



The invasive plant data landscape: a synthesis of spatial data and applications for research and management in the United States

Emily J. Fusco · Evelyn M. Beaury · Bethany A. Bradley · Michelle Cox · Catherine S. Jarnevič · Adam L. Mahood · R. Chelsea Nagy · Ty Nietupski · Jessica E. Halofsky

Received: 27 September 2022 / Accepted: 27 February 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

Context An increase in the number and availability of datasets cataloging invasive plant distributions offers opportunities to expand our understanding, monitoring, and management of invasives across spatial scales. These datasets, created using on-the-ground observations and modeling techniques, are made both for and by researchers and managers.

Objectives The large number and variety of data types and associated datasets can be difficult to navigate, require high levels of data literacy, and can overwhelm the intended end-users. By providing a synthesis of available data types and datasets, this work may facilitate data understanding and use among researchers and managers.

Methods We synthesize types of invasive plant distribution data sources, highlighting publicly available datasets and their potential applications and limitations for research and management.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10980-023-01623-z>.

E. J. Fusco (✉)
ORISE Fellow Hosted at the U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Olympia, USA
e-mail: efusco@uw.edu

E. M. Beaury
High Meadows Environmental Institute and Ecology and Evolutionary Biology, Princeton University, Princeton, USA

B. A. Bradley
Department of Environmental Conservation, University of Massachusetts Amherst, Amherst, USA

M. Cox
U.S. Department of Agriculture, Forest Service, Northern Region, Missoula, USA

C. S. Jarnevič
U.S. Geological Survey, Fort Collins Science Center, Fort Collins, USA

A. L. Mahood · R. C. Nagy
Earth Lab, University of Colorado Boulder, Boulder, USA

A. L. Mahood
U.S. Department of Agriculture, Agricultural Research Service, Water Resources, Fort Collins, USA

T. Nietupski
US Department of Agriculture, Forest Service, Geospatial Technology and Applications Center, Fort Collins, USA

T. Nietupski
ORISE Fellow Hosted at the U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Corvallis, USA

J. E. Halofsky
Northwest Climate Hub and Western Wildland Environmental Threat Assessment Center, U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Olympia, USA

Results Eight data types and their potential applications for research and management are described. We also describe gaps in current invasive species distribution data usability and outline a path forward for improving the use of invasive plant data in future research and management.

Conclusions Accessible and usable invasive plant spatial data are needed for developing landscape scale analysis and management plans. By synthesizing the invasive plant data available, with examples and limitations for application, this work will serve as a guide to facilitate appropriate and efficient data choices in current and future research and management.

Keywords Invasive plants · Spatial data · Land management · Data use · Data accessibility · Data types

Introduction

Negative impacts from invasive plants have been demonstrated across spatial scales and include altering biodiversity (Powell et al. 2011; Vilà et al. 2011), disturbance (D'Antonio and Vitousek 1992), and biogeochemical cycles (e.g., Castro-Diez et al. 2014; Mahood et al. 2022). Non-native invasive species are extensive across geographies (Early et al. 2016; Pyšek et al. 2020) and are considered a major driver of global change (IPBES 2019). Invasive plants also interact with other global change drivers, in some cases compounding negative impacts (Lopez et al. 2022), making impact mitigation a high priority for protecting ecosystems and economies (IPBES 2019). A rapidly expanding and sprawling invasive plant spatial data landscape can help inform robust prevention and mitigation strategies but could also stymie success if these data are not accessible or interpretable. Consequently, a better understanding of the spatial data landscape could promote more efficient and effective invasive plant research and management.

Many aspects of plant invasion research and management depend on knowledge of invasive plant localities. In research, invasive plant distribution data can shed light on global patterns of invasion (e.g., Early et al. 2016), ecological niche theory (e.g., Bradley et al. 2015), dispersal dynamics (e.g., Flory and Clay 2009),

and potential distributions under climate change (e.g., Allen and Bradley 2016). In management, interventions early in the invasion process are the most effective when considering cost and efficacy of treatment (Keller et al. 2007; Ahmed et al. 2022), and assessing which species to regulate to prevent initial introductions often requires knowledge of current distributions (Bradley et al. 2022). Once plants are introduced, Early Detection and Rapid Response (EDRR) requires that managers are made aware of new populations so they can be treated before the population spreads (Martinez et al. 2020). The breadth of these research and management issues necessitates a variety of high-quality datasets of invasive plant distributions.

In recent years, several factors have facilitated a rise in the variety and availability of invasive plant spatial data, including an increase in ecological data sharing (Michener 2015), opportunities for public participation in invasive plant data acquisition (Howard et al. 2022), and increased availability of remotely sensed data products (Vaz et al. 2018). However, many of these datasets are imperfect for addressing conservation goals and may lack appropriate spatial resolution, be out of date, or have high uncertainty regarding accuracy (Joppa et al. 2016). Further, these datasets can simply be too difficult for end users to find and use (Culina et al. 2018; Sofaer et al. 2019). Researchers and managers observing a well-populated spatial data landscape could achieve their respective goals most effectively if they have a clear understanding of the applications and limitations of these datasets, as well as if and how they relate to one another.

To address the need for invasive plant spatial data across research and management communities, a multitude of data sources have been created (Joppa et al. 2016). Here we describe the spatial data landscape for invasive plants in the United States. We identify four types of primary data sources which are collected through on-the-ground observations: community science, inventory and monitoring, vegetation community, and original research data. We also describe two forms of modeled datasets: remotely mapped distributions and potential distributions. Finally, we describe a third type of data source, data repositories, which consist of data aggregates and data archives (Table 1). For each category of invasive plant spatial data, we identify the associated originators and users, and discuss how these datasets are interrelated. We

also include examples to describe demonstrated and potential data uses and limitations (Table 2). Finally, we discuss the ways that data access and use can be facilitated among researchers and managers.

Primary data

Invasive species data in this category are collected through on-the-ground observations and consist of community science, inventory and monitoring, vegetation community, and original research data (Table 1). These data may be collected by managers, researchers, professionals, or non-professionals and may encompass any scale, from a single management area to the globe.

Community science

Community science, or citizen science, data refers to data contributions made by volunteers in a non-professional capacity. Many members of the public are motivated to contribute to community science initiatives by a desire to facilitate scientific research (Maund et al. 2020). There are several apps available

for the public to report invasive species, which typically integrate data collection, reporting, identification, and user engagement into their interface (Howard et al. 2022). Apps for community scientists may prompt users to contribute information such as occurrence, abundance, absence, and photographs, along with a georeferenced location (Howard et al. 2022). Depending on where the observation took place, and to whom it is being reported, users may be encouraged to report sightings via a state reporting form or through an app.

Several popular apps and programs are used by community scientists to collect invasive plant data and can be chosen based on their needs. For example, Wild Spotter (wildspotter.org/), which feeds into the Early Detection and Distribution Mapping System (EDDMapS; a separate mobile app and a data repository; see section on “Aggregates”), is limited to data collection on participating federal and tribal lands, and community scientists can contribute to organized projects. In some states, community scientists may report sightings to government departments through an online web form or state app (e.g., Washington, invasivespecies.wa.gov/report-a-sighting/, shared with EDDMapS). Alternatively, iNaturalist (www.inaturalist.org) users can contribute and download

Table 1 Data types and categories with examples

Data type	Data category	Example datasets discussed
Primary data	Community science	WildSpotter (see EDDMapS/Aggregates); iNaturalist (inaturalist.org)
	Inventory and monitoring	iMapInvasives/EDDMapS (see Aggregates); Threatened, Endangered, Sensitive Plant—Invasive Species (TESP-IS; fs.usda.gov/about-agency/enterprise-data-warehouse)
	Vegetation community	National Ecological Observatory Network (NEON; data.neonscience.org/), Forest Inventory Analysis (FIA; fs.usda.gov/research/products/data-and-tools/forestinventorydata)
	Research	Munro et al. in review (https://doi.org/10.7275/m1an-m171), (Nagy et al. 2021a) (https://doi.org/10.5061/dryad.4mw6m9082)
Modeled data	Remotely mapped distributions	Bradley (2017) (https://doi.org/10.7275/R5XW4GZR); (Nietupski 2021) (https://doi.org/10.2737/RDS-2022-0011)
	Potential distributions	Invasive Species Habitat Tool (INHABIT; gis.usgs.gov/inhabit/)
Data repositories	Aggregates	Early Detection and Distribution Mapping System (EDDMapS; eddmapp.org); iMapInvasives (imainvasives.org); Standardized Plant Community with Introduced Status database (SPCIS; https://doi.org/10.6084/m9.figshare.19593373.v7)
	Archives	Environmental Data Initiative (edirepository.org/); Dryad (datadryad.org/); Knowledge Network for Biocomplexity (knb.ecoinformatics.org/); Vegbank (vegbank.org); U.S. Geological Survey Science Data Catalog (sciencebase.gov/catalog/); USDA Forest Service Research Data Archive (fs.usda.gov/rds/archive/); Shyvers et al. (2022) (https://doi.org/10.5066/P9VW97AO), Simpson et al. (2020) (https://doi.org/10.5066/P9CNVBYR)

Table 2 Best uses, benefits, and limitations of each data type

Data category	Scale	Best uses	Benefits	Limitations
Community science	Local to regional	Management: Real-time reporting for Early Detection and Rapid Response (EDRR) Research: Creating current and future distribution models	Cost effective	Some records may not be verified; Spatial biases based on where people visit
Inventory and monitoring	Local to landscape	Management: Tracking management priorities and accomplishments within group Research: Supplement open-access data sets, Answer questions related to treatment	Collected by professionals or trained personnel; May include ancillary data	Variable data collection methods and storage; Availability varies; Update frequency varies
Vegetation community	Local to regional	Management: Supplement inventory and monitoring data, Understanding vegetation community susceptibility to invasion Research: Answering community invasion questions, answering temporal invasion questions if repeat sampling occurs	Collected by professionals or trained personnel; Includes rich ancillary data	Data collection is not targeted to invasive plants
Research	Local to landscape	Management: Supplement inventory and monitoring data, Consider treatment efficacy in specific localities Research: Supplement data for creating models, Reviews, Meta-analyses	Collected by professionals or trained personnel; Includes rich ancillary data	Difficult to find and extract, Few data-points available, Spatial biases near research institutions
Remotely mapped distributions	Landscape to regional	Management: Long-term monitoring, Land management planning, Prioritizing areas for monitoring and treatment, Detecting impacts Research: Understanding mechanisms, processes, patterns, and invasion impacts over broad spatial scales	Provide estimates of invasion over large, sometimes difficult to survey, areas	Not all invasive plant species are suitable for mapping, Lack of technology transfer
Potential distributions	Landscape to regional	Management: Prioritizing areas for surveys, monitoring, and treatment; Supporting regulatory decisions, Develop climate-informed watchlists Research: Answering biogeography questions, Predicting species range shifts under climate change	Provide estimates of invasion over large, sometimes difficult to survey, areas; Provide projections for invasion under climate change scenarios	Lack of accuracy in predicting species abundance, Difficult to interpret, Lack of technology transfer

Table 2 (continued)

Data category	Scale	Best uses	Benefits	Limitations
Aggregates	Landscape to regional	Management: Share invasive plant occurrence and abundance records, Inform EDRR activities Research: Inform data models, Inform initial field research planning, Identify local experts	Low barriers to access for many databases, Broad geographic scope	High variability in data collection protocols, Some records may not be verified
Archives	Local to regional	Management: Curated archives may provide highly usable options and instruction for use Research: Support reproducible science, Store data	Curated archives provide more comprehensive data options, Other benefits are consistent with the associated original datasets	Decentralized, May be difficult to find relevant data, Other limitations are consistent with the associated original datasets

observations from any location and optionally affiliate their data with project(s) in the location. Records in iNaturalist do not require expert identification before inclusion in the database, but “research grade” observations require additional documentation, such as a photo, and concurrence of the identification by other members. Verification may not be important for community scientists who want to contribute to a community forum. However, non-verified records may be prone to error (Howard et al. 2022), and community scientists motivated by a desire to contribute to research could prioritize reporting to apps that include expert verification and contribute to an aggregate dataset, such as EDDMapS. Verified records alleviate concerns among researchers about including localities where a species has been mistakenly reported as present.

Data contributions from the public are useful for supplementing real-time reporting of invasive species, which is essential for EDRR (Martinez et al. 2020). In research, these data are widely used to create species distribution models in both current and future climate conditions (see section on “Modeled data”). Data contributed by community scientists are also relatively cost effective to obtain for both researchers and managers, especially when an app for collection already exists (Encarnação et al. 2021). However, community science data have limitations when considering use in research and management. These data are not systematically collected and are likely to be spatially biased towards areas with high human populations (Geldmann et al. 2016), although similar biases may also be present in professionally collected data (Kosmala et al. 2016).

Inventory and monitoring data

Invasive plant inventory and monitoring is common for managers who need to track and treat invasions within their landholdings. A variety of groups employ invasive species managers for this work, including local, state, and federal governments, Tribal nations, and non-governmental organizations such as land trusts and conservation groups (Beaury et al. 2020). Inventory and monitoring data are typically collected by managers and volunteers in the field, and may include plant occurrence, abundance, treatment, and phenology information (Wallace et al. 2018).

In the U.S., state agencies typically track invasive plant distribution data, though the specific agency varies by state and often by aquatic and terrestrial ecosystems. For example, in Washington state, the Department of Agriculture (WSDA) tracks and records invasive plants and is the main repository for invasive plant data. Several additional entities send invasive plant monitoring data to the WSDA, including county and state noxious weed boards and the Department of Ecology (aquatics). Both the WSDA and Department of Ecology have public data viewing platforms (agr.wa.gov/departments/insects-pests-and-weeds/weeds; apps.ecology.wa.gov/coastalatlas/tools/LakeDetail.aspx), and the WSDA also periodically reports data to EDDMapS (W. DesCamp, Washington State Department of Agriculture, and W. Glisson, Washington State Department of Ecology, written communication, 2022). Some states (e.g., Arizona, Maine, New York, Oregon, and Pennsylvania) store their data using the third-party aggregator iMapInvasives (imapinvasives.org), while many others use EDDMapS. Still other states do not report to any public database and keep internal records that are available upon request (W. McAvoy, Delaware Department of Fish and Wildlife, written communication, 2022). Although states may be rich resources for invasive plant inventory and monitoring data, highly variable data collection methodologies and storage practices necessitate consideration for any landscape-scale ecological analysis or management planning across state borders.

At the federal level, land management agencies collect substantial invasive plant inventory and monitoring data. Many agencies have their own monitoring systems, such as the National Invasive Species Information Management Systems (NISIMS) used by the National Park Service Invasive Plant Management Teams and the Bureau of Land Management (Wallace et al. 2018). For national forests and grasslands, U.S. Forest Service managers follow the Invasive Species Management directive (Forest Service Manual 2900; USDA Forest Service 2011), which emphasizes the importance of adhering to National Forest System protocols, business rules, and standards when collecting, maintaining, and reporting accurate spatial and tabular records for invasive species management activities. These records are required to be entered into the U.S. Forest Service national database of record, currently the Threatened, Endangered,

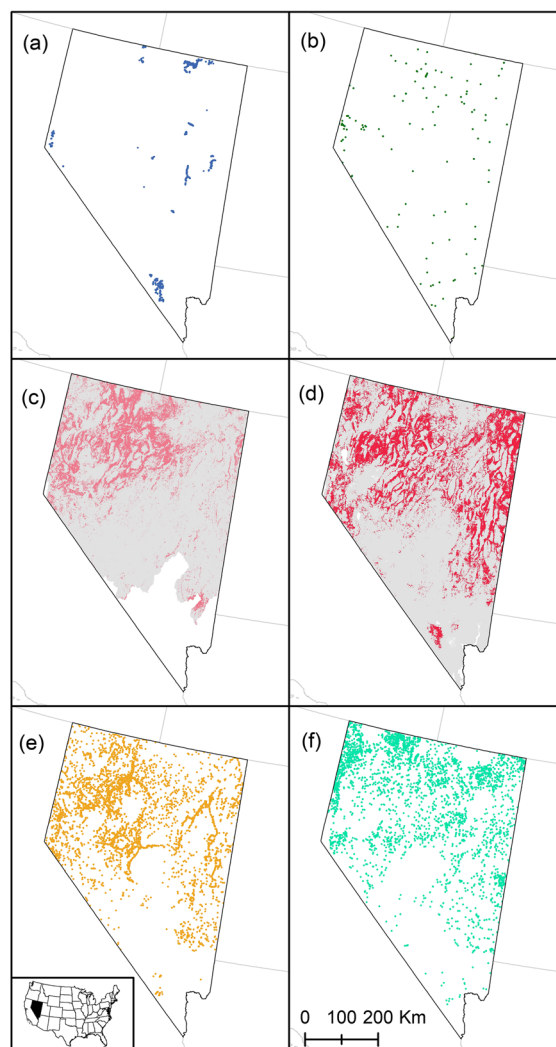


Fig. 1 Invasive cheatgrass (*Bromus tectorum*) is found in Nevada. Cheatgrass occurrence/abundance is shown as inventory and monitoring data from the U.S. Forest Service Threatened, Endangered, and Sensitive Plant—Invasive Species (TESP-IS) database (a), research data from Munro et al. in review (b), remotely mapped data (c; 250 m resolution; > 15% abundance; Bradley et al. 2017), and predicted distribution (d; 100 m resolution; at least 9/10 model agreement; the Invasive Species Habitat Tool (INHABIT)). Data from Early Detection and Distribution Mapping System (EDDMapS), an aggregate dataset, also includes community science observations (e) and aggregate data from the Standardized Plant Community Introduced Status (SPCIS) database (f; Petri et al. 2022) includes vegetation community data

Sensitive Plant—Invasive Species (TESP-IS) database. Inventory records are available to the public on the Enterprise Data Warehouse website (Fig. 1;

fs.usda.gov/about-agency/enterprise-data-warehouse) and may be shared with other national databases through state natural heritage programs.

There are several strengths and limitations to using federal inventory and monitoring data for landscape-scale management and analysis. A major strength of these databases is confidence in data quality and consistency due to strict protocols, rules, and standards for recording tabular and spatial information at the national level. On the other hand, these databases may not be useful for EDRR activities, as they may be infrequently updated (Wallace et al. 2018). In addition, invasive plant data may be stored separately from other inventory and monitoring data, as is the case for National Park Service inventory and monitoring data if they are not collected by the Invasive Plant Management Teams (Wallace et al. 2018). When datasets are stored separately within an entity, users face an additional hurdle in acquiring and integrating data or may miss related datasets completely. Finally, a primary purpose of inventory and monitoring data collected by the U.S. Forest Service (USFS) TESP-IS database is to track and report (to agency leads and Congress) accomplishments in invasive species management, and data collection methods have been developed accordingly. As a result, ancillary information related to treatment may be included and can be applied by management to track species distribution trends and treatment effectiveness. USFS-specific data collection objectives can lead to several difficulties in using these data in spatial analyses and for sharing data outside the agency.

Additional groups tasked with land management and stewardship, including local entities (e.g., county noxious weed boards), non-governmental organizations, and Tribal nations, may also collect invasive plant inventory and monitoring data. For example, The Nature Conservancy (TNC) in Massachusetts collects invasive species data at the state chapter level. Although optional data collection standards have been designed to be compatible with repositories like EDDMapS, and the state chapter may choose to upload these data, there are currently no requirements for the state chapter to do so (K. Lombard, The Nature Conservancy, and A. Sirois-Pitel, The Nature Conservancy, written communication, 2022). Tribal nations in the Pacific Northwest reported that they both work

with local, state, and federal entities, and conduct their own surveys to monitor invasive plant populations (Hoard et al. 2020). Although accessing these data may be difficult, landscape- and regional-scale invasive plant research and management may benefit from coordinating with these groups to supplement open-access spatial datasets.

Vegetation community data

Vegetation community data are typically collected as plots or transects and provide a species census within an area. It is unusual for this class of data to focus exclusively on invasive plants; however vegetation community data include incidental invasive plant data collection when invasive plants are present. Across the United States, several vegetation community datasets have been created for both research and monitoring purposes.

Vegetation community datasets have substantial variation in their collection schedules, habitats sampled, and spatial distribution schemes (Petri et al. 2022). Long-term repeated data collection is exemplified in the Forest Inventory Analysis (FIA) program. The FIA program began in the 1930's, with a more modern version of the program developed in the 1990's, and includes resampling of gridded permanent forested plots located across the country every 5–10 years (Bechtold and Patterson 2005; Tinkham et al. 2018). Invasive plant data are a relatively recent addition to the FIA program (Tinkham et al. 2018), and how invasive plants are recorded often depends on the plant life form, percent cover, and regional or state lists of target invasives developed by experts (Oswalt et al. 2012; Gray et al. 2021). More recently, the National Ecological Observatory Network (NEON) was implemented, which will conduct repeated sampling of plant diversity data yearly or sub-yearly for 30 years (Keller et al. 2008). Vegetation community data may also be collected as part of a sampling protocol such as FEAT/FIREMON Integrated (FFI), which is a sampling protocol used by federal agencies for fire effects monitoring where data are collected at regular intervals within a project area (Lutes 2009).

In addition to variation in sampling collection schedules, focal habitats also vary among datasets. For example, FIA data are focused on forested areas (Tinkham et al. 2018), Bureau of Land Management

Assessment, Inventory, and Monitoring (AIM) data are mostly collected in shrublands of the southwest (Petri et al. 2022), while NEON data are collected across 20 ecological regions (Barnett et al. 2019) and several terrestrial and aquatic habitat types. However, NEON plant diversity sampling plots are only located in dominant land cover classes at each site and may not detect early invasion in rare vegetation types (Gill et al. 2021). Vegetation community datasets also have substantial variation in how sampling sites are chosen, or how plots within a site are distributed, from spatially unbiased equal probability sampling (e.g., FIA; Tinkham et al. 2018) to stratified across representative land cover types (e.g., NEON; Barnett et al. 2019).

Vegetation community data allow researchers to investigate invasions from a ‘whole-community perspective’, and answer questions such as why recipient native communities may vary in vulnerability to invasions (Stohlgren et al. 2002) or how invasion impacts vary across space (Latzka et al. 2016). The rich ancillary data provided in these datasets are perhaps the greatest asset to researchers interested in how invasive plants influence landscape-scale processes and patterns. For example, consistent inventory methods of size class across non-native tree species in the FIA data collection protocol allow for more direct comparisons of invasiveness among inventoried species (e.g., Potter et al. 2022). There may also be opportunities to compare invaded versus uninvaded locations among plots, sites, and regions, and track how these invasions affect ecosystem services with long-term community datasets. For managers, vegetation community datasets are most useful for supplementing inventory and monitoring data when exact plot locations are available, and for understanding broader vegetation community susceptibility and invasion trends.

One major limitation to these datasets is that they are not focused on collecting invasive plant data, resulting in high variation in invasive plant data inclusion. For example, of NEON’s 47 Terrestrial sites (each with around 40 plots), 4 are completely uninvaded, 4 sites have over 50 non-native plant species (max 97), and most sites have low relative cover of non-native species (Figure S1). In these instances, uninvaded plots with long-term collection schedules may provide information for baseline communities

and invasion over time, should invasion occur, but their current use for invasion research may be limited.

Research data

Invasive plant research relies heavily on field data collection and often includes species locality information. However, these data can be challenging to find and extract. In the Global Plant Invaders database (Laginhas and Bradley 2022), a database of all invasive plants mentioned in the scientific literature, more than half of all unique records of invasions documented (57%; 12,854 of 22,356) included a geographic location recorded as coordinates, a map, or both. Data from the scientific literature tended to provide a richer story beyond spatial data, including information about invader impacts, management efficacy, and/or temporal trends in invasions (Munro et al. in review). However, there were very few data points compared to other forms of data collection and aggregation (Fig. 1); between 1999 and 2018, there were 148 studies on *Phragmites australis* with geographic locations in the lower 48 states, compared to over 28,000 observations on the Early Detection and Distribution Mapping System (EDDMapS, see below; Barger and Moorehead 2007). Scientific studies also tended to cluster near universities and research institutes, but tended to be farther from roads, potentially in more pristine habitats that might indicate invader spread independent of disturbance.

Although the vast majority of environmental researchers are interested in sharing their data (Tenopir et al. 2011), it is unlikely that invaded site locations from research are incorporated into management and decision-making because most invasive species managers do not get their information primarily from scientific publications (Beaury et al. 2020) and extracting data from publications is time consuming. In some cases, managers might benefit from meta-analyses where data from previous research papers have already been compiled and made available (e.g., Nagy et al. 2021a). They might also consider management actions informed by research on treatment efficacy in specific localities. Publications including source data are becoming increasingly commonplace in ecology, so future scientists and managers may have an easier time accessing this research data.

Despite these limitations, drawing geographic information from the scientific literature can provide

a useful complement to existing geographic data, particularly for in-depth investigations into single species or for research synthesis projects, such as reviews and meta-analyses. For example, West et al. (2015) predicted future distributions of cheatgrass in Rocky Mountain National Park in Colorado, USA, using data from four different previously collected field surveys. Records from the scientific literature do not greatly expand the geographic scope already sampled by manager records, but often are rich in ancillary information such as invasive plant abundance, the invaded community, and/or the invaded environment (Munro et al. in review). These ancillary data are useful for conducting landscape- and regional-scale analyses related to ecological theory and processes. For example, Stohlgren et al. (2006) combined modified-Whittaker multi-scale plot data from several different published studies in several different states to conduct a regional analysis from 48 different vegetation associations.

Modeled data

Modeled data are a form of invasive species distribution data that are derived from the previously described primary observations and are transformed into a new data source (Fig. 2). These data describe invasive plant distributions across a range of scales

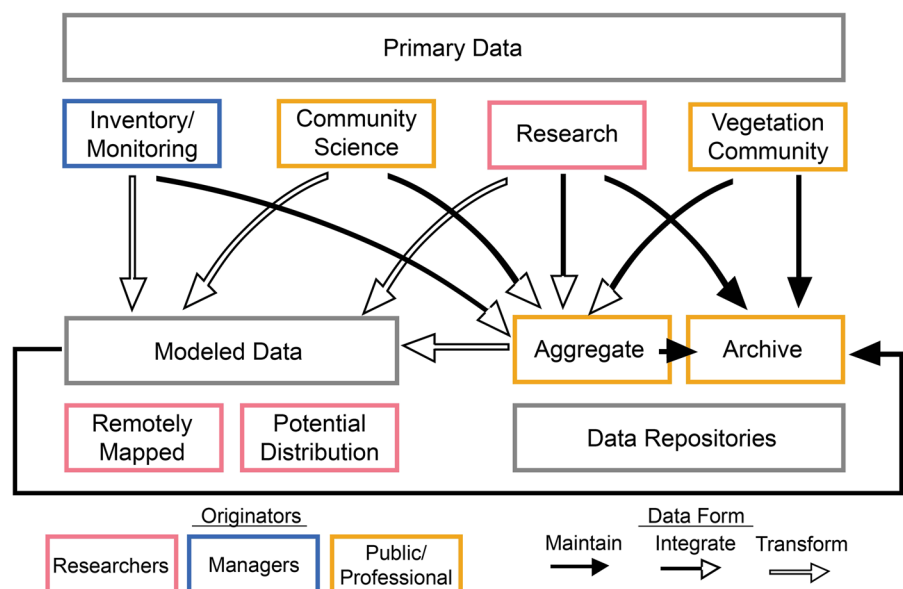
and can be used to augment our understanding of spatial and temporal patterns of plant invasion (Table 2). There are two types of modeled data: (1) Remotely Mapped Distributions and (2) Potential Distributions.

Remotely mapped distributions

Remotely sensed invasive species distribution products (hereafter RS-distribution products) use space-craft- or aircraft-mounted sensor data to create landscape- to regional-scale estimates of target species' distributions. To directly map species distributions using RS data, two conditions must be met: (1) the invasive plant must have unique spectral or temporal traits that are measurable with available sensors (Bradley 2014), and (2) the information measured by a sensor must be available at an appropriate resolution, time, and location (Huang and Asner 2009). Plant traits that allow for direct detection of invasive species include physiological, biochemical, phenological, and morphological properties (e.g., leaf nitrogen content or early greening; Niphadkar and Nagendra 2016). Although it can be argued that remotely sensed plant traits are a type of primary observation, these data are used in a modeling framework with primary invasive species distribution data sources and are used for training and testing (Fig. 2).

There are many research applications for RS-distribution products in invasion ecology, including

Fig. 2 Eight invasive species data categories are interrelated, with primary datasets contributing to the creation of modeled and repository datasets. These datasets can be created and curated by researchers, managers, and the public/professionals



enhancing our fundamental understanding of mechanisms, processes, patterns, and impacts of invasion over broad spatial scales (Vaz et al. 2019). For example, a RS-distribution product was created (Bradley 2017) and used to quantify the impact of invasive annual grass species on fire regimes in the Great Basin (Fig. 1), furthering our understanding of regional-scale disturbance dynamics related to invasion (Bradley et al. 2018a). In circumstances where relatively little is known about an invading species, RS-distribution products can provide early information about the location and trends in invasion over time. In the Pacific Northwest USA, a RS-distribution product was created and helped characterize the spread of a recently problematic invasive annual grass (*Ventena dubia*) over a decade, identifying habitats associated with the invasion, spatial patterns of spread, and potential associations with wildfire (Nietupski 2021). In Hawai'i, a RS-distribution data product was created and used to demonstrate how invasive tree species changed forest structure (Asner et al. 2008).

In management, RS-distribution products have been used for myriad purposes, including biosecurity (Sheffield and Dugdale 2020), long-term monitoring (Cordell et al. 2017; Parker et al. 2021), and land management planning (Pascual et al. 2022). RS-distribution products can help improve the efficiency of invasive plant monitoring through either the direct mapping of invasive populations over time (e.g., Dahal et al. 2022), or the detection of negative impacts to the health of surrounding plant communities (Parker et al. 2021) by reducing management time spent setting up field surveys to track the progression or impacts of an invasion. In addition to monitoring and mapping, RS-distribution products can be used to prioritize invasive species treatment. For example, the *V. dubia* product developed by Nietupski (2021) is being used to analyze the invasion impact to grazing allotments and to identify high-priority areas for population control (J. Brunson, U.S. Forest Service Malheur National Forest, written communication, 2022).

However, there are limitations to the development, interpretation, and curation of products derived from RS. One of the greatest limitations in the development of RS-distribution products is that not all invasive plants are detectable by RS, precluding some species from product development. This could be because the target invasive plant's traits are not sufficiently unique

compared to surrounding communities (Huang and Asner 2009), because its distinguishing traits are not yet known, or because its populations are too small to be detected at the available data resolution. Once plant distributions have been mapped, interpretation of distributions can present unique challenges. For example, if environmental properties (e.g., soil type and surface moisture) of the invaded habitat are incorporated into the predictive model, the output will be a combination of habitat suitability and species' presence, a nuance that may not be understood by an end user (Bradley et al. 2012). A final limitation to the widespread use of RS-distribution products is the lack of a central repository, which can make locating the appropriate dataset difficult and time consuming (see "Data Repositories" section).

Potential distributions

Habitat suitability models, a type of correlative species distribution model (SDM), combine species occurrence information from primary data sources (Fig. 2) with environmental predictors such as temperature, precipitation, and elevation. These inputs are used to develop mathematical relationships describing occurrence or abundance, which can be used to determine a potential distribution for a target invasive species. Some SDM approaches use several model algorithm types (e.g., Maxent, random forest, generalized linear models; Jarnevich et al. 2021a) to create more robust predictions using a model ensemble approach. The resulting data products are typically spatially continuous estimates of presence/abundance, or maps that describe model agreement on suitability for invasion when an ensemble approach is used and can be landscape to regional in scale.

Species distribution models can also be used to predict potential future species distribution in a changing climate. In these applications, SDMs use the statistical relationship between an invasive species occurrence and environmental variables such as current climate to predict the future distribution of the species using climate projections (Jeschke and Strayer 2008; Iverson and McKenzie 2013; Srivastava et al. 2019). Several different statistical approaches are used in these analyses (e.g., linear and logistic regression, climate envelope techniques, classification and regression tree). However, SDMs under future climate conditions are focused on predictions using

climate space and lack consideration of other abiotic factors such as soils, dispersal, biotic interactions, and non-equilibrium conditions, and assume climate conditions are the factors controlling distributions (Jeschke and Strayer 2008; Srivastava et al. 2019).

Invasive plant SDMs have basic science applications such as exploring biogeographical questions related to species range infilling (the expansion of a species within its fundamental range; e.g., Bradley et al. 2015) and predicting range shifts under climate change (e.g., Allen and Bradley 2016). Similar to RS-distribution products, managers and practitioners can use SDMs to answer questions related to designing surveys for new populations and individuals, identifying priority locations and actions for conservation and management, and supporting regulatory decisions and streamlining compliance (Sofaer et al. 2019). Using invasive plant habitat suitability models to target population searches is particularly useful for practitioners who are hampered by the scope of the invasive species problem compared to available resources to manage invasive species (Engelstad et al. 2022). In addition to facilitating monitoring for current invasive plant populations, SDMs that include future climate scenarios can be used to guide the development of climate-informed watchlists (e.g., Coville et al. 2021).

However, there are several limitations to SDMs including potentially poor performance in predicting actual species occurrence, and even more often, difficulty in accurately predicting species abundance (Lee-Yaw et al. 2022). This may largely be because the species occurrence data used are seldom collected for the purpose of generating robust SDMs, so biases inherent to the data may be present in the final models (Lee-Yaw et al. 2022). Particularly for species early in the invasion process, SDMs may be limited if the species is not in equilibrium in its environment (Václavík and Meentemeyer 2012), though this may not be as important for larger scale analyses (Bradley et al. 2015). Barriers also exist for SDM implementation in management because there is often a divide between researchers creating these models and practitioners who may find them useful in informing land management actions (Sofaer et al. 2019).

Data repositories

Data repositories are a third category of data that include aggregates and archives (Table 1). Data repositories are the result of multiple inputs from the primary and modeled sources discussed previously. Here we distinguish data aggregates from archives by classifying aggregates as integrated input data that result in a new product, while data archives are collections of datasets in their original form that may include lists and access information for a group of thematically related data sources (Fig. 2). Data in repositories may either be uploaded by the originator or collected by a third party.

Aggregates

Data aggregates integrate invasive species data from primary data sources into a new, often more comprehensive, data product (Fig. 2). They include, but are not limited to, data collected by government entities, community scientists, and non-governmental organizations. Two common aggregates in the United States include the Early Detection and Distribution Mapping System (EDDMapS; eddmaps.org/; Barger and Moorhead 2007; Fig. 1) and iMapInvasives (imapsinvasives.org/). Together, these two aggregates house millions of observations of many invasive plants collected over the last decade. While most records from these sources are species occurrences (Bradley et al. 2018b), both aggregates are increasingly making it easy for managers to report invasive plant abundance (typically in cover classes such as trace, 1–5%, 5–10%, 10–25%, > 25%). Collectively, these data provide information about the geography of invasions across the United States and are accessible to anyone.

Importantly, occurrence records from both EDDMapS and iMapInvasives are verified by experts, leading to high confidence in the species identity and location. Researchers may use these aggregated datasets to inform initial field research planning and to identify local species experts if the data collector's name or organization is included in the dataset. Managers frequently use these datasets to alert them to new occurrences of invasive species observed within their state, county, or management area, making this data collection extremely valuable for supporting the monitoring efforts needed to inform early detection and rapid response to manage new infestations.

However, contributions from managers to these data aggregates are not uniform, with substantial variability in state protocols for whether, where, and which data should be reported. This could lead to spatial biases when using data across political boundaries. It is also important to note that not all apps collecting species data require expert verification (Howard et al. 2022), so records from non-verified databases should be used with caution.

In addition to aggregates of individual observations from managers and the public, vegetation community data can also be aggregated. For example, the Standardized Plant Community with Introduced Status database (SPCIS) is a collated database of nine datasets that recorded plant percent cover in either plots or along transects (Petri et al. 2022). It includes 83,391 sampling units and cover measurements for 14,056 plant taxa observed in nearly all ecosystem types of the 50 United States, and Puerto Rico. The primary focus of SPCIS was to compile and provide community-level data on ecosystems invaded by non-native plants, including standardized taxonomy and measures of percent cover. SPCIS also includes substantial information on uninvaded areas that can be used to understand plant community dynamics more generally.

The broad geographic scope of the data included in SPCIS, and in other similar data aggregates (e.g., sPlot; Bruelheide et al. 2019), provides opportunities for novel research into invasion patterns across taxa, scales, and environments. Curated sets of community data have the advantage of circumventing geographic limitations of any single input source and standardizing some of the inconsistencies discussed for individual data products (e.g., taxonomy). However, users must carefully consider differences across input data, such as the biases already discussed across ‘Vegetation Community Data’ sources, several of which are aggregated in SPCIS. In SPCIS for example, the area sampled per plot or transect varies from 1 to 10,000 m² and was sometimes adjusted to reflect differences in the dominant vegetation type (e.g., larger plot sizes in forested ecosystems). Area sampled influences many ecological patterns (Levin 1992; Powell et al. 2011) and should be accounted for in many types of analyses. Differences in sampling schemes across the SPCIS database could lead to high spatial autocorrelation in some regions but insufficient

sampling effort elsewhere (e.g., whether plots are clustered within National Park Service or NEON sites or are randomly distributed across space). So, although SPCIS provides substantial data that can be used to study invasions across environments, it does not represent a random sample of invaded ecosystems. These caveats should not be seen as a barrier to using aggregated community data, but rather factors that need to be carefully considered and accounted for in analyses.

Archives

Data archives may include primary data and/or modeled data which are curated into a list of datasets. Archives typically have a common theme, or can be searched based on a theme, and include a description of and information for accessing the original data source. Archives differ from previously discussed repositories in that archival data are not integrated into a new data product and instead keep their original form. The main purposes of databases in this category are to keep a record of data and to facilitate availability to researchers and managers.

Several data archives exist for storing research data with the aim of supporting reproducible and accessible science. These include third party storage such as the Environmental Data Initiative (edirepository.org/), the Knowledge Network for Biocomplexity (knb.ecoinformatics.org/), Dryad (datadryad.org/stash), and Vegbank (vegbank.org), which is an open-source repository for vegetation surveys, many of which contain similar information to the inventory and monitoring datasets described above. In addition, government archives such as the U.S. Geological Survey’s Science Data Catalog (sciencebase.gov/catalog/) and the USDA Forest Service Research Data Archive (fs.usda.gov/rds/archive/) include data produced by government agencies. Researchers and managers can access these data either through a link in an associated publication, or by searching the archives directly. Although these archives contain spatial datasets, they are not limited to spatial data.

A major limitation for archived datasets is finding relevant data due to the often-decentralized nature of relevant datasets spread across multiple data archives (Culina et al. 2018). However, this can be mitigated when archives are already curated for a specific purpose with suggestions for choosing an appropriate

dataset. For example, Tarbox et al. (2022) evaluated invasive annual grass spatial data sources for the western United States and created an archive database with descriptions and links to the evaluated products (Shyvers et al. 2022). Simpson et al. (2020) curated a list of federal invasive species databases and tools that could aid researchers and managers in EDRR activities. Other applications and limitations of using these data in research and management are consistent with those associated with the original datasets they reference.

Facilitating data usability

None of the discussed spatial datasets can improve research and management if they are not used or if the infrastructure that holds them is not updated and maintained (Millerand and Baker 2020). To maximize use of these existing datasets, data creators should aim to increase findability, accessibility, interoperability, and reusability (FAIR data principles; Wilkinson et al. 2016), wherever possible. This is particularly important in invasive species research and management where data are generated and stored across many different entities, and following FAIR principles could facilitate data integration. In addition, invasive species managers report funding and personnel limit their work (Beaury et al. 2020), and efforts to follow FAIR data principles could alleviate some of the workload associated with finding and using existing data for management applications.

Many data creators are aware of limitations to data use, and several efforts are underway to encourage more standardized data products. For example, the North American Invasive Species Management Association (NAISMA) mapping standards (NAISMA 2018) define a minimum standard for data collection. These standards were created with the goal of facilitating invasive plant map sharing across jurisdictions and recommend the inclusion of five basic elements: (1) Species, (2) Location, (3) Area infested, (4) Collection date, and (5) Collector (NAISMA 2018). Standardization during data collection would allow the many already existing invasive species data products to be more interoperable and integrated, and relevant decision support tools to be created (Wallace et al. 2020).

To increase (re)usability, data products, online tools, and user interfaces can be created, ideally with manager involvement, or co-produced (Morelli et al. 2021). For example, the Invasive Species Habitat Tool (INHABIT; gis.usgs.gov/inhabit/) integrates SDM outputs into a web application displaying visual and statistical summaries of these habitat suitability models. The design and content of INHABIT was co-produced with managers who provided feedback on the models themselves and on INHABIT's features. The content and functionality of INHABIT are designed to provide practical information leading to enhanced land management actions, including mapped products with interactive thresholds to define suitability based on management objectives (with geospatial download options; Jarnevič et al. 2021b), information on modeled environmental relationships, and tabular proximity summaries that compare relative risk of invasion.

Accessing and using large, archived datasets requires both the data science skills to work with data and the ecological or environmental science domain knowledge to understand and interpret them (Nagy et al. 2021b). An ecosystem of software and derived data products has been developing over the past several years via efforts for NEON, FIA, and other products. For NEON data, *neonUtilities* (Lunch et al. 2021) can be used for acquiring raw NEON data, and *ecocomDP* (O'Brien et al. 2021; Smith et al. 2022) and *neonDivData* (Li et al. 2022) for converting raw NEON data into standard data formats. For FIA data, the *rFIA* packages can acquire, process, and analyze data (Stanke et al. 2020). More generally, *ecocomDP* (O'Brien et al. 2021) and *VegX* (Wiser 2016; De Cáceres 2018) can be used to convert vegetation plots from a myriad of sources into standardized formats.

Additional tools for increasing data reuse include instruction materials for researchers and managers such as tutorials, workshops, and factsheets. If data creators want to encourage use, creating these companion materials for datasets can facilitate uptake. For NEON data, users can view tutorials and upcoming workshops within the online "Learning Hub" (neon-science.org/resources/learning-hub). Tutorials may also be created by third party organizations, as is exemplified by the Northeast Regional Invasive Species and Climate Change (RISCC) Network, which created video tutorials and summaries for several invasive species distribution tools (risccnetwork.org/tool-summaries). Further facilitating use, these

materials were sent directly to a network of managers and researchers. While still considered useful, invasive species managers reported that online decision-making tools are less useful than management documents/ technical reports (Beaury et al. 2020), indicating that factsheets to accompany data products (e.g., Van Schmidt et al. 2022) are a high-priority format when engaging with managers.

Choosing spatial data

With so many invasive plant spatial data sources currently available, potential users may become

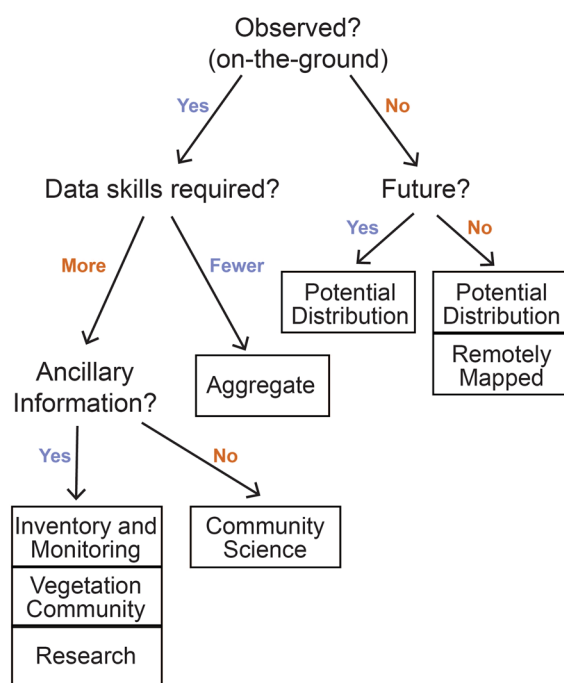


Fig. 3 Users may consider several questions based on their unique needs and abilities when choosing a dataset. “Observed” refers to whether on-the-ground observations are needed as opposed to whether a modeled dataset is sufficient. Users may also consider their own data skills. Where more skills are required, special software or analysis abilities may be required, whereas fewer skills might be required for access and interpretation for other datasets. “Ancillary information” refers to whether information, in addition to presence/abundance/absence, may be available in the dataset. “Future” refers to datasets that provide distribution estimates under potential future conditions rather than estimates of distributions with current conditions. Note here that potential distribution data can provide information on potential current and future distributions

overwhelmed with options. However, each type of data source has a unique set of characteristics, and there are several questions that users might consider to determine which data type is most appropriate for their research or management needs (Fig. 3). Users might contemplate whether they need a dataset that provides primary (on-the-ground) observations or whether predicted or modeled distributions will suffice. For managers, modeled distribution data may be useful for determining places to monitor or search for new populations of invasive plants. For researchers, modeled datasets may be sufficient for landscape- or regional-scale analyses. Users might also consider whether they are interested in current distributions or future projections. If interested in current distributions, users might consider predicted potential distribution data (e.g., SDMs) or remotely mapped data (Fig. 3). However, climate change adaptation planning or climate change research might require the use of future projections based on potential distribution datasets. Modeled datasets are also the only ones that provide spatially continuous estimates across landscapes.

The scale for each data type also varies, with on-the-ground observations typically available at smaller local scales and modeled data at larger spatial scales (Table 2). On-the-ground observations of invasive plants may be useful for EDRR activities and creating treatment plans for managers and might be more appropriate for local scale analysis or planning potential field sampling efforts for researchers. Of these data types, data aggregates are often the most user friendly and may only require standard computer software and basic GIS skills. These datasets often include a species identification, and occurrence or abundance information at a georeferenced location. However, users with more advanced GIS, data analysis, and coding abilities may find the rich ancillary data (e.g., vegetation community composition, phenology, treatment) included in research, inventory and monitoring, or vegetation community datasets worth the additional processing efforts (Fig. 3).

When available data products are not sufficient, users may also choose to integrate existing data products both within and across data types. In addition to the technical expertise required for integrating data, users must understand the component datasets to avoid misrepresenting species’ distributions. This requires reading associated publications and user

guides and may necessitate contacting the data originators. Data integration also requires careful review and cleaning of data records. For example, Allen and Bradley (2016) integrated invasive plant distribution data from 33 sources, removing duplicate records, erroneously located records, and spatially co-located records, to create species distribution models. Another barrier to data integration that must be considered is inconsistent definitions (Richardson et al. 2000) and lists of invasive species (e.g., Oswalt et al. 2012) that determine which species data are collected. For example, if not all states consider a plant invasive, it may be difficult to accurately integrate distribution data for that species across state borders. However, if datasets are integrated with care, they can provide new avenues for answering questions in macroecology.

Conclusions

Understanding and managing current and future invasions across landscapes will require high-quality and accessible data on current and potential invasive plant distributions. Although many datasets are currently available, a lack of coordination between the entities that acquire and create these data can stymie use and access, particularly across spatial scales and jurisdictions. For example, potential users may find it prohibitive to locate all relevant data products, and data aggregators may face insurmountable obstacles when trying to integrate products.

Improving access and understanding of invasive plant spatial data will help managers plan search activities and prioritize current invasive plant treatment and future climate change adaptation actions, including EDRR and other restoration activities that may lower the risk of invasion (e.g., seeding native species). Increasing access and usability will also aid landscape scale invasive plant research and further our understanding of ecological theory, including community resilience, invasion risk, and landscape scale impacts on ecological processes. This explanation of the invasive plant data landscape, with examples for application and limitations, should facilitate appropriate and efficient data choices in current and future research and management.

Acknowledgements We would like to thank Andrew Gray and two anonymous reviewers for their helpful comments which greatly improved this manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author contributions EJF conceptualized the manuscript, led the manuscript, and created Figs. 1–3. ALM created the supplemental figure. All authors wrote the first draft, edited subsequent drafts, and approved the final version of the manuscript.

Funding This research was supported in part by appointments (E.J.F. and T.N.) to the United States Forest Service (USFS) Research Participation Program administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy (DOE) and the U.S. Department of Agriculture (USDA). ORISE is managed by ORAU under DOE contract number DE-SC0014664. All opinions expressed in this paper are the author's and do not necessarily reflect the policies and views of USDA, DOE, or ORAU/ORISE. Funding for R.C.N. was provided by Earth Lab through the University of Colorado Boulder's Grand Challenge Initiative, the Cooperative Institute for Research in Environmental Sciences (CIRES), and the U.S. Geological Survey Award #G21AC10278-00. B.A.B. was supported by U.S. Geological Survey Award #G21AC10233-00.

Declarations

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

References

- Ahmed DA, Hudgins EJ, Cuthbert RN, Kourantidou M, Diagne C, Haubrock PJ, Leung B, Liu C, Leroy B, Petrovskii S, Beidas A, Courchamp F (2022) Managing biological invasions: the cost of inaction. *Biol Invasions* 24:1927–1946
- 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
- Asner GP, Hughes RF, Vitousek PM, Knapp DE, Kennedy-Bowdoin T, Boardman J, Martin RE, Eastwood M, Green RO (2008) Invasive plants transform the three-dimensional structure of rain forests. *Proc Natl Acad Sci USA* 105(11):4519–4523
- Bargeron CT, Moorhead DJ (2007) EDDMapS—early detection and distribution mapping system for the southeast exotic pest plant council. *Wildl Weeds* 10:4–8
- Barnett DT, Adler PB, Chemel BR, Duffy PA, Enquist BJ, Grace JB, Harrison S, Peet RK, Schimel DS, Stohlgren TJ, Vellend M (2019) The plant diversity sampling design for the national ecological observatory network. *Ecosphere* 10(2):e02603
- Beaury EM, Fusco EJ, Jackson MR, Laginhas BB, Morelli TL, Allen JM, Pasquarella VJ, Bradley BA (2020)

- Incorporating climate change into invasive species management: insights from managers. *Biol Invasions* 22:233–252
- Bechtold WA, Patterson PL (2005) The enhanced Forest Inventory and Analysis program—national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. <http://www.treesearch.fs.fed.us/pubs/20371>
- Bradley BA (2014) Remote detection of invasive plants: a review of spectral, textural and phenological approaches. *Biol Invasions* 16(7):1411–1425
- Bradley BA (2017) Cheatgrass (*Bromus tectorum*) percent cover data, data and datasets. 1 <https://doi.org/10.7275/R5XW4GZR>
- Bradley BA, Olsson AD, Wang O, Dickson BG, Pelech L, Sesnie SE, Zachmann LJ (2012) Species detection vs. habitat suitability: are we biasing habitat suitability models with remotely sensed data? *Ecol Model* 244:57–64
- Bradley BA, Early R, Sorte CJ (2015) Space to invade? Comparative range infilling and potential range of invasive and native plants. *Glob Ecol Biogeogr* 24(3):348–359
- Bradley BA, Allen JM, O'Neill MW, Wallace RD, Bargerion CT, Richburg JA, Stinson K (2018b) Invasive species risk assessments need more consistent spatial abundance data. *Ecosphere* 9(7):e02302
- Bradley BA, Curtis CA, Fusco EJ, Abatzoglou JT, Balch JK, Dadashi S, Tuanmu MN (2018a) Cheatgrass (*Bromus tectorum*) distribution in the intermountain Western United States and its relationship to fire frequency, seasonality, and ignitions. *Biol Invasions* 20:1493–1506
- Bradley BA, Beaury EM, Fusco EJ, Munro L, Brown-Lima C, Coville W, Kesler B, Olmstead N, Parker J (2022) Breaking down barriers to consistent, climate-smart regulation of invasive plants: a case study of US Northeast states. *Ecosphere* 13:e4014
- Bruehlheide H, Dengler J, Jiménez-Alfaro B, Purschke O, Hennekens SM, Chytrý M, Pillar VD, Jansen F, Kattge J, Sandel B, Aubin I, Biurrun I, Field R, Haider S, Jandt U, Lenoir J, Peet RK, Peyre G, Sabatini FM, Schmidt M, Schrodte F, Winter M, Ačić S, Agrillo E, Alvarez M, Ambarlı D, Angelini P, Apostolova I, Arfin Khan MAS, Arnst E, Attorre F, Baraloto C, Beckmann M, Berg C, Bergeron Y, Bergmeier E, Bjorkman AD, Bondareva V, Borchardt P, Botta-Dukát Z, Boyle B, Breen A, Brisse H, Byun C, Cabido MR, Casella L, Cayuela L, Černý T, Chepinoga V, Csiky J, Curran M, Čušterevska R, Dajić SZ, De Bie E, de Ruffray P, De Sanctis M, Dimopoulos P, Dressler S, Ejrnæs R, El-Sheikh MAERM, Enquist B, Ewald J, Fagúndez J, Finckh M, Font X, Forey E, Fotiadis G, García-Mijangos I, de Gasper AL, Golub V, Gutierrez AG, Hatim MZ, He T, Higuchi P, Holubová D, Hölzel N, Homeier J, Indreica A, Işık GD, Jansen S, Janssen J, Jedrzejek B, Jiroušek M, Jürgens N, Kački Z, Kavğacı A, Kearsley E, Kessler M, Knollová I, Kolomiychuk V, Korolyuk A, Kozhevnikova M, Kozub Ł, Krstonošić D, Kühl H, Kühn I, Kuzemko A, Küzmič F, Landucci F, Lee MT, Levesley A, Li C, Liu H, Lopez-Gonzalez G, Lysenko T, Macanović A, Mahdavi P, Manning P, Marcenò C, Martynenko V, Mencuccini M, Minden V, Moeslund JE, Moretti M, Müller JV, Munzinger J, Niinemets Ü, Nobis M, Noroozi J, Nowak A, Onyshchenko V, Overbeck GE, Ozinga WA, Pauchard A, Pedashenko H, Peñuelas J, Pérez-Haase A, Peterka T, Petřík P, Phillips OL, Prokhorov V, Rašomavičius V, Revermann R, Rodwell J, Ruprecht E, Rūsiņa S, Samimi C, Schaminée JHJ, Schmiedel U, Šibík J, Šilc U, Škvorec Z, Smyth A, Sop T, Sopotlieva D, Sparrow B, Stančić Z, Svenning J, Swacha G, Tang Z, Tsiripidis I, Turtureanu PD, Uğurlu E, Uogintas D, Valachovič M, Vanselow KA, Vashenyak Y, Vassilev K, Vélez-Martin E, Venanzoni R, Vibrans AC, Violle C, Virtanen R, von WH, Wagner V, Walker DA, Wana D, Weiher E, Wesche K, Whitfield T, Willner W, Wiser S, Wohlgemuth T, Yamalov S, Zizka G, Zverev A (2019) sPlot—a new tool for global vegetation analyses. *J Veg Sci* 30(2):161–186
- Castro-Díez P, Godoy O, Alonso A, Gallardo A, Saldaña A (2014) What explains variation in the impacts of exotic plant invasions on the nitrogen cycle? A meta-analysis. *Ecol Lett* 17(1):1–12
- Cordell S, Questad EJ, Asner GP, Kinney KM, Thaxton JM, Uwolo A, Brooks S, Chynoweth MW (2017) Remote sensing for restoration planning: how the big picture can inform stakeholders. *Restor Ecol* 25:S147–S154
- Coville W, Griffin BJ, Bradley BA (2021) Identifying high-impact invasive plants likely to shift into northern New England with climate change. *Invasive Plant Sci Manag* 14:57–63
- Culina A, Baglioni M, Crowther TW, Visser ME, Woutersen-Windhoeuw S, Manghi P (2018) Navigating the unfolding open data landscape in ecology and evolution. *Nat Ecol Evol* 2(3):420–426
- Dahal D, Boyte SP, Parajuli S, Pastick NJ, Oimoen MJ, Shrestha D (2022) Early estimates of exotic annual grass (EAG) in the Sagebrush Biome, USA, 2022 (ver 6.0, July 2022): U.S. Geological Survey data release. <https://doi.org/10.5066/P9FVYOGD>
- D'Antonio CM, Vitousek PM (1992) Biological invasions by exotic grasses, the grass/fire cycle, and global change. *Annu Rev Ecol Syst* 23:63–87
- De Cáceres M (2018) Sobre el intercambio de datos de vegetación: el estándar 'Veg-X' y el paquete de R 'VegX'. *Ecosistemas* 27(1):128–129
- Early R, Bradley BA, Dukes JS, Lawler JJ, Olden JD, Blumenthal DM, Gonzalez P, Grosholz ED, Ibáñez I, Miller LP, Sorte CJ, Tatem AJ (2016) Global threats from invasive alien species in the twenty-first century and national response capacities. *Nat Commun* 7:1–9
- Encarnação J, Teodósio MA, Morais P (2021) Citizen science and biological invasions: a review. *Front Environ Sci* 8:602980
- Engelstad P, Jarnevich CS, Hogan T, Sofaer HR, Pearse IS, Sieracki JL, Frakes N, Sullivan J, Young NE, Prevý JS, Belamaric P, LaRoe J (2022) INHABIT: A web-based decision support tool for invasive plant species habitat visualization and assessment across the contiguous United States. *PLoS ONE* 17(2):e0263056
- Flory SL, Clay K (2009) Effects of roads and forest successional age on experimental plant invasions. *Biol Conserv* 142(11):2531–2537
- Geldmann J, Heilmann-Clausen J, Holm TE, Levinsky I, Markussen BO, Olsen K, Rahbek C, Tøttrup AP (2016)

- What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. *Divers Distrib* 22(11):1139–1149
- Gill NS, Mahood AL, Meier CL, Muthukrishnan R, Nagy RC, Stricker E, Duffy KA, Petri L, Morissette JT (2021) Six central questions about biological invasions to which NEON data science is poised to contribute. *Ecosphere* 12(9):e03728
- Gray AN, Baer KC, Witt C (2021) Chapter 8: Invasive Plants. In: Barrett TM, Robertson GC. [eds]. *Disturbance and sustainability in forests of the Western United States*. Gen. Tech. Rep. PNW-GTR-992. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 161–190.
- Hoard S, Sanders C, Anderson B (2020) *Invasive Species Capabilities and Capacity Survey—Tribal Governments*. Washington Invasive Species Council Report
- Howard L, van Rees CB, Dahlquist Z, Luikart G, Hand BK (2022) A review of invasive species reporting apps for citizen science and opportunities for innovation. *NeoBiota* 71:165–188
- Huang CY, Asner GP (2009) Applications of remote sensing to alien invasive plant studies. *Sensors* 9(6):4869–4889
- IPBES (2019) *Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. E. S. Brondizio, J. Settele, S. Díaz, and H. T. Ngo (editors). IPBES secretariat, Bonn, Germany. 1148
- Iverson LR, McKenzie D (2013) Tree-species range shifts in a changing climate: detecting, modeling, assisting. *Landsc Ecol* 28:879–889
- Jarnevich CS, LaRoe JM, Engelstad PS, Sullivan JR (2021a) INHABIT species potential distribution across the contiguous United States. U.S. Geological Survey data release. <https://doi.org/10.5066/P92476V6>
- Jarnevich CS, Sofaer HR, Engelstad P (2021b) Modelling presence versus abundance for invasive species risk assessment. *Divers Distrib* 27(12):2454–2464
- Jeschke JM, Strayer DL (2008) Usefulness of bioclimatic models for studying climate change and invasive species. *Ann N Y Acad Sci* 1134(1):1–24
- Joppa LN, O'Connor B, Visconti P, Smith C, Geldmann J, Hoffmann M, Watson JE, Butchart SH, Virah-Sawmy M, Halpern BS, Ahmed SE, Balmford A, Sutherland WJ, Harfoot M, Hilton-Taylor C, Foden W, Minin ED, Pagad S, Genovesi P, Hutton J, Burgess ND (2016) Filling in biodiversity threat gaps. *Science* 352:416–418
- Keller RP, Lodge DM, Finnoff DC (2007) Risk assessment for invasive species produces net bioeconomic benefits. *Proc Natl Acad Sci USA* 104:203–207
- Keller M, Schimel DS, Hargrove WW, Hoffman FM (2008) A continental strategy for the National Ecological Observatory Network. *Front Ecol Environ* 6(5):282–284
- Kosmala M, Wiggins A, Swanson A, Simmons B (2016) Assessing data quality in citizen science. *Front Ecol Environ* 14:551–560
- Laginhas BB, Bradley BA (2022) Global plant invaders: a compendium of invasive plant taxa documented by the peer-reviewed literature. *Ecology* 103(2):e03569
- Latzka AW, Hansen GJ, Kornis M, Vander Zanden MJ (2016) Spatial heterogeneity in invasive species impacts at the landscape scale. *Ecosphere* 7(3):e01311
- Lee-Yaw JA, McCune J, Pironon S (2022) Species distribution models rarely predict the biology of real populations. *Ecography* 2022:e05877
- Levin SA (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur Award lecture. *Ecology* 73:1943–1967
- Li D, Record S, Sokol ER, Bitters ME, Chen MY, Chung YA, Helmus MR, Jaimes R, Jansen L, Jarzyna MA, Just MG, LaMontagne JM, Melbourne BA, Moss W, Norman KEA, Parker SM, Robinson N, Seyednasrollah B, Smith C, Spaulding S, Surasinghe TD, Thomsen SK, Zarnetske PL (2022) Standardized NEON organismal data for biodiversity research. *Ecosphere* 13:e4141
- Lopez BE, Allen JM, Dukes JS, Lenoir J, Vilà M, Blumenthal DM, Beaury EM, Fusco EJ, Laginhas BB, Morelli TL, O'Neill MW, Sorte CJB, Maceda-Veiga A, Whitlock R, Bradley BA (2022) Global environmental changes more frequently offset than intensify detrimental effects of biological invasions. *Proc Natl Acad Sci* 119(22):e2117389119
- Lunch C, Laney C, Mietkiewicz N, Sokol E, Cawley K (2021) Utilities for Working with NEON Data. R package version 2.1.3 <https://CRAN.R-project.org/package=neonUtilities>
- Lutes DC, Benson NC, Keifer M, Caratti JF, Streetman SA (2009) FFI: a software tool for ecological monitoring. *Int J Wildl Fire* 18(3):310–314
- Mahood AL, Jones RO, Board DI, Balch JK, Chambers JC (2022) Interannual climate variability mediates changes in carbon and nitrogen pools caused by annual grass invasion in a semiarid shrubland. *Glob Change Biol* 28(1):267–284
- Martinez B, Reaser JK, Dehgan A, Zamft B, Baisch D, McCormick C, Giordano AJ, Aicher R, Selbe S (2020) Technology innovation: advancing capacities for the early detection of and rapid response to invasive species. *Biol Invasions* 22:75–100
- Maund PR, Irvine KN, Lawson B, Steadman J, Risely K, Cunningham AA, Davies ZG (2020) What motivates the masses: understanding why people contribute to conservation citizen science projects. *Biol Conserv* 246:108587
- Michener WK (2015) Ecological data sharing. *Eco Inform* 29:33–44
- Millerand F, Baker KS (2020) Data infrastructures in ecology: an infrastructure studies perspective. *Oxford Research Encyclopedia of Environmental Science*. 27. Oxford University Press. Date of access 14 Dec. 2022. <https://doi.org/10.1093/acrefore/9780199389414.013.554>
- Morelli TL, Brown-Lima CJ, Allen JM, Beaury EM, Fusco EJ, Barker-Plotkin A, Laginhas BB, Quirion BR, Griffin B, McLaughlin B, Munro L, Olmstead N, Richburg J, Bradley BA (2021) Translational invasion ecology: bridging research and practice to address one of the greatest threats to biodiversity. *Biol Invasions* 23(11):3323–3335
- Munro L, Griffin B, Laginhas B, Bradley BA (in review). Does invasion science encompass the invaded range? A comparison of the geographies of invasion science versus management in the U.S
- Nagy RC, Fusco EJ, Balch JK, Finn JT, Mahood AL, Allen JM, Bradley BA (2021a). Data from: a synthesis of the effects

- of cheatgrass invasion on U.S. Great Basin carbon storage. Dryad Digital Repository, <https://doi.org/10.5061/dryad.4mw6m9082>
- Nagy RC, Balch JK, Bissell EK, Cattau ME, Glenn NF, Halpern BS, Ilangakoon N, Johnson B, Joseph MB, Marconi S, O'Riordan C, Sanovia J, Swetnam TL, Travis WR, Wasser LA, Woolner E, Zarnetske P, Abdulrahim M, Adler J, Barnes G, Bartowitz KJ, Blake RE, Bombaci SP, Brun J, Buchanan J D, Chadwick KD, Chapman MS, Chong SS, Chung YA, Corman JR, Couret J, Crispo E, Doak TG, Donnelly A, Duffy KA, Dunning KH, Duran SM, Edmonds JW, Fairbanks DE, Felton AJ, Florian CR, Gann D, Gebhardt M, Gill NS, Gram WK, Guo JS, Harvey BJ, Hayes KR, Helmus MR, Hensley RT, Hondula KL, Huang T, Hundertmark WJ, Iglesias V, Jacinthe P-A, Jansen LS, Jarzyna MA, Johnson TM, Jones KD, Jones MA, Just MG, Kaddoura YO, Kagawa-Vivani AK, Kaushik A, Keller AB, King KBS, Kitzes J, Koontz MJ, Kouba PV, Kwan W-Y, LaMontagne JM, LaRue EA, Li D, Li B, Lin Y, Liptzin D, Long WA, Mahood AL, Malloy SS, Malone SL, McGlinchy JM, Meier CL, Melbourne BA, Mielkiewicz N, Morissette JT, Moustapha M, Muscarella C, Musinsky J, Muthukrishnan R, Naithani K, Neely M, Norman K, Parker SM, Perez Rocha M, Petri L, Ramey C A, Record S, Rossi MW, SanClements M, Scholl VM, Schweiger AK, Seyednasrollah B, Sihi D, Smith KR, Sokol ER, Spaulding SA, Spiers AI, St Denis LA, Staccone AP, Stack Whitney K, Stanitski DM, Stricker E, Sursinghe TD, Thomsen SK, Vasek PM, Xiaolu L, Yang D, Yu R, Yule KM, Zhu K (2021b) Harnessing the NEON data revolution to advance open environmental science with a diverse and data-capable community. *Ecosphere* 12(12):e03833
- NAISMA (2018) Mapping Standards for Program Managers. North American Invasive Species Management Association. Milwaukee, WI <https://www.naisma.org/images/mappingstandardsfinal.pdf>. accessed 18 July 2022
- Nietupski TC (2021) Characterizing an annual grass invasion and its link to environmental and disturbance factors using remote sensing: New tools and applications. Oregon State University. Dissertation, Oregon State University
- Niphadkar M, Nagendra H (2016) Remote sensing of invasive plants: incorporating functional traits into the picture. *Int J Remote Sens* 37(13):3074–3085
- O'Brien M, Smith CA, Sokol ER, Gries C, Lany N, Record S, Castorani MCN (2021) ecocomDP: a flexible data design pattern for ecological community survey data. *Ecol Inform* 64:101374
- Oswalt CM, Oswalt SN, Zimmerman L (2012) Updating the southern nonnative plant watch list: the future of NNIP Monitoring in the south. In Morin, Randall S, Liknes, Greg C., comps. Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012; 2012 December 4–6; Baltimore, MD. Gen. Tech. Rep. NRS-P-105. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station. [CD-ROM]: 274–277. pp 274–277
- Parker K, Elmes A, Boucher P, Hallett RA, Thompson JE, Simek Z, Bowers J, Reinmann AB (2021) Crossing the great divide: bridging the researcher-practitioner gap to maximize the utility of remote sensing for invasive species monitoring and management. *Remote Sens* 13(20):4142. <https://doi.org/10.3390/rs13204142>
- Pascual A, Giardina CP, Povak NA, Hessburg PF, Heider C, Salminen E, Asner GP (2022) Optimizing invasive species management using mathematical programming to support stewardship of water and carbon-based ecosystem services. *J Environ Manag* 301:113803
- Petri L, Beaury EM, Corbin J, Peach K, Sofaer H, Pearse IS, Early R, Barnett DT, Ibáñez I, Peet RK, Schafale M, Wentworth TR, Vanderhorst JP, Zaya DN, Spyreas G, Bradley BA (2022) SPCIS: standardized plant community with introduced status database. *Ecology*. <https://doi.org/10.1002/ecy.3947>
- Potter K, Riitters K, Guo Q (2022) Non-native tree regeneration indicates regional and national risks from current invasions. *Front for Global Change* 5:966407
- Powell KI, Chase JM, Knight TM (2011) A synthesis of plant invasion effects on biodiversity across spatial scales. *Am J Bot* 98(3):539–548
- Pyšek P, Hulme PE, Simberloff D, Bacher S, Blackburn TM, Carlton JT, Dawson W, Essl F, Foxcroft LC, Genovesi P, Jeschke JM, Kühn I, Liebhold AM, Mandrak NE, Meyerson LA, Pauchard A, Pergl J, Roy HE, Seebens H, van Kleunen M, Vilà M, Wingfield MJ, Richardson DM (2020) Scientists' warning on invasive alien species. *Biol Rev* 95(6):1511–1534
- Richardson DM, Pyšek P, Rejmanek M, Barbour MG, Panetta FD, West CJ (2000) Naturalization and invasion of alien plants: concepts and definitions. *Divers Distrib* 6(2):93–107
- Sheffield K, Dugdale T (2020) Supporting urban weed biosecurity programs with remote sensing. *Remote Sensing* 12(12):2007
- Shyvers JE, Tarbox BC, Van Schmidt ND, Saher DJ, Heinrichs JA, Aldridge CL (2022) Database of invasive annual grass spatial products for the western United States January 2010 to February 2021: U. S. Geological Survey data release. <https://doi.org/10.5066/P9VW97AO>
- Simpson A, Morissette JT, Fuller P, Reaser J, Guala GF (2020) Catalog of U.S. Federal Early Detection/Rapid Response Invasive Species Databases and Tools: Version 2.0 (ver. 2.0, 2020): U.S. Geological Survey data release. <https://doi.org/10.5066/P9CNVBYR>
- Smith C, Sokol E, O'Brien M (2022) ecocomDP: Tools to Create, Use, and Convert ecocomDP Data. R package version 1.2.1 <https://CRAN.R-project.org/package=ecocomDP>
- Sofaer HR, Jarnevich CS, Pearse IS, Smyth RL, Auer S, Cook GL, Edwards Jr TC, Guala GF, Howard TG, Morissette JT, Hamilton H (2019) The development and delivery of species distribution models to inform decision-making. *Bio-science* 69(7):544–557
- Srivastava V, Lafond V, Griess VC (2019) Species distribution models (SDM): applications, benefits and challenges in invasive species management. *CABI Reviews* 1–13. <https://doi.org/10.1079/PAVSNNR201914020>
- Stanke H, Finley AO, Weed AS, Walters BF, Domke GM (2020) rFIA: an R package for estimation of forest attributes with the US Forest Inventory and Analysis database. *Environ Model Softw* 127:104664
- Stohlgren TJ, Chong GW, Schell LD, Rimar KA, Otsuki Y, Lee M, Kalkhan MA, Villa CA (2002) Assessing vulnerability

- to invasion by nonnative plant species at multiple spatial scales. *Environ Manage* 29(4):566–577
- Stohlgren TJ, Barnett D, Flather C, Fuller P, Peterjohn B, Kartesz J, Master LL (2006) Species richness and patterns of invasion in plants, birds, and fishes in the United States. *Biol Invasions* 8(3):427–447
- Tarbox BC, Van Schmidt ND, Shyvers JE, Saher DJ, Heinrichs JA, Aldridge CL (2022) Bridging the gap between spatial modeling and management of invasive annual grasses in the imperiled sagebrush biome. *Rangel Ecol Manage* 82:104–115
- Tenopir C, Allard S, Douglass K, Aydinoglu AU, Wu L, Read E, Manoff M, Frame M (2011) Data sharing by scientists: practices and perceptions. *PLoS ONE* 6(6):e21101
- Tinkham WT, Mahoney PR, Hudak AT, Domke GM, Falkowski MJ, Woodall CW, Smith AM (2018) Applications of the United States Forest Inventory and Analysis dataset: A review and future directions. *Can J for Res* 48(11):1251–1268
- USDA Forest Service Manual 2900 – Invasive Species Management. Chapter Zero Code (2011) Washington D.C. <https://www.fs.usda.gov/im/directives/dughtml/fsm2000.html>
- Václavík T, Meentemeyer RK (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Divers Distrib* 18(1):73–83
- Van Schmidt ND, Shyvers JE, Saher DJ, Tarbox BC, Heinrichs JA, Aldridge CL (2022) A user guide to selecting invasive annual grass spatial products for the western United States (No. 2022–3001). US Geological Survey. <https://doi.org/10.3133/fs20223001>
- Vaz AS, Alcaraz-Segura D, Campos JC, Vicente JR, Honrado JP (2018) Managing plant invasions through the lens of remote sensing: a review of progress and the way forward. *Sci Total Environ* 642:1328–1339
- Vaz AS, Alcaraz-Segura D, Vicente JR, Honrado JP (2019) The many roles of remote sensing in invasion science. *Front Ecol Evol* 7:1–10
- Vilà M, Espinar JL, Hejda M, Hulme PE, Jarošík V, Maron JL, Pergl J, Schaffner U, Sun Y, Pyšek P (2011) Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecol Lett* 14(7):702–708
- Wallace RD, Bargerion IV CT, Moorhead DJ, LaForest JH (2018) Information management relevant to invasive species early detection and rapid response programs. National Invasive Species Council Contractor's Report
- Wallace RD, Bargerion CT, Reaser JK (2020) Enabling decisions that make a difference: guidance for improving access to and analysis of invasive species information. *Biol Invasions* 22(1):37–45
- West AM, Kumar S, Wakie T, Brown CS, Stohlgren TJ, Laituri M, Bromberg J (2015) Using high-resolution future climate scenarios to forecast *Bromus tectorum* invasion in Rocky Mountain National Park. *PLoS ONE* 10(2):e0117893
- Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J, da Silva SLB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJG, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PAC, Hoofst R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone S, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3(1):1–9
- Wiser SK (2016) Achievements and challenges in the integration, reuse and synthesis of vegetation plot data. *J Veg Sci* 27(5):868–879

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.