A data driven method for prioritizing invasive species to aid policy and management

Dylan Finley
State University of New York College of Environmental Science and Forestry at Syracuse: SUNY College of Environmental Science and Forestry
Martin Dovciak
State University of New York College of Environmental Science and Forestry at Syracuse: SUNY College of Environmental Science and Forestry
Jennifer Dean (jennifer.dean@dec.ny.gov)
State University of New York College of Environmental Science and Forestry at Syracuse: SUNY College of Environmental Science and Forestry https://orcid.org/0000-0003-4158-9467

Research Article

Keywords: Invasive species, management, prioritization, iNaturalist, iMapInvasives

Posted Date: March 25th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1435367/v1

Version of Record: A version of this preprint was published at Biological Invasions on April 5th, 2023. See the published version at https://doi.org/10.1007/s10530-023-03041-3.

Abstract

Natural resource managers overseeing large regions are often challenged by an overwhelmingly long list of invasive species to prioritize for management and surveys. Often, managers determine priorities through subjective experience and not regional data, contributing to a lack of objectivity, consistency, and transparency. Using the invasion curve as a guiding principle, we developed a data-driven process to guide expert input in creating regionally specific invasive species lists based on management priorities. The invasive species tiers framework uses a standardized set of definitions, data from locational databases and impact assessments, and expert review to categorize high-impact invasive species present in and surrounding the target regions. The analysis process was evaluated and improved by feedback from the structured network of invasive species managers in New York State. Changes between the first and second version of the data tiers analysis increased total number of species represented in the tier lists, accounting for 40% of the differences between the versions. Results of the invasive species tiers process for eight management regions and at the state-scale were made publicly available, and demonstrated variation in invasive species diversity across the management landscape. The approach developed here can be replicated in and scaled to other regions of the U.S. or other countries with comparable data, and it can provide a common management language to better coordinate invasive species management efforts.

Introduction

Non-native invasive species were estimated to cost US$47 billion to US$163 billion in 2017 worldwide, with mean annual cost tripling every 10 years as they outcompete, predate, or parasitize native species, agricultural crops and livestock, clog pipes, damage houses, decrease water quality, and produce a host of other harmful impacts (Diagne et al. 2021). In the coming decades, the economic losses associated with these species will increase in regions such as the northeastern United States, where the climate will be suitable to hundreds of new species by 2050 (Allen and Bradley 2016). There are approximately 11,400 non-native species already established in the United States (Simpson et al. 2018), although, only a fraction of these established non-native species exhibit invasive behaviors that may cause negative ecological impacts and economic costs (Williamson and Fitter 1996). Even so, the number of invasive species can be regionally very large and it can easily outstrip financial, technical, and personnel resources to manage them. Accordingly, invasive species managers need tools to help them decide which species deserve the limited management resources most.

For decades, a guiding framework for invasive species managers has been the invasion curve, which suggests that preventing invasive species from establishing is the most cost-effective management strategy (if feasible), followed by eradication, while the containment and long-term management of successful invaders are more difficult and costly over time (Hobbs and Humphries 1995). However, the invasion curve provides only a very general management guideline, rather than specific management metrics that would help prioritize which invasive species out of many should be allocated the limited resources. For example, since there is no specific population threshold at which a species shifts from an eradication to a more costly containment phase (Hobbs and Humphries 1995), managers have to interpret where a species may be along the invasion curve, leaving the door open to subjective bias and an overreliance on expert opinion. Decision making processes that rely on expert opinions alone risk neglecting relevant criteria and are rarely documented sufficiently (Hiebert and Stubbendieck 1993; Fox and Gordon 2009).

Seeking a more objective and more quantitative framework for invasive species management decisions, natural resource agencies are increasingly turning to prioritization schemes to allocate species into categories according to the urgency of management. These schemes are generally structured around a series of quantitative and qualitative questions about the species' distribution, ecological and socioeconomic impacts, history of invasion, difficulty of control, and other questions with answers provided that correspond to scores then used to produce a prioritization ranking of the species (McGeoch et al. 2016). Species with the highest rankings are typically ones that pose the highest risk and are the easiest to manage (Dawson et al. 2009). In practice, however, implementation of these prioritization rankings may fall into the trappings of less structured approaches due to two common and interrelated issues. First, data on the abundance, distribution, and impacts of invasives may be unavailable, requiring time consuming and expensive surveys impossible across large geographies with many species. Second, relying on various assumptions about populations to generate prioritization rankings when data are lacking may allow bias to enter prioritization. However, combining existing data with expert knowledge can make prioritization schemes more objective and consistent than an expert-only alternative (Hiebert and Stubbendieck 1993; Cipollini et al. 2005).

To help resource managers prioritize invasive species management across different regions, we had two objectives. The first objective was to develop a general process for combining publicly available data on species locations and impacts with expert knowledge to classify invasive species into categories (tiers) that would allow regionally variable management priorities (referred to hereafter as invasive species tiers). The second objective was to apply this approach to different regions of New York State in the U.S. to both demonstrate the applicability of the approach and to aid regional invasive species management efforts. The resulting general framework and approach can be replicated in other regions of the U.S. or other countries with comparable data, and it can drive the development of necessary data sources and expert networks elsewhere.

Methods

Study area

The study area encompasses the entire New York State (NYS) in northeastern U.S., ranging from New York Harbor in the south to the St. Lawrence Seaway in the north. NYS is a hub of global commerce and one of the most heavily trafficked states in the country, offering multiple paths of entry for new species (Taylor 2013). Of nearly 1,500 non-native plants documented in NYS, 726 are naturalized, and 249 have an unknown status (Werier 2017). With regards to invasive forest pests, the Northeast U.S. has the highest number of damaging non-native insects and pathogens, with NYS leading with over 40 species established (Liebhold et al. 2013). Currently, 155 species are listed as invasive species under the NYS Prohibited and Regulated Invasive Species regulations (NYS Department of Environmental Conservation 2015).

In response to the economic and ecological threats of invasive species, NYS has created robust invasive species management infrastructure. Guided by the NYS Invasive Species Council and funded primarily by the NYS Environmental Protection Fund, a suite of government agencies and other organizations work in tandem to combat invasives (NYS Department of Environmental Conservation 2018). An essential element of the state framework is the network of Partnerships for Regional Invasive Species Management (PRISM) (Fig. 1). Each of the eight PRISMs covering NYS coordinate regionally-specific invasive species issues while maintaining elements of consistency through oversight by the state's Department of Environmental Conservation. This infrastructure, along with a centralized locational database for the state, was particularly well-suited for the developing, testing, revising, and evaluating of a data-driven approach to invasive species prioritization.

Conceptual Species Tiers Framework

The invasive species tiers framework offers a standardized set of definitions to categorize species into regionally specific lists based on impacts and feasibility of control. The general framework follows the concepts of the 'invasion curve' (Hobbs and Humphries 1995), which describes the phases of invasion over time and priorities for action at each phase. Tier 1 species are not yet reported in the region of interest but have reasonable potential for introduction and high impacts elsewhere. Tier 1 species can further be divided into Tier 1a (present within a 160km (100 mile) buffer), Tier 1b (not within the buffer, but present in Eastern North America), and Tier 1c (not present in Eastern North America, but with a viable pathway of invasion). Tiers 2 through 4 represent high-impact species currently present, but at differing abundances within the area for which the list is created. Tier 2 species have a low enough abundance that their eradication may be feasible, Tier 3 species may be contained from spreading further, and Tier 4 species are widespread across the region, making local management efforts only reasonable when protecting valued resources. Tier 5 species are present in the area of concern, but their impacts remain unknown, and they require further monitoring and research. These working definitions were developed from the general concepts of Hobbs & Humphries (1995) by the New York Natural Heritage Program, revised with input from the PRISM coordinators and NYS DEC agency staff, and accepted by these stakeholders in 2017 (Fig. 2).

While the tier ranking definitions provided a useful language for categorizing and communicating invasive species priorities, the initial lists were subjective and based solely on expert opinion. To be more objective, consistent, and transparent in these rankings, a data-driven process that blended expert opinion with data from multiple sources was needed.

Data sources

The data used in this study can be divided into two categories: (1) data on the occurrence of invasive species and (2) data on the impacts of species.

Occurrence Data

Occurrence data came from four species observation databases, spanning a mix of professionally collected data to reports from the general public. Species locations from iMapInvasives (NatureServe 2021) were used to assemble baseline species lists. iMapInvasives is an online, collaborative, GIS-based database and mapping tool which has served as the official invasive species database for NYS since 2010, with data for the state managed by the New Yok Natural Heritage Program (NYS Department of Environmental Conservation 2018). Throughout NYS, trained volunteers and natural resource professionals use iMapInvasives to report invasive species, record treatment effectiveness, and document non-detection results (Jewitt et al. 2021). New reports through the mobile applications or the online interface are marked as unconfirmed until the submitted photo of the species is reviewed by a designated taxonomic expert or verified by project leaders based on professional expertise. Reports of exact species locations are typically mapped as points, lines, or polygons. In some cases, records have locations approximated (e.g., centroids of a counties or waterbody) when sourced from herbarium or museum records, or other historical archives.

Three additional data sources complemented species occurrence data outside of NYS (within a 160km buffer surrounding the state). The Early Detection and Distribution Mapping System (EDDMapS) is a national, web-based mapping system that documents invasive species and pest distributions aggregated from multiple databases and public reports (EDDMapS 2021). We also included observations from Nonindigenous Aquatic Species (NAS), a database from the United States Geological Survey that monitors, analyzes, and records sightings of introduced aquatic species throughout the U.S. (U.S. Geological Survey 2021). A third dataset, iNaturalist, supplied data for both the 160 km buffer area around the state as well as for the areas within NYS. iNaturalist is a global online social network, composed of naturalists, professionals, and the public, who map observations of all types of species, including non-native ones (iNaturalist 2021). We included all observations in iNaturalist that were classified as introduced, not captive/cultivated, and research grade (includes reported location, available photographs, and identification verified by at least two additional users). Using these multiple data sources meant the data included multiple scientific names for the same species. We used a power query in Excel to match species with synonymic scientific names between systems, then manually matched species synonyms to align with the name used in iMapInvasives, which follows the taxonomic standards of NatureServe. Only observations recorded during the 21st century were included in the downloads.

Invasive Species Impact Ranks

We used the impact ranks for non-native species that were generated under the direction of NYS DEC with the standardized assessment forms developed for the state's invasive species regulatory system (NYS Department of Environmental Conservation 2015). Over 600 non-native species have been assessed since 2010 by invasive species biologists and other experts for ecological 'invasiveness' and impacts on social and economic values. Ecological invasiveness scores were based off ecological impacts (0–40 points); biological characteristics and dispersal ability (0–25 points); distribution within both its native landscape and other places it has been introduced (0–20 points); and difficulty of detection and control (0–10 points) (Jordan et al. 2012). Assessors compiled individual scores from each of these categories for a relative maximum score on a scale of 0-100 points. Species with a score of 80 or above were given a 'very high' impact rank, between 70 and 79.99 a 'high' rank, between 50 and 69.99 a 'moderate' rank, between 40 and 49.99 a 'low' rank, and below 40 an 'insignificant' rank (Jordan et al. 2012). A second assessment quantified each species' socio-economic impact based on its human health, economic, and cultural benefit (positive score) or detriment (negative score). Scores from these three socio-economic components were added together for a net score on a scale of -100 to +100, with negative scores meaning the species had an overall negative socio-economic impact and positive scores meaning the species had an overall positive socio-economic impact (NYS Invasive Species Council 2010).

Data tiers

To provide a provisional baseline set of regional species tier lists for expert review, we generated standardized tier rankings based solely on the occurrence and impact data, henceforth 'data tiers,' for all nonnative species. We produced lists of these data tiers for each of the eight PRISMs and for the entirety of NYS, with the region of interest henceforth referred to as the 'target region.' We did this for all target regions twice between January

2020 and January 2021, with adjustments to the process between the first and second versions based on expert feedback (Fig. 3).

Within Target Regions (Tiers 2 through 4)

For both versions, we began by counting the number of documented populations for all non-native species located within the target regions (to approximate their abundance in each region), using observation records in iMapInvasives in the first version, and adding observations from iNaturalist in the second version. This generated population estimates of each species that were often inflated due to overreporting of the same species in the same locations. To account for this, we employed a separation distance of 100 meters in ArcMap that consolidated individuals of the same species within 100 meters of each other into a single population (initially following the 50-meter separation distance suggested by (Rew and Pokorny 2006), but we amended it to 100 meters for Version 2 based on input from PRISM leaders). In both versions the separation distance generated a more realistic number of populations for each species. We only used species observations with known coordinates when applying the separation distance. Observations with 'approximate' locations in iMapInvasives were counted separately for each species and added together with the populations generated through the separation distance to get final estimate of the number of populations for each species in each target region.

Any non-native species that had either a ‘high’ or ‘very high’ ecological impact score, or a ‘significant negative,’ ‘high negative,’ or ‘very high negative’ socio-economic score according to the assessments (referred to collectively as “high impact”) were placed in the data tiers 2, 3, or 4 lists. These species were then sorted by population count, with the list divided evenly into tiers 2 through 4 based on its population count, with the lowest 33% of populations landing a species in Tier 2, the middle 33% of populations in Tier 3, and the upper 33% of populations in Tier 4 (Fig. 4). We placed non-native species that did not have a high impact (either unassessed or scoring insignificant, low. or moderate) in an ‘untiered’ list. We did not generate data tier 5 species (those with unknown invasiveness), as this category was to be developed solely from expert input (given the lack of data).

Outside of Target Region – 160km Buffer (Tier 1a)

Next, we created initial Tier 1a lists by analyzing the 160km buffer around each target region and selecting species that have high impact ranks. For the buffer around NYS, data came from iMapInvasives, EDDMapS, and USGS NAS in Version 1, with the addition of iNaturalist data in Version 2. Any high impact species present in a buffer surrounding the target region, but not in the region itself was placed in Tier 1a. Any unassessed species or ones with moderate or low impact ranks present in this buffer received an ‘untiered in buffer’ rank. We did not include any species that exclusively fell outside of this 160 km buffer; these species were considered for Tiers 1b and 1c in the expert review.

Expert Review of Data Tiers

Invasive species and taxonomic experts were recruited to review the data tiers and provide feedback on the analysis process and create finalized ‘expert tiers.’ For the eight PRISMs, the expert groups consisted of PRISM coordinators and staff for that target region, with some creating their own partner workgroups extending beyond the PRISM host organization. For the statewide tiers we assembled a 'Statewide Committee' of invasive species professionals from across NYS to provide feedback. We divided the members into three groups based on expertise: terrestrial plants, aquatic organisms, and plant pests/terrestrial animals.

For all target regions we convened webinar meetings with the experts to explain the process for providing feedback in their tier review spreadsheets. Experts then assessed each species and determined an expert tier, either agreeing or disagreeing with the data tier. If the experts disagreed with the data tier, they chose one of the options from a specified list of potential reasons for change (Table 1). The following target regions participated in a full expert review of Version 1: LIISMA, SLELO, CRP, and the Statewide Committee. All target regions participated in the Version 2.

Table 1
Reasons for change that experts would use to adjust a data tier within one of the data tier categories. Checkboxes indicate the data tier sheet in which the reason was an available option. Italicized reasons were only present in the second version.

	Data Tier Categories			
Reason for Change	**2–4**	**1a**	**Untiered**	**Buffer**
Data populations higher than estimated	ü			
Under-reported in the data	ü			
Higher ecological impact than NYS Assessment			ü	ü
Lower ecological impact than NYS Assessment	ü	ü		
Other reason (please explain)	ü	ü	ü	ü
Higher socio-economic impact than NYS Assessment			ü	ü
Lower socio-economic impact than NYS Assessment	ü	ü		
Distributed over a broad geographic range	ü			
Distributed over a limited geographic range	ü			
Present in region		ü		ü
Absent from region	ü		ü	

We embedded into the tier review spreadsheets a standardized method for providing feedback. Many PRISMs had previously determined tiers for species and published them on their websites. We placed these tiers adjacent to the data tier (derived from our analyses above) in a column called 'Old Expert Tier.' Adjacent, a 'New Expert Tier' column provided a place for experts to offer an updated and final tier rank. A 'Reason for Change' column allowed experts to explain why the new expert tier might have differed from the data tier (Table 1). For example, experts might suggest that a species with a data tier of 2 was more abundant than reported in the online data, and may in turn give it an expert tier of 3. Or they might disagree with the state impact ranks and consider a species not as invasive in their region and reclassify it with an expert tier of '5' (i.e., species requires more monitoring or research to understand its invasiveness). Experts could also add species to Tiers 1b and 1c based on their knowledge of species to watch for beyond the state boundaries. We also included an ‘Other Reason’ column to allow for alternative reasons for change, and a ‘Notes’ column to justify any of these changes in greater detail. We designed this feedback process to encourage experts to think critically about each species tier placement, to incorporate the data into their decision making, and to provide documentation on their reasoning.

Between the first and second versions we made some minor changes to the feedback process. We added an additional reason for change, ‘Species is distributed over a broad geographic range’ or ‘Species is distributed over a limited geographic range’ to allow experts to change a species’ tier based on its distribution within the target region, a different criteria from abundance (Table 1). For target regions that had participated in the first version we retained the feedback they had originally provided in the ‘Expert Tier,’ ‘Reason for Change,’ ‘Other Reason,’ and ‘Notes’ columns for reference in the second version. This expedited allocating species into tiers since much of their original decision-making process had been documented.

Once all target regions provided expert tiers, we made the tiers publicly viewable in an online table (www.nynhp.org/invasives/species-tiers-table), allowing anyone to see how invasive species are being prioritized across the state. We also created an Esri StoryMap explaining in detail the tiering process to accompany the table (www.arcgis.com/apps/MapJournal/index.html?appid=c7af93ee62314f789b2bfc1802a5cc4a).

Evaluating Additional Data Tier Variables

To assess the Version 1 method of creating data tiers 2 through 4 by dividing the species list into thirds by population count, we generated additional data tiers using three different variables and compared alignment of each method to the expert tiers. For species with a data tier of 2 through 4, we counted the number of species that had differing data and expert tiers, henceforth ‘tier mismatches,’ a sign that the data tiers were not aligned with expert opinion. Some PRISMs did not review every data tier species, and these were omitted from the tallies. If one or more of the new variables improved alignment between data and expert tiers, then a new method using those variables would replace the thirds by population count method of generating data tiers for use in the second version. The variables tested were:

1) Using a measure of regional occupancy instead of population counts as the proxy for abundance: In the first version, we did not consider a species’ distribution across the target area when generating a data tier, potentially overlooking an important factor in prioritization. In this test, instead of using population counts across the target region, we estimated occupancy of a species across a region using the percentage of the region that had observations reported in it. For this, we divided the state into 5 x 5 km grid cells and calculated in ArcMap the number of these cells that contained observations of each species. We used this proportion of grid cells in a region containing at least one observation as the distribution estimate.

2) Inclusion of a 'difficulty of control' rating separate from the ecological impact ranking: The difficulty required to control a species may be another important factor in determining management priorities. Difficulty of control ratings were taken from the New York non-native ecological invasiveness assessments. These ratings were comprised of four subsections, each broken down into a series of questions. Together they added up to a score between 0–10. A species’ maximum possible score was 10/10, or 1.0. If assessors did not have

enough information on a species' difficulty of control information to answer all the questions, we used the relative maximum score. For example, if the maximum possible points for the questions that could be answered are 8, and the species received an 'outcome score' of 6, then the species' 'relative maximum score' would be 6/8 or 0.75. We gave species with difficulty of control ratings of > 0.8 a lesser priority data tier (e.g. bumping a species in Tier 3 to Tier 4). Species with difficulty of control ratings of < 0.3 were given a higher priority data tier. We used 0.8 as a baseline because it equates to 80/100, which was the score used to classify a species' impact as 'very high' in the assessments. 30/100 equated to 'insignificant' in the assessments.

3) Inclusion of an abundance comparison to adjacent regions: Even if a species is abundant in one PRISM, it may not be in an adjacent PRISM, and managers may want to devote extra resources to keeping it that way. We gave species in a PRISM with 300% or greater abundance (based on population counts) than in any neighboring PRISMs a higher priority data tier.

We tested each of these variables separately by generating new data tier lists using each variable and comparing with the Version 1 expert tier results. Whichever variables resulted in a net positive matching of data tiers and expert tiers were selected for inclusion into the next round. The changes were to be employed in the order of the highest percentage of positive matching between data and expert tiers, with any variables that did not result in a net positive matching of data and expert tiers omitted. At each stage, the included variable would be retested to ensure it was still improving the overall matching and eliminated if not.

Evaluating the overall process

The expert feedback process and discussions with regional managers prompted changes to the overall data tier generation process between the first and second versions: inclusion of iNaturalist data within and surrounding the target region, splitting out genus-level groupings to individual species, increasing the separation distance for individual observations to be grouped into populations, and the inclusion of socio-economic ranks. After all the expert tiers had been completed for both versions, we analyzed tier 2 through 4 lists for the following: (1) The reasons species shifted data tiers between Version 1 and 2, (2) The proportion of species with mismatched data and expert tiers vs matched tiers, (3) The most commonly cited reasons for tier mismatches, and (4) The species data tiers which had the most tier mismatches. We also compared the number of mismatches in the Version 2 to the number of mismatches in the Version 1 to determine if the changes incorporated into Version 2 had improved alignment between the data and expert tiers. We broadly disseminated these results back to our expert panels to provide feedback on their rankings and facilitate further discussion.

Results

Within NYS, 167 species were considered high impact invasive species and consequently placed in either Tier 2, 3, or 4 for the regions in which they were documented (Version 2 of the Data Tiers; Table 2). Terrestrial plant species represented 63.5% of these species, with the remainder comprised of aquatic animals (16.5%), aquatic plants (10.0%), and terrestrial animals (10.0%).

Table 2
The number of species and populations, and population density for all target regions according to the Version 2 data tiers. Tiers 2 through 4 refer to high impact species present in each target region. Sq km is the area of the target region, and pops / sq km is the population density. Tier 1a refers to high-impact species present in a 160-km buffer around the target region but not present in the target region. V1 Mismatch refers to the percentage of Tiers 2 through 4 species that had conflicting data and expert tier ranks. V2 Mismatch is the same but for Version 2.

Tiers 2 through 4							**Tier 1a**	
Region (PRISM)	**Species**	**Population**	**Sq km**	**Pops / sq km**	**V1 Mismatch**	**V2 Mismatch**	**Species**	**Population**
NYS	167	79,639	141,191	0.6	37.2%	28.9%	22	317
LH	133	22,751	8,073	2.8		44.4%	54	1,358
LIISMA	120	14,765	8,275	1.8	64.0%	42.5%	40	1,073
FL	115	11,902	32,056	0.4		8.9%	54	3,273
WNY	102	7,899	21,085	0.4		44.1%	47	1,005
CRP	96	8,784	12,761	0.7	31.7%	38.5%	73	6,226
CRISP	85	4,625	13,338	0.3		27.4%	86	15,076
SLELO	78	4,277	19,062	0.2	39.4%	26.9%	66	4,809
APIPP	74	4,457	26,537	0.2		47.3%	74	7,081

Importantly, the regional data tiers also showed broad differences between the regions (PRISMs) in the number of populations and diversity of non-native species (Table 2). While some regions (e.g., APIPP, SLELO, and CRISP) were relatively less invaded (with less than 5,000 mapped populations of tier 2 through 4 species), others (e.g., LH) were heavily invaded (with over 20,000 mapped populations of tier 2 through 4 species). While one of the least invaded regions (CRISP) had the lowest counts of non-native tier 2 through 4 species populations, its proximity to the most heavily invaded region (LH) contributed to its very large number of Tier 1 species population (over 15,000). The PRISMs with the highest populations of non-native species were located in the southern part of NYS, close to New York City. While the regions (PRISMs) are far from equal in size, their non-native species population density bears out a similar result, with both of the two regions with the highest density of invasive species in the southern portion of the state (LH and LIISMA).

Reducing mismatches between Data Tier and Expert Tier

An analysis of the expert feedback for species with data-based tiers 2 through 4 showed broad differences in the types of reasons experts chose to differ from the data tier (Fig. 5). The most commonly cited reason was that a species was ‘under-reported in the data’ (55.6% of reasons listed in Version 1 and 45.7% of reasons in Version 2). In both versions this reason was cited far more often than the reverse: a tier shift attributed to the data population over-reporting the species (5.6% in Version 1 and 11.5% in Version 2). ‘Other reason’ was also frequently cited (21.3% in Version 1 and 25.7% in Version 2) as a reason for tier shifts. ‘Other reasons’ listed by experts included that the species was common in neighboring regions, biocontrol options were in development, or management programs for it had already begun. Sometimes, reviewers moved species that were in the data tiers 2 through 4 list into the untiered list. This was typically when they considered the species to have a lower ecological or socio-economic impact than assessed or there was disagreement about the species' nativity. Difficulty of control did not factor heavily into reviewers' tiering decisions in either version.

Though there were similar numbers of species in each of the data tiers 2 through 4 (approximately 33.3% each), mismatches between expert and data tiers were not proportional. In the second version (with similar proportions found in the first version), species placed in Data Tier 3 were most likely to be moved by an expert reviewer, representing 49.2% of the mismatches. Species given a Data Tier 3 were much more likely to be given an Expert Tier 4 (32.7% of mismatches; indicating a higher abundance than revealed by the data), than an Expert Tier 2 (13.0%). Data Tier 2 accounted for 42.1% of the mismatches, with over half of these moved off the Tiers 2–4 list by the experts, with the species either being moved to the tier 5 category (unknown impact) or deemed low impact and thus should not be prioritized for management. Data Tier 4 had the smallest portion of the total mismatches (8.7%).

The total amount of tier mismatches varied widely between the regions assessed, ranging from expert disagreement with the data tier in 8.9% of assessed species in the FL PRISM to 47.3% of species in APIPP (Table 2). Across all target regions, 32.9% of the species placed in data tiers 2 through 4 were changed during the expert review of the second version. Only four target regions participated in the first version expert review, and of these, three had fewer mismatches with the new methods of Version 2 (Table 2), suggesting overall improved alignment between the data and expert tiers.

Changes to the data tiers process based on expert feedback increased total number of species represented in the tier lists

Between the first and second versions, a total of 422 data tier shifts (a species moving from one data tier to another or added as a new species) occurred across all regions based on the updated methodology (Fig. 6). Most of the data tier shifts (40%) occurred due to the addition of observation data ("New to the Data"), either from added inclusion of iNaturalist data, or new observations added into iMapInvasives. In the NYS tiers analysis, 16 species were added as "New to the Data" in the second version. Some species were added to data tiers 2 through 4 because they had high or very high socio-economic scores (6 species added to NYS tiers), which in the first version had not qualified them for inclusion in these data tiers. Other species were added because they had previously only been identified to the genus level, such as *Lonicera spp* and were now identified to the species level, like *Lonicera maackii* and *L. morrowii.* Among species which shifted between data tiers 2 through 4, most shifted towards a higher data tier, likely because the inclusion of the iNaturalist data increased their population estimate. Some however, shifted towards a lower data tier because the separation distance which defined distinct populations shifted from 50 meters to 100 meters. Counterintuitively, a few species that gained populations sank to a lower tier because the percentiles which determine data tiers 2 through 4 had shifted. Some species which lost populations increased to a higher tier because of the same reason.

A simple approach to creating the data tiers outperformed more complex methods

None of the additional variables tested with the Version 1 data improved tier matchups between the data tiers and expert tiers; incorporating regional occupancy, difficulty of control, and abundance in adjacent regions each increased the number of mismatches as compared to dividing the population counts into thirds (Table 3). Therefore, we did not incorporate any of these variables going forward, and kept the simplistic method of dividing the lists into thirds based on population counts to bin into data tiers in the second version.

Table 3

Effects of adding other variables to the process of dividing species lists into data tiers 2 through 4, as compared to the simple method dividing the population counts into thirds. Regional occupancy refers to replacement of population as the measure of abundance with number of occupied 5 x 5 km grid cells; Difficulty of control refers to the inclusion of a difficulty of control rating separate from the ecological invasiveness ranking; Abundance in Adjacent Regions refers to the incorporation of an abundance comparison to adjacent PRISMs. Regional Occupancy and Difficulty of control were based off data and expert tier comparisons from the target regions that participated in Version 1: SLELO, LIISMA, CRP, and NYS. Abundance in Adjacent Regions was assessed for SLELO, LIISMA, and CRP, but did not include analyses from the Statewide tiers, which is why there are less mismatches than the other variables. A positive net difference (increase) in the number of mismatches indicates that the additional variable did not result in better matchup with the expert tiers than the simple thirds method.

Additional Variables	Mismatches w/ thirds rule	Mismatches w/ new variables	Net difference	Incorporate into Version 2
Regional Occupancy	90	93	+ 3	No
Difficulty of Control	90	118	+ 18	No
Abundance in Adjacent Regions	58	63	+ 5	No

Discussion

The invasive species tiers framework was designed to help regional and state managers think strategically about which species to assign to which management priorities from overwhelmingly long lists of non-native and potentially invasive species. The data driven process developed here brings efficiency and standardization to the task of categorizing invasive species into priorities for a target region, which is often done with varying degrees of subjectivity by resource managers. However, since data gaps over large regions and lack of published information on impacts are common, the expertise of professionals dealing with invasive species was invaluable to informing the finalized tiers. By interacting with the experts, our work further informed expert opinions in different regions while also taking into account information not available in existing datasets.

Though the tier lists generated are particular to the invasive species of NYS, this process can be replicated with other similar sources of locational data and invasiveness information. Federal and non-governmental databases in the United States are available for focused invasive species reporting, such as USGS NAS, EDDMapS, and iMapInvasives (Reaser et al. 2020). However, in the absence of resources and political will for invasive species mapping projects, a region or state might have limited data entered into these targeted systems.

With the rise of simple and mobile mapping interfaces, species location data from the public has become more abundant, particularly as unstructured observation reports into systems such as iNaturalist (Rapacciuolo et al. 2021). There are at least 26 community science-focused databases containing nonnative species observations with most reporting in North America and Western Europe (Johnson et al. 2020). In Version 1 of the tiers process, iNaturalist data was not included, since the initial plan was to only use data from systems focused on invasive species reporting. However, after the first round of expert review, it became apparent that many species, particularly ones in low abundance, were underreported in the state invasive species database, yet captured by public reports in iNaturalist. Upon recommendations by the expert reviewers, iNaturalist data was incorporated into the tiering process for Version 2.

Bringing iNaturalist data into the tier analysis increased the number of high impact species and overall populations of non-native species. There were, however, some limitations important to note. Since much of the information in iNaturalist is crowdsourced, such as species identifications in reports and whether a species is flagged as introduced to a location, there is the possibility of errors in the dataset. Also, the comprehensive list of non-native species in iNaturalist greatly increased the number of species names to match to the baseline from iMapInvasives, which focuses on species known or suspected to be invasive. Species that were not flagged as "high impact" were categorized as untiered, making this list longer and more difficult for experts to review. And while the iNaturalist data was useful for new presumed invasive species, such as ornamental species that had been noted by the expert reviewers as growing aggressively outside of cultivation, many reports were obviously intentional plantings from the photos submitted. The field in iNaturalist for "captive / cultivated" species is underutilized; none of the approximately 200,000 introduced species records in New York downloaded for the analysis were marked as captive. Finally, not all species are reported evenly in unstructured community science databases (Ward 2014; Rapacciuolo et al. 2021; Callaghan et al. 2021). Expert reviewers noted improved alignment of data and expert tiers with iNaturalist data for showy species, such as Japanese primrose (*Primula japonica*), but not for less detectable species, such as grasses. Awareness of these and other caveats are important, yet in the absence of other data sources, iNaturalist data can offer a means to generating lists of invasive species observed in a region and watchlists of species in a buffer surrounding the region of interest (Young et al. 2021).

In addition to spatial information, the data tier analysis relied on ecological invasiveness and socioeconomic assessments for each species. It can take months, if not years of work to produce comprehensive assessments for regions with many species (Heikkilä 2011). We had the luxury of using impact assessments for over 600 species that were previously completed by the NYS Department of Environmental Conservation over the course of a decade using a standardized protocol (Jordan et al. 2012). Other standardized systems for evaluating invasiveness could be used, such as the IUCN Environmental Impact Classification for Alien Taxa (Hawkins et al. 2015).

The final critical element in this prioritization scheme is expert feedback. As previously stated, expert feedback filled in gaps in the data and provided a counterbalance to our automated approach. Expert feedback was responsible for several improvements to the data tiering process between the first and second versions, including lengthening the separation distance to 100 meters, using socioeconomic impacts as a criterion, and including observation records from iNaturalist. Also, comparing data tiers to expert tiers allowed us to evaluate the effectiveness of different variables to bin the high impact species into appropriate tiers; we found the simplistic method of binning the species into thirds by number of populations outperformed more complex approaches. Also, the expert feedback on species that were "underreported in the data" prompted a push for new observations to be reported to iMapInvasives so future tier analyses will be more accurate.

These new tier lists are currently being used in New York by state and regional invasive species managers in their prioritization decisions. Seven of the eight NY PRISMs incorporate the tier framework into their priority species webpages, and PRISM and statewide tier rankings are informing which species projects receive management contracts and where managers direct their resources. At the statewide level, the Tier 1 lists are being used for a 'horizon scanning' committee that will assess risks posed by new incoming invasive species, which is a recommended action set forth in the NYS Invasive Species Comprehensive Management Plan (NYS Department of Environmental Conservation 2018).

The online searchable tiers table and StoryMap developed to explain the process will help NYS managers communicate invasive species efforts with national invasive species organizations and neighboring states, as well as between NYS agencies and regions. These resources will also help connect and coordinate management efforts between PRISMs. And they will educate the public about how invasive species management decisions are made across the state.

The tiers process will be repeated each year by NYNHP with updated data from the previous field season. Scripts are being developed to automate the processing steps of Version 2, such as taxonomy matching and separation distance analyses. The expert review spreadsheets will have the previous year's data and expert tiers, with changes in the new data tiers highlighted to facilitate expert review. Any changes to the expert tiers will be reflected on the publicly available online tiers table.

This process developed and tested for NYS and the NY PRISMs is replicable and scalable to other regions of interest. Widespread adoption of similar data-driven tiers would allow neighboring states and provinces to share a common management language and better coordinate invasive species management efforts.

Declarations

Acknowledgments

We are grateful for the many expert reviewers who contributed time and feedback on the tier process and outputs, including staff of the eight NY PRISMs and NYS DEC and NYS AGM. Colleagues at NYNHP were invaluable throughout this project, including Steve Young, John Marino, Mitchell O'Neill, and Meg Wilkinson. Funding was provided by the New York State Environmental Protection Fund as administered by the New York State Department of Environmental Conservation.

Funding

This work was supported by the New York State Environmental Protection Fund as administered by the New York State Department of Environmental Conservation.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Data Availability

The datasets of invasive species locations analyzed during this current study are publicly available at the following online databases:

iMapInvasives (2021) iMapInvasives: NatureServe's online data system supporting strategic invasive species management. Arlington, Virginia. http://www.imapinvasives.org

EDDMapS (2021) Early Detection & Distribution Mapping System. The University of Georgia - Center for Invasive Species and Ecosystem Health. https://www.eddmaps.org/

iNaturalist (2021) iNaturalist. https://www.inaturalist.org

U.S. Geological Survey (2021) Nonindigenous Aquatic Species Database. Gainesville, Florida. https://nas.er.usgs.gov/

The New York State socio-economic and ecological invasiveness rank for each species analyzed in this study can be found at: https://www.nynhp.org/invasives/species-tiers-table/

References

1. Allen JM, Bradley BA (2016) Out of the weeds? Reduced plant invasion risk with climate change in the continental United States. Biol Conserv 203:306–312. https://doi.org/10.1016/j.biocon.2016.09.015
2. Callaghan CT, Poore AGB, Hofmann M et al (2021) Large-bodied birds are over-represented in unstructured citizen science data. Sci Rep 11:19073. https://doi.org/10.1038/s41598-021-98584-7
3. Cipollini KA, Maruyama AL, Zimmerman CL (2005) Planning for Restoration: A Decision Analysis Approach to Prioritization. Restor Ecol 13:460–470. https://doi.org/10.1111/j.1526-100X.2005.00057.x
4. Dawson W, Burslem DFRP, Hulme PE (2009) The suitability of weed risk assessment as a conservation tool to identify invasive plant threats in East African rainforests. Biol Conserv 142:1018–1024. https://doi.org/10.1016/j.biocon.2009.01.013
5. Diagne C, Leroy B, Vaissière A-C et al (2021) High and rising economic costs of biological invasions worldwide. Nature 592:571–576. https://doi.org/10.1038/s41586-021-03405-6
6. EDDMapS (2021) Early Detection & Distribution Mapping System. The University of Georgia - Center for Invasive Species and Ecosystem Health. https://www.eddmaps.org/
7. Fox AM, Gordon DR (2009) Approaches for Assessing the Status of Nonnative Plants: A Comparative Analysis. Invasive Plant Science and Management 2:166–184. https://doi.org/10.1614/IPSM-08-112.1
8. Hawkins CL, Bacher S, Essl F et al (2015) Framework and guidelines for implementing the proposed IUCN Environmental Impact Classification for Alien Taxa (EICAT). Divers Distrib 21:1360–1363. https://doi.org/10.1111/ddi.12379
9. Heikkilä J (2011) A review of risk prioritisation schemes of pathogens, pests and weeds: principles and practices. Agricultural and Food Science 20:15–28. https://doi.org/10.2137/145960611795163088
10. Hiebert RD, Stubbendieck JL (1993) Handbook for ranking exotic plants for management and control. U.S. Dept. of the Interior. National Park Service, Natural Resources Publication Office

11. Hobbs RJ, Humphries SE (1995) An integrated approach to the ecology and management of plant invasions. ConservBiol 9:761–770
12. iNaturalist (2021) iNaturalist. https://www.inaturalist.org
13. Jewitt A, Antolos E, Lutz C, Dean J (2021) Targeted species projects for volunteers to increase early detection capacity: the water chestnut mapping challenge. naar 41:203–208. https://doi.org/10.3375/043.041.0306
14. Johnson BA, Mader AD, Dasgupta R, Kumar P (2020) Citizen science and invasive alien species: An analysis of citizen science initiatives using information and communications technology (ICT) to collect invasive alien species observations. Global Ecol Conserv 21:e00812. https://doi.org/10.1016/j.gecco.2019.e00812
15. Jordan M, Moore G, Weldy T (2012) New York State Ranking System for Evaluating Non-Native Plant Species for Invasiveness. The Nature Conservancy, Cold Spring Harbor, NY. http://nyis.info/wp-content/uploads/2017/10/New_York_State_Invasive_Plant_Ranking_System_Rev_2012.pdf
16. Liebhold AM, McCullough DG, Blackburn LM et al (2013) A highly aggregated geographical distribution of forest pest invasions in the USA. Divers Distrib 19:1208–1216. https://doi.org/10.1111/ddi.12112
17. McGeoch MA, Genovesi P, Bellingham PJ et al (2016) Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. Biol Invasions 18:299–314. https://doi.org/10.1007/s10530-015-1013-1
18. NatureServe (2021) iMapInvasives: NatureServe's online data system supporting strategic invasive species management. Arlington, Virginia. http://www.imapinvasives.org
19. NYS Department of Environmental Conservation (2015) Part 575 Prohibited and Regulated Invasive Species (6 CRR-NY 575)
20. NYS Department of Environmental Conservation (2018) New York State Invasive Species Comprehensive Management Plan. https://www.dec.ny.gov/docs/lands_forests_pdf/iscmpfinal.pdf
21. NYS Invasive Species Council (2010) A Regulatory System for Non-Native Species. https://www.dec.ny.gov/docs/lands_forests_pdf/invasive062910.pdf
22. Rapacciuolo G, Young A, Johnson R (2021) Deriving indicators of biodiversity change from unstructured community-contributed data. Oikos 130:1225–1239. https://doi.org/10.1111/oik.08215
23. Reaser JK, Simpson A, Guala GF et al (2020) Envisioning a national invasive species information framework. Biol Invasions 22:21–36. https://doi.org/10.1007/s10530-019-02141-3
24. Rew LJ, Pokorny ML (2006) Inventory and survey methods for nonindigenous plant species. Montana State University Extension
25. Simpson A, Eyler MC, Sikes D et al (2018) A comprehensive list of non-native species established in three major regions of the United States: Version 3.0, (ver. 3.0, 2020): U.S. Geological Survey data release, https://doi.org/10.5066/P9E5K160
26. U.S. Geological Survey (2021) Nonindigenous Aquatic Species Database. Gainesville, Florida. https://nas.er.usgs.gov/
27. Ward DF (2014) Understanding sampling and taxonomic biases recorded by citizen scientists. J Insect Conserv 18:753–756. https://doi.org/10.1007/s10841-014-9676-y

28. Werier D (2017) Catalogue of the vascular plants of New York State. Mem Torrey Bot Soc 27:1–542. http://www.jstor.org/stable/26279839
29. Williamson M, Fitter A (1996) The Varying Success of Invaders. Ecology 77:1661–1666. https://doi.org/10.2307/2265769
30. Young BE, Lee MT, Frey M et al (2021) Using Citizen Science Observations to Develop Managed Area Watch Lists. Nat Areas J 41:307–314. https://doi.org/10.3375/21-8

Figures

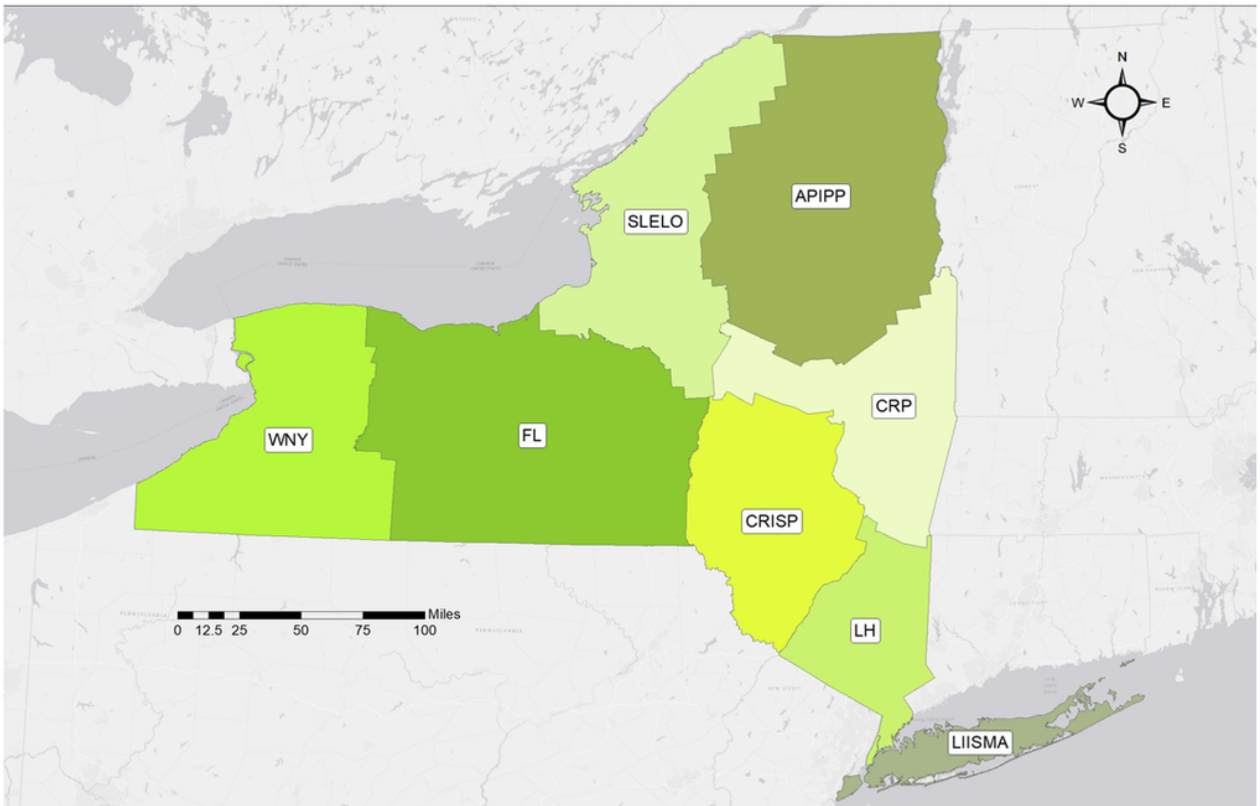

Figure 1

The regional Partnerships for Regional Invasive Species Management (PRISMs) in New York State in northeastern United States reflect differences in environmental factors and invasive species composition and history (from west to east): WNY (Western NY), FL (Finger Lakes), SLELO (Saint Lawrence and Eastern Lake Ontario), APIPP (Adirondack Park Invasive Plant Program), CRP (Capital Region PRISM), CRISP (Catskill Regional Invasive Species Partnership), LH (Lower Hudson), and LIISMA (Long Island Invasive Species Management Area)

		Difficultly of Eradication / Cost of Control Abundance (In Region plus Buffer)			
		None in Target Region	**Low** (Eradication/ Full containment may be feasible)	**Medium** (Strategic management to contain infestations and slow spread in PRISM)	**High** (Established/widespread in PRISM; only strategic localized management)
Impact (current and future)	**Very High or High**	**TIER 1** *Early Detection/Prevention* Highest level of early detection survey efforts. Should conduct delineation surveys and assign to appropriate Tier if detected. a) Inside buffer, but not in PRISM b) Outside PRISM and Buffer, but close (eastern North America) c) Far outside PRISM and buffer (not in east NA), but introduction pathway exists	**TIER 2** *Eradication* Highest level of early detection response efforts. High impact species with low enough abundance and suitable treatment method available to make eradication feasible within the PRISM. Need delineation surveys to determine extent.	**TIER 3** *Containment* Target strategic management to slow the spread, as likely too widespread for eradication, but many surrounding regions could be at risk if left unattended. For plants, use the IPMDAT. Possible eradication candidate only if adequate resources and effective control methods available.	**TIER 4** *Local Control* Eradication from PRISM not feasible; focus on localized management over time to contain, exclude, or suppress to protect high-priority resources like rare species or recreation assets. Be strategic when deciding if / where to control.
	Medium	*Evaluate (Medium Impact)* Further evaluate impacts and PRISM resources to see if the species should be assigned to one of the other lists. If this species could feasibly become high impact with climatic or other environmental changes, consider moving to the appropriate High Impact row based on abundance. If too little is known, consider moving to "Monitor".			
	Unknown	X	**TIER 5** *Monitor* Species that need more research, mapping, and monitoring to understand their invasiveness. This includes naturalized species and cultivated-only species that are known to be invasive in other regions but are not yet invasive here. Invasiveness may change with environmental or genetic changes. Should monitor populations on a regular basis to see if they are starting to become invasive and assign to appropriate Tier if invasive infestations detected.		

Figure 2

Invasive Species Tiers definitions, formalized in 2017 by the NYS invasive species network of PRISMs, state agencies, and NYNHP. Only species considered to have high negative impacts were considered for inclusion in Tiers 1 through 4. Impacts were determined by NYS invasiveness and socio-economic ranks. For species that are not ranked yet, or region-specific adjustments of state ranks are deemed necessary, expert opinion was used and documented. Low-impact species were not included since it cannot be justified to spend resources to control these. The buffer applied in the data tiers analyses for Tier 1 is 160 km surrounding the region of interest

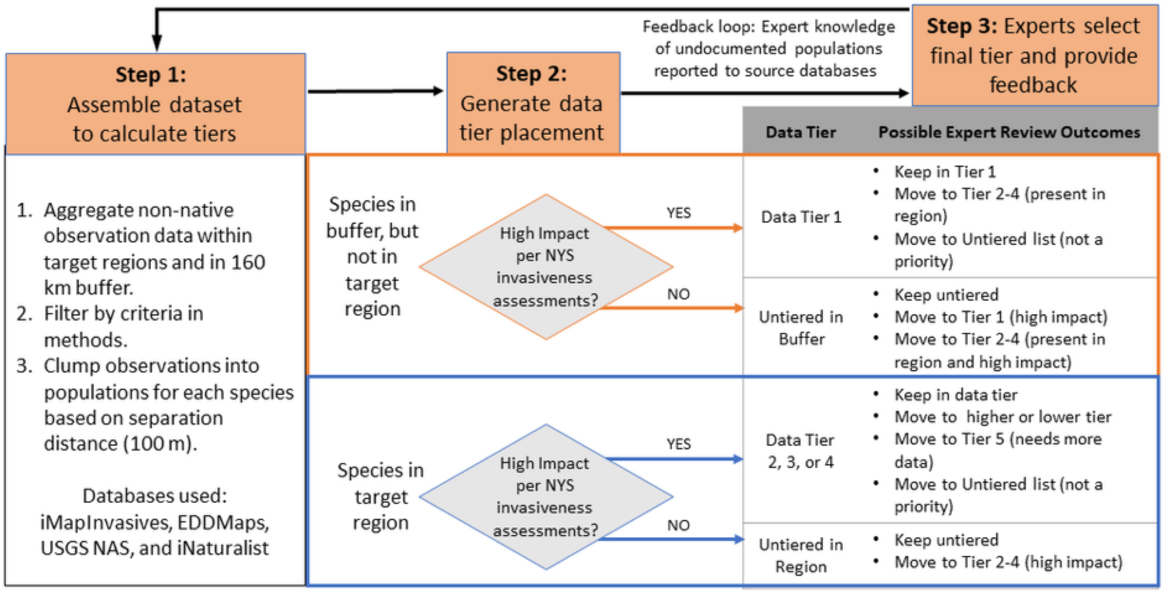

Figure 3

Generalized process of creating data tiers, which then undergo expert review

Tiers 2-4: All high impact species in PRISM

	Scientific name	Common name	Species Type	Ecol. Rank	Ecol. Score	Socio-ec. Rank	Populations	Data Tier
3	*Alliaria petiolata*	Garlic mustard	Terrestrial Plant	Very High	84	Insignificant Negative	775	4
4	*Phragmites australis ssp*	Common reed	Terrestrial Plant	Very High	92	Insignificant Positive	720	4
5	*Reynoutria japonica var*	Japanese knotweed, (	Terrestrial Plant	Very High	97.94	Insignificant Negative	518	4
6	*Ampelopsis brevipedun*	Porcelain berry	Terrestrial Plant	High	71.26		405	4
7	*Lythrum salicaria*	Purple Loosestrife	Terrestrial Plant	Very High	91	Insignificant Negative	145	4
8	*Acer pseudoplatanus*	Sycamore maple	Terrestrial Plant	High	71.11		143	4
9	*Pueraria montana var lo*	Kudzu	Terrestrial Plant	Very High	84.44		70	4
10	*Vincetoxicum louiseae*	Black swallow-wort	Terrestrial Plant	Very High	89.69		58	3
11	*Lymantria dispar*	Gypsy Moth	Terrestrial Animal	High	73		53	3
12	*Phellodendron amurens*	Amur corktree	Terrestrial Plant	High	75		32	3
13	*Rhamnus cathartica*	Buckthorn	Terrestrial Plant	Very High	81		22	3
14	*Vincetoxicum rossicum*	European Swallow-wo	Terrestrial Plant	Very High	87.63		22	3
15	*Lycorma delicatula*	Spotted Lanternfly	Terrestrial Animal	High	77		13	3
16	*Viburnum plicatum*	Japanese Snowball	Terrestrial Plant	High	78	Insignificant Negative	12	3
17	*Anthriscus sylvestris*	Wild chervil	Terrestrial Plant	High	78.75		9	2
18	*Procambarus clarkii*	Red Swamp Crayfish	Aquatic Animal	High	76	Equal outcome	8	2
19	*Actinidia arguta*	hardy kiwi, taravine, ta	Terrestrial Plant	High	75	Insignificant Positive	4	2
20	*Photinia villosa*	Oriental photinia, Ori	Terrestrial Plant	High	77.14	Insignificant Positive	4	2
21	*Pistia stratiotes*	water-lettuce, water ca	Aquatic Plant	Very High	80.65	Insignificant Negative	3	2
22	*Macleaya cordata*	Plume Poppy	Terrestrial Plant	High	75	Equal outcome	2	2
23	*Mahonia bealei*	Beale's Oregon-grape	Terrestrial Plant	High	73	Equal outcome	1	2
24	*Viviparus georgianus*	Banded Mystery Snail,	Aquatic Animal	High	78	Insignificant Negative	1	2

Tiers Divided by 33% percentiles:
Tier 4: x > 62
Tier 3: 10 - 62
Tier 2: x < 10

Figure 4

Example of Data Tiers 2 through 4 output spreadsheet for expert review. Ecol. Rank is the ecological assessment rank derived from the invasiveness assessments. Socio-ec. Rank is the socio-economic rank

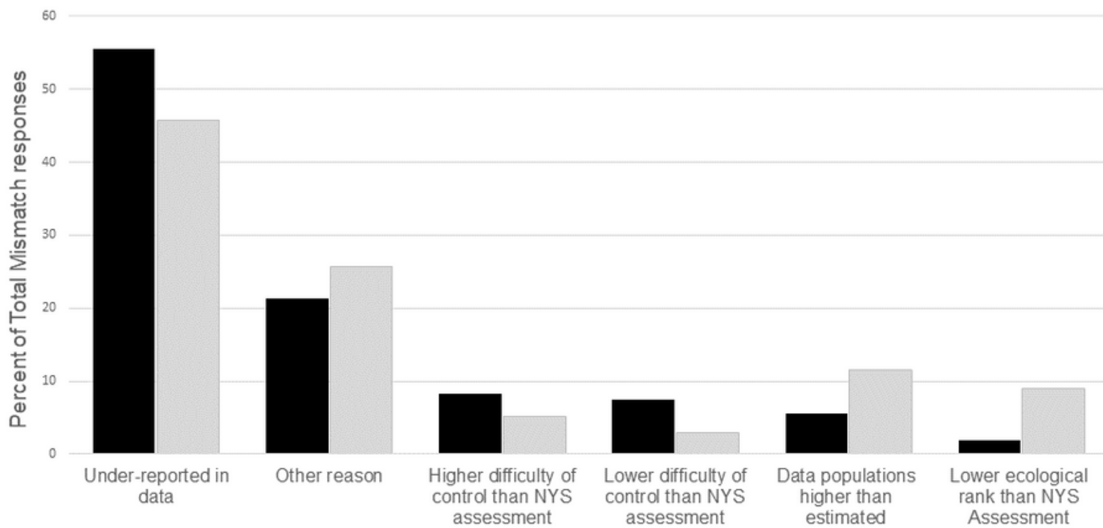

Figure 5

Percent that each reason for tier change was listed from expert review (i.e., “mismatches”), in data tiers 2 through 4 species, combined across all target regions. Black bars are responses from version 1 (with a total of 108 mismatches) and grey bars represent responses from Version 2 (269 total mismatches). Multiple reasons could be listed for a single species. In Version 2, a few reasons for change were added that were not in Version 1. To directly compare reasons for change between the two versions, this analysis used ‘Other reason’ as an umbrella term for the following reasons: ‘Lower socio-economic rank than NYS Assessment,’ ‘Species distributed over a limited geographic range,’ and ‘Species distributed over a broad geographic range’. Only four target regions participated in version 1, while version 2 of the data tiers was evaluated by all nine target regions

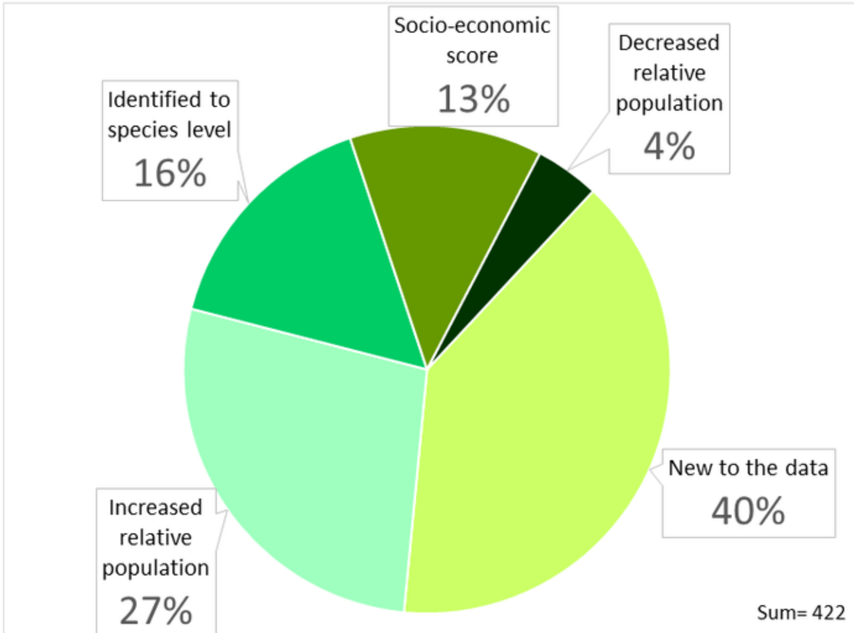

Figure 6

Frequency of causes for species tier shifts in data tiers 2 through 4 between Version 1 and Version 2. ‘New to the data’ refers to species that were previously absent in the data but were added because of newly included iNaturalist or iMapInvasives records. ‘Increased relative population’ refers to species who went from a low tier to a high tier because their population increased due to the new data, or the tier divisions shifted. ‘Identified to species level’ were species previously identified only to the genus level. ‘Socio-economic score’ are species that hadn’t been considered data tier 2 through 4 in the first version but had high or very high socio-economic scores. ‘Decreased relative population’ are species which decreased in population because of the larger separation distance or because the tier divisions shifted. Many of the 146 shifting species were counted in multiple target regions, accounting for the 422 overall shifts that the figure and associated percentages represent