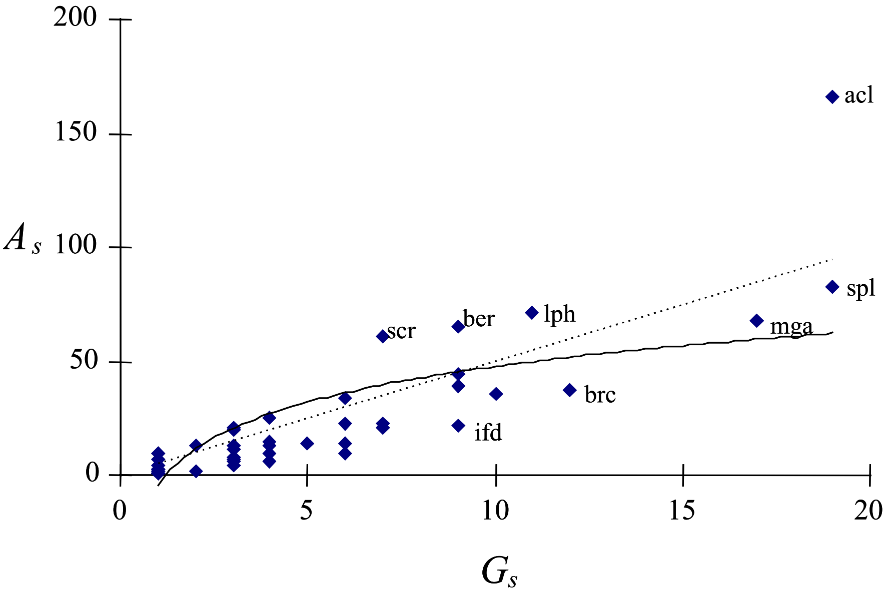


Figure 3.3 Number of accessions per wild potato species (*Solanum* spp.), A_S, versus number of grid cells per species, G_S (Solid line, $y = 24.2 \times \text{Ln}(x) - 7.3$, $R^2 = 0.54$; dashed line, $y = 5.0 \times x$, $R^2 = 0.76$; acl, *S. acaule*; spl, *S. sparsipilum*; lph, *S. lepthophyes*; mga, *S. megistacrolobum*; ber, *S. berthaultii*; scr, *S. sucrense*; brc, *S. brevicaule*; ifd, *S. infundibuliforme*).

Hotspot bias

The Bolivian wild potato collection is the result of 18 collecting expeditions. The number of accessions per cell increased with the number of expeditions to the cell. *A* linear approximation of this relationship is:

$$A_G = 5.1 \times E_G$$

($R^2 = 0.45$, $p < 0.01$). This relationship does not necessarily imply the presence of a hotspot bias; it might just reflect that the cells in which wild potatoes are more diverse have been visited more often. This is supported by the fact that the number of species collected within a grid cell, S_G, also increases with the number of expeditions to the cell:

$$S_G = 0.96 \times E_G$$

($R^2 = 0.13$, $p = 0.04$). This relation is rather weak, however. Moreover, the ratio of S_G and A_G, (i.e., the number of accessions per species in a cell) also goes up with the number of expeditions to the cell:

$$A_G / S_G = 1.4 + 0.4 \times E_G$$

($R^2 = 0.52$, $p < 0.01$), (Fig. 3.4). This suggests the presence of a hotspot bias. In terms of diversity, there has been excessive sampling, relatively speaking, in areas with high species diversity, unless one assumed that intraspecific diversity is higher in areas with more species.

Infrastructure bias

Most accessions have been collected near main roads that connect major Bolivian cities and surrounding areas (Fig. 3.5). Sixty percent of all accessions were collected within 2 km of the nearest road, and 78% were collected within 4 km (Fig. 3.6). This is in sharp contrast with the distance from random points to the major road in each area: 22% within 2 km and 44% within 4 km. Forty-four percent of the accessions were found within 10 km of a main road, rather than the expected 27%. Thirty-three percent of the accessions were found within 40 km of a department capital, as opposed to the expected 21%. In all three cases, the distributions differed significantly from a random distribution (χ^2 test; $p < 0.01$).

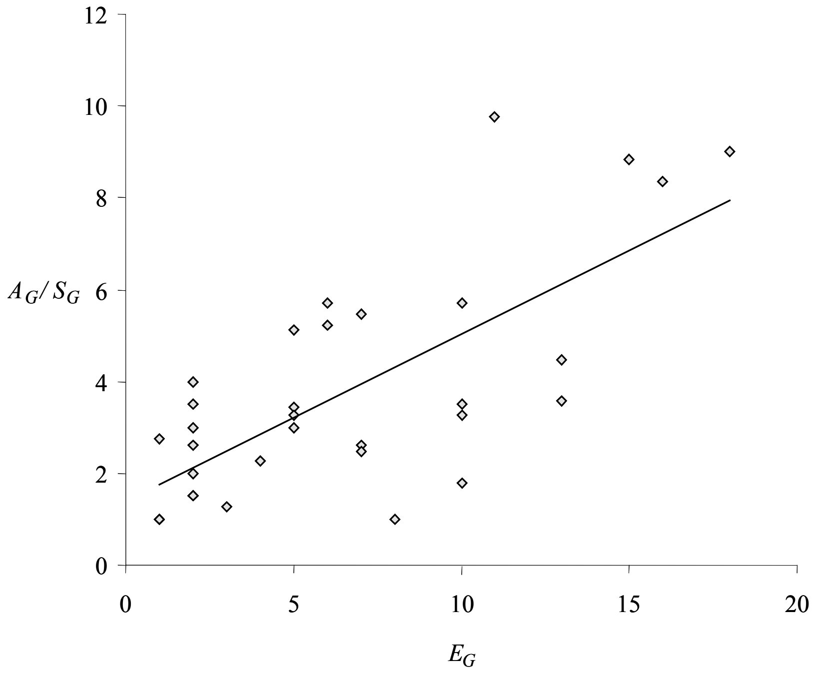


Figure 3.4 Ratio of the number of accessions per species (A_G/S_G) of wild potato species (*Solanum* spp.) in the Intergenebank Potato Database for Bolivia versus the number of expeditions (E_G) per 100 × 100 km grid cell.

3.4 Discussion

The species and species-area biases we found may have originated for several reasons. Collectors may have had a preference for a particular species. Some species may be more abundant than others, irrespective of the size of their distribution area, increasing the likelihood of collection. Some species are more common in disturbed areas such as roadsides, and are therefore more likely to be spotted and collected. The most frequently collected species, *S. acaule*, is an abundant inbreeder that is almost always found carrying seeds. Other species are rare, endemic, and often found without seeds and are therefore hard to collect. On the other hand, collection expeditions expeditions are sometimes justified by the absence of certain species in genebanks (e.g., Spooner *et al.,* 1999), and a lot of effort may be expended to obtain samples of these species, which may be difficult to collect.

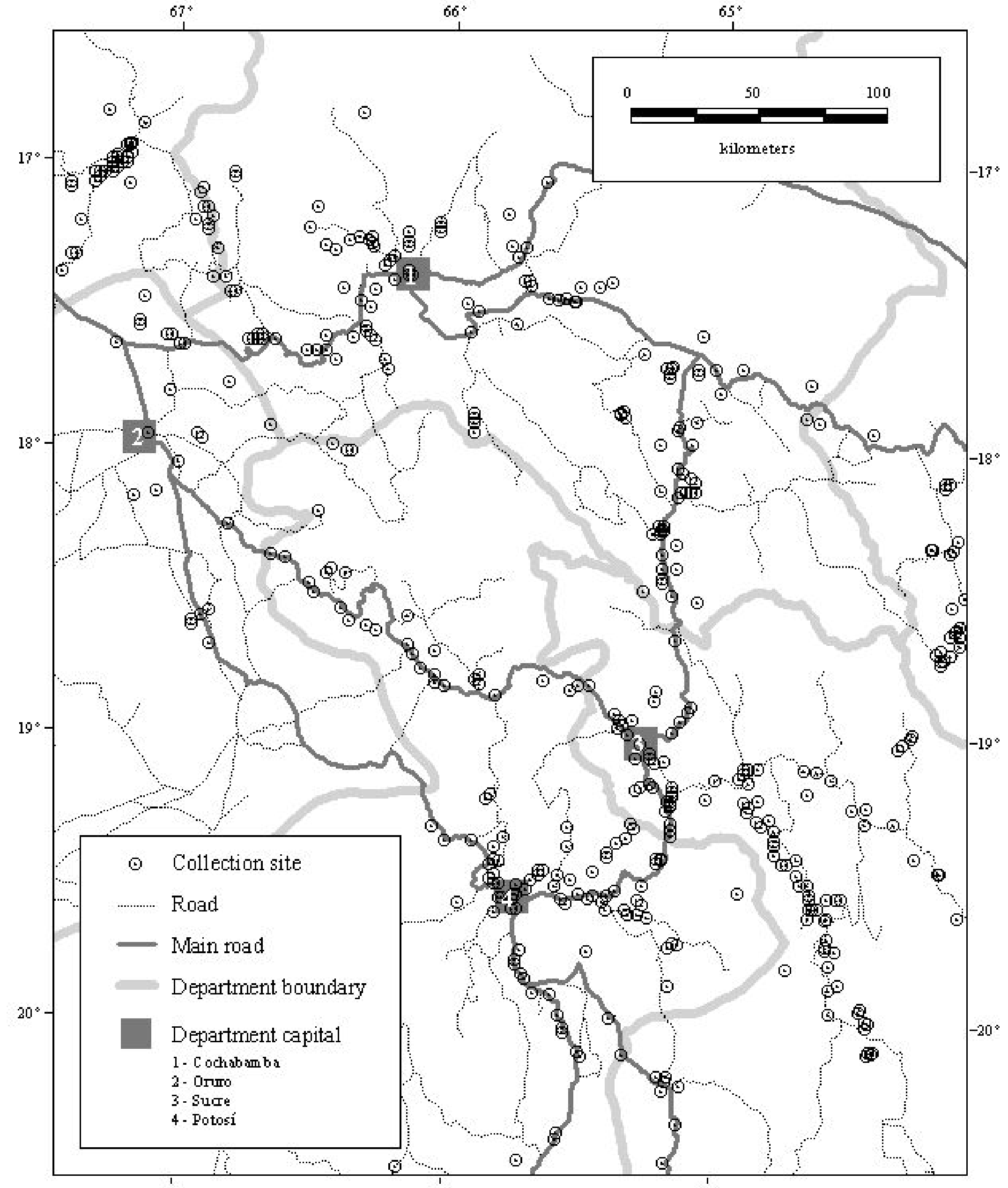


Figure 3.5 Location of road, towns and collections sites of wild potatoes (*Solanum* spp.) in central Bolivia.

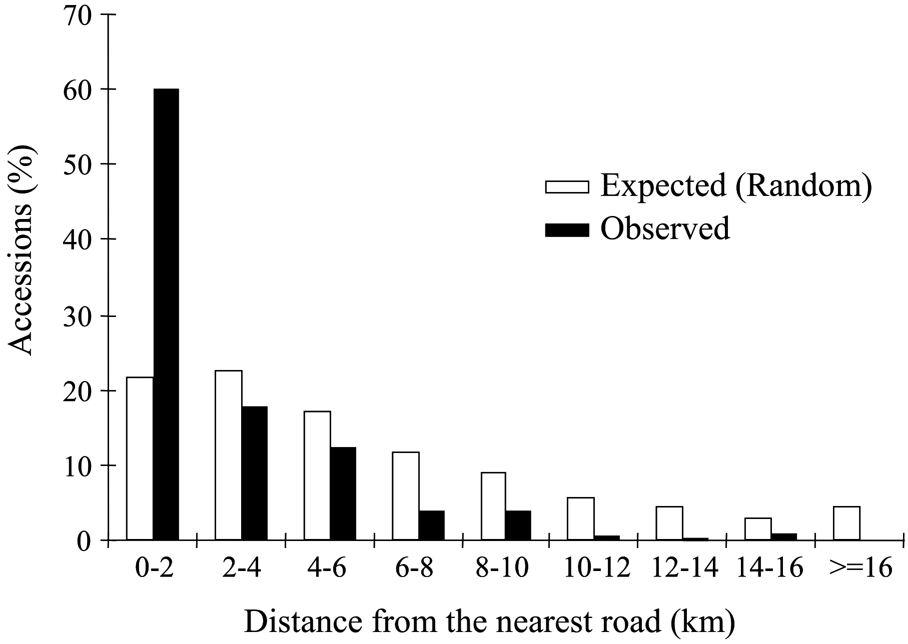


Figure 3.6 The distance from the nearest road in Bolivia for the accessions in the genebanks (observed) and for random points in the area where wild potatoes (*Solanum* spp.) occur.

If there is no species-area bias, a stricter definition of species-area bias might be considered in which a representative collection would contain equal numbers of accessions for each species in each grid cell where that species occurs. This definition is stricter, because it requires an equal allocation of accessions per grid cell within a species as well as an allocation per species that is proportional to the number of grid cells a species occupies.

Species and species-area data depends on the definition of species boundaries. *S. megistacrolobum* and *S. toralapanum*, Numbers 4 and 7 in the abundance ranking (Fig. 3.2), respectively, can be considered related varieties (Ochoa, 1990) or subspecies (Giannattasio and Spooner 1994a, 1994b). If they were treated as one entity instead of two, that one species would be the second most dominant in the collection.

The presence of a hotspot bias suggests that collectors tend to collect in areas where previous expedition reports indicate the presence of many taxa. The fact that the IPD comprises several different genebank collections may have amplified this bias, because all genebank curators would seek some representation of the predominant species.

One weakness in hotspot bias analysis is that it is not known if unsuccessful attempts were made to collect wild potatoes (see white areas in Fig. 3.5). These locations are not recorded in genebank databases, and there are probably few collectors who systematically include observations on the absence of certain species in certain areas in their records. It would be useful if collectors would develop a more systematic documentation system that would include absence/presence and abundance data for all sites they explore, whether germplasm was collected or not. It would be useful if route maps indicated the stretches along which active exploration took place, to distinguish these from stretches that were, e.g., deemed not relevant, or passed by night. The more information recorded about additional factors such as habitat and land-use patterns, the more useful both the collection and accompanying data will be (Brush 1989). In 1970, Bennett noted that "the need to formulate basic and standard procedural patterns with the capacity to absorb, integrate, preserve and make available the data of exploration missions is now urgent". Unfortunately, this statement is still valid.

Species-area and hotspot biases are scale-dependent. Final scale selection is usually a compromise between the conflicting objectives of a high number of grid cells and a high number of observations per cell. As our main objective is screening genebank data for use in GIS-based studies, it is reasonable to check for these two biases at the same scale that will be used in subsequent studies. In terms of maximizing diversity in the actual genebanks, at least for wild potato, the optimal collection scale has not been determined. Multi-scale studies on the distribution of crop genetic diversity would be useful in this respect.

An important factor leading to infrastructure bias is time limitations, particularly for expeditions in vast, rugged areas like the Bolivian Andes. Maximum coverage of such areas precludes roaming far from main roads. Therefore, in practice, explorers can not sample randomly, particularly for endemic species. The fact that wild potatoes have been over-sampled near roads and cities may render the genebank less representative, but it may also have a positive side. Genetic erosion is mainly due to increased human activities such as building and agriculture, which often occur near roads and cities. Therefore, it may be that the infrastructure bias has inadvertently led to a collection bias toward genepools that are the most endangered. The opposite might also be true: infrastructure bias may have favored collection of common weedy species that tend to grow near roads. For these species, an apparent infrastructure bias would be expected even if collection were done randomly throughout an area. Other species may be reduced near roads and cities because of land-use patterns such as grazing. For these species, an infrastructure bias would lead to under-collecting.

Since the primary objective of genebanks is to conserve diversity, sampling should be carried out accordingly. By standards such as those suggested by Brown and Marshall (1995), all species in the collection are underrepresented in the absolute sense. Many wild potato species are endemic, rare, and hard to locate. For these species, theoretical sampling schemes do not have much practical relevance.

For studies of the spatial distribution of crop diversity, other sampling approaches such as the grid-based data collection system used in Great Britain that was analyzed by Prendergast *et al*. (1993a) would be very useful. It is unlikely, however, that this type of data could be obtained for many areas or species, so researchers will have to rely on data from genebanks. Checking for the four types of biases described here can be a useful point of departure for studies based on genebank data. Depending on the objectives of the study, it may be necessary to adjust the methods and scales used. For example, if an infrastructure bias is present and could potentially skew data, the subsequent study may have to be carried out at a scale that minimizes this bias (generally a low-resolution scale). In the case of a species-area bias, one might consider eliminating some data, to remove the bias..

In formal comparisons of the observed allocation of accessions to the predicted allocation, assuming no biases, the following should be noted. Since the information on the collection accessions is, in effect, a complete census, statistical tests are not needed to determine the presence of the four types of biases described in this paper. Therefore, the collection can be examined to determine whether the accessions are allocated in the proportions expected in the absence of a bias. Decision on how much deviation from this expectation is needed to consider a bias important then becomes a matter of judgement, based on how the information will be used. The statistical tests used in this study, comparing observed allocations to predicted allocations in the absence of bias, can be viewed as indicators of inherent trends toward bias. That is, assuming that past and future processes for adding accessions to collections are the same, these tests can help indicate whether or not, disproportionate allocations are likely to be maintained.

3.5 Conclusion

We have defined four types of biases that may be present in genebank collections: species, species-area, hotspot, and infrastructure, and have evaluated them for a collection of wild potatoes from Bolivia. Since all four types of bias were present to some extent, the genebank may not be an adequate representation of the existing diversity or of the actual geographic spread of wild potato species in Bolivia. Collecting expeditions did not use a spatially unbiased statistical sampling procedure to collect a representative sample of the wild potatoes in Bolivia. Instead, it seems that collectors often went to sites where they expected to find a particular species, perhaps based on the results of former expeditions, and that they tended to collect samples near major roads.

As genebanks are mainly compiled to achieve maximum diversity and usefulness (e.g. crossability, presence of resistance genes), spatial biases may be acceptable. However, the methods and findings presented in this paper may help prevent redundancy and omissions in genebanks. Future collections could focus on under-represented species and/or on those from under-represented areas, depending on whether the genetic diversity (in useful traits) was a function of species, region, or both. This would increase the probability of conserving a more balanced collection of potentially useful diversity.

Our findings also have important implications for the use of genebank data in GIS-based exploration and conservation studies (e.g., Jones *et al.,* 1997). An underlying assumption in such studies may be that data from collections are a representative random sample of the existing diversity of a species. We have shown that this assumption may not always be justified. Although we have only analyzed data from one genebank, the factors considered in this study may be relevant in other settings. For example, McAllister *et al.* (1994) mention a sampling intensity bias that complicates species density mapping of coral reef fish. Therefore, analyses of spatial representativeness may be important for any biological collection being used in GIS-based studies.

4. Distribution and diversity of wild potatoes

Published as: Hijmans, R.J. and D.M. Spooner, 2001. Geographic distribution of wild potato species. American Journal of Botany 88:2101-2112.

Abstract

The geographic distribution of wild potatoes (Solanaceae sect. Petota*) is analyzed using a database of 6073 geo-referenced observations. Wild potatoes occur in 16 countries, but 88% of the observations are from Argentina, Bolivia, Mexico, and Peru. Most species are rare and narrowly endemic: for 77 species the largest distance between two observations of the same species is less then 100 km. Peru has the highest number of species (93), followed by Bolivia (39). A grid of 50 × 50 km cells and a circular neighborhood with a radius of 50 km to assign points to grid cells was used to map species richness. High species richness occurs in northern Argentina, central Bolivia, central Ecuador, central Mexico, and south and north-central Peru. The highest number of species in a grid cell (22) occurs in southern Peru. To include all species at least once, 59 grid cells need to be selected (out of 1317 cells with observations). Wild potatoes occur between 38°N and 41°S, with more species in the southern hemisphere. Species richness is highest between 8°and 20°S and around 20°N. Wild potatoes typically occur between 2000 and 4000 m altitude.*

4.1 Introduction

The wild tuber-bearing *Solanum* L. species (Solanaceae sect. *Petota* Dumort.) and outgroup relatives in *Solanum* sect. *Etuberosum* (Bukasov and Kameraz) A. Child, hereinafter called wild potatoes, are relatives of the cultivated potato (Hawkes, 1990; Contreras and Spooner, 1999). The cultivated potato (*Solanum tuberosum* L. and six other cultivated species that are only grown in the Andes) is one of the world's principal food crops (Walker *et al.*, 1999).

In addition to these seven cultivated potato species, there are 199 wild potato species (Spooner and Hijmans, 2001). These wild species all grow in the Americas, from the southwestern United States to central Argentina and Chile. They have been collected intensively to contribute to potato genebanks, and they have been used in breeding programs to improve the cultivated potato (Ross, 1986; Hawkes, 1990). There is no evidence that wild potatoes are currently being domesticated or traded, or that their geographic distribution is deliberately altered by man, except for the case of *S. chacoense*, a species that is grown as an ornamental plant in Lima, Peru (Ochoa, 1962). However, in some parts of the Andes there is ongoing gene flow between wild and cultivated potato species (Hawkes, 1979; Watanabe and Peloquin, 1989; Rabinowitz *et al.*, 1990).

Despite a purported great variability in wild potatoes (Hawkes, 1990), many of the species have a general appearance similar to the cultivated potato (Spooner and Van den Berg, 1992). A typical wild potato has pinnately dissected leaves (a few species have entire leaves), corollas in various shades of purple (colors vary from white to blue to purple and pink), corollas pentagonal to rotate (some are stellate), and fruits spherical to ovoid (some are conical). This similarity has led to widely conflicting taxonomic treatments of wild potatoes. Hawkes (1990) and Spooner and Hijmans (2001) provide the most recent taxonomic overviews of the entire group.

Wild potatoes have been the object of intensive study. Spooner and Hijmans (2001) recently reviewed wild potato systematics and germplasm collection. Jansky (2000) reviewed the value of wild and cultivated potatoes in breeding for disease resistance. However, there has been no comprehensive analysis of the geographic distribution of wild potatoes, which we provide here for the first time. We used geographic information systems (GIS) to analyze a large geo-referenced database of locations where wild potatoes were observed.

We computed country- and species-level statistics. For each species, we estimated the geographic area over which it occurs and mapped the number of observations and species richness, using grid cells. We determined the minimum number of grid cells needed to include all species, and related species richness to latitude and altitude. Species richness is used because it is a simple, widely used, well-understood, and useful measure of taxonomic diversity (Gaston, 1996), and because it is less sensitive than diversity indices to the problems of unsystematic sampling intensities and procedures (Hijmans *et al.*, 2000b).

This study explores the use of Geographic Information Systems (GIS) to describe the geographic distribution of wild crop relatives. This type of study can provide baseline data for further GIS analysis for exploration, conservation, and use of germplasm of wild crop relatives (Guarino *et al.*, 2002), as well as for studies of the factors that explain the geographic distribution of these species.

4.2 Materials and methods

Wild potato distribution data

Our sources of wild potato distribution data were the following: (1) the Inter-genebank Potato Database, which has data from seven genebanks in the USA, Peru, The Netherlands, Germany, Argentina, the UK, and Russia (in order of size of contribution) (Huamán *et al.*, 2000). (2) Data from 16 collecting expeditions in 12 countries by D.M. Spooner and coworkers (Spooner and Hijmans, 2001). These included records that are not in genebanks because they were collected as herbarium species or were lost as living specimens after collection. (3) A database of herbarium records developed by J.G. Hawkes (Hawkes, 1997). (4) Hawkes and Hjerting (1969; Argentina, Brazil, Paraguay and Uruguay). (5) Hawkes and Hjerting (1989; Bolivia). (6) Ochoa (1990; Bolivia). (7) Ochoa (1999, Peru). (8) Spooner *et al.* (1998; Guatemala). (9) Spooner *et al.* (1999; Peru). (10) Spooner *et al.* (2001a, Costa Rica). (11) Spooner *et al.* (2001b, Mexico).

We used all records from sources 1-3 that included a species name and passport data. Passport data include a description of the location of origin (such as locality name), administrative units (such as departments and districts), and geographic coordinates. However, coordinate data were absent for many records. For the genebank databases, coordinates were assigned using the locality description where possible. For Hawkes' (1997) database this was only attempted for species for which we had fewer than five observations with coordinate data. Sources 5-7 were used to verify and improve the geographic coverage of our species distribution data (see below). Additional herbarium records were taken from sources 8-11.

The presence of coordinate data allowed the analysis of the database with GIS; we used ArcView-GIS (Environmental Systems Research Institute, 1999) and DIVA-GIS software (Hijmans *et al*., 2001b). Coordinate data in genebank databases often lack precision, and were checked and modified following procedures described by Hijmans *et al.* (1999). First, we checked for gross errors, such as accessions located in the oceans. Then, we made overlays (simultaneous spatial queries) of the collection sites and administrative boundary databases (first level subdivision for Mexico and Central America; first and second level for the United States and South America; and first, second and third level subdivision for Peru). We compared the names of the administrative units according to the wild potato distribution database with those of the administrative boundary database. In case of discrepancies between the two databases, the coordinates were checked against the locality description and new coordinates were assigned where needed.

Dot maps of the distribution of all species were compared with published species distribution maps from the floristic sources 5-7. When general areas of occurrence were already represented on our maps, we did not include additional points, because of a possible lack of precision of many of these maps, and the risk of duplicating records. However, if it appeared that our distribution maps did not include all major areas where a species was reported to occur, we did copy additional observations for these areas. Species names follow Spooner and Hijmans (2001), who list 196 wild species in sect. *Petota* and three outgroup species in sect. *Etuberosum*. Taxonomic groups below the species level (subspecies, varieties, and forms) are all treated within their component species.

Country- and species-level distribution

The number of observations and species in the database were tabulated by country. This was done separately for rare species, here defined as species for which we had fewer than five observations. The number of observations per species was calculated and plotted. The average number of observations per species was calculated to assess intensity of collection by country, given the species richness it harbors.

Area of distribution

For each species we estimated the area over which it occurs, using two statistics: (1) maximum distance (MaxD) between two observations of a single species was calculated as the largest distance (in metres) between all possible pairs of observations of one species; and (2) we assigned a circular area circular area (CA_r), with a radius r, to each observation and calculated the total area of all circles per species. Areas where circles of a species overlap are only included once. Area is expressed as the area relative to the area of one circle, i.e., the number of circular areas covered. We decided to use a radius of 50 km (i.e., CA_{50}).

The assumption is that each point observation represents a group of plants that covers a circular area with a 50-km radius. Expressing CA_r as the number of circles instead of the absolute area makes it more easily comparable across different studies and scales (when a radius other than 50 km is chosen).

The CA_{50} statistic was plotted against the number of observations to explore differences in abundance between species. A species with a relatively high number of observations per CA_{50} would be abundant within its area of distribution, whereas a low number would indicate that a species was more scattered over the range in which it occurs.

To describe species distributions we use the terms “endemic” and “rare”. We use “endemic” for species that occur in relatively small areas (have a small range size) (Rabinowitz, 1981; Gaston and Williams, 1996) and “rare” for species that have been observed in relatively few cases.

Grid based distribution

We compared the number of observations and species using a grid with 50 × 50 km cells, and summarized the results by country. We used ArcView-GIS (ESRI, Redlands, CA, USA) to transform the coordinate data to the Lambert equal-area azimuthal projection, with 80°W as the central meridian and the equator as the reference latitude. Because the origin of a grid is arbitrary but can influence the results, it may not be accurate to assign a point to one cell only, if the point is located near one or three other cells. Therefore, the data were assigned to grid cells using a circular neighborhood (Cressie, 1991; Bonham-Carter, 1994) with a radius of 50 km, using the DIVA-GIS software. All the observations within that neighborhood were assigned to its respective grid cell, and an observation can, therefore, be assigned more than once. The result is a smoother grid, which is less biased by the origin of the grid and also less sensitive to small changes (errors), in the coordinate data. When we discuss grid cells in this paper, these refer to circles with an area of $\pi r^2 = 7854 \text{ km}^2$ with their center in the middle of the grid cells with an area of 2500 km^2.

To assess the distribution of species over grid cells, we plotted the number of species per grid cell against the number of grid cells. Data on plant distributions can be biased due to spatial differences in recorder (a person who takes data) effort (Rich and Woodruff, 1992; Gaston, 1996; Hijmans *et al*., 2000b). In extreme cases, differences in number of species between areas would reflect the amount of time spent there by recorders, and not reflect actual differences in distribution. We assessed the extent to which the number of observations predicts the number of species by plotting the number of species vs. the number of observations per grid cell.

Complementarity analysis

To further analyze aspects of species distribution and endemicity we identified the smallest area (number of grid cells) needed to capture all wild potato species. This type of complementarity analysis is typically used in studies for optimal reserve selection (Csuti *et al*., 1997). We used the algorithm described by Rebelo (1994; see also Rebelo and Siegfried, 1992), and implemented in the DIVA-GIS software that selects grid cells so as to identify the minimum set of cells that captures a maximum amount of species. The algorithm selects the cell with most species in it, and then, step by step, selects cells that contain the highest number of additional (not previously included) species. In the case of cells having the same number of additional species, a random cell is selected from such cells. Selecting these complementary cells is a nonlinear optimization problem for which Rebelo's (1994) algorithm finds a near optimal solution. We determined the minimum number of grid cells needed to include all species, and mapped the location of these grid cells.

Distribution by latitude and altitude

To summarize the species distribution data, the number of species was tabulated by latitude and altitude. First, the number of species that occur in strips of 1° latitude was determined. Then, to obtain a smoothed line, for each 1° latitude zone the moving average was calculated, using five adjacent zones (two at each side). GTOPO30, a 30" grid (each cell is approximately 0.8 km^2) with altitude data (United States Geological Survey, 1998) was used to estimate altitude for all wild potato localities. This estimate was only used for the records for which the passport data did not include altitude. Observations were then grouped in classes of 250 m altitude and the number of species per class was plotted. The five-observation moving average was calculated and plotted.

4.3 Results

Country and species level distribution

Wild potatoes occur in 16 countries (Table 4.1). Four countries (Argentina, Peru, Bolivia, and Mexico) account for 88% of the records in the database. Peru has by far the highest number of species (93 species; 47% of the total). Peru also has the highest absolute and relative (over all species in a country) number of rare species (here defined as species with five or fewer observations), and 15 Peruvian species occur only once in our database (out of 17 species that occur once).

The distribution of the number of observations by species is far from uniform (Fig. 4.1). The most frequently observed species are *S. acaule* (630 observations), *S. leptophyes* (337), *S. megistacrolobum* (320), *S. bukasovii* (252), and *S. chacoense* (205). These five species account for 29% of the records and *S. acaule* alone accounts for 10%. The 72 species (36% of all species) with the least number of observations (five or less) make up only 3% of the records. A similarly skewed distribution has been described for Bolivian genebank accessions by Hijmans *et al*. (2000b) and for the interpotato genebank by Huamán *et al*. (2000).

The ratio between the number of observations and the number of species varies strongly across countries (Table 4.1). The ratio is very high in species-poor USA as well as in species-rich Argentina, indicating that these two countries have been explored more intensively for wild potatoes than have other countries, relative to their species richness. The ratio is low in many countries, some of which have low wild potato species richness. However, other countries in this group have an intermediate level of species richness, such as Ecuador and Colombia. Because the number of species tends to go up with collecting effort, the countries with a low ratio between species and observations would be the most likely places to find species that have not yet been discovered.

Table 4.1 Wild potato distribution by country. Number of observations (obs), species, rare species (obs ≤ 5), and the ratio of observations to species.

Country	obs	Species	Rare species	obs/species
Argentina	1688	28	4	60.3
Bolivia	1303	39	5	33.4
Brasil	24	3	0	8.0
Chile	100	4	1	25.0
Colombia	144	13	7	11.1
Costa Rica	24	1	0	24.0
Ecuador	142	15	7	9.5
Guatemala	69	6	0	11.5
Honduras	1	1	0	1.0
Mexico	926	36	6	25.7
Panama	15	2	0	7.5
Paraguay	22	2	0	11.0
Peru	1420	93	42	15.3
USA	158	3	0	52.7
Uruguay	13	2	0	6.5
Venezuela	24	3	1	8.0
Total	6073	199	72	30.5

Area of distribution

Most species are narrowly endemic and only 39 species occur in two or more countries. Most of these country co-occurrences are from Bolivia and Argentina (13 species in common), and from Bolivia and Peru (ten species in common). *Solanum chacoense* is the only species that occurs in five countries, followed by *S. acaule* and *S. commersonii* which occur in four countries. The average greatest distance between two observations of the same species (MaxD) is 411 km. For 77 species, MaxD is less than 100 km, and for 100 species (50% of the total), MaxD is less than 200 km (Fig. 4.2). The greatest MaxD observed was for *S. acaule*. At the time of this research, this number had recently increased by 732 km to a total of 3253 km, due to the recent discovery of this species in Ecuador (Spooner et al. [1992] who identified it as *S. albicans*; but Kardolus [1998], recognized it as an anomalous new hexaploid variety of *S. acaule*). Average circular area (CA_{50}) over all species is 6.7, but its distribution is strongly skewed. Seventy-two species have a CA_{50} of less than 2, and 124 have a CA_{50} of less than 5 (Fig. 4.2).

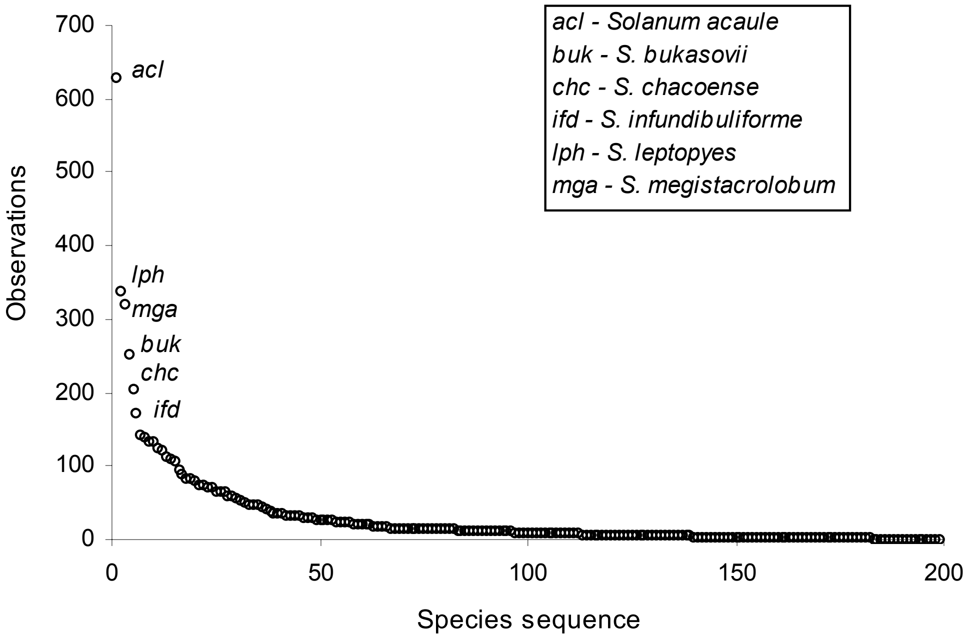


Figure 4.1 Number of observations of wild potatoes by species.

Clearly, both maximum distance and circular area go up with the number of observations. On average, a species has a CA_{50} of 0.16 times its number of observations (Fig. 4.3, regression line). There are, however, important differences among species. For example, *S. acaule* and *S. chacoense* occur in an area of comparable size (CA_{50} = 73 for *S. acaule* and 74 for *S. chacoense*), but *S. acaule* has been observed about three times more often, suggesting that *S. acaule* is much more abundant. *Solanum commersonii* is third in terms of CA_{50}, but is only 17th in terms of number of observations. The CA_{50} of *S. commersonii* is 0.36 times its number of observations while for *S. acaule* it is 0.12 times its number of observations. This suggests that *S. commersonii* is less abundant within its area of distribution than *S. acaule*.

Grid based distribution

The grid based maps showing the number of observations and species richness (Plate 1[*]) give a much more refined picture than the country summaries presented in Table 4.1. Species richness is clearly not homogeneously distributed within countries. There are few areas with many species, and many areas with few species (Plate 1; Fig. 4.4). The number of species follows a similar pattern to that of the number of observations. There is a strong positive correlation between the number of observations and species richness per grid cell (Fig. 4.5; compare Plate 1A and B). On average over all grid cells, there are 4.2 observations per species. Important deviations from this average are some areas in the USA and Argentina. As mentioned above, the number of observations in these two countries is high in comparison to the number of species. There is a relatively high number of samples in Argentina, particularly in the areas with high diversity. For example, there are 421 observations in the cell in Argentina with the highest number of species (17), a much higher ratio than for Peru (213 observations:22 species) or Bolivia (141 observations:20 species). However, in Argentina wild potatoes occur over a larger area than in any other country (295 grid cells) and the number of observations averaged out over that area has an intermediate value (17 per grid cell), lower than that of Peru (21) and of Bolivia (36; Table 4.2).

Species richness is particularly high in the southern and central Andes, and in central Mexico (Plate 1B). Going from north to south, the principal areas with high species richness are (1) the central Mexican highlands (México and Michoacán states); (2) a small area in central Ecuador (Chimborazo province); (3) a stretch from northern to central Peru (in Ancash, southern Cajamarca, La Libertad, and Lima departments); (4) southern Peru (in Cusco department); (5) central Bolivia (in Cochabamba, Chuquisaca, and Potosí and to a lesser extent La Paz an Tarija departments); and (6) northern Argentina (Jujuy and Salta provinces).

There are few cells with many species (Plate 1). Cells with more than 15 species are only found in Peru, Bolivia, and Argentina; Ecuador and Mexico are the only two additional countries that have cells with nine or more species (Table 4.2). Only 5% of the cells have more than ten species, while 52% of the cells only have one species (Fig. 4.4).

---

[*] Color plates are grouped in the Appendix.

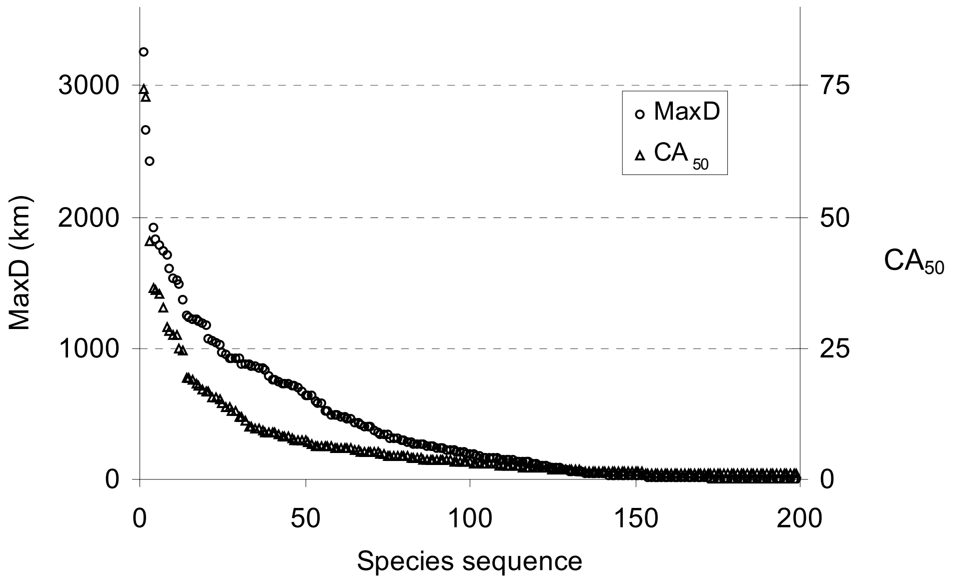


Figure 4.2 Maximum distance between two observations of one wild potato species (MaxD) and circular area (CA$_{50}$, a circular area with a 50 km radius was assigned to each observation. Areas where circles of a species overlap were only counted once. The area is expressed relative to the area of one circle). Species sequence is not necessarily the same for MaxD and CA.

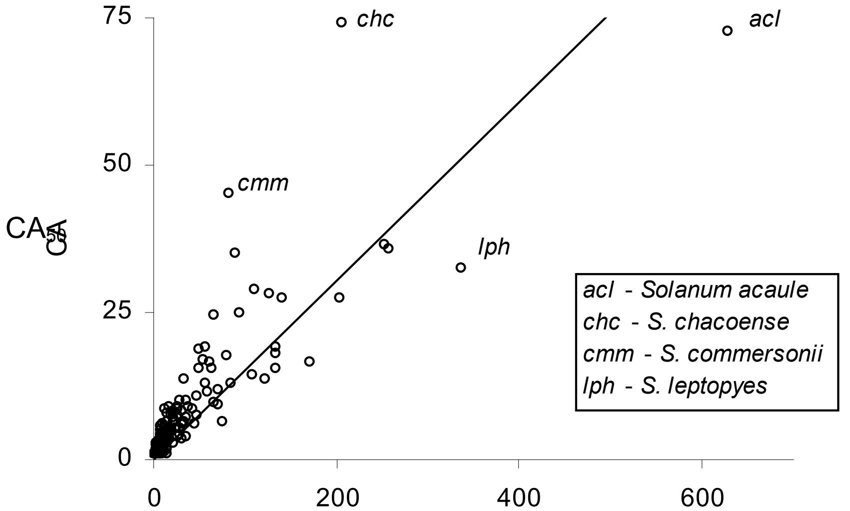


Figure 4.3 Circular area (CA$_{50}$) vs. number of observations of wild potato species. Each dot refers to one species. A circular area with a 50 km radius was assigned to each observation. Areas where circles of a species overlap were only counted once. The area is expressed relative to the area of one circle. Regression line: $y = 0.15x$, $R^2 = 0.68$. The two dotted lines ($y = 0.1x$ and $y = 0.5x$) are included for comparison only.

The highest number of species in a single grid cell is 22, and occurs in the department of Cusco in south Peru. Two cells have 20 species, one in the Bolivian department of Potosí (on the border with Chuquisaca), and one in the Peruvian department of Ancash. There are two cells in the Peruvian department of Cusco with 19 species, and one cell in the Bolivian department of Chuquisaca with 18 species. Although Peru has more species, its most species-rich areas are comparable in species richness to those of Bolivia. However, Peru has more cells with a high number of species, and its most species-rich cell only has 24% of all species present in the country. This again illustrates the high number of endemic species in Peru. In Bolivia, in contrast, the most species-rich grid cell has 51% of all Bolivian species (Table 4.2). There are also occurrences of relatively species-rich areas in Ecuador (60% of all species in that country), Argentina (61%), and in all countries with only a few species, but to a lesser extent in Colombia (31%) and Mexico (36%).

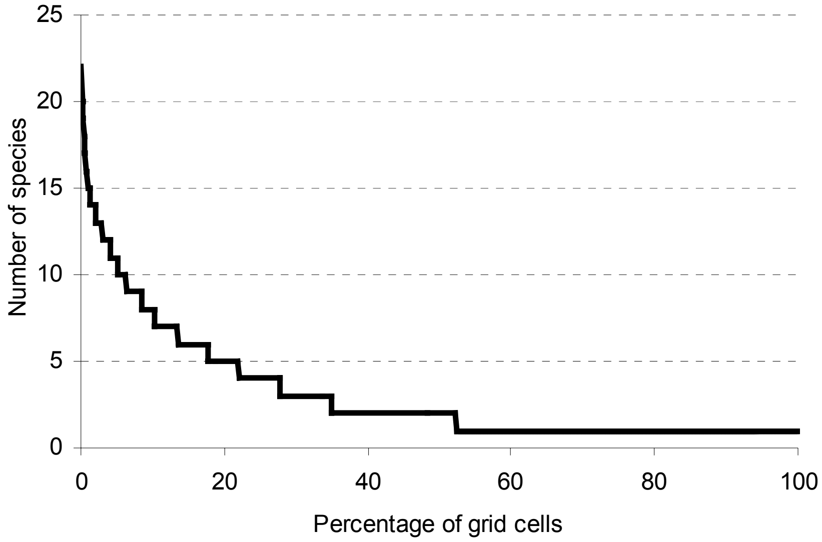


Figure 4.4 Frequency distribution of the number of wild potato species per 50 × 50 km grid cell. A circular neighborhood with a radius of 50 km was used to assign observations to a grid cell.

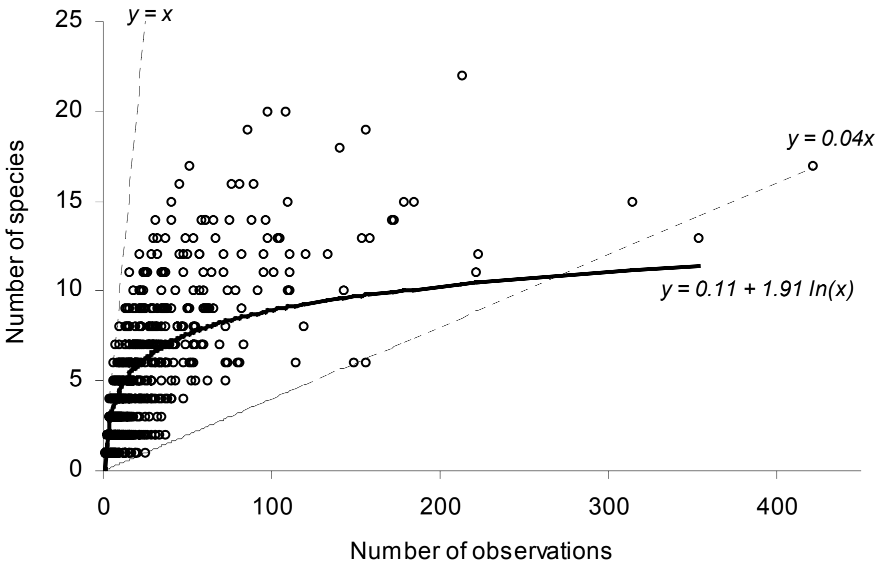


Figure 4.5 Ratio of the number of wild potato species to number of observations for each grid cell (obs. = 1247). Correlation coefficient = 0.74. Regression line: $y = 0.11 + 1.91 \ln(x)$, $R^2 = 0.65$.

Complementarity analysis

Although nine grid cells are enough to capture 51% of all wild potato species, the minimum number of grid cells needed to capture all species at least once is 59 (out of 1317 total cells) (Figs. 4.6 and 4.7). Twenty-three cells contribute only one additional species each (Fig. 4.7). The locations of the first 15 grid cells that get selected by Rebelo's (1994) algorithm follow a pattern that can only partly be inferred from Plate 1B. The early appearance of cells from Mexico and Ecuador (grid cells number 4 and 7 in Fig. 4.6) may seem surprising because they have only an intermediate level of species richness (Plate 1B), but the species in these countries are all different from those observed in Peru and Bolivia (the single observation of *S. acaule* mentioned above is the only exception). Areas in Argentina are selected later (numbers 6 and 13 in Figure 4.6) than might be expected on the basis its high number of species. This is because some of the species in these cells were already included in area number 3 (Fig. 4.6) in southern Bolivia.

Table 4.2 Grid-based species richness statistics by country.

Country	No. of grid cells with one or more obs.	Mean no. of spp. per grid cell	Mean no. of obs. per grid cell	Highest no. of spp. in one cell	Total no. of spp. in the cell with highest no. of spp. (%)
Argentina	295	2.4	16.6	17	61
Bolivia	124	6.2	36.4	20	51
Brazil	56	1.0	1.5	2	67
Chile	49	1.2	5.2	2	50
Colombia	65	1.9	6.9	4	31
Costa Rica	11	1.0	6.2	1	100
Ecuador	37	4.1	12.3	9	60
Guatemala	19	3.5	11.6	6	100
Honduras	2	1.0	1.0	1	100
Mexico	269	3.2	10.7	13	36
Panama	5	1.8	11.6	2	100
Paraguay	21	1.4	3.2	2	100
Peru	211	5.3	21.2	22	24
USA	119	1.2	4.1	2	67
Uruguay	22	1.2	1.7	2	100
Venezuela	12	2.1	6.5	3	100

In northern Peru and southern Ecuador, there is a group of four nearby cells that are selected at an early stage (within the first 15 iterations). This means that in these areas, there are not only individual cells with a high level of species richness, but that there is also a high turnover of species composition between nearby grid cells. Despite its second rank (tied with Ecuador) for rare species (Table 4.1), Colombia is not included in the selection at an early stage because its species are not geographically clustered (see also Plate 1B; Table 4.2).

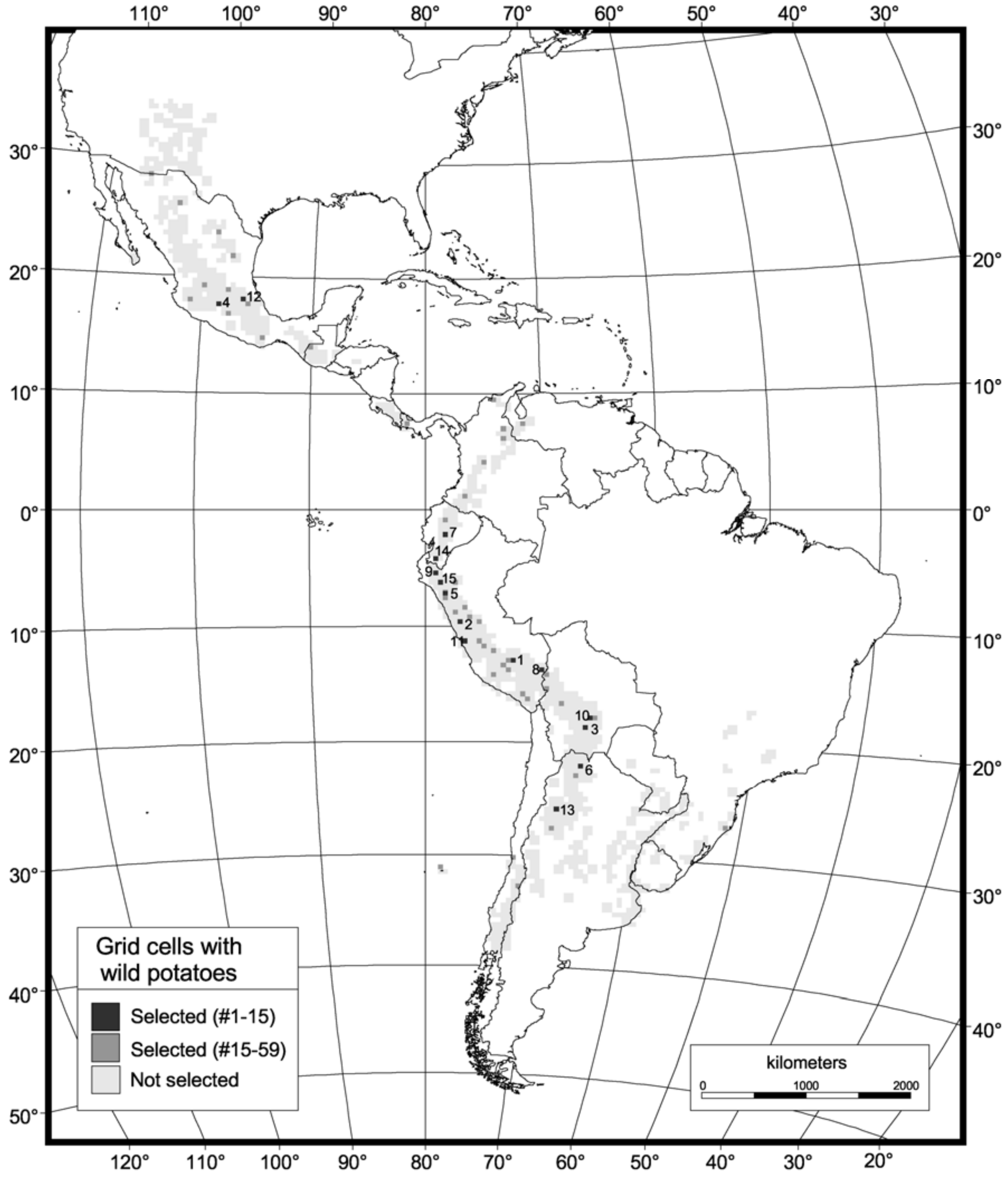


Figure 4.6 The location of the first 15 grid cells selected, and locations of the other 44 grid cells needed to include each wild potato species at least once.

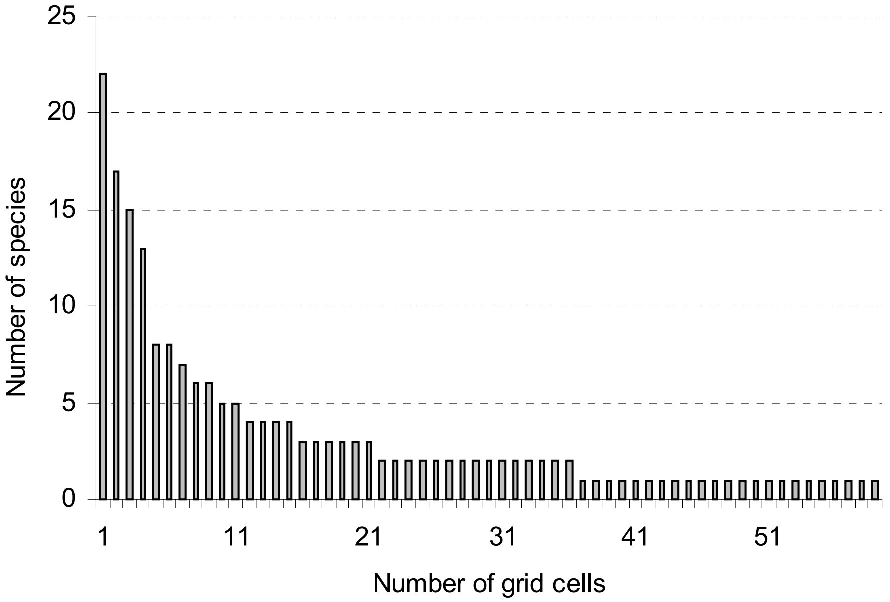


Figure 4.7 Number of additional species included per grid cell, when selecting grid cells with the objective to select all species in as few grid cells as possible. The first 15 sites correspond to the numbered grid cells on Figure 4. 8.

Distribution by latitude and altitude

Wild potatoes occur between 38°N and 41°S. The highest number of species per degree latitude (>20) occurs between 8°S and 20°S, i.e., from north-central Peru to central Bolivia, and around 20°N, in the central Mexican highlands (Fig. 4.8). The distribution of the number of species by latitude follows a bimodal distribution. There is a remarkably similar pattern between 20° and 40° in both hemispheres. However, in the zone between 20°N and 20°S, and particularly the zones between 8°N and 15°N and 8°S and 15°S, the number of species is rather different, with a conspicuously higher number of species in the southern hemisphere.

Wild potatoes are most common in the tropical highlands (compare Plate 1B and Fig. 4.9), particularly between 2000 and 4000 m (Fig. 4.10). The average elevation for all species is 2770 m when weighted by species, and 2890 m when giving equal weight to all observations in the database. Ninety-one percent of the wild potato species occur, on average, above 1750 m. Of all observations, 75% appear in areas above 2300 m. Almost all of the lower elevation species and observations are from the plains and hills in Argentina, Brazil, Mexico, Paraguay, Uruguay, and USA, i.e., from high latitudes.

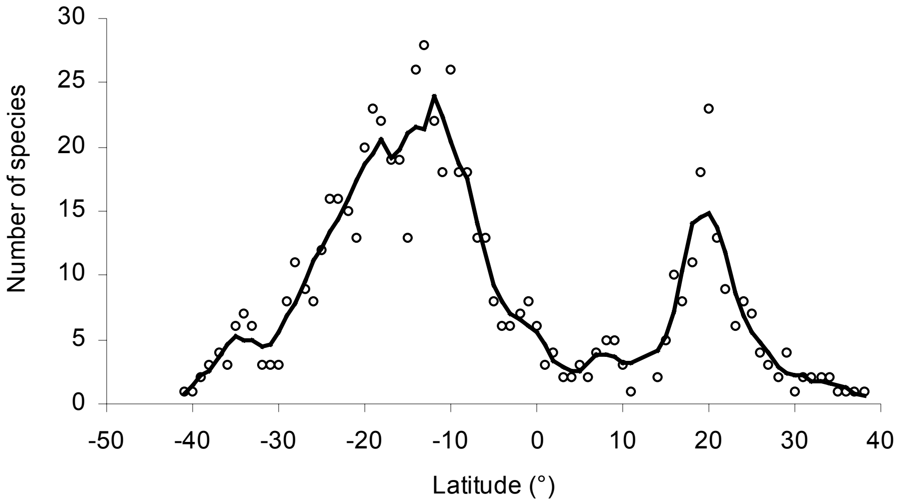


Figure 4.8 Wild potato species richness by latitude. Each observation represents the number of species found in a 1° latitude wide area. The line is the five-observations moving average.

4.4 Discussion

Although distribution maps of wild crop relatives are common (e.g., Zeven and Zhukovsky, 1975), this is the first study in which a group of closely related wild crop relatives is systematically analyzed using GIS. This study is also unique in its use of a very large number of geo-referenced observations for a single group of closely related wild species.

Wild potatoes occur between 38°N and 41°S. Species richness of wild potatoes is particularly high in the Central and South American tropical highlands, with clear peaks between 8°S and 20°S and around 20°N, i.e., areas in the Andes of northern Argentina, Bolivia, Ecuador, and Peru, and in central Mexico. Peru stands out for the high number of wild potato species as well as for the high number of rare wild potato species.

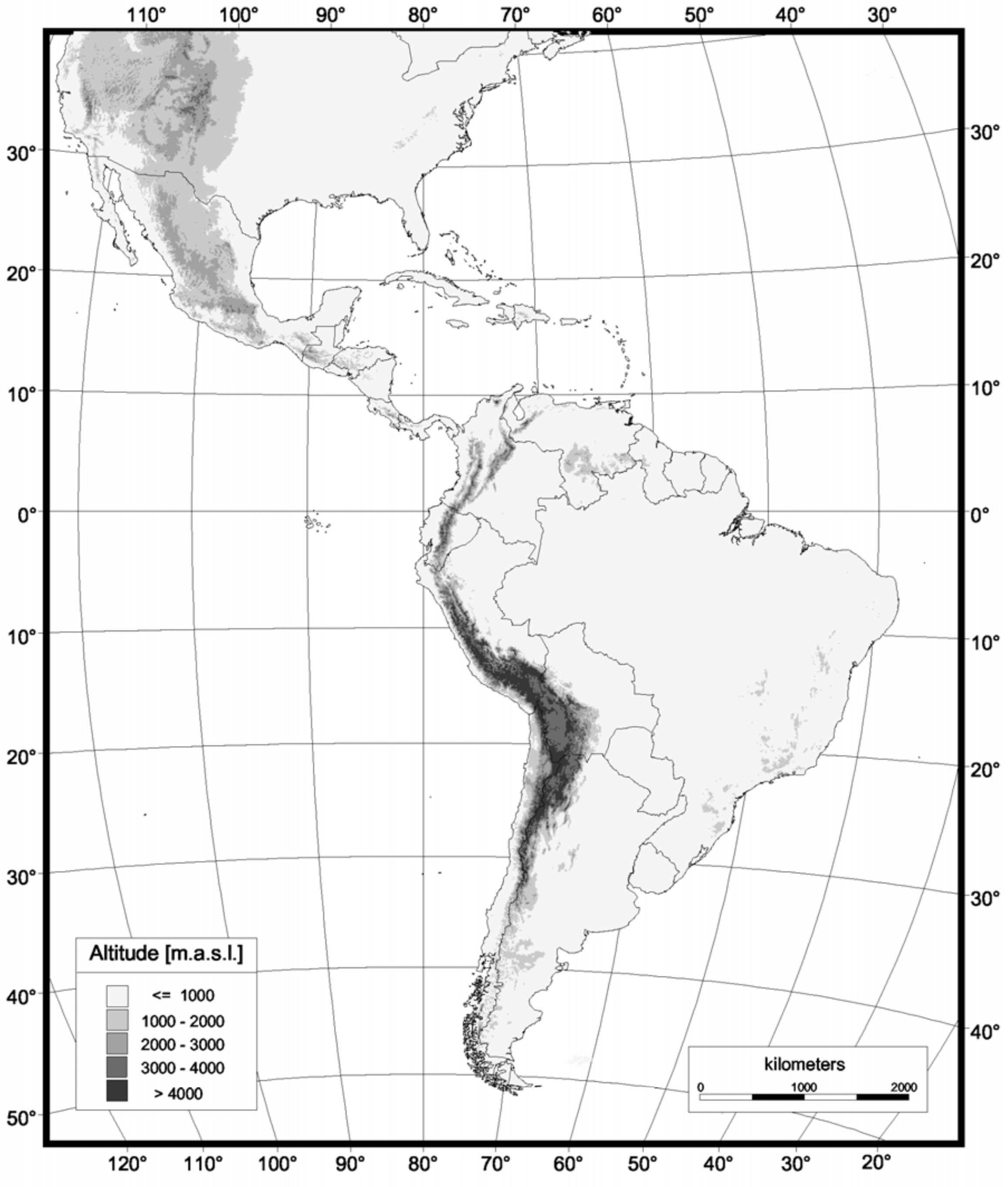


Figure 4.9 Elevation in Latin America and parts of the USA. Data source: United States Geological Survey (1998).

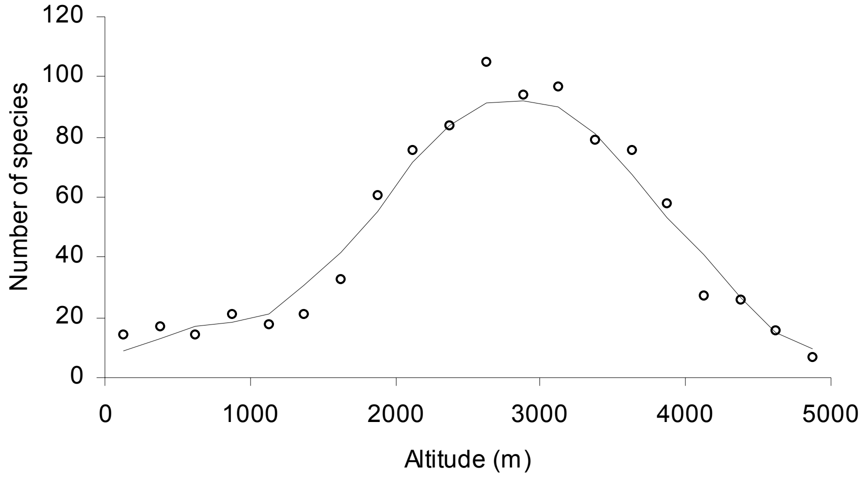


Figure 4.10 Wild potato species richness by altitude. Each dot represents the number of species observed in an area covering 250 m of difference in altitude. The line is the five-observations (1250 m altitude difference) moving average.

Many wild potatoes species are narrowly endemic, and yet a selection of only nine grid cells was needed to include 51% of the species, which emphasizes the presence of areas of high species richness. In a number of countries, the most species-rich grid cell has a high percentage all species. This might facilitate the design of in situ conservation reserves to protect these species (as called for by Huamán *et al.*, 2000). However, the clustering of species on the scale used in this study is not directly meaningful for conservation programs, which typically would operate in considerably smaller areas.

The lower species richness around the equator, particularly in the northern hemisphere, as compared with higher tropical latitudes, contrasts with the general pattern of increasing species richness (of all flora and fauna) towards the equator (e.g., Blackburn and Gaston, 1996; Gaston and Williams, 1996). The absence of cool tropical highlands appears to be an important factor that explains the paucity of wild potato species around the equator, particularly in the northern hemisphere. The climate in these equatorial areas is also more humid and less seasonal. The absence of a clear dry (or cold) season could diminish the relative fitness of tuber-bearing perennials such as wild potatoes. At higher latitudes, where the data are more similar for both hemispheres, there is a considerable stretch of high mountains in central Mexico (the

Mexican transvolcanic belt). At even higher latitudes, wild potatoes mainly occur below 1500 m altitude.

Our data were gathered from several sources (expeditions) and this may have led to some redundancies. Particularly type localities of rare species were visited by different expeditions, as these species may not be found elsewhere. Hence, some of these species are even more rare and endemic than appears from our data. Overall, however, it may make our data more reliable given the timing dependency of the results of wild potato exploration: there are differences within and among years in the likelihood of finding certain species in certain locations.

Some of our records are recent, but many date back many years. In some cases, the habitat in which the species occurred has now disappeared, and the species may no longer occur there. For example, Spooner *et al.* (1998) describe a rapid rate of loss of wild potato habitat in upland forests in Guatemala. However, our recent experience in Peru, and elsewhere, indicates that this is not always the case. For example, Spooner *et al.* (1999) and Salas *et al.* (2001) collected many wild potato species in Peru in the exact location, often at the type locality, where they had been collected many years before. In some cases, it was not possible to collect at documented localities, but this was often attributed to phenology, as wild potatoes often have a short growing period. In other cases, incomplete locality data hindered collections.

Recorder effort (bias) (Rich and Woodruff, 1992; Gaston, 1996; Hijmans *et al*., 2000b) also influences the results. Nevertheless, the consistency of the results (there are no sudden gaps in the distributions) leads us to believe that we have presented a good representation of overall wild potato distribution, and this large database is one of the most comprehensive for any group of plants. Although the number of observations per grid cell is a reasonably good predictor of the number of species in that cell, we do not think that a high number of species follows causally from a high number of observations.

Wild potatoes have been the object of intensive exploration over many years (Spooner and Hijmans, 2001). While there will still be some areas where further exploration would discover additional species, a low number of observations in a cell is more likely to be the result than the cause of low species richness. Hence, the logic might be reversed: much collection has taken place in areas with a known high species diversity, referred to by Hijmans *et al*. (2000b) as “hotspot bias.” Attempts could be made to correct for recorder effort through rarefaction (estimating how many species would have been observed given a constant sample size per grid cell; Sanders, 1968;

Prendergast *et al.*, 1993b; Gaston, 1996). However, this would also lead to a great loss of information, because the observed number of species would be replaced by an estimate.

Distributions of wild relatives of crop plants have previously been summarized by country (e.g., Huamán *et al.*, 2000). However, countries (or their subsidiary administrative units) have different shapes and sizes. Hence, they have only limited value in comparing geographic distributions of wild plants, despite the advantage of being familiar entities. Equal-area grid cells as used in this study are clearly to be preferred. Nevertheless, there are a number of methodological issues that need to be considered when using grids.

Resolution (cell size) of the grid affects the results. Number of species per grid cell will increase with the size of the grid-cell, but this increase will be different among areas. We used a 50 × 50 km grid (and a circular neighborhood with a radius of 50 km) to strike a balance between the desire for high resolution and geographic sampling bias (Hijmans *et al.*, 2000b). This bias becomes less important when grid cell size increase. We used a neighborhood of 7854 km^2; most other studies of species distribution on a continental (or global) scale have used much larger grid cells. For example, Gaston and Williams (1996) review various global studies using grids of 611,000 km^2. The same was also used by Blackburn and Gaston (1996) for a study on birds in the Americas. Given the high density of observations, it was not necessary to use such large grid cells in our study. This would have amounted to serious information loss, and even smaller grid cells will be more appropriate for design of in situ reserves or for planning collecting expeditions to specific areas.

We used a circular neighborhood to assign values to grid cells because this gives a smoother result, particularly for areas with few observations (Cressie, 1991; Bonham-Carter, 1994) and is less sensitive to the origin of the grid and small errors in the locality data. However, this does include yet other factors (size, shape, and method used to define the neighborhood), that need to be considered when interpreting the results, in addition to the scale effect (effect of grid cell size). Research is needed to better understand the effects of different gridding methods and scales, in relation to data density and quality and the objectives of the study.

A complication to using species richness is the existence of conflicting taxonomic classifications (Gaston, 1996). Wild potatoes are a classic case in this respect (Harlan and De Wet, 1971; Spooner and Van den Berg, 1992). We use the list of species provided by Spooner and Hijmans (2001), which is a compilation of taxonomic names that updates Hawkes (1990). Nevertheless, future changes in species circumscription will likely change the results presented here. For example, the results of ongoing research described by Van den Berg *et al*. (1998) and Miller and Spooner (1999) on the 30 taxa of the *S. brevicaule* complex suggest a need for reduction of the number of species in this group. These species occur in southern Peru and Bolivia and taxonomic revision would reduce species richness here.

Although Peru seems to be reasonably well explored (number of species over observations), it has an extraordinarily high number of apparently rare species. This indicates that Peru may still harbor unknown species, as illustrated by the ten new Peruvian wild potato species described by C.M. Ochoa between 1998 and 1999 (Ochoa, 1999; Spooner and Hijmans, 2001). Peruvian species are also underrepresented in genebanks, and a collecting program is currently under way to fill this important gap (Spooner *et al*., 1999; Salas *et al*., 2001).

5. Locating useful traits

Submitted for publication as: Hijmans, R.J., M. Jacobs, J.B. Bamberg, and D.M. Spooner. Frost tolerance in wild potatoes: unraveling the predictivity of taxonomic, geographic, and ecological factors.

Abstract

The use of genetic resources could be more effective and efficient if we were able to predict the presence or absence of useful traits in different populations or accessions. We analyzed the extent to which taxonomic, geographic and ecological factors can predict the presence of frost tolerance in wild potatoes. We used screening data for 1646 samples from 87 species that had been collected in 12 countries in the Americas. There was a strong association of frost tolerance with species and to a lesser extent with taxonomic series. There was significant geographic clustering of wild potatoes with similar levels of average frost tolerance. Areas with a high level of frost tolerance are the central and southern Peruvian Andes, the lowlands of Argentina and adjacent areas, and a small area in the central Chilean Andes. There is a greater chance of finding wild potatoes with high levels of frost tolerance in areas with a yearly average minimum temperature below 3 ℃ than there is in warmer areas. However, temperature is only a weak predictor of frost tolerance. Temperature data alone did not predict observed frost tolerance in eastern Argentina/Uruguay and falsely predicted it in the southwestern United States. Because many wild potato species occur over small areas, taxonomic, ecological, and geographical factors are often confounded, and these factors should be analyzed simultaneously to interpret patterns in the distribution of traits.

5.1 Introduction

Wild crop relatives may have traits that can be useful for crop improvement. Particularly important traits are tolerances to biotic (insects, pathogens) and abiotic (e.g., cold, drought) stresses. Identifying wild populations or genotypes that possess such useful traits typically involves screening accessions from genebanks. There are usually many more populations in the wild than can be sampled, and screening all samples in genebanks is limited by available funding. It would, therefore, be valuable to be able to predict which populations would most likely possess specific traits of interest.

Taxonomic, ecological, and geographic factors could be used for prediction based on prior evidence of association, or on a priori assumptions. Taxonomic classification of organisms is based on morphological similarity and/or evolutionary relatedness. Ideally, groups of closely related taxa would also have certain useful traits in common that were not used to construct the taxonomic classification. Plant breeders have implicitly used taxonomy in this predictive sense, by linking traits to particular species (Ross, 1986; Hawkes, 1990).

Certain ecological factors may also serve for predicting traits because the presence of these traits might reflect adaptation of wild plants to ecological conditions prevailing in their area of occurrence. For example, tolerance to drought might be likely in populations growing in dry areas (Rick, 1973; Nevo *et al*., 1982). Geographic factors could play a role in prediction because certain traits may have arisen in an area and spread among taxa and ecologies in that area, but not reached areas farther away. Such geographic effects may be due to chance, but also to coevolution. For example, resistance to a certain disease may be present in areas where the pathogen is endemic, but absent in areas that, although similar from an ecological and/or taxonomic perspective, do not have the pathogen.

There are 199 wild potato species (196 in *Solanum* sect. *Petota,* and 3 in *Solanum* sect. *Etuberosum*; Spooner and Hijmans, 2001). They all occur in the Americas, from Colorado (United States) to Chile and Uruguay. Species richness is high around 20°N (Mexico), but much higher in the southern hemisphere, particularly in the Andean highlands between 8° and 20°S (Hijmans and Spooner, 2001). The wild potato species in sect. *Petota* have been grouped in 20 taxonomic series (Hawkes, 1990; Ochoa, 1999; Hijmans *et al*., 2002b).

Researchers of wild potato have often associated traits with certain species (e.g., Ross, 1986; Hawkes, 1990; Ochoa 1999). Geographical and ecological factors have also been associated with traits. For example, associations have been found between the altitude of origin and the frost tolerance in accessions of *Solanum acaule* (see Table 5.1 for full species names, including authors) (Li *et al*., 1980); between altitude and resistance to potato leaf hopper (Flanders *et al*., 1992); and between altitude and glycoalkaloid content (Ronning *et al*., 2000).

Van Soest *et al*. (1983) found that there was a concentration of wild potato species with cyst-nematode resistance near Potosí, Bolivia. Van Soest *et al*. (1984) concluded that wild potatoes with resistance to *Phytophthora infestans* (Mont.) De Bary occur near the tropics of Capricorn and Cancer. Flanders *et al*. (1992) found that species from hot and arid areas had resistance to Colorado potato beetle, potato flea beetle, and potato leafhopper. Species from cool or moist areas tended to be resistant to potato aphid. Flanders *et al*. (1997) found statistically significant differences between geographic areas for the presence of insect resistance in wild potatoes.

However, in most of these studies, factors that might explain geographic clustering of the level of a trait have been studied in isolation, and the validity of the associations found was not critically tested. For example, Van Soest *et al*. (1983) point out that in Bolivia, there are many wild potato species with cyst-nematode resistance near the city of Potosí. However, as this is also the area in Bolivia with highest species richness (Hijmans and Spooner, 2001), this would be expected when assuming a random distribution of this trait across species (wherever there are many species, there would be a relatively high number of species with a certain trait). In this case, evidence is insufficient as it is based on absolute, and not on relative numbers of species or accessions with a certain level of resistance. Another complicating factor is the rather limited size of the area of distribution of wild potato species, which makes it difficult, if not impossible, to rigorously separate the distribution of a trait over species, geographic, and ecological space (Flanders *et al*., 1997).

Frost damage is an important constraint in potato production at high latitudes and in high areas at low altitudes. Compared to *Solanum tuberosum* L., the common cultivated potato, some wild potato species have high levels of frost tolerance (Li, 1977; Estrada, 1982; Barrientos *et al*., 1994; Vega and Bamberg, 1995). The objectives of the present study were to investigate the extent to which taxonomic, ecological and geographic factors can be used to predict frost tolerance in wild potato species is investigated. In contrast with previous studies we study these factors together, while attempting to disentangle their separate contributions.

5.2 Materials and methods

Frost tolerance data for wild potatoes reported by Vega and Bamberg (1995) were used in this study. They screened 2635 accessions from 101 species in one field experiment in 1992 in Sturgeon Bay, Wisconsin, USA. The plants were scored twice for frost damage, first after two light frosts of about –2°C and later after a more severe frost of –5°C. Frost damage was assessed visually, using a scale with six classes that could be consistently distinguished by visual inspection from 0 (no damage) to 6 (all leaves and stems killed). The damage was averaged over the two readings, and linearly transformed the data to percentages (a reading of 0 equals 0%, and a reading of 6 equals 100% survival). In this paper we refer to this percentage score as frost tolerance. We used these data because they represented the largest single trait screening dataset for wild potato that we could find, and because of the reasonable a priori expectation that the presence of frost tolerance could be predicted from temperature data.

The accessions were originally collected in 12 countries in the Americas, covering most of the distribution area of wild potatoes. We only used data from accessions for which we had geographic coordinates. The correctness of the coordinates was checked following procedures described by Hijmans *et al.* (1999). In the case of errors, coordinates were changed where possible but accessions were deleted when it was not possible to assign precise coordinates. This left a total of 1646 accessions from 87 species (Table 5.1). The species with most observations was *Solanum acaule* (*n*=320). On average there were 19 observations per species, but the median was only 4 observations per species. The species belonged to 17 series with an average of 5.1 and a median of 3.0 species per series.

The difference between mean and median was largely due to the representation of the ser. *Tuberosa*, for which data for 36 species were available. The second and third largest series, in terms of species for which data were available, were *Pinnatisecta* (seven species), *Conicibaccata* and *Megistacroloba* (six species).

Species and Series

The species names used by Vega and Bamberg (1995) are all still current according to the review by Spooner and Hijmans (2001), except for *S. polytrichon* that was renamed *S. wightianum* (Hijmans *et al.*, 2002b). Series membership follows Hawkes (1990) except for subsequent changes for Peruvian species by Ochoa (1999). For the analysis in this paper we treat sect. *Etuberosa* as a series in sect. *Petota*.

Average frost tolerance was calculated over species and series. χ^2 tests were used to test the hypothesis that the occurrence of frost tolerance and species or series are associated. The average frost tolerance of a series was compared with that of its constituent species using linear regression.

Table 5.1 Number of observations and average frost tolerance for 87 wild potato species from 17 series.

Series	Species	obs[1]	Frost tolerance[2]
Acaulia Juz.	*S. acaule* Bitter	320	100
	S. albicans (Ochoa) Ochoa	7	100
Bulbocastana (Rydb.) Hawkes	*S. bulbocastanum* Dunal	22	10
	S. clarum Correll	3	17
Circaeifolia Hawkes	*S. circaeifolium* Bitter	2	21
Commersoniana Bukasov	*S. commersonii* Dunal	27	99
Conicibaccata Bitter	*S. agrimonifolium* Rydb.	2	17
	S. chomatophilum Bitter	1	83
	S. colombianum Bitter	4	27
	S. moscopanum Hawkes	1	33
	S. subpanduratum Ochoa	2	17
	S. tundalomense Ochoa	2	42
Cuneoalata Hawkes	*S.* ×*blanco-galdosii* Ochoa	2	33
	S. infundibuliforme Phil.	78	42
Demissa Bukasov	*S. brachycarpum* Correll	25	16
	S. demissum Lindl.	83	92
	S. hougasii Correll	5	13
	S. schenckii Bitter	3	28
Etuberosa Juz.	*S. etuberosum* Lindl.	22	82
	S. fernandezianum Phil.	1	33
	S. palustre Poepp.	63	58
Lignicaulia Hawkes	*S. lignicaule* Vargas	1	8
Longipedicellata Bukasov	*S. fendleri* A. Gray	33	21
	S. hjertingii Hawkes	4	42
	S. papita Rydb.	14	19
	S. stoloniferum Schltdl. and Bouchet	80	18

	S. wightianum Rydb.	31	17
Megistacroloba Cárdenas and Hawkes	*S. boliviense* Dunal	7	56
	S. dolichocremastrum Bitter	2	29
	S. megistacrolonum Bitter	94	83
	S. raphanifolium Cárdenas and Hawkes	16	56
	S. sanctae-rosae Hawkes	4	86
	S. sogarandinum Ochoa	1	67
Morelliformia Hawkes	*S. morelliforme* Bitter and G. Muench	4	0
Pinnatisecta (Rydb.) Hawkes	*S. brachistotrichum* (Bitter) Rydb.	13	23
	S. cardiophyllum Lindl.	14	22
	S. jamesii Torr.	7	20
	S. nayaritense (Bitter) Rydb.	1	17
	S. pinnatisectum Dunal	12	20
	S. tarnii Hawkes and Hjert.	1	50
	S. trifidum Correll	2	17
Piurana Hawkes	*S. acroglossum* Juz.	3	0
	S. albornozii Correll	2	8
	S. hypacrarthrum Bitter	1	8
	S. paucissectum Ochoa	2	92
Polyadenia Bukasov	*S. lesteri* Hawkes and Hjert.	2	8
	S. polyadenium Greenm.	10	20
Tuberosa (Rydb.) Hawkes	*S. acroscopicum* Ochoa	1	50
	S. alandiae Cárdenas	6	17
	S. ambosinum Ochoa	1	33
	S. andreanum Baker	3	14
	S. avilesii Hawkes and Hjert.	1	17
	S. berthaultii Hawkes	37	19
	S. brevicaule Bitter	11	38
	S. bukasovii Juz.	33	71
	S. cajamarquense Ochoa	1	25
	S. candolleanum P. Berthault	2	63
	S. ×*doddsii* Correll	1	17
	S. gandarillasii Cárdenas	2	17
	S. hoopesii Hawkes and K.A. Okada	1	42
	S. huancabambense Ochoa	2	38
	S. immite Dunal	1	25
	S. incamayoense K.A. Okada and A.M. Clausen	5	53
	S. kurtzianum Bitter and Wittm.	44	33
	S. leptophyes Bitter	137	49

	S. marinasense Vargas	6	18
	S. medians Bitter	3	17
	S. microdontum Bitter	14	17
	S. multiinterruptum Bitter	8	24
	S. neocardenasii Hawkes and Hjert.	1	17
	S. neorosii Hawkes and Hjert.	1	33
	S. okadae Hawkes and Hjert.	12	17
	S. oplocense Hawkes	38	27
	S. orophilum Correll	2	21
	S. pampasense Hawkes	2	46
	S. scabrifolium Ochoa	1	17
	S. sparsipilum (Bitter) Juz. and Bukasov	39	29
	S. spegazzinii Bitter	40	38
	S. ×*sucrense* Hawkes	20	31
	S. tarapatanum Ochoa	1	42
	S. venturii Hawkes and Hjert.	1	17
	S. vernei Bitter and Wittm.	23	57
	S. verrucosum Schltdl.	19	24
	S. vidaurrei Cárdenas	12	56
Yungasensa Correll	*S. arnezii* Cárdenas	2	8
	S. chacoense Bitter	28	31
	S. tarijense Hawkes	48	17
	Total	1646	

[1] obs = number of accessions tested.

[2] Frost tolerance = percent non-damaged tissue after two frosts at -2°C and one at -5°C in Sturgeon Bay, Wisconsin, USA in 1992.

Geography

A map of square 100 × 100 km grid cells was made of the average and maximum observed frost tolerance per grid cell, using the DIVA-GIS software (Hijmans *et al.*, 2002a). Calculations were made with each individual observation weighted equally, irrespective of species or series. We used a Lambert equal-area azimuthal projection, with 80°W as the central meridian and the equator as the reference latitude.

To investigate spatial autocorrelation, Moran's *I* statistic was calculated for the grid using IDRISI v32 software (Clark Labs, Worcester, MA, USA). We used the king's case (i.e., considering the eight adjacent cells of each cell). Positive spatial autocorrelation means that locations close to each other are more similar than locations farther apart, and negative autocorrelation means the opposite. Moran's I is basically positive for positive autocorrelation and negative for negative autocorrelation, except that the expected value for no autocorrelation is slightly negative (Bonham-Carter, 1994).

The degree to which taxonomic and geographic factors are confounded was investigated in two ways. We plotted the average observed frost tolerance in a grid cell versus the relative amount of observations from the nine most frost-tolerant wild potato species in that cell. We also plotted the drop in average frost tolerance in a grid cell when these most tolerant species are not considered.

Temperature

The ecological variable that a priori would be expected to be associated most with frost tolerance is frost incidence. Because of a lack of frost incidence (or daily minimum temperature) data we used average monthly temperature data instead. Monthly minimum and maximum temperature data were assigned to all accessions, using data on interpolated climate grids by Jones (1991) for Latin America and New *et al*. (1999) for the USA. For each accession, the average minimum and maximum temperature during the year, and during the estimated growing season, were calculated.

The dominant growing season for all locations where wild potatoes have been observed was estimated using the wild potato distribution database described by Hijmans and Spooner (2001). For a grid of 50 by 50 km cells, and a circular neighborhood with a 100 km diameter, the mode (most frequent observation) of the month of collecting was determined, using DIVA-GIS. A modal filter was then used to remove spatial outliers, using IDRISI. Some manual editing was carried out, for example for the coast of Peru, where a small strip (with 'lomas' vegetation) has a different growing season than that in the nearby Andes (Spooner *et al*., 1999). Because most wild potatoes are collected as seed, the growing season was assumed to be the month of collection and the three previous months.

Temperatures were plotted against frost tolerance. The GLM procedure (Type III sum of squares) in SAS was used to determine whether there were effects of temperature, species and series on frost tolerance, and whether there were temperature effects on frost tolerance within species and within series. A map was made predicting the presence of frost-tolerant wild potatoes on the basis of temperature using the climate data described above.

5.3 Results

Species and Series

There is a significant (χ^2=1508; df 86; p <0.001) association between species and frost tolerance (Fig. 5.1). Few species have high frost tolerance, whereas many species have a relatively low frost tolerance. Out of the 87 species, only 5 species have a frost tolerance score of more than 90%, and 9 species have a score of 82% or higher. Highest frost tolerance was for *S. acaule* (100%; n=320 observations) and *S. albicans* (100%; n=7), and for *S. commersonii* (99%; n=27).

There also is a significant association between series and frost tolerance (χ^2=357; df 17; p<0.001). Species from ser. *Acaulia* and *Commersoniana* had the highest frost tolerance, followed at some distance by ser. *Megistacroloba*, and sect. *Etuberosa* (Fig. 5.2). For most series there are important differences in frost tolerance among the constituent species (Fig. 5.3), as illustrated by the low coefficient of determination (R^2) of 0.51 between the average (over species) frost tolerance by series versus that of the series' constituent species. Variation is particularly high in ser. *Concibaccata*, *Commersoniana*, *Demissa*, *Piurana*, and *Tuberosa*, and in sect. *Etuberosa*, which are all series with an intermediate to low average frost tolerance (Fig. 5.2).

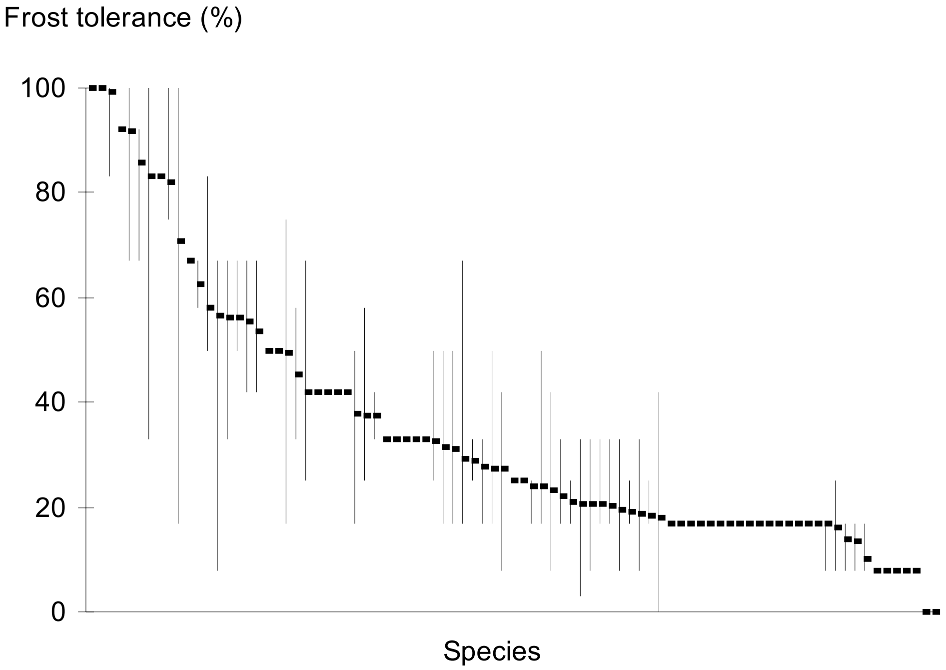


Figure 5.1 Mean frost tolerance (percentage non-damaged tissue) by wild potato species (lines indicate maximum and minimum values). Data for 87 species and 1646 accessions (average of 19, and median of 4 accessions per species).

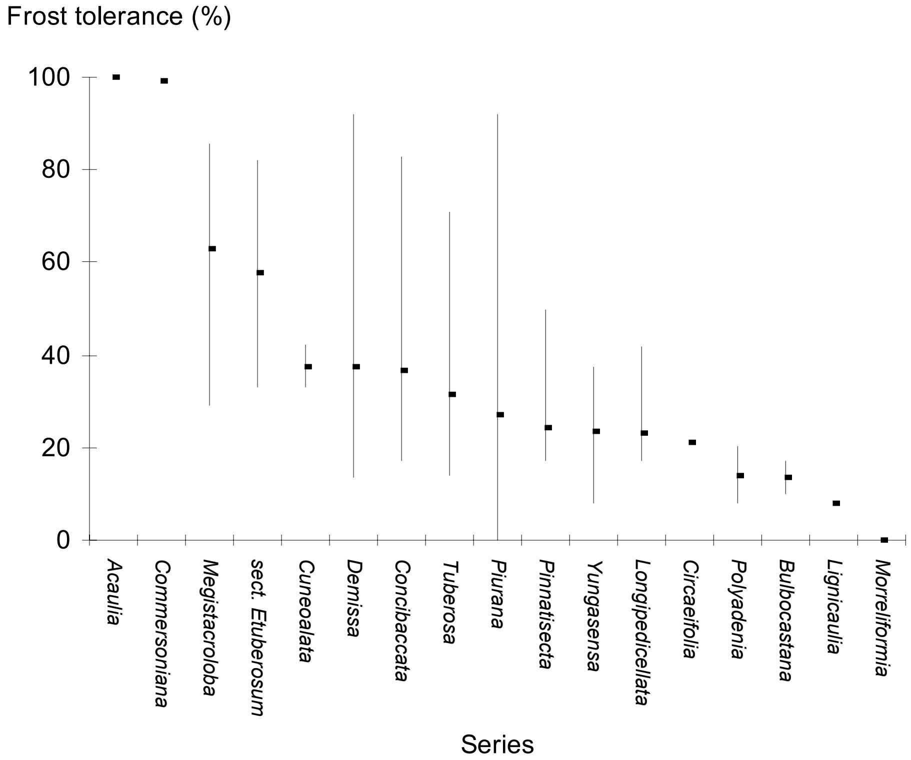


Figure 5.2 Mean frost tolerance (percentage non-damaged tissue) by wild potato series (lines indicate maximum and minimum values). Data for 17 series containing 5 species on average (in our sample).

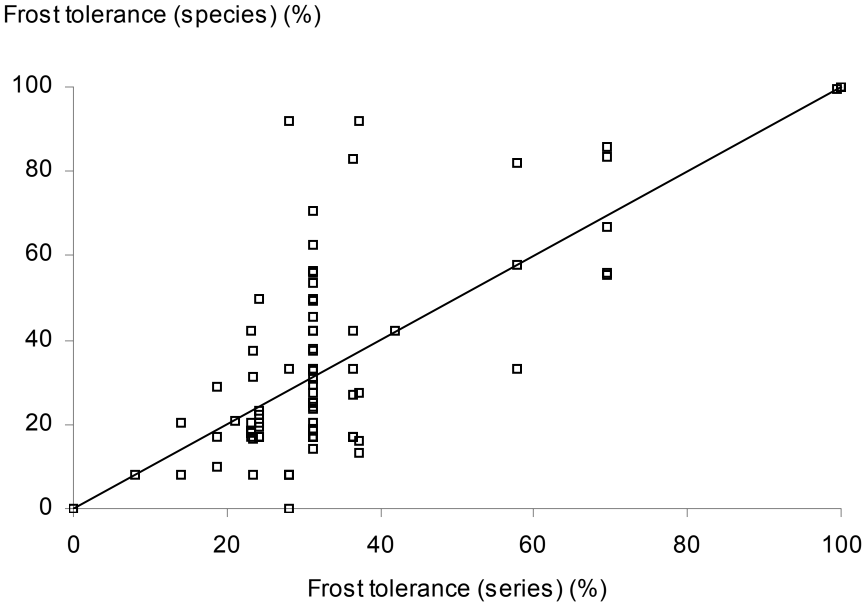


Figure 5.3 Mean frost tolerance (percentage non-damaged tissue) of wild potato series (averaged over species) versus the frost tolerance of the species in those series. R^2 = 0.513. Data for 18 series containing 4.8 species on average (in our sample). Regression line: $y = x$ (by definition).

Geography

Our sample contains wild potatoes from most of the areas where wild potatoes occur (Hijmans and Spooner, 2001). Important exceptions are areas in Central America, Colombia, and Ecuador (Plate 2). Areas where high mean levels of frost tolerance were observed constitute a zone from central to south Peru and a small part of adjacent northern Bolivia; and a zone stretching south and east from Paraguay into adjacent Argentina, Brazil and Uruguay, and a small area in central Chile (Plate 2A). Northern Argentina has a zone with an intermediate level of frost tolerance. There are many grid cells with a high maximum observed frost tolerance (> 99%; Plate 2B) that did not have a very high mean frost tolerance (< 75%; Plate 2A). These cells are in central Mexico, north-central Peru, Bolivia, and north Argentina.

Spatial autocorrelation between grid cells is positive and highly significant. Moran's $I = 0.521$ for mean frost tolerance, and $I = 0.498$ for maximum frost tolerance (number of cells = 185, $p<0.001$ in both cases). This indicates the presence of geographic clustering of areas with wild potatoes that have similar levels of frost tolerance.

Geography and Species

The relative abundance of one or more of the nine most tolerant species (above 80%; Fig. 5.1) in a grid cell is a reasonably good predictor of overall frost tolerance in that grid cell (Fig. 5.4). When discarding these nine most frost tolerant species, there are cells in which the average frost tolerance decreases sharply. This decrease is conspicuous in areas with medium to high levels of frost tolerance. This was to be expected because there cannot be much decrease in areas where tolerance is already low. Nevertheless, some of the cells with very high levels of frost tolerance maintain high levels even without the most tolerant species (Fig. 5.5). These cells are all located in southern Peru.

Solanum acaule, the most frost tolerant species in our sample, is also one of the most common and widespread wild potato species (Hijmans and Spooner, 2001). The high maximum tolerance scores observed across Peru, Bolivia, and north Argentina (Plate 2B) largely coincide with the distribution of this species. The distribution of mean frost tolerance (Plate 2A) is less associated with the presence of *S. acaule*.

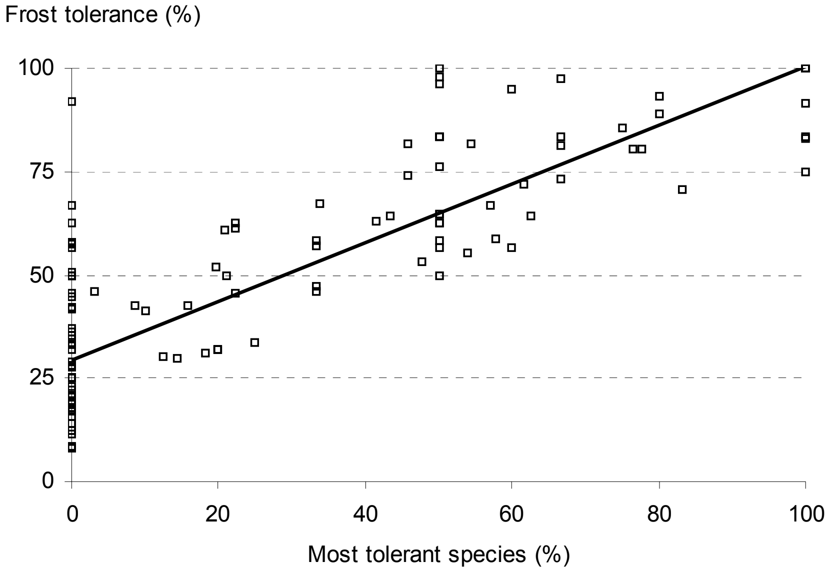


Figure 5.4 Percentage of observations per grid cell for the nine most frost tolerant wild potato species (out of a total 87 species in the sample) versus the average frost tolerance (percentage non-damaged tissue) for all observations. Each dot on the graph represents one 100 by 100 km grid cell. Regression line: $y = 0.71x + 29.4$; $R^2 = 0.767$.

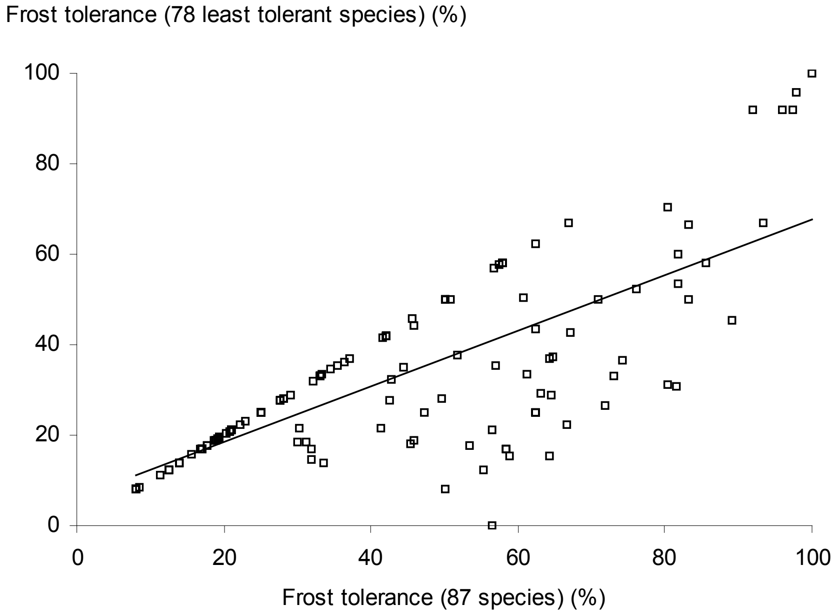


Figure 5.5 Mean frost tolerance (percentage non-damaged tissue) per grid cell with and without the nine most tolerant wild potato species (out of 87 species in the sample). The number of observations decreased from 1646 to 1086. Each dot on the graph represents one 100 by 100 km grid cell. Regression line: $y = 0.61x + 6.45$; $R^2 = 0.589$.

Temperature

Temperature data of the areas where the accessions were collected predicts frost tolerance well at low levels of tolerance. If only the scores between 0 and 42% are considered, there is a strong association between yearly average minimum temperature and frost tolerance (R^2=0.87; Figure 5.6). In this range, wild potatoes collected from warmer places are less likely to have frost tolerance than those collected from colder places. However, at higher levels of tolerance (between 42 and 100%), there is a weak (R^2=0.31) and even positive association (more frost tolerance in wild potatoes from warmer areas). Thus there is only a weak overall association between frost tolerance and temperature. The results do not change whether minimum or maximum temperature is used, or whether temperature is averaged over the whole year or only over the apparent growing season.

Within series or species, the predictivity of temperature data was low. For the series, the coefficient of determination between frost tolerance and average annual minimum temperature was below 0.14 and not significant (p>0.05) except for ser. *Conicibaccata* (R^2=0.88), *Piurana* (R^2=0.54; but with a positive relationship between temperature and frost tolerance!), *Demissa* (R^2=0.38), and *Tuberosa* (R^2=0.18). For individual species, there was a significant association (p<0.05) of minimum temperature with frost tolerance for only two species: *S. colombianum* (R^2=0.95) and *S. verrucosum* (R^2=0.45).

Temperature, Species and Geography

In a single factor statistical analysis, average annual minimum temperature is significantly associated with frost tolerance (F=279; df=1; p<0.001; R^2=0.15). However, species (F=304; df=86; p<0.001; R^2=0.94) and series (F=345; df=17; p<0.001; R^2=0.78) effects are much stronger. In combined models of species or series and minimum temperature, the minimum temperature effect is not significant. Very similar results were obtained when all the observations with 100% frost tolerance scores were deleted, in order to have a more normal distribution of the data.

The map of areas with an average minimum temperature below 3°C (typical for frost tolerant accessions, Fig 5.6) 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 (Fig. 5.7). The map correctly identifies the Central Andes and parts of Chile, but it misses the lowland area of Argentina and Uruguay (where *S. commersonii* occurs) and wrongly predicts the southwestern USA as an area with frost tolerant wild potatoes. The species present in the United States are *S. jamesii* (7 accessions; 20% frost tolerance score) and *S. fendleri* (33 accessions; 21% frost tolerance score).

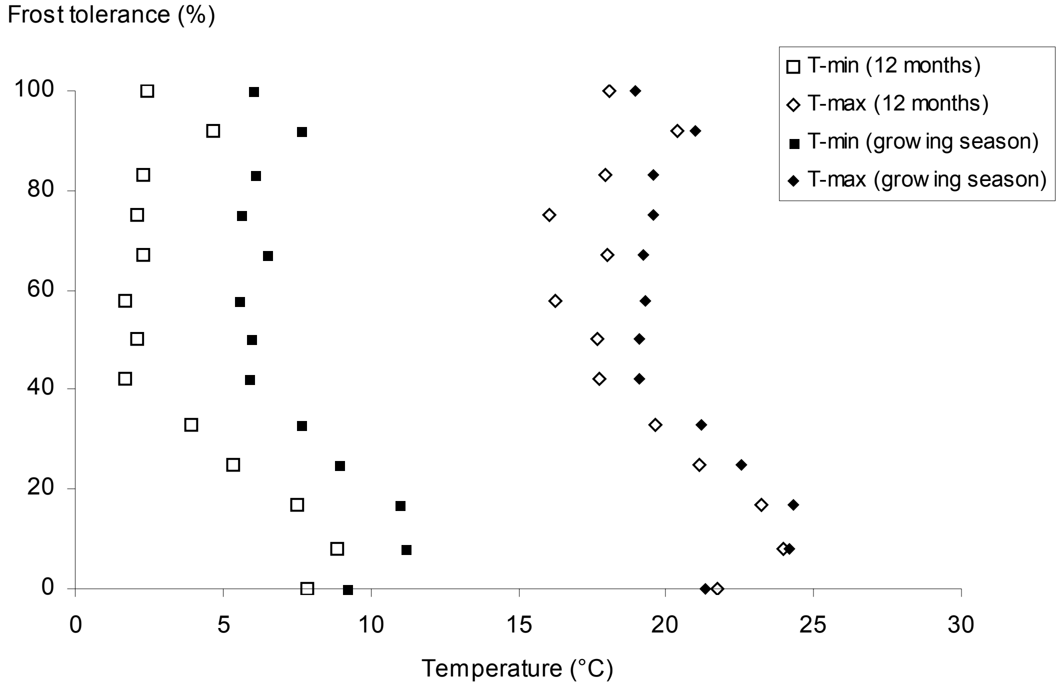


Figure 5.6 Average frost tolerance versus temperature of the locations where the wild potato accessions in our sample were collected, for the monthly minimum temperature (T-min), maximum temperature (T-max) averaged over the year (12 months), or over the most likely growing season of four months (growing season).

5.4 Discussion

This study illustrates the need for simultaneously analyzing different factors to predict the presence of an agronomic trait in wild species. We showed that taxonomic categories, particularly species, are strongly related to frost tolerance in wild potatoes. Ecology and geography were also associated with frost tolerance, but probably not to the extent that they would be very useful to guide further screening or collecting. If further screening of wild potatoes is warranted, priority could be given to those species not yet tested, particularly those from series with high frost tolerance, and that occur in southern Peru, and to species with a high level of frost tolerance for which only a few accessions were tested (e.g., *S. paucissetcum*). It would also be important to test the most tolerant accessions (particularly those with 100% frost tolerance scores) to more severe frosts.

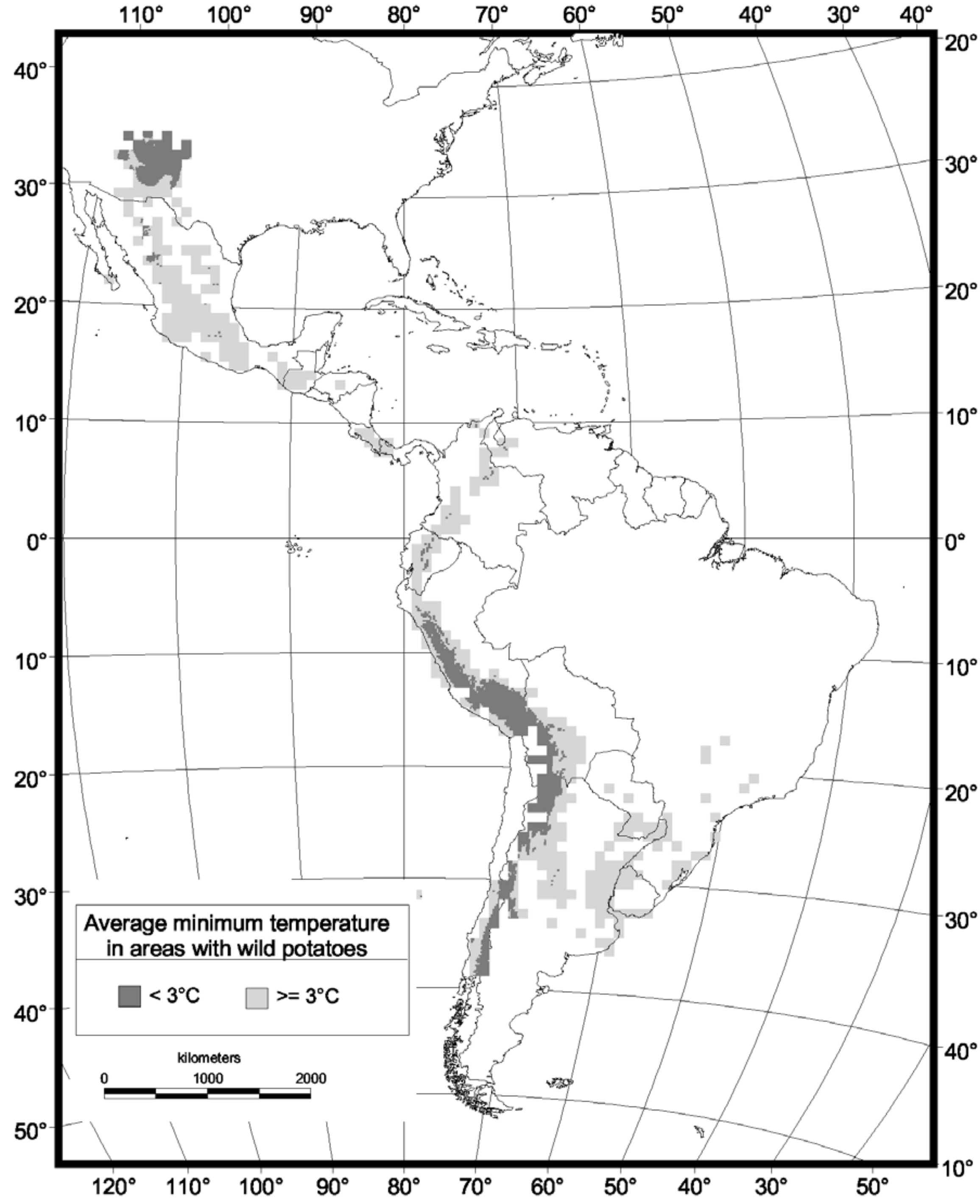


Figure 5.7 Average annual minimum temperature in areas where wild potatoes have been observed.

Species were associated with different levels of frost tolerance. This is of value for prediction within, but not beyond the species examined. The latter was important in our study for which data for fewer than half the existing wild species were available. Taxonomic series might be used for prediction of frost tolerance in species that have not yet been tested. However, there was much variation within those series with a relatively high number of observations. Other series may be too small to be of much use in prediction. For example, we examined data for two species (*S. acaule* and *S. albicans*) of ser. *Acaulia,* the most frost tolerant series. There are only two additional species in this group *S.* ×*indunii* and *S.* ×*viirsoii*, both of which are rare and assumed to be of hybrid origin, with *S. acaule* as one of the parents. The other series associated with very high frost tolerance, ser. *Commersoniana*, only has two species: *S. commersonii*, for which we had data, and *S. calvescens*, for which we did not.

The species and series conclusions also depend on the accuracy, consistency, and relevance of the taxonomic system used. Ongoing research is continuing to refine wild potato taxonomy, and in the future there will likely be a reduction in the number of species and changes in series memberships (Van den Berg *et al*., 1998; Spooner and Hijmans, 2001). The commonly used series classification by Hawkes (1990) has received little support in any molecular marker data set used to date (Spooner and Hijmans, 2001). For example, chloroplast DNA restriction site data supported only four clades within sect. *Petota*, and not the 19 series (of 21) that were examined (Spooner and Sytsma, 1992; Spooner and Castillo, 1997). Other nuclear molecular markers such as nuclear restriction fragment length polymorphisms (Bonierbale *et al*., 1990) and amplified fragment length polymorphism data (Debener *et al*., 1990; Kardolus, 1998) also fail to support many traditional series. The predictivity of series classifications may be greatly improved with new data. For example, there is compelling support for including *S. demissum* in ser. *Acaulia*, rather than in ser. *Demissa* (Spooner *et al*., 1995; Spooner and Hijmans, 2001). It is noteworthy that *S. demissum* had a high frost tolerance (92%), comparable to the members of ser. *Acaulia*, but unlike the other members of ser. *Demissa*.

Our findings corroborate the weak association of genetic variation with eco-geographic factors (Chapman, 1984; Peeters *et al*., 1990; Del Rio *et al*., 2001). However, our findings are discordant with the results of Li *et al*. (1980), who found a strong and simple positive linear relation between frost tolerance and altitude of origin for 15 accessions of *S. acaule*. Perhaps this relation is specific to *S. acaule*, but we could not test this because we did not have differentiation within *S. acaule* (for which all accessions had a 100% frost tolerance score).

We had many observations originating from a large geographic area. The frost tolerance screening data used here constitutes the largest single evaluation of the US Department of Agriculture potato genebank database. Yet data quality problems in the coordinates or in the interpolated climate data may have hampered our ability to find relationships between temperature and frost tolerance to some extent, particularly because most of the wild potatoes that were evaluated occur in highly dissected mountain ranges, with steep climate gradients. Detailed studies with a few species from a small area and with more precise temperature data would be useful to validate our results. We used average monthly minimum temperature instead of perhaps more informative data such as monthly extreme temperatures. This may have weakened the relationships found. Yet a likely more important factor is the low resolution of the interpolated weather data. This may be particularly problematic in the Andes where there are large changes in altitude (and hence temperature) over relatively small distances. In all areas there may be micro-climatic differences that play a role which cannot be captured with our data. Nevertheless, we used the best data available, and these data did not have a strong predictive value.

In previous studies, altitude has sometimes been used as a proxy for ecological factors (Flanders *et al.*, 1992). However, altitude is only a good proxy for temperature in small areas. It is less useful for studies of large areas such as considered in the present study. Also, variation is sometimes compared using countries or groups of countries (Peeters *et al.*, 1990). These areas are often rather different in size and shape and may not allow for appropriate comparison. Grid cells, as used in this study, are more appropriate for these comparisons.

Other trait/crop combinations could be studied to re-evaluate the emphasis that is typically given to eco-geographic stratification in genetic resources collection (e.g., Brown and Marshall, 1995; Von Bothmer and Seberg, 1995). It is noteworthy that wild potato collectors in particular have given much more emphasis to taxonomic considerations (Spooner *et al.*, 1999). Yet this taxonomic bias is compensated by the fact that many wild potato taxa occur in small areas. Consequently taxonomy and geography (and hence also ecology) of wild potatoes are somewhat confounded. With some species being differentiated on the basis of the area they occur (Spooner and Van den Berg, 1992), sampling by species is to some extent also sampling by geographic area.

Part II

Agro-ecological zoning

6. Global distribution of the potato crop

Published as: Hijmans, R.J., 2001. Global distribution of the potato crop. American Journal of Potato Research 78: 403-412.

Abstract

The global distribution of potato area is described using country level statistics and a new geo-referenced database. There are two main peaks in global potato distribution by latitude. The major peak is between 45°N and 57°N and represents potato production zones in the temperate climates where potato is a summer crop. The other peak is between 23°N and 34°N, mainly represents production zones in the subtropical lowlands, where potato is a winter crop. Between 1950 and 1998 potato production area increased at low latitudes and decreased at high latitudes, particularly around 53°N (this zone includes parts of Belarus, Germany, Poland, Russia and Ukraine). The northern limit of potato production coincides with the boundaries of agriculture and the presence of human population. The peak between 23°N and 34°N coincides with the area of highest population density (per area of land and per area of arable land). About 25% of the global potato area is in the highlands (above 1000 m).

6.1 Introduction

Crop distribution data are readily available at the country level through the database of the Food and Agriculture Organization of the United Nations (FAO, 2000a). To study the distribution of a particular crop at the global or at lower aggregation levels, however, more disaggregated data are needed to account for the considerable differences within countries in crop distribution and other variables such as climate. While databases have been compiled for some other factors such as global altitude (USGS-EDC, without date), population (CIESIN, 2000), and climate (e.g., New *et al*., 1999), crop distribution databases at a lower aggregation level than country are an important missing link for studies of global agriculture.

For the major crops, distribution maps may exist, but these are often outdated. Another shortcoming of these maps is that they are generally not based on digital databases, which complicates analysis and further use in research. Recent work that aims to fill this important information gap is reported by Huke (1982) for rice in Asia, Carter *et al*. (1992) for cassava in Africa, Wortmann *et al*. (1998) for common bean in Africa, Hyman (1999) for all major crops in Latin America, and Huaccho and Hijmans (2000) for sweetpotato worldwide. Global potato distribution maps were published in Van Royen (1954), the Oxford Economic Atlas of the World (EIU and CDCP, 1954, 1972), and Bertin *et al*. (1971). Finch and Baker (1917) published what is probably the first map of global potato distribution. Their map is rather incomplete, however, despite detailed data for Europe and North America. For example, it does not include any potato area in China, India, and in the Andes.

In this paper a new global geo-referenced database of the distribution of potato (*Solanum tuberosum*) is described. An earlier version of this database (Huaccho and Hijmans, 1999) was used for a dot map in Walker *et al*. (1999). The database was developed at the International Potato Center (CIP), a research institute that seeks to improve potato production in the developing world. To be able to do so effectively, CIP needs base-line information on where potatoes are produced and what the major constraints are in those areas. Agro-ecological zoning studies can be useful in this respect (e.g., FAO, 1978-81; Stol *et al*., 1991), but the relative importance of different zones cannot be assessed accurately without high-resolution potato distribution data. The database described here is suitable for that purpose as was illustrated by a study of global late blight severity (Hijmans *et al*., 2000a).

In the present study the potato distribution database is used to provide a detailed description of the current distribution of the potato crop, and relate this to the other factors (latitude, altitude, land area and cropland area, and population). Haverkort (1990) described potato production for the zone between 0 and 50°N and 20°W and 50°E. This paper covers global potato production, and focuses on a quantitative description of the distribution of production area, while largely following Haverkort's organizing principle of latitude and altitude. Changes in potato area over time, by latitude, are described as well. The present study is intended to provide base-line data for further research into the current and possible future distribution of the potato, its agro-ecology, and production constraints. Walker *et al.* (1999) described aspects of the economic geography of global potato production, emphasizing the effect of diversification in consumption and specialization in production. Scott *et al.* (2000) provided projections of future global potato production.

6.2 Materials and methods

Country level data.

Potato area, relative potato area over cropland, and potato area per capita were computed at the country level. The FAO database (FAO, 2000a) was used for potato area, and cropland area. Country level population data were taken from ESRI (1999).

Geo-referenced database of potato area.

For each country of the world where potato is produced, a geo-referenced database was made of potato distribution using ArcView-GIS version 3.1 (ESRI, Redlands, CA, USA) software. Potato area statistics were obtained from national and some international sources, such as the European Union's statistical office (EUROSTAT), and from the GIEWS/GeoWeb database (FAO, 2000b). The latter was the source for 47 countries, mainly countries in Africa and Asia with little potato area. Potato production zones were delineated following administrative boundaries, or followed specific potato production zones taken from other maps, or by using a combination of both. Administrative boundaries were taken from the ADMIN98 global database that comes with ArcView-GIS 3.1, a database for Africa provided by Corbett and O'Brien (1997), and various other country-level sources.

Potato area by administrative units were used for those countries for which highly disaggregated statistical data (i.e., for small units) were available. For example, for the USA data were available at county level for most states, and for the European Union statistical data were available at the NUTS-2 level (NUTS-0 is country level, NUTS-2 is two levels lower). Other countries with highly disaggregated data included Brazil, China, India, and Peru.

For many other countries, however, only data aggregated at large administrative units were available. To increase precision, the spatial units from these sources were adapted using other information such as previous maps, particularly those from Rhoades (1987). When administrative units were very small, or had a potato area less than 100 ha, they were in some cases aggregated into larger areas. For example, for Peru district-level data was aggregated by province, and for Brazil, municipality data was aggregated by state (but excluding spatial units without potato area). For some countries, the location of potato production zones was available, but not the distribution of potato area over these zones. In those cases, it was assumed that the total national potato production area was evenly distributed over the potato production zones.

To create a consistent database representative for one time period, the data was adjusted using average country-level potato area for 1997-1999, according to FAO (2000a). First the fraction of the national potato area in each production zone was estimated, using the national statistical data when available. This fraction was then multiplied by the FAO estimate of the total national potato area for 1997–1999. The only country for which the FAO estimate was not used was Malawi. Only about 20% of the potato area reported for Malawi by FAO actually is potato, the remainder being sweetpotato (Peter Ewell, CIP, personal communication 1999). The potato area in Malawi was estimated at 10,000 ha.

<u>Spatial and temporal change.</u>

The change in spatial allocation of potato area over time was assessed. Data from three time periods, about 25 years apart, were used: 1948—52, 1974—76, and 1997—99. Hereinafter, these time-periods are referred to as 1950, 1975, and 1998, respectively. FAO estimates of the potato area at the country level were used. For the within-country distribution the relative distribution data for 1998, as described in the previous section were employed. Hence, the assumption was made that there was no change in the relative potato distribution over time within countries (and within groups of countries if a country split into smaller countries during these time periods, such as happened with the Soviet Union).

<u>Data conversion to grid.</u>

The polygon-based geo-referenced potato area data were transferred to a grid. Relative potato area over total land area (RPA) was calculated for each production zone (polygon), using IDRISI (Clark Labs, Worcester, MA, USA) software. RPA data were transferred to a 1 by 1 minute grid, which was subsequently aggregated to a coarser resolution grid (30 by 30 minute and 1 by 1 degree), calculating the average RPA for each cell. This two-step process prevents data loss and creates a smooth grid. The 30 by 30 minute resolution seemed an adequate compromise between the desire for high resolution and the high uncertainty of some of the potato distribution data, which do not justify too high a resolution.

Description of potato area distribution

The 1 by 1 degree grid was used to summarize the data by degree latitude (see below). By multiplying the area of each grid cell with the RPA, the total potato area for a grid cell was calculated. The potato area was computed for bands of 1-degree of latitude wide for the current situation (1998) and for 1950 and 1975. The average latitude of potato production was computed. For this computation, data for the northern and southern hemisphere were not separated.

Current potato area distribution was described in relation to the distribution of land area, cropland area, and population density. Relations between these variables were visually assessed. The global distribution of cropland data described by Wood *et al.* (2000) was used. They adapted the 30-second (~1 km2) resolution global land-cover database produced by the EROS Data Center of the US Geological Survey (USGS-EDC, 1998). Wood *et al.* (2000) included three primary agricultural land cover categories each with a different range of agricultural area intensity (30-40, 40-60, and greater than 60% agriculture that was simplified to 35, 50 and 80% respectively). Cropland area from two remaining categories, in which agriculture might occur on 30% or less of the area, was not included. Population data were taken from the Gridded Population of the World (GPW) database for the year 1995 (CIESIN, 2000). Altitude data were taken from the ETOPO5 (5-minute) grid database from the U.S. Geological Survey (USGS-EDC, without date).

6.3 Results

Country level distribution of potato area

Global potato area has decreased from 23 million ha in 1950 to 18 million ha in 1999, with a rather stable annual decrease of 92,000 ha per year on average. Whereas the potato area has decreased in the developed countries, it has increased in developing countries (Fig. 6.1). Thanks to overall increasing yields, total production has been stable at around 30 million tons per year.

About half the potato area is in Europe, while Asia has about one third, and most of the remainder is about equally divided between North America, South America and Africa (Fig. 6.2). Potato is grown in 149 of the 227 countries listed in the FAO database (FAO, 2000a). China currently is the country with the world's highest potato area and production, followed by Russia, Ukraine, Poland and India (Table 6.1). Among the major producing countries, Netherlands is most specialized (20% of the cropland is planted with potato), followed, distantly, by Belarus (11%), Poland (9%), Colombia (8%) and Peru (7%) (Table 6.1). Per capita potato area is very high in Belarus (650 m^2) and high in its neighboring countries Lithuania, Poland, Ukraine and Russia.

Note the low potato area per capita in the big Asian potato producing countries (China, India, Bangladesh). The relative importance of the potato crop in these areas, where potato is an increasingly important vegetable, does not compare to its role in eastern European countries, where potato is a staple crop, and also used for feed (Walker *et al.,* 1999). There is no African country with more than 100,000 ha of potato area. The largest producer in Africa is Egypt (82,561 ha). In Sub-Saharan Africa, South Africa (61,000 ha) and Uganda (52,000 ha) are the leading potato producing countries.

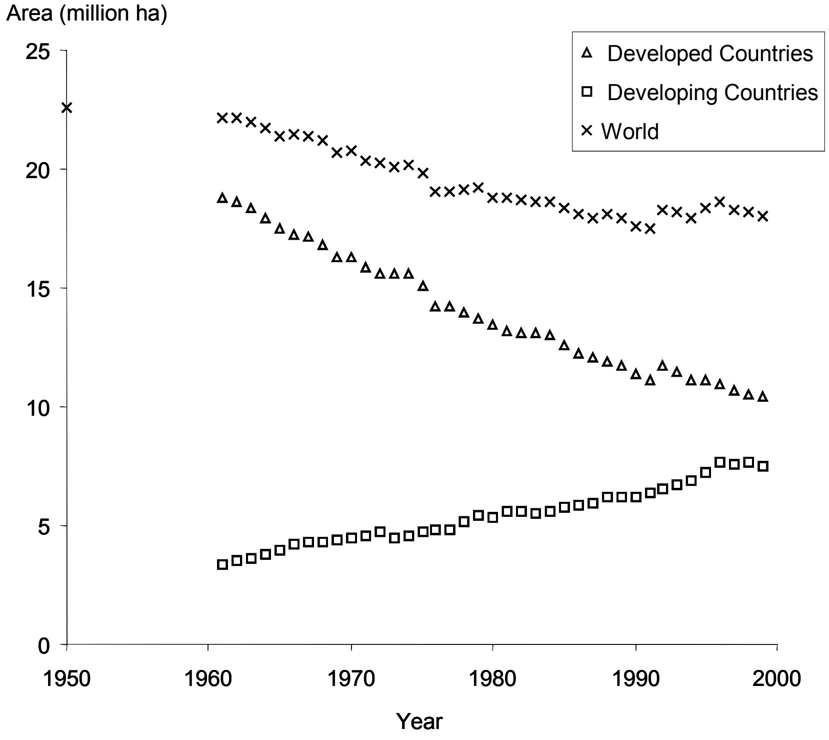


Figure 6.1 Global potato area over time: total and for developed and developing countries. Source: FAO, 2000a

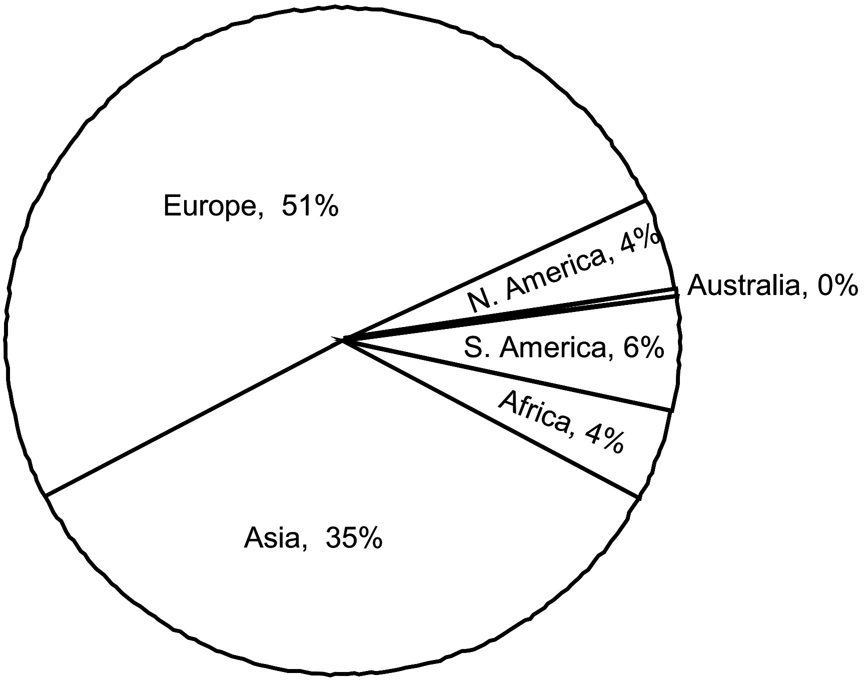


Figure 6.2 Distribution of potato area by continent. Source: FAO, 2000a.

Geo-referenced distribution of potato area.

The map of global potato distribution (Plate 3) confirms the country level statistics, but it provides more detailed information. A belt that goes from northwest to south China and continues into the plains of the Ganges river in Bangladesh and northern India dominates potato production in Asia. This is the area where the greatest expansion of potato production has taken place over the last decades (Walker *et al.*, 1999). In India and Bangladesh potatoes are mainly planted as a lowland winter crop. In south and central China, potato is a highland crop grown in all seasons (winter crops are only grown in the south), depending on altitude. In northern China, potato is a summer crop.

RPA is very high in the “traditional” European potato-growing areas, notably in Belarus, Poland, and Ukraine. In Western Europe there is a concentration of potato production around the North Sea, in Belgium, France, Netherlands, and United Kingdom, and in northwest Spain (Galicia) and Portugal. In North America potato production is much less contiguous than in Europe. The main zones of high RPA in South America are in the Andean highlands. RPA in Africa is generally low; highest in parts of North Africa, Ethiopia, the East African Highlands, and in South Africa.

Table 6.1 Potato area, potato area as percentage of total cropland area, and potato area per 1000 inhabitants by country (only for countries with more than 100,000 ha of potato).

Country	Potato area (1000 ha)[1]	Potato area as percentage of total cropland area[2]	Potato area per 1000 inhabitants (ha)[3]
China	3,430	2.8	2.7
Russia	3,289	2.6	21.7
Ukraine	1,534	4.7	28.9
Poland	1,290	9.2	34.0
India	1,253	0.8	1.4
Belarus	692	11.2	65.7
USA	548	0.3	2.1
Germany	300	2.5	3.7
Peru	263	7.2	10.7
Romania	262	2.8	11.1
Turkey	207	0.8	3.4
Netherlands	181	19.9	11.7
Brazil	177	0.3	1.2
United Kingdom	169	2.7	3.0
France	168	0.9	2.9
Colombia	167	8.0	4.8
Kazakhstan	165	0.6	9.6
Iran	161	1.0	2.5
Canada	155	0.3	5.4
Spain	142	1.0	3.6
Bangladesh	140	1.7	1.2
Bolivia	131	6.7	17.2
Lithuania	126	4.3	33.2
Argentina	115	0.5	3.4
Nepal	115	4.0	5.8
Japan	102	2.2	0.8

[1] Average for 1997-99. Source FAO (2000a)

[2] Calculated after FAO (2000a)

[3] Potato area from FAO (2000a), population data from ESRI(1999)

Distribution by latitude

Potato area is distributed between 47°S and 65°N. Only 6.9% of the potato area is in the southern hemisphere (Fig. 6.3). This is also illustrated by the data in Table 6.1: only four out of the 26 countries with more than 100,000 ha of potato are from the southern hemisphere. These four countries are from South America, and two of them (Bolivia and Peru) have most of their potato area in the Andes. Virtually all (90%) potato area is between 22°N and 59°N. There are two remarkable peaks in the distribution of potato area by latitude. The first peak is between 23°N and 34°N (19% of the total potato area), and represents the potato area in the Ganges plain, southern China, and Egypt (Plate 3 and Fig. 6.4), where potato is grown as a winter crop. The second and most important peak is between 44°N and 58°N (52% of the total potato area) and represents the summer crop production zones in western and eastern Europe: a band from the North Sea area through Poland, Belarus and Russia. The area in between these two peaks, also with a large potato area, is made up by parts of central and northern China, southern Europe and parts of the USA and Canada.

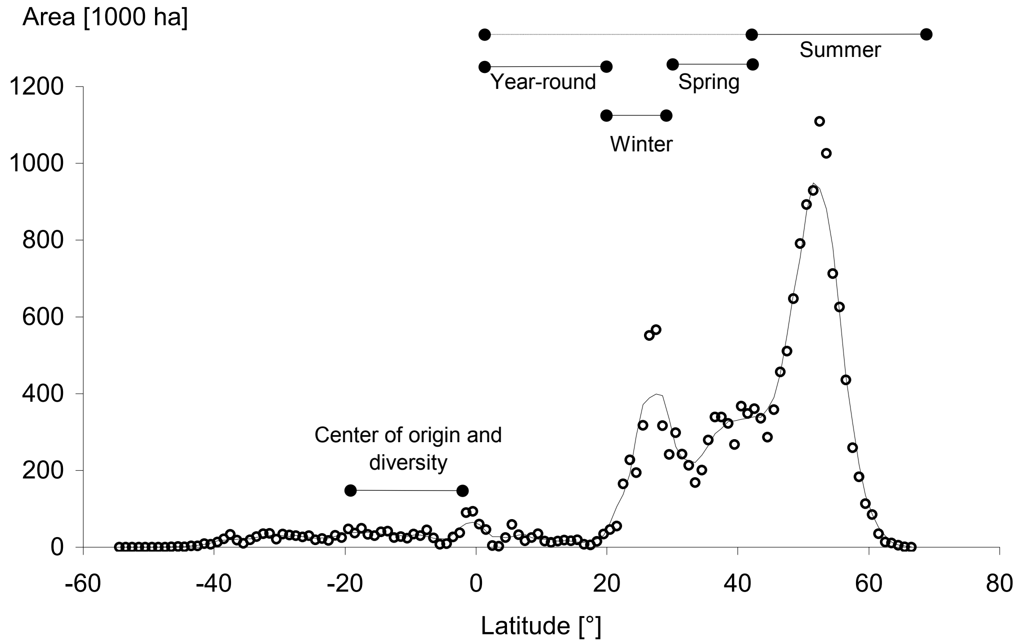


Figure 6.3 Global distribution of potato area by latitude. Each dot represents one-degree latitude. Latitude in the southern hemisphere is indicated with a minus (-) sign. The line is the five-observations moving average. The location of the center of origin and diversity (the Andes) is indicated. The dominant growing seasons are indicated for the northern hemisphere.

Change in distribution by latitude

Between 1950 and 1998, potato area increased at all latitudes except for the area between 33°S and 39°S and the area north of 42°N (Fig. 6.4). In the latter area, the decrease of potato area was dramatic. At 53°N (i.e., the area between 52.5 and 53.5 °N that includes parts of Poland, Belarus, Ukraine, and Russia), potato area decreased by 1.4 million ha over the past 50 years. This decrease is related to diversification in the diet and a decreased use of potato for feed (Walker *et al.*, 1999). Most increase of potato area was observed between 20°N and 40°N (China, India, North Africa). In relative terms, important expansion of potato area was also found in the area between 20°N and 8°S: at these latitudes where potato is mostly grown in the highlands, potato area doubled between 1950 and 1998.

The direction and amount of change was comparable over the periods 1950 to 1998 and the 1975 to 1998 periods. But change was strongest during the last 25 years: average latitude of potato production decreased from 47.5° in 1950 to 45.2° in 1975 and 41.2° in 1998.

Potato area, land and cropland area, and population

The dearth of potato area in the southern hemisphere is partly due to the relative absence of land and people compared to the northern hemisphere (Fig. 6.5). The southern hemisphere has 26% of the land area (excluding Antarctica), 12% of the global population, and 23% of the cropland area.

Between 60°S and 30°N, land area, cropland area and number of people both increase. Although there is a peak in population at 7-8°S (particularly on Java), global population increases sharply north of the equator. Population reaches a maximum around 30°N. This is also a peak in the potato area. This area includes parts of North Africa (Nile valley) and the Indo-Gangetic plain in India and Bangladesh where the potato is a winter crop in a multiple cropping system. Potato's potential to produce a lot of high-value food in a short time makes it an important component of the intensive agricultural systems in these areas of high population density (Walker *et al.*, 1999).

In the southern hemisphere, the relative amount of crop land devoted to potato peaks around 40°S and around the equator (Fig. 6.6). The latter area consists of potato area in the tropical highlands of Ecuador, Colombia, Uganda and Kenya, where the crop is grown year round.

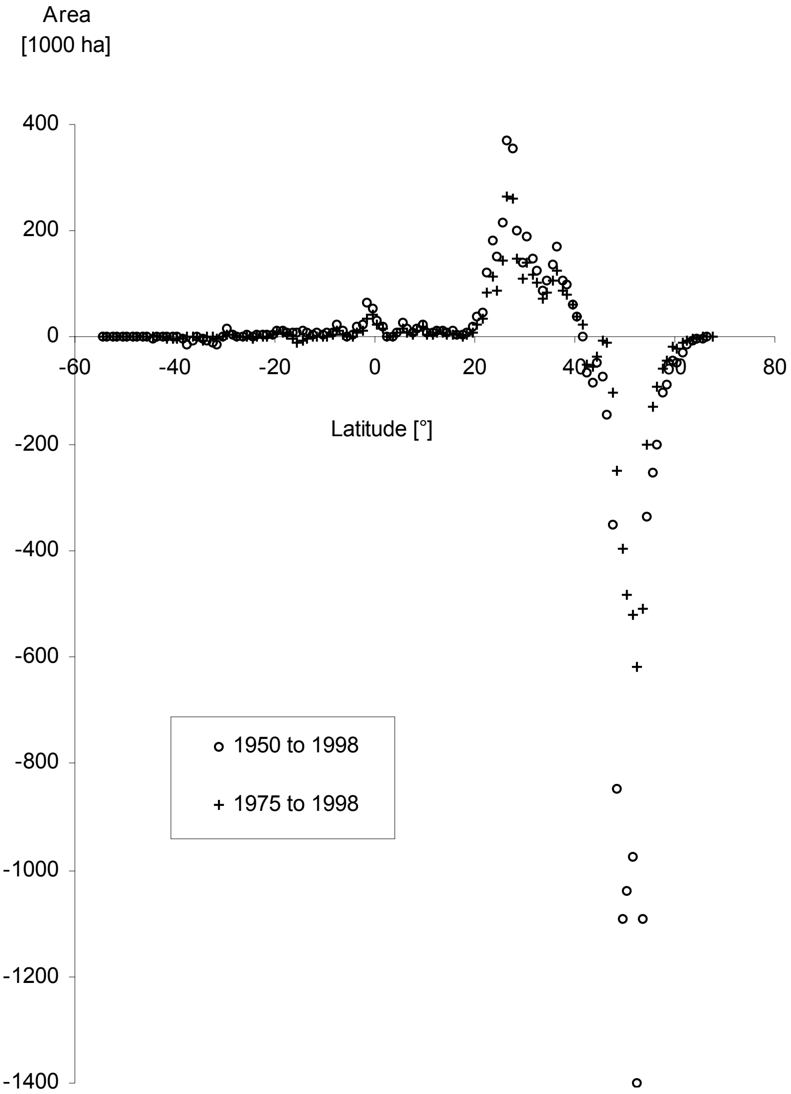


Figure 6.4 Change in potato area between 1950 and 1998, and between 1975 and 1998, by latitude. Each dot represents one-degree latitude. Latitude in the southern hemisphere is indicated with a minus (-) sign.

The zone between 30°N and 45°N is one of transition. Population goes down while land area and the percentage with crops goes up. North of 45°N, the potato area increases rapidly and peaks at 55°N, despite a decrease in population. This peak consists of parts of Canada, and the North Sea region, but particularly important is the “potato-basket” of Poland, Belarus and western Russia. North of 55°N, potato area falls, together with the total agricultural area and population.

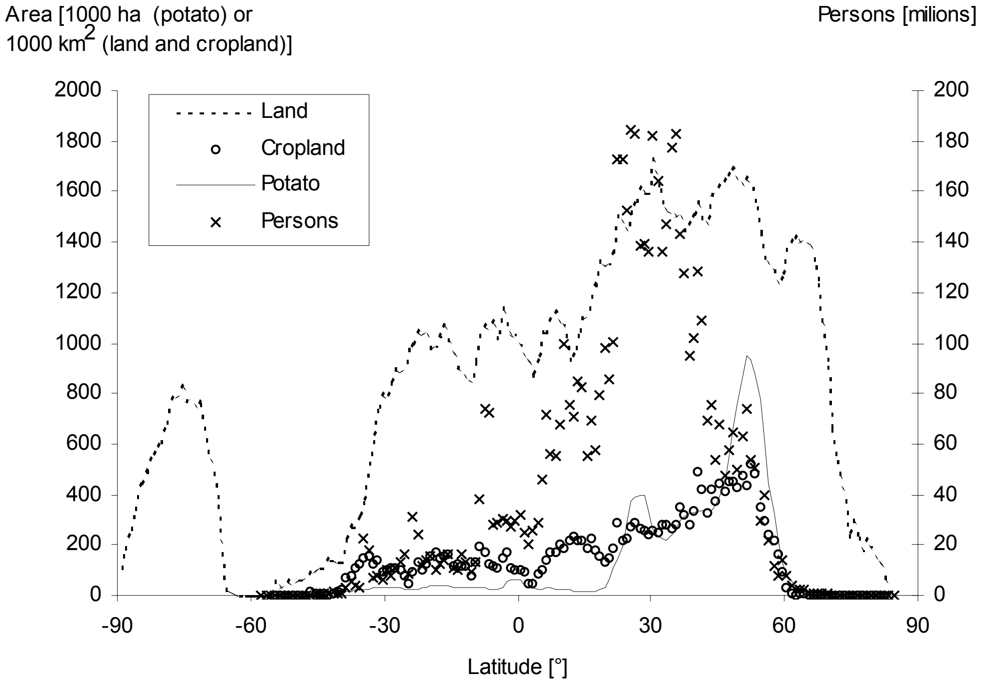


Figure 6.5 Distribution of potato area, land area, cropland, and population by degree latitude. Each dot represents one-degree latitude. Potato area is the five-observations moving average. Latitude in the southern hemisphere is indicated with a minus (-) sign.

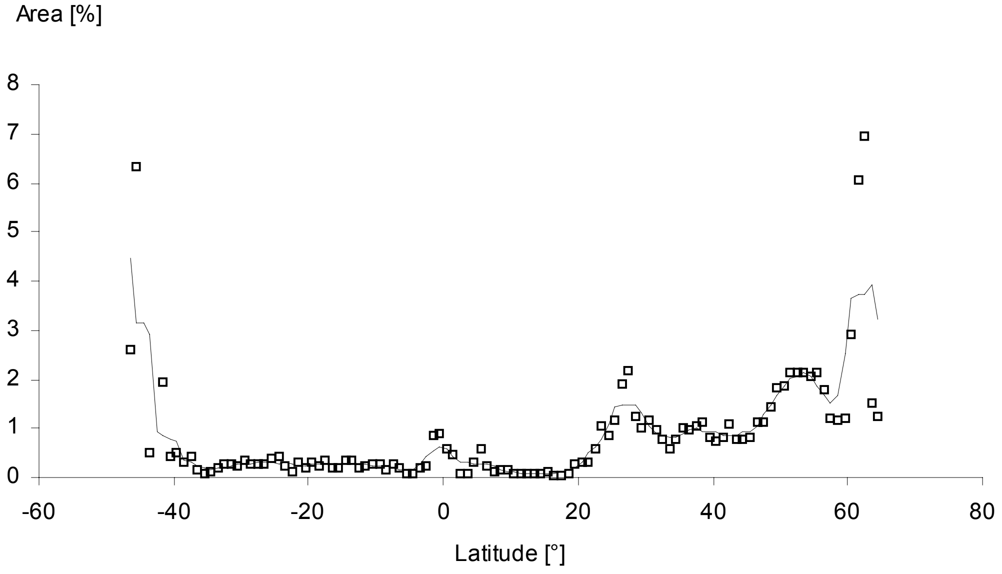


Figure 6.6. Distribution of potato area as percentage of cropland by latitude. Each dot represents one-degree latitude. The line (cropland) is the five-observations moving average. Latitude on the southern hemisphere is indicated with a minus (-) sign.

The relative importance of potato is comparable in the sub-tropical lowlands (between 26–30°N) and at the main latitudes of potato production (50-55°N). However, this obscures the fact that in the Nile valley of Egypt, and in the subtropical lowlands of India and Bangladesh, potato is part of a rotation of two, and sometimes three, crops per year. In these areas, the total area planted with crops in one year is at least double the cropland area; at higher latitudes typically only one crop is grown per year on a given piece of land.

The relative potato area over cropland is highest at the extremes of the area with cropland (south of 40°S and north of 60°N; Figure 6.6). Potato is well adapted to these areas because it can produce in a short growing season, and when the growing season ends prematurely due to frost some of the harvest can be still be recovered. Yet, the estimate of relative potato area at high latitudes should be interpreted with care as there is only little cropland and potato at these latitudes, and the data are, hence, rather uncertain.

Potato area distribution over altitude

Of the global potato area 25% is located above 1000 m. Distribution of potato area over altitude, by degree latitude, is rather different between the northern and southern hemisphere (Fig. 6.7). South of 30°S, almost all area is below 1000 m. Between 22°S and 9°N, in the tropical highland areas of the Andes, central and East Africa, Ethiopia and Indonesia, most potato area is above 1000 m. In this zone most potatoes are grown much higher than 1000 m, especially in the Andes, where most potatoes are grown between 3000 and 4000 m.

Further north, however, the percentage of the potato area above 1000 m decreases sharply, due to the lowland potato area of the Indo-Gangetic plains, where potato is a winter crop. In the southern hemisphere there is not much lowland (sub-) tropical potato (exceptions include production zones in Mozambique, Angola, Brazil, and the coast of Peru).

Between 32°N and 42°N (mainly central China) more then 50% of the potato area is above 1000 m. In this area, potato can be a winter, spring, summer or autumn crop, depending on altitude and cropping system (He, 1997). Further north, in the temperate zone, most of the potato area is below 1000 m.

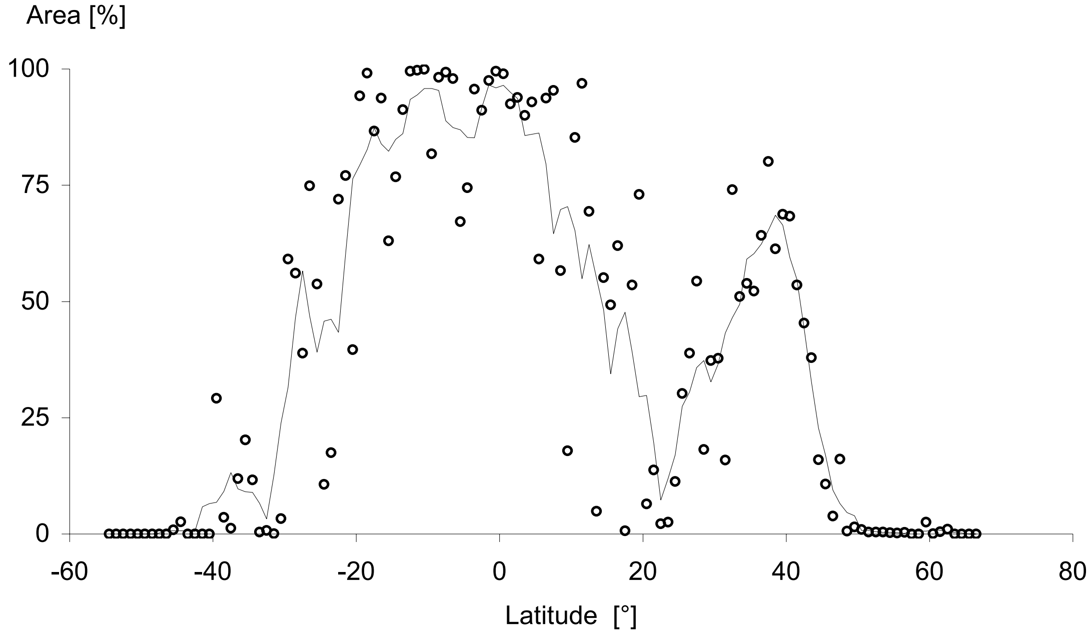


Figure 6.7 Percentage of the potato area that is located above 1000 m, by latitude. Each dot represents one-degree latitude. Latitude on the southern hemisphere is indicated with a minus (-) sign. The line is the five-observations moving average.

6.4 Discussion

Potato is grown over all latitudes where there are crops and (large groups of) people. Because potato does not produce well under conditions of high temperature, in the tropics it is grown in the cool highlands, and in the subtropics in the cool (winter, autumn or spring) season, or at mid-elevations. In the temperate zone it is a lowland crop.

The potato area distribution is highly skewed towards the northern hemisphere, and particularly to the temperate zone in Europe. Although potato area in Europe has decreased strongly over the past 50 years, its share is still high at 51% of the total global area. There has been an increase of potato area in the tropical highlands, and particularly in the subtropical lowlands of India and Bangladesh, and in China. This trend is likely to continue (Walker *et al.*, 1999; Scott *et al.*, 2000).

Shifts in potato production by latitude were assessed assuming that there were no changes in the spatial distribution of potato area within countries. This is a simplification of reality, notably for the USA. In the early 1900s potatoes were produced in almost all counties in the USA (Finch and Baker, 1917). Local producers could compete with those in areas of specialization because of high transportation costs of the bulky and perishable potatoes. This situation of "where there are people are potatoes" has changed radically in the USA and, to a lesser extend, in western Europe (Walker *et al*., 1999).

Whereas there are good maps of potato production in North America and Europe that can be used to assess these within-country changes (Walker *et al*., 1999), little is known about changes in the spatial distribution of potato production in Asian countries like China, and to what extent the current loci of production are determined by bio-physical conditions and transportation costs. With the availability of the current within country potato distribution data, as described in this paper, this type of question can now be addressed.

It should be noted, however, that the data sources used for the global database of potato area described here differed greatly in detail and quality and the overall accuracy of the global potato distribution database is difficult to assess. Whereas in some countries accurate data are disseminated at a low spatial aggregation level, for other countries data are scant and uncertain. The database for China was a special case; it probably has good spatial accuracy, in a relative sense, but the total area is highly uncertain. County-level data for 1987–88 (2420 counties of which 1134 have potato) were used. The total potato-growing area sums to 4.1 million ha. For the same year the FAO gives an estimate of about 2.6 million ha for China. The difference of 1.5 million ha is more than 50% of the FAO estimate for China and almost 10% of the FAO estimate of the global potato area.

Observations during a visit to southwest China in May 2000 indicated that the FAO estimate of potato area is probably too low, as in the case of other crops and total crop area (Crook, 1993). Local agriculture officials customarily talked about the "official" potato area, and the "real" (and higher) potato area. Reasons mentioned were tax evasion and problems with accounting for area in multiple and mixed cropping systems. In southwest China, potato is often grown as an intercrop with maize (He, 1997). In Sichuan province there was no space for the (relatively new) autumn potato crop on the census form. Nevertheless, for reasons of consistency, the FAO estimate for 1997-99, for the total potato area in China was used.

There are prospects of using more remotely sensed data such as the high-resolution (~1km2) global land cover data (USGS EDC 1998). However, it is currently difficult to use these databases to identify the spatial extent of global agriculture (Wood *et al*., 2000), let alone that of specific crops. At present, detailed census type data are indispensable for the development of global crop distribution databases.

For this type of database, documenting the sources used for each country is very important, as this can guide efforts to update and help users assess the data quality. The sources of a previous version of the database discussed here were documented by Huaccho and Hijmans (1999), but better examples of documentation can be found in Carter *et al*. (1992), Wortmann *et al*. (1998) and Huaccho and Hijmans (2000). Having the data in a digital format has the dual advantage of increasing its usefulness for research of global potato production, and making it relatively easy to update when newer or better data become available.

7. Constraint specific agro-ecological zoning

Published as: Hijmans R.J., G.A. Forbes and T.S. Walker, 2000. Estimating the global severity of potato late blight with GIS-linked disease forecast models. Plant Pathology 49: 697-705.

Abstract

Global severity of potato late blight was estimated by linking two disease forecast models, Blitecast and Simcast, to a climate database in a Geographic Information System (GIS). The disease forecast models indirectly estimate late blight severity by determining how many sprays were needed during a growing season, as a function of the weather. Global zones of estimated late blight severity were similar for both forecast models, but Blitecast generally predicted a lower number of sprays needed. With either forecast model, there were strong differences between potato production zones. Zones of high late blight severity include the tropical highlands, Western Europe, the east coast of Canada and the northern USA, southeast Brazil, and central-south China. Major production zones with a low late blight severity include the western plains in India, where irrigated potato is produced in the cool dry season, north-central China, and the northwestern USA. Using a global GIS database of potato production, the average number of sprays was calculated by country. These averages were compared with estimates of current fungicide use. The results using Blitecast and Simcast were correlated but only Blitecast estimates correlated with observed data for developed countries. The estimated number of sprays, whether from Blitecast or Simcast, did not correlate with the observed number of sprays in developing countries and in a number of developing countries the predicted optimal number of sprays was much higher then the actual number observed. In these countries, the increased access to host resistance and fungicides could have a strong economic impact.

7.1 Introduction

Late blight (LB) of potato (*Solanum tuberosum)*, caused by the oomycete pathogen *Phytophthora infestans* (Mont.) de Bary, is considered to be the most important disease of potato worldwide. In most parts of the world, potato production depends on the routine use of fungicides to control LB. Poor farmers in developing countries, however, often cannot control the disease, resulting in heavy losses and even crop abandonment (Haverkort, 1990; Walker and Collion, 1997; Forbes *et al*., 1998).

The International Potato Center (CIP) and research partners provide resistant cultivars and disease management training to farmers in developing countries. The global extent of LB poses a problem for allocating the limited resources available for this work. Ideally, allocations should be made to regions where LB research would contribute most to increased potato production, food security, and poverty alleviation. Unfortunately, the information currently available on LB severity in different parts of the world is incomplete and highly uncertain. Cox and Large (1960) wrote an insightful overview of potato LB around the world; however, their study is somewhat outdated and lacks detail about the situation in developing countries.

A potential solution to this problem is agro-ecological zoning, which has been used to guide resource allocation for agricultural research and development (e.g., Gryseels *et al*., 1992). However, generic agro-ecological zones may not be very useful in dealing with specific production constraints such as potato LB, because factors that affect disease severity may cut across or vary within agro-ecological zones. Zoning that is specific to a particular technology or production constraint can give better estimates of intensity and variation of the factor under study (Corbett, 1998; Wood and Pardey, 1998). This type of zoning has recently been facilitated by strong technological progress in computer hardware and geographic information system (GIS) software, which has made the manipulation of large geo-referenced databases relatively straightforward.

However, the use of these databases for crop– or constraint–specific zoning requires the development or identification of zoning criteria. Criteria that can link climatic data with crop growth and production constraints are encapsulated in simulation and forecast models, permitting the rather straightforward use of these models in GIS. Crop growth simulation models have been linked to GIS to examine potential crop growth (Stol *et al*., 1991; Han *et al*., 1995; Van Keulen and Stol, 1995), effects of abiotic constraints such as drought (Van Keulen and Stol, 1995), and effects of biotic constraints (Seem, 1995). Effects of global climate change on plant disease severity were evaluated for rice leaf blast caused by *Pyricularia oryzae* using disease simulation models linked to GIS (Luo *et al*., 1998).

In this paper we describe the use of two LB forecast models linked to GIS to develop a LB specific agro-ecological zonation, based on estimates of the number of fungicide sprays needed for effective disease management. The number of sprays needed for disease management was estimated, aggregated by country and compared with estimates of current fungicide use.

7.2 Materials and methods

LB forecast models

The LB forecast models used in this study were Blitecast (Krause *et al*., 1975) and Simcast (Fry *et al*., 1983; Grünwald *et al*., 2000; see also Hansen *et al*., 1995). Both forecast models have a daily time-step and use rainfall, temperature, and leaf-wetness duration (estimated to occur when relative humidity equals or exceeds 90%) data to predict the need for protectant fungicide sprays. Blitecast predicts only for susceptible cultivars while Simcast predicts for susceptible and moderately resistant cultivars, but only predictions for susceptible cultivars were used in this study.

Simcast recommends spraying after reaching either a threshold of accumulated blight units (BUs) or fungicide units (FUs). BU accumulation is based on the daily average temperature and hours of leaf wetness. FUs accumulate with time and rainfall, and represent fungicide washing or chemical degradation of the active ingredient. BU accumulation and thresholds and FU thresholds are dependent on cultivar resistance, but FU accumulation is not. We assumed initial occurrence of the pathogen to be at emergence, and the first spray needed when the accumulated BUs reached its threshold for the first time. Blitecast estimates both the timing of the initial fungicide spray, as well as the need for subsequent sprays. Blitecast either indicates not to spray or to use a five-day or a seven-day spray schedule (Krause *et al*., 1975). The need to spray depends on a combination of accumulated severity values, calculated from temperature and hours of leaf wetness, and “rain-favorable days” during the previous seven days. For both Simcast and Blitecast, there is a minimum interval of 5 days between sprays.

Climate data

Climate data were taken from a database with global monthly climate data (New *et al.*, 1999). The original database consists of interpolated climate data on a grid with a cell size of 30-minutes in both longitude and latitude. This database was disaggregated, using linear interpolation, to a 10-minute grid to allow for smoother changes in (tropical) highland areas. Minimum and average temperature, precipitation, number of days with precipitation, and vapor pressure were used. Minimum and maximum relative humidity (RH) were calculated from minimum and maximum vapor pressure using the formulas also used for this transformation by New *et al.* (1999). Daily temperature data were derived by linear interpolation between the monthly averages. Daily leaf wetness was estimated from the minimum and maximum RH (see below). Daily precipitation was estimated with a stochastic rainfall generator (Supit *et al.*, 1994).

Determining growing seasons

In temperate regions, there may be one clearly distinguishable season in which potatoes are grown. However, in subtropical and tropical areas, potatoes may be grown at different times of the year, or even year-round (Haverkort, 1990; Forbes *et al.*, 1998). There is no comprehensive database on planting dates and growing seasons of potato worldwide. Therefore, we determined the most likely growing season(s) for each 10-minute grid cell following rules adapted from previous studies (Stol *et al.*, 1991; Van Keulen and Stol, 1995). In each grid cell there were between 0 and 12 growing seasons a year, with emergence on the 15th of each month. Potato could be grown in any month when the minimum temperature was above 5°C and average temperature below 20°C. At temperatures below 5°C, frost risk is high; at average temperatures above 20°C, tuberization is reduced (Stol *et al.*, 1991). A growing season could start in any month that permitted an accumulation of at least 1,250 degree days (°Cd) in consecutive months without reaching either of the temperature limits. Degree days were calculated as the sum of daily average temperatures minus a base temperature of 2°C. At some sites, potato production was not possible according to these rules, although a database of global potato production (Huaccho and Hijmans, 1999) indicated the presence of potato. For these areas, which were at the cold or hot extremes of the potato production zones, the temperature limits and minimum duration of the growing season rules were relaxed. The growing season ended when 1,750 °Cd had been accumulated, or at 140 days after emergence, or when either of the temperature limits was reached. These rules reflect the fact that growing seasons are generally longer in cooler climates due to slower crop development. Precipitation was not taken into account to determine length of growing seasons, because potato is an irrigated crop in many areas. Irrigation was assumed to be available where needed.

The LB forecast models were run for each grid cell and for each growing season, from emergence until 1 week before the end of the growing season. The number of sprays predicted to protect a potato crop against LB was recorded. Because of the stochastic character of the leaf wetness and rainfall generators, each run was repeated 30 times, and the number of sprays/run was averaged. For grid cells with more than one growing season, the model was run 30 times for each season and the average of all runs was used.

Effect of using monthly weather data.

The forecast models use daily leaf wetness duration (i.e., the number of hours that RH equals or exceeds 90%) as an input variable. Available global geo-referenced weather databases only have monthly data. Therefore, we assessed the effect of using leaf wetness estimated from monthly RH averages. To do this, we used a database of hourly weather data from 262 sites in the US, taken over a 6-year period (NOAA and US-EPA, 1997). First, monthly averages of these data were calculated over the six-year period. Distributions of daily leaf wetness duration were then generated from these monthly averages, using a relationship between monthly average minimum and maximum RH and leaf wetness duration, and assuming a normal distribution around the mean leaf wetness for a given day. We then calculated the optimum number of sprays needed to control LB (using Simcast) using both the six years of observed hourly RH data and 30 years of the generated daily leaf wetness duration data. Daily temperature and rainfall from the original data set were used in both cases. The results were compared using regression.

Of the three weather variables needed for the forecast models, the effect of deriving hourly data from monthly data was assessed only for RH for the following reasons. First, temperature and rainfall data are used as daily averages in the forecast models. Therefore, the degree of interpolation (from monthly to daily) is less than required for RH (monthly to hourly). Second, much work has been done to develop stochastic models for generating daily rainfall values from monthly values (e.g., Geng *et al*., 1986). Finally, BU accumulation is much less sensitive to small changes in temperature than to changes in leaf wetness (Fry *et al*., 1983).

Aggregation of model results and comparison with farmer practices.

Averages of predicted optimum number of sprays were calculated for each country. First, the average number of sprays/ha was calculated for the potato production zones in each country, using a geo-referenced database of potato production that delineates the potato production zones in each country and indicates the area planted with potato in each zone (Huaccho and Hijmans, 1999). Then, for each country, the average number of sprays/ha, weighted by the potato area in each production zone, was calculated. We compared the optimum number of sprays predicted by the model with the actual number of sprays used by farmers in large areas (country or region within a country). This was done separately for developed countries (parts of the USA and the EU) and for developing countries.

7.3 Results

Effect of using monthly weather data

For the 262 US stations used in this analysis, the prediction of LB severity using RH data generated from monthly averages was similar to the prediction when using observed hourly RH data (Fig. 7.1). There did not appear to be any systematic effects related to LB severity as data fit a linear regression model fairly well ($R^2 = 0.89$).

Global zonation based on potential LB severity.

Both Blitecast and Simcast, linked to the GIS based interpolated climate data, identified striking differences in LB severity among different locations, expressed as the number of sprays needed for disease control (Plate 4). The optimum number of sprays was especially high in the tropical highlands of Latin America, Africa, and Asia; in western Europe, the east coast of Canada, and the northern USA; southeast Brazil, central-south China, and in many coastal areas. Major potato producing areas with low LB severity included north-central China; the western plains in India, where irrigated potato is produced in the cool dry season; and the north-western USA.

Comparison of forecast models

Simcast and Blitecast were compared for the number of sprays predicted, aggregated at the country level. The predictions of the two forecast models were highly correlated (Fig. 7.2), but on the average Simcast predicted about twice as many sprays as Simcast (Fig. 7.2, Tables 7.1 and 7.2).

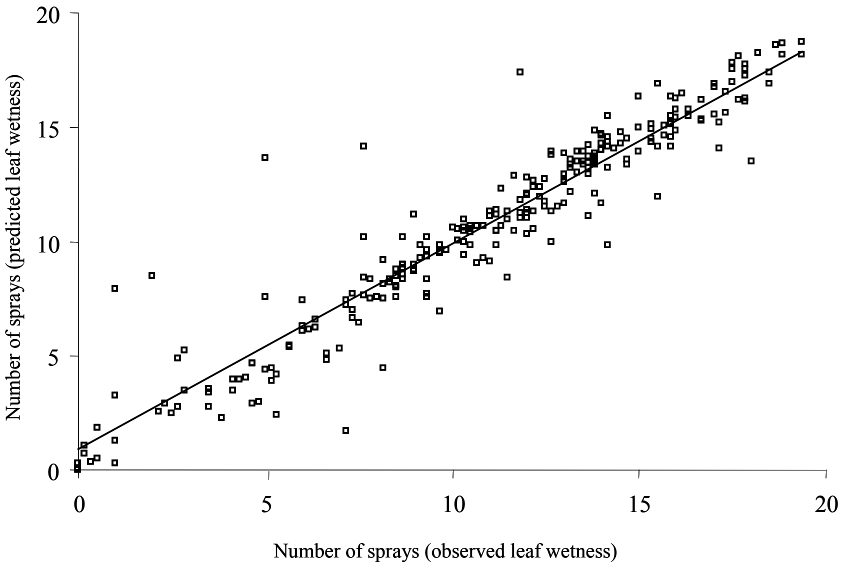


Figure 7.1 Estimated average number of sprays needed to control potato late blight for 262 USA weather stations, using six years of observed data versus 30 runs with generated leaf-wetness duration (i.e., hours when RH≥90%) data. Daily data was generated from mean monthly data derived from the six-years hourly data. Regression line: $y = 0.89 + 0.9x$; $R^2 = 0.89$.

Comparison of predicted and observed values

The average observed number of sprays in selected developed countries was more closely predicted by Blitecast than by Simcast (Table 7.1). The observed number of sprays was poorly correlated with the number predicted by Simcast. The correlation between observed and predicted by Blitecast, however, was significant at the 10% level ($p = 0.059$, Table 7.1). The average observed number of sprays for developing countries was also predicted better by Blitecast than by Simcast (Table 7.2). In this analysis, however, neither prediction was significantly correlated with observed values.

Plotting the observed number of sprays versus those predicted by Blitecast for developing countries permitted evaluation of particular countries of interest (Fig. 7.3). Countries above the $y = x$ axis are those in which fungicides may be under-utilized. Many of these are African countries where under-utilization of fungicides has long been suspected. Countries below the line included Ecuador, Costa Rica and Indonesia, an indication of potentially excessive fungicide use in these countries.

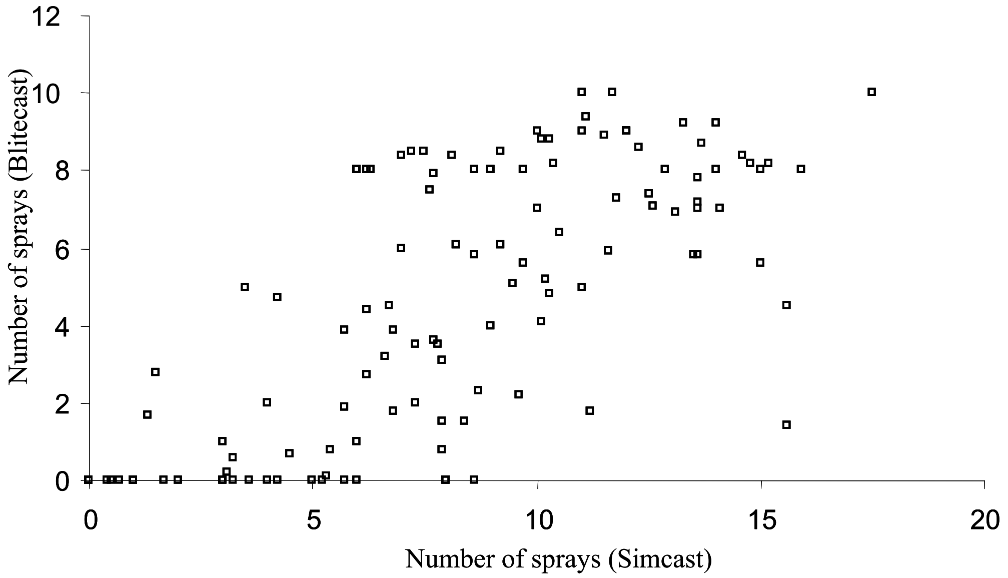


Figure 7.2 Average number of protectant fungicide sprays for potato producing countries needed to control potato late blight as predicted by Blitecast and Simcast, two late blight forecasting models, within a geographic information system. Each dot represents one country. Correlation coefficient = 0.75.

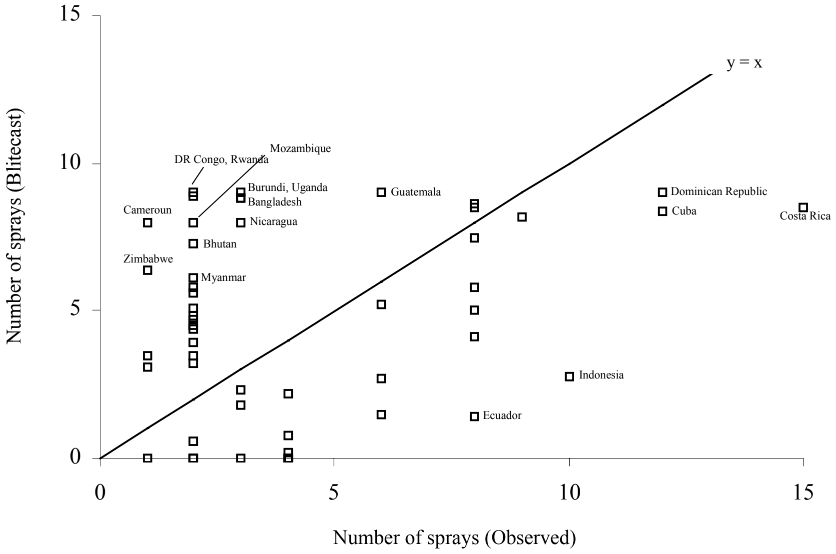


Figure 7.3 Average number predicted and observed protectant fungicide sprays needed to control potato late blight in selected developing countries. Predictions based on Blitecast, a late blight forecasting model, run within a geographic information system. Observed numbers of sprays are primarily from surveys. Each dot represents one country.

Table 7.1 Observed number of fungicide sprays per crop used by farmers to control late blight in several develop countries, and predicted optimum number of sprays for a susceptible cultivar according to Simcast and Blitecast.

Country	Location	Observed No. sprays[a]	Predicted No. sprays Simcast	Predicted No. sprays Blitecast
USA	Maine	13	13.5	8.4
	Red River Valley	8	11	9
	Washington	5	0.6	0.2
Austria		4.5	11.6	5.9
Belgium		10	14.0	9.2
Switzerland		6.5	13.6	7.8
Germany		6	13.5	5.8
Denmark		6	15.2	8.2
France		14	12.9	8.0
Italy	North	5.5	6.8	1.8
Norway		2.9	8.4	1.5
Netherlands		7	14.6	8.4
Sweden	South	10	15.4	8.0
UK	Northern Ireland	4	18	10
	England and Wales	7	15.9	8.3
	Scotland	5	15.6	5.8
Average		7.15	12.54	6.64
Correlation[b]			0.22	0.48
p-value			> 0.1	0.059

[a] Data for the USA are for 1996 and were calculated from data provided by the U.S. Department of Agriculture's National Agricultural Statistics Service. Data for Europe were taken from Schepers *et al.* (1997).

[b] Correlation coefficient and P-values between observed and predicted a number of sprays for Simcast and Blitecast.

Table 7.2 Observed number of fungicide sprays per crop used by farmers to control late blight, as estimated by Walker and Collion (1997) and predicted optimum number of sprays for a susceptible cultivar, according to Simcast and Blitecast. Aggregate values for potato producing developing countries with more then 30,000 ha of potato.

Country	Observed	Optimal No. sprays		Potato area [a]
	No. sprays	Simcast	Blitecast	(1000 ha)
China	2	6.6	3.2	3,409
India	2	7.3	3.5	1,380
Peru	6	7.9	1.5	242
Turkey	4	3.1	0.2	209
Brazil	8	12.3	8.6	181
Colombia	9	14.8	8.2	169
Iran	4	0.4	0.0	161
Bolivia	2	3.2	0.6	135
Bangladesh	3	10.3	8.8	134
Egypt	3	8.7	2.3	115
Nepal	2	9.7	5.6	108
Argentina	8	10.1	4.1	99
Pakistan	2	5.7	0.0	90
Kenya	6	10.2	5.2	86
Algeria	4	3.6	0.0	78
Ecuador	8	15.6	1.4	66
Morocco	4	9.6	2.2	63
Mexico	6	6.2	2.7	63
Chile	3	11.2	1.8	61
South Africa	1	7.9	3.1	61
Indonesia	10	1.5	2.8	56
Uganda	3	10.1	8.8	55
Malawi	2	6.8	3.9	52
Ethiopia	2	6.7	4.5	45
Tanzania	2	8.6	5.8	36
Viet Nam	4	5.4	0.8	33
Average	4.23	7.83	3.45	
Correlation[b]		0.27	0.15	
P-value		> 0.1	> 0.1	

[a] Average annual potato area for 1996-1998, as estimated by the Food and Agriculture Organization of the United Nations (http://apps.fao.org).

[b] Correlation coefficient and P-value between observed and predicted number of sprays for Simcast and Blitecast.

7.4 Discussion

In this paper we demonstrate the successful integration of two LB forecasting models within a GIS. This approach allowed us to estimate global LB severity, based on the number of fungicide sprays that would be needed to control the disease. We think that the GIS-linked forecast models provides a useful approximation of relative LB severity for two reasons. First, both Blitecast and Simcast have been found to give reasonably good LB forecasts (e.g., Grünwald *et al.*, 2000). The second reason is related to the nature of the forecast models. Both are relatively simple models based on the logical premise that LB develops when there are extended periods of leaf wetness within certain temperature ranges. Neither model attempts to simulate complex components of the disease process. Rather, they simply identify, and to some extent quantify, weather periods conducive to disease development. Simcast also estimates loss of fungicide based on time and rainfall. It may be because of the low level of complexity of these models that similar results were obtained with monthly and hourly RH data. It is also because of their simplicity that they are appropriate for use in this type of global studies.

The two forecast models gave different results for the absolute number of sprays needed to control LB, but similar results for the relative number of sprays. Information on relative LB severity was sufficient for our zonation purposes. Nevertheless, the difference between the two forecast models illustrates that there is scope for research aimed at developing more globally valid LB forecast models, or for establishing the geographic domain within which particular forecast models are valid. There is a large, and growing, number of LB forecast models available (see, e.g., Schepers *et al.*, 1997; Harrison, 1992), but generally only little is known about their validity outside the locations where they were developed, and information about their validity in developing countries (i.e., tropical and subtropical climates) is rare.

Our late blight severity zoning method is based on a number of assumptions. We think they are transparent and simple, and helped us to present a close approximation to reality. They are, however, subject to discussion and improvement. For example, our method of estimating planting dates should be more formally tested. Greater knowledge of how farmers adapt to LB pressure would help in the interpretation of results. In the African highlands, for example, farmers tend to plant early or late to avoid the wettest season, which is most conducive to LB, thereby accepting the increased risk of drought stress (Devaux and Haverkort, 1987).

Nonetheless, the results of our zonation seem reasonable. The observed number of sprays surpasses the predicted number of sprays in countries with intensive potato production, such as Costa Rica and Indonesia. In contrast, the observed number of sprays in most African countries was below the numbers predicted (Fig. 7.3). Surveys done by the International Potato Center (Table 7.2, and unpublished data) indicate that many farmers in sub-Saharan Africa frequently do not use fungicides. Potatoes are often grown in low-input systems in these countries and frequent fungicide use may not be economically feasible. As a result of this discussion, it becomes clear that our zonation exercise can improve the interpretation of survey data. Surveys provide information about how farmers manage or do not manage late blight, but this information is not necessarily correlated with actual late blight severity.

Zonation of LB severity can be used in decision making about resource allocation for and prioritisation in LB research. Although host resistance should be the primary component of all management strategies, the rate of adoption of resistant varieties is generally low. For this reason, and given the gap between the optimal and observed number of sprays, projects that help farmers gain access to and safely use fungicides may have rapid and significant economic impact in sub-Saharan Africa. This approach should not replace but complement the long-term goal of introducing resistant cultivars.

In Costa Rica, Indonesia, and Ecuador, however, resources might better be directed toward an overall reduction of fungicide use by optimisation methods such as decision support systems and better application technology. Here, as in many developed countries, introduction of host resistance would be beneficial from an environmental and health point of view, but adoption may be less likely, as the disease is apparently well controlled with fungicides, allowing the production of susceptible varieties with high market demand.

8. Ex-ante impact assessment

Published as: Hijmans, R.J., B. Condori, R. Carillo and M.J. Kropff. A quantitative and constraint-specific method to assess the potential impact of new agricultural technology: the case of frost resistant potato for the Altiplano (Peru and Bolivia). Agricultural Systems (in press).

Abstract

A quantitative and constraint-specific approach to assess the potential impact of new agricultural technology is described and applied to frost resistant potato cultivars for the Altiplano (Peru and Bolivia). The approach uses geo-referenced databases and a simulation model. Calculations are made for small grid cells, and no arbitrary delimitation of agro-ecological zones is needed. The LINTUL potato growth simulation model was adapted to incorporate the effect of frost damage on yield. High-resolution grids of monthly climate data were created for a number of variables, including absolute minimum temperature and its standard deviation, and used as input for the simulation model. The model was run for each grid cell, using a standard potato cultivar for which frost resistance parameters were changed in increments of 1 ℃. A geo-referenced database of potato distribution was used to process the output of the simulation model to calculate potato-area weighted results. When frost resistance increases from –1 ℃ (current level) to –2 ℃ or –3 ℃, average potato yield increases by 26 and 40%, respectively. After that, the effect tapers off and a further increase in resistance leads to only a small increase in simulated potato yield.

8.1 Introduction

Declining support for agricultural research has sharpened the need to economize and heightened the need to systematically evaluate the impacts of past research and the prospects for current and proposed work (Wood and Pardey, 1997). Prioritization of agricultural research should be based on the likelihood of success of developing new technology, and of its adoption by farmers, and its consequential economic and environmental impact. Adoption and impact of new agricultural technology is site-specific; it depends to a great extent on environmental factors (particularly on land characteristics such as climate) but also on social and economic circumstances. This paper describes a new approach to account for the effect of environmental factors on the potential impact of new agricultural technology, using a case-study on frost resistant potato cultivars for the Altiplano (Peru and Bolivia).

Agro-ecological zonation

Agro-ecological zones (AEZs) stratify an area into environmentally homogeneous domains usually derived from climate and soil data (e.g., FAO, 1978-81). AEZs can be useful for selecting test sites, interpreting experimental data, targeting technology, setting research priorities, and ex-ante impact assessment (Wood and Pardey, 1998). For example, Gryseels *et al.* (1992) used AEZs to set research priorities for the Consultative Group on International Agricultural Research (CGIAR). AEZs can be classified as generic or specific, and as quantitative or qualitative. Qualitative AEZs can be ordinal or not. The International Potato Center (1991) uses a generic, qualitative, non-ordinal zonation for global potato production, with six different classes: semiarid tropics; humid tropics; subtropical lowlands; arid and Mediterranean; highlands; and temperate. Generic zonation was also used in a FAO (1978-81) study in which production areas were divided into four ordinal zones designated “very suitable” to “not suitable” for rainfed potato. Stol *et al.* (1991) and Van Keulen and Stol (1995) describe a generic and quantitative zonation for global potato production, using a simulation model and a climate and soil database within a Geographic Information System (GIS) to calculate potential and water-limited yield.

The generic character of the AEZs described above makes them useful as a general reference that can be understood intuitively. However, they are difficult to use when addressing specific research questions. Facilitated by the progress in information technology and the development of geo-referenced databases, approaches to agro-ecological zonation are emerging that are more flexible and are specific to the constraint under study (Wood and Pardey, 1998; Corbett, 1998). For example, Hijmans (1999) produced a specific zonation for frost-risk in potato on the Altiplano (Peru and Bolivia) and concluded that if frost resistance in potato increases by 1°C, the percentage of the current potato area with a frost damage event less than once every 10 years nearly doubles, increasing from 18 to 32%. This study does not, however, quantify the potential effect on potato production of the introduction of a frost resistant potato cultivar.

To estimate the potential impact of a new agricultural technology Wood and Pardey (1998) advocate the use of specific agro-ecological zonation for the problem at hand and then estimate changes in yield or production costs for each zone. But estimating, for example, yield loss due to frost damage is difficult because it depends on the probability, the severity and the timing during the growing season of a frost event. A frost event of a specific magnitude in the middle of the growing season will have a sharply different effect on yield than one that occurs at the end of the growing season as most potatoes will have been formed by then. Eliciting estimates on frost-induced changes in yield from experts or farmers is one alternative (Valdivia *et al.*, 1997), but may result in highly subjective data. Producing a good estimate for a large heterogeneous area can prove to be very difficult.

This study demonstrates how crop growth simulation models provide an alternative to eliciting techniques or classification criteria in determining yield change from frost damage. Crop growth simulation models are a mathematical description of a crop's response to the environment. They encapsulate our knowledge of eco-physiological processes, and they can be used to process environmental data to produce easy-to-interpret output such as yield. By comparing the output of different model runs, e.g., representing current and new technology, the effect of technology adoption can be estimated. Instead of estimating differences for predefined agro-ecological zones, the model is run for small grid cells and the results can be aggregated by administrative unit or by production zone. A general framework for quantitative and constraint-specific agro-ecological zonation is presented and then applied to assess the potential impact of potato cultivars with increased frost resistance.

Potato production on the Altiplano

The Altiplano is a high plateau in the Andes of Peru and Bolivia. In this study we focus on the part of the Altiplano called the TDPS system, named after the catchment of Lake Titicaca, the Desaguadero river, Lake Poopó, and the Salt Lake of Coipasa (OEA, 1996) (Fig. 8.1). That excludes the southernmost part of the Altiplano, which is rather arid, sparsely populated, and less important for agriculture. Seventy-five percent of the TDPS system, hereinafter called the Altiplano, is between 3,600 and 4,300 m above sea level; the other 25% is higher. The area comprises 149,000 km^2. Lake Titicaca, covering 8,400 km^2, is a conspicuous part of the topography that greatly influences local precipitation and temperature.

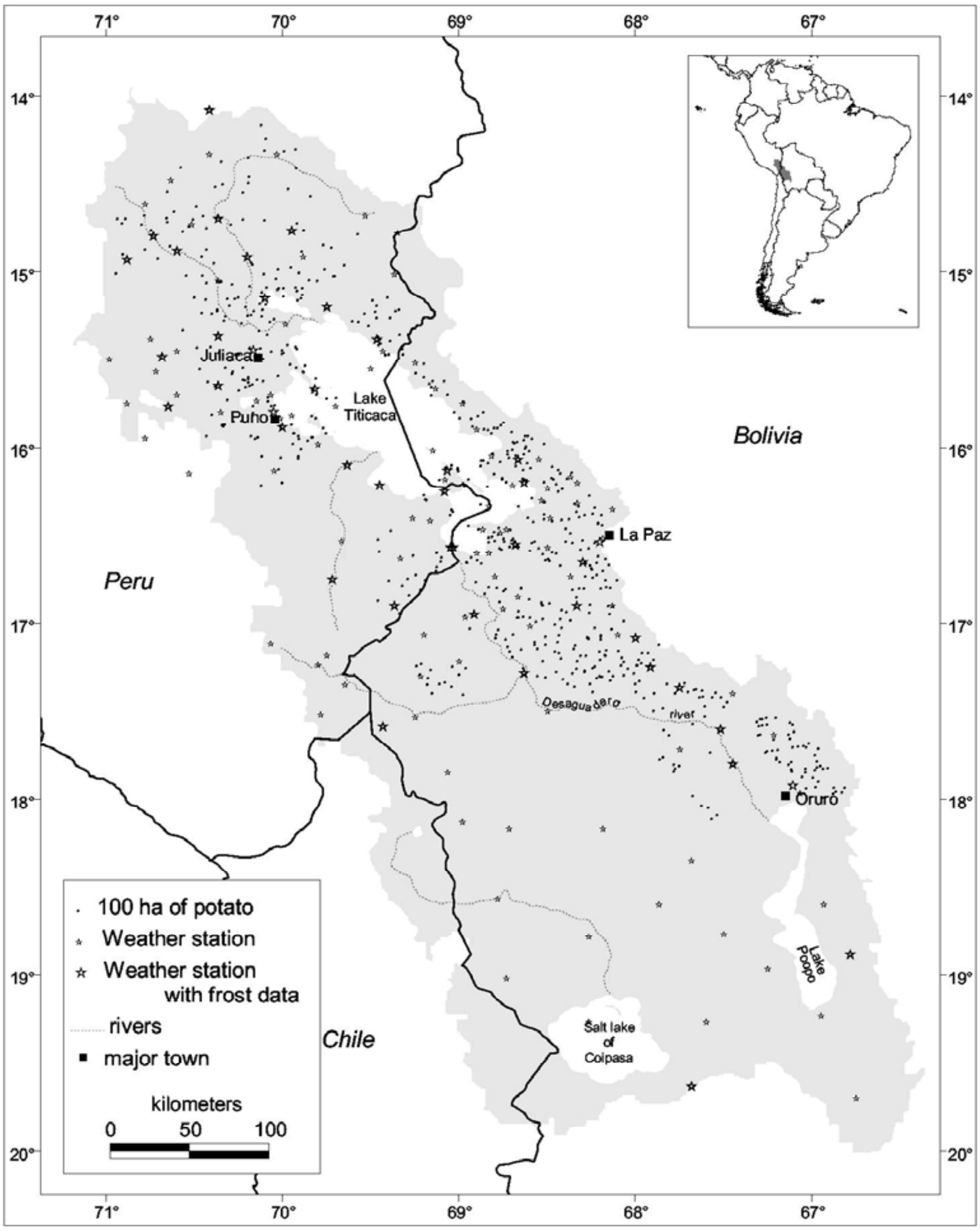


Figure 8.1 The Altiplano (TDPS system) in the Andes of Peru and Bolivia. Topography and distribution of potato production and locations of weather stations that provided data used for the construction of the climate surfaces.

The Altiplano has about 2.2 million inhabitants (OEA, 1996) and is one of the poorest areas of the Americas. About 65% of the economically active population is engaged in agriculture (OEA, 1996). Most of the cropland is located below 4000 m; above that elevation land is mainly used for grazing. Potato is by far the most important crop, accounting for 63% of the gross value of crop production (OEA, 1996). The area planted to potato is about 63,000 ha (G-DRU, 1996; INEI, 1993). Reported potato yields are low, at 5.2 t/ha in the Peruvian and northern Bolivian sections of the Altiplano, and 3.6 t/ha in the southern part of the Bolivian section (OEA, 1996; G-DRU, 1996). However, an extensive survey in four departments of Bolivia (outside the Altiplano) indicates that government statistics have underreported yields by as much as 50% (Terrazas *et al.*, 1998).

The growing season in the Altiplano extends between October and March, when maximum annual temperature coincides with the rainy season. In the agricultural zones, maximum temperature is around 18°C and minimum temperature around 4°C during the growing season (Frère *et al.*, 1975; INTECSA, 1993). Precipitation is highest, around 800 mm/year, in the northeast and in peripheral areas of Lake Titicaca, and lowest, about 200 mm/year, in the southwest. Production risk for potato is high due to a variety of factors, particularly drought, hail, and frost. Frost-free periods average about 140 days for the northern Altiplano and 110 days for the southern Altiplano (Le Tacon, 1989), frosts caused by radiative cooling on clear nights may occur at any time during the growing season. The mid-season frost problem is common throughout the Andean highlands: see, e.g., Knapp (1988), for a description of frost incidence and potato production in Ecuador. This is unlike potato production conditions in temperate regions, where frosts occur only at the beginning and end of the growing season (e.g., Booy, 1961).

Frost damage on potato

Frost occurs when air temperature near the Earth's surface drops below 0°C (Kalma *et al.*, 1992). When frost damage is described in the literature, it is not always clear whether the reported air temperatures refer to conditions at screen height (1.5 to 2 m) or at crop canopy height, which may be as much as 1°C lower (De Bouet du Portal, 1993). In this study, temperature refers to screen height conditions.

The temperature at which frost damage occurs depends on the crop species and the cultivar. For *Solanum tuberosum* subsp. *andigena*, the most commonly grown potato taxon in the Andes, frost damage is likely to occur when the temperature drops to –2°C or lower (Carrasco *et al.*, 1997). Higher frost resistance exists in other cultivated and in wild potato species. For example, cultivated potato species such as *S. ajanhuiri* and

S. curtilobum incur damage at –3 to –5°C (Huanco, 1992; Tapia and Saravia, 1997), whereas *S. juzepzuckii* generally resists temperatures down to –5°C and perhaps even lower (Huanco, 1992; Canahua and Aguilar, 1992; Tapia and Saravia, 1997). With the exception of *S. ajanhuiri*, the tubers of the species with higher frost resistance tend to be bitter due to high levels of glycoalkaloids and therefore require processing before consumption. Henceforward, non-*tuberosum* cultivated potato species will be referred to as "bitter potatoes". Hijmans (1999) estimated that at least 25% of total area planted to potato in the Altiplano is planted with bitter potatoes. This was supported by Canahua and Aguilar (1992) and Huanco (1992), who estimate that about one-third of total potato area of the Peruvian Altiplano is planted with bitter potatoes, of which 60% are *S. juzepzuckii*, and 33% are *S. curtilobum*. It was also supported by Rea's (1992) estimate that bitter potatoes make up 15% of total potato area in Bolivia, where more bitter potatoes are found in the Altiplano than in most other zones.

Frost can cause partial or complete loss of leaf area of a potato crop, leading to a reduction in photosynthesis and hence yield. In turn, crop failure caused by frost damage may lead to a decrease in the total area planted to potato in the subsequent season due to seed shortage (Morlon, 1989). The high production risks presented by frost and other factors may also lead to less investment in agriculture, resulting in decreased production, in spite of the weather conditions in a given year.

Farmers can prevent or reduce frost damage by planting potatoes on "warm" soil (with a high thermal conductivity (cf. Booy, 1961)) and on slopes, where frost incidence is lower than on the valley floor (De Bouet du Portal, 1993); by applying frost-related management practices such as the use of smoke, rustic greenhouses (Aguirre *et al*., 1999), and raised beds (Sánchez de Lozada *et al*., 1998); and by planting frost resistant potato cultivars. Because the latter method is the most practical, a breeding program has been established that aims to produce frost resistant potato cultivars that are otherwise similar to *S. tuberosum* (Carrasco *et al*., 1997). Successful breeding for frost resistance has been reported in the USA (Dearborn, 1969).

8.2 Material and methods

General framework

The general framework for constraint-specific and quantitative agro-ecological zonation is illustrated in Fig. 8.2. As opposed to generic zonation approaches, the study is driven by a specific question related to the production constraint under study. The framework is flexible in the sense that the different types of data used, and how they are combined, depends on the problem at hand as well as on data availability. Model choice depends on what is available and on the type and scale of corresponding input data that are needed. In most cases, weather and soil data would be needed. Once models are selected, different model runs can be compared using ancillary data to interpret and aggregate output. Ancillary data will typically include crop distribution and administrative boundaries. Question and Answer are grouped because we assume that in many cases the answer will serve as a feedback loop that leads to adjustments in the model, data and methods used.

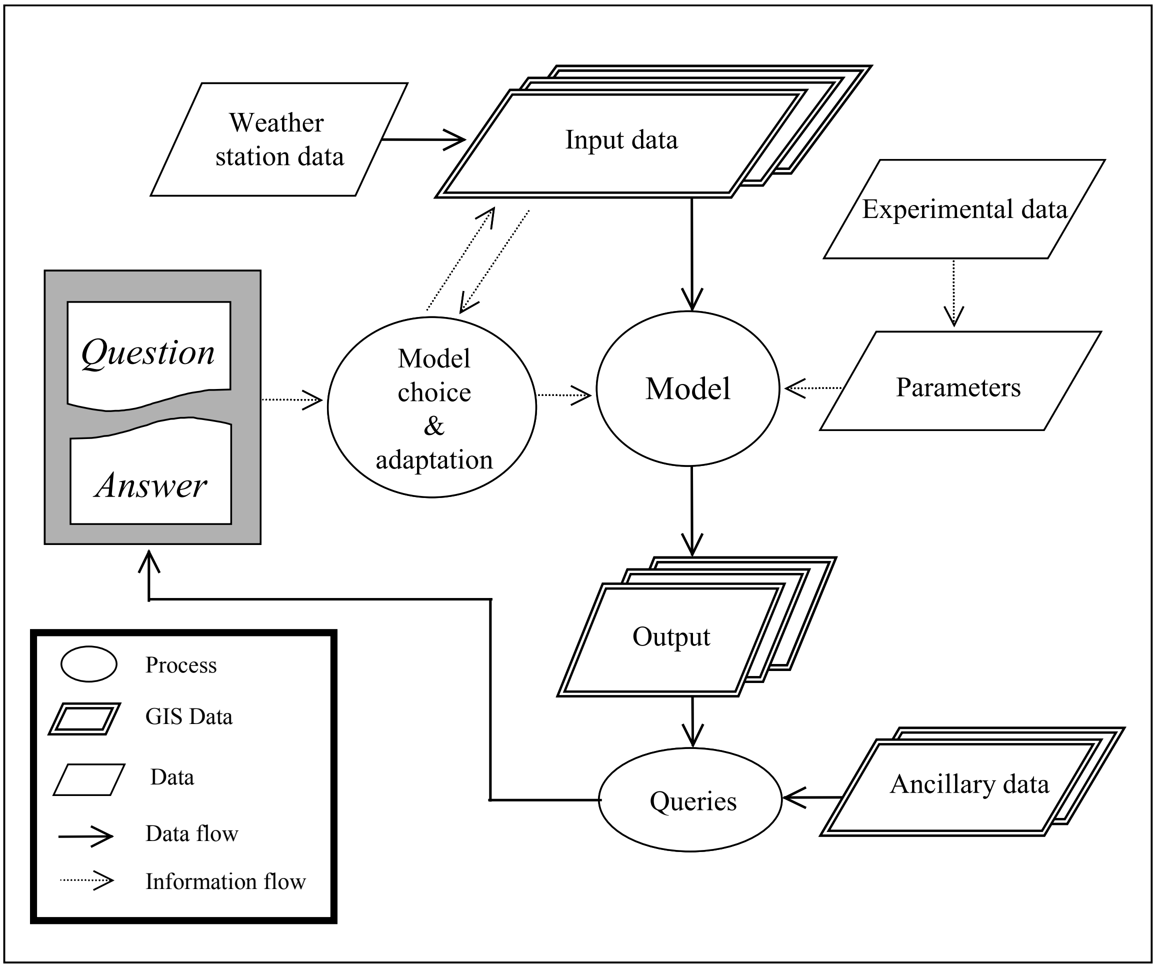


Figure 8.2 Flow diagram for quantitative and constraint-specific agro-ecological zonation using a simulation model and geo-referenced databases.

Simulation model

A slightly modified version of the LINTUL potato growth simulation model (Spitters, 1987; Stol *et al.*, 1991) was used. LINTUL uses a daily time-step and thermal-time dependent ground cover. Ground cover is used to calculate intercepted radiation, and a constant radiation-use efficiency (RUE) parameter is used to calculate dry matter production. Allocation of dry matter to the various organs is thermal-time dependent. The effect of frost was modeled using a simple damage function in which loss of ground cover is described as a linear function of minimum temperature between two temperatures: the "critical temperature" (T_{cr}) and the "leaves-dead" temperature T_{ld} (Fig. 8.3). Above T_{cr} there is no frost damage. If temperature drops below T_{ld}, all leaves are lost. This type of relation was described by Sukumaran and Weiser (1972) for excised leaflets of different potato cultivars and species. Subsequent ground cover expansion is also reduced by frost, depending on a linear function between T_{ld} and the regrowth temperature, T_{rg}. Even at T_{ld} when the crop loses all its leaves, the crop may continue to grow. However, when the temperatures drop below T_{rg}, growth ceases. We assume that frost does not have an effect on the radiation-use efficiency of the remaining foliage.

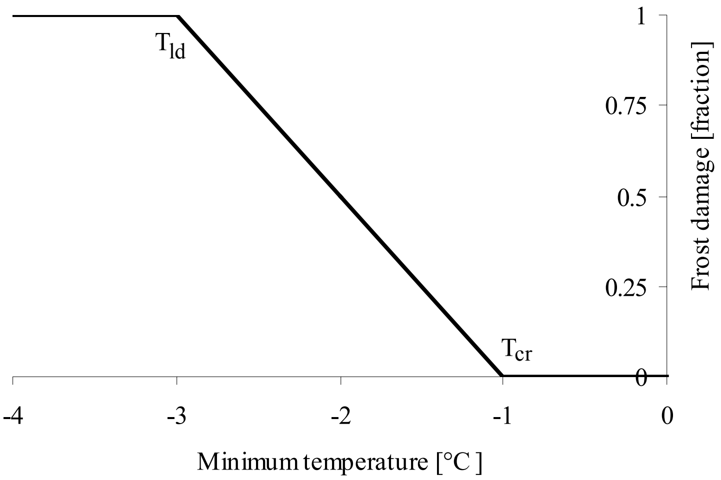


Figure 8.3 Description of frost damage (relative fraction of ground cover loss) as a function of minimum temperature for a standard *S. tuberosum* subsp. *andigena* potato cultivar from the Altiplano (T_{cr} = –1, T_{ld} = –3°C). Estimated from Carrasco *et al.* (1997) and Enrique Carrasco (personal communication).

The simulation model was calibrated for the native potato cultivar 'Gendarme' using data collected during field trials in Patacamaya (La Paz Department, Bolivia) in the 1998/99 season. In this experiment the potato crop did not suffer any major stress, although the growing season ended somewhat prematurely due to a number of frosts. Ground cover expansion parameters were derived by fitting biweekly observations. Radiation-use efficiency was calculated using observed ground cover and biomass from four sequential harvests. The parameters for dry matter distribution between foliage and tubers were also estimated from the sequential harvest data. The model was validated with data from Pomani (Department of La Paz, Bolivia) in 1996/97. In Pomani there was a mid-season, –2°C frost on 1 February 1997 that led to considerable damage of crop foliage. Planting distance (0.7 × 0.3 m) and all other crop management practices, such as fertilization, were similar in both trials.

Weather data

A weather database with data from 139 weather stations was used (INTECSA, 1993). For most weather stations there are at least 30 years of monthly records of total precipitation and average minimum and maximum temperature data. There is also a considerable amount of data on monthly extreme minimum temperature (44 stations) and on other climate variables (Fig. 8.1).

Monthly climate surfaces (grids) were generated for minimum and maximum temperature, absolute minimum temperature and its standard deviation, and solar radiation. The ANUSPLIN program developed by Hutchinson (1995, 1997) was used to interpolate climate data from weather stations and produce high-resolution climate surfaces. ANUSPLIN fits Laplacian smoothing spline functions of two or more independent variables (longitude, latitude, and usually elevation) through the climate observations. This method relies on the strong dependence of climate (especially temperature) on elevation, but allows the size of this dependence to vary over time and space (Corbett, 1998). Elevation data were taken from the US Geological Survey's GTOPO30 database, and used as an independent co-variable. The GTOPO30 and climate surface data have a resolution of 30-arc seconds (approximately 1 km^2). There are 181,757 grid cells in our study zone.

As the simulation model needs daily weather data, these were generated from the monthly climate data. For each grid cell 100 years of synthetic weather were generated to run the simulation model. Daily minimum and maximum temperatures were calculated by linear interpolation of the monthly averages. On one random day of each month, minimum temperature was simulated, using a random value drawn from a normal distribution described by the monthly average extreme minimum temperature and its standard deviation.

Potato distribution

A geo-referenced potato distribution database was created for the Altiplano. District level census data (INEI, 1993) were used for Peru (96 districts on the Altiplano, 62 with potato). District data were adjusted for relatively large districts with major differences in altitude, taking out the areas above 4500 m above sea level. For Bolivia, only departmental level census data (G-DRU, 1996) were available. In our study area there are only 2 Departments, La Paz and Oruro. To obtain a higher spatial resolution, the potato area in Bolivia was disaggregated using a satellite image derived land use map developed by ZONISIG (1998).

Estimating the effect of resistance

The potential yield of 'Gendarme' was compared with the potential yield of mentally constructed genotypes (“ideotypes”) that differed from ‘Gendarme’ in their level of frost resistance only. Frost resistance was described with functions like the one shown in Figure 8.3. The parameters used are presented in Table 8.1.

Table 8.1 Different levels of frost resistance used in the simulations. No damage occurs above critical temperature T_{cr} (°C). At T_{ld}, 100% of the foliage is damaged. When temperatures drop below T_{rg}, the crop stops growing. R–1 is the current level of resistance in *Solanum tuberosum* subsp. *andigena*.

Resistance level	T_{cr}	T_{ld}	T_{rg}
R0	0	–2	–3
R–1	–1	–3	–4
R–2	–2	–4	–5
R–3	–3	–5	–6
R–4	–4	–6	–7
R–5	–5	–7	–8

In the Altiplano, most potatoes are planted in October or November and harvested in April or May. For this study, we fixed emergence arbitrarily at 25 November. Because of the stochasticity of the weather generator, the model was run 100 times per grid cell, once for each generated "year". The model outputs were averaged by grid cell and treatment. Then, the relative yield differences between these averages were calculated by grid cell. These results were weighted by the potato area in each grid cell and tabulated.

8.3 Results

Model calibration

Based on the 1998 experiment, crop growth duration of 'Gendarme' (between emergence and senescence) was estimated at 1250°Cd (base temperature = 0°C), with tuberization starting at 500°Cd after emergence. Radiation-use efficiency was estimated at 2.5 g (dry weight) MJ^{-1} (PAR), a value comparable to those reported in the literature (Stol *et al.*, 1991).

The simulation model somewhat overestimates ground cover and biomass production in the 1996 experiment used for validation (Figs. 8.4 and 8.5). This can partly be ascribed to growth reduction due to excess water in the experimental field, leading to a relatively low ground cover at 50 days after emergence, before the frost event at 58 days after emergence. The difference between observed and simulated tuber yield is smaller then would be expected given the difference in observed and simulated ground cover. More data would be needed to further calibrate the model, but we considered it sufficiently accurate for use in this explorative study.

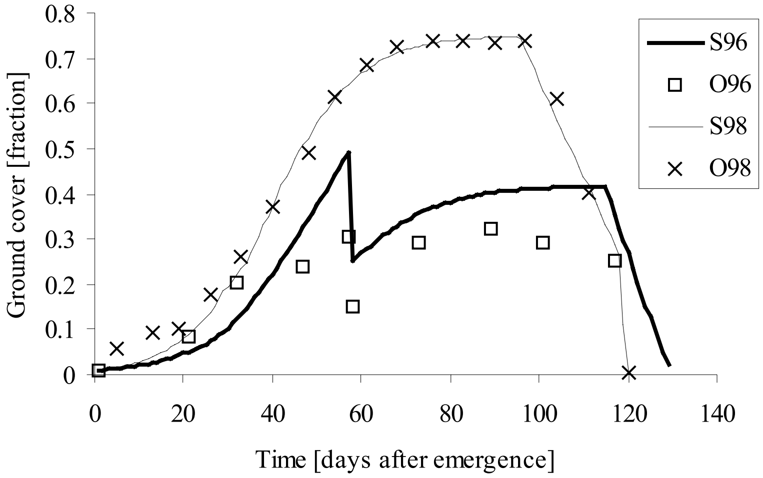


Figure 8.4 Simulated (S) and observed (O) ground cover over time for the 1996 and 1998 experiments. The model was calibrated with the 1998 data. In the 1996 experiment, there was a –2°C frost 58 days after emergence, and in the 1998 experiment, there were frosts of –2 to –4°C 119, 120 and 121 days after emergence.

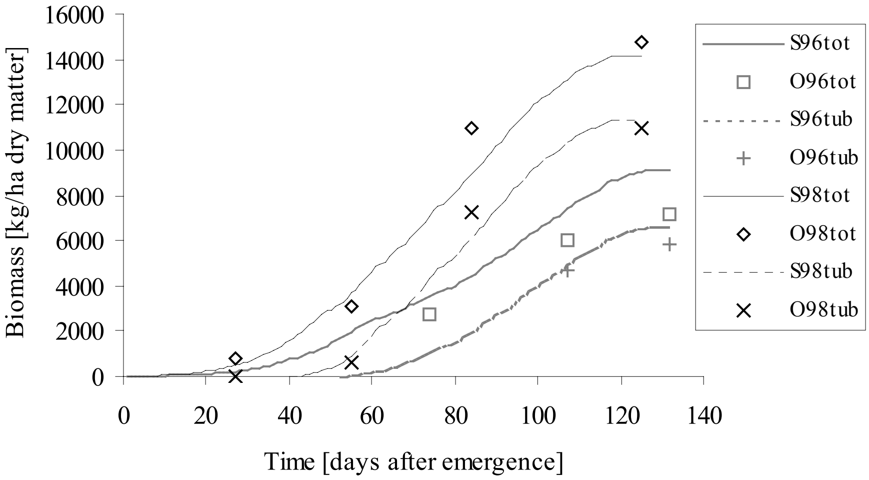


Figure 8.5 Simulated (S) and observed (O) total (tot) and tuber (tub) dry matter biomass production over time for the 1996 and 1998 experiments. The model was calibrated with 1998 data. In the 1996 experiment, there was a –2°C frost 58 days after emergence, and in the 1998 experiment, there were frosts of –2 to –4°C 119, 120 and 121 days after emergence.

Effect of increased resistance

For large stretches of the potato production area on the Altiplano, changing frost resistance in 'Gendarme' leads to large changes in simulated potential yield. The change in simulated potential potato yield when frost resistance (T_{cr}) in 'Gendarme' is increased from –1 to –2°C is shown in Fig. 8.6. Only the area with the highest 75% yield at current resistance levels is taken into account, as the other 25% is considered to be predominantly planted with bitter potatoes. Simulated yield increase is lowest in the areas around Lake Titicaca, that have a relatively mild climate, and in the lower, and hence warmer, parts of the Altiplano, where the Desaguadero river flows into Lake Poopó. The likelihood of adoption of frost resistant cultivars would be highest in areas with a high estimated yield increase. These areas would also be a good location for field trials with new genotypes.

When the data are aggregated over the potato area in the Altiplano, yield response is strong when frost resistance changes from R0 to R–3 (Table 8.2; Figs. 8.7 and 8.8). When the area with bitter potatoes is eliminated, the median simulated potato yield for the Altiplano goes up 37% from 27 to 37 t / ha (fresh weight) (Fig. 7); the average goes up 26% from 31 to 39 t/ha (Table 2) between resistance levels R–1 and R–2. Mean yield goes up an additional 14% between R–2 and R–3 but then the effect levels off. Between R–3 and R–4, the yield gain is only from 43 to 46 t/ha for the median, and from 44 to 47 t/ha for the average yield.

If potato varieties with increased frost resistance (from R–1 to R–2) would be adopted over the whole area with non-bitter potato species we would predict a yield increase of 26%. For the Altiplano, this would amount to a yield increase from 6 to about 7.6 t/ha. The 1.6 tons/ha increase over 47,250 ha (eliminating 25% of the total area, which is planted with bitter potatoes) would lead to an average yearly increase in potato production of 16%. It is unlikely, however, that frost resistant varieties would be adopted over the whole potato area, but further refining this estimate is outside the scope of this paper.

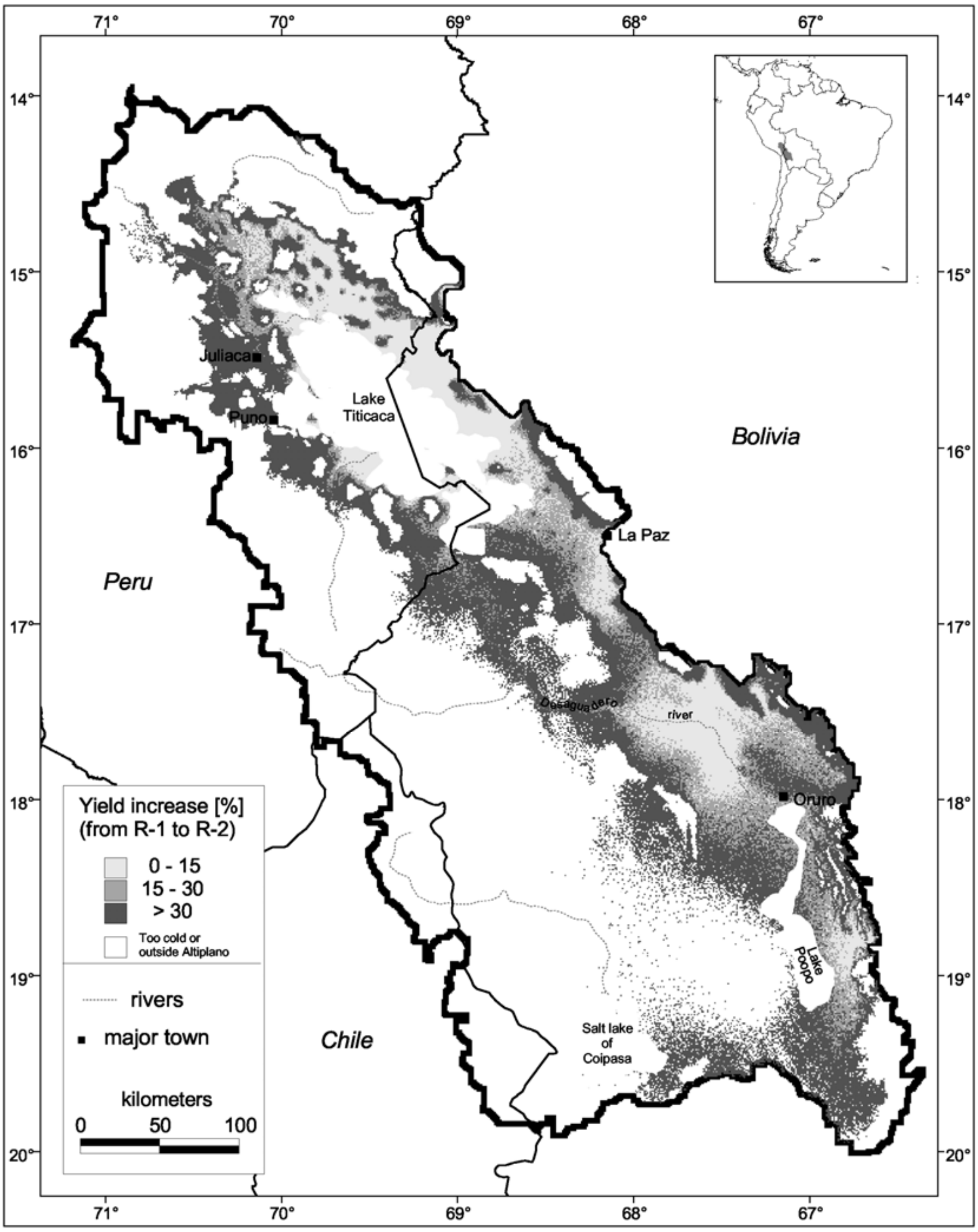


Figure 8.6 The Altiplano (TDPS system) in the Andes of Peru and Bolivia. Simulated yield increased for ‘Gendarme’ when frost resistance increased from –1 to –2°C.

Table 8.2 Simulated potential potato yield (t/ha) for 'Gendarme' with different levels of imposed frost resistance (R0 is resistant (T_{cr}) to 0°C, R–1 to –1°C, etc.). Averages calculated for 75% of the total potato area (excluding areas with a predominance of bitter potatoes).

Resistance level	T_{cr}	Yield
R0	0	20.6
R–1	–1	30.6
R–2	–2	38.6
R–3	–3	43.7
R–4	–4	46.6
R–5	–5	48.0

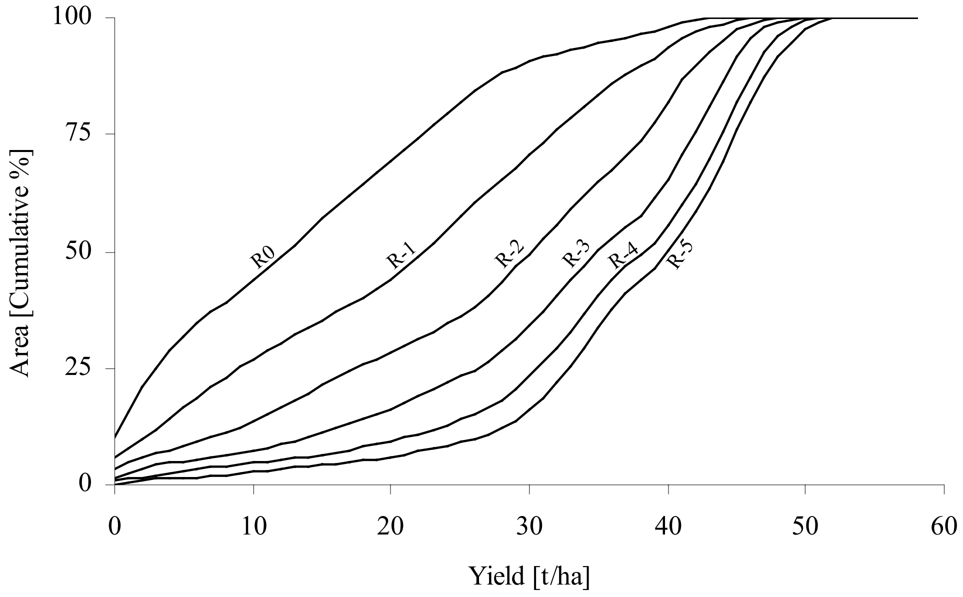


Figure 8.7 Cumulative distribution over total potato area in the Altiplano of simulated potential potato yield (fresh weight) for 'Gendarme' with different imposed levels of frost resistance. R0 is resistant to 0°C, R–1 to –1°C, etc.

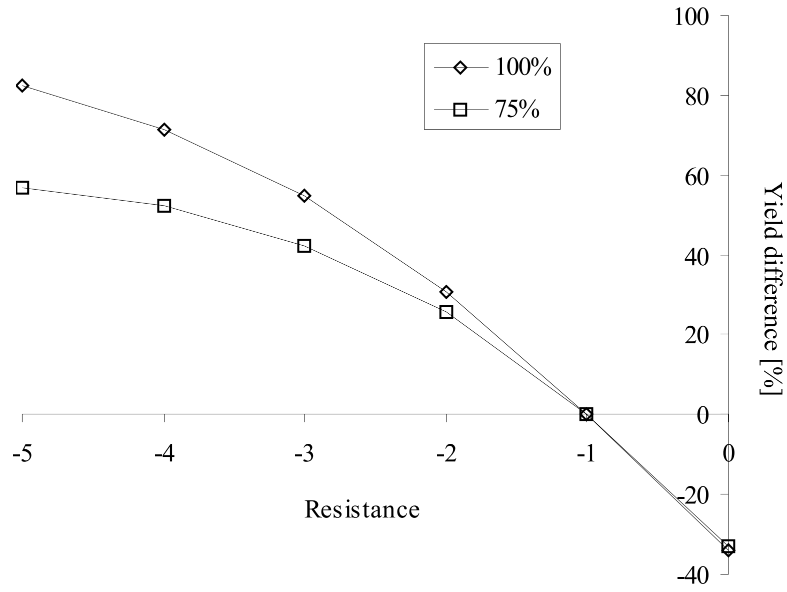


Figure 8.8 Simulated change in potential potato yield (fresh weight) for 'Gendarme' with imposed different levels of frost resistance (T_{cr}= 0°C to –5°C), compared to that at the cultivar's actual level resistance (T_{cr}= –1°C). Averages calculated for total potato area in the Altiplano (100%), and for the area excluding sections with predominance of bitter potatoes (75%).

8.4 Discussion

Frost is an important constraint to potato production in the Altiplano and the adoption of potato cultivars with increased frost resistance would lead to a strong decrease in yield loss. Hence, breeding for increased frost resistance seems to be a viable goal. Based on the data presented here, we suggest that Altiplano breeding programs aim to develop cultivars with –2°C frost resistance (i.e., an increase of resistance by 1°C). This would have a major impact on yield, and the probability of success in developing such cultivars seems high, given the relatively low increase in resistance required and the high levels of resistance available in wild and cultivated potatoes. After increasing current levels of resistance by more then 2°C, the return in investment levels off, while the research cost will probably increase.

This study only considered the current potato area. However, the introduction of new frost resistant potatoes could lead to relative shifts (within the same area) and absolute shifts (to other areas) in the Altiplano potato area. With the exception of the shores of Lake Titicaca, however, the Altiplano is not densely populated and agriculture is not intensive, so land availability does not seem to be a limiting factor in potato production. Therefore, rather than shifting production to new, colder areas, farmers would probably choose to increase production in current production zones. This would result in more efficient and less risky potato production. On the other hand, potential production areas at higher altitudes might be associated with better (less eroded) soil and more precipitation. Hence, some expansion of potato into higher areas could be expected with the introduction of cultivars with increased frost resistance. More in-depth knowledge of farmer strategies in relation to potato production and frost risk on the Altiplano is needed to predict their response to frost resistant varieties.

The simulation-based method used in this study is a more objective way to estimate the potential effect of technological change than classic AEZ approaches in which environmental data are aggregated prior to analysis. Moreover it provides quantitative estimates of changes in production, which are needed in ex-ante impact assessment studies (Alston *et al.*, 1995; Wood and Pardey, 1998). Nevertheless, there are still a number of assumptions and potential sources of systematic error. Systematic error may have occurred because of the way the weather data were used but also because of other reasons. Outcomes are sensitive to changes in model parameters, notably in T_{cr} and T_{ld} and in the parameters describing phenology. We assume that the percent change in simulated yield is a good estimator of the effect on yield in farmers' fields. Yields in farmers' fields tend to be much lower than simulated potential potato yields, particularly in areas like the Altiplano where much production is for subsistence and external inputs such as quality seed or fertilizers are rarely used, and where there are occasional droughts. For some production constraints, changes in potential yield might not be indicative of changes in actual yield. However, the effect of frost is rather simple, it takes away foliage, and not strongly influenced by the absence or presence of other yield limiting factors. The relation between leaf area, intercepted radiation and yield are robust (e.g., Monteith, 1977; Jeffries and MacKerron, 1989) and in this case there is thus no reason to assume that the effect of frost would be very different at low yield levels.

Outcomes are highly dependent on the quality of the crop distribution data and this may very well be the greatest source of error in our estimates. Crop distribution data are important to obtain meaningful results from GIS-linked models. National-level crop distribution data are available through FAO, but data at a lower level of aggregation are often hard to get. In order to obtain more precise results in future studies, efforts to assemble crop distribution databases should be intensified, probably increasingly relying on remote sensing. Ideally reliable yield data of a high spatial resolution would also become available.

Our method is an improvement over other studies, which used site-specific weather station data, assuming simulation results were representative of each respective surrounding area (e.g., Stol *et al.*, 1991). A major disadvantage of that approach is that the results become meaningless if weather stations are far apart and in different climatic conditions, which is often the case in mountain regions. If climate data are interpolated before use in a simulation model, more meaningful results can be obtained. A disadvantage of spatial interpolation of weather data is that the gain in spatial resolution generally comes at the cost of temporal resolution. The choice between long-term, daily weather data and interpolated monthly average climate data has important implications, because simulated yield from monthly climate data (as per average weather patterns) does not necessarily coincide with the average simulation calculated from daily data (De Wit and Van Keulen, 1987; Nonhebel, 1994). We interpolated monthly average extreme temperature and its standard deviation to generate daily weather data.

Although we used a very fine resolution grid, there can still be important differences in micro-climate that we have ignored. For example, frost is often worse on valley bottoms than on slopes, because of nocturnal cold air drainage. In some areas, this is reflected by preferential use of hillsides for potato production, despite their shallower soils. In the future, remotely sensed temperature data might prove useful for improve the quality of the extreme minimum temperature surfaces (François *et al.*, 2000), although it would be more important to get good estimates of the future climate in the coming decades for the area of study, and particularly how global warming might affect frost incidence. Our method of creating daily minimum temperature through linear interpolation with one random extreme temperature per month is simple and could be improved. However, our study lacked time series of daily minimum temperature data, a prerequisite for testing more elaborate approaches. This is an important limitation because the results of this study are strongly influenced by even small variation in minimum temperature data. One should realize, however, that ex-ante impact assessment studies could not always be postponed long enough to allow for collection of all the data required for optimal analysis. In this type of study, researchers will have to strike a balance between ideal procedures and availability of data and models (Fig. 8.2).

9. The effect of climate change

Submitted for publication as: Hijmans, R.J. The impact of global warming on potato production.

Abstract

Possible climate change induced changes in global potential potato yields were assessed. Monthly climate data on a 1-degree grid for current (1961-90) and future (2010-2039 and 2040-2069) conditions and a simulation model were used to calculate potential yields. The results were mapped and aggregated weighted by the current distribution of potato area. Between 1961-90 and 2040-69 the global (terrestrial excluding Antarctica) average temperature is predicted to increase between 2.1 and 3.2℃, depending on the climate scenario. The temperature increase is smaller when changes are weighted by the potato area and particularly when adaptation of planting time and varieties is considered (a predicted temperature increase between 1 and 1.4℃). For this period, potato-area-weighted global potential potato yield decreases by 18 to 32% (without adaptation) and by 9 to 18% (with adaptation). At high latitudes, global warming will likely lead to changes in the time of planting, the use of later maturing varieties, and a shift of the location of potato production. In many, but not all, of these areas changes in potato yield are likely to be relatively small, and sometimes positive. Shifting planting time or location is less feasible at lower latitudes, and in these areas global warming could have a strong negative effect on potato production. To mitigate some of the effects of global warming, heat-tolerant potato varieties could be identified or developed for use in (sub)tropical regions.

9.1 Introduction

It is likely that the currently observed trend of global warming, which has been 0.6°C ± 0.2 since 1900, will continue and that the average global temperature will increase by between 1.4 and 5.8°C over the period 1990 to 2100 (Houghton *et al.*, 2001). The impact of this type of climate change will probably lead to a decrease in crop productivity, but with important differences between regions (Rosenzweig and Liverman, 1992; McCarthy *et al.*, 2001).

The effects of climate change on crop production can be complex. Phenological development of plants is generally related to temperature. At higher temperatures crops develop faster and their potential production will be generally lower (this does not apply to plants that are mainly photoperiod sensitive). Yet a shorter growth cycle can also be beneficial, e.g., to escape drought or frost, and the use of late maturing varieties could offset the effect of higher development rates. In environments where low temperatures now limit production, global warming could lead to a beneficial lengthening of the growing season. Moreover, global warming is related to the increase of atmospheric CO_2 concentration, which is likely to increase crop yields, particularly when water limits crop production (Nonhebel, 1993).

Potato is grown in many different environments, but it is best adapted to temperate climates (Haverkort, 1990). At high temperatures (above 17°C; Stol *et al.*, 1991) tuberization diminishes, but potato is also frost sensitive, and severe damage may occur when temperature drops below 0°C. Over the past 50 years, potato production has decreased in Europe and the USA, and increased in developing countries in Latin America and Africa and particularly in Asia (Walker *et al.*, 1999; Hijmans, 2001).

Various authors have used simulation models to study the effect of global warming on potato production. Higher temperatures are predicted to increase potato yields in England and Wales (Davies *et al.*, 1996), Scotland (Peiris *et al.*, 1996) and Finland (Carter *et al.*, 1996), primarily because of a longer growing season. However, an overall yield decrease was predicted for the USA (Rosenzweig *et al.,* 1996).

All these studies on potato were conducted for small areas at high latitudes and their results are difficult to extrapolate to other regions. In this paper, some of the possible effects of climate change on potato production are studied at the global level. A simulation model was used to calculate potential potato yield for the current climate and for predicted future climate of the 2020s and 2050s, using seven future scenarios from five different climate models. The goal of the paper is to identify areas where there is likely to be a strong decline in productivity due to the increase in temperature, and to determine the extent to which heat tolerant potato varieties would be useful to mitigate the effect of climate change in those areas. Only the effect of changes in temperature and solar radiation was considered and not the effect of changes in rainfall or of increased levels of atmospheric CO_2. The results are presented in relation to the current global distribution of the potato crop.

9.2 Materials and methods

Climate data

Average monthly climate data for 1961-1990 (hereinafter referred to as “current climate”), for 2010-2039 (referred to as the “2020s”), and for 2040-2069 (referred to as the “2050s”) were used. For the current climate, data from New *et al.* (1999) were used; for the future climate, seven scenarios from five climate models were used (Table 1). Data were supplied by the Intergovernmental Panel on Climate Change Data Distribution Center (1999).

Predicted changes in climate between the current climate and the 2020s and 2050s were available for each scenario. These changes were superimposed on the data for the (observed) current climate. Hence, there was only one dataset for the current climate, and all changes are relative to that dataset. Daily temperature and radiation data were derived by linear interpolation between the monthly averages. Average temperature was calculated as the mean of the minimum and maximum temperature.

The current and future climate data were on grids with different resolutions (Table 1). All datasets were resampled (statistically disaggregated by interpolation) to a 1 by 1 degree resolution, using IDRISI software (Clark Labs, Worcester, MA, USA) (the data for the current climate were available at a 0.5 by 0.5 degree resolution and were aggregated). Global average temperatures for the current climate and for all scenarios were calculated for terrestrial cells only (16,862 cells), without considering Antarctica, taking the size of each square degree grid cell into account (the size of the one square degree cells decreases with increasing latitude).

Table 9.1 Climate scenarios used in this study.

Scenario	Institute	Code	Forcing details(1)	Original reso-lution in degrees	Solar radiation(2)	Min. and max. temperature
I	Canadian Centre for Climate Modelling and Analysis	CGCM1 (3)	GS	3.75 x 3.75	Yes	Yes
II	Australian Commonwealth Scientific and Industrial Research Organisation	CSIRO-Mk2	GG	5.625 x 3.214	Yes	Yes
III	Australian Commonwealth Scientific and Industrial Research Organisation	CSIRO-Mk2	GS	5.625 x 3.214	Yes	Yes
IV	German Climate Research Centre	ECHAM4	GG	2.8125 x 2.8125	Yes	Yes
V	US Geophysical Fluid Dynamics Laboratory	GFDL-R15	GS	7.5 x 4.5	Yes	No
VI	UK Hadley Centre for Climate Prediction and Research	HadCM2 (3)	GG	3.75 x 2.75	No	Yes
VII	UK Hadley Centre for Climate Prediction and Research	HadCM2 (3)	GS	3.75 x 2.75	No	Yes

(1) Climate change model forcing details: GG = Greenhouse Gas; GS = Greenhouse Gas and Aerosols. All scenarios used assume an increase of 1% atmospheric CO2 per annum (ISO92a).

(2) Yes – data were present; No – data were absent. If solar radiation data were absent these were calculated from extraterrestrial radiation and cloud cover. If minimum and maximum temperature were absent these were calculated from mean temperature (see text). Parameters for these calculations were taken from the current climate in a grid cell.

(3) Mean of four ensembles (identical model experiments performed with the same historical changes and future changes in greenhouse gases, but initiated from different points on the control run).

Simulation model

A slightly adapted version of the LINTUL simulation model as described by Stol *et al.* (1991) and Van Keulen and Stol (1995) was used to calculate potential yield (see below). The model has a temperature-dependent development of the canopy (green ground cover). Biomass production is the product of the fraction green ground cover, incident solar radiation and radiation use efficiency.

Stol *et al.* (1991) estimated tuber yield as a temperature-dependent percentage of the total biomass accumulated during the growing season. In this study, however, the allocation of biomass to tubers was calculated on a daily basis. Allocation to the tubers is initially 0%. After a thermal time threshold is reached, allocation starts increasing linearly with thermal time until the next threshold after which 100% of new biomass goes to the tubers. The values of these parameters were estimated so that harvest index of a mature crop is 80% under normal circumstances (no frost or heat stress), as in the original model. This daily allocation procedure avoids overestimating yield for a prematurely killed crop or for a crop with a very warm end of the growing season. The allocation of biomass to the tubers is also dependent on daily average temperature: it decreases when temperatures are above 15°C and becomes zero at an average temperature above 28°C (Stol *et al.*, 1991). A "heat-tolerant" potato variety was defined by shifting this curve with two degrees (Fig. 9.1).

In the adapted model, radiation use efficiency (RUE) was made dependent on daily average temperature, following Kooman (1995) and Kooman and Haverkort (1995). RUE is highest (2.9 g MJ^{-1} (PAR)) between 15 and 21°C and zero below 2°C and above 34°C, with intermediate values inbetween. Decrease of RUE at high temperatures is due to increasing respiration (Kooman and Haverkort, 1995). Radiation (PAR) above 12 MJ (PAR) m^{-2} day^{-1} was not considered, to account for light saturation (Kooman and Haverkort, 1995).

For each grid cell, the model was run for 12 planting times (with planting at the first day of each month) and for five maturity classes of potato, representing different varieties with early to late senescence. This was repeated for the "heat-tolerant" potato. Maturity classes used were 1000, 1200, 1400, 1600, or 1800 °Cd, expressed as the temperature sum (thermal time) between emergence and harvest, with a base temperature of 2°C. The optimal planting time for a location (grid cell) was determined after the simulations, selecting the month/variety combination that led to the highest yield. Average temperature during the optimal planting time was calculated for each grid cell. To distinguish between the effect of changes in radiation and in temperature, the model was also run for the future climate, while using radiation data for the current climate.

Current and future potential yield were compared for two cases: with and without adaptation. Adaptation is narrowly defined as changes in the month of planting or in the maturity class of the variety. This is sometimes referred to as “autonomous adaptation” in the sense that these are inexpensive and can be carried out at the farm level (McCarthy *et al.*, 2001). In the case “without adaptation”, potential yield for future conditions is calculated for the combination of variety and month of planting that gave the highest yield for the current climate. In the case of “with adaptation”, the highest yield is taken from the 60 (5 varieties × 12 months) months of planting/variety combinations for the future climate scenarios. Hence, in this latter case, the month of planting and variety type in a location (grid cell) can be different for current and future climates.

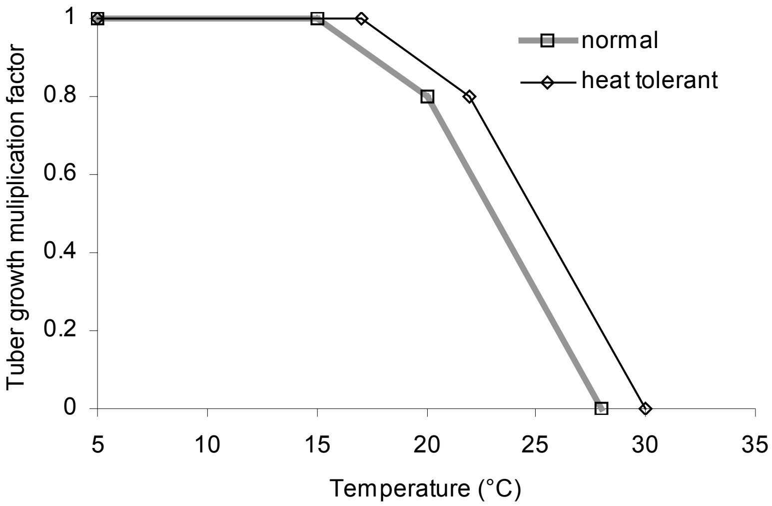


Figure 9.1 The effect of average temperature on yield. The daily amount of biomass allocated to the tubers (in itself a function of development stage) is multiplied with the temperature dependent multiplication factor. For a “normal” (Stol *et al.*, 1991) and a “heat tolerant” potato variety.

Maps and potato distribution data

For each grid cell, the mean potential yield (over all climate scenarios) was compared with the yield as calculated for the current climate. Maps were made of changes in potential yield for the "with adaptation" and "without adaptation" cases for the 2020s and the 2050s. For the 2020s "with adaptation" the changes relative to the minimum and maximum estimated yield (depending on the climate scenarios) were also mapped. Results were also summarized on maps indicating the relative contribution of adaptation to yield and of the potential relative contribution to yield of heat tolerant potatoes.

The maps only include data for the current areas with potato production according to a global 1-degree grid described by Hijmans (2001) (7004 cells with potato area; 42% of the global land area excluding Antarctica; Plate 5A). These potato distribution data were also used to weigh changes in temperature and yield by potato area. Such potato-area-weighted results were obtained by multiplying the grid of potato area with the grids of current and future temperature and yield, and by dividing these with the total global potato area. Thus, in the aggregate results, an area (grid cell) with, for example, 10,000 ha of potatoes would count twice as compared to an area with 5000 ha.

Results were summarized for countries with more than 100,000 ha of potato area. Change in yield and the percentage of grid cells where climate change would lead to higher yields was calculated for these countries.

9.3 Results

Temperature

According to the climate scenarios considered in this study, the increase in global average temperature will be between 1.2 and 1.8°C in the 2020s and between 2.1 and 3.2°C in the 2050s (Fig. 9.2). This increase is higher than the predicted temperature change weighted by potato area, which is between 0.9 and 1.7°C for the 2020s and between 1.6 and 3.0°C for the 2050s. When no adaptation of variety type and month of planting is allowed between current and future conditions, the potato-area-weighted average temperature change during the potato growing period is only a little lower than the average temperature change over the whole year (averaged over all climate scenarios a difference of 0.1°C for the 2020s and 0.2°C for the 2050s). When adaptation of the planting time and variety choice is allowed, however, average temperature change during the potato growing period is much lower than change over the whole year: between 0.6 and 1.1°C for the 2020s and between 1.0 and 1.4°C for the 2050s (Fig. 9.2).

Yield

When no adaptation is allowed, overall simulated global potato yields decrease between 10 and 19% in the 2020s, and between 18 and 32% in the 2050s (Fig. 9.3). With adaptation yields still go down, but the decrease is about 40% less: between 5 and 11% in the 2020s and between 9 and 18% in the 2050s. Adaptation typically consists of a shift of one or two months of the planting time and the use of varieties that are later maturing in terms of thermal time.

These global aggregate data mask differences between regions. Although simulated potato yields decrease almost everywhere the crop is currently grown (Plates. 5BC, 6, and 7A), the magnitude of change differs sharply among potato production areas, and strongly depends upon whether adaptation is considered or not (compare Plates 5B and 6B, and Plates 4C and 7A), and on the climate scenario (Plate 6). Without adaptation, calculated decrease in potential yield is large (> 25%) for many areas, particularly in the 2050s (Plate 5C). When adaptation is allowed, the effect on yield is lower, except for areas in the tropics.

Changes in yield can in most cases be attributed to temperature change alone (in 90% of the grid cells, radiation contributed less than 10% of the change in yield). Although in some restricted areas changes in yield are strongly influenced by changes in radiation, sometimes induced by changes in cloudiness, but more often because of a shift in planting time.

The moderating effect of adaptation on the temperature during the growing period, and hence on yield, is illustrated in Plate 7B. In some areas temperature during the potato growing period will likely go down between the current situation and the 2050s! Yet in other areas temperature increases by more than 2°C. The temperature increase is lowest in areas with most scope for adaptation (compare Plates 7B and C).

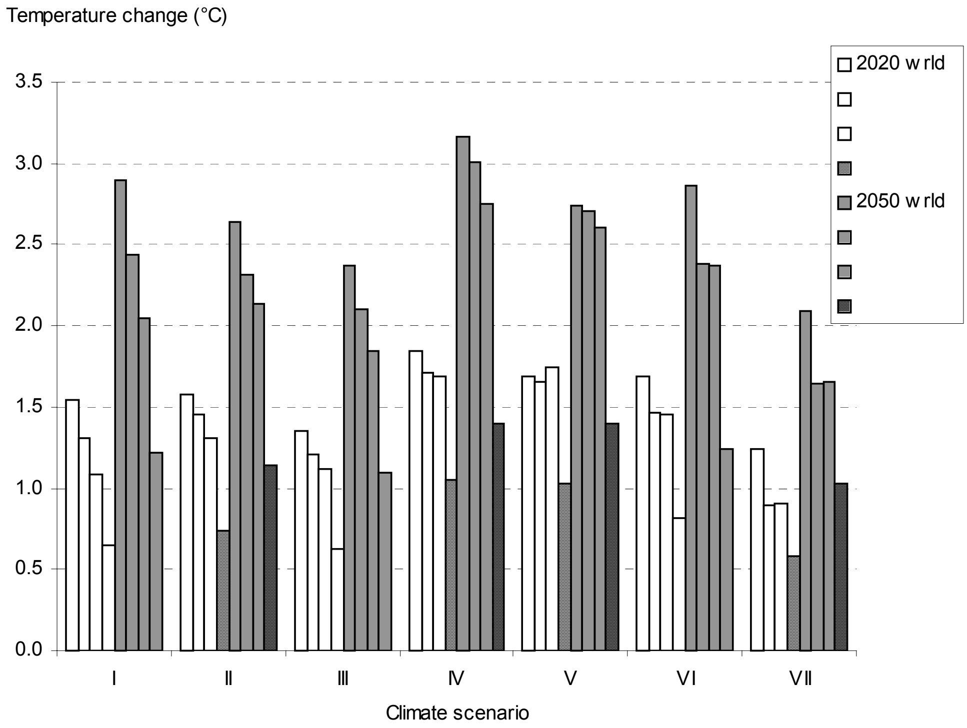


Figure 9.2 Average temperature change in the 2020s and 2050s, relative to the current climate, for seven climate scenarios (see Table 1). For the whole year and world (wrld; the terrestrial areas except Antarctica); the whole year weighted by potato area (pyr); during the potato growing season without adaptation (pgs-c); and during the potato growing season with adaptation of the growing season and varieties (pgs-a).

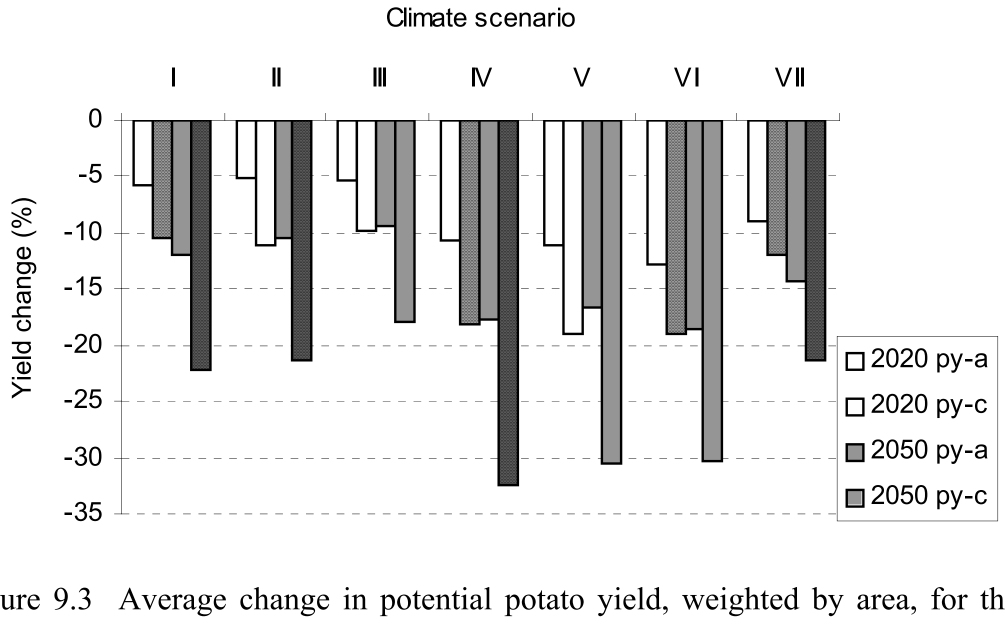


Figure 9.3 Average change in potential potato yield, weighted by area, for the 2020s and 2050s, relative to the current climate, for seven climate scenarios (see Table 1). Without adaptation (py-c); and with adaptation of the growing season and varieties (py-a).

In general, the potato areas in the (sub)tropics will suffer the largest decline of potential potato yield, and there is not much scope for adaptation in these areas (Plate 7C). Note that some of the worst hit areas (yield decline > 50%) have only very little potato area. An area with much potato area and a strong predicted yield decline is a zone from south-east Europe through Russia and Kazakhstan. The areas where global warming would not be a very serious problem, and might be even beneficial, are mostly at high latitudes, such as areas in Canada, Russia (Siberia) and Scandinavia, where global warming will result in longer (frost-free) growing periods, or at very high altitudes in the tropics, such as in the cold highland areas of the Peruvian/Bolivian Altiplano. These are also the areas where there is much land that is currently climatically unsuitable for potato production that will become suitable with global warming. Adaptation is particularly important in parts of southern China where higher temperature increases the opportunity for winter cropping.

Heat tolerance

The current value of increased heat tolerance would be rather small (less then 5% increase in potential yield in most areas), except for some zones in the lowland tropics, with little potato area (Fig. 9.4). However, this situation changes strongly with climate change. In the 2020s and 2050s, in most production zones heat tolerance would increase potential yields with more than 5%. In most production zones, potential yield increase would be over 10% in many zones, particularly in the tropics, but also in a large stretch in eastern Europe and west Asia, and in parts of the USA and Canada. In these northern areas, both adaptation and heat tolerance would be important to allow for high yields.

Results by country

Of the major potato producing countries, most would suffer great losses in potential potato yield, when adaptation is not considered. Bolivia is the only country where potential yield would go up without adaptation, and with adaptation it is predicted to go up with a staggering 77%. In most other major potato producing countries, adaptation mitigates a large part of the climate change induced yield loss. For example for Iran, yield loss decreases from –48% to –13%. China, Peru, Russia and the USA are other notable examples of countries were adaptation could mitigate much of the negative effects of global warming. When considering adaptation, Bangladesh, Brazil, Colombia and Ukraine have the largest decrease in potential yield (more than 20% in 2050). The percentage of areas (grid cells) with yield increase differs strongly between countries. This statistic reflects the possibility to mitigate the effect of climate change by shifting the location of production within existing potato growing areas. It is particularly high (>30%) in Argentina, Canada, China, Japan, UK, Russia, and Spain.

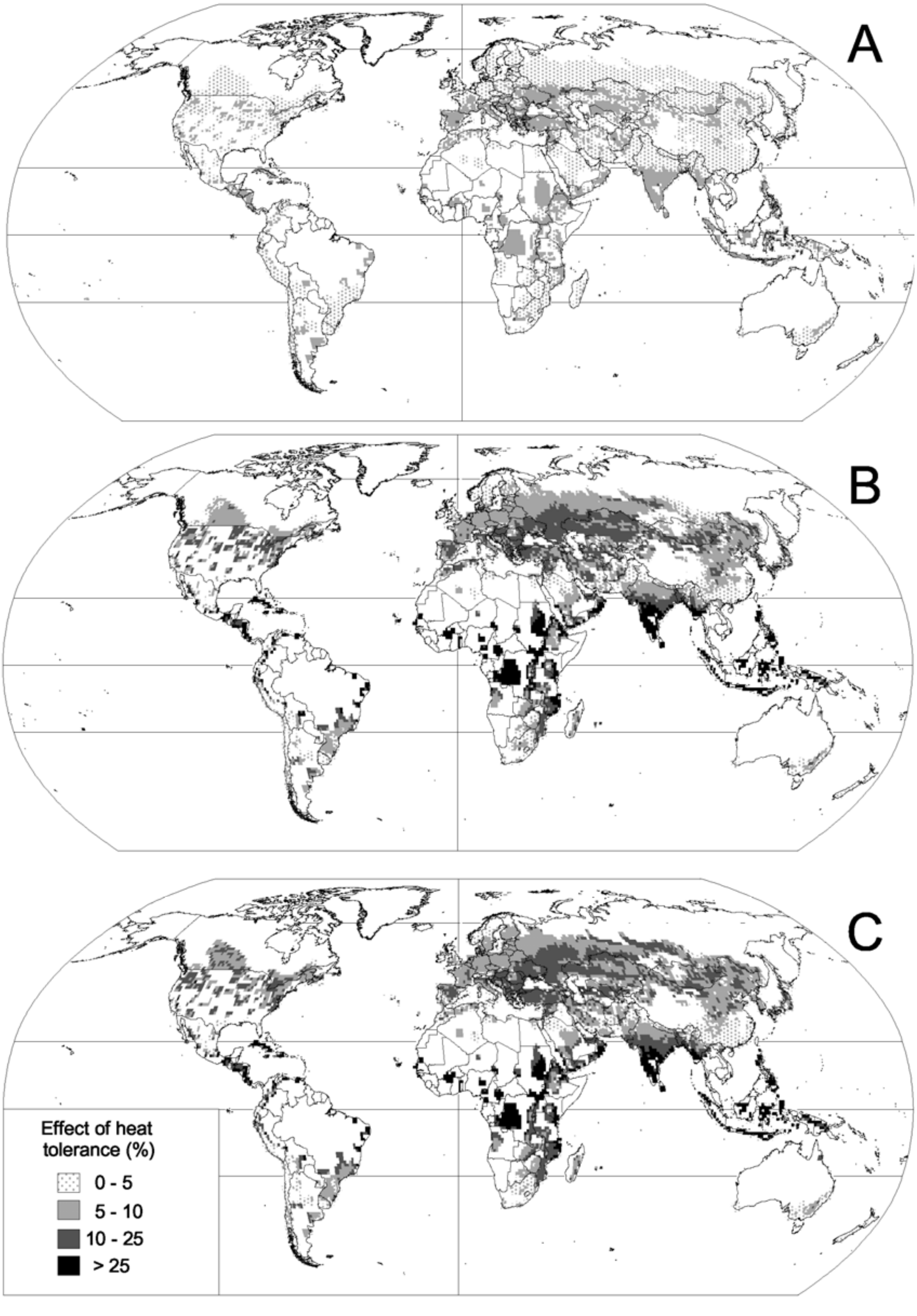


Figure 9.4 The value of heat increased heat tolerance (a 2°C shift in the temperature-tuberization curve; Fig. 9.1) for the 1970s (A), 2020s (B), and 2050s (C).

Table 9.2 Potato area and changes in potential potato yield induced by climate change in the 2050s, and the percentage of the potato area (grid cells) in a country where potential potato yield will increase. Yield changes are weighted by potato area in their respective grid cells.

		Change in potential yield (%)		Areas with yield increase (% of cells)	
Country	Potato area (1000 ha)	Without adaptation	With adaptation	Without adaptation	With adaptation
China	3430	-22.2	-2.5	8.5	30.7
Russia	3289	-24.0	-8.8	12.4	48.4
Ukraine	1534	-30.3	-24.8	0.0	2.7
Poland	1290	-19.0	-16.1	0.0	2.4
India	1253	-23.1	-22.1	0.4	2.0
Belarus	692	-18.8	-16.6	0.0	0.0
United States	548	-32.8	-5.9	1.4	20.1
Germany	300	-19.6	-15.5	0.0	0.0
Peru	263	-5.7	5.8	8.3	13.9
Romania	262	-26.0	-9.9	0.0	19.2
Turkey	207	-36.7	-17.1	0.0	10.4
Netherlands	181	-20.0	-10.9	0.0	0.0
Brazil	177	-23.2	-22.7	0.0	0.0
United Kingdom	169	-6.2	8.1	50.0	57.1
France	168	-18.7	-6.9	4.5	29.9
Colombia	167	-32.5	-30.6	4.5	4.5
Kazakhstan	165	-38.4	-12.4	2.3	9.4
Iran	161	-48.3	-13.3	0.0	21.4
Canada	155	-15.7	4.6	17.9	55.5
Spain	142	-31.4	-6.6	0.0	37.5
Bangladesh	140	-25.8	-24.0	0.0	0.0
Bolivia	131	8.4	76.8	22.6	29.0
Lithuania	126	-13.7	-9.2	0.0	0.0
Argentina	115	-12.9	0.5	11.4	35.2
Nepal	115	-18.3	-13.8	0.0	16.7
Japan	102	-17.4	-0.9	8.8	41.2

9.4 Discussion and conclusions

Because of increases in temperature, potato yields could decrease in many regions. In some areas, mainly in temperate regions, yield decline can partly be avoided through adaptation. Yields may even go up at high latitudes because of a lengthening of the growing season. In some areas, like in parts of Algeria, Morocco, China, and South Africa yield may increase because a warmer climate would allow growing a winter crop (instead of an autumn or spring crop).

There is not much scope for adaptation of potato production in the tropics, where there is not much temperature variation during the year, and in the warmer parts of the subtropics, where potatoes are already grown in the coolest season. In much of the tropical highlands of Africa, temperature is relatively high and stable throughout the year. In the river plains in India and Bangladesh, potato is a winter crop already grown in the coldest season and climate change might slow down the impressive expansion of potatoes in Asia (Walker *et al*., 1999) that is otherwise expected to continue (Scott *et al*., 2000). Warm summers can become problematic in many areas with continental climates, such as Kazakhstan.

The differences between the results obtained when adaptation is taken into account or not, show how there can be important differences between general global climate change and the climate change that a particular crop will experience. These differences were due to adaptation of varieties and planting time. It should be noted, however, that in practice some of these “autonomous” types of adaptation may be not that straightforward. The planting season of a crop also depends on other factors like other crops (particularly in production systems with multiple cropping), water availability, pests and diseases, and markets. Varieties that are better adapted to a changed climate may exist, but are perhaps not available to farmers in a specific region. Many potato varieties are photoperiod sensitive, and this might decrease the temperature sensitivity of their development rate. Changes in the planting time also lead to changes in the photoperiod, and these have not been taken into account.

The identification or the development of potato varieties with increased heat tolerance appears to be important to cope with climate change. Heat tolerance in this context refers to the effect of temperature on tuberization. The effect of temperature on development is also important, but there is much variation for that and increased earliness can be very useful, particularly in intensive cropping systems. Moreover maturity can also be photoperiod regulated and hence less temperature sensitive than assumed in this paper. Given the long time it takes to develop new potato varieties, breeding programs should take future climate change into account. Breeding populations could be tested in warmer environments similar to the future climate of the areas that they are being bred for. Heat tolerance has been found in wild potato species (e.g., Reynolds and Ewing, 1989) and progress in selecting and breeding for heat tolerance in cultivated potato has been reported in the literature (Khanne, 1966; Levy, 1984; Van der Zaag and Demagante, 1988; Tai *et al*., 1994).

In addition to adaptation, there could be a shift of location of potato production between production zones in a country and there could also be a shift toward zones where there is currently no potato production (cf. Leemans and Solomon, 1993). In some (tropical) highland areas potato area could expand into higher areas (e.g., into the Puna and Paramo zones of the Andes). There would also be considerable scope for area expansion in Russia and Canada, but whether this is likely to happen depends on many factors outside the scope of this paper.

The results presented in this paper are somewhat pessimistic because the direct effect of increased atmospheric CO_2 concentration on crop growth was not accounted for. The magnitude and persistence of this effect under field conditions is highly uncertain. Summarizing three studies on the effect of doubling of CO_2 on potato yield, Rosenzweig and Hillel (1998) calculated an expected 51% yield increase. Miglietta *et al*. (1998) found a 40% increase but recent research in multiple locations across the European Union found an average yield increase of 20% (De Temmerman *et al*., 2000). When assuming a 1% annual increase of atmospheric CO_2 (as was used for the climate scenarios in this study), these levels of double CO_2 (~680 ppm) will be reached around 2080. It would seem on the safe side to count on 10-15% higher yields due to increased levels of CO_2 by 2050. This would be enough to offset most of the temperature induced yield decline in many areas.

However, accounting for the effect of CO_2 could obscure the prospect of exploiting the increase of CO_2 for crop production. The most relevant question is not so much where, and to what extent, the positive CO_2 effect compensates for negative temperature effects, but rather how the positive effect of CO_2 on production can be captured to the largest extent.

Unlike in previous studies, detailed crop distribution data were used to interpret the result in this study. This avoided giving too much weight to yield changes in areas with relatively little potato area. Instead of climate data for a limited number of weather stations as used in most studies of the impact of climate change on crop production (e.g., Rosenzweig and Parry, 1994), in this study a comprehensive grid was used. This provides a more complete picture because results from simulation models cannot be easily interpolated. Using monthly climate data on a grid instead of daily weather data may mask the effect of extreme weather events on crop production. Yet this does not seem to be very important for this study because, with the exception of frost incidence, potato (and the model used) is not is highly sensitive to short time fluctuations of the weather (for example, there is no equivalent of heat stress during flowering in cereals).

To improve this study, variation between cultivars in the effect of temperature on tuberization, and the current geographic spread of these cultivars, could be considered. It would be difficult to accommodate all this type of variation in a global study. A next step would be to zoom in to priority areas and evaluate these taking into account local cropping systems and production constraints. In studies of smaller areas it will also be easier to include the effect of (changes in) rainfall because more will be known about the degree to which the crop is irrigated. Additional information about the interaction of drought stress and increased levels of CO_2 on potato growth would also be needed.

Whether changes in potential yield accurately reflect changes in actual yield is also uncertain. There can be many other factors that diminish potato yield, such as lack of water and nutrients, and damage from pests and diseases. For example, the range of the Colorado potato beetle in Europe is expected to increase with global warming (Jeffree and Jeffree, 1996). In Finland, the area infected with the potato cyst nematode, and losses caused by this pest are expected to increase (Carter *et al.*, 1996). The longer growing season would lead to increased late blight problems and fungicide use (Kaukoranta, 1996). These findings support a sustained investment in knowledge-intensive technologies such as integrated pest management and breeding for potatoes with pest and disease resistance.

10. General discussion

10.1 Introduction

Maps are one of the most powerful means of human communication. This is illustrated by the fact that the word “map” is often used as a metaphor (Robinson and Petchenik, 1976). Yet for a long time map making was an activity almost exclusively reserved for specialists, often housed in large and perhaps inaccessible government institutes, such as the defense ministries. Over the last decade, with the advent of geographic information systems (GIS), this situation has changed dramatically.

Given the importance of maps, and their necessarily subjective nature (Monmonier, 1999), the development of GIS has been of scientific, social and economic relevance. It has empowered a large number of people, by letting them integrate various types of data, and make maps. But GIS is about more than just making maps. Its principal strength is that it allows for spatial analysis, i.e., analyzing and describing spatial variation and relationships, and developing spatial models to extrapolate, explore and predict.

Spatial variation is a fundamental characteristic of agriculture, but crop researchers have largely ignored it, particularly at levels beyond the field scale (White *et al.*, 2002). This thesis provides examples of how to deal explicitly with spatial variation over larger areas. New data, methods and tools have been used to address classic questions of crop geography, such as “Where is this crop grown?”, “Where is it most genetically diverse?” and “Where is a production constraint important?”. The availability of GIS allowed tackling these old questions more efficiently and effectively.

This thesis explored the use of GIS for the management of genetic resources, and for crop improvement related agro-ecological zoning. In the final chapter, the contribution made to these two areas is discussed in relation to further research needs. This is followed by a discussion of challenges to the further use of GIS in crop science, particularly in the context of international agricultural research.

10.2 Managing genetic resources

Genetic resources of wild and cultivated plants are the building blocks of our future crop varieties. To assure access to genetic resources, and to avoid their irreversible loss through “genetic erosion”, national and international efforts have been made to safeguard crop genetic resources in genebanks. However, considerable effort is still needed to make these more complete and effective.

Spatial analysis can guide genetic resource collecting efforts, assess the state of conservation of plant genetic resources, and help guide the use of such resources (Guarino *et al*., 2002). Much of the spatial analysis possible depends on the use of genebank databases. These databases are increasingly available—for example, the SINGER database (singer.cgiar.org) which assembles the collections of the CGIAR institutes.

Genebank databases have often been regarded as an administrative repository without much scientific value, a fact which has affected their quality. A persistent problem is that many of the records do not have coordinate data (Hazekamp, 2002). For example, only 9% of the accessions of six major genebanks at the United States Department of Agriculture genebanks have coordinate data (Greene and Hart, 1996). However, 50% of these accessions have a locality description. This means that there is scope for assigning coordinates to at least another 41% of the accessions.

Assigning, checking and improving the coordinates of a germplasm database is tedious and time consuming, and this may be a major barrier to the use of such data in spatial analysis. However, given the dramatic increase possible in data quality (as shown by the data from the case study in Chapter 2) it seems that the effort would be justified for most genebanks. Further research is needed on methods used to assign coordinates on the basis of locality descriptions, and on documenting uncertainty of these coordinates (Wieczorek, 2001).

In some cases fruitful analysis will not be possible, because of insufficient data or because the data has a strong (spatial) sampling bias (Chapter 3). Analysis of bias in the spatial distribution of a collection is useful, as it can help identify gaps in the collection that may need to be filled in the future. However, it can seriously hamper analysis of the data for some other purposes. To correct for biases in the data, extrapolation methods could be used. Collwell and Caddington (1994) describe methods for estimating diversity (species richness) from imperfect data sets; their methods could be integrated into a GIS framework. Jones *et al*. (1997) describe an elaborate method to predict the presence of species in the wild on the basis of the climate distribution of a sample of observations. Because these methods are potentially very useful, their validity merits more research. They could, for example, be tested for data from data-rich areas, or well-studied groups of plants by deducting parts of the data, whether at random or systematically, and comparing the predicted with the “real” distributions.

Imperfect as these databases may be, they can be an important source of information, as this thesis showed by identifying areas of high diversity (Chapter 4). By expanding such diversity studies to cultivated potatoes, and to other crops and their wild relatives, fresh insights might be gained about the process of domestication and about the origin of crop genetic diversity. In the end, this would allow for more refined quantification and testing of Vavilov’s (1926) theories about areas of diversity and origins of agriculture.

There also exist a number of more immediately practical needs that call for further elaboration. An example of such is the need to identify areas where geneflow between a cultivated species and wild relatives might occur, or to identify areas where crops or their wild relatives are likely to have useful traits for use in breeding programs. Based on assumptions about co-evolution and environmental adaptation, hypotheses of the presence of traits in certain areas have often been formulated. However, these have rarely been rigorously tested. Chapter 5 contributes towards a better understanding of how this could be done. More work in this area is needed in the case of potato, and other crops, to allow for generalization about predictivity of ecological and other factors. In this context, there is also a need for further analysis of the effect of differences in germplasm evaluation data, as these are dependent on the timing, location and methods applied.

Hodgkin and Ramanta Rao (2002) singled out GIS and biotechnology, as the two major new fields of research in plant genetic resource management. However, most publications in this area still ignore the spatial dimension of the data they describe. While it is quite common to present distribution maps of samples that were analyzed, for example with molecular markers, it is uncommon to spatially analyze the results. In many cases spatial analysis would be hindered because of poor spatial sampling, as this is often not considered at the outset of the investigation.

Grid based analysis approaches are to be preferred above most other spatial techniques to summarize point distributions. Nevertheless, there are some problems associated with the technique that need further elaboration. Particularly important is that the results are sensitive to the size of the grid cell (scale effect), and methods are needed to objectively determine the best scale of analysis or to develop indices or methods that allow the presentation of the results across scales.

10.3 Agro-ecological zoning

Spatial variation seriously complicates a number of related activities, such as site selection for experiments and the extrapolation of experimental results, as well as the targeting of technologies, and the prediction and evaluation of the impact of agricultural research.

Agro-ecological zoning (AEZ) approaches can be used to come to grips with environmental heterogeneity and to define quasi-homogeneous domains. AEZ has been used to divide the world into generic zones that will remain of value as a general reference tool. However, flexible zoning approaches are emerging to address specific questions for specific crops (as illustrated in Part II of this thesis). Such new approaches are promising, but there are still many assumptions and uncertainties associated with them.

To improve AEZ approaches, there is a need for more systematic data about crop management (e.g., timing of planting and harvesting, varieties used) and for models that can predict these aspects of crop management. This is particularly difficult in complex cropping systems in the tropics, where a calculated “optimal production season” may not be the season actually used. This may be true for a number of reasons, including the needs of other crops in the cropping system, availability of irrigation, or disease pressure. Insight into the validity of this type of AEZ study, which is often carried out at a global or continental level, would be best obtained through more detailed country- or region-level studies, for which detailed data is available, or where such databases can be more easily assembled.

In Chapter 8 it was shown how simulation models and GIS can be used for ex-ante impact assessment, as a way to estimate objectively the potential effect of technological change on production. It should be noted, however, that for many technologies, applicable simulation models are not available. Depending on the goal of the study, simulation models can be replaced by simpler decision models or by clustering approaches (Duchateau *et al.*, 1997), although this does not solve the problem of estimating yield changes. In cases where a new technology (variety) does not cause yield changes, but gives other benefits, (e.g., nutritional benefits), this would not be a limitation (Low *et al.*, 2001).

When simulation models are available, they may be too complex to use because of a lack of documentation or shared programming approach. Also, they could produce invalid results when used for areas far removed from the areas for which they were developed. To verify their validity, and improve the models, comprehensive databases of experimental results are needed. For example, to improve the climate change study (Chapter 9), data on variation between cultivars in terms of the effect of temperature on tuberization, and the current geographic spread of these cultivars, would be important.

The use of GIS and simulation modeling approaches should thus not be considered as a method by which experimental fieldwork could be replaced. On the contrary, they highlight the need for systematic, documented, and accessible experimental data. In general, this type of secondary data is scarce. Available experimental data are often incomplete, in that there are only a few variables available, and often there are no weather or soil data. When trials are carried out to investigate disease resistance, basic growth data are often not taken, and when general adaptation is investigated disease incidence is often not recorded. To make the best use of the many experiments that are carried out, further use of “minimum data sets” concepts such as developed by the IBSNAT project (Tsuji et al., 1998) would be useful. This would include recording observations on ground cover (or leaf area index), biomass accumulation, quantitative information about important pests and diseases, inputs used, on-site recorded weather data, and basic soils data. There is, of course, a cost attached to this scheme, but this amount is likely to be small in comparison to the overall costs of running an experiment.

Experimental data storage, exchange standards and software should be improved and used more. A laudable example of an attempt at the compilation of experimental meta-data is that of crop networks of the Global Change and Terrestrial Ecosystem (GCTE) project of the International Geosphere-Biosphere Programme (www.gcte.org)—for potato data see GCTE (1997). In the Ecocrop II database of the FAO, actual experimental data is collated for different crops.

In addition to ex-ante impact assessment, as explored in this thesis, spatial analysis has a largely unexplored role to play in ex-post impact assessment (Nelson, 2002). Differences in the adoption and impact of a technology between different sites cannot be adequately explained if environmental variables are not taken into account. In addition, spatial autocorrelation (or neighbor effects) of adoption needs to be taken into account, because innovations often spread from neighbor to neighbor (Case, 1992). A recent example of both ex-ante and ex-post impact assessment in which GIS was used for environmental stratification is a study on maize technology development and transfer in Kenya by Hassan (1998).

Another area of research in which spatial analysis could play a larger role is the study of adaptation of crop varieties. Such studies are often carried out through multilocational trials, and differences in the response (yield) of the varieties between sites are attributed to an environmental effect or, when responses are complex, to genotype by environment interaction. However, in few cases is there sufficient attention given to what the differences in environment really are, and which factors, climatic or otherwise, determine the different responses.

10.4 Challenges for further GIS adoption

There is a large gap between the theoretical usefulness and the practical use of spatial analysis and GIS; a number of reports have indicated that GIS has great potential (it is a "powerful tool") but that the rate of its use is lower than expected in agricultural research and related areas of work. This discrepancy has been noted for geographic targeting of poverty alleviation (Bigman and Fofack, 2000); for agronomy (White *et al.*, 2002); and for agricultural research impact assessment (Nelson, 2002). Similarly, Guarino *et al.* (2002) underlined the potential that GIS-based analysis has for the conservation of crop genetic resources but they noted that GIS technology is not being used much by crop genetic resource conservation programs. Reasons mentioned for this apparent lack of adoption include issues of data quality and access to data, and lack of appropriate analytical methods and software. Aspects of these issues are discussed below.

Data availability

For the type of GIS studies discussed in this thesis, one is critically dependent on many different types of secondary data. To expand and improve this type of studies, there is a need to continue improving global climate, soils and land use databases. For example, the studies in Part II of this thesis showed how crop distribution data is needed for the meaningful interpretation of agro-ecological zoning studies. As crop distribution data is increasingly used for impact assessment and allocation of funds, it is becoming more important to strive for higher quality data, or at least for insight into the uncertainty associated with these data. For more precise results in future studies, international collaborative efforts to assemble crop distribution databases should be intensified. The use of remotely sensed data to improve the spatial resolution of crop distribution databases, by disaggregating statistical data, might also prove an important approach (EUROSTAT, 2001).

Even when data exists, they may not be readily available. In many cases, access is restricted by a lack of organization. Data exists in many different institutions that do not, by themselves, have the capacity, or the will, to make that data widely available. Fortunately, there are some initiatives intended to make this type of spatial data available (e.g., Hodson *et al.*, 2002). These initiatives can flourish, as the incentive exists to share data among different sectors that have data needs in common (and the data may be used for very different applications within these sectors).

Funding agencies have an important role to play to increase data availability, for example, by stipulating that funding for the compilation of (spatial) databases is provided only under the condition that the data will be placed in the public domain, and by stimulating the creation of national and international clearing houses. Ethical reasons aside, it is poor use of public funds to invest in GIS databases without ensuring their availability and use on a wider scale. Mechanisms need to be found and used to further stimulate data exchange. An obvious role for international organizations, such as the FAO and CGIAR institutes, is the compilation of global databases, which should then be made widely available.

There is a need for international collaboration, and data compilation and exchange mechanisms. Although this has been discussed by many, not much progress has been made. At the national level, very different approaches have been taken in legislation governing database ownership and protection (Cho, 1998). The USA stands out because of its large number of national and international digital database development activities and a very high degree of open access to that data. The situation in the EU is not quite as bright, with great variation in legislation between countries. An issue that particularly needs to be addressed in many countries is the difficult or costly access to public data. A lack of access to public data constitutes a barrier to the movement towards a democratic “information society”. Analysis of data leads to information; information leads to knowledge, and knowledge is power.

Methods

The GIS practitioner who aims to contribute to the program of a crop improvement institute will have to provide timely answers to urgent questions. This fact implies the need to strike a balance between ideal procedures and availability of data and methods, as was argued in Chapter 7. At the same time there exists a need to continue improving the methods used, through more in depth methodological research and database development. In a similar vein, Kristjanson and Thornton (2002) pointed out that each of their GIS based impact assessment studies of livestock technology highlighted a lack of data, but also improved their understanding of the system studied.

The studies in this thesis follow this pattern. They are not the final word on the topic, but constitute incremental improvements of our understanding of complex issues. This is illustrated by the differences between Chapters 8 and 9. In Chapter 8 it was shown that potential existed for frost tolerant potato varieties to have a significant impact in the Altiplano of Peru and Bolivia. However, in Chapter 9 it was shown that this area will benefit strongly from global warming, which suggests that an investment in frost tolerant potato varieties is much less warranted!

Besides the argument that a need exists for the provision of quick answers to practical problems, there is a need for further research into the science of GIS. In the worst case scenario, GIS could become wizardry, impossible to scrutinize, no more than colorfully illustrated guesswork. There is a need to further strengthen the scientific foundations of the different GIS applications that are emerging, and a number of issues have been mentioned above, and in the previous chapters. Additional research should include the development of basic building blocks, such as improved weather generators and interpolation methods (e.g., Hartkamp, 2002), but also scrutinization of analytical methods through sensitivity analysis and analysis of error propagation (e.g., Heuvelink, 1998).

Software

Data and methods alone are not enough; little can be done without software. Over the past ten years there has been a revolution in GIS software development. GIS software has been transformed from cumbersome command line programs that ran on mainframe computers, to programs with graphical interfaces on PCs and these are being continually refined. We are currently also witnessing a movement towards distributed mapping applications over the Internet, which are accessed and manipulated with a browser.

While commercial GIS software can still be very expensive, and is out of reach of many scientists (particularly those in developing countries), there are a number of (simple) GIS software programs that are available free, or at low cost. Such software has boosted the wider use of GIS. These software packages are relatively easy to learn, can improve spatial awareness amongst researchers, and are an entry point for the efficient use of more complex GIS software when funds permit (Hodson and White, 2002).

Particularly relevant for the crop science community is the development of GIS software for specific analysis such as FloraMap (Jones and Gladkov, 1999) and DIVA-

GIS (Hijmans *et al.*, 2001a), for the analysis of different aspects of biodiversity, and the Almanac Characterization Tool (Corbett *et al.*, 2000), an example of a more agro-ecological, zoning oriented GIS software program which is low in cost. Without such efforts, is unlikely that many non-specialists would enter the GIS arena.

For a number of reasons, such as idiosyncratic data-handling, and different programming approaches, it can still be problematic to link models, like crop growth models, to GIS databases. In this respect much more can still be done in terms of standardization of scientific software, and the development of software components that can be re-used.

People and institutional setting

Despite all the shortcomings and needs described above, conditions for GIS use are improving steadily. The extent to which GIS will be used in crop science will depend greatly on the people in the profession. Because, no matter how cheap and easy to use the software becomes, for the useful application of GIS, practitioners will also need a combination of computer skills, scientific background in the specific subject studied, and a dose of creativity to develop methods and communicative maps. Thus there will always be a price, in terms of acquiring these skills, and this may very well be the prime reason for GIS being less readily adopted than might be expected.

While the application of spatial analysis will always remain a considerable investment, with improvements in software, better-understood and more refined methods, and more data, the shape of the learning curve is changing. GIS may soon, like statistics, become integrated into mainstream research, allowing every crop scientist to make maps and do basic spatial analysis, while relying on the input of specialists for more complex analysis and advice.

The use of GIS is still rare in many crop science programs in both universities and national research institutes, but use of GIS is now well established in the international research centers of the CGIAR. It must be noted, however, that in some cases "GIS units" have been somewhat divorced from the mainstream research agenda of the centers. Most GIS related research responds to the relatively recent interest in natural resource management, and often is not greatly integrated with the commodity (crop science) research. In those cases, GIS (and natural resources) is only an addition to the research agenda, and an opportunity to strengthen crop improvement and management research agenda is missed. This situation may have partly been caused by the background of most GIS specialists (which is not crop science), as well as by a lack of interest in spatial analysis shown by the crop scientists. It appears, however, that this situation is changing; it seems that geographers have learned about crops, and that the supply of spatial analysis is creating the demand for more.

10.5 Conclusion

We have advanced greatly since Klages (1949) advocated "ecological crop geography". Although GIS and spatial analysis are not used in the day-to-day decision making of most crop research institutes, we seem to be moving in that direction—although much remains to be done to arrive at that goal. Wide availability of GIS is a relatively recent phenomenon, of roughly the last five years. Thus, use of GIS could simply be "lagging" behind, due to lack of knowledge and capacity amongst the current generation of researchers. Alternatively, the expectations of some GIS aficionados could simply be too high. It is probably too early to tell.

GIS adoption is, of course, not a goal in itself. The goals of international agricultural research include improving food security, reducing poverty, protecting the environment, and conserving and using biodiversity. People do not eat maps, so the crucial question is whether agricultural research can be made more efficient and effective using spatial analysis and GIS. To answer this question, the "added value" provided by these approaches must be evaluated. The value of information is difficult to assess, but it would be of interest to compare decision-making with and without spatial analysis.

The perceived "barriers to entry" for GIS use by crop scientists may still be high, but the current revolution in GIS technology is putting both the data and the analytical tools within the reach of rapidly increasing numbers and types of users. The crop science community needs to use the analytical capacities of the new computer tools to their fullest extent, in order to have the strongest possible positive impact on biodiversity conservation and agricultural development.

References

Aguirre, G., J. Calderon, D. Buitrago, V. Iriarte, J. Ramos, J. Blajos, G. Thiele, and A. Devaux, 1999. Rustic seed bed, a bridge between formal and traditional potato seed systems in Boliva. In: International Potato Center. Impact on a changing world: Program report, 1997-98. International Potato Center, Lima, Peru. p.195-203.

Allard, R.W., 1970. Population structure and sampling methods. In: O.H Frankel and E. Bennett (eds.). Genetic resources in plants: their exploration and conservation. F.A. Davis Company, Philadelphia, USA. p. 97-107.

Alston, J.M., G.W. Norton, and P.G Pardey, 1995. Science under scarcity: principles and practice for agricultural research evaluation and priority setting. Cornell University Press, Ithaca, New York, USA.

Bacon, L.B., R.G. Hainsworth, N. Jasny, C.M. Purves, L. Violin, and C.E. Whipple, 1948. Agricultural Geography of Europe and the Near East. Miscellanneous publication No. 665. Office of Foreign Agricultural Relations, United States Department of Agriculture. United States Government Printing Office, Washington D.C., USA.

Barrientos, M., E. Mol, G. Perruzo, A. Contreras, and M. Alberdi, 1994. Responses to cold of Chilean wild *Solanum* species. Environmental and Experimental Botany 34:47-54.

Bennett, E., 1970. Tactics of plant exploration. In: O.H. Frankel and E. Bennett (eds.). Genetic resources in plants: Their exploration and conservation. F.A. Davis Company, Philadelphia, USA. p. 157-179

Bertin, J., J.J. Hémardinquer, M. Keul, and W.G.L. Randles, 1971. Atlas of food crops. Geographical and chronological survey for an atlas of world history. École Pratique des Hautes études-Sorbonne. VIe Section: Sciences économiques et Sociales. Centre de Recherches Historiques et Laboratoire de Cartographie, Paris, France.

Bigman, D. and H. Fofack (eds.), 2000. Geographical targeting for poverty alleviation: methodology and applications. World Bank regional and sectorial studies. World Bank, Washington DC, USA.

Blackburn, T.M., and K.J. Gaston, 1996. Spatial patterns in the species richness of birds in the New World. Ecography 19: 369-376.

Bonham-Carter, G., 1994. Geographic information systems for geoscientists. Modelling with GIS. Computer Methods in the Geosciences 13. Pergamon/Elsevier, London, UK.

Bonierbale, M., W. Ganal, and S.D. Tanksley, 1990. Applications of restriction fragment length polymorphisms and genetic mapping in potato breeding and molecular genetics. In: M.E. Vayda and W.D. Park (eds.), The molecular and cellular biology of the potato. CABI, Wallingford, UK. p 13-24.

Booth, T.H., J.A. Stein, H.A. Nix, and M.F. Hutchinson, 1989. Mapping regions climatically suitable for particular species: an example using Africa. Forest ecology and management 28: 19-31.

Booy, A.H., 1961. Verband tussen nachtvorstschade aan aardappelen en bodem-gesteldheid op Drentse dalgrond. Landbouwvoorlichting (januari 1961): 10-18.

Brown, A.H.D. and D.R. Marshal,. 1995. A basic sampling strategy: theory and practice. In: L. Guarino, V. Ramanatha Rao, and R. Reid (eds.), Collecting plant genetic diversity, technical guidelines. pp. 75-91. CABI, Wallingford, UK.

Brush, S.B., 1989. Rethinking crop genetic resource conservation. Conservation Biology 3:19-29.

Brush, S.B, 1995. In situ conservation of landraces in centers of crop diversity. Crop Science 35:346-354.

Canahua, A., and P.C. Aguilar, 1992. Agroecologica de las papas amargas en Puno. In: J. Rea, J.J. Vacher (eds.). La papa amarga. Primera mesa redonda: Perú - Bolivia, La Paz 7-8 Mayo 1991. ORSTOM, La Paz. p. 57-61.

Carrasco, E., A. Devaux, W. García, and R. Esprella, 1997. Frost tolerant potato varieties for the Andean Highlands. In: International Potato Center, Program Report 1995-1996. International Potato Center, Lima, Peru. p. 227-232.

Carter, S.E., L.O. Fresco, and P.G. Jones, with J.N. Fairbairn, 1992. An atlas of cassava in Africa. Historical, agroecological and demographic aspects of crop distribution. Centro Internacional de Agricultura Tropical, Cali, Colombia.

Carter, T.R., R.A. Saarikko, and K.J., Niemi, 1996. Assessing the risks and uncertainties of regional crop potential under a changing climate in Finland. Agriculture and Food Science in Finland 5: 329-350.

Case, A., 1992. Neighborhood Influence and Technological Change. Regional Science and Urban Economics 22(3): 491-508

Chapman, C.D.G., 1984. On the size of a genebank and the genetic variation it contains. In: J.H.W. Holden and J.T. Williams (eds.). Crop genetic resources: Conservation and evaluation. George Allen and Unwin, London, UK. p. 102-119.

Chapman, C.D.G., 1989. Collection strategies for the wild relatives of field crops. In: A.H.D. Brown, O.H. Frankel, D.R. Marshall, and J.T. Williams. The use of plant genetic resources. Cambridge University Press, Cambridge, UK. p. 263-279.

Cho, G., 1998. Geographic information systems and the law: Mapping the legal frontiers. John Wiley and Sons, Chichester, UK.

CIESIN, 2000. Global gridded population database. Available at http://www.ciesin.org/

Cochran, W.G., 1977. Sampling techniques. John Wiley and Sons, Chichester, UK.

Contreras M.A., and D.M. Spooner, 1999. Revision of *Solanum* section *Etuberosum* (subgenus *Potatoe*). In: M. Nee, D.E. Symon, R.N. Lester, and J. P. Jessop (eds.), Solanaceae IV. Royal Botanic Gardens, Kew, UK. p. 227-245

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

Corbett, J.D., and R.F. O'Brien. 1997. The spatial characterization tool – Africa v 1.0. Texas Agriculture Experiment Station, Texas A&M University System, Blackland Research Center Report No.97-03, CDROM Publication, Blackland, TX, USA.

Corbett, J.D., S.N. Collis, B.R. Bush, E.I. Muchugu, R.Q. Jeske, R.E. Martinez, M.F. Zermoglio, E. Martinez-Romero, J.W. White, and D.P. Hodson. 2001. Almanac Characterization Tool: a resource base for characterizing agricultural, natural and human environments. Texas Agricultural Experiment Station, Texas AandM University System, Blackland Research and Extension Center Report No. 99-06. Temple, TX, USA.

Cox A.E., and E.C. Large, 1960. Potato late blight epidemics throughout the world. Agriculture Handbook Series 174. United States Department of Agriculture. Washington DC, USA.

Cressie, N.A.C., 1991. Statistics for spatial data. John Wiley and Sons, New York, New York, USA.

Crook, F.W., 1993. Under reporting of China's cultivated area: implications for world agricultural trade. In: China, international agriculture and trade reports. Situation and outlook series, RS-93-4. United States Department of Agriculture, Economic Research Service, Washington, DC, USA.

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

Davies, A., T. Jenkins, A. Pike, J. Shaq, I. Carson, C.J. Pollock, and M.I. Parry, 1996. Modelling the predicted geographic and economic response of UK cropping systems to climate change scenarios: the case of potatoes. Aspects of Applied Biology 45: 63-69.

De Bouet du Portal, D., 1993. Etude des gelees sur l'Altiplano Bolivien. ORSTOM and ENSAM, Montpellier, France.

De Candolle, A., 1884. Origin of cultivated plants. Trench, London, UK.

De Temmerman, L., M. Bindi, J. Craigon, A. Fangmeier, A. Hacour, H. Pleijel, K. Vandermeiren, V. Vorne, and J. Wolf, 2000. Changing climate and potential impacts on potato yield and quality. Veterinary and Agrochemical Research Centre, Tervuren, Belgium.

De Wit, C.T., and H. van Keulen, 1987. Modeling the production of field crops and its requirements. Geoderma 40: 253-265.

Dearborn, C.H., 1969. Alaska Frostless: an inherently frost resistant potato variety. American Potato Journal 46, 1-4.

Debener, T., F. Salamini and C. Gebhardt, 1990. Phylogeny of wild and cultivated *Solanum* species based on nuclear restriction fragment length polymorphisms (RFLPs). Theoretical and Applied Genetics 79:360-368.

Del Rio, A.H., J.B. Bamberg, Z. Huamán, A. Salas, and S.E. Vega, 2001. Association of ecogeographical variables and RAPD maker variation in wild potato populations of the USA. Crop Science 41: 870-878.

Devaux A., and A.J. Haverkort, 1987. The effects of shifting planting dates and mulching on late blight (*Phytophthora infestans*) and drought stress on potato crops grown under tropical highland conditions. Experimental Agriculture 23, 325-333.

Dixon, J., A. Gulliver, with D. Gibbon, (Principal editor: M.Hall). Farming systems and poverty. Improving farmers' livelihoods in a changing world. FAO and World Bank, Rome, Italy.

Duchateau, L., R.L. Kruska, and B.D. Perry, 1997. Reducing a spatial database to its effective dimensionality for logistic-regression analysis of incidence of livestock disease. Preventive veterinary medicine 32: 207-218.

Duckham, A.N., and G.B. Masefield, 1970. Farming Systems of the World. Chatto and Windus, London, UK.

EIU and CDCP (Economist Intelligence Unit and the Cartographic Department of the Clarendon Press), 1954. Oxford economic atlas of the world. Oxford University Press, Oxford, UK.

EIU and CDCP (Economist Intelligence Unit and the Cartographic Department of the Clarendon Press), 1972. Oxford economic atlas of the world. Oxford University Press, Oxford, UK.

ESRI (Environmental Systems Research Institute), 1999. ESRI data and maps. Database. ESRI, Redlands, California.

Estrada, R.N., 1982. Breeding wild and primitive potato species to obtain frost-resistant cultivated varieties. In: P.H. Li and A. Sakai (eds.). Plant cold hardiness and freezing stress: Mechanisms and crop implications, Volume 2. Academic Press. New York. p. 615-633.

EUROSTAT, 2001. Towards agri-environmental indicators – Integrating statistical and administrative data with land cover information. European Commission, Brussels, Belgium.

FAO (Food and Agriculture Organization of the United Nations). 1978-81. Report on the agro-ecological zones project. World soils resources report 48, Vol 1-4. FAO, Rome, Italy.

FAO, 2000a, 2002. Database at http://apps.fao.org

FAO, 2000b. Database at http://geoweb.fao.org

Finch, V.C. and O.E. Baker, 1917. Geography of the world's agriculture. Government printing office, Washington, DC, USA.

Flanders, K.L., J.G. Hawkes, E.B. Radcliffe and F.I. Lauer, 1992. Insect resistance in potatoes: sources, evolutionary relationships, morphological and chemical defenses, and ecogeographical associations. Euphytica 61:83-111.

Flanders, K.L., E.B. Radcliffe and J.G. Hawkes, 1997. Geographic distribution of insect resistance in potatoes. Euphytica 93:201-221.

Forbes G.A., R.J. Hijmans, and R.J. Nelson, 1998. Potato blight: a world problem. Abstracts of the 7th International Congress of Plant Pathology, 9-16 August 1998. British Society for Plant Pathology, Edinburgh, UK On-line at [www.bspp.org.uk/icpp98/abstracts/5.1/9S.html].

François, C., R. Bosseno, J.J. Vacher, and B. Seguin, 2000. Frost risk mapping derived from satellite and surface data over Bolivian Altiplano. Agriculture and Forest Meteorology, 95(2): 113-137.

Frankel, O.H., A.H.D. Brown, and J.J. Burdon, 1995. The conservation of plant biodiversity. Cambridge University Press, Cambridge, UK.

Frère, M., J.Q. Rijks, and J. Rea, 1975. Estudio agroclimatológico de la zona Andina. FAO, Rome, Italy.

Fry W.E., A.E. Apple, and J.A. Bruhn, 1983. Evaluation of potato late blight forecasts modified to incorporate host resistance and fungicide weathering. Phytopathology 73: 1054-1059.

G-DRU (Grupo Interinstitucional de Desarrollo Rural), 1996. Anuario Estadístico del Sector rural 1995-1996 de Bolivia. Centro de Información para el Desarrollo, La Paz.

Gaston, K.J., 1996. Species richness: measure and measurement. In: K. J. Gaston (ed.). Biodiversity, a biology of numbers and diference. Blackwell Science, London, UK. p. 77-113.

Gaston, K.J., and P.H. Williams, 1996. Spatial patterns in taxonomic diversity. In: K. J. Gaston (ed.). Biodiversity, a biology of numbers and diference. Blackwell Science, London, UK, p. 202-229.

GCTE, 1997. Report No. 9. GCTE Focus 3 Potato Network: Model and Experimental Metadata. GCTE Focus 3 Office, Wallingford, UK.

Geng S., F.W.T. Penning de Vries, and I. Supit, 1986. A simple method for generating daily rainfall data. Agricultural and Forest Meteorology 38: 363-376.

Giannattasio, R.B., and D.M. Spooner. 1994a. A reexamination of species boundaries between *Solanum megistacrolobum* and *S. toralapanum* (*Solanum* sect. *Petota*, series *Megistacroloba*): morphological data. Systematic Botany 19: 89-105.

Giannattasio, R.B., and D.M. Spooner. 1994b. A reexamination of species boundaries and hypotheses of hybridization concerning *Solanum megistacrolobum* and *S. toralapanum* (*Solanum* sect. *Petota*, series *Megistacroloba*): molecular data. Systematic Botany 19: 106-115.

Gregor, H.F. 1970 Geography of agriculture. Themes in research. Prentice-Hall. Engelewood Cliff, USA.

Greene, S.L., and T. Hart, 1996. Plant Genetic Resource Collections: an Opportunity for the Evolution of Global Data Sets. Third International Conference/Workshop on Integrating GIS and Environmental Modeling Santa Fe, New Mexico, USA. (CD-ROM and web publication;http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/ greene_stephanie/sgreene.html).

Grigg, D.B., 1995. An introduction to agricultural geography (2nd edition). Routledge, London, United Kingdom.

Grünwald N.J., O.A. Rubio-Covarrubias, and W.E. Fry, 2000. Potato late-blight management in the Toluca Valley: Forecasts and resistant cultivars. Plant Disease 84: 410-416.

Gryseels G., C.T. de Wit, A. McCalla, J. Monyo, A. Kassam, E. Crasswell, and M. Collinson, 1992. Setting agricultural research priorities for the CGIAR. Agricultural Systems 40(1): 59-103.

Guarino, L., 1995. Mapping the ecogeographic distribution of biodiversity. In: Guarino, L., V. Ramanatha Rao and R. Reid (eds.), Collecting Plant Genetic Diversity, Technical Guidelines, CABI. Wallingford. p. 287-328.

Guarino, L., N. Maxted, and M. Sawkins, 1999. Analysis of geo-referenced data and the conservation and use of plant genetic resources. In: S.L. Greene and L. Guarino (editors). Linking genetic resources and geography: emerging strategies for conserving and using crop biodiversity. American Society for Agronomy Special Publication. ASA, CSSA, and SSSA, Madison, Wisconsin, USA.

Guarino, L., A. Jarvis, R.J. Hijmans, and N. Maxted, 2002. Geographic information systems (GIS) and the conservation and use of plant genetic resources. In: J.M.M. Engels, V. Ramantha Rao, A.H.D. Brown, and M.T. Jackson (eds.). Managing plant genetic diversity. CABI publishing, Wallingford, UK. p 387-404.

Hall, J.B., 1994. Mapping for monographs: Baselines for resource development. In: Miller, R.I. (ed.). Mapping the Diversity of Nature. Chapman and Hall, London, UK. p. 21-35.

Han S., R.G. Evans, T. Hodges, and S.L. Rawlins, 1995. Linking a geographic information system with a potato simulation model for site-specific crop management. Journal of Environmental Quality 24: 772-777.

Hanotte, O., D.G. Bradley, J.W. Ochieng, Y. Verjee, E.W. Hill, and J.E.O. Rege, 2002. African pastoralism: genetic imprints of origins and migrations. Science 296:336-339.

Hansen J.G., B. Andersson, and A. Hermansen, 1995. NEGFRY - A system for scheduling chemical control of late blight in potatoes. In: L.J. Dowley, E. Bannon, L.R. Cooke, T. Keane, and E. O'Sullivan (eds.). *Phytophthora infestans* 150. Boole Press, Dublin, Ireland. p. 201-208.

Harlan, J.R., 1992. Crops and man. American Society of Agronomy and Crop Science Society of America, Madison, WI, USA.

Harlan, J.R. and J.M.J. de Wet, 1971. Towards a rational classification of cultivated plants. Taxon 20: 509-517.

Hartkamp, A.D., 2002. Learning from biophysical heterogeneity: inductive use of case studies for Maize cropping systems in Central America. PhD thesis. Wageningen University, Wageningen, Netherlands.

Harrison J.G., 1992. Effects of the aerial environment on late blight of potato foliage: a review. Plant Pathology 41, 384-416.

Hassan, R.M. (ed.), 1992. Maize technology development and transfer. A GIS application for research planning in Kenya. CABI, Wallingford, UK.

Haverkort, A.J., 1990. Ecology of potato cropping systems in relation to latitude and altitude. Agricultural Systems 32: 251-272.

Hawkes, J.G., 1979. Evolution and polyploidy in potato species. In: J. G. Hawkes, R. N. Lester, and A. D. Skelding (eds.). The biology and taxonomy of the Solanaceae Academic Press, London, UK. p. 637-645.

Hawkes, J.G., 1990. The potato: evolution, biodiversity and genetic resources. Belhaven Press, London, UK.

Hawkes, J.G., 1997. A database for wild and cultivated potatoes. Euphytica 93: 155-161.

Hawkes, J.G., and J.P. Hjerting, 1969. The potatoes of Argentina, Brazil, Paraguay and Uruguay: a biosystematic study. Clarendon Press, Oxford, UK.

Hawkes, J.G., and J.P. Hjerting, 1989. The potatoes of Bolivia, their breeding value and evolutionary relationships. Clarendon Press, Oxford, UK.

Hayek, L.A.C., and M.A. Buzas, 1997. Surveying natural populations. Columbia University Press, New York, USA.

Haverkort, A.J., 1990. Ecology of potato cropping systems in relation to latitude and altitude. Agricultural Systems 32: 251-272.

Hazekamp, Th. and L. Frese, 1992. Application of mapping systems for the analysis of the geographical origin of collected material. In: L. Frese (ed.) International *Beta* Genetic Resources Network. A report on the 2nd International *Beta* Genetic Resources Workshop held at the Institute for Crop Science and Plant Breeding, Braunschweig, Germany, 24-28 June 1991. International Crop Network Series 7. International Board for Plant Genetic Resources, Rome, Italy. p. 69-70.

Hazekamp, Th., 2002. The potential role of passport data in the conservation and use of plant genetic resources. In: J.M.M. Engels, V. Ramantha Rao, A.H.D. Brown, and M.T. Jackson (eds.). Managing plant genetic diversity. CABI publishing, Wallingford, UK. p. 185-194.

He, W., 1997. Agronomic and ecological studies on the potato (Solanum tuberosum L.) in southwest China: Seed and crop management. PhD Thesis. Wageningen Agricultural University, Netherlands.

Hermann, M., 1988. Progress report of the IBPGR research project on the genetic resources of Andean tuber crops, May-Sept 1988. AGP-IBPGR report 89/2. International Board for Plant Genetic Resources, Rome, Italy.

Heun, M., R. Schaefer-Pregl, D. Klawan, R. Castagna, M. Accerbi, B. Borghi and F. Salamini, 1997. Site of einkorn wahet domestication identified by DNA fingerprinting. Science 278:1312-1314.

Heuvelink, G.B.M., 1998. Error propagation in environmental modelling with GIS. Taylor and Francis. London, UK.

Hijmans, R.J., 1997. Diversity of wild potato species. In: Denisov, N., C. Heberlein, L. Czaran and O. Simonnet, GIS in Agricultural Research: Awareness Package. UNEP/DEI/TR.97-9, Case-study 5.

Hijmans, R.J., 1999. Estimating frost risk in potato production on the Altiplano using interpolated climate data. In: International Potato Center, Impact on a changing world, Program Report 1997-1998. International Potato Center, Lima, pp. 373-380.

Hijmans, R.J., 2001. Global distribution of the potato crop. American Journal of Potato Research 78: 403-412.

Hijmans, R.J., and D.M. Spooner, 2001. Geographic distribution of wild potato species. American Journal of Botany 88:2101-2112.

Hijmans, R.J., M. Schreuder, J. de la Cruz, and L. Guarino, 1999. Using GIS to check coordinates of genebank accessions. Genetic resources and crop evolution 46:291-296.

Hijmans R.J., G.A. Forbes, and T.S. Walker, 2000a. Estimating the global severity of potato late blight with a GIS-linked disease forecaster. Plant Pathology 49: 697-705.

Hijmans, R.J., K.A. Garrett, Z. Huamán, D.P. Zhang, M. Schreuder, and M. Bonierbale, 2000b. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. Conservation Biology 14:1755-1765.

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

Hijmans, R.J., M. Cruz, E. Rojas, and L. Guarino, 2001b. DIVA, version 1.4. A geographic information system for the management and analysis of genetic resources data. Manual. International Potato Center and International Plant Genetic Resources Institute. Lima, Peru.

Hijmans, R.J., L. Guarino, C. Bussink, and E. Rojas, 2002a. DIVA-GIS, version 2. A geographic information system for the analysis of biodiversity data. Manual. International Potato Center, Lima, Peru.

Hijmans, R.J., D.M. Spooner, A. Salas, L.Guarino, and J. de la Cruz, 2002b. Atlas of wild potatoes. Systematic and ecogeographic studies of crop genepools. International Plant Genetic Resources Institute, Rome, Italy.

Hodgkin, T., and V. Ramanatha Rao, 2002. People, plants and DNA: Perspectives on the scientific and technical aspects of conserving and using plant genetic resources. In: Engels, J.M.M., V. Ramanatha Rao, A.H.D. Brown, and M.T. Jackson (eds.). Managing plant genetic diversity. CABI publishing, Wallingford, UK. p. 469-480.

Hodson, D.P., and J.W. White, 2002. GIS Tools: They're not just for experts anymore. Paper presented at the international conference: Impacts of international research and development: why has impact assessment research not made more of a difference? February 4-7, 2002, San José, Costa Rica.

Hodson, D.P., E. Martinez-Romero, J.W. White, J.D. Corbett, and M. Bänziger. 2002. Africa Maize Research Atlas (v3.0). Mexico, D.F.: CIMMYT CDROM publication.

Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, D. Xiaosu, , K. Maskell, and C.A. Jonhnson, C.A. (eds.), 2001. Climate change 2001. The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK.

Huaccho L., and R.J. Hijmans, 1999. A global geo-referenced database of potato production for 1995-1997, GPOT97. Production Systems and Natural Resource Management Department Working Paper 1. International Potato Center, Lima, Peru.

Huaccho L., and R.J. Hijmans, 2000. A geo-referenced database of global sweetpotato production. Production Systems and Natural Resource Management Working Paper 4. International Potato Center, Lima, Peru.

Huamán, Z., R. Hoekstra and J.B. Bamberg, 1996. The Intergenebank Potato Database. In: Abstracts of Conference Papers, Posters and Demonstrations of the 13th Triennial Conference of the EAPR, Veldhoven, The Netherlands, 14-19 July 1996, p. 315.

Huamán Z., R. Hoekstra, and J.B. Bamberg, 2000. The Inter-genebank potato database and the dimensions of wild potato germplasm. American Journal of Potato Research 77: 353–362

Huanco, S., 1992. Potencial de las papas amargas en el Altiplano de Puno. In: Rea J., Vacher, J.J. La papa amarga. Primera mesa redonda: Perú - Bolivia, La Paz 7-8 Mayo 1991. ORSTOM, La Paz.pp. 25-26.

Huke, R.E., 1982. Rice area by type of culture: south, southeast, and east Asia. International Rice Research Institute, Los Baños, Phillipines.

Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems 9(4), 385-403.

Hutchinson, M.F., 1997. ANUSPLIN version 3.2. User Guide. Australian National University. Centre for Resource and Environmental Studies. Canberra.

Hyman, G., 1999. Crop distribution mapping: applications and techniques for broad-scale analysis of crop geography. In: Pande, S., C. Johansen, J. Lauren, and F.T. Bantilan Jr. 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.

INEI, 1993. Tercer Censo Nacional Agropecuario. Database. INEI, Lima.

INTECSA, 1993. Estudio de Climatologia. Plan director global binacional de proteccion – prevencion de inundaciones y aprovechamiento de los recursos del lago Titicaca, rio Desaguadero, lago Poopó y lago Salar de Coipasa (Sistema T.D.P.S). INTECSA, AIC, CNR.

Intergovernmental Panel on Climate Change Data Distribution Center, 1999. Providing climate change and related scenarios for impact assessment. CD-ROM. Version 1.0. Climate Research Unit, University of East Anglia, Norwich, UK.

International Potato Center, 1991. Meeting the Challenge: The International Potato Center's strategy for the 1990s and beyond. International Potato Center, Lima, Peru.

Jansky, S.H., 2000. Breeding for disease resistance in potato. Plant Breeding Reviews 19: 69-155.

Jeffree, C.E., and E.P. Jeffree, 1996. Redistribution of the potential geographical ranges of mistletoe and Colorado beetle in Europe in response to the temperature component of climate change. Functional Ecology 10(5): 562-577.

Jones, P.G., 1991. The CIAT Climate Database Version 3.41. Machine readable dataset. CIAT, Cali.

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

Kalma, J.D., Laughlin, G.P., Caprio, J.P., Hamer, P.J.C., 1992. The bioclimatology of frost. Advances in Bioclimatology 2. Springer-Verlag, Berlin.

Kaukoranta, T., 1996. Impact of global warming on potato late blight: Risk, yield loss and control. Agricultural and Food Science in Finland 5: 311-327.

Kardolus, J.P., 1998. A biosystematic study of *Solanum acaule*. Thesis, Wageningen Agricultural University, The Netherlands.

Khanna, M.L., 1966. Breeding potato varieties tolerant to higher thermoperiods. Current Science 35(6): 143-144.

Klages, K.H.W., 1949. Ecological crop geography. MacMillan, New York, USA.

Kleinpenning, J.M.G., 1968. Geografie van de landbouw. Aula 373. Het Spectrum, Utrecht, Netherlands.

Knapp, G., 1988. The effects of variations in mean temperature and frost risk. In: M.L. Parry, T.R. Carter, and N.T. Konijn. The impact of climatic variations on agriculture. Volume 2: assessments in semi-arid regions. Part IV, The impact of climatic variations on agriculture in the Sierra of Ecuador. Kluwer, Dordrecht, Netherlands. p. 443-459.

Kooman, P.L., 1995. Yielding ability of potato crops as influenced by temperature and daylenght. PhD thesis. Wageningen Agricultural University. Wageningen, Netherlands.

Kooman, P.L., and A.J. Haverkort. 1995. Modelling development and growth of the potato crop influenced by temperature and daylength: LINTUL-POTATO. In: A.J. Haverkort and D.K.L. MacKerron (eds.). Potato ecology and modeling of crops under conditions limiting growth. Kluwer Academic Publishers, Dordrecht, Netherlands. p. 41-60.

Kostrowicki, J., 1974. The typology of world agriculture. Principles, methods and model types. International Geographical Union, Commission on Agricultural Typology, Warsaw, Poland.

Krause R.A., L.B. Masie, and R.A. Hyre, 1975. Blitecast: a computerized forecast of potato late blight. Plant Disease Reporter 59, 95-98.

Kristjanson, P. and P. Thornton, 2002. Herding cats: is impact assessment the ultimate exercise in futility? Paper presented at the international conference: Impacts of international research and development: why has impact assessment research not made more of a difference? February 4-7, 2002, San José, Costa Rica.

Le Tacon, Ph., 1989. Manifestation des risques climatiques a l'échelle de l'exploitation agricole, consequenes sur les pratiques paysannes. Cas de l'Altiplano Bolivien. Memoire de stage. ORSTOM, ENSSAA, CNEARC, ENSAM, Montpellier.

Leemans, R. and A.M. Solomon, 1993. Modeling the potential change in yield and distribution of the Earth's crops under a warmed climate, Climate Research 3: 79-96.

Levy, D., 1984. Cultivated *Solanum tuberosum* L. as a source for the selection of cultivars adapted to hot climates. Tropical Agriculture 61(3): 167-170.

Li, P.H., 1977. Frost killing temperatures of 60 tuber-bearing *Solanum* species. American Potato Journal 54:452-456.

Li, P.H., J.P. Palta and J.G. Hawkes, 1980. A scientific note: interrelationship between frost hardiness and elevation of genotype origin. American Potato Journal 57:184-185.

Luo, Y., P.S. Teng, N.G. Fabellar, and D.O. Tebeest, 1998. The effects of global temperature change on rice leaf blast epidemics: A simulation study in three agroecological zones. Agriculture Ecosystems and Environment 68:187-196.

Magurran, A.E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey.

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

McAllister, D.E., F.W. Schueler, C. M. Roberts and J.P. Hawkins, 1994. Mapping and GIS analysis of the global distribution of coral reef fishes on an equal area-grid. In: R.I. Miller (ed.). Mapping the diversity of nature. Chapman and Hall, London, UK. p. 155-175.

McCarthy, J.J., O.F. Canziani, N.A. Leary, D.J. Dokken, and K.S. White, 2001. Climate change 2001. Impacts, adaptation, and vulnerability. Contribution of Working Group II to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK.

Miglietta, F., V. Magliulo, M. Bindi, L. Cerio, F.P. Vaccari, V. Loduca, and A. Peressotti, 1998. Free Air CO_2 Enrichment of potato (*Solanum tuberosum* L.): development, growth and yield. Global Change Biology 4: 163-172.

Miller, J.T., and D.M. Spooner, 1999. Collapse of species boundaries in the *Solanum brevicaule* complex (Solanaceae, Sect. *Petota*): molecular data. Plant Systematics and Evolution 214: 103-130.

Monmonier, M., 1999. How to lie with maps. The University of Chicago Press. Chicago, USA. 2nd ed.

Morlon, P., 1989. Du climat à la commercialisation: l'example de l'Altiplano péruvien. In: M. Eldin and P. Milleville. Le risque en agriculture. ORSTOM, Paris, France. p.187-224.

Nabhan, G.P., 1990. Wild Phaseolus ecogeography in the Sierra Madre Occidental, Mexico: aerographic techniques for targeting and conserving species diversity. Systematic and ecogeographic studies on crop genepools 5. International Board for Plant Genetic Resources, Rome, Italy.

Nelson, M.R., T.V. Orum, R. Jaime-Garcia, and A. Nadeem, 1999. Applications of geographic information systems and geostatistics in plant disease epidemiology and management. Plant Disease 83:308-319.

Nelson, G.C., 2002. What GIS can (and can't) bring to impact assessment: novel data, analysis, and insights. Paper presented at the international conference: Impacts of international research and development: why has impact assessment research not made more of a difference? February 4-7, 2002, San José, Costa Rica. In press.

New, M., M. Hulme, and P. Jones. 1999. Representing twentieth-century space-time climate variability. Part I: development of a 1961-1990 mean monthly terrestrial climatology. Journal of Climate 12: 829-856

Nevo, E., 1998. Molecular evolution and ecological stress at global, regional and local scales: the Israeli perspective. The Journal of Experimental Zoology 282:95-119.

Nevo, E., E. Golenberg, A. Beiles, A.H.D. Brown, and D. Zohary, 1982. Genetic diversity and environmental associations of wild wheat, Triticum dicoccoides, in Israel. Theoretical and applied Genetics 62:241-254.

New, M., M. Hulme, and P. Jones, 1999. Representing twentieth-century space-time climate variability. Part I: development of a 1961-1990 mean monthly terrestrial climatology. Journal of Climate 12:829-856.

NOAA and US-EPA, 1997. Hourly United States Weather Observations 1990-1995. Asheville, NC, USA: National Oceanic and Atmospheric Administration; Raleigh, NC, USA: U.S. Environmental Protection Agency. (CD-ROM).

Nonhebel, S., 1993. The importance of weather data in crop growth simulation models and assessment of climatic change effects. Ph.D. thesis. Wageningen Agricultural University, Wageningen, Netherlands.

Nonhebel, S., 1994. The effects of use of average instead of daily weather data in crop growth simulation models. Agricultural Systems 44:377-396.

Ochoa, C.M., 1962. Los Los *Solanum* tuberíferos silvestres del Perú (*Tuberarium* subsecc. *Hyperbasarthrum*). Privately published.

Ochoa, C.M., 1990. The potatoes of South America: Bolivia. Cambridge University Press, Cambridge, United Kingdom.

Ochoa, C.M. 1999. Las papas de Sudamérica: Peru (Parte I). International Potato Center, Lima, Peru.

OEA, 1996. Diagnostico ambiental del sistema Titicaca-Desaguadero-Poopo-Salar de Coipasa (Sistema TDPS), Bolivia–Peru. Organización de los Estados Americanos, Washington DC, USA.

Øllgaard, B., 1995. Diversity of Huperzia (Lycopodiaceae) in neotropical montane forests. Pages 349-358 in S.P. Churchill, H. Balslev, E. Forero, and J. L. Luteyn, editors. Biodiversity and Conservation of Neotropical Montane Forests. The New York Botanical Garden. New York.

Peeters, J.P., H.G. Wilkes and N.W. Galwey. 1990. The use of ecogeographical data in the explotation of variation from genebanks. Theoritcal and Applied Genetics 80:110-112.

Peiris, D.R., J.W. Crawford, C. Grashoff, R.A. Jefferies, J.R. Porter, and B. Marshall. 1996. A simulation study of crop growth and development under climate change. Agricultural and Forest Meteorology 79:271-287.

Prendergast, J.R., R.M. Quinn, J.H. Lawton, B.C. Eversham, and D.W. Gibbons, 1993a. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature 365:335-337.

Prendergast, J.R., S.N. Wood, J.H. Lawton, and B.C. Eversham, 1993b. Correcting for variation in recording effort in analyses of diversity hotspots. Biodiversity Letters 1: 39-53.

Rabinowitz, D., 1981. Seven forms of rarity. In: H. Syne (ed.). The biological aspects of rare plant conservation, 205-217. John Wiley and Sons, New York, New York, USA.

Rabinowitz, D., C.R. Linder, R. Ortega, D. Begazo, H. Murguia, D.S. Douches, and C.F. Quiros, 1990. High levels of interspecific hybridization between *Solanum sparsipilum* and *S. stenotomum* in experimental plots in the Andes. American Potato Journal 67:73-81.

Rea, J., 1992. Vigencia de las papas nativas en Bolivia. In: J. Rea and J.J. Vacher, La papa amarga. Primera mesa redonda: Perú - Bolivia, La Paz 7-8 Mayo 1991. ORSTOM, La Paz. p.15-23.

Rebelo, A.G., 1994. Iterative selection procedures: centres of endemism and optimal placement of reserves. Strelitzia 1: 231-257

Rebelo, A.G. and W.R. Siegfried, 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.

Rediscovering Geography Committee, Board on earth sciences and resources, Commission on geosciences, environment, and resources, National research council, 1997. Rediscovering geography: new relevance for science and society. National Academy Press, Washington DC, USA.

Reynolds, M.P., and E.E. Ewing. 1989. Heat tolerance in tuber bearing *Solanum* species: A protocol for screening. American Potato Journal, 66(2): 63-74.

Rhoades, R.E. 1987. Potato production zones of developing countries, computerized maps (unpublished draft). International Potato Center.

Rich, T.C.G., and E.R. Woodruff, 1992. Recording bias in botanical surveys. Watsonia 19: 73-75.

Rick, C.M., 1973. Potential genetic resources in tomato species: clues from observations in native habitats. In: A.M. Srb (ed.). Genes, enzymes, and populations. Plenum, New York, USA. p. 255-269.

Robinson, A.H., and B.B. Petchenik, 1976. The nature of maps: essays toward understanding maps and mapping. University of Chicago Press, Chicago, IL, USA.

Ronning, C.M., S.P. Kowalsky, L.L. Sanford, and J.R. Stommel, 2000. Geographic variation of solanidane aglycone glycoalkaloids in the wild potato species *Solanum chacoense* Bitter. Genetic Resources and Crop Evolution 47:359-369.

Rosenzweig, C., and D. Hillel. 1998. Climate change and the global harvest. Potential impacts of the greenhouse effect on agriculture. Oxford University Press, New York, USA.

Rosenzweig, C., and D. Liverman, 1992. Predicted effects of climate change on agriculture: A comparison of temperate and tropical regions. In: S.K. Majumdar (ed.). Global climate change: Implications, challenges, and mitigation measures. The Pennsylvania Academy of Sciences, Philadelphia, USA. p. 342-61.

Rosenzweig, C., and M.L. Parry, 1994. Potential impact of climate change on world food supply. Nature 367:133-138.

Rosenzweig, C., J. Phillips, R. Goldberg, J. Carroll, and T. Hodges, 1996. Potential impacts of climate change on citrus and potato production in the US. Agricultural Systems 52(4): 455-479.

Ross, H., 1986. Potato breeding: Problems and perspectives. Advances in Plant Breeding 13. Paul Parey Verlag, Berlin, Germany.

Salas, A.R., D.M. Spooner, Z. Huamán, R.V. Torres-Maita, R. Hoekstra, K. Schüler, and R.J. Hijmans, 2001. Taxonomy and new collections of wild potato species in central and southern Peru in 1999. American Journal of Potato Research 78(3): 197-207.

Sánchez de Lozada, D., P. Baveye, and S. Riha, 1998. Heat and moisture dynamics in raised field systems of the lake Titicaca region (Bolivia). Agricultural and Forest Meteorology 92(4): 251-265.

Sanders, H.L., 1968. Marine benthic diversity: a comparative study. American Naturalist 102: 243-282.

Sauer, J.D., 1993. Historical geography of crop plants - a select roster. CRC Press, Boca Raton, FL, USA.

Schepers, H.T.A.M., E. Bouma, and C.B Bus, 1997. State of the art of *Phytophthora infestans* control in Europe. In: H.T.A.M. Schepers and E. Bouma (eds.). Proceedings of the workshop on the European network for development of an integrated control strategy of potato late blight. PAV-Report 1. Applied Research for Arable Farming and Field Production of Vegetables, Lelystad, Netherlands. p. 7-11.

Scott, G.J., M.W. Rosegrant, and C. Ringler, 2000. Global projections for root and tuber crops to the year 2020. Food Policy 25(5): 561-597.

Seem, R.C., 1995. Geographic information systems for localized pest predictions. Bulletin OEPP 23: 639-646.

Spooner, D.M., and R.T. Castillo, 1997. Reexamination of series relationships of South American wild potatoes (Solanaceae: *Solanum* sect. *Petota*): evidence from chloroplast DNA restriction site variation. American Journal of Botany 84:671-685.

Spooner, D.M., and R.J. Hijmans, 2001. Potato systematics and germplasm collecting 1989-2000. American Journal of Potato Research 78: 237-268; 395.

Spooner, D.M., and K.J. Sytsma, 1992. Reexamination of series relationships of Mexican and Central American wild potatoes (*Solanum* sect. *Petota*): evidence from chloroplast DNA restriction site variation. Systematic Botany 17:432-448.

Spooner, D.M., and R.G. van den Berg, 1992. An analysis of recent taxonomic concepts in wild potatoes (*Solanum* sect. *Petota*). Genetic Resources and Crop Evolution 39:23-37.

Spooner, D.M., R. Castillo, and L. López, 1992. Ecuador, 1991 potato germplasm collecting expedition: taxonomy and new germplasm resources. Euphytica 60: 159-169.

Spooner, D.M., R.G. van den Berg, W. Garcia, and M.L. Ugarte, 1994. Bolivia potato germplasm collecting expeditions 1993, 1994: Taxonomy and new germplasm resources. Euphytica 79:137-148.

Spooner D.M., R.G. Van den Berg, and J. Bamberg, 1995. Examination of species boundaries of *Solanum* series *Demissa* and potentially related species in series *Acaulia* and series *Tuberosa* (sect. *Petota*). Systematic Botany 20:295-314.

Spooner, D.M., R. Hoekstra, R.G. van den Berg, and V. Martínez, 1998. *Solanum* sect. *Petota* in Guatemala: taxonomy and genetic resources. American Journal of Potato Research 75: 3-17.

Spooner, D.M., A. Salas, Z. Huamán, and R.J. Hijmans, 1999. Potato germplasm collecting expedition in southern Peru (Departments of Apurímac, Arequipa, Cusco, Moquegua, Puno, Tacna) in 1998: Taxonomy and new genetic resources. American Journal of Potato Research 76:103-119.

Spooner, D.M., R. Hoekstra, and B. Vilchez, 2001a. *Solanum* section *Petota* in Costa Rica: taxonomy and genetic resources. American Journal of Potato Research 78: 91-98.

Spooner, D.M., R.G. van den Berg, A. Rivera-Peña, P. Velguth, A. del Rio, and A. Salas, 2001b. Taxonomy of Mexican and Central American members of *Solanum* Series *Conicibaccata* (sect. *Petota*). Systematic Botany 26: 743-756.

Stol, W., G.H.J. de Koning, P.L. Kooman, A.J. Haverkort, H. van Keulen, and F.W.T. Penning de Vries, 1991. Agro-ecological characterization for potato production. A simulation study at the request of the International Potato Center (CIP), Lima, Peru. CABO-DLO, Report 155. CABO-DLO, Wageningen, Netherlands.

Spitters, C.J.T, 1987. An analysis of variation in yield among potato cultivars in terms of light absorption, light utilization, and dry matter partitioning. Acta Horticulturae 267: 71-84.

Sukumaran, N.P., and C.J. Weiser, 1972. An excised leaflet test for evaluating potato frost tolerance. Hortscience 7(5): 467-468.

Supit I, A.A. Hooijer, and C. van Diepen, 1994. System description of the WOFOST 6.0 crop simulation model implemented in CGMS. Vol. I: Theory and algorithims. Joint Research Center, Commission of the European Communities, Luxembourg, Luxembourg.

Tai, G.C.C., D. Levy, W.K. Coleman, 1994. Path analysis of genotype-environment interactons of potatoes exposed to increasing warm-climate constraints. Euphytica 75(1-2): 49-61.

Tapia, N., and G. Saravia, 1997. Biodiversidad en papas amargas. Prov. Tapacarí, Dpto. de Cochabamba. AGRUCO, Cochabamba.

Teng, P.S., 1983. Estimating and interpreting disease intensity and loss in commercial fields. Phytopathology 73: 1587-1590.

Terrezas, F., V. Suarez, G. Gardner, G. Thiele, A. Devaux, and T. Walker, 1998. Diagnosing potato productivity in farmers' fields in Bolivia. Social Science Department Working Paper 1998-5. International Potato Center, Lima, Peru.

Travasso, M.I., D.O. Caldiz, and J.A. Saluzzo, 1996. Yield prediction using the SUBSTOR-POTATO model under Argentinean conditions. Potato Research 39:305-312.

Times Books, 1988. The Times' Atlas of the World - Comprehensive ed. Times Books, London.

Tsuji, G.Y., G. Hoogenboom, and P.K. Thornton (Eds.), 1998. Understanding options for agricultural production. System approaches for sustainable agricultural development 7. Kluwer Academic Publishers, Dordrecht, the Netherlands. 325 pp.

United States Geological Survey. 1998. Database available at: http://edcwww.cr.usgs.gov/landdaac/gtopo30/gtopo30.html.

USGS-EDC (United States Geological Survey. EROS Data Center), without date. 30-seconds (GTOPO30) and 5-minute (ETOPO5) global altitude databases. Online at: http://edcdaac.usgs.gov/.

USGS-EDC (United States Geological Surveys EROS Data Center), 1998. 1-km global land cover characterization database. Online at: http://edcdaac.usgs.gov/glcc/glcc.html.

Valdivia -A., R., R. Quiroz, R. Valdivia -F., and V. Choque-Huanca, 1997. Risk analysis of potato production in the Altiplano: quantifying farmers' beliefs. CIP Program Report 1995-96. International Potato Center, Lima. p. 29-33.

Van Den Berg, R.G., J.T. Miller, M.L. Ugarte, J.P. Kardolus, J. Villand, J. Nienhuis, and D.M. Spooner, 1998. Collapse of morphological species in the wild potato *Solanum brevicaule* complex (Solanaceae: sect. *Petota*). American Journal of Botany 85: 92-109.

Van der Zaag, P., and A.L. Demagante, 1988. Potato (*Solanum* spp.) in an isohyperthermic environment 3: Evaluation of clones. Field Crops Research 19: 167-181.

Van Keulen H., and W. Stol, 1995. Agro-ecological zonation for potato production. In: A.J. Haverkort and D.K.L. MacKerron (eds). Potato ecology and modeling of crops under conditions limiting growth. Kluwer Academic Publishers, Dordrecht, the Netherlands. p. 357-379.

Van Royen, W., 1954. The agricultural resources of the world. Atlas of the world's resources. Volume 1. Prentice Hall, New York, USA.

Van Soest, L.J.M., H.J. Rumpenhorst, and C.A. Huijsman, 1983. Resistance to potato cyst-nematodes in tuber-bearing *Solanum* species and its geographical distribution. Euphytica 32:65-74.

Van Soest, L.J.M., B. Schober, and M.F. Tazelaar, 1984. Resistance to *Phytophthora infestans* in tuber-bearing species of *Solanum* and its geographical distribution. Potato Research 27:393-411.

Vavilov, N.I., 1926. Studies on the origin of cultivated plants. Bulletin of Applied Botany and Plant Breeding 16: 1-294.

Vega, S.E., and J.B. Bamberg. 1995. Screening the U.S. potato collection for frost hardiness. American Potato Journal 72:13-21.

Von Bothmer, R., and O. Seberg. 1995. Strategies for the collecting of wild species. In: L. Guarino, V. Ramanatha Rao and R. Reid (eds.). Collecting plant genetic diversity. Technical guidelines. CABI, Wallingford, UK. p. 93-111.

Von Thünen, J.H., 1842. Der isolierte Staat in Beziehung auf Landwirtschaft und Nationaloekonomie - 2 Vols. Reprint, 1921. Fischer, Jena, Germany.

Walker T.S., and M. Collion, 1997. Priority setting at CIP for the 1998-2000 medium-term plan. Lima, Peru: International Potato Center.

Walker, T.S., P.E. Schmiediche, and R.J. Hijmans. 1999. World trends and patterns in the potato crop: An economic and geographic survey. Potato Research 42: 241-264.

Watanabe, K., and S.J. Peloquin, 1989. Occurrence of 2n pollen and *ps* gene frequencies in cultivated groups and their related wild species in tuber-bearing solanums. Theoretical and Applied Genetics 78: 329-336.

Wiegers, E.S., R.J. Hijmans, D. Hervé, and L.O. Fresco, 1999. Land use intensification and desintensification in the Upper Cañete valley, Peru. Human Ecology 27(2):319-339.

Wilsie, C.P., 1962. Crop adaptation and distribution. W.H. Friedman and Company, San Francisco, CA, USA.

Weltzien, 1978. Geophytopathology. In: J.G. Horsfall and E.B. Cowling. Plant disease – an advanced treatise. Academic Press, New York, USA. p.339-358.

White, J.W., J.D. Corbett, and A. Dobermann. 2002. Insufficient geographic characterization and analysis in the planning, execution and dissemination of agronomic research? Field Crops Research (in press).

Wieczorek, J., 2001. MaNIS project. http://dlp.cs.berkeley.edu/manis/

Wood, S., and P.G. Pardey, 1997. Agroecological aspects of evaluating agricultural R&D. ETPD Discussion Paper 23. IFPRI, Washington.

Wood, S. and P.G. Pardey, 1998. Agroecological aspects of evaluating agricultural R&D. Agricultural Systems 57: 13-42.

Wood, S., K. Sebastian, and S.J. Scherr, 2000. Pilot analysis of global ecosystems: agroecosystems. A joint study by International Food Policy Research Institute and World Resources Institute. International Food Policy Research Institute and World Resources Institute, Washington, DC, USA.

Wortmann, C.S., R.A. Kirkby, C.A. Eledu, and D.J. Allen, 1998. Atlas of common bean (*Phaseolus vulgaris* L.) production in Africa. Centro Internacional de Agricultura Tropical, Cali, Colombia.

Zeven, A.C., and P.M. Zhukovsky, 1975. Dictionary of cultivated plants and their centres of diversity. Excluding ornamentals, forest trees and lower plants. Centre for Agricultural Publishing and Documentation (PUDOC), Wageningen, The Netherlands.

Zadoks J.C. and F.H. Rijsdijk, 1984. Agro-ecological atlas of cereal growing in Europe: Vol. 3. Atlas of cereal diseases and pests in Europe. Backhuys, Leiden, Netherlands.

Zhang, D., G. Rossel, A. Kriegner, and R.J. Hijmans, 2002. AFLP assessment of diversity in sweetpotato from Latin America and the Pacific region: Its implications on the dispersal of the crop. Genetic Resources and Crop Evolution. (In press).

ZONISIG, 1998. Zonificación agroecológica y socio-económica de la Cuenca del Altiplano del Departamento de La Paz. Ministerio de Desarrollo Sostenible y Planificación, Prefectura del Departamento de La Paz, DHV-ITC, La Paz, Bolivia.

Summary

This thesis aims at contributing to our knowledge of the potato crop and its wild relatives, and to methodological progress in the use of geographic information in crop science. Spatial variation is a fundamental characteristic of agriculture, but crop scientists have largely ignored it, particularly at levels beyond the field scale. This thesis provides examples of how to deal explicitly with spatial variation over larger areas. New data, methods and tools have been used to address classic questions of crop geography, such as “where is this crop grown?”, “where is it most genetically diverse?” and “where is a production constraint important?”. The availability of GIS allows tackling these old questions in new ways and spatial analysis can increasingly be used by crop scientists to provide the information they need.

In Part I (Chapters 2–5), the use of GIS for the management of genetic resources is explored. Part II (Chapters 6–9) deals with the use of agro-ecological zoning in the context of a crop improvement and management research program. In the General Introduction (Chapter 1) the contribution of geographers and crop scientists to spatial analysis in agriculture is briefly reviewed. This is followed by an introduction of the two main themes of this thesis: management of plant genetic resources, and agro-ecological zoning. Finally, a brief outline of this thesis is given. In the final chapter (General Discussion, Chapter 10), the contribution to the two research themes made in this thesis is discussed in relation to further research needs. This is followed by a discussion of challenges to the further use of GIS in crop science, particularly in the context of international agricultural research.

Part I

Genebank databases are the principal source of information on the spatial distribution of crop genetic diversity. However, coordinate data in genebank databases are notorious for their scarcity, and, for those records that do have coordinate data, for occasionally being inaccurate. This makes spatial analysis of genebank data more complicated, and the results more uncertain. In Chapter 2, a protocol is described to check and improve the data quality of the geographical coordinates in genebank databases. Applying these procedures increases the capability for analysis of crop genetic diversity data.

As genebanks were not primarily contemplated as a source of geo-referenced information about a crop, but rather at maximizing diversity, the analysis of these data can be problematic, and issues of information quality need to be considered. Due to

spatial differences in recorder effort, data on plant distributions are typically biased. In extreme cases, differences in the number of species between areas would reflect the amount of time spent there by recorders, and not actual differences in distribution. In Chapter 3, methods are described and used to assess the information quality of genebank data on wild potatoes in Bolivia.

Four types of biases were defined: species, species-area, hotspot, and infrastructure bias. Species bias is the sampling of some species more often than others. Species-area bias is a sampling that is disproportionate to the total area in which a species is found. Hotspot bias is the disproportionate sampling of areas with high levels of diversity. Infrastructure bias is the disproportionate sampling of areas near roads and towns. Each of these biases is present in the Bolivian wild potato collection. The infrastructure bias is especially strong: 60% of all wild potato accessions were collected within 2 km of a road, as opposed to 22%, if collections had been made randomly. This analysis can serve as a guide for future collecting trips. It can also provide baseline information for the application of genebank data in GIS-based studies.

Chapter 4 provides the first comprehensive and quantitative description of the distribution of the 199 currently accepted wild potato species (Solanaceae sect. Petota). Understanding the distribution of genetic diversity in this group of plants can be useful for future collecting and conservation efforts. Wild potatoes occur in 16 countries, between 38°N and 41°S, with more species in the southern hemisphere, and are most common in Argentina, Bolivia, Mexico, and Peru. Peru has the highest number of species (93), followed by Bolivia (39). Areas with high species richness occur in northern Argentina, central Bolivia, central Ecuador, central Mexico, and south and north-central Peru. Species richness is highest between 8°and 20°S and around 20°N. Wild potatoes typically occur between 2000 and 4000 m altitude. Most species are rare and narrowly endemic: for 77 species the largest distance between two observations is less then 100 km.

Conservation and collection of wild potato species has been done mostly on a taxonomic basis. From a crop improvement perspective, however, the presence of useful agronomic traits, such as disease resistance is more important then taxonomic diversity per se. The use of taxonomic diversity could be justified if it were a proxy for trait diversity, because there are many useful traits and they often can not be observed when collecting. An alternative approach might entail the use of geographical and ecological information to stratify and sample wild potato populations. The relative merits of these different factors for predictive purposes are put to the test in Chapter 5, with a case study on frost tolerance in wild potatoes.

Frost tolerance screening data was used for 1646 wild potato collections representing 87 species that had been collected in 12 countries in the Americas. There is a strong association of frost tolerance with species and to a lesser extent with taxonomic series. There is also a significant geographic clustering of wild potatoes with similar levels of average frost tolerance. Areas with a high level of frost tolerance are the central and southern Peruvian Andes, the lowlands of Argentina and adjacent areas, and a small area in the central Chilean Andes. There is a greater chance of finding wild potatoes with high levels of frost tolerance in areas with a yearly average minimum temperature below 3°C than there is in warmer areas. However, temperature is only a weak predictor of frost tolerance. Temperature data alone did not predict observed frost tolerance in eastern Argentina/Uruguay and falsely predicted it in the southwestern United States. Because many wild potato species occur over small areas, taxonomic, ecological, and geographical factors are often confounded, and these factors should be analyzed simultaneously to interpret patterns in the distribution of traits.

Part II

The studies in the second part of this thesis contribute to the use of agro-ecological zonation (AEZ) for research planning. An important limitation to many AEZ studies is that the results are difficult to interpret and use because they are too generic and not related to the current area of distribution of the crop. This may have been caused by a lack of detailed crop distribution data. To fill this gap for potato, a global potato area distribution database was developed and discussed in relation to other variables (latitude, altitude, population density) in Chapter 6.

There are two main peaks in global potato distribution by latitude. The major peak is between 45 and 57°N and represents potato production zones in the temperate climates where potato is a summer crop. The other peak is between 23 and 34°N, mainly represents production zones in the subtropical lowlands, where potato is a winter crop. Between 1950 and 1998 potato production area increased at low latitudes and decreased at high latitudes, particularly around 53°N. The northern limit of potato production coincides with the boundaries of agriculture and the presence of human population. The peak between 23°N and 34°N coincides with the area of highest population density (per area of land and per area of arable land). About 25% of the global potato area is in the highlands (above 1000 m).

Potato late blight, caused by *Phytophthora infestans*, is one of the most devastating crop diseases—as was illustrated by the Irish Famine of the 1840s. As late blight severity is highly weather dependent, particularly to air humidity, agro-ecological zonation can be useful for understanding the spatial pattern in the distribution of this disease. In Chapter 7, global climate databases and disease forecast models are used to develop a quantitative and constraint-specific AEZ for potato late blight. The main purpose of the study is to assist in targeting technology (resistant varieties, improved fungicide use) for late blight control in developing countries.

Two disease forecast models, Blitecast and Simcast, to a climate database in a Geographic Information System (GIS). The disease forecast models indirectly estimate late blight severity by determining how many sprays were needed during a growing season, as a function of the weather. Global zones of estimated late blight severity were similar for both forecast models, but Blitecast generally predicted a lower number of sprays. Zones of high late blight severity include the tropical highlands, Western Europe, the east coast of Canada and the northern USA, southeast Brazil, and central-south China. Major production zones with a low late blight severity include the western plains in India, where irrigated potato is produced in the cool dry season, north-central China, and the northwestern USA. Using the global database of potato area, the average number of sprays was calculated by country and compared with estimates of current fungicide use. The results using Blitecast and Simcast were correlated but only Blitecast estimates correlated with observed data for developed countries. The estimated number of sprays, whether from Blitecast or Simcast, did not correlate with the observed number of sprays in developing countries and in a large number of developing countries the predicted optimal number of sprays was much higher then the actual number observed. In these countries, the increased access to host resistance and fungicides could have a strong economic impact.

A new method for estimating potential impact of proposed research is described in Chapter 8. The potential impact of the use of frost tolerant potato varieties on the Altiplano of Peru and Bolivia is quantified using a potato growth simulation model, and high-resolution interpolated climate data in a GIS. To avoid information loss, zonation in the strict sense is not used, as each small grid cell is treated individually, i.e., without aggregation prior to the calculations.

The LINTUL potato growth simulation model was adapted to incorporate the effect of frost damage on yield. Grids of monthly climate data were created for a number of variables, including absolute minimum temperature and its standard deviation, and used as input for the simulation model. The model was run using a standard potato

cultivar for which frost resistance parameters were changed in increments of 1°C. A geo-referenced database of potato distribution was used to process the output of the simulation model to calculate potato-area weighted results. When frost resistance increases from –1°C (current level) to –2°C or –3°C, average potato yield increases by 26 and 40%, respectively. After that, the effect tapers off and a further increase in resistance leads to only a small increase in simulated potato yield.

In Chapter 9, the effect of climate change on global potato production is quantified using current and predicted future climate data and a crop growth simulation model. Whereas most studies of this type study analyze a limited number of locations, this study is truly global. Particular emphasis is given to the potential to adapt to climate change by shifting the growing season or by using varieties with a different maturity.

Between 1961-90 and 2040-69 the global (terrestrial excluding Antarctica) average temperature is predicted to increase between 2.1 and 3.2°C, depending on the climate scenario. The temperature increase is smaller when changes are weighted by the potato area and particularly when adaptation of planting time and varieties is considered (a predicted temperature increase between 1 and 1.4°C). For this period, potato-area-weighted global potential potato yield decreases by 18 to 32% (without adaptation) and by 9 to 18% (with adaptation). At high latitudes, global warming may lead to earlier planting, the use of later maturing varieties, and a shift of the location of potato production; and changes in potato yield are likely to be small, and sometimes positive. Shifting planting time or location is less feasible at lower latitudes, and in these areas global warming could have a strong negative effect on potato production. To mitigate some of the effects of global warming, in (sub)tropical regions heat-tolerant potato varieties could be identified or developed.

Samenvatting

Het doel van dit proefschrift is bij te dragen aan de kennis van het aardappelgewas en de hieraan verwante wilde aardappels en bij te dragen aan methodologische vooruitgang in het gebruik van geografische informatie in de gewaswetenschappen. Ruimtelijke variatie is een belangrijk kenmerk van de landbouw. Nochtans hebben gewaswetenschappers hier maar weinig aandacht aan geschonken, vooral aan variatie over grotere gebieden dan velden. Dit proefschrift geeft voorbeelden van hoe ruimtelijke variatie over grotere gebieden kan worden bestudeerd. Nieuwe gegevens, methoden en computer programma's werden gebruikt om klassieke vragen te beantwoorden uit de gewasgeografie, zoals "Waar komt dit gewas voor?", "Waar heeft het de meeste genetische diversiteit?" en "Waar is een productie limiterende factor van belang?". De beschikbaarheid van Geografische Informatie Systemen (GIS) maakt het mogelijk om deze oude vragen op nieuwe manieren te beantwoorden. Hierdoor wordt de analyse van ruimtelijke variatie in toenemende mate mogelijk en nuttig voor gewaswetenschappers.

In Deel I (Hoofdstukken 2-5) wordt het gebruik van GIS verkend voor het beheer van genetische hulpbronnen van planten. Deel II (Hoofdstukken 6-9) gaat over het gebruik van agro-ecologische zonering in de context van een gewasonderzoekprogramma. In de Algemene Introductie (Hoofdstuk 1) wordt in het kort de bijdrage besproken van geografen en gewaswetenschappers aan ruimtelijke analyse in de landbouw. Na deze introductie worden de twee hoofdthema's van dit proefschrift ingeleid: het beheer van genetische hulpbronnen en agro-ecologische zonering. In het laatste hoofdstuk (Algemene Discussie, Hoofdstuk 10) wordt de bijdrage van dit proefschrift besproken, in relatie met de behoefte aan verder onderzoek. Dit wordt gevolgd door een bespreking van de uitdagingen en moeilijkheden bij een verdere toepassing van GIS in de gewaswetenschappen, met name in de context van internationaal landbouwkundig onderzoek.

Deel I

Databases van genenbanken zijn de voornaamste bron van gegevens over de ruimtelijke spreiding van genetische diversiteit van gewassen. Helaas hebben veel records geen geografische coördinaten, en, als ze aanwezig zijn, zijn ze soms onnauwkeurig. In Hoofdstuk 2 wordt een protocol beschreven om de kwaliteit van de geografische coördinaten in genenbanken te verifiëren en te verbeteren. Door toepassing van deze procedures neemt de mogelijkheid van verdere analyse van de gegevens sterk toe.

Genenbanken zijn bedoeld om zo veel mogelijk genetische diversiteit te bewaren, en niet in de eerste plaats als een bron voor geografische gegeven over de verspreiding van planten. Dit moet bij de analyse onderkend worden, en het kan zinvol zijn de "informatiekwaliteit" van deze gegevens te evalueren. Vanwege ruimtelijke verschillen in de intensiviteit van het waarnemingen doen, zijn gegevens over de distributie van planten meestal statistisch onzuiver. In extreme gevallen zouden verschillen tussen gebieden in het aantal soorten dat er voorkomt een afspiegeling kunnen zijn van de tijd die daar gespendeerd is door verzamelaars, en niet van de werkelijke verschillen in het aantal soorten.

In Hoofdstuk 3 worden methoden beschreven en toegepast om de informatie kwaliteit te bepalen van gegevens over de verspreiding van wilde aardappels (*Solanum* sect. *Petota*) in Bolivia. Vier soorten onzuiverheden werden gedefinieerd: soort, soort-gebied, hotspot, en infrastructuur. Soortonzuiverheid is het meer verzamelen van sommige soorten ten opzichte van andere. Soort-gebiedonzuiverheid is het onevenredig verzamelen ten opzichte van de grootte van het gebied waarin de soort voorkomt. Infrastructuuronzuiverheid is het onevenredig verzamelen in gebieden dichtbij wegen en steden. Al deze onzuiverheden waren aanwezig in de gegevens over Boliviaanse wilde aardappels. Vooral de infrastructuuronzuiverheid was groot: 60% van alle wilde aardappelen werden verzameld binnen 2 km afstand van een weg, terwijl de (random) verwachting 22% is. Dit soort analyse kan zinvol zijn bij de voorbereiding van toekomstige verzamelexpedities, en kan informatie verschaffen die van belang is voor de verdere ruimtelijke analyse van gegevens van genenbanken.

Inzicht in de distributie van planten is nuttig voor het plannen van verzamel- en bescherming activiteiten. Hoofdstuk 4 verschaft een kwantitatieve analyse van de distributie van alle 199 wilde aardappelsoorten. Wilde aardappelen komen voor in 16 landen, tussen 38°N en 41°Z. Er zijn meer soorten op het zuidelijk halfrond, en ze worden vooral gevonden in Argentinië, Bolivia, Mexico en Peru. Peru heeft de meeste soorten (93), gevolgd door Bolivia (39). Gebieden met een grote soortenrijkdom komen voor in Noord-Argentinië, Centraal-Bolivia, Centraal-Ecuador, Centraal-Mexico, en het zuiden en noorden van Centraal-Peru. De soortenrijkdom is het hoogst tussen 8° en 20°Z en rond 20°N. Wilde aardappels komen vooral voor tussen 2000 en 4000 m hoogte. De meeste soorten zijn zeldzaam en endemisch: voor 77 soorten is de grootste afstand tussen twee observaties minder dan 100 km.

Het verzamelen en behouden van wilde aardappelen vindt veelal plaats op basis van taxonomische criteria. Voor de plantenveredeling is echter de aanwezigheid van landbouwkundig bruikbare eigenschappen, zoals resistentie tegen ziektes, in principe belangrijker dan taxonomische verschillen. Nochtans zou het gebruik van taxonomische criteria gerechtvaardigd zijn als taxonomische diversiteit een goede indicator voor diversiteit in eigenschappen zou zijn. Er zijn immers veel bruikbare eigenschappen die vaak niet waargenomen kunnen worden tijdens het verzamelen. Een alternatieve benadering zou kunnen bestaan uit het gebruik van geografische en ecologische informatie om populaties van wilde aardappelen te groeperen en te verzamelen.

De voorspellende waarde van deze verschillende variabelen wordt getest in Hoofdstuk 5, met een studie over vorsttolerantie in wilde aardappelen. Gegevens werden gebruikt over 1646 verschillende populaties, waaronder 87 verschillende soorten, uit 12 landen in Noord- en Zuid-Amerika. Er bleek een sterke associatie te zijn van vorsttolerantie met soort, en in mindere mate met serie (een hoger taxonomisch niveau). Er bleek ook een significante geografische clustering te zijn van wilde aardappels met vergelijkbare niveaus van vorsttolerantie. Gebieden met hoge niveaus van vorsttolerantie bevinden zich in de centrale en zuidelijke Peruviaanse Andes, de laaglanden van Argentinië en aangrenzende gebieden, en een klein gebied in de Andes in Centraal-Chili. Er was een grotere kans op het vinden van wilde aardappels met hoge vorsttolerantie in gebieden met een jaarlijkse minimum temperatuur beneden de 3°C dan in warmere gebieden. Nochtans was temperatuur maar een matige voorspeller van vorsttolerantie. Met alleen temperatuurgegevens zou de geobserveerde vorsttolerantie in Oost-Argentinië en Uruguay niet voorspeld worden, en zou er ten onrechte vorsttolerantie voorspeld kunnen worden in wilde aardappels uit het zuiden van de VS. Omdat veel wilde aardappel soorten in een klein gebied voorkomen, zijn taxonomische, geografische, en ecologische factoren vaak met elkaar verstrengeld. Deze factoren dienen tegelijkertijd bestudeerd te worden om de patronen in de verspreiding van eigenschappen te kunnen interpreteren.

Deel II

De studies in het tweede deel van dit proefschrift dragen bij aan het gebruik van agro-ecologische zonering (AEZ) voor het sturen van onderzoek aan gewassen. Een belangrijke tekortkoming van veel AEZ studies is dat de resultaten moeilijk te interpreteren en te gebruiken zijn omdat ze te algemeen zijn en niet gerelateerd aan de huidige verspreiding van het gewas. Dit kan deels veroorzaakt zijn door een gebrek aan gedetailleerde gegevens over de geografische verspreiding van gewassen. Om deze leemte te vullen voor de aardappel werd een mondiale aardappelverspreidings-database ontwikkeld. In Hoofdstuk 6 wordt deze beschreven en geanalyseerd in relatie met onder andere breedtegraad, hoogte, en bevolkingsdichtheid.

In de verdeling van de aardappel over de breedtegraden zijn er twee gebieden met opmerkelijk veel areaal. Aardappel komt het meest voor tussen 45° en 57°N, in de gematigde klimaatsgebieden waar het gewas in de zomer verbouwd wordt. De andere piek in de verspreiding is tussen de 23° en 34°N, waar de productie vooral in de subtropische laaglanden plaatsvindt en de aardappel in de winter verbouwd wordt. Tussen 1950 en 1998 is het aardappelareaal toegenomen op lage breedtegraden en afgenomen op hoge breedtegraden, met name rond 53°N. De noordelijke productiegrens van de aardappel komt overeen met die van de landbouw en van de menselijke bewoning. Het gebied met veel aardappelareaal tussen 23° en 34°N, valt samen met het gebied van de hoogste bevolkingsdichtheid (per eenheid land en landbouwgrond). Ongeveer 25% van het mondiale aardappelareaal bevindt zich in de hooglanden (boven 1000 m).

De "aardappelziekte", veroorzaakt door *Phytophthora infestans*, is een van de meest verwoestende plantenziektes—zoals geïllustreerd werd door de Ierse hongersnood in jaren 1840. Omdat de aardappelziekte sterk van het weer afhankelijk is, met name van de luchtvochtigheid, kan agro-ecologische zonering zinvol zijn om inzicht te krijgen in de ruimtelijke variatie van de ziekte en de implicaties daarvan voor de gewasbescherming. In Hoofdstuk 7 worden mondiale weersgegevens en ziektevoorspellingsmodellen gebruikt voor kwantitatieve en specifieke AEZ voor de aardappelziekte. Het doel van deze studie is bij te dragen aan het beter inzetten van technologie, zoals resistente variëteiten of het gebruik van fungiciden, voor de beheersing van de aardappelziekte in ontwikkelingslanden.

Twee modellen, Blitecast en Simcast, werden gekoppeld aan klimaatsgegevens in een GIS. De modellen geven een indirecte schatting van de ziektedruk door te voorspellen hoe veel fungicide bespuitingen nodig zijn tijdens het groeiseizoen om de aardappelziekte te bestrijden, afhankelijk van de weersgesteldheid. De zones van gelijke ziektedruk waren vergelijkbaar bij de twee modellen, maar Blitecast voorspelt in het algemeen een lager aantal benodigde bespuitingen. Zones met een hoge ziektedruk zijn o.a. de tropische hooglanden, West-Europa, de oostkust van Canada en van het noorden van de VS, Zuidoost-Brazilië, en centraal-zuid China. Belangrijke aardappel productiegebieden waar de ziektedruk niet hoog is zijn o.a. de westelijke riviervlaktes van India, waar aardappels tijdens het droge seizoen worden verbouwd, noord-centraal China, en het Noordwesten van de VS. Met behulp van de gegevens over de mondiale aardappelverspreiding werd het gemiddeld aantal benodigde bespuitingen per land berekend. Deze werden vergeleken met schattingen van het huidige gebruik van fungiciden. Ofschoon de resultaten verkregen met Blitecast en

Simcast correleerden, correleerde alleen de resultaten verkregen met Blitecast goed met de schatting van het huidige gebruik in ontwikkelde landen. Er was geen correlatie tussen de voorspelde (optimale) en het geschatte huidige gebruik van fungiciden in ontwikkelingslanden. In een groot aantal ontwikkelingslanden was het voorspelde optimale aantal bespuitingen veel hoger dan het huidige gebruik daarvan. In deze landen zou toegang tot resistente variëteiten en fungiciden een groot positief economisch effect kunnen hebben.

Een nieuwe methode om het potentiële effect van voorgesteld onderzoek te evalueren wordt beschreven in Hoofdstuk 8. Het effect van het gebruik van vorsttolerante aardappelvariëteiten in de Altiplano van Peru en Bolivia werd geschat met behulp van een simulatiemodel en geïnterpoleerde klimaatsgegevens. Om verlies van informatie te voorkomen is er strikt gesproken geen spraak van zonering. De berekeningen werden voor iedere individuele grid cel uitgevoerd, zonder daaraan voorafgaande aggregatie.

Het aardappel simulatiemodel LINTUL werd aangepast om het effect van vorst op de opbrengst te kunnen simuleren. Grids met maandelijkse klimaatsdata, waaronder absolute minimumtemperatuur en de standaardafwijking daarvan, werden gebruikt als input voor het model. Het model werd gebruikt voor een standaard aardappelvariëteit waarvoor de vorstresistentieparameters werden veranderd in stappen van 1°C. Een aardappelverspreidings database werd gebruikt om de resultaten te wegen over het aardappelareaal. Indien vorstresistentie toeneemt van –1°C (huidig niveau) tot –2°C of –3°C, dan neemt de gemiddelde opbrengst toe met respectievelijk 26 en 40%. Een verdere toename van resistentie leidt tot slechts een geringe toename van de opbrengst.

Het effect van klimaatsverandering op de aardappelproductie werd geanalyseerd in Hoofdstuk 9, met behulp van gegevens over het huidige en over toekomstige klimaten, en een simulatiemodel. Meestal wordt dit soort studies voor een beperkt aantal gebieden gedaan, maar deze studie dekt de hele wereld. In de analyse wordt met name gekeken naar mogelijkheden tot aanpassing aan klimaatsverandering door het verschuiven van groeiseizoenen of het gebruik van variëteiten met een verschillende ontwikkelingsduur.

Tussen de periodes 1961–1990 en 2040–2060 gaat de gemiddelde globale temperatuur (zonder inachtneming van de temperatuur boven de oceanen en op Antarctica) tussen de 2,1 en 3,2°C omhoog, afhankelijk van het klimaatsscenario. De temperatuurstoename gewogen over het aardappelareaal is kleiner, met name wanneer aanpassing van poottijdstip en variëteiten wordt toegelaten (in dat geval is de voorspelde verandering in de temperatuur tussen de 1 en 1,4°C). Op hoge breedtegraden kan aanpassing aan klimaatsverandering er toe leiden dat boeren vroeger in het voorjaar gaan poten en variëteiten met een langere groeiduur gaan gebruiken. Het zou ook kunnen leiden tot een verschuiving van de locatie van de aardappelproductie. Veranderingen in opbrengst zullen hier waarschijnlijk klein zijn, en soms positief. Aanpassing van poottijdstip en locatie is in veel mindere mate mogelijk op lagere breedtegraden en in deze gebieden zou klimaatsverandering een sterke negatieve invloed op de aardappelproductie kunnen hebben. Om het effect van klimaatsveranderingen te verminderen in deze gebieden zouden warmte-tolerante aardappel variëteiten kunnen worden geselecteerd of ontwikkeld.

Selected publications by the author

Papers in refereed journals

Ghislain, M., D. Zhang, D. Fajardo, Z. Huamán and R.J. Hijmans, 1999. Marker-assisted sampling of the cultivated Andean potato *Solanum phureja* collection using RAPD markers. Genetic Resources and Crop Evolution 46:547-555.

Hijmans, R.J., 2001. Global distribution of the potato crop. American Journal of Potato Research 78: 403-412.

Hijmans, R.J., B. Condori, R. Carillo and M.J. Kropff, 2002. A quantitative and constraint-specific method to assess the potential impact of new agricultural technology: the case of frost resistant potato for the Altiplano (Peru and Bolivia). Agricultural Systems (in press).

Hijmans R.J., G.A. Forbes, and T.S. Walker, 2000. Estimating the global severity of potato late blight with GIS-linked disease forecast models. Plant Pathology 49: 697-705.

Hijmans, R.J., K.A. Garrett, Z. Huamán, D.P. Zhang, M. Schreuder, and M. Bonierbale, 2000. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. Conservation Biology 14: 1755-1765.

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

Hijmans, R.J., L. Huaccho, and D. Zhang, 2002. Description and analysis of a geo-referenced database of the global distribution of sweetpotato area. Acta Horticultura (in press).

Hijmans, R.J., and D.M. Spooner, 2001. Geographic distribution of wild potato species. American Journal of Botany 88:2101-2112.

Hijmans, R.J., M. Schreuder, J. de la Cruz, and L. Guarino, 1999. Using GIS to check coordinates of germplasm accessions. Genetic Resources and Crop Evolution 46: 291-296.

Hijmans R.J., and M.K. van Ittersum, 1996. Aggregation of spatial units in linear programming models to explore land use options. Netherlands Journal of Agricultural Science 44: 145-162.

Low, J., T. Walker, and R. Hijmans. The potential impact of orange-flesh sweetpotatoes on vitamin A intake in Sub-Saharan Africa (submitted).

Quispe, C., A. Devaux, S. Gonzales, C. Tourneux and R. Hijmans, 1997. Evaluación comparativa del desarrollo y crecimiento de papa, oca e isaño en Cochabamba, Bolivia. Revista Latinoamericana de la Papa 9/10 (1): 140-155.

Salas, A.R., D.M. Spooner, Z. Huamán, R.V. Torres-Maita, R. Hoekstra, K. Schüler, and R.J. Hijmans. 2001. Taxonomy and new collections of wild potato species in central and southern Peru in 1999. American Journal of Potato Research 78(3): 197-207.

Spooner, D.M. and R.J. Hijmans, 2001. Potato systematics and germplasm collecting 1989-2000. American Journal of Potato Research 78: 237-268; 395.

Spooner, D.M., A. Salas, Z. Huamán, and R.J. Hijmans, 1999. Potato germplasm collecting expedition in southern Peru (Departments of Apurímac, Arequipa, Cusco, Moquegua, Puno, Tacna) in 1998: Taxonomy and new genetic resources. American Journal of Potato Research 76:103-119.

Walker, T.S., P.E. Schmiediche and R.J. Hijmans, 1999. World trends and patterns in the potato crop: An economic and geographic survey. Potato Research 42: 241-264.

Wiegers, E.S., R.J. Hijmans, D. Hervé, and L.O. Fresco, 1999. Land use intensification and desintensification in the Upper Cañete valley, Peru. Human Ecology 27(2):319-339.

Zhang, D.P., M. Ghislain, Z. Huamán, A. Golmirzaie, and R. Hijmans, 1998. RAPD variation in sweetpotato [*Ipomoea batatas* (L.) Lam] cultivars from South America and Papua New Guinea. Genetic Resources and Crop Evolution 45 (3): 271-277.

Zhang, D., G. Rossel, A. Kriegner, and R.J. Hijmans, 2002. AFLP assessment of diversity in sweetpotato from Latin America and the Pacific region: Its implications on the dispersal of the crop. Genetic Resources and Crop Evolution. (In press).

Books

Hijmans, R.J., D.M. Spooner, A. Salas, L.Guarino, and J. de la Cruz, 2002. Atlas of wild potatoes. Systematic and ecogeographic studies of crop genepools 10. International Plant Genetic Resources Institute, Rome, Italy.

Spooner, D.M., R.G. van den Berg, A. Rodríguez, J. Bamberg, R.J. Hijmans, and S.I. Lara-Cabrera. Wild potatoes (*Solanum* section *Petota*) of North and Central America. Systematic Botany Monographs (submitted).

Book chapters, reports, and proceedings

Bussink, C.B., and R.J. Hijmans, 2001. Land use change in the Cajamarca catchment, Peru, 1975-1996. Program Report 1998-2000. International Potato Center, Lima, Peru. p. 421-428.

De la Cruz, J., P. Zorogastúa y R.J. Hijmans, 1999. Atlas digital de los recursos naturales de Cajamarca. Production Systems and Natural Resource Management Department Working Paper 2. International Potato Center, Lima, Peru.

Forbes, G.A., G. Chacón, M. Taipe, and R.J. Hijmans, 2001. Simulating potato late blight in the highland tropics. Program Report 1998-2000. International Potato Center, Lima, Peru. p. 87-93.

Guarino, L., A. Jarvis, R.J. Hijmans, and N. Maxted, 2002. Geographic information systems (GIS) and the conservation and use of plant genetic resources. In: J.M.M. Engels, V. Ramantha Rao, A.H.D. Brown, and M.T. Jackson (eds.). Managing plant genetic diversity. CABI, Wallingford, UK. p. 387-404.

Hijmans, R.J., I. Guiking-Lens and C.A. van Diepen, 1994. WOFOST, user guide for the WOFOST 6.0 crop growth simulation model. Technical Document 12. SC-DLO, JRC, Wageningen, Netherlands.

Hijmans, R.J., 1999. Global climate surfaces at a 10-minute resolution (GCLIM10). Version 1.0. Production Systems and Natural Resource Management Department Working Paper 3. International Potato Center, Lima, Peru.

Juárez, H.S, J.R. Amaro, M.D. Rivera, A. Párraga and R.J. Hijmans, 2001. The effect of nitrogen fertilization on potato late blight in the field. Program Report 1998-2000. International Potato Center, Lima, Peru. p. 69-76.

Perez, W., S. Gamboa, M. Coca, R. Raymundo, R. Hijmans and R. Nelson, 1999. Characterization of *Phytophthora infestans* populations in Peru. Impact on a changing world. Program Report 1997-1998. International Potato Center, Lima, Peru. p. 31-38.

Posner, J., C. Bussink, R. Hijmans, J. de la Cruz, H. Willet, and P.A. Sanchez, 2001. Putting natural resource management on the map: using GIS as a tool for soil and water conservation planning in two small Andean watershed. Program Report 1998-2000. International Potato Center, Lima, Peru. p. 413-420.

Van Ittersum, M.K., R.J. Hijmans, and D.Scheele, 1995. Description and user guide of GOAL-QUASI: an IMGLP model for the exploration of future land use. Quantitative Approaches in Systems Analysis 2. AB-DLO, TPE-WAU, Wageningen, Netherlands.

About the author

Robert Jan Hijmans was born on 4 July 1963 in Haarlem, the Netherlands. In 1983, he obtained a secondary school (VWO) diploma from the Rijnlands Lyceum in Sassenheim. From 1983 to 1984 he studied Political Science at the University of Amsterdam and obtained the Propaedeuse degree. From 1986 to 1992 he studied Tropical Crop Science at the Wageningen Agricultural University and obtained a Master's degree. During this period he spend 8 months of practical training in Côte d'Ivoire. He wrote theses on The influence of ash, wood and fertilizers on soil fertility in shifting cultivation (experimental work in Côte d'Ivoire); The influence of photoperiod and temperature on growth and development of Bambara groundnut (*Vigna subterranea*) (experimental work in Wageningen); and on Sustainability, productivity and efficiency of land use in Andalucía (Literature research).

Between March 1991 and May 1994 he worked at several departments of the Wageningen Agricultural University. At the Department of Tropical Crop Science he developed introductionary reading material for a field practical on land use in Andalucía, Spain. He was assistant to the professor in courses given by the Departments of General and Regional Agricultural Science, Mathematics, and Theoretical Production-ecology. At the department of Theoretical Production-ecology he worked on translating, simplifying and aggregating GOAL, a linear programming model for analysis for land use in the EU, and developed exercises for the course "Quantitave analysis of agro-ecosystems at higher integration levels". He researched the effect of aggregation on the output of the GOAL model.

He developed course material for the World Food Production course of the Dutch Open University (October—December 1992). At the International Agricultural Center (IAC), Wageningen, he developed case studies on integrated pest management in cabbage and potato production in tropical areas for use in the International Course on Vegetable Production (April—May 1993). He wrote a user guide for the WOFOST crop growth simulation model while employed by the Winand Staring Centre, also in Wageningen (May—July 1994). From January 1991 to December 1993 he was contributor to and editor (as of Oct 1992) of Ekoland, a periodical on organic agriculture.

From September 1994 to August 2002 he worked at the International Potato Center (CIP) in Lima, Peru, where he initiated and led the GIS laboratory. Since 2001 he was also the head of CIP's Bioinformatics laboratory, and a member of the CIP project management team that evaluates CIP's research program. At CIP he initially focused his research on natural resource use in the Andes. Later he shifted his attention to aspects of the conservation and use of root and tuber crop genetic diversity; potato and sweetpotato characterization and impact assessment; and integrated management of potato late blight. Parts of his work on these topics resulted in this thesis. His additional activities at CIP included database, software, and website development, and providing training.

On 1 September 2002 he joined the Museum of Vertebrate Zoology of the University of California at Berkeley where he investigates spatial variation of biodiversity.

Appendix: Color plates

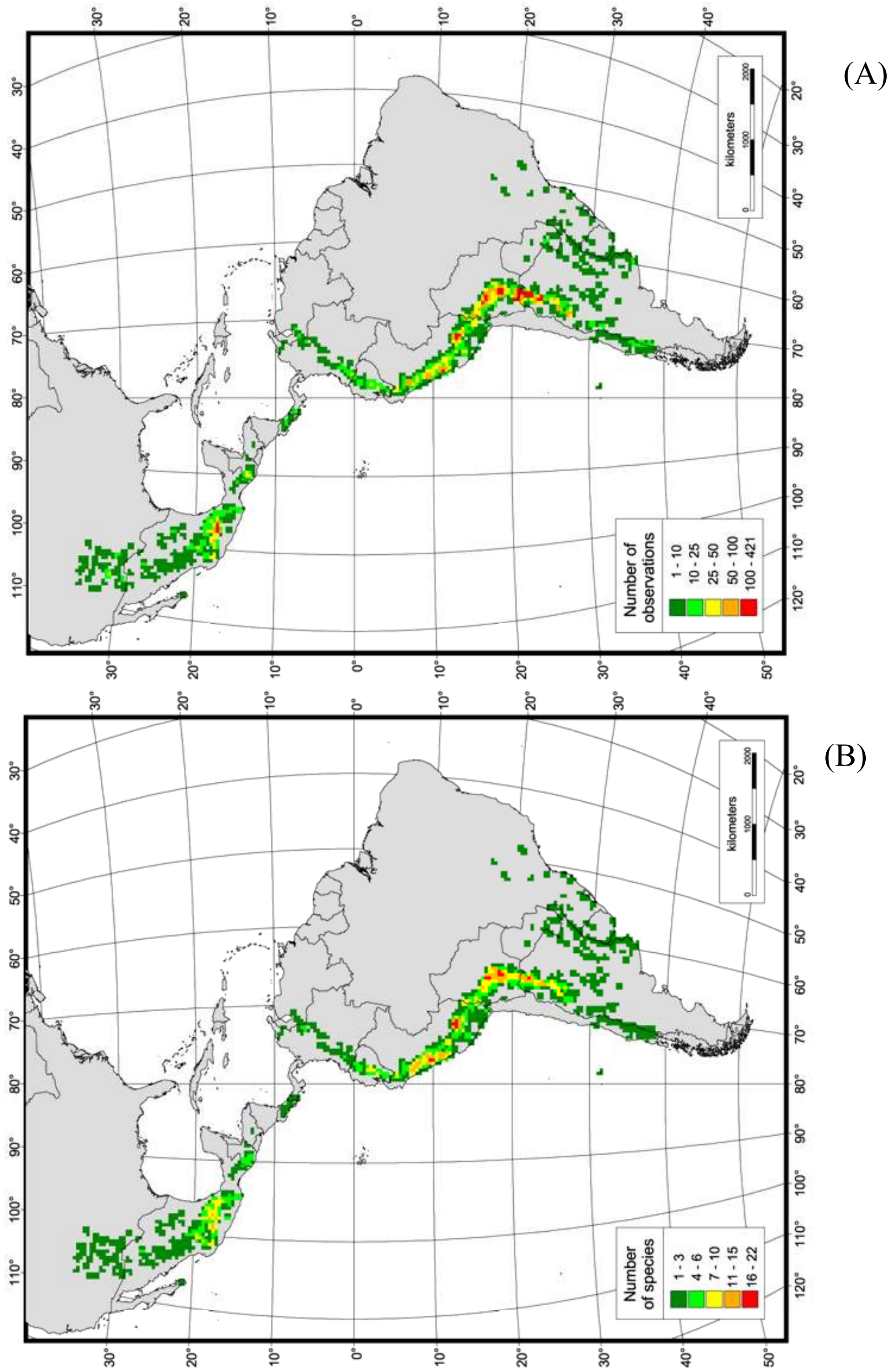


Plate 1. Number of observations (A) and species (B) of wild potato species per 50 × 50 km grid cell. A circular neighborhood with a radius of 50 km was used to assign observations to a grid cell. There are 1317 grid cells with observations. (Chapter 4).

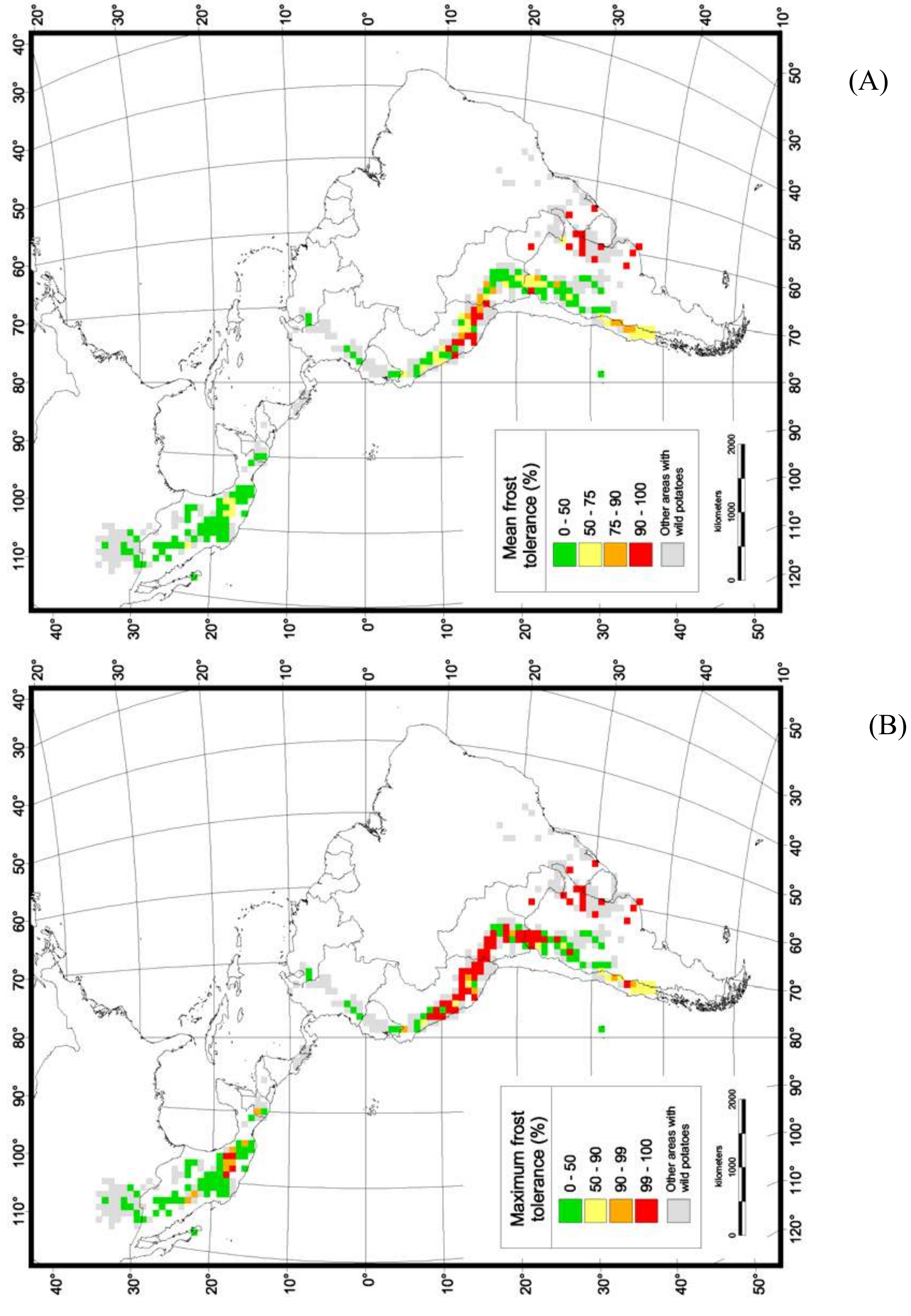


Plate 2 Spatial distribution of the mean (A) and maximum (B) observed frost tolerance (percentage non-damaged tissue) in wild potatoes in 100 by 100 km grid cells. Each observation (*n* = 1646) referred to the locality where a wild potato species was collected. (Chapter 5).

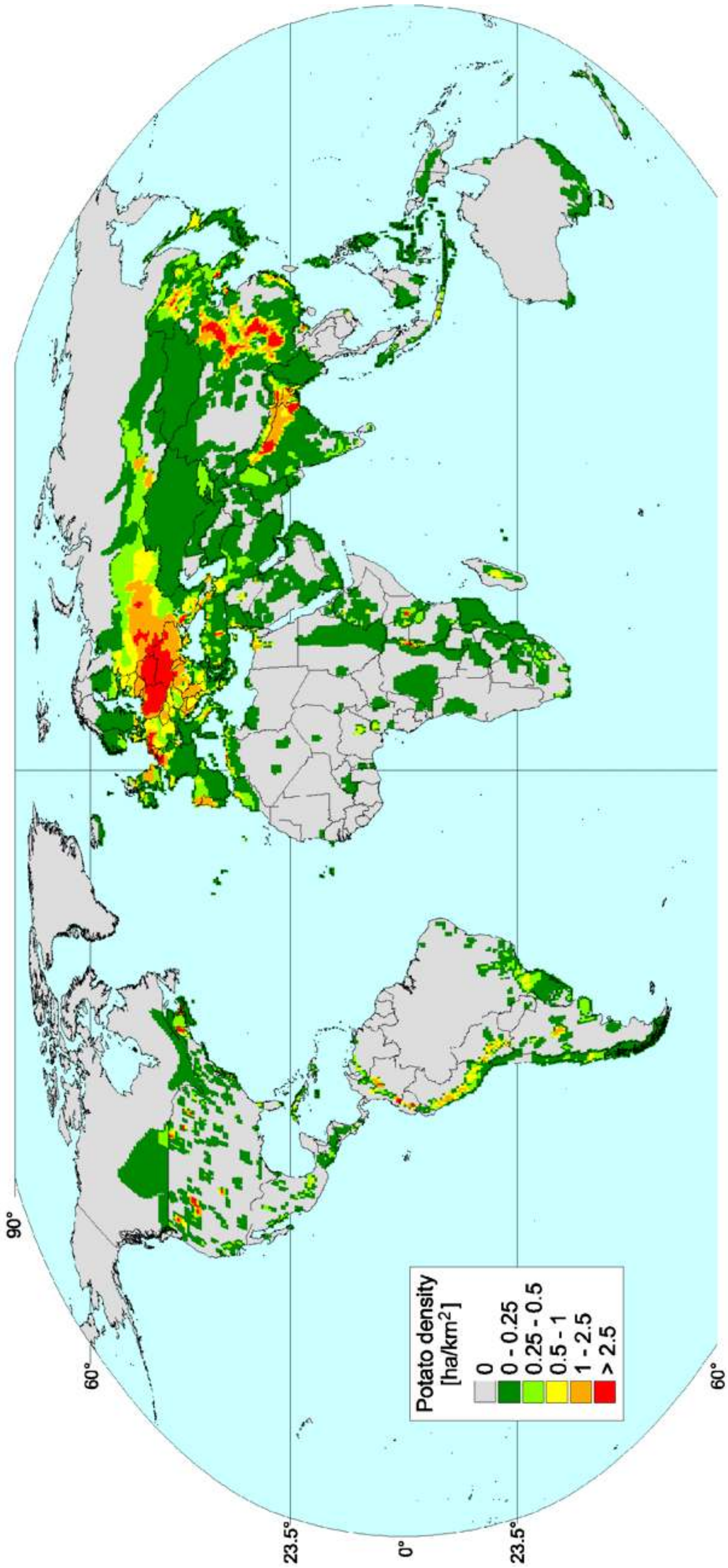


Plate 3 Global distribution of potato area. Expressed as relative potato area (ha/km^2) (30 by 30 minute resolution). (Chapter 6).

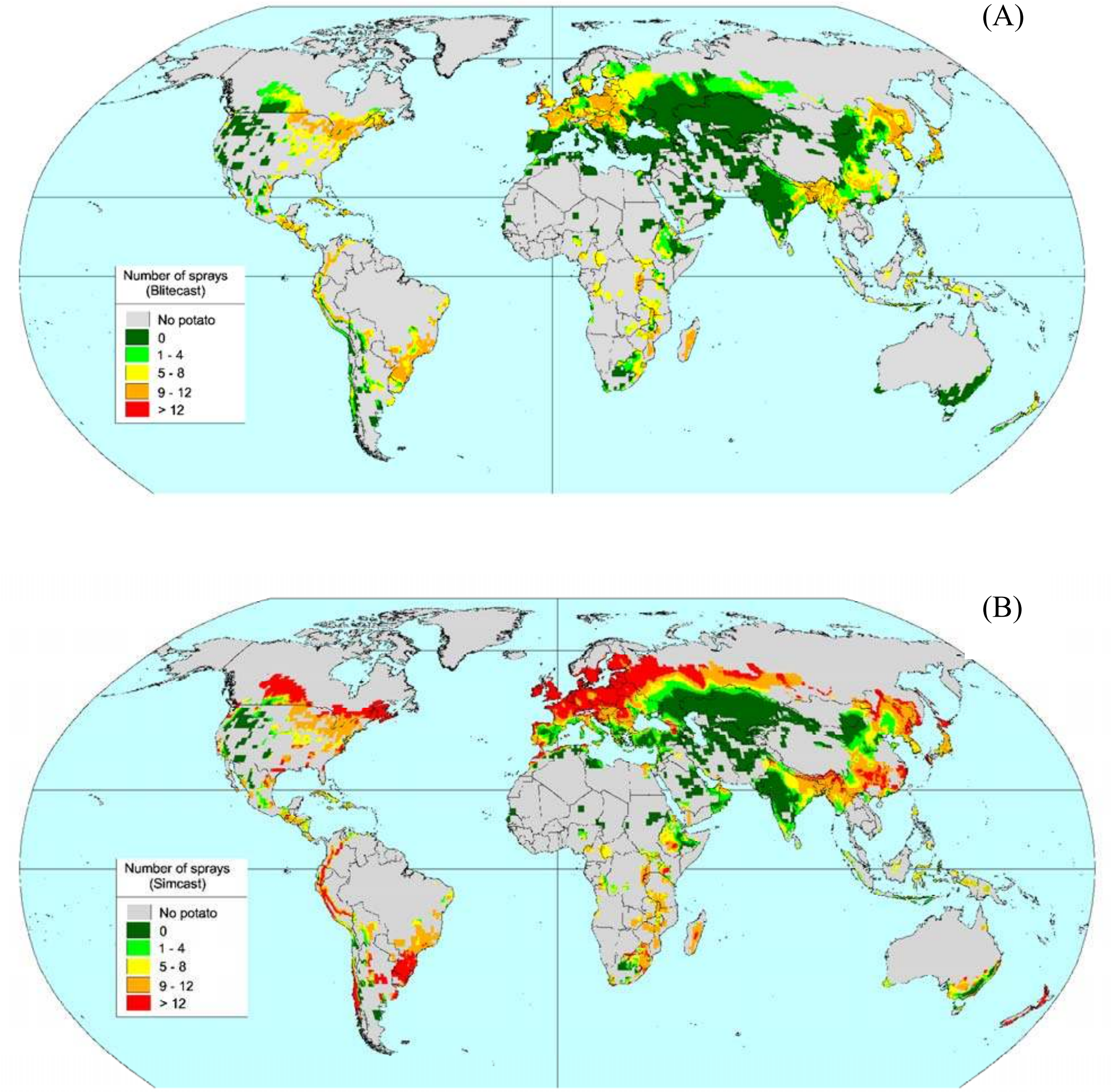


Plate 4 Predicted global late blight severity for potato production zones expressed as the number of protectant fungicide sprays needed to control late blight. Predictions based on a late blight forecast models Blitecast (A) and Simcast (B). (Chapter 7).

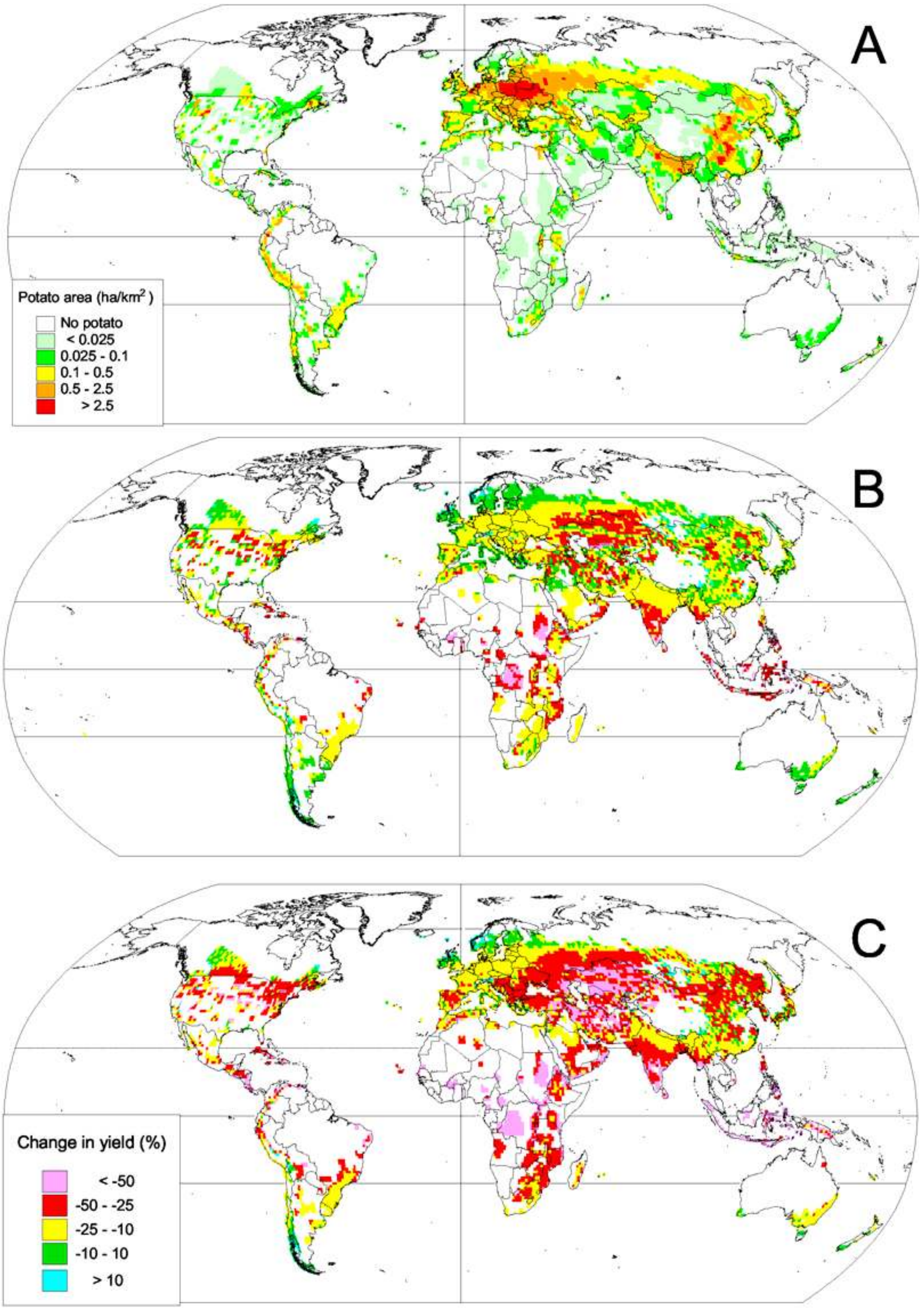


Plate 5 Global potato distribution (A, source: Hijmans, 2001); and mean change in potential potato yield between the 1970s and the 2020s (B) and the 2050s (C) without adaptation of the growing season and maturity type of the varieties. (Chapter 8).

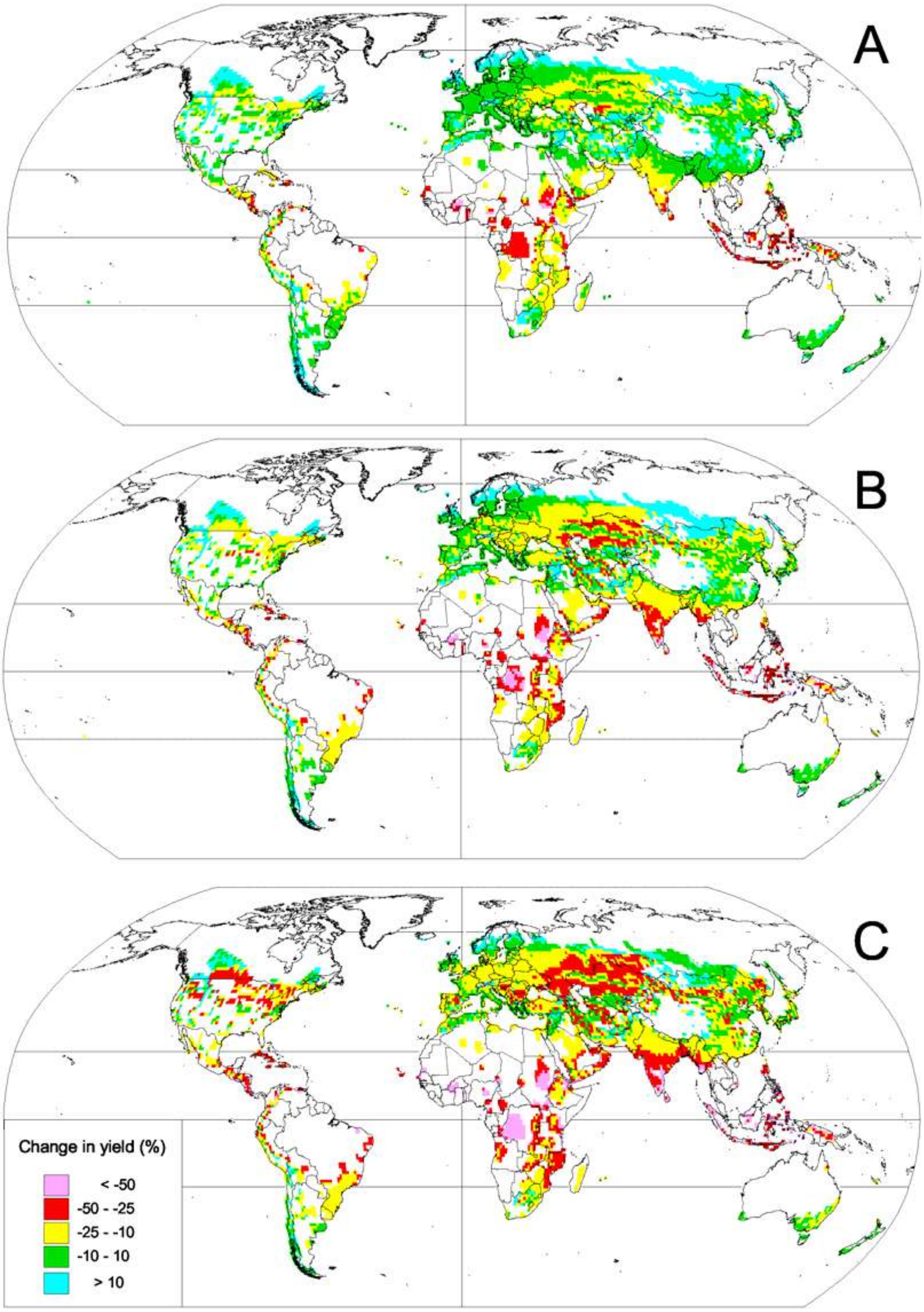


Plate 6 Change in potential potato yield between the 1970s and the 2020s with adaptation of the growing season and maturity type of the varieties. For the highest (A), mean (B) and lowest (C) yield estimates in 2020 over the seven climate scenarios used. (Chapter 8).

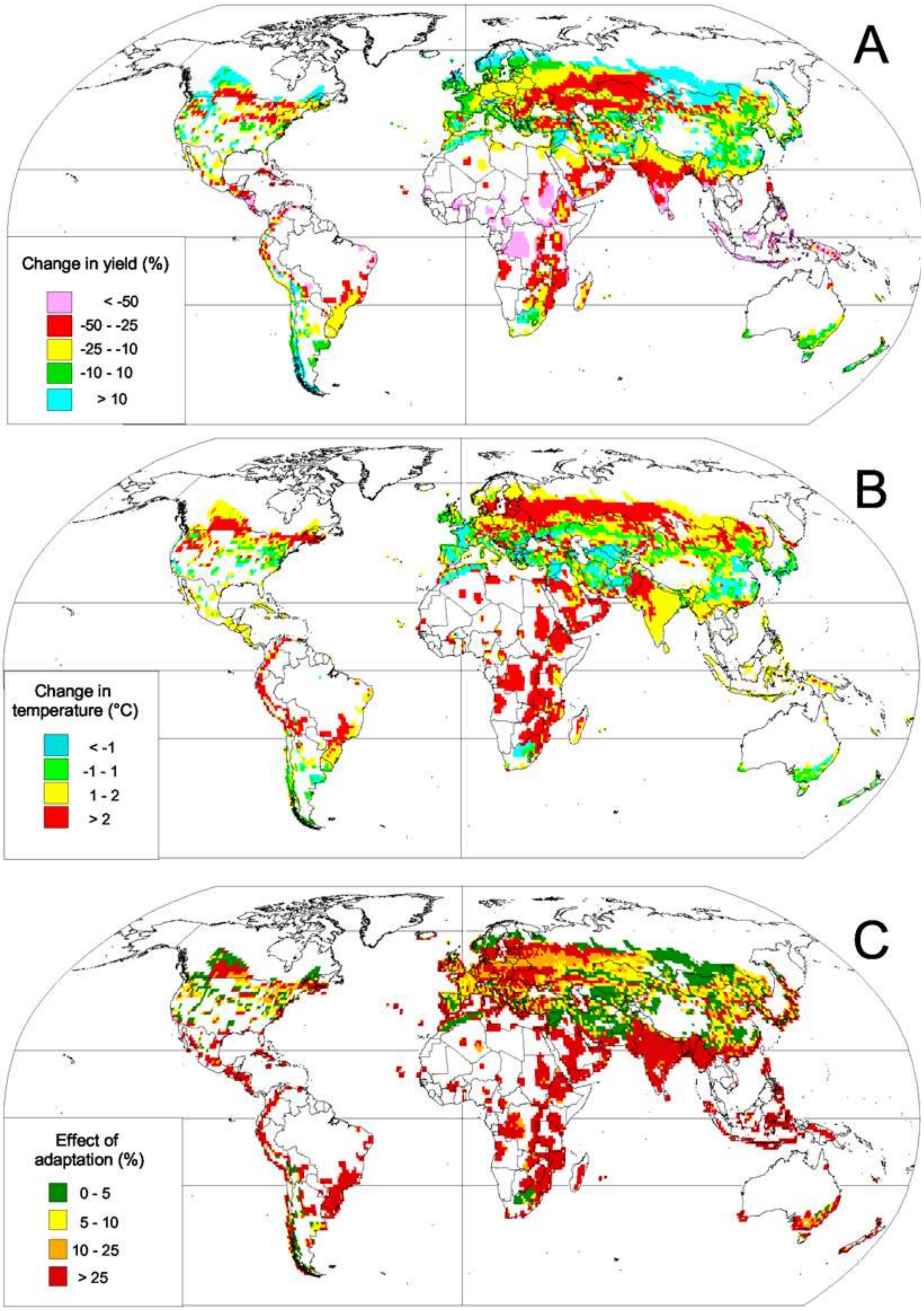


Plate 7 Average predicted changes in (A) potential potato yield; and (B) temperature during the growing season, between the 1970s and 2050s with adaptation of the growing season and maturity type of the varieties; and (C) the relative contribution to potential yield of adaptation over this period. (Chapter 8).

The research reported in this thesis was partly funded by the Development Cooperation branch (DGIS) of the Dutch Ministry of Foreign Affairs, and by additional donors of the International Potato Center.