*Biodiversity Informatics*, 17, 2022, pp. 96-107

# BIODIVERSITY INFORMATICS FOR PUBLIC POLICY: THE CASE OF CONABIO IN MEXICO

Jorge Soberón[1]

[1]*Biodiversity Institute and Department of Ecology & Evolutionary Biology, University of Kansas, 1345 Jayhawk Blvd., Lawrence, Kansas 66045, USA (ORCID https://orcid.org/0000-0003-2160-4148)*

*Abstract*. In this paper, I present and review the development of the biodiversity information system that was developed in Mexico. I describe briefly the organization that made the system possible and some of its history. Then, I focus on the principles of design of the information system, and a few of its major uses. I provide data on costs and usage, and end with some reflections on the fragility of such institutional systems.

*Key words:* Biodiversity information systems; Mexico

Biodiversity is the aggregate of ways in which life manifests itself in the planet (Brooks et al. 2006). Biodiversity is a complex concept that can be defined from multiple perspectives (MacLaurin and Sterelny 2008; Sarkar 2002). A comprehensive perspective is to regard biodiversity as an aggregate of elements, how are they structured, and how they function, at scales from the sub-individual to the planetary (Noss 1990). For instance, at certain scales, the elements of biodiversity are individuals of species, the structure is their spatiotemporal locations, and the functioning is their interactions. At a different scale, elements of biodiversity may be biomes, structure would be their spatial extents, and functioning would be the biogeochemical processes taking place in them.

From this comprehensive perspective, conservation of biodiversity requires actions and policies at multiple scales. Historically, however, biodiversity has been managed mostly at relatively local scales (i.e., at the scale of activities of human groups of small size), by indigenous peoples, farmers, fishermen and such local actors (Gadgil et al. 1993). This "management" has taken place for thousands of years, such that, overall, indigenous and traditional cultures generally have deep knowledge of their environments and respectful attitudes towards nature (Toledo 2001). This proximity leads to a mostly sustainable management of components of biodiversity (Gadgil et al. 1993), since many of the impacts were spatially concentrated, and were reversible in nature. Moreover, traditionally, natural resources were often the subject of strict governance (Ostrom et al. 1999), as opposed to the naïve view of traditionally managed resources as open-access "commons" (Hardin 1968). Traditional governance is, in the end, highly conducive to sustainable use (Gadgil et al. 1993).

In modern times (i.e., over the last ~400 years), however, the rate at which human activities have impacted biodiversity has accelerated (Butchart et al. 2010; Ehrlich 1995; McNeely et al. 1990; Steffen 2015). Actors beyond the local now exert substantial impacts on different components of biodiversity, sometimes in ways that are spatially very extended or have long-term effects, and that sometimes are irreversible. Governance of common-pool resources of global extent is challenging (Ostrom et al. 1999). Managing and conserving biodiversity, therefore, requires participation of stakeholders at many different levels, which creates problems of obtaining and assembling the required information. At first, emphasis was placed on spatially structured information, in effect "putting biodiversity on the map" (Bibby 1992; Edwards et al. 2002; Reid 1998; Scott 1993). In practice, however, this emphasis meant putting the biodiversity *of developed countries* on the map, and biodiversity loss was not abated elsewhere (Peterson and Soberón 2018).

Still, some voices have insisted that, without biodiversity data, management would be difficult or impossible (Balmford et al. 2005). Indeed, when viewed from a multilevel perspective, management of the multitude of entities and processes comprising biodiversity is impossible without an overarching perspective. In the context of widespread loss of the components, structure, and functioning of biodiversity, the countries of the world negotiated a "Convention on Biological Diversity" (Koester 2002), which stressed a dire need for globally relevant biodiversity data to be made available openly to the broadest community (Laihonen 2004).

What are "biodiversity data" then, how can biodiversity data be managed in accessible ways, and what can they be used for? The core of this paper is an attempt to answer these questions, from a mainly historical perspective, using the case of the Mexican national biodiversity agency (the *Comisión Nacional para el Uso y Conocimiento de la Biodiversity*, or CONABIO) as an example. From its creation in 1992 until 2005, I served as the Executive Secretary of CONABIO. It is from this perspective that I write this paper.

## Biodiversity Data

As stated above, "biodiversity" is a complex concept, being both multi-scale and multi-perspective. Numerous perspectives on biodiversity have been documented in countless books, papers, images, recordings, and databases regarding protein structure, genetic sequences, species diversity, community ecology, etc. However, in practice, the key, focal concept has been that of records of occurrence of a species, otherwise known as primary biodiversity data (Peterson et al. 2010; Soberón and Peterson 2004; Sousa-Baena et al. 2014).

The key idea of primary biodiversity data is that each record comprises a date, a description of a locality, and a taxonomic identity (Johnson 2007; Soberón and Peterson 2004). The locality data allow linking to geographic information, and the taxonomic identity provides an index to genetic, demographic, systematic, or cultural data. The importance of the taxonomic identity in linking databases cannot be overemphasized. Solving all the "knowledge shortfalls" described for biodiversity (Hortal et al. 2015) is predicated on having a consistent and stable system of names, which in biology is based on Linnean taxonomic schemes. The names constitute a "hinge feature" of primary biodiversity data, in fact linking geography with a multiplicity of perspectives, via the name, which is of fundamental importance (Chapman 1991; Peterson et al. 2010). In what follows, I will be focusing on this core of primary biodiversity data, mainly because in practice it has been the focus of most large-scale biodiversity informatics initiatives (Coetzer 2012; CONABIO 2012; Sandlund 1991).

## The Beginnings of CONABIO

In June of 1992, the United Nations organized the conference on Environment and Development (also known as the "Earth Summit"). This took place in Rio de Janeiro, Brazil. In preparation for this, the then-president of Mexico asked the Chancellor of the National University, José Sarukhán, the foremost ecologist of Mexico, to provide some possible initiatives to present in Rio. In February 1992, a meeting was organized in Mexico (Sarukhán and Dirzo 1992) to begin designing a national initiative on biodiversity for the country. As a consequence of this meeting of international experts (mostly in biodiversity conservation), two of the most prominent ecologists of Mexico (Daniel Piñero and Rodolfo Dirzo, both researchers in the Institute of Ecology of the National University) worked with Sarukhán to propose to the President of Mexico to create a high-level government agency in charge of biodiversity. In 1992, an inter-ministerial commission was created (CONABIO), composed of ten cabinet-level ministers, and presided ex-officio by the President of Mexico. I was appointed Executive Secretary of CONABIO, a role in which I served for 13 years.

Although CONABIO is formally a multi-ministry federal government agency, it operates via an Executive Secretariat that was allowed to establish a private trust fund via which to operate. This hybrid structure, combining private and public aspects, gave CONABIO not only the capacity to address challenging technical tasks, but also to act as a trusted and necessary government interlocutor. CONABIO was given a number of tasks. The most important was: "*To synthesize information relative to the biological resources of the country, in a database that should be kept permanently updated*." This activity was the initial and major focus of CONABIO: to this end, the first step was to take stock of similar initiatives elsewhere in the world.

The CONABIO team obtained information by visiting four existing organizations around the world. First, we consulted one in India, now extinct, that had worked entirely based on secondary information (bibliography). Although the system was open to the public, it was entirely based on secondary data, making that consultation a dead end. A map was a page in a publication (as opposed to a machine-readable geospatial dataset), and a list of occurrence localities was an image of some text. This system was essentially a bibliographic consult system, and was not a useful lesson for Mexico.

The second system that we studied was that of the Heritage Methodology of *The Nature Conservancy* (Groves 1995). This system was based on *primary data*, obtained from public museums in the United States and Canada, among other sources. The fact that it used primary data meant that a variety of operations could be performed (e.g., performing statistical analyses or visualization of patterns in maps and graphs) on the data (Stein et al. 2000), but the data were not available to the public. However, it was regularly used in for-profit consultations, leading to widespread resentment among museum officials, who had provided the data for free, without imagining a for-profit use. Therefore, eventually, many sources of data closed to this system, and it clearly was not a model that we wanted to follow in Mexico. This unfortunate situation has seldom been discussed in the literature, but anecdotally it is well known in the community. The experience led to another principle in CONABIO: if the data were to be used for a for-profit purpose, the user would need to consult with the original sources.

A third system was that of Costa Rica's *Instituto Nacional de Biodiversidad* (InBio). This database, which was still in a design phase when we visited, was based on primary biodiversity data, mostly obtained from *de novo* collections performed and maintained by InBio (Tangley 1990). At the time of our visit, the system was still in incipient stages. Also, although the system was based on primary data, it had a rather narrow focus on bioprospecting for pharmaceutical products (Sittenfeld and R.Villers 1993).

Finally, in 1992, personnel of CONABIO, as well as an international group including Kenyan, Indonesian, Costa Rican, and U.S. American scientists (Chapman 2001), visited the Environmental Resources Information Network (ERIN), in Australia (Kaye et al. 1997). ERIN has since disappeared, although many of its capabilities were replaced by the Atlas of Living Australia (Belbin 2021). In the 1990s, Australia was without a doubt the most advanced country in the world in biodiversity informatics. Their system was based on primary biodiversity data, provided in largest part by the network of Australian herbaria and museums. They had sophisticated bioinformatics capacities for taxonomic descriptions (Dallwitz 1993), species distribution modeling (Booth 2018; Busby et al. 1991; Nix 1986), prioritizing sites for conservation (Pressey et al. 1993), and more generally for organization, visualization and analysis of large databases of primary biodiversity data. The ERIN system was open to the public (even at a time when HTML was not yet operational), which in practice was principally academic users, though the users were many, and the types of applications were varied (e.g., designing conservation plans, and surveying poorly explored localities).

The Australian experience, compared with the others, suggested great potential for a biodiversity information system based on two key principles:

**Primary biodiversity data**. That is, the data should be as little interpreted as possible. Essentially a name, a date, and a locality associated with a physical specimen. Combining the data, interpreting, visualizing, and analyzing the data is the responsibility of the users (Soberón and Peterson 2004).

**Data publicly available**. Data should be completely and openly accessible to everyone. When CONABIO was launched, in 1992, the World Wide Web was just being developed (Berners-Lee 1992), but computer scientists at CONABIO were already aware of it, and appreciated its potential to allow efficient public access to what was going to be a large amount of data.

Neither of these two fundamental points had been obvious at that time. Regarding the utility of primary data, there were many expressions of doubt. Most advisors to CONABIO were used to reading books and papers, not to performing their own analysis using large databases (recall that large-scale, publicly available databases of primary data were basically non-existent at this point in time). Nevertheless, CONABIO opted for primary data, following the experience of Australia, and what would eventually become the case in Costa Rica.

On public access, at the time at which CONABIO was starting, attention to the problem of so-called biopiracy (Reid 1996; ten Kate 1999) was at its most intense. The authorities of CONABIO were under considerable pressure not to allow public access to the information, in case commercial agents might misuse it. Moreover, some museum curators opposed releasing collections-associated data (Graves 2000) for other reasons. For instance, it was argued that scientists working with vertebrates might be targeted for animal-rights concerns, or that the data were of monetary value. Several prominent Mexican biologists were similarly adamant in their refusal to share data. After almost a year of such intense discussions, CONABIO convened a meeting of Mexican museum curators and directors, to discuss the issue of public access to data via the internet. In November 1993, in Oaxaca, Mexico, a declaration was issued[1] stating that the Mexican museums and herbaria were committed to computerizing and distributing biodiversity data. As such, an important political battle had been won.

However, at that time, in Mexico (and indeed worldwide), very few biological collections had been digitized. What is more, no effective implementations existed for sharing data on the internet. Finally, despite having signed the Oaxaca Declaration, many curators still had serious misgivings (expressed in private) about public access to biodiversity data! Nevertheless, the signed commitment by Mexican scientists gave CONABIO the legitimacy to start computerizing collections and developing technologies by which to share the data.

## The Sistema Nacional de Información de la Biodiversidad (SNIB)

Building a robust and stable computer system capable of dealing with the millions of data elements about biodiversity took CONABIO more than 10 years (Sarukhán et al. 2014; Soberón and Koleff 2000). The cost was substantial, since most of the data were not yet digitized, and that process required resources to pay experts to travel to collections, acquire computers, and curate data. The cost of digitizing specimens (Figure 1) was on the order of millions of dollars, paid by the Mexican federal government. The figure shows the cost and yield (i.e., number of biodiversity records) for each of 221 projects supported by CONABIO between 1993 and 2000 that digitized or produced records for the main database.

The other major element making up the SNIB was remote sensing, mostly oriented toward monitoring at the ecosystem level. To this end, the Mexican government purchased a satellite dish and associated hardware and software, capable of downloading images from the Moderate Resolution Imaging Spectroradiometer (MODIS) in the Terra and Aqua satellites in real time. This purchase was an investment on the order of many hundreds of thousands of dollars (paid by the Mexican government), and required hiring foreign experts familiar with remote-sense technology. The foreign experts were paid mostly by the German

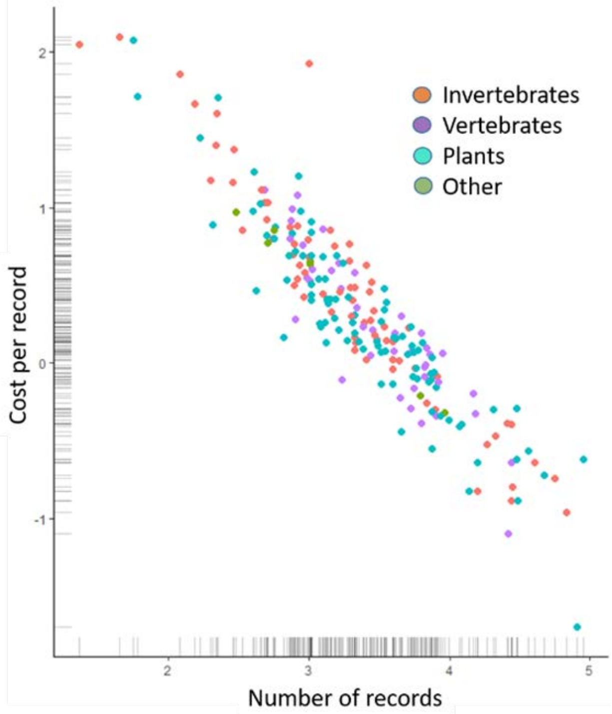


Figure 1. Cost (in contemporary US dollars) of digitizing biodiversity collections, as a function of the number of specimens digitized (note that data are on a log-log scale). Each point is a digitization project. The data for this figure were sourced from internal CONABIO reports. The "rugs" along each axis show the distribution of points.

GTZ cooperation agency, with a symbolic contribution from CONABIO. The German experts came to work at CONABIO under the "Shared Experts" scheme of the GTZ, which guaranteed several years of work in the host country. This long-term participation was key to the success of the project. Although acquiring the remote-sensing infrastructure was costly, delays inherent in acquiring the same images from commercial or noncommercial foreign sources made the purchase necessary, mainly for initiatives to monitor disasters such as wildfires.

By 2005, CONABIO had spent about US$10M of taxpayer's money in acquiring data and remote-sensing hardware, and the computers and system engineers required to run the system. The cost of acquiring primary biodiversity data remained constant per project, on average, at US$5,500 per project. But the cost per specimen is inversely related to the size of the collection (Figure 1), which means that is more efficient to computerize large collections. On the other hand, the experience in Mexico was often that the larger institutional collections tended to be less willing to participate in these initiatives.

[1] http://www.conabio.gob.mx/remib/doctos/declaracion.html.

Table 1. Main informational elements in the *Sistema Nacional de Información de la Biodiversidad* of Mexico (SNIB, based on the 2017-2019 CONABIO Activities Reports[2])

| Data type | Number | Link |
|---|---|---|
| Primary data records | 14,000,000 | https://www.snib.mx/ejemplares/descarga/ |
| Images | 155,000 | http://www.conabio.gob.mx/otros/cgi-bin/herbario.cgi |
| Taxonomy controlled vocabularies | 103,000 | https://www.snib.mx/taxonomia/descarga/ |
| Remote sensing images | 582,000 | http://www.conabio.gob.mx/informacion/gis/ |
| Digital maps | 14,000 | http://www.conabio.gob.mx/informacion/gis/ |
| Technical data about species | 4000 | https://www.gob.mx/conafor/documentos/fichas-tecnicas-especies-exoticas-invasoras; https://enciclovida.mx/ |

By 2005, CONABIO had accumulated a substantial storehouse of data comprising primary biodiversity records, satellite images, photographs, maps, and textual data (Table 1). The SNIB is the computer system that organizes all of these information resources, to assure both efficient access and open sharing.

An outline of the technical details of the system has been published elsewhere (Sarukhán and Jiménez 2016), but stressing that the system is based on the two principles stated above: the backbone is primary biodiversity data, and all data are openly available.

The sheer amount of data means that the expenditures involved are substantial, in terms of hardware and human resources. More precisely, the Mexican taxpayer, and some foreign agencies (the German GTZ, specifically) invested more than US$10M in the system. For comparison, the Convention on Biological Diversity spent $12,300 per country on its "Biodiversity Clearing House Mechanism" (Reed 2017). The resources spent by CONABIO included not only expenses involved in capturing, organizing, and analyzing data, but also in design and implementation of the computer system to manage it (Soberón et al. 2010).

The need to keep the data updated means that hundreds of Mexican (and some foreign) scientists' participation was crucial to the success of the system. Maintaining such participation requires money, time, and effort.

Despite the fact that much was developed in-house, SNIB is compliant with important international efforts. Specifically, the data architecture follows the "Darwin Core" (Wieczorek et al. 2012). Data quality control was influenced by the work of Chapman (2005) and Wieczorek et al. (2004); and the primary data can be accessed via the Global Biodiversity Information Facility (Lane and Edwards 2007). Digitizing data on the labels of millions of specimens was accomplished mostly by hand, often (mostly in herbaria) by taking photographs of the specimen sheets and capturing the data in Mexico. Digitizing specimens is now a major activity all over the world (Asase et al. 2020; Canhos 2017; Nelson and Ellis 2019; Siebert and Smith 2004), one that is increasingly technological (Beaman and Cellinese 2012; Tegelberg et al. 2014).

The SNIB is more than just a data repository, complex as this task is. There are serious analytical capacities developed in the area of biodiversity informatics. Among the principal skills are those related to visualizing data (Stephens et al. 2017), predicting species' geographic distributions (CONABIO 2012), assembling complex remote-sensing products (Gonzalez et al. 2014; Hruby et al. 2016), monitoring wildfires (Ressl et al. 2009) and others. Biodiversity informatics, in a wide sense, is now a major activity in CONABIO, with engineers, mathematicians, taxonomists, and remote-sensing experts collaborating in the activities.

## Usage of SNIB

The primary data that CONABIO has assembled have been used regularly for many government purposes. This is also the case in other parts of the world (Guisan et al. 2013), but the Mexican examples are very illustrative.

Before discussing some examples of use of data for policy, it is interesting to mention that much of the data are used without CONABIO knowing the purpose. That is, the primary data of CONABIO are accessed very frequently. Indeed, CONABIO's website is accessed many thousands of times per week

[2] https://www.gob.mx/cms/uploads/attachment/file/548546/informe-conabio-2017-2019.pdf.

(Figure 2), with data being downloaded at the level of gigabytes (internal communication), although the organization is not aware of the purpose of the use of data downloads.

One concern at the beginning of CONABIO was that most users of open biodiversity data would be foreign "biopirates" (ten Kate 1999). In Table 2, I show the data on access, over the last four years, by country domain. It shows that (by a factor of ~100-fold), most users are Mexicans, not foreigners. Anecdotally, it is known that most users of CONABIO data are researchers, NGOs, or Mexican government agencies.

*Planting permits for GMOs*

In Mexican legislation, planting genetically modified organisms (GMOs) is forbidden if there is a risk of introgressions of modified sequences into wild relatives. CONABIO implemented a system of predicting the risk which is based on ecological niche modeling (a computational method used to predict areas of distribution) applied to wild relatives

Figure 2. Number of unique users of CONABIO website who had at least one session within 7-day time periods between April 2018 and August 2022.

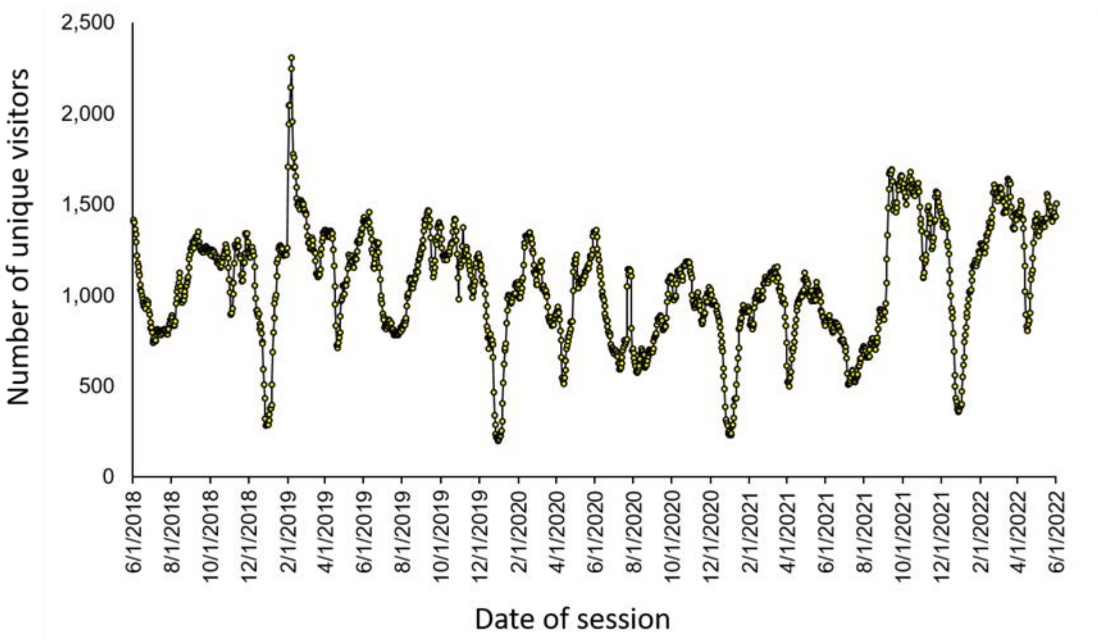


Table 2. Statistics on visits to CONABIO's website over the last four years, with data sourced from Google Analytics in August 2022. Note that most users of CONABIO databases are in Mexico.

| Country | Users | Sessions | Average time (s) |
|---|---|---|---|
| Mexico | 173,613 | 409,283 | 104 |
| United States | 2,501 | 3,908 | 67 |
| Colombia | 1,099 | 1,421 | 53 |
| Peru | 900 | 1,171 | 60 |
| Spain | 644 | 918 | 70 |
| Ecuador | 588 | 724 | 45 |
| Argentina | 424 | 586 | 68 |
| Canada | 298 | 555 | 141 |
| Guatemala | 294 | 417 | 81 |
| Total (4 years) | 184,148 | 424,825 | 103 |

of candidate species (Soberón et al. 2002). This system has proved to have predictive ability (Wegier et al. 2011), it is transparent and empirical (i.e., based on data), and was adopted by the Ministries of the Environment and of Agriculture of Mexico. The system is complicated, in the sense that it uses a variety of databases, predictive algorithms and software tools (Acevedo et al. 2016). However, it is practical, and it has been accepted by major stakeholders. By 2005, more than 1000 permit applications had been assessed with the corresponding recommendations issued to the authority in the Ministry of Agriculture.

*Invasive species*

A major use of CONABIO's databases and capabilities in biodiversity informatics has been in assessing the risk of invasive species, mostly plants of economic importance (Goettsch et al. 2021). The first example originated with an information request from the U.S. Department of Agriculture, about any known occurrences of the moth *Cactoblastis cactorum*, a well-known pest of cacti (Zimmermann et al. 2000) in Mexico. This request (via Mexico's Ministry of Agriculture) lead to one of the first niche modeling exercises (Simonson et al. 2005; Soberón et al. 2001) performed by CONABIO. After several attempts at convincing the Mexican Government about the importance of the problem, the Ministry of Agriculture of Mexico finally organized a campaign of monitoring and control for this pest species (Hernández et al. 2007).

*Wildfire monitoring*

Mexico is a large country, with complex topography and large forested and inaccessible regions. Monitoring of wildfires is done by CONABIO via its remote sensing capabilities (CONABIO 2011). The system, entirely developed at CONABIO (Ressl et al. 2009), uses daily data from the MODIS sensor, and state of the art algorithms, to produce maps (published daily online) of "hot points" across Mexico, Central America and the southern United States. The software automatically issues emails to relevant local authorities in areas of Mexico where wildfires are spotted.

It may be interesting to note that the capacities of CONABIO for remote sensing, as applied to wildfires, were the first test of the power and promise of a biodiversity informatics-focused organization. The daily data about the occurrence of wildfires over the entirety of Mexico was a test not only of the technical capacities of the organization, but also of its political clout, since data about wildfires involved major budget investments, issues of federalism, and even issues of national security. CONABIO was, on a daily basis, monitoring the entire country, and issuing daily reports of direct relevance. One of the first tests of CONABIO's commitment to open data was the wildfires system, since many powerful agents in the federal government were staunchly opposed to what eventually happened: the wildfires reports were made public, daily, over the internet. Wildfires monitoring was also one of the first occasions for using biodiversity informatics in a diplomatic context, since CONABIO was monitoring wildfires also in Central America. Whether or not to share such information required diplomatic negotiations.

*Ecosystem Monitoring*

The capacity to monitor wildfires lead quickly to other monitoring initiatives. Specifically, CONABIO initiated efforts to monitor mangrove cover (Valderrama et al. 2014), marine photosynthetic activity (Cerdeira-Estrada and López-Saldaña 2008), and ecosystem health (García-Alaniz et al. 2017; Gebhardt et al. 2014). The capacity to use remote sensing to monitor functioning of ecosystems is of great utility to government agencies. However, since biodiversity is a multi-scale phenomenon, the components and processes at the local scales should not be forgotten. Monitoring at the scale of populations and their interactions is a significant challenge, as I outline in the next section.

*Wildlife Monitoring*

Recently, CONABIO has started attempts to monitor wildlife. In 2010, working as partners of the National Commission of Forestry (CONAFOR, *Comisión Nacional Forestal*) and of the National Commission of Protected Areas (CONANP, *Comisión Nacional de Areas Naturales Protegidas*). CONAFOR runs a forestry monitoring scheme, and CONABIO began adding recorders and infrared cameras to >3000 of the 25,000 monitoring sites that CONAFOR maintains (Medellín and Corrales 2019)[3]. Although 3000 monitoring sites appears to be a large number, Mexico is a large country, with nearly 2M km$^2$, so the density is only 0.0015 sites/km$^2$. Despite this low density, hundreds of thousands of sound or

[3] https://sipecamdata.conabio.gob.mx/mapa.

image files have been processed (Dirzo et al. 2021)[4]. Processing the deluge of data produced by cameras and recorders has required that CONABIO recruit experts in artificial intelligence and pattern recognition. Moreover, the system requires active participation of local stakeholders, of NGOs, and of government agencies at federal and state levels. This effort is at the level of pioneer, and its applications to policy are still in the future.

*Biodiversity exploration*

Where to conduct biodiversity explorations, which are expensive in funds, time, and personnel, was one of the first questions that CONABIO had to answer, to use public resources in an efficient way. This work was accomplished using the primary data repositories, combined with remote-sensing information about land use (Soberón et al. 2004). Essentially, CONABIO worked to identify areas that were simultaneously poorly sampled and with low human impact, to prioritize for exploration. For instance, there were large regions in the Western Sierra Madre that were both unexplored (i.e., no specimens reported in any of the databases) and relatively well preserved, being very mountainous areas with few human settlements and no roads. This region was highlighted as a priority for exploration, and a call for relevant projects was issued in 2000. Figure 3 illustrates the case for the state of Durango (much of it covered in montane Sierra Madre ecosystems), which was identified as of high priority for retrospective data capture and digitization and *de novo* biodiversity explorations (Soberón et al. 2004). The red line shows the point at which CONABIO began assigning priorities for funding based on existing databases. With a delay, key data started pouring into the system.

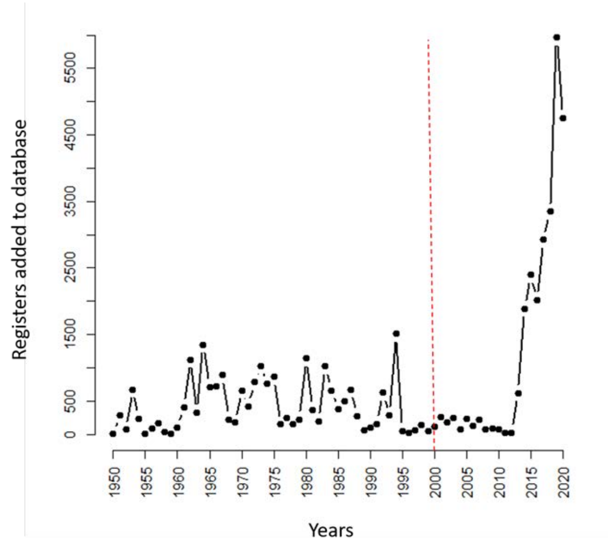


Figure 3. Number of specimens in CONABIO's databases for the state of Durango, identified as a high priority in the year 2000 (red dashed line)

*Scientific articles*

One last use of CONABIO's data that should not be forgotten is to enhance capacity for research by the Mexican biodiversity science community. This research community has taken good advantage of the massive, new, and unprecedented availability of data (Peterson et al. 2016; Rodríguez et al. 2017). This effect is illustrated in the graphs in Figure 4.

## Conclusions

The national biodiversity agency of Mexico performs a large variety of functions, including diplomatic, legislative and educational (Sarukhan 2018). However, the core of its capacities, what truly distinguishes it from other government agencies in Mexico, is its solid empirical grounding in primary data. The time, money, and human effort (the result of literally hundreds of years of biological research about Mexico, nationally and internationally) spent building a powerful, comprehensive data system provide the agency with its credibility. This credibility is one of the keystones of the process of translating from science to policy-making (Cash et al. 2003; Soberón 2004). When the scientists and negotiators of CONABIO argue in Mexico's congress, or negotiate in an international forum, they have the credibility that comes from positions solidly grounded on primary, verifiable, open data.

Moreover, the amount of research that the data made available by CONABIO has enabled is difficult to quantify. One can count number of papers published, but the number of internal reports in government agencies, dissertations, and other "gray" uses of data is impossible to quantify. Anecdotally, however, it is known that the system of CONABIO is widely used.

CONABIO was made possible by the vision of pioneers, and a very singular political environment that allowed Mexico to create a politically and economically independent organization, capable of issuing science-based opinions at a high governmental level. Political circumstances have changed, however, such that now CONABIO has been deprived of its

[4] https://sipecamdata.conabio.gob.mx/manual.

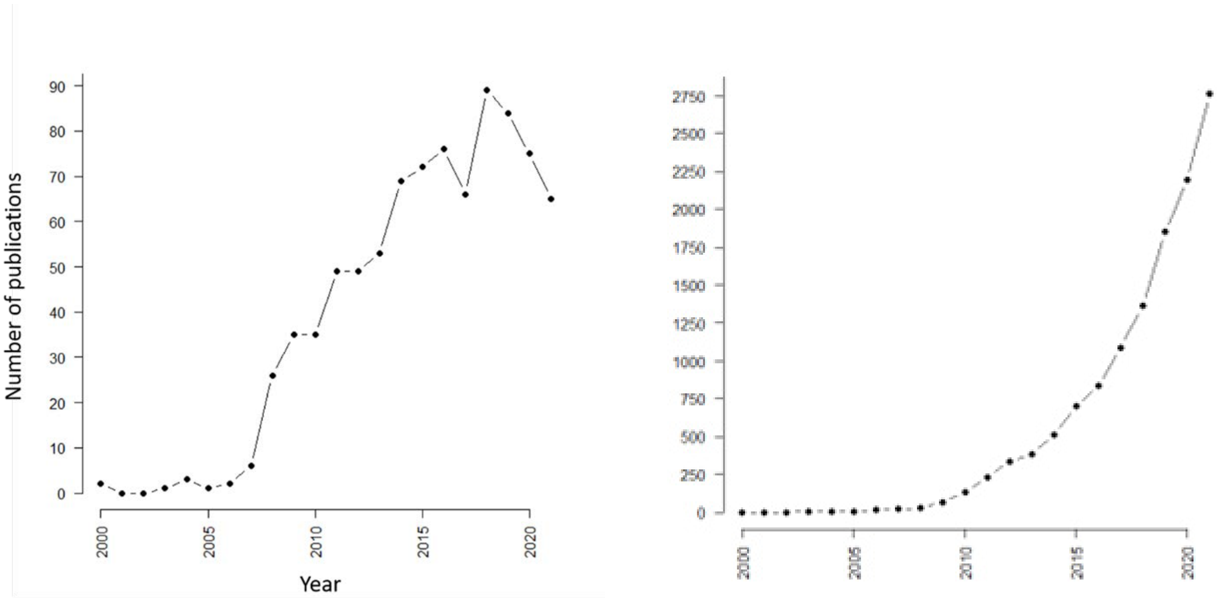


Figure 4. Impact of CONABIO on biodiversity research outputs. Numbers of publications and citations were drawn from Web of Science, using the key word CONABIO in Title or in Funding Source. Left graph, number of papers published. Right graph, citations to papers published, since 2000.

economic independence. It may be in the process of losing its political independence as well. The Costa Rican InBio has also disappeared, or collapsed (Fonseca 2015), and the Indian initiative on bioinformatics is also non-existent. Of the original biodiversity institutions that visited ERIN in 1992, only the Australian initiative survives, in the form of the Atlas of Living Australia project.

The long-term survival of any institution depends on a combination of political, economic, and social factors. CONABIO was created by the fortunate combination of a diplomatic need for Mexico to have something to present at the Earth Summit conference, and the fact that the most prominent ecologist of Mexico was also the chancellor of the national university at the time. Given its hybrid private-public design, CONABIO was able to build an impressive capacity to assemble, organize, and analyze biodiversity data. Moreover, the organization was acting as a bridge (Cash et al. 2003; Soberón 2004) between academia and decision-making in the federal government.

This combination, however, has not survived changes in the political world of Mexico. It is difficult to speculate what combination of factors could have maintained CONABIO as an independent, fully funded government agency. CONABIO'S hybrid design allowed it to maintain some of its assets (i.e., computing cluster, remote-sensing capacities, databases…) as private, thus providing some degree of permanence, but the cross-cutting multiple ministries character and CONABIO's budgetary and political independence are probably gone for good. It is to be hoped that the huge data resources of CONABIO, still openly available on-line, will remain so, via mirrors like GBIF and others, although even a multinational initiative like GBIF is vulnerable to budgetary constraints. It is now clear that if scientists want to keep primary data openly available, databases probably will need to be spread over many independent organizations, to minimize the risk of collapse due to failure of one main participant. This perhaps will protect, at least, the purpose of the data, if not the organizations as such.

## Literature Cited

Acevedo, F., E. Huerta, and C. Burgeff. 2016. Biosafety and Environmental Releases of GM Crops in Mesoamerica: Context Does Matter *in* R. Lira, A. Casas and J. Blancas, eds. Ethnobotany of Mexico. Interactions of People and Plants in Mesoamerica. Springer, New York.

Asase, A., M. Sainge, R. Radji, U. Omokafe, and T. Peterson. 2020. A new model for efficient, need-driven progress in generating primary biodiversity information resources. Applications in Plant Sciences 8:e11318.

Balmford, A., P. Crane, A. Dobson, R. Green, E., and G. Mace. 2005. The 2010 challenge: data availability, information needs and extraterrestrial insights. Philosophical Transactions of the Royal Society B 360:221-228.

Beaman, R., and N. Cellinese. 2012. Mass digitization of scientific collections: New opportunities to transform the use of biological specimens and underwrite biodiversity science. ZooKeys 209:7-17.

Belbin, L., Wallis, E. Hobern, D., Zerger, A. 2021. The Atlas of Living Australia: history, current state and future directions. Biodiversity Data Journal 9:e65023.

Berners-Lee, T. 1992. The world-wide web. Computer networks and ISDN systems 25:454-459.

Bibby, C. J. e. a. 1992. Putting Biodiversity on the Map: Priority Areas for Global Conservation. International Council for Bird Preservation, Cambridge, UK.

Booth, T. 2018. Why understanding the pioneering and continuing contributions of BIOCLIM to species distribution modelling is important. Austral Ecology 43:852-860.

Brooks, T. M., R. Mittermeier, G. da Fonseca, J. Gerlach, M. Hoffmann, J. Lamoreux, C. Mittermeier, J. D. Pilgrim, and A. S. Rodrigues. 2006. Global biodiversity conservation priorities. Science 313:58-61.

Busby, J. R., C. R. Margules, and M. P. Austin. 1991. BIOCLIM - A bioclimate analysis and prediction system. Pp. 64 *in* C. R. Margules and M. P. Austin, eds. Nature Conservation. Cost-effective Biological Surveys and Data Analysis, Canberra, Australia.

Butchart, S. H. M., M. Walpole, B. Collen, A. van Strien, J. P. W. Scharlemann, R. E. A. Almond, J. E. M. Baillie, B. Bomhard, C. Brown, J. Bruno, K. E. Carpenter, G. M. Carr, J. Chanson, A. M. Chenery, J. Csirke, N. C. Davidson, F. Dentener, M. Foster, A. Galli, J. N. Galloway, P. Genovesi, R. D. Gregory, M. Hockings, V. Kapos, J.-F. Lamarque, F. Leverington, J. Loh, M. A. McGeoch, L. McRae, A. Minasyan, M. H. Morcillo, T. E. E. Oldfield, D. Pauly, S. Quader, C. Revenga, J. R. Sauer, B. Skolnik, D. Spear, D. Stanwell-Smith, S. N. Stuart, A. Symes, M. Tierney, T. D. Tyrrell, J.-C. Vié, and R. Watson. 2010. Global Biodiversity: Indicators of Recent Declines. Science 328:1164-1168.

Canhos, D. 2017. Data Management Plan: Brazil's Virtual Herbarium. Research Ideas and Outcomes 3:e14675.

Cash, D. W., W. C. Clark, F. Alcock, N. M. Dickson, N. Eckley, D. H. Guston, J. Jager, and R. B. Mitchell. 2003. Science and technology for sustainable development special feature: knowledge systems for sustainable development. Proceedings of the National Academy of Sciences USA 100:8086-8091.

Cerdeira-Estrada, S., and G. López-Saldaña. 2008. Automatic processing of near-real time operational MODIS Ocean Products applied to Mexico seas monitoring. 2008 5th International Conference on Electrical Engineering, Computing Science and Automatic Control, 545-549.

Chapman, A. 1991. The role of specimen-backed information in environmental decision making -- The Australian experience. Pp. 1. Symposium at the Australian National Botanical Gardens, Canberra, Australia.

Chapman, A. 2001. Biodiversity informatics, Biota/FAPESP and the future: a personal view. Biota Neotropica 1:1-9.

Chapman, A. 2005. Principles and Methods of Data Cleaning-Primary Species and Species-Occurrence Data. Global Biodiversity Information Facility, Copenhagen.

Coetzer, W. 2012. A new era for specimen databases and biodiversity information management in South Africa. Biodiversity Informatics 8:1-11.

CONABIO. 2011. Sistema de alerta temprana de incendios forestales en México y Centroamérica. National Commission on Biodiversity, Mexico

CONABIO. 2012. CONABIO: Two Decades of History, 1992-2012. Pp. 1-36 *in*. ed. Comision Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico D. F., Mexico.

Dallwitz, M. J. 1993. DELTA and INTKEY. Pp. 287-296 *in* R. Fortuner, ed. Advances in computer methods for systematic biology: artificial intelligence, databases, computer vision. Johns Hopkins University Press, Baltimore, USA.

Dirzo, R., O. López, P. Maeda, R. Mejía, M. Munguía-Carrara, E. Robredo, and M. Schmidt. 2021. Manual de monitoreo: Sitios Permanentes de Calibración y Monitoreo de la Biodiversidad. Pp. 85. CONABIO, Mexico City.

Edwards, J. L., M. Lane, and E. Nielsen. 2002. Interoperability of biodiversity databases: biodiversity information on every desktop. Science 289 2312.

Ehrlich, P. 1995. The scale of the human enterprise and biodiversity loss. Pp. 233 *in* J. H. Lawton and R. M. May, eds. Extinction Rates. Oxford University Press, Oxford.

Fonseca, P. 2015. A major center of biodiversity research crumbles. Scientific American e The Sciences Section. http://www. scientificamerican. com/article/a-major-center-of-biodiversityresearch-crumbles

Gadgil, M., F. Berkes, and C. Folke. 1993. Indigenous knowledge for biodiversity conservation. AMBIO 22:151-156.

García-Alaniz, N., M. Equihua, O. Pérez-Maqueo, J. E. Benítez, P. Maeda, F. P. Urrutia, J. J. F. Martínez, S. A. V. Gaytán, and M. Schmidt. 2017. The Mexican national biodiversity and ecosystem degradation monitoring system. Current Opinion in Environmental Sustainability 26:62-68.

Gebhardt, S., T. Wehrmann, M. A. M. Ruiz, P. Maeda, J. Bishop, M. Schramm, R. Kopeinig, O. Cartus, J. Kellndorfer, and R. Ressl. 2014. MAD-MEX: Automatic wall-to-wall land cover monitoring for the Mexican REDD-MRV program using all Landsat data. Remote Sensing 6:3923-3943.

Goettsch, B., T. Urquiza-Haas, P. Koleff, F. Acevedo, A. Aguilar-Melendez, and e. al. 2021. Extinction risk of Mesoamerican crop wild relatives. Plants People Planet 3:775-795.

Gonzalez, C., E. Mora, and M. Munguia. 2014. Modeling Ecological Integrity with Bayesian belief networks. Pp. 1-3 *in* CONABIO, ed. CONABIO, Mexico.

Graves, G. 2000. Costs and benefits of Web access to museum data. Trends in Ecology and Evolution 15:374.

Groves, C., Klein, M. Breden, T. 1995. Natural heritage programs: public-private partnerships for biodiversity conservation. Wildlife Society Bulletin 23:784-790.

Guisan, A., R. Tingley, J. B. Baumgartner, I. Naujokaitis-Lewis, P. R. Sutcliffe, A. I. T. Tulloch, T. J. Regan, L. Brotons, E.

McDonald-Madden, C. Mantyka-Pringle, T. G. Martin, J. R. Rhodes, R. Maggini, S. A. Setterfield, J. Elith, M. W. Schwartz, B. A. Wintle, O. Broennimann, M. Austin, S. Ferrier, M. R. Kearney, H. P. Possingham, and Y. M. Buckley. 2013. Predicting species distributions for conservation decisions. Ecology Letters 16:1424-1435.

Hardin, G. 1968. The tragedy of the commons. Science 162:1243-1248.

Hernández, J., H. Sánchez, A. Bello, and G. González. 2007. Preventive programme against the cactus moth Cactoblastis cactorum in Mexico. Pp. 345-350 *in* M. J. B. Vreysen, A. S. Robinson and J. Hendricks, eds. Area-Wide Control of Insect Pests. IAEA, Vienna, Austria.

Hortal, J., F. de Bello, J. A. F. Diniz-Filho, T. M. Lewinsohn, J. M. Lobo, and R. J. Ladle. 2015. Seven shortfalls that beset large-scale knowledge of biodiversity. Annual Review of Ecology, Evolution, and Systematics 46:523-549.

Hruby, F., S. Melamed, R. Ressl, and D. Stanley. 2016. Mosaicking Mexico. The big-picture of big-data. Pp. 407-411. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, Prague, Czech Republic.

Johnson, N. F. 2007. Biodiversity Informatics. Annual Review of Entomology 52:421-438.

Kaye, P., S. Noble, and W. Slater. 1997. Environmental Information for Intelligent Decisions. Pp. 245-258. Intelligent Environments. Elsevier.

Koester, V. 2002. The five global biodiversity-related conventions: a stocktaking. Review of European Community & International Environment Law 11:96-103.

Laihonen, P. K., R. Salo, J. 2004. The biodiversity information clearing-house mechanism as a global effort. Environmental Science and Policy 7:99-108.

Lane, M., and J. Edwards. 2007. The Global Biodiversity Information Facility. Pp. 1-4 *in* C. Humphries, ed. Systematics Association Special Volume. CRC Press, Boca Raton, CA.

MacLaurin, J., and K. Sterelny. 2008. What is Biodiversity? The University of Chicago Press, Chicago.

McNeely, J., K. R. Miller, W. V. Reid, R. Mittermeier, and T. Werner. 1990. Conserving the World's Biological Diversity. The World Bank, Washington, D. C.

Medellín, C., and L. Corrales. 2019. Sistemas de monitoreo forestal en México. Pp. 84. Serie Técnica. Boletín Técnico. CATIE, Turrialba, Costa Rica.

Nelson, G., and S. Ellis. 2019. The history and impact of digitization and digital data mobilization on biodiversity research. Philosophical Transactions of the Royal Society B 374:20170391.

Nix, H. A. 1986. A biogeographic analysis of Australian elapid snakes *in* R. Longmore, ed. Atlas of Elapid Snakes of Australia. Australian Government Publishing Service, Canberra.

Noss, R. F. 1990. Indicators for monitoring biodiversity: A hierarchical approach. Conservation Biology 4 355-364.

Ostrom, E., J. Burger, C. B. Field, R. B. Norgaard, and D. Policansky. 1999. Revisiting the commons: local lessons, global challenges. Science 284:278-282.

Peterson, A. T., S. Knapp, R. Guralnick, J. Soberón, and M. T. Holder. 2010. The big questions for biodiversity informatics. Systematics and Biodiversity 8:159-168.

Peterson, A. T., A. G. Navarro-Sigüenza, and A. Gordillo-Martínez. 2016. The development of ornithology in Mexico and the importance of access to scientific information. Archives of natural history 43:294-304.

Peterson, A. T., and J. Soberón. 2018. Essential biodiversity variables are not global. Biodiversity and Conservation 27:1277-1288.

Pressey, R. L., C. J. Humphries, C. R. Margules, R. I. Vane-Wright, and P. H. Williams. 1993. Beyond opportunism: key principles for systematic reserve selection. Trends in Ecology & Evolution 8 124-128.

Reed, G. 2017. The Clearing-House Mechanism: An Effective Tool for Implementing the Convention on Biological Diversity? Pp. 115-126. Governing Global Biodiversity. Routledge.

Reid, W. L., S., Meyer, C. Gamez, R. Sittenfeld, A., Janzen, D. Gollin, D. Juma, C. 1996. Biodiversity Prospecting. Pp. 142-173 *in* W. Reid, ed. Biodiversity Prospecting. Using Genetic Resources for Sustainable Development. World Resources Institute, Washington, DC.

Reid, W. V. 1998. Biodiversity Hotspots. Trends in Ecology & Evolution 13:275-280.

Ressl, R., G. Lopez, I. Cruz, R. Colditz, M. Schmidt, S. Ressl, and R. Jimenez. 2009. Operational active fire mapping and burnt area identification applicable to Mexican Nature Protection Areas using MODIS and NOAA-AVHRR direct readout data. Remote Sensing of Environment 113:1113-1126.

Rodríguez, P., F. Villalobos, A. Sánchez-Barradas, and M. Correa-Cano. 2017. La macroecología en México: historia, avances y perspectivas. Revista mexicana de biodiversidad 88:52-64.

Sandlund, O. T. 1991. Costa Rica's INBio: towards sustainable use of natural biodiversity. Pp. 1-25 *in* N. I. f. Naturforskning, ed, Trondheim.

Sarkar, S. 2002. Defining biodiversity; assessing biodiversity. The Monist 85:131-155.

Sarukhan, J. 2018. CONABIO, 25 years of evolution. Pp. 160. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico.

Sarukhán, J., and R. Dirzo. 1992. Mexico ante los Retos de la Biodiversidad. Pp. 343. Universidad Autonoma de Chapingo, Mexico DF.

Sarukhán, J., and R. Jiménez. 2016. Generating intelligence for decision making and sustainable use of natural capital in Mexico. Current opinion in environmental sustainability 19:153-159.

Sarukhán, J., T. Urquiza-Haas, P. Koleff, J. Carabias, R. Dirzo, E. Ezcurra, S. Cerdeira-Estrada, and S. Jorge. 2014. Strategic actions to value, conserve, and restore the natural capital of megadiversity countries: the case of Mexico. BioScience 65:164-173.

Scott, J. M. 1993. Gap analysis: a geographic approach to protection of biological diversity. Wildlife Monographs 123:1-41.

Siebert, S., and G. Smith. 2004. Lessons learned from the SABONET Project while building capacity to document the botanical diversity of southern Africa. Taxon 53:119-126.

Simonson, S. E., T. Stolhgren, L. Tyler, W. P. Gregg, M. Rachel, and J. Garrett. Lynn. 2005. Preliminary assessment of the potential impacts and risks of the invasive cactus moth, *Cactoblastis cactorum* Berg, in the U.S. and Mexico International Atomic Energy Agency, Vienna, Austria.

Sittenfeld, A., and R.Villers. 1993. Exploring and preserving biodiversity in the tropics: The Costa Rican case. Biotechnology 4 280-285.

Soberón, J. 2004. Translating life's diversity: can scientists and policymakers learn to communicate better? Environment 46:10-20.

Soberón, J., P. Dávila, and J. Golubov. 2004. Targeting sites for biological collections *in* R. R. Smith, J. B. Dickie, S. Linington, H. Pritchard and R. Probert, eds. Seed storage: turning science into practice. Royal Botanic Gardens, Kew, UK.

Soberón, J., J. Golubov, and J. Sarukhan. 2001. The importance of *Opuntia* in Mexico and routes of invasion and impact of *Cactoblastis cactorum* (Lepidoptera: Pyralidae). Florida Entomologist:486-492.

Soberón, J., E. Huerta, and L. Arriaga. 2002. The use of databases to assess the risk of gene flow: the case of Mexico. Pp. 61-67 *in* C. R. Roseland, ed. LMOs and the Environment. Organization for Economic Cooperation and Development, Paris.

Soberón, J., R. Jiménez, P. Koleff, and J. Golubov. 2010. La informática sobre la biodiversidad: datos redes y conocimiento. Pp. 354 *in* V. M. Toledo, ed. La Biodiversidad de México. El Fondo de Cultura Económica, México D. F.

Soberón, J., and P. Koleff. 2000. The national biodiversity information system of Mexico. Pp. 625 *in* P. Raven, ed. Nature and Human Society: Proceedings of the 1997 Forum on Biodiversity. National Academies Press, Washington, D. C.

Soberón, J., and A. T. Peterson. 2004. Biodiversity informatics: managing and applying primary biodiversity data. Philosophical Transactions of the Royal Society B 35:689-698.

Sousa-Baena, M. S., R. García, L. Couto, and A. T. Peterson. 2014. Completeness of digital accessible knowledge of the plants of Brazil and priorities for survey and inventory. Diversity and Distributions 20:369-381.

Steffen, W. W. B., L. Deutsch, O. Gaffney, C. Ludwig. 2015. The trajectory of the Anthropocene: The Great Acceleration. The Anthropocene Review 2:81-98.

Stein, B., L. Kutner, and J. S. Adams. 2000. Precious Heritage. The Status of Biodiversity in the United States. The Nature Conservancy and Oxford University Press, New York.

Stephens, C., R. Sierra-Alcocer, C. González-Salazar, J. Barrios, J. C. Salazar, E. Robredo, and E. del Callejo. 2017. SPECIES: A platform for the exploration of ecological data. Ecology and Evolution 9:1638-1653.

Tangley, L. 1990. Cataloging Costa Rica's diversity. BioScience 40:633-636.

Tegelberg, R., T. Mononen, and H. Saarenmaa. 2014. High-performance digitization of natural history collections: automated imaging lines for herbarium and insect specimens. Taxon 63:1307-1313.

ten Kate, K. 1999. Legal Aspects of Regulating Access to Genetic Resources and Benefit-sharing: the Convention on Biological DIversity, National and Regional Laws and Material Transfer Agreements. Earthscan, London, UK.

Toledo, V. M. 2001. Indigenous peoples, and biodiversity. Pp. 1181-1203 *in* S. Levin, ed. Encyclopedia of Biodiversity. Academic Press.

Valderrama, L., C. Troche, M. Rodriguez, D. Marquez, B. Vázquez, S. Velázquez, A. Vázquez, M. Cruz, and R. Ressl. 2014. Evaluation of mangrove cover changes in Mexico during the 1970–2005 period. Wetlands 34:747-758.

Wegier, A., A. Piñeyro-Nelson, J. Alarcón, A. Gálvez-Mariscal, E. R. Álvarez-Buylla, and D. Piñero. 2011. Recent long-distance transgene flow into wild populations conforms to historical patterns of gene flow in cotton (*Gossypium hirsutum*) at its centre of origin. Molecular Ecology 20:4182-4194.

Wieczorek, J., D. Bloom, R. Guralnick, S. Blum, M. Doring, R. Giovanni, T. Robertson, and D. Vieglais. 2012. Darwin Core: an evolving community-developed biodiversity data standard. PLoS One 7:e29715.

Wieczorek, J., Q. Guo, and R. J. Hijmans. 2004. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. International Journal of Geographic Information Science 18:745-767.

Zimmermann, H., V. Moran, and J. Hoffmann. 2000. The renowned cactus moth, *Cactoblastis cactorum*: its natural history and threat to native *Opuntia* floras in Mexico and the United States of America. Diversity and Distributions 6:259-269.