Survey of Non-Native Species Databases

April 2006

Prepared for:
The State of the Nation's Ecosystems Project
The H. John Heinz III Center for Science, Economics and the Environment
1001 Pennsylvania Ave, NW
Suite 735 South
Washington, DC 20004

Prepared by:
The Natural Resource Ecology Laboratory
Colorado State University
Natural and Environmental Sciences Building
Fort Collins, CO 80523

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

Non-native species adversely affect many United States ecosystems, threatening biodiversity, ecosystem functioning, human health, agriculture, and the economy. There is an increasing need for early detection, rapid response and long-term monitoring of these invasive species. Many governmental, academic and private institutions currently collect non-native species data, but to date there has been only limited coordination, data sharing, and merging of these datasets.

The Natural Resource Ecology Laboratory of Colorado State University, at the request of The Heinz Center, conducted this review of existing non-native species data in the United States from state, multi-state region, national, and global scales. The goal of this effort is to provide a better understanding of what data currently exist for non-native species and to determine where data gaps exist (taxonomically, spatially, and temporally) to guide future survey, research, and spatial predictive modelling efforts. Metadata on non-native species databases were collected through an on-line survey to provide additional information on data type, data quality, and data availability. The actual database containing this information can be accessed at www.heinzctr.org/ecosystems.

Two hundred fifty-two databases were identified through this process and assessed for their taxonomic, geographic, and temporal completeness, as well as their data quality. The assessment revealed a number of key findings. Most non-native species database contacts were affiliated with federal agencies or universities. Very few were affiliated with private institutions. Nearly all databases are currently available to the public or will be in the future, but some have conditions for access. Many but not all of the publicly funded databases are available to the public, while all of the privately funded were publicly available. Researchers, academics, field technicians and taxonomists collected most non-native species data.

The databases vary greatly in terms of spatial scale. Most ecosystem types are represented in a hundred or more databases; however, tundra, deserts, and coasts and oceans are somewhat less well represented. Almost two-thirds of the databases have geo-referenced data. Many of the databases cover only one taxon; very few cover all taxa. Most databases have information on non-native species locations, but very few have data on control efforts.

More than half of the databases are not updated or updated only irregularly. However, almost half are updated at least annually. Sampling design varied widely across the databases, and many contained data gathered using more than one design. It is hoped that this initial assessment of these databases will lead to other, more comprehensive analysis and cooperation among the various database managers.

Over a period of approximately one year (2004-2005), the Heinz Center's Non-native Species Task Group has also developed a suite of non-native species indicators to report on four major taxonomic groups of non-native species: plants, vertebrates, invertebrates and pathogens. The information and networks described in this database survey should facilitate the population of those indicators with data from the broad range of U.S. non-native species databases.

Preface: Context for the Survey

Over a period of approximately one year (2004-2005), the Heinz Center's Non-native Species Task Group (Appendix A) developed a suite of non-native species indicators to report on four major taxonomic groups of non-native species: plants, vertebrates, invertebrates and pathogens. The Task Group also developed a hierarchy of preference for the non-native species indicators. This hierarchy acknowledged the need for and importance of indicators that report on the pattern and distribution of non-native species (the data that are primarily available at the present time) but stressed the importance of collecting and reporting national-scale information on the impacts of non-native species.

While these "impact" indicators may not be fully populated with data right away, their establishment should encourage agencies, institutions, and individuals to fund and design sampling schemes and monitoring programs to collect long-term data on the effects on non-native species on ecosystems, the economy, and human health. For the purposes of this report and its indicators, *non-native species* include plants, vertebrates, invertebrates, and pathogenic organisms that affect plants, animals, and humans; and are defined as organisms that are not indigenous to the ecosystem to which they were introduced and which are capable of surviving and reproducing without human intervention. This definition is consistent with Executive Order 13112 and the National Invasive Species Council Management Plan (NISC, 2001).

The initial suite of indicators developed by the Task Group was sent out for peer-review and was presented at a series of professional conferences and agency meetings. Many helpful comments on the indicators provided guidance for revising the first set of indicators, and the Heinz Center staff continued to consult with individual experts and with the Non-native Species Task Group. A revised set of seven indicators was established. Full indicator descriptions are out for peer-review with the draft final report on the non-native species indicators.

Approach

The Natural Resource Ecology Laboratory of Colorado State University, at the request of The Heinz Center, conducted this review of existing non-native species data in the United States from state, multi-state region, national, and global scales. The goal has been to provide a better understanding of what data currently exist for non-native species and to determine where data gaps exist (taxonomically, spatially, and temporally) to guide future survey, research, and spatial predictive modelling efforts. Metadata on non-native species databases were collected through an on-line survey to provide additional information on data type, data quality, and data availability. We have provided a framework that we hope will increase collaboration among various research organizations to begin to more efficiently tackle the non-native species problem.

This review has also provided the Heinz Center with a list of datasets that could potentially populate their set of non-native species indicators. These indicators describe the overall condition of U.S. ecosystems with respect to non-native species, both by taxa and by specific ecosystem types and may be included in the Center's *State of the Nation's Ecosystems* report 2007 report.

Introduction

Invasion by non-native species has adversely affected many ecosystems in the United States, threatening biodiversity, ecosystem functioning, human health, and the economy (Vitousek et al. 1997; Wilcove et al. 1998; Mack et al. 2000; Pimentel et al. 2000). Because organisms continue to be introduced from other countries via trade and transportation, there is a growing need for early detection and rapid response to new invaders (Vitousek et al. 1997). Synthesis of existing data on non-native species abundance and distributions is an important first step. New data can be then added to existing data to provide the most up-to-date and accurate information on non-native species locations, moving us from a reactive to a proactive control strategy (Ricciardi et al. 2000). However, little is currently known regarding what data exist on non-native species, and there have been few efforts to improve collaboration and data synergy among governmental agencies, non-governmental organizations, industry, academic researchers, and other non-native species networks (Crosier and Stohlgren 2004). Therefore, a primary goal of non-native species research should be to facilitate non-native species data sharing among these different research groups.

Data for non-native species are collected using various research methods, spatial and temporal scales, and data quality procedures. Merging these many disparate datasets would have several benefits. First, data synergy would help improve area species lists. A study by Crosier and Stohlgren (2004) found that merging disparate datasets within Colorado was a cost-effective way of improving existing species lists in the state. Second, having a mix of presence/absence and distribution data for a particular species in a study area would enable production of spatially predictive models on current and potential invasions, and identify data gaps to guide future research efforts (Crosier and Stohlgren 2004). Quantifying species distributions and richness patterns should lead to the development of better research questions.

Third, data synergy could aid land managers in controlling and monitoring new invasions by providing watch lists to manage lands adjacent to currently invaded areas. By providing watch lists, an invasion can be stopped before expensive control methods have to be implemented (Rejmànek and Pitcairn 2002). Finally, many organizations are limited by resource constraints (i.e., time, money, personnel). Therefore, combining available datasets from an area would capitalize on these limited resources with minimal additional cost and effort.

Although some steps have already been taken to facilitate data sharing, those efforts are either still in their infancy or have not yet been successful over large scales (Ricciardi et al. 2000; Simpson 2004; Stohlgren et al. 2006). An example of such an effort is the Global Invasive Species Information Network (GISIN). This project is developing a registry of all on-line non-native species databases worldwide to provide an outlet for easily obtaining non-native species information (Simpson 2004). The GISIN project is still in its preliminary stages, but this effort will be a valuable asset in providing a global network for invasive species data sharing (Simpson 2004).

Although the internet provides a good way to store, analyze, and rapidly distribute non-native species data, there are many additional electronic data sources (e.g., spreadsheets, databases, GIS) that are not yet available online. Some have not been made available to the public or are only used locally. In other cases, databases owners do not have the technical capabilities or do not perceive the need to put their databases online. Therefore, there is also a need to collect and evaluate metadata in each of these databases to determine the type of data they house and its quality and public availability.

Database Review

The purpose of the database review was to create a comprehensive list of non-native species databases in the United States. Several search strategies were implemented to find existing non-native species data, including gathering pre-existing lists of databases, conducting a comprehensive web search, and conducting a literature review of related publications. Websites were included only if they contain information on non-natives directly related to this study at an appropriate scale. From this preliminary search, 192 databases were identified.

Pre-existing Database Lists

The largest source for the initial non-native species database list was the pre-existing Global Invasive Species Information Network's (GISIN) list of Invasive and Alien Species (IAS) Online Databases (Sellers et al. 2004). Ninety-two of the 192 (48%) initially identified databases were taken from this list. The Heinz Center also provided an additional list of 28 databases found through previous work of their State of the Nation's Ecosystems project efforts. Other lists were found through our literature review, as described in the literature review section.

Comprehensive Web Search

The GISIN database list helped to initiate the comprehensive web search. Many of the websites listed for the GISIN contained links to related sites. Although many of these links contained duplicates of databases listed in the original GISIN list, some new sites were found through this line of inquiry. Further web searches were performed to provide additions and fill apparent gaps to the pre-existing database lists. For instance, databases on non-native vertebrates were scarce, so specific searches were designed to locate these types of databases. Web searching was considered complete when the related links portion of each website no longer contained new databases. It is reasonable to assume that there are additional on-line databases available that were not found through our efforts. However, the cost of time spent searching for these less prominent databases exceeded the apparent gain in finding new ones.

Literature Review

Following the web search, a library journal search was conducted to find publications related to non-native species databases. Literature searches located one article and two conference proceedings that provided lists of multiple databases (Jacono and Boydstun 1997; Ridgway et al. 1998; Ricciardi et al. 2000). Only one of the databases listed was not a duplicate of those found through the GISIN list and comprehensive web search. We also found 13 additional articles related to individual databases. Of these 13, five were new to our list (Binggeli 1996; Despain et al. 2001; Kimberling 2004; Semmens et al. 2004; Unmack and Fagan 2004).

Undoubtedly, there are non-native species databases in existence in addition to those found through our efforts. Many of these databases may be off-line, unpublished, or not available to the general public (e.g., some lists of crop pests from APHIS are sensitive and unavailable to the public). To locate some of these other sources, we contacted approximately 1,500 experts in the field of non-native species science. We added any relevant data on non-native species found through these queries.

Compiling Non-Native Species Contacts

Expert contacts were compiled from conference rosters, agency lists, and the aforementioned literature searches. Agency lists were provided for the US Geological Survey, the Fish and Wildlife Service, and US Forest Service Wilderness Areas. A contact list was also obtained from the National Science Foundation that identified researchers receiving grants for non-native species research. We attempted to acquire lists for the National Park Service, other US Forest Service land managers, and the Bureau of Land Management, but these agencies were unable to provide such lists. However, many contacts from each of these agencies were obtained through roster lists and literature searches.

Survey Development

To determine the type, quality, and availability of non-native species data in the United States, an on-line survey was created to collect metadata on each database that had been identified (Appendix B). Metadata included information related to geographic scope, data collection methods, taxonomic focus, spatial and temporal scale, and data quality.

Two survey request letters were created to send out to our list of contacts. The first letter was to persons affiliated with a specific database (Appendix C), and the second letter was sent to persons known to be conducting research in non-native species (Appendix C). The Heinz Center also sent out a general request letter to the Ecolog and Alien list serves. A specific contact list from these list serves was not available, so these contacts were not added to the final contact list.

Each survey request letter directed the contact to an online survey linked to the National Institute of Invasive Species Science web site (see http://www.niiss.org). From this link, each contact was given a prompt to provide a unique username and password. This guaranteed that only persons within our contact list could take the survey and that all participants' responses could be tracked. Other persons wanting to take the survey had to send an email requesting access. As survey participants took the survey, their responses were automatically uploaded into a database linked to each survey question (Appendix B). The database design was organized in a way that could easily be queried across all fields relevant to our analysis.

Database Analysis

Once the survey was closed, analysis began on the generated database. Contact information (i.e., name, email, phone, affiliation) was collected for each individual taking the survey. Using the affiliation response provided, we classified each contact within one of the following affiliation categories: Animal and Plant Health Inspection Service (APHIS); Agricultural Research Service (ARS); Bureau of Land Management (BLM); Centers for Disease Control (CDC); Cooperative State Research, Education and Extension Service (CSREES); Environmental Protection Agency (EPA); Fish and Wildlife Service (FWS); Military; Museum/Herbarium; National Aeronautics and Space Administration (NASA); National Oceanic and Atmospheric Administration (NOAA); Non-Profit; National Park Service (NPS); Natural Resource Conservation Service (NRCS); Other; Private; State Agencies; University; United

States Forest Service (USFS); or United States Geological Survey (USGS). This made it easy to query each contact by affiliation, and we were able to generate a count of the number of databases provided by each of these research organizations.

A similar approach was taken with respect to the agency funding each database. Using the response given in the survey, we classified each database as being funded by public, private, both (i.e., public and private), or none (i.e., no funding available). Again, this made it easy to query this field and provide a count of the number of databases funded publicly or privately.

Using other information gained from the survey, we analyzed databases within each of the following additional categories: database type, data completeness (taxonomic, geographic, and temporal), and data quality. This process allowed us to assess databases on the extent to which the information they contain captured the essential components of the systems to which they relate.

Database Type

There are many websites and databases dedicated to non-native species, but not all of these websites and databases contain actual data on non-native species useful for specific research objectives or for the Heinz Center's non-native species indicators. To determine which databases contained applicable information, databases were assigned one or more of the following classifications:

- General Species List—Database contains a list of species including native, non-native, and invasive species in a specified study area. Not specific to non-native species.
- Non-Native Species List—Database contains a list of non-native species in a specified study area (e.g., National Park, county, state).
- Species Distribution—Database contains maps or any georeferenced data for non-native species distributions.
- Track/Control of Species—Database contains information on the control of a non-native species. All data related to species presence or distribution is directly related to control and restoration efforts.
- Distributed—Database queries other databases; data-dependent.
- Species Information—Database contains general life history information about specific non-native species.
- Bibliographic—Database of references related to non-native species.
- Other—Database did not fit into any of the above categories.

Non-native species lists, species distribution, general species lists, track/control of species, and distributive databases were sought specifically for their ability to provide non-native species location information and to populate the Heinz Center non-native species indicators. Species information and bibliographic databases were included when found, but were not specifically sought.

Data Completeness

To provide information on data completeness, we examined survey responses to classify each database into low, medium, and high categories for taxonomic, geographic, and temporal completeness. Although these descriptions are subjective, they were necessary to provide a means to query and classify the databases according to their completeness in these various fields.

Taxonomic Completeness

Survey participants initially categorized their databases by taxonomic focus. Taxonomic classifications included plants, fungi, vertebrates (fish, amphibians, reptiles, birds, mammals), invertebrates (insects, echinoderms, tunicates, crustaceans, arachnids, centipedes and millipedes, annelids, molluscs, cnidarians, sponges), and pathogens (fungi, bacteria, virus, insects). We then examined taxa records for each database to determine taxonomic completeness based on the proportion of taxa in the study area that were captured by the database.

- High taxonomic completeness was defined as a database covering all taxa or many biological groups.
- Medium taxonomic completeness was defined as a database covering more than one, but not all taxa in a biological group, or few biological groups (e.g., plants, birds, fishes).
- Low taxonomic completeness was defined as a database covering only one taxon.

Geographic Completeness

A full assessment of the impact of non-native species on a system requires that the sampling process adequately capture the study area's spatial variability. In light of this, survey participants classified their databases by spatial extent (i.e., smaller than county, county, state, multi-state, national, global). For databases that covered an area smaller than the national scale, the states the data covered were specified. We then used a Geographic Information System (GIS) to locate gaps in the spatial coverage of non-native species databases that we had identified. This allowed us to categorize states into low, medium, and high levels of geographic coverage by state.

To classify the geographic completeness of each database, we looked at survey responses related to how well the study area had been sampled, if the study crossed all environmental gradients within the study area, and the sampling design that was used in data collection for each database (i.e., complete random, stratified random, complete systematic, random systematic, opportunistic). Each database was classified according to the following definitions:

- High geographic completeness was defined as the entire study area well surveyed, using either complete random, stratified random, complete systematic, or random systematic sampling designs; few data gaps existed within the study area; and data were collected across all major environmental gradients.
- Medium geographic completeness was defined as having some but not all of the attributes of "high geographic completeness
- Low geographic completeness was defined as the study area surveyed opportunistically (e.g., easily accessible roads and trails); many data gaps existed within the study area; and data were not collected across major environmental gradients within the study area.

Temporal Completeness

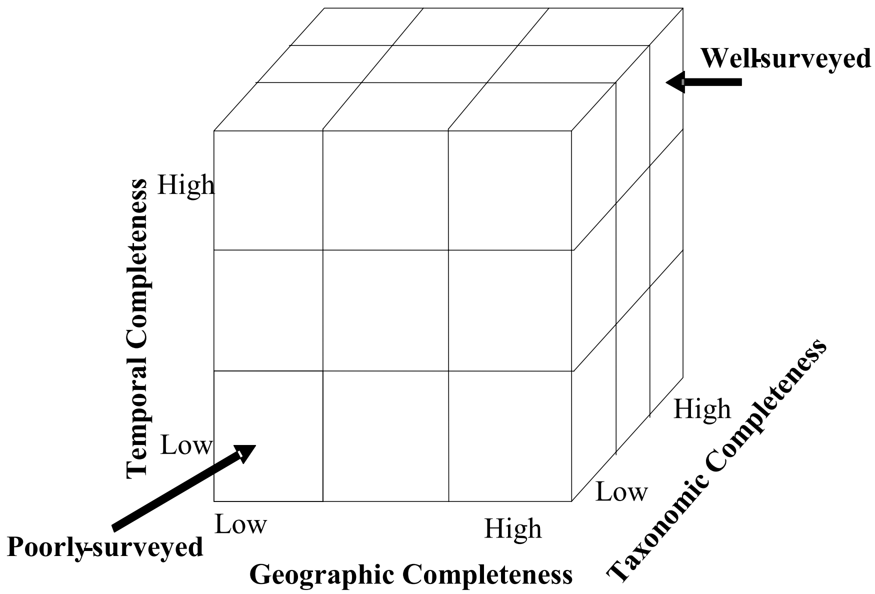


Figure 1: Model demonstrating how categories of low, medium, and high temporal, geographic, and taxonomic completeness can give insight to how well a site, area, or region has been surveyed.

Long-term studies are essential in determining the degree of establishment of non-native species populations in a given environment and changes in establishment and populations through time. We assessed the temporal completeness of databases using information on the time over which data were collected, and the update frequency of data in the database.

- High temporal completeness was defined as data collected continuously for more than 10 years; data collection completed or ongoing.
- Medium temporal completeness was defined as data collected for more than one year but spanning a time frame of ten or less years or data collected for five or less years with data collection still ongoing.
- Low temporal completeness was defined as data collected at one point in time or data collected for five or less years with data collection not ongoing.

Using this approach, we were able to identify taxonomic completeness along with spatial and temporal completeness of non-native species knowledge in the United States. A database could have high taxonomic completeness (nearly all taxa surveyed), high geographic completeness (the study area extremely well surveyed), and high temporal completeness (many surveys over many years). Conversely, a database could have poor completeness taxonomically, geographically, or temporally. Thus, there were 27 possible levels of completeness for which each database was classified (Fig. 1).

Data Quality

We assessed data quality using information obtained from the survey related to the skill level of those who participated in data collection, the presence of a quality assurance/quality control procedure, and the description of this procedure.

- High data quality was defined as data that have been put through a standardized, rigorous quality assurance/quality control process; data were collected by experts in the field (e.g., data collected by graduate student, researcher, taxonomist).
- Medium data quality was defined as data that have had some quality assurance/quality control, but not rigorous and standardized; data were collected by people with some knowledge, but were not experts in the field (i.e., data collected by undergraduate student, field technician, land manager).
- Low data quality was defined as data that were never subjected to a quality assurance/quality control process; data were not collected by experts in the field (e.g., data collected by K-12 student, K-12 educator, naturalist/hobbyist).

Other Database Classifications

Other database analyses were conducted solely on the survey participant's choices from the survey questions. This included information related to ecosystem type, data availability, data collection methods, georeferenced data, and published data.

Ecosystem types, as described in the Heinz Center's State of the Nation's Ecosystems report, were provided in the survey for participants to choose (The Heinz Center 2002). These types included coasts and oceans, farmlands, forests, freshwaters, grasslands and shrublands, and urban and suburban. Additional categories were added to include deserts, arctic and alpine tundra, and an "other" category. A GIS coverage for these defined ecosystem types was not available for spatial analysis.

After determining what data exists, it is necessary to know which data are currently available to the public. Therefore, a survey question was created to provide information on data availability. Choices included currently available, currently available with conditions on access, available in the future, available in the future with conditions on access, and unavailable.

We also collected information on the methods used to collect data within the database. Survey participants could choose from the following categories: field study, publication review, distributed data warehouse, human health surveillance, survey, or other. In addition, survey participants were asked if their data have been georeferenced or published.

Queries

Once the databases were classified, we queried them for useful information that could be used to meet our objectives. We obtained counts of databases by availability, collector, sampling design, ecosystem type, taxon, method, scale, geographic completeness (i.e., low, medium, high), taxonomic completeness (i.e., low, medium, high), temporal completeness (i.e., low, medium, high), and database type. We also determined the number of databases with geo-referenced data and published data.

Results

Survey Response

From the list of 1,284 contacts known to have received our survey request letter (not including list serves), we had a 46% response rate. From these respondents, we ultimately compiled 252 usable database entries from 214 survey participants.[1] It is critically important to note that the conclusions in this report result completely from the information gathered from the survey respondents and do not include all existing non-native species databases in the U.S. However, we believe these results provide reasonable overview of the state non-native species data in the U.S.

Database Contacts and Funding

Key Findings: Most non-native species database contacts were affiliated with federal agencies or universities. Very few were affiliated with private institutions.

Of the 252 databases, we found that a majority of the survey contacts were affiliated with federal agencies (104), with the Fish and Wildlife Service having the most contacts (41). Universities had the next highest number of contacts affiliated with databases (69), followed by state departments (25), non-profit affiliates (21), museum/herbarium affiliates (11), private affiliates (9), and other affiliates (9; See Table 2).

---

[1]In our initial database research efforts, we identified 192 databases with 169 contacts. Of these 169 contacts, 81 participated in the survey and only 43 of the contacts did not respond to our survey request letter, giving a 75% response rate. After we sent out the survey request letter to our additional contacts not known to be affiliated with a database, we added 155 databases to this list to give a total of 343 databases with 315 contacts. The number of total databases dropped from 343 to 319 after receiving word from 24 of our initial database contacts saying they had no database.

From the entire list of 1,284 contacts known to have received our survey request letter (not including list serves), we had a 46% response rate. We received no response from 700 of these contacts, 271 responded that they have no database, and 43 passed the survey along to be filled out by another researcher. We had nine contacts that responded to our letter but never filled out the online survey, and we only had 17 e-mail failures. We closed the survey with 277 database entries from 230 survey participants. After removing duplicate entries for the same database (i.e., different contacts responded for the same database) and records for databases not related to the United States, we had 252 database entries from 214 survey participants (Note: numbers differ because some participants completed the survey for more than one database). Therefore, we collected metadata for 79% of the existing databases we found through our research. The remainder of the results will deal specifically with the 252 databases that were entered into our online survey.

Table 1. Number of databases by contact affiliation.

Contact Affiliation	Number of Databases
Animal and Plant Inspection Service (APHIS)	2
Agricultural Research Service (ARS)	4
Bureau of Land Management (BLM)	2
Centers for Disease Control (CDC)	1
Cooperative State Research, Education, and Extension Service (CSREES)	1
Environmental Protection Agency (EPA)	1
Fish and Wildlife Service (FWS)	41
Military	4
Museum/Herbarium	11
National Aeronautics and Space Administration (NASA)	1
National Oceanic and Atmospheric Administration (NOAA)	2
Non-profit Organization	21
National Park Service (NPS)	9
Natural Resources Conservation Service (NRCS)	1
Private	9
State Agencies	25
University	69
USDA Forest Service (USFS)	15
US Geological Survey (USGS)	24
Other	9

Information provided as to who provided funding for each database was variable, and there were usually several funding groups listed. A majority of database funding came from public sources (165). Forty-one databases were funded privately, 38 were funded both publicly and privately, four were not funded, and four were classified as unknown because we did not know the status of the funding source (Fig. 2).

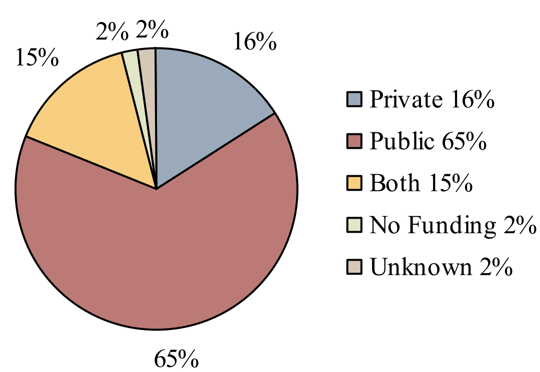


Figure 2. Percentage of databases by funding category.

Of the 165 publicly funded databases, 75 were currently available (45%) and 9 (5%) were unavailable. All of the privately funded databases were available to the public.

Database Type

Key Findings: Most databases have information on non-native species locations, but very few have data on control efforts.

Each database was classified by database type to determine how each one could be used for various research objectives and to determine which ones were useful for the Heinz Center non-native species indicators. Of the eight database types that we assigned to each database, a majority had data on species locations and distribution (137). Species information, general species lists, and non-native species lists were the next most common database types with 77, 64, and 58 databases, respectively. There were also 41 databases that tracked control of non-native species, 13 bibliographic databases, nine distributed databases, and eight databases that did not fit into the other categories (Fig. 3).

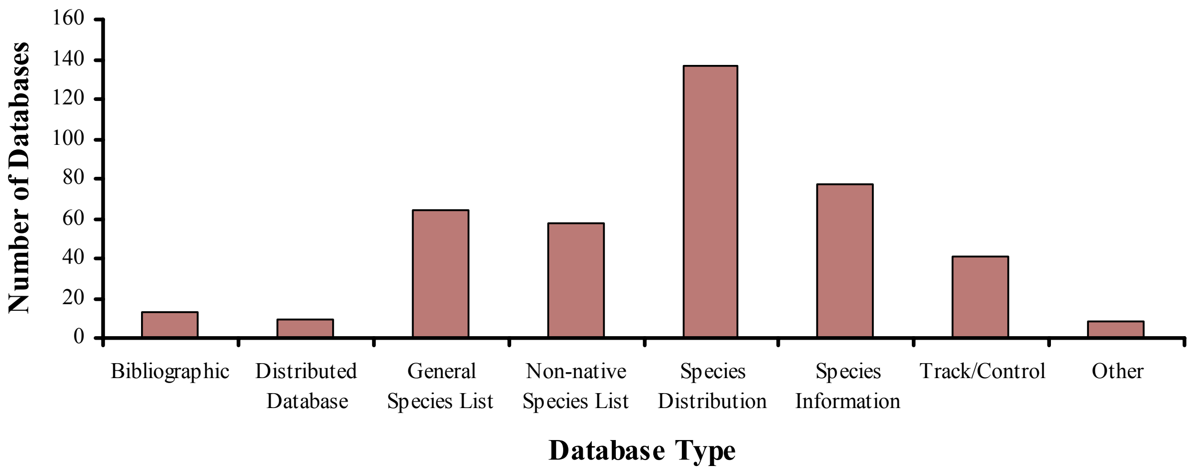


Figure 3. Number of databases by database type.

Data Completeness

Taxonomic Completeness

Key Finding: Many of the databases cover only one taxon; very few cover all taxa.

A majority of databases covered plant species (193). Vertebrates were included in 96 databases, followed by invertebrates (77), pathogens (36), and fungi (22). The total number of databases covering each taxon (424) adds up to more than the total number of databases (252) because many databases covered more than one taxon. This was important in our analysis of taxonomic completeness. We found that 170 databases fell into the low taxonomic completeness category (database covering one taxon), while 67 were found to be of medium taxonomic completeness (database covering more than one, but not all taxa), and 15 were found to be of high taxonomic completeness (database covering all taxa; Fig. 4).

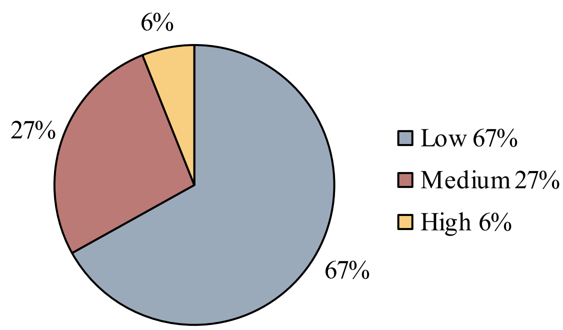


Figure 4. Percentage of databases by taxonomic completeness.

Geographic Completeness

Key finding: The databases vary greatly in terms of spatial scale.

In our analysis, we found that the databases covered a range of spatial scales, with a fairly even distribution among the six categories (Fig. 5). The highest number of databases was at the smaller than county scale (55). We found that 36 covered the global scale, 39 databases covered the national scale (i.e., the United States), 51 covered the state scale, and 44 covered multiple states. Fewer databases covered the county scale, with 27 databases represented (Fig. 5).

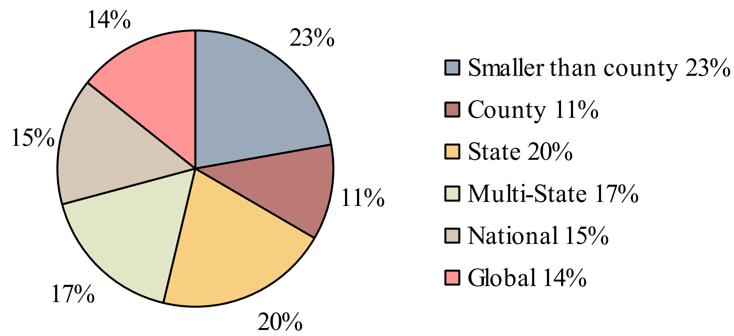


Figure 5. Percentage of databases classified by scale.

Nine states had five or FEWER databases that contained non-native species information specific to their state (Alabama, Arkansas, Georgia, Kansas, Missouri, North Dakota, Oklahoma, South Dakota, and Tennessee). Only six states (California, Colorado, Florida, Idaho, New Mexico, New York) had more than 15 databases, while the rest fell into the medium category with database numbers ranging from six to 15 (Fig. 6). Note that these numbers do not include global or national databases that cover all or almost all of the states.

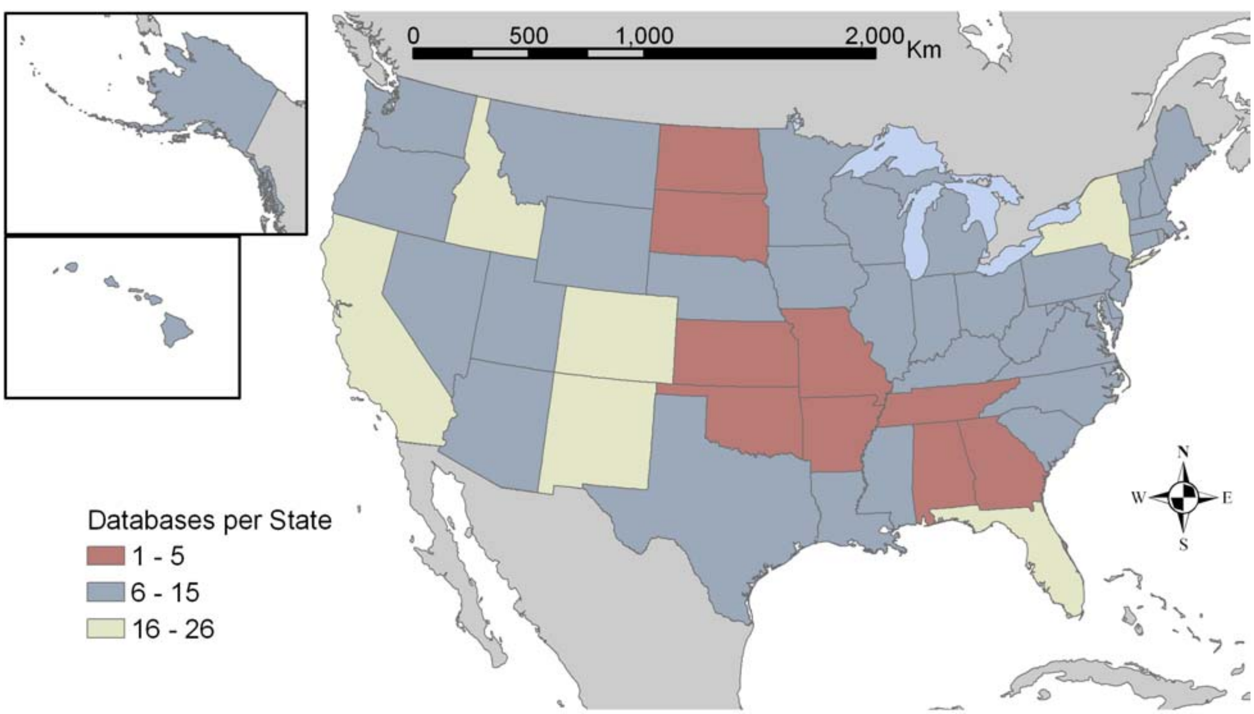


Figure 6. States classified as having a low (1-5), medium (6-15), or high (16-26) number of databases.

Temporal Completeness

Key Finding: More than half of the databases are not updated or updated only irregularly. However, almost half are updated at least annually.

Information pertaining to the time over which data were collected and how often data are updated was used to classify databases by temporal completeness. Survey participants classified their database by how often it is updated as follows: annually (65), monthly (27), daily (18), weekly (10), not regularly (105), not updated (27). Twenty-two databases were ranked as having low temporal completeness, 133 were ranked as having medium temporal completeness, and 97 were ranked as having high temporal completeness (Fig. 7).

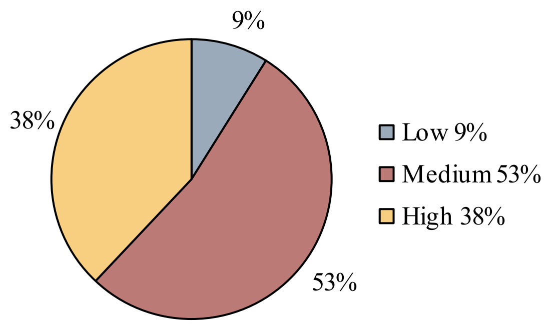


Figure 7. Percentage of databases by temporal completeness.

Key Finding: Sampling design varied widely across the databases; many contained data gathered using more than one design.

We also asked survey participants to describe the sampling design used for each database. The database designs are as follows: complete random (23), stratified random (44), complete systematic (41), random systematic (41), opportunistic (86), other (38), and not applicable (101). (The total number of databases for each design category (374) added up to more than the total number of databases (252) because some databases used more than one sampling design.)

Using this information and the survey responses to how well the study area had been surveyed and whether the study crossed all environmental gradients, it was determined that 37 databases had low geographic completeness, 104 had medium geographic completeness, and 47 had high geographic completeness (Fig. 8). Only four databases could not be classified with the information provided, and 60 survey participants responded that geographic completeness was not applicable to their database design. This typically occurred when the database did not contain data from a field study or studies.

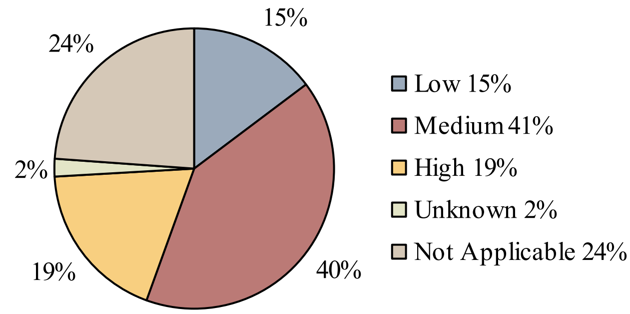


Figure 8. Percentage of databases by geographic completeness.

Data Quality

Key Findings: Most non-native species data were collected by researchers, academics, field technicians and taxonomists.

Information obtained from the survey related to the skill level of those who participated in data collection, the presence of a quality assurance/quality control (QA/QC) procedure, and the description of that procedure was used to place each database into a data quality category. Survey participants classified their databases by collector as follows: researcher (180), field technician (159), graduate student (123), land manager (102), taxonomist (77), undergraduate student (73), naturalist/hobbyist (69), K-12 educator (11), and K-12 student (8). The total number of databases for each collector category (801) added up to more than the total number of databases (252) because most databases had more than one collector type.

Of the 252 databases entered into the survey database, 55% (141) had a QA/QC procedure in place. From our analysis, we found that only 91 of the databases had a high quality assurance/quality control ranking from the definitions we developed. Most databases had a low QA/QC ranking (103), and 57 databases had a medium ranking. There was only one database that we could not classify with the information provided through the survey.

Ecosystem Type[2]

Key findings: Most ecosystem types are represented in a hundred or more databases; however, tundra, deserts, and coasts and oceans are somewhat less well represented.

All ecosystem types were fairly well represented in the databases (Fig. 9). Forests, grasslands/shrublands, and freshwater had similar representation with 166, 162, and 140 databases, respectively. Farmlands and urban and suburban areas had almost equal

[2] Definitions for the ecosystem types used in the *State of the Nation's Ecosystems* are in Appendix D of this document

representation with 110 and 108 databases. Coasts and oceans were represented by 95 databases, deserts by 85, and tundra by 58. Forty-five databases were classified as not falling into any of the ecosystem type categories.

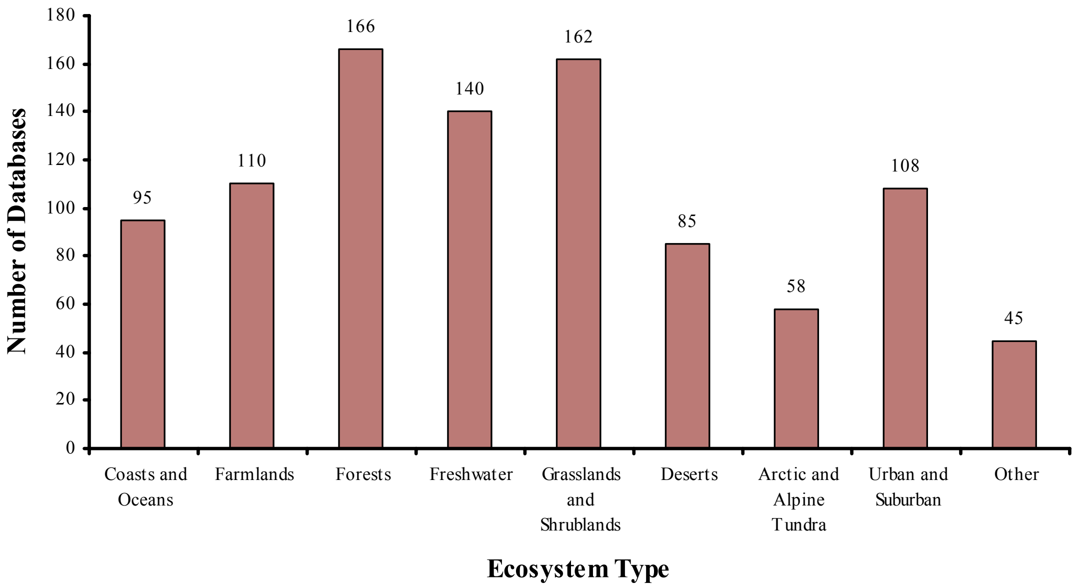


Figure 9. Number of databases within each ecosystem type.

Accessibility and Availability of Databases

Key Findings: Nearly all databases are currently available to the public or will be in the future; some have conditions for access. Almost two-thirds have geo-referenced data.

Additional metadata were collected on each database. Survey participants classified their databases by collection method as follows: distributed data warehouse (12), field study (141), human health surveillance (1), publication review (86), survey (96), and other (84). Survey participants also classified their databases by availability as follows: currently available (116), currently available with conditions on access (57), available in the future (40), available in the future with conditions on access (30), and unavailable (9). It was also found that 162 databases had geo-referenced data, while 90 did not. Seventy-three databases had been published, while 179 had not.

Identification of gaps in non-native species datasets

Additional analysis detailing the state, taxonomic, and ecosystem focus of databases enabled us to identify potential gaps in non-native species data to guide future surveying and spatial predictive modelling efforts.[3]

State Coverage

The state and multi-state databases provide a complete coverage of the United States, but some states were less studied or documented than others. If possible, more surveys should be conducted in these less studied areas to provide more consistent coverage at a national scale. Two of the states that had high coverage, California and Florida, had 16 databases containing data solely for their state. These states may have more intensive invasive species programs, greater invasive species problems, or both. The other states ranked as having a high number of databases (New York, Idaho, Colorado, and New Mexico) are more likely to be involved in a greater number of multi-state efforts because they contain fewer databases dedicated solely to their own individual state.

Hawaii is a state known for having an extensive invasive species problem; however it was classified as having a medium number of databases. Our research shows that this may be the result of greater collaboration among research groups in Hawaii, reducing the number of redundant databases generated. For example, five of the databases listed for Hawaii are part of the Hawaiian Ecosystems at Risk (HEAR) project. Hawaii may prove to be a good example for other states to follow as they begin to improve their invasive species program.

Taxonomic Coverage

We used information on the taxonomic coverage of the databases to identify gaps in non-native species information specific to particular taxa. A majority of databases covered plants, twice as many as the next most covered taxa group (i.e., vertebrates). Of all the databases entered into our survey, 124 covered plants only, and only 60 did not include plants and only focused on other taxa. Fourteen databases looked at only one species, and six of the 14 were plant databases. These results suggest that more studies need to be done on a wider range of taxa.

There are many reasons why plants tend to be more studied than other taxonomic groups. Primarily, plants are easier to study. Also, there appear to be a greater percentage of non-native plants detected in the total plant species pool (60%) relative to other taxa (Pimentel et al. 2001).

Ecosystem Coverage

It is not yet well understood which factors make an ecosystem vulnerable to invasion. For example, a long held theory of plant invasion states that disturbed, species-poor communities are more susceptible to invasion by non-natives due to a lack of biotic resistance from such factors as competition or predation (Elton 1958; Simberloff 1986). However, this theory has been challenged recently as new research on plant invasions has found a higher risk of invasion into highly diverse vegetation types with intermediate levels of disturbance, such as tallgrass prairies, wet meadows, and riparian zones (Robinson et al. 1995; Planty-Tabacchi et al. 1996; Wiser et al. 1996; Stohlgren et al. 1999; Stohlgren et al. 2001).

---

[3] An important caveat is that our results are solely dependent on our survey responses. Therefore, our findings pertain specifically to the databases entered into our survey and are not meant to provide broad generalizations pertaining to all non-native species data.

Therefore, it is important to look at those systems most vulnerable to invasion as well as those that are less vulnerable to fully assess the invasion patterns of all taxa. From the databases collected through the survey, all ecosystem types were covered fairly well. In fact, the numbers of databases for each ecosystem type appears to be fairly proportional to the area of land that each of these systems covers within the United States (Table 2). For example, forest ecosystems had the most database representation, and this ecosystem type covers more area than other ecosystem type on the national landscape. However, the size, coverage and quality of databases vary widely, so their quantity may not be the most important factor.

Table 2. Ecosystem types and the number of databases within each ecosystem type.

Ecosystem Type	Number of Databases
Forests	166
Grasslands and shrublands	162
Deserts	85
Arctic and alpine tundra	58
Coasts and oceans	95
Freshwater	140

The influence of spatial and temporal scale

Scale, both temporal and spatial, is one of the most important concepts in the impact assessment of non-native species. As temporal or spatial scale increases, both the number of processes and their importance in influencing local populations and communities will change, increasing the variability encompassed by the study (Hewitt et al. 2001). Most ecological studies are conducted over short time periods within small spatial regions because long-term and large-scale studies are very costly. However, invasion patterns tend to occur across broad scales especially once the species has become widespread enough to be considered problematic. Therefore, smaller scale studies are unable to provide data that can meet the needs of all land managers dealing with the same species.

The databases that were included in our survey effort showed that non-native species data covered a range of spatial scales, with a fairly even distribution among all the scale categories. This is beneficial considering that invasion patterns are influenced by different factors at different scales. Although smaller scale studies can provide greater detail about the physiological mechanisms that control patterns of invasion, larger scale studies can provide a means to form broad generalizations about landscape scale patterns (Wiens 1989). For land managers, surveys conducted at multiple spatial scales account for all these various patterns and prove to be most beneficial when managing invasions (Stohlgren et al. 2002).

Our understanding of ecological dynamics is also directly related to the temporal scale at which we measure system attributes. A full understanding of the nature of an ecological process may often only be gained after several years or decades of study. Systems that seem highly variable, or chaotic over short time scales may reveal more stable dynamics when observed over longer time periods, as is found for many mammal populations (Hansson 1994; Clutton Brock et al. 1997; Fryxell et al. 1998), nearshore fish and macroinvertebrate assemblages (Jackson and Jones 1999), and microfauna in inter-tidal soft sediments (Olabarria and Chapman 2002). Thus, more long-term studies are essential in determining the degree of establishment of non-native species populations in a given environment. For plants, research findings are dependent on the vegetation type's stage of invasion succession (i.e., temporal scale). Because vegetation surveys only record one point in time, their findings detect current native and non-native species richness, which may have changed since initial invasion and may be different at a future point in time (Levine and D'Antonio 1999).

Although not as equally distributed as spatial scale, databases did cover a range of temporal scales. Only 9% of databases had low temporal completeness, suggesting that a majority of databases have been generating new data for five or more years since their establishment. It was surprising to find that 38% of databases have been generating new data for over ten years since there are few long-term studies in existence. Therefore, it appears from our finding that variations in spatial and temporal scale have been included in current non-native species databases.

The importance of data quality

Regardless of data completeness or the number of records within a database, non-native species data is not very useful if it is not of good quality. Data quality is tightly linked to data analysis because the quality of the data determines the importance and value of the results that are gathered through mining the data (AT&T Corporation 2004). Poor data quality can affect the findings of any study, produce inaccuracies in spatial predictive models, and misguide management efforts, costing land managers both time and money. Therefore, data quality has to be monitored and managed from the very beginning to encompass data gathering, data delivery, data storage, data integration, data retrieval, publishing, and analysis (AT&T Corporation 2004). A majority of the data were collected by people with some experience in the field. However, only 55% of the databases had a QA/QC procedure. It would be desirable to establish a standardized and rigorous quality assurance/quality control procedure for the many non-native species databases currently in existence. This would facilitate data sharing and synthesis.

Conclusions

We have taken an initial step to facilitate non-native species data sharing and collaboration among the many diverse research organizations within the United States. We identified 252 databases within the United States that dealt with non-native species. We are unaware of any other formal investigation of this kind of data quantity, quality, and completeness prior to a national assessment of invasive species.

Of the total number of databases we found, 43% were not available online. This demonstrates the importance of looking for data sources offline to comprehensively determine what data are currently in existence. To improve area species lists and the ability to produce accurate spatial predictive models, it is important to use all electronic data sources currently available, whether online or not. Collaboration among research groups could therefore be improved by providing the capabilities for offline data sources to provide their data online through existing non-native species data source networks (e.g., NIISS). In general, it will be beneficial to encourage all research groups to enter into collaborative efforts and provide all data in an electronic format that will increase their data-sharing capabilities.

The reported datasets are often large and sometimes in archaic formats. Pooling data into a standardized database will be a challenge, but it is not impossible. Current computer systems and standard software packages (SQL-server, Oracle) at the Natural Resource Ecology Laboratory (i.e., NIISS) and most other universities or agencies can handle this volume of data. However, maintaining current data and information on this volume of databases (and even metadata) is more than a full-time job, and a team of database and GIS technicians will be needed.

Our survey response rate was extremely high considering the short duration of the project (4 months) and the usual difficulty in achieving survey participation (Sheehan 2001). We attribute our high survey response rate to the personalized letters we sent out to contacts we knew to be affiliated with a database, the extensive contact list we developed for researchers conducting work in non-native species, and the online survey that was easy to access and not time-consuming. Although we only received a 46% response rate from our entire list of contacts, it is our hope that this initial effort will increase awareness of this project and increase participation in such surveys in the future. Our database survey was completed by a diverse group of researchers from a wide range of state and federal agencies, and we feel that our results provide a strong indication of what information is currently available.

We are now aware of the major existing datasets and those that could be accumulated, formatted, and synthesized to address research issues and to develop and populate the Heinz Center indicators for invasive species. We are also aware of the inherent limitations of the datasets in terms of geographic, taxonomic, and spatial completeness. “Summing” information from various taxa and at various scales to meet various research objectives will require great care since data quality and quantity vary and noticeable data gaps occur in key taxa and in key areas of the country. We are hopeful that our survey will allow for a preliminary formal investigation of the Heinz Center indicators with related information on the uncertainty of the final values presented.

Acknowledgments

Many people volunteered to share data and many more people volunteered time to complete the survey. We were overwhelmed by the positive responses. The Heinz Center State of the Nation's Ecosystems Program provided funding, and the Natural Resource Ecology Laboratory at Colorado State University provided logistic support. To all we are grateful.

References

AT&T Corporation. 2004. Research in Data Quality (DQ). Updated: Internet. Available:http://www.dataquality-research.com/index.html.

Binggeli, P. 1996. A taxonomic, biogeographical and ecological overview of invasive woody plants. Journal of Vegetation Science 7(1):121-124.

Clutton Brock, T.H., A.W. Illius, K. Wilson, B.T. Grenfell, A.D.C. Maccoll, and S.D. Albon. 1997. Stability and instability in ungulate populations: An empirical analysis. American Naturalist 149:195-219.

Crosier, C.S., and T.J. Stohlgren. 2004. Improving biodiversity knowledge with data set synergy: a case study of nonnative plants in Colorado. Weed Technology 18:1441-1444.

Despain, D.G., T. Weaver, and R.J. Aspinall. 2001. A rule-based model for mapping potential exotic plant distribution. Western North American Naturalist 61:428-433.

Elton, C.S. 1958. The ecology of invasions by animals and plants. London: Methuen. 181 p.

Fryxell, J.M., J.B. Falls, E.A. Falls, and R.J. Brooks. 1998. Long-term dynamics of small-mammal populations in Ontario. Ecology 79:213-225.

Hansson, L. 1994. Spatial dynamics in relation to density variations of rodents in a forest landscape. Polish Ecological Studies 20:193-201.

Hewitt, J.E., S.E. Thrush, and V.J. Cummings. 2001. Assessing environmental impacts: Effects of spatial and temporal variability at likely impact scales. Ecological Applications 11:1502-1516.

Jackson, G., and G.K. Jones. 1999. Spatial and temporal variation in nearshore fish and macroinvertebrates assemblages from a temperate Australian estuary over a decade. Marine Ecology Progress Series 182:253-268.

Jacono, C.C., and C.P. Boydstun. Proceedings of the Workshop on Databases for Nonindigenous Plants, 24-25 September; 1997; Gainesville, FL. USGS.

Kimberling, D.N. 2004. Lessons from history: predicting successes and risks of intentional introductions for arthropod biological control. Biological Invasions 6:301-318.

Levine, J.M., and C.M. D'Antonio. 1999. Elton revisited: a review of evidence linking diversity and invasibility. Oikos 87:1-11.

Mack, R.N., D. Simberloff, W.M. Lonsdale, H. Evans, M. Clout, and F.A. Bazzaz. 2000. Biotic invasions: causes, epidemiology, global consequences, and control. Ecological Applications 10:689-710.

National Invasive Species Council. 2001. Meeting the Invasive Species Challenge: National Invasive Species Management Plan. 80 pp.

Olabarria, C., and M.G. Chapman. 2002. Inconsistency in short-term temporal variability of microgastropods within and between two different intertidal habitats. Journal of Experimental Marine Biology and Ecology 269:85-100.

Pimentel, D., L. Lach, R. Zuniga, and D. Morrison. 2000. Environmental and economic costs of nonindigenous species in the United States. BioScience 50:53-65.

Pimentel, D., S. McNair, J. Janecka, J. Wightman, C. Simmonds, C. O'Connell, E. Wong, L. Russel, J. Zern, T. Aquino, and T. Tsomondo. 2001. Economic and environmental threats of alien plant, animal, and microbe invasions. Agriculture, Ecosystems and Environment 84:1-20.

Planty-Tabacchi, A., E. Tabacchi, R.J. Naiman, C. Deferrari, and H. Decamps. 1996. Invasibility of species-rich communities in riparian zones. Conservation Biology 10:598-607.

Rejmànek, M., and M.J. Pitcairn. 2002. When is eradication of exotic pest plants a realistic goal? In: Veitch, C.R., and M.N. Clout, editors. Turning the tide: the eradication of invasive species. Switzerland and Cambridge, UK: IUCN SSC Invasive Species Specialist Group. p. 249-253.

Ricciardi, A., W.W.M. Stiener, R.N. Mack, and D. Simberloff. 2000. Toward a Global Information System for Invasive Species. BioScience 50:239-244.

Ridgway, R.L., W.P. Gregg, R.E. Stinner, and A.G. Brown. Invasive species databases. Proceedings of a workshop 12-13 November; 1998; Las Vegas, NV. Charles Valentine Riley Memorial Foundation.

Robinson, G.R., J.F. Quinn, and M.L. Stanton. 1995. Invasibility of experimental habitat islands in a California winter annual grassland. Ecology 76:786-794.

Sellers, E., A. Simpson, and S.C. Hetrick. List of Invasive Alien Species (IAS) Online Databases and Databases Containing IAS Information; 2004 6-8 April 2004; Baltimore, MD.

Semmens, B.X., E.R. Buhle, A.K. Salomon, and C.V. Pattengill-Semmens. 2004. A hotspot of non-native marine fishes: evidence for the aquarium trade as an invasion pathway. Marine Ecology-Progress Series 266:239-244.

Sheehan, K. 2001. E-mail Survey Response Rates: A Review. Journal of Computer-Mediated Communication 6.

Simberloff, D. 1986. Introduced insects: A biogeographic and systematic perspective. In: Mooney, H.A., and J.A. Drake, editors. Ecology of Biological Invasions of North America and Hawaii. New York: Springer-Verlag.

Simpson, A. 2004. The Global Invasive Species Information Network: What's in It for You? BioScience 54:613-614.

Stohlgren, T.J., D. Barnett, C. Flather, P. Fuller, B. Peterjohn, J. Kartesz, and L.L. Master. 2005. Species richness and patterns of invasion in plants, birds, and fishes in the United States. Biological Invasions. Biological Invasions 18:427-447.

Stohlgren, T.J., D. Binkley, G.W. Chong, M.A. Kalkhan, L.D. Schell, K.A. Bull, Y. Otsuki, G. Newman, M. Bashkin, and Y. Son. 1999. Exotic plant species invade hot spots of native plant diversity. Ecological Monographs 69:25-46.

Stohlgren, T.J., G.W. Chong, L.D. Schell, K.A. Rimar, Y. Otsuki, M. Lee, M.A. Kalkhan, and C.A. Villa. 2002. Assessing Vulnerability to Invasion by Nonnative Plant Species at Multiple Spatial Scales. Environmental Management 29:566-577.

Stohlgren, T.J., Y. Otsuki, C.A. Villa, M. Lee, and J. Belnap. 2001. Patterns of plant invasions: a case example in native species hotspots and rare habitats. Biological Invasions 3:37-50.

The Heinz Center. 2002. The State of the Nation's Ecosystems: Measuring the Lands, Waters, and Living Resources of the United States. New York, NY: Cambridge University Press. 270 p.

Unmack, P.J., and W.E. Fagan. 2004. Convergence of differentially invaded systems toward invader-dominance: time-lagged invasions as a predictor in desert fish communities. Biological Invasions 6:233-243.

Vitousek, P.M., C.M. D'Antonio, L.L. Loope, M. Rejmanek, and R. Westbrooks. 1997. Introduced species: a significant component of human caused global change. New Zealand Journal of Ecology 21:1-16.

Wiens, J.A. 1989. Spatial scaling in ecology. Functional Ecology 3:385-397.

Wilcove, D.S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos. 1998. Quantifying threats to imperiled species in the United States. BioScience 48:607-615.

Wiser, S.K., R.B. Allen, P.W. Clinton, and K.H. Platt. 1996. Invasibility of species-poor forest by a perennial herb over 25 years. Bulletin of the Ecological Society of America 77:488.

APPENDIX A

<u>The Non-native Species Task Group</u>

Ann M. Bartuska (*Chair*)
Deputy Chief for Research and Development
U.S. Forest Service

Jerome Beatty
Deputy Director, Forest Health Protection
U.S. Forest Service, RPC 7

Faith Campbell
Senior Policy Representative
The Nature Conservancy

Gabriela Chavarria
Vice President, Conservation Policy
Defenders of Wildlife

Pam Fuller
Biologist, USGS/BRD

Nelroy E. Jackson
Consultant, Monsanto (*Retired*)

Terri Killeffer
Botanical Research Associate, NatureServe

Richard N. Mack
Professor, School of Biological Sciences
Washington State University

Gary C. Matlock, Ph.D.
Director, National Centers for Coastal Ocean Science (NCCOS), NOAA

Sarah Reichard
Assistant Professor
University of Washington
College of Forest Resources, Center for Urban Horticulture

Peter M. Rice
Project Director, Invaders Data Base
Division of Biological Sciences
University of Montana

Gregory Ruiz
Senior Scientist
Smithsonian Environmental Research Center

Thomas Stohlgren
Science Program Director
USGS/BRD

David Thomas
Director, Illinois Natural History Survey

<u>Liaison to the National Invasive Species Council and the Invasive Species Advisory Council</u>

Chris Dionigi, Ph.D.
Assistant Director for Domestic Policy, Science, and Cooperation
National Invasive Species Council

Heinz Center Staff

Robin O' Malley
Senior Fellow and Program Director

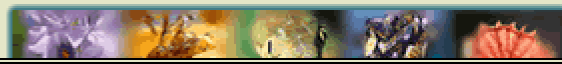
Laura Meyerson
Staff Scientist
(*Currently an Assistant Professor at the University of Rhode Island, Department of Natural Resources Science*)

Caroline Cremer
Project Assistant

Appendix B

THE HEINZ CENTER

NON-NATIVE SPECIES DATABASE SURVEY



THE HEINZ CENTER

The Heinz Center, in collaboration with the USGS National Institute of Invasive Species Science, is working to build an exhaustive list of non-native species databases within the United States. The following survey has been developed to collect essential metadata for each of these databases. If any survey question does not specifically apply to your database, please address this in the comments field. If you have any questions about the survey, or if you prefer to answer the survey questions person-to-person, please feel free to contact Alycia Waters (970) 491-2302.

Please enter the following information	
What is the name of your data set? (If your data set does not have a name, please create one and enter it here.)	
What is your data set's acronym? (Leave blank if not applicable)	
If your data set is online, give the URL. If not online, give its physical location. Example URL: http://www.niiss.org Example Physical Location: The National Institute of Invasive Species Science A219 NESB/NREL; Colorado State University Fort Collins, CO 80523-1499	
What was the purpose of data collection? (Please be concise)	
Enter contact information for a person to contact regarding your data set.	**Last Name** **First Name** **Email** **Work Phone** **Affiliation**
What method was used to collect your data? (Choose all that apply)	☐ Field Study

	☐ Publication review ☐ Distributed data warehouse ☐ Human health surveillance ☐ Survey ☐ Other
If a field study, what was the sampling design used? (Choose all that apply)	☐ Not Applicable ☐ Complete Random ☐ Stratified random ☐ Complete systematic ☐ Random systematic ☐ Opportunistic ☐ Other
What is the taxonomic focus of your data set? (Choose all that apply) Note: Insects and fungi are listed twice because they can be either pathogenic or non-pathogenic. For the purposes of the Heinz Center non-native species indicators, the term "pathogens" applies to fungi, bacteria and viruses, as well as other pathogen-like invertebrates, such as hemlock woolly adelgid.	**Plants** ☐ Vascular & non-vascular **Vertebrates** ☐ Fish ☐ Amphibians ☐ Reptiles ☐ Birds ☐ Mammals **Invertebrates** ☐ Insects ☐ Echinoderms ☐ Tunicates ☐ Crustaceans ☐ Arachnids ☐ Centipedes ☐ Annelids ☐ Mollusks ☐ Cnidarians

	☐ Sponges **Pathogens** ☐ Fungi ☐ Bacteria ☐ Viruses ☐ Insects **Fungi** ☐ Fungi
What ecosystem(s) does your data set cover? (Choose all that apply)	☐ Coasts and oceans ☐ Farmlands ☐ Forests ☐ Freshwaters ☐ Grasslands and shrublands ☐ Deserts ☐ Arctic and alpine tundra ☐ Urban and suburban ☐ Other
Specify the approximate number of species within each taxon covered by your data set.	**Plants** 0 **Vertebrates** 0 **Invertebrates** 0 **Pathogens** 0 **Fungi** 0
What is the scale of your data set?	○ Global ○ National ○ Multi-state ○ State ○ County ○ Smaller than county
What is the name of your study area? Please specify the politically defined region, including state(s) if finer than the national scale.	

(e.g., Rocky Mountain National Park, Colorado)	
When did data collection begin? (Year)	pre-1980
When did data collection end? (Year or Ongoing)	pre-1980
How often are the data updated?	Daily Weekly Monthly Annually Updated, but not regularly Not updated
Please comment on how well the study area has been sampled. (Please be concise) (If not applicable, please leave blank.)	
Does your data encompass the environmental heterogeneity of your study area? (i.e., Was data collected across major environmental gradients?)	No Yes Not applicable (N/A)
Are the data geo-referenced?	No Yes
Who collected the data? (Choose all that apply)	K-12 student K-12 educator Undergraduate student Graduate student Researcher Field technician Land manager Taxonomist Naturalist/hobbyist
Has the data been taken through a quality assurance/quality control process?	No Yes

If yes, give a brief description of the QA/QC process (Please be concise): (If no QA/QC process, please leave blank.)	
Who funded the data set or the research that resulted in the data?	
Will this data collection continue in the future?	○ No ○ Yes
What is the availability of your data?	○ Currently available ○ Currently available with conditions on access ○ Available in the future ○ Available in the future with conditions on access ○ Unavailable
Have your data been published?	○ No ○ Yes
Give the proper citation for referencing your data set. (If unpublished, please fill in the appropriate name(s) or institution(s), date, and the phrase "unpublished data".)	
Additional comments:	

Submit

Appendix C

Heinz Center Non-Native Species Database Survey Request

Dear **(insert contact)**,

The Heinz Center's *State of the Nation's Ecosystems* project, in collaboration with Dr. Tom Stohlgren, Science Program Director at the Fort Collins Science Center, is conducting a review of existing non-native species databases in the United States. This effort is part of the Center's work on the next edition of the *State of the Nation's Ecosystems* report (http://www.heinzctr.org/ecosystems/), scheduled for publication in 2007, and will support work on identifying a consistent set of national indicators for describing non-native species.

Through our preliminary research, we have identified you as the primary contact for the **(insert database name)** database. We request that you participate in a short survey (5-10 minutes total) designed to collect general information about your database (see below for web address). The overarching goal of the survey is to gather information on the types of data that are collected on non-native species in the United States. Two other goals of this effort are to ascertain the availability of data for populating non-native species indicators under development by the Heinz Center and its Non-Native Species Task Group (see below) and to contribute to the development of a larger meta-database of non-native species databases, now being developed by the National Institute of Invasive Species Science (NIISS; http://www.niiss.org/).

With your cooperation, we hope to obtain metadata associated with all major non-native species databases to accurately assess currently available data and identify where gaps exist. Data contributors will be given the opportunity to establish links to the entire "shared database" through the NIISS website, thus providing access to other non-native species data sets and facilitating data-sharing among researchers, agencies, and organizations. All survey participants will be acknowledged for their contribution on the NIISS website, all necessary permissions will be obtained, and all data will be properly cited and attributed. Please specify in the comments portion of the survey if you would prefer to remain anonymous, and we will gladly honor your request. After reviewing the survey results, we may ask your permission to use selected data for inclusion in the 2007 *State of the Nation's Ecosystems* report or other publications.

To complete the survey, please go to www.niiss.org. Click on "Fill out Heinz Center database survey" at the bottom of the page. Your login information will be as follows:

Login: **(insert login)**
Password: **(insert password)**

After your initial login, you can change your login and password if you would like. We request that you please complete the survey by Friday the 24^{th} of September. If you have any questions, concerns, or insights about the survey, please contact Alycia Waters at mawaters@nrel.colostate.edu.

If you have questions about the Heinz Center Non-native Species indicators, please contact Laura Meyerson, meyerson@heinzctr.org, 202-737-6307.

Thank you for your help and cooperation,

Laura A. Meyerson
Staff Scientist
The H. John Heinz III Center for Science, Economics and the Environment

The Heinz Center Non-Native Species Task Group

- Ann Bartuska (Chair), Deputy Chief, Research and Development, USDA Forest Service
- Jerome Beatty, Deputy Director, Forest Health Protection, USDA Forest Service
- Faith Campbell, Senior Policy Representative, The Nature Conservancy
- Gabriela Chavarria, Vice President, Conservation Policy, Defenders of Wildlife
- Pam Fuller, Biologist, USGS, Biological Research Division, Nonindigenous Aquatic Species Program, Center for Aquatic Resource Studies
- Nelroy E. Jackson, Consultant, Monsanto (Retired)
- Terri Killeffer, Botanical Research Associate, NatureServe
- Richard N. Mack, Professor, School of Biological Sciences, Washington State University
- Gary C. Matlock, Director, National Centers for Coastal Ocean Science (NCCOS), National Oceanic and Atmospheric Administration, National Ocean Service
- Sarah Reichard, Assistant Professor, University of Washington, College of Forest Resources
- Peter M. Rice, Project Director, Invaders Data Base, Division of Biological Sciences, University of Montana
- Gregory Ruiz, Senior Scientist, Smithsonian Environmental Research Center, Smithsonian Institution
- Thomas Stohlgren, Science Program Director, Biological Resources Division, USGS, Natural Resource Ecology Laboratory, Colorado State University
- David Thomas, Director, Illinois Natural History Survey
- Chris Dionigi, Assistant Director for Domestic Policy, Science, and Cooperation, National Invasive Species Council

Heinz Center Non-Native Species Database Survey Request

Dear **(insert contact name)**,

The Heinz Center's *State of the Nation's Ecosystems* project, in collaboration with Dr. Tom Stohlgren, Science Program Director at the Fort Collins Science Center, is conducting a review of existing non-native species databases in the United States. This effort is part of the Center's work on the next edition of the *State of the Nation's Ecosystems* report, scheduled for publication in 2007, and will support work on identifying a consistent set of national indicators for describing non-native species.

Through our preliminary research, we have identified you as an expert in the area of non-native species research. If you currently have or are working on a database that includes non-native species, we request that you participate in a short survey (5-10 minutes total) designed to collect general information about your database (see below for web address). The overarching goal of the survey is to gather information on the types of data that are collected on non-native species in the United States. Two other goals of this effort are to ascertain the availability of data for populating non-native species indicators under development by the Heinz Center and its Non-Native Species Task Group (see below) and to contribute to the development of a larger meta-database of non-native species databases, now being developed by the National Institute of Invasive Species Science (NIISS).

With your cooperation, we hope to obtain metadata associated with all major non-native species databases to accurately assess currently available data and identify where gaps exist. Data contributors will be given the opportunity to establish links to the entire "shared database" through the NIISS website, thus providing access to other non-native species data sets and facilitating data-sharing among researchers, agencies, and organizations. All survey participants will be acknowledged for their contribution on the NIISS website, all necessary permissions will be obtained, and all data will be properly cited and attributed. Please specify in the comments portion of the survey if you would prefer to remain anonymous, and we will gladly honor your request. After reviewing the survey results, we may ask your permission to use selected data for inclusion in the 2007 *State of the Nation's Ecosystems* report or other publications.

To complete the survey, please go to www.niiss.org. Click on "Fill out Heinz Center database survey" at the bottom of the page. Your login information will be as follows:

Login: heinzctr
Password: nndb

After your initial login, you can change your login and password if you would like. We request that you please complete the survey by Friday October 15. If you do not have a non-native species database, please respond to this email with 'NO DATABASE' in the subject line. If you have any questions, concerns, or insights about the survey, please contact Alycia Waters at mawaters@nrel.colostate.edu.

If you have questions about the Heinz Center Non-native Species indicators, please contact Laura Meyerson, meyerson@heinzctr.org, 202-737-6307.

Thank you for your help and cooperation,

Laura A. Meyerson
Staff Scientist
The H. John Heinz III Center for Science, Economics and the Environment

The Heinz Center Non-Native Species Task Group

- Ann Bartuska (Chair), Deputy Chief, Research and Development, USDA Forest Service
- Jerome Beatty, Deputy Director, Forest Health Protection, USDA Forest Service
- Faith Campbell, Senior Policy Representative, The Nature Conservancy
- Gabriela Chavarria, Vice President, Conservation Policy, Defenders of Wildlife
- Pam Fuller, Biologist, USGS, Biological Research Division, Nonindigenous Aquatic Species Program, Center for Aquatic Resource Studies
- Nelroy E. Jackson, Consultant, Monsanto (Retired)
- Terri Killeffer, Botanical Research Associate, NatureServe
- Richard N. Mack, Professor, School of Biological Sciences, Washington State University
- Gary C. Matlock, Director, National Centers for Coastal Ocean Science (NCCOS), National Oceanic and Atmospheric Administration, National Ocean Service
- Sarah Reichard, Assistant Professor, University of Washington, College of Forest Resources
- Peter M. Rice, Project Director, Invaders Data Base, Division of Biological Sciences, University of Montana
- Gregory Ruiz, Senior Scientist, Smithsonian Environmental Research Center, Smithsonian Institution
- Thomas Stohlgren, Science Program Director, Biological Resources Division, USGS, Natural Resource Ecology Laboratory, Colorado State University
- David Thomas, Director, Illinois Natural History Survey
- Chris Dionigi, Assistant Director for Domestic Policy, Science, and Cooperation, National Invasive Species Council

Appendix D

Ecosystem definitions used in the State of the Nation's Ecosystems report and for identifying databases for the non-native species indicators

Coasts and oceans consist of three components: estuaries, ocean waters under U.S. jurisdiction, and the shoreline along both estuaries and oceanfront areas.

Farmlands include lands used for production of annual and perennial crops and livestock, field borders and windbreaks, small woodlots, grassland or shrubland areas, wetlands, farmsteads, small villages and other built-up areas, set-aside lands, and similar areas not used for production.

Forests are defined using the USDA Forest Service definition, i.e., any lands at least 10% covered by trees of any size and at least one acre in extent. This includes both heavily treed areas and areas where trees are intermingled with other cover, such as the chaparral and pinyon–juniper areas of the Southwest and includes both naturally regenerating forests and areas planted for future harvest (plantations or "tree farms")—that is, areas that may not have mature trees now, but that will in the future, are classified as forest.

Freshwater ecosystems include rivers and streams, including those that flow only intermittently, lakes, ponds, and reservoirs, from the Great Lakes to small farm ponds, groundwater, freshwater wetlands, including forested, shrub, and emergent (marsh) wetlands.

Grasslands and shrublands are the parts of the terrestrial landscape that are not generally recognized as forests, cropland, or urban and suburban areas. Examples of grasslands and shrublands include tall, mid-, and shortgrass prairies of the Midwest and Great Plains, sagebrush steppes of the northern Rockies, palouse prairies of Oregon and Washington, Florida scrublands, coastal grasslands of the Atlantic and Gulf coasts, chaparral and savanna in California, deserts of the Southwest and intermountain West, mountain shrublands, shrubland and tundra in Alaska, and pastures, as long as they are not cultivated.

Urban and suburban ecosystems are defined using a newly developed approach. It relies upon the physical characteristics of the land, rather than population density, as is commonly done, and employs two basic criteria. First, a substantial portion of the land must be covered with buildings, roads, concrete, and the like, and second, these areas must be sufficiently large (about 270 acres or more) to be considered "urban / suburban." This method excludes scattered or isolated areas such as small settlements, large parking lots, or single residences, but includes large "natural" areas, such as city parks, which are surrounded by otherwise-urban lands.