Invasive species risk assessments need more consistent spatial abundance data

Bethany A. Bradley  
*University of Massachusetts Amherst*

Jenica M. Allen  
*University of New Hampshire, Durham*

Mitchell W. O’Neill  
*University of New Hampshire, Durham*

Rebekah D. Wallace  
*University of Georgia*

Charles T. Bargeron  
*University of Georgia*

*See next page for additional authors*

Follow this and additional works at: https://scholarworks.umass.edu/nrc_faculty_pubs

Part of the [Environmental Monitoring Commons](https://scholarworks.umass.edu/nrc_faculty_pubs) and the [Natural Resources and Conservation Commons](https://scholarworks.umass.edu/nrc_faculty_pubs)

Recommended Citation

https://doi.org/10.1002/ecs2.2302

This Article is brought to you for free and open access by the Environmental Conservation at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Environmental Conservation Faculty Publication Series by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
Authors

This article is available at ScholarWorks@UMass Amherst: https://scholarworks.umass.edu/nrc_faculty_pubs/403
Invasive species risk assessments need more consistent spatial abundance data

BETHANY A. BRADLEY,1,† JENICA M. ALLEN,2 MITCHELL W. O’NEILL,2 REBEKAH D. WALLACE,3 CHARLES T. BARGERON,7 JULIE A. RICHBURG,4 AND KRISTINA STINSON1

1Department of Environmental Conservation, University of Massachusetts, Amherst, Massachusetts 01003 USA
2Department of Natural Resources and the Environment, University of New Hampshire, Durham, New Hampshire 03824 USA
3Center for Invasive Species and Ecosystem Health, University of Georgia, Tifton, Georgia 31793 USA
4The Trustees of Reservations, 100 Main Street, Suite 100, Florence, Massachusetts 01062 USA


Abstract. Spatial abundance information is a critical component of invasive plant risk assessment. While spatial occurrence data provide important information about potential establishment, abundance data are necessary to understand invasive species’ populations, which ultimately drive environmental and economic impacts. In recent years, the collective efforts of numerous management agencies and public participants have created unprecedented spatial archives of invasive plant occurrence, but consistent information about abundance remains rare. Here, we develop guidelines for the collection and reporting of abundance information that can add value to existing data collection efforts and inform spatial ecology research. In order to identify the most common methods used to report abundance, we analyzed over 1.6 million invasive plant records in the Early Detection and Distribution Mapping System (EDDMapS). Abundance data in some form are widely reported, with 58.9% of records containing qualitative or quantitative information about invasive plant cover, density, or infested area, but records vary markedly in terms of standards for reporting. Percent cover was the most commonly reported metric of abundance, typically collected in bins of trace (<1%), low (1–5%), moderate (5–25%), and high (>25%). However, percent cover data were rarely reported along with an estimate of area, which is critical for ensuring accurate interpretation of reported abundance data. Infested area is typically reported as a number with associated units of square feet or acres. Together, an estimate of both cover and infested area provides the most robust and interpretable information for spatial research and risk assessment applications. By developing consistent metrics of reporting for abundance, collectors can provide much needed information to support spatial models of invasion risk.

Key words: abundance; citizen science; Early Detection and Distribution Mapping System (EDDMapS); field data; field survey; invasive plant; species distribution model.

Received 11 May 2018; accepted 14 May 2018. Corresponding Editor: Debra P. C. Peters.
Copyright © 2018 The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
† E-mail: bbradley@eco.umass.edu

INTRODUCTION

With readily available and inexpensive global positioning system (GPS) units (including smartphones), spatial data describing the range and abundance of species are increasingly being collected and archived (e.g., Bargeron and Moorhead 2007). Spatial data are invaluable for scientific applications, which often leverage contributed datasets to understand the distributions of species (e.g., Franklin 2010). For invasive plants, contributed spatial data can inform habitat models and identify invasion risk (e.g., Bradley 2013). A recent analysis of invasive plant risk
in the United States identified hotspots of vulnerability based on spatial occurrences of nearly 900 species, which were derived primarily from contributed datasets (Allen and Bradley 2016). Similarly, at landscape scales, occurrence datasets are often used to understand how invasive plants are influenced by local disturbance and development (Vilà and Ibáñez 2011). Collectively, these spatial analyses improve our understanding of the climate conditions, land cover, and disturbance regimes that may make ecosystems more susceptible to invasive species establishment.

However, establishment, or the likelihood that an invasive species will persist in a given location, is only a portion of overall invasion risk (Blackburn et al. 2011). Of equal importance is the likelihood that a species will become abundant in a given location (Parker et al. 1999). For example, changes in the abundance of invasive species (invader demography) are central to most conceptual frameworks that aim to understand invasion ecology (Gurevitch et al. 2011). The size of invasive species populations is hypothesized to drive all other invasion processes and underlie interactions with native species and communities (Gurevitch et al. 2011). As a result, recommended metrics for measuring degree of invasion include not only presence/absence, but also abundance (e.g., Guo et al. 2015). Invasive species that are likely to become abundant in a given location are also of the highest priority for management (Hulme 2006, McDonald et al. 2009). Despite the importance of abundance for understanding invader demography and predicting risk, the vast majority of spatial models of invasion risk currently rely on occurrence data alone (Bradley 2013, but see Curtis and Bradley 2015, Kulhanek et al. 2011, Ibáñez et al. 2009).

The reliance of spatial models on occurrence data stems from the widespread availability of these data and the relative lack of accessible and spatially extensive abundance data. Datasets such as the USDA Forest Service’s Forest Inventory and Analysis and the National Park Service’s Inventory and Monitoring (I&M) systematically measure invasive plant abundance, but focus on forest habitat and the National parks, respectively (Fancy et al. 2009, Oswalt et al. 2015), and can be challenging to access. In contrast, curated museum and herbarium records are easy for scientists to access (e.g., downloadable from the Global Biodiversity Information Facility), but these data do not report abundance. Similarly, invasive species databases contributed by managers and individuals have a primary goal of recording occurrences. The lack of spatially extensive estimates of invasive plant abundance creates a challenge for spatial risk assessment.

Spatial models built from occurrence data often include a score for each spatial location (e.g., a numeric suitability value ranging from 0 to 1), and several researchers have asked whether this suitability for occurrence can be used to estimate abundance (Weber et al. 2017). In tests of modeled suitability for occurrence vs. measured abundance, relationships tend to be positive and significant when measured abundance includes absence data (Pearce and Ferrier 2001, VanDerWal et al. 2009, Weber et al. 2017). However, when measured abundance includes only locations where a species is present, relationships between suitability for occurrence and abundance are rarely significant (Pearce and Ferrier 2001, Sakai et al. 2001, Jiménez-Valverde et al. 2009, Filz et al. 2013, Bradley 2016). In other words, spatial models of suitability for occurrence can differentiate between presence and absence, but, given species presence, cannot differentiate between low, medium, or high abundance.

The poor relationship between modeled suitability for occurrence and measured abundance may be due to spatial biases in the way that occurrence data are collected. Marvin et al. (2009) tested the relationship between invasive plant abundance reported by managers and the frequency of point occurrences found in contributed datasets for three problematic species in the southeast United States. Marvin et al. (2009) found no relationship between reported abundance and frequency of point occurrences for kudzu (Pueraria montana var. lobata) and cogongrass (Imperata cylindrica) and a counter-intuitive negative relationship for privet (Ligustrum spp.). That is, contributors to spatial databases collected the same number or even more occurrence points in areas where the invasive species was rare than where the species was abundant. Cross et al. (2017) found a similar result for 13 problematic invasive plants in the northeast United States and hypothesized that the lack of relationship between occurrence frequency and reported abundance could stem from a focus on early...
detection and rapid response to emerging invasions. Lacking abundance information, the spatial pattern of occurrence data collection may cause models to misidentify landscapes with the potential for abundant infestations. As a result, spatial occurrence data alone are unlikely to be useful for predicting invasive plant abundance and ecological impacts.

In addition to lack of abundance information, lack of absence data is also an ongoing problem for modeling habitat suitability. Species distribution models focused on presence or potential occurrence of a species are much more accurate when absence data are available (Václavík and Meentemeyer 2009, Guillera-Arroita et al. 2015). Mapped predictions of species potential occurrence based only on presence points tend to over-predict area of occupancy, potentially increasing survey effort and associated costs (Guillera-Arroita et al. 2015). Absence data are also important for identifying the lower bounds of abundance in spatial abundance models. Similarly, data recording changes from absence to presence through time could improve our understanding of species dispersal and range dynamics.

In order to more effectively model and predict the spatial patterns of invasion risk, better information about invasive species abundance is needed (Howard et al. 2014). Fortunately, many contributors to spatial databases are already reporting abundance information. Cross et al. (2017)’s survey of spatial data for 13 invasive plants found that 30% of contributed spatial records included some form of abundance information, but inconsistent reporting of abundance data remains a major challenge for scientific assessment. Here, we analyze invasive plant records contributed to the Early Detection and Distribution Mapping System (EDDMapS; Bargeron and Moorhead 2007), an aggregate database widely used by management agencies and public participants. By analyzing abundance data already collected by managers, we aim to identify commonly used metrics for reporting abundance as well as potential pitfalls that make contributed data harder to interpret. This analysis provides recommendation for future data collection that will increase the consistency of abundance information in contributed datasets which, in turn, will improve the accuracy and applicability of spatial risk assessments.

**METHODS**

*Early detection and distribution mapping system*

The University of Georgia’s Center for Invasive Species and Ecosystem Health (Bugwood) launched EDDMapS in 2005 as a means for natural resource managers and public contributors to report spatial occurrences of invasive plants. The initial focus of EDDMapS was on early detection of invasive species in the southeast United States, with contributors uploading spatial data through the EDDMapS website. Since then, EDDMapS has grown to encompass the full United States and Canada and has over 14,000 individual contributors as well as partnerships with state and federal agencies (e.g., USDA Forest Service, National Park Service, Florida Natural Areas Inventory Database, Alaska Exotic Plants Information Clearinghouse, and Texas Intruders) that use EDDMapS as a permanent repository. EDDMapS also archives data from projects that are no longer active such as the Invasive Plant Atlas of New England, What’s Invasive, and the Southwest Exotic Plant Mapping Program. Recently, EDDMapS’ data collection has expanded to accept spatial data collected through iPhone and Android apps, including IveGot1 (Wallace et al. 2016), EDDMapS West, Outsmart Invasives (Starr et al. 2014), and EDDMapS Midwest. These smartphone apps have been downloaded over 200,000 times, and smartphone users have contributed over 93,000 unique reports.

Given the variety of partner institutions and large number of contributors to EDDMapS, there are a range of methods used to collect and report abundance data. Different agencies and individuals have different goals in terms of data collection, which leads to a variety of measurement protocols, measurement units, and ancillary information. In order to accommodate the range of data and formats, bulk data entry into the EDDMapS database is done manually, and all data are reviewed for completeness and consistency as they are entered. Nonetheless, consistency is a major concern, and the more that standard protocols and assessments can be adopted by contributors, the more interpretable and useful the database will ultimately be. Consistency is particularly important for information such as abundance, which extends the scope of the initial EDDMapS goals.
All data contributed to EDDMapS are evaluated for completeness to ensure that minimum standards are met (reporter, observation date, location, and species) as agreed upon by the North American Invasive Species Management Association. Ancillary data, including abundance and/or absence information, are added to the database in either new or existing data columns. Bugwood staff communicate with data contributors to ensure that the data are correctly interpreted and formatted. However, due to the range of contributors, there are multiple data columns within the database potentially containing some form of abundance information.

**Data analysis**

We analyzed data columns in the EDDMapS database representing cover, stem count, and density of invasive species as well as area of infestation. We also assessed commonly reported information in the comments column for information about abundance. A description of the data columns and definitions in the EDDMapS database is presented in Table 1. We analyzed reported abundance information for all species in the EDDMapS database. For each data column, we summarized the data values to identify commonly reported quantitative and qualitative metrics of abundance. We used this information to develop recommendations for collecting and reporting spatial abundance data.

**RESULTS**

**Summary of data from EDDMapS**

We analyzed all records for invasive plants in the EDDMapS database as of 18 October 2016. This dataset consisted of 1,663,768 unique spatial records associated with 2022 plant species. The majority of these records (873,507; 52.5%) contained some form of information for abundance (density, stem count, percent cover, or number observed). A larger majority of records (980,647; 58.9%) included abundance and/or infested area. There are strong spatial patterns in both occurrence and abundance data within the EDDMapS database (Fig. 1). Some occurrence data were reported for most states, although occurrence data were sparser in midwest states, California, and Washington due to lack of collection and/or archiving in different spatial databases (e.g., California primarily archives data to CalFlora, New York primarily archives to iMapInvasives). EDDMapS is increasingly implementing data sharing agreements with new agencies, so occurrence data are likely to expand. Abundance data appear more strongly defined by state boundaries, with only a handful of agencies consistently including some form of abundance information (Fig. 1). Absence data are rarely reported (Fig. 1).

**Percent cover**

Percent cover was the most commonly reported metric of invasive plant abundance. There were a total of 445,812 records (26.8%) containing information about invasive plant percent cover. Cover data were commonly reported as integers (Appendix S1), but were also commonly reported as ranges of quantitative cover values (e.g., 1–5%) or qualitative descriptions (low, moderate, high). The percent cover bins presented in Table 2 are currently being adopted in smartphone apps linked to EDDMapS and would provide a useful metric of abundance if integer estimates cannot be collected.

<table>
<thead>
<tr>
<th>Column name</th>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>percentcover</td>
<td>Cover</td>
<td>Percent cover of invasive species</td>
</tr>
<tr>
<td>NumberObserved</td>
<td>Count</td>
<td>Number or estimation of number of subjects observed</td>
</tr>
<tr>
<td>stemcount</td>
<td>Count</td>
<td>Approximate number of stems for infestation, range of numbers</td>
</tr>
<tr>
<td>Grossarea</td>
<td>Area</td>
<td>Entire area that a large or discontinuous infestation covers</td>
</tr>
<tr>
<td>Grossareaunits</td>
<td>Area</td>
<td>Unit of measure (acres, sq feet, etc.)</td>
</tr>
<tr>
<td>Infestedarea</td>
<td>Area</td>
<td>Actual amount of infested area within the gross area</td>
</tr>
<tr>
<td>Infestedareaunits</td>
<td>Area</td>
<td>Unit of measure (acres, sq feet, etc.)</td>
</tr>
<tr>
<td>Comments</td>
<td>General</td>
<td>Anything that is relevant to the subject, environment, mapping</td>
</tr>
</tbody>
</table>

*Note: EDDMapS, Early Detection and Distribution Mapping System.*
There were a surprisingly high number of integer values between 1 and 3 in the EDDMapS database. It appears likely that some agencies use these integers to indicate ranges of cover values (e.g., 1 = <1%, 2 = 1–5%, 3 = 5–25%). While this practice makes sense at a local scale, when the data are compiled into a large repository like EDDMapS, the underlying metadata information about how to interpret values is often lost or difficult to access. As a result, rankings (1,2,3) could easily be misinterpreted as percent cover values (1%, 2%, 3%). Similarly, quantitative values that are reported as ranges (e.g., 1–5%) are more read-
ily interpretable for research than if those ranges are transformed into a mean (3%), which implies a higher level of specificity than was actually measured. Reporting ranges or including a % symbol after percent cover values would make this information easier to interpret correctly.

Misinterpretation is also possible when cover values are reported as proportional cover rather than percent cover (i.e., numbers between 0 and 1 rather than 0 and 100). Lacking metadata on the reporting methods, it is unclear whether a value of 0.5 is a percent cover estimate corresponding to <1% or a proportional cover estimate corresponding to 50%. Percent cover estimates reported as integers between 0% and 100% or binned quantitative/qualitative ranges like the ones suggested in Table 2 are less likely to be misinterpreted by researchers.

Table 2. Commonly reported percent cover values in the EDDMapS database.

<table>
<thead>
<tr>
<th>Quantitative cover (%)</th>
<th>Qualitative cover</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>Trace</td>
</tr>
<tr>
<td>1–5</td>
<td>Low</td>
</tr>
<tr>
<td>5–25</td>
<td>Moderate</td>
</tr>
<tr>
<td>25–100</td>
<td>High</td>
</tr>
</tbody>
</table>

Note: EDDMapS, Early Detection and Distribution Mapping System.

There were a surprisingly high number of integer values between 1 and 3 in the EDDMapS database. It appears likely that some agencies use these integers to indicate ranges of cover values (e.g., 1 = <1%, 2 = 1–5%, 3 = 5–25%). While this practice makes sense at a local scale, when the data are compiled into a large repository like EDDMapS, the underlying metadata information about how to interpret values is often lost or difficult to access. As a result, rankings (1,2,3) could easily be misinterpreted as percent cover values (1%, 2%, 3%). Similarly, quantitative values that are reported as ranges (e.g., 1–5%) are more readily interpretable for research than if those ranges are transformed into a mean (3%), which implies a higher level of specificity than was actually measured. Reporting ranges or including a % symbol after percent cover values would make this information easier to interpret correctly.

Misinterpretation is also possible when cover values are reported as proportional cover rather than percent cover (i.e., numbers between 0 and 1 rather than 0 and 100). Lacking metadata on the reporting methods, it is unclear whether a value of 0.5 is a percent cover estimate corresponding to <1% or a proportional cover estimate corresponding to 50%. Percent cover estimates reported as integers between 0% and 100% or binned quantitative/qualitative ranges like the ones suggested in Table 2 are less likely to be misinterpreted by researchers.

Stem count

There were a total of 113,795 records (6.8%) of stem count (reported as either stemcount or NumberObserved) in the EDDMapS database. The vast majority of these records were quantitative (107,992), and quantitative values were typically presented as ranges (Appendix S1). Although many such ranges are reported in EDDMapS,
breadth of values of stem count suggest that a log scale, as those shown in Table 3, would best encompass the data. Qualitative stem counts in the EDDMapS database tended to be challenging to interpret (e.g., common, uncommon). If qualitative metrics are preferred for reporting, we suggest those listed in Table 3, which provide more levels of information about abundance. Note that qualitative metrics do not necessarily correspond to quantitative stem counts.

Infested area

Estimating the spatial extents of an infestation is a critical part of understanding local abundance. A total of 840,764 records (50.5%) contained information about infested area as well as a unit of infested area. By far the most commonly reported area units were acres or square feet. Practitioners should use whatever area units are most intuitive to them to estimate, as it is easy to standardize units later for research purposes. Infested area is defined in the EDDMapS database as the total extent of the infestation, whereas gross area delimits the extents within which the invasive species is found and/or the total area of the surveyed park or management area (Table 1). Only 198,468 points (24% of those reporting infested area) also reported gross area. As a result, it is possible that some of the records of infested area were instead reporting the extents of the survey or management area (gross area), which would lead to an overestimate of the total extents of infestation. For research purposes, creating unique records for the locations where data were collected is much more useful than aggregated information across a larger management area because aggregation removes smaller-scale spatial details.

Combining infested area with an estimate of stem count or percent cover is critical for understanding overall abundance and provides the highest quality data for research. A total of 366,706 records (22% of the EDDMapS database) included both infested area and information about percent cover or stem count. Infested area and percent cover are each commonly reported individually. Including both pieces of information in assessments of invasion would greatly improve the interpretability of these data.

Table 3. Suggested quantitative and qualitative estimates of stem count.

<table>
<thead>
<tr>
<th>Quantitative stem count</th>
<th>Qualitative stem count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Single plant</td>
</tr>
<tr>
<td>2–10</td>
<td>Scattered plants</td>
</tr>
<tr>
<td>11–100</td>
<td>Scattered dense patches</td>
</tr>
<tr>
<td>101–1000</td>
<td>Dominant cover</td>
</tr>
<tr>
<td>&gt;1000</td>
<td>Dense monoculture</td>
</tr>
</tbody>
</table>

Contributor comments

A total of 573,798 records (34.5%) contained additional information in the comments section. Comments included information about infested area, stem count, and percent cover and included both qualitative and quantitative information that would be useful for researchers interested in invader abundance. However, 182,884 comments were unique. As a result, it is challenging and time-consuming to use comments as metrics of abundance. Abundance information is more readily accessible for research purposes when specific data entry fields for abundance are available rather than relegating this information to a comment or other category.

Absence data

Absence data are rarely collected and reported to EDDMapS. The option to report absences as negative occurrences has recently been added as a web and app reporting option, so the capability to archive absences is fully operational. A total of 52,297 absence records have been reported (3.1% of the EDDMapS database). While rare, absence information is extremely valuable for spatial risk assessments, which typically provide relative suitability scores, but not probabilities of occurrence because absence data are unavailable (Merow et al. 2013). In addition to their use in spatial models (Ibáñez et al. 2009b, Václavík and Meentemeyer 2009, Guillera-Arroita et al. 2015), absence data are important for constraining when invasions start (e.g., the length of lag time between establishment and population expansion; Crooks et al. 1999) as well as rates of dispersal across the landscape. Although absence data could be reported as 0 stem count, infested area, or percent cover, reporting absence in this way is potentially unusable because some contributors use zeros to indicate no data rather than a true absence of the species. To ensure that
absence data are correctly identified and interpreted, we suggest including a separate column where absences are labeled using text (negative/absent) rather than a numeric zero. Currently, EDDMapS compiles information about presence and absence into a single status column using the following text descriptions: (1) positive/found, (2) treated/managed, (3) eradicated, and (4) negative/absent. Absence data would be most useful to research if an estimate of the total area (gross area) surveyed were also included. EDDMapS currently compiles total survey area in two data columns describing gross area and gross area units. We strongly encourage managers to consider adding a negative/absent description coupled with an estimate of gross survey area to their data collection.

**DISCUSSION**

Although EDDMapS was initially designed as an aggregate database for invasive species occurrence information, the majority of records (58.9%) also contain information about abundance and/or infested area. Thus, a remarkable amount of abundance information is already being collected and compiled. This information is a boon for research, potentially enabling invasion risk assessments to move beyond modeling distributions to include prediction of where abundant infestations and associated ecological impacts are likely. However, abundance as currently reported comes in a variety of formats that can be hard to interpret or even unusable for research. The more consistently abundance data together with an estimate of infested area are collected and reported, the more directly useable these data will be for research.

As our summary of the EDDMapS database reveals, contributors are already moving toward standardized range bins for percent cover and stem count (Tables 2 and 3). These range bins are also being adopted in smartphone apps created by EDDMapS developers. Although integer estimates of percent cover or stem count represent the highest quality of abundance information for research, standardized range bins are still extremely useful. For example, habitat suitability models trained on locations of the highest cover show a much stronger relationship with continuous percent cover than models built on all occurrence data (Estes et al. 2013, Bradley 2016). Similarly, the availability of multiple range bins allows for the development of ordered regression models (e.g., Guisan and Harrell 2000), which can be used to estimate increasing ranks of abundance. Contributors to EDDMapS already spend a lot of effort reporting various metrics of abundance and adding comments about sites. This is particularly true for stem count data, which can be very time-consuming to collect. Directing this effort toward standardized estimates of percent cover and infested area instead would enable easier data entry while remaining scientifically useful.

Despite the increasing amount of abundance information reported in EDDMapS, absence data remain extremely rare (Fig. 1C). Early Detection and Distribution Mapping System smartphone apps increasingly contain a negative/absent option with the possibility of listing several species at once. Agencies that send data to EDDMapS in bulk should also consider reporting absence of species within their survey or management area. Increasing the availability of absence data will allow researchers to create spatial models that are less likely to overpredict the species’ potential range (Václavík and Meentemeyer 2009, Guillera-Arroita et al. 2015). Improved maps of potential range will reduce the spatial extents of necessary monitoring effort and associated costs (Guillera-Arroita et al. 2015).

In addition to the need to expand absence information and make abundance data collection more consistent, contributors should also consider undersampled regions (Fig. 1). While many states report both occurrence and abundance to EDDMapS, others report primarily occurrence information or have low amounts of reporting overall. Spatial models rely on the assumption that occurrence data are representative of all suitable environmental conditions (Hutchinson 1957), but invasion ecology suffers from well-known spatial biases in data collection (Pyšek et al. 2008), which can reduce the reliability of invasion risk assessments. Even in well-sampled regions, the focus of data collection on early detection and rapid response could create spatial biases. For example, contributed datasets tend to undersample the most extreme environments.
abundant infestations (Marvin et al. 2009, Cross et al. 2017) and rarely include absence data. The lack of both types of these data can also lead to less reliable risk assessments. State and regional management agencies that currently collect occurrence data should consider adding abundance to their data collection protocols and include data collection in areas of absence as well as heavy infestations.

Finally, with over 14,000 agencies and individuals archiving data with EDDMapS, contributors should assume that any metadata pertaining to data collection or reporting will be lost or unavailable to the end user. For example, several data columns contained an unusually high number of integers between 1 and 3, suggesting that some contributors use numbers to represent bins (e.g., 1 = low, 3 = high) rather than reporting qualitative abundance estimates or quantitative ranges (Tables 2 and 3). A text description or range of abundance values is much less likely to be incorrectly interpreted as a numeric estimate of percent cover or stem count. Similarly, a cover category recorded as percent cover (0–100%) rather than proportional cover (0–1) is much less likely to be misinterpreted. Lastly, including units along with numeric values with each data entry is essential for correct interpretation. Units could be reported in their own data field or could be appended to a numeric value (10% is more clear than 10). Contributed data that are as intuitive and interpretable as possible will be the most readily useable for research.

CONCLUSIONS

Contributed spatial observations from managers and the public are becoming an important resource for scientific risk assessments and invasive species research. But, the addition of consistently reported abundance information would greatly improve existing risk assessments, enabling scientists to model the potential magnitude of infestation as well as the overall distribution. Abundance information that includes an estimate of percent cover combined with an estimate of infested area would provide the best information for research. Similarly, locations where a species is absent combined with an estimate of the survey area would provide a critical missing piece in existing spatial models. We encourage individual contributors and management agencies to adopt standardized collection methods like the ones outlined here.

ACKNOWLEDGMENTS

BAB, JA, and MWO were supported by the National Science Foundation Award BCS 1560925. BAB was also supported by the National Institute of Food and Agriculture, U.S. Department of Agriculture, the Massachusetts Agricultural Experiment Station, and the Department of Environmental Conservation under Project Number MAS00033.

LITERATURE CITED


Curtis, C. A., and B. A. Bradley. 2015. Climate change may alter both establishment and high abundance of red brome (Bromus rubens) and African Mustard (Brassica tournefortii) in the semiarid Southwest United States. Invasive Plant Science and Management 8:341–352.

Estes, L. D., B. A. Bradley, H. Beukes, D. G. Hole, M. Lau, M. G. Oppenheimer, R. Schulze, M. A.


VanDerWal, J., L. P. Shoo, C. N. Johnson, and S. E. Williams. 2009. Abundance and the environmental

**Supporting Information**

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.2302/full