esa ECOSPHERE

Taxonomic error rates affect interpretations of a national-scale ground beetle monitoring program at National Ecological Observatory Network

Lauren Egli,[1,†] Katherine E. LeVan,[2] and Timothy T. Work[1]

[1]*Département des sciences biologiques, Université du Québec à Montréal, CP 8888, succursale Centre-ville, Montréal, Québec H3C 3P8 Canada*
[2]*National Ecological Observatory Network, 1685 38th Street #100, Boulder, Colorado 80301 USA*

Citation: Egli, L., K. E. LeVan, and T. T. Work. 2020. Taxonomic error rates affect interpretations of a national-scale ground beetle monitoring program at National Ecological Observatory Network. Ecosphere 11(4):e03035. 10.1002/ecs2.3035

Abstract. Parataxonomists are responsible for taxonomic identifications in large-scale biodiversity monitoring programs. However, they may lack formal taxonomic training, and thus, quantifying error rate in identification is paramount for evaluating data quality of larger biomonitoring efforts. In large-scale biomonitoring in particular, parataxonomist error rate could vary among regions with different species richness and composition. Here, we tested whether error rates in identification of ground beetles (Coleoptera: Carabidae) by parataxonomists increased in regions with greater species richness throughout the National Ecological Observatory Network (NEON), a national biomonitoring network spanning the United States. We compared identifications made by both parataxonomist and experts of 33,516 specimens collected between 2013 and 2017 from 18 ecoclimatic regions and analyzed error rates across ecoclimatic regions as a function of total richness identified by taxonomic experts. We then compared the additional level of taxonomic support that would be required to resolve identifications to species-level identifications. We demonstrated the extent to which parataxonomist error rate can affect interpretation of common objectives of biomonitoring results, such as comparisons of species richness between ecoclimatic regions and capacity to identify target species of interest such as non-indigenous species. Overall parataxonomist error rate was 11.1% and did not increase in regions with greater species richness. Expert taxonomists were required to resolve parataxonomist identifications to species in an additional 16% of specimens. With an average error rate of 11.1%, species richness estimates based on parataxonomists generally mirrored richness determined by experts. However, parataxonomist error rates as low as 5% were sufficient to misrepresent the gradient of species richness across ecoclimatic regions. Parataxonomist errors also led to false detection/missed detections of non-indigenous species. As error rates were not influenced by increasing species richness, our study suggests that parataxonomists may be used consistently in large-scale biomonitoring efforts to amplify the abilities of taxonomists by increasing the quantity and speed in which specimens are processed. However, our study also highlights that due to parataxonomist limitations, a subset of their materials must be regularly verified by professionals to ensure the quality of data collected.

Key words: biomonitoring; ground beetles; invasive species monitoring; National Ecological Observatory Network; parataxonomists.

Received 12 July 2019; revised 27 September 2019; accepted 10 December 2019. Corresponding Editor: Sudeep Chandra.

† **E-mail:** egli.lauren@courrier.uqam.ca

Introduction

Comprehensive and accurate biodiversity monitoring is essential for documenting decline and loss of species (Pereira et al. 2005) but can be difficult to implement over large spatial scales or long-term timeframes. As a consequence, national-scale monitoring efforts have been slow to develop. However, such programs are in place for several countries including the United States, Sweden, and Denmark (Friberg et al. 2005, Ståhl et al. 2011, Kao et al. 2012). In addition to extensive logistical challenges related to sampling and data collection, large-scale monitoring programs face significant constraints related to the availability of taxonomic expertise and timely identification of the large number of specimens collected (Gewin 2002, Abadie et al. 2008, Didham et al. 2013). These limitations have led to the use of parataxonomists in large-scale biodiversity monitoring of speciose taxa such as terrestrial invertebrates (Janzen 1991, 2004, Oliver and Beattie 1993, Derraik et al. 2002).

Parataxonomists do not hold formal academic training often associated with expert taxonomists and are responsible for the initial determination of specimens in monitoring programs. Parataxonomists typically learn to identify specimens by working through collected material with the aid of reference collections, database images, and published identification materials, often with limited formal instruction or oversight by taxonomic experts (Janzen 1991, 2004, Janzen et al. 1993). Parataxonomists have been implicated in biodiversity monitoring efforts encompassing a variety of invertebrate taxa including moths and butterflies (Janzen 1988, 2000, Janzen et al. 1993, 1998, Janzen and Hallwachs 2011), leaf-chewing and sap-sucking insects (Novotny and Basset 1998, 1999, Basset et al. 2000, 2001, Novotny et al. 2002), ants (Oliver and Beattie 1993, Longino and Colwell 1997), spiders (Oliver and Beattie 1993), aquatic macroinvertebrates (Cranston and Hillman 1992, Fore et al. 2001), and parasitoids (Schauff and Janzen 2001).

In 2013, a national-scale biodiversity monitoring program was initiated throughout 18 ecoclimatic zones (termed domains) throughout the United States as part of the National Ecological Observatory Network (Kao et al. 2012, Schimel 2013, NEON 2018). Within this program, ground beetles (Coleoptera: Carabidae) are one group of study (Hoekman et al. 2017). Ground beetles are often chosen for biodiversity monitoring programs because they are easily sampled, they are implicated in a wide range of ecological functions, and they often respond rapidly to environmental changes (Work et al. 2008, Kotze et al. 2011, Cameron and Leather 2012). Relatively low-cost sampling schemes involving pitfall trapping frequently result in a large sample size and relatively long species lists. As with other large-scale monitoring efforts, the initial identifications of ground beetle specimens collected by NEON staff are made by parataxonomists.

While the work of parataxonomists can result in accurate identifications (Basset et al. 2004, Janzen and Hallwachs 2011), there will necessarily be limitations. Increased sampling in an extensive biodiversity monitoring scheme will likely encounter undescribed species or rarer taxa that may not be included in regional keys or reference collections provided to parataxonomists. In addition, unlike expert taxonomists, parataxonomists do not benefit from an extended period of taxonomic training before they are employed to begin identifications. For these reasons, parataxonomists are expected to have error rates in identification that are higher than expert taxonomists. Parataxonomists are additionally expected to encounter higher instances of limitations in their abilities to identify a specimen to species, restricting some identifications to a broader taxonomic rank, such as genus or family level. However, despite the limitations, the use of parataxonomists may allow large-scale biodiversity monitoring programs to process a larger quantity of specimens more quickly than relying on expert taxonomists alone. This distinction will become increasingly important as the need to monitor serious threats to biodiversity, such as introduction and spread of invasive species, gains attention (Karatayev et al. 2009, Kenis et al. 2009).

Here, we report on the overall accuracy of parataxonomists, as well as the cumulative limitations in parataxonomist identifications (those which are only identified to family or genus level), of ground beetles collected throughout NEON. We hypothesized that identification error rate by parataxonomists would increase in more species-rich domains as would parataxonomist limitations in identifications.

Materials and Methods

Ground beetle collection and identification

We evaluated accuracy of parataxonomist identifications from specimens collected between 2013 and 2017 which were sampled across 18 domains of the NEON project. Sampling periods varied among domains depending on seasonal constraints on carabid activity. In each domain, ground beetles were sampled from either two or three sampling sites using 40 pitfall traps per site (Fig. 1). The NEON project provides parataxonomists with a series of detailed protocols describing the procedure by which samples should be collected and processed (LeVan et al. 2018). In brief, samples are collected on a biweekly basis throughout the growing season at each site. Samples are collected and sorted by temporary field technicians to separate non-carabid invertebrate and vertebrate bycatch from ground beetles. For each sample, all ground beetles then receive further identifications by parataxonomists.

National Ecological Observatory Network parataxonomists hold at a minimum a bachelor's degree in ecology, environmental sciences, or a related scientific discipline and have one or more years of relevant experience. However, staff may not have any specific entomology or taxonomy training prior to working at NEON. Training in carabid identification largely occurs on the job

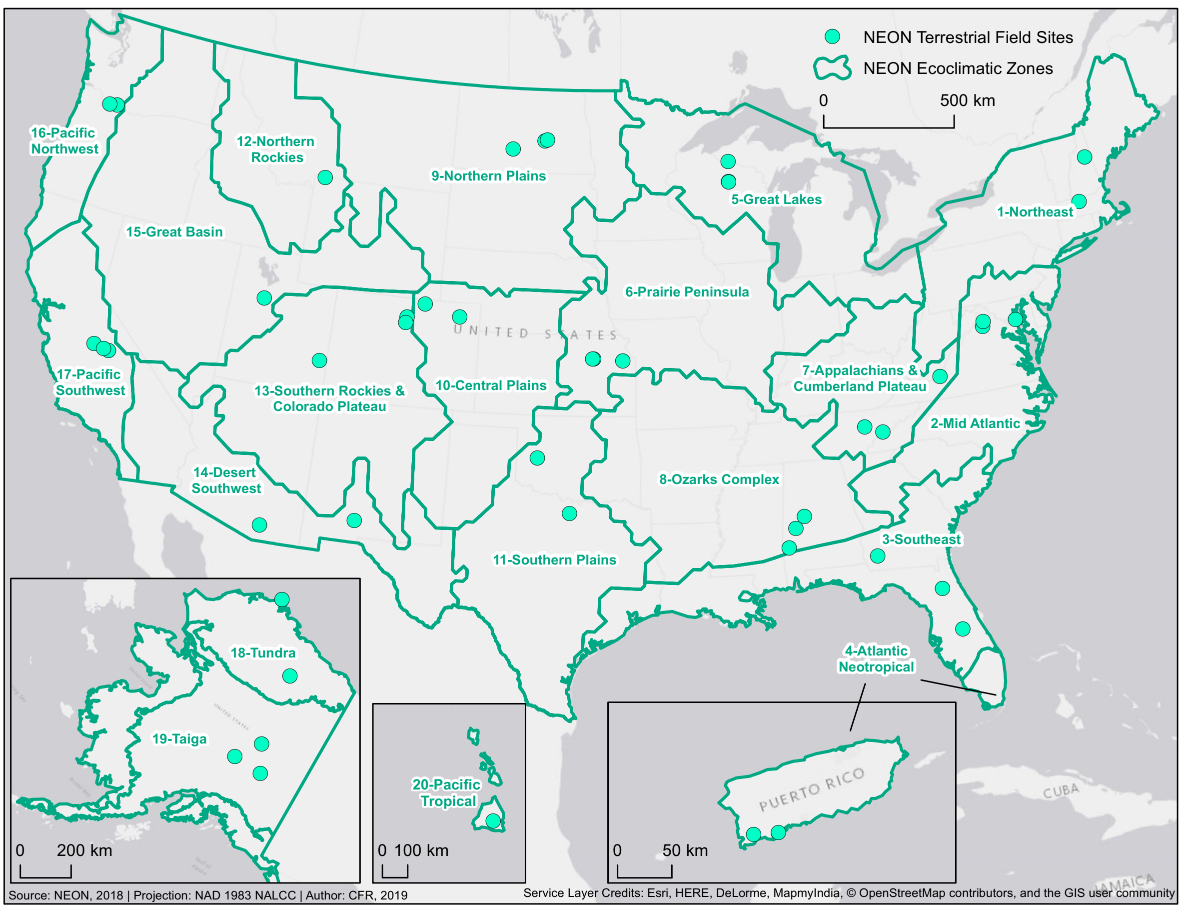

Fig. 1. Distribution of ecoclimatic zones and field site locations for the National Ecological Observatory Network project.

and is supplemented by access to regional keys, identification guides, and access to local reference collections of verified specimens. After identification by a parataxonomist, a subset of identified carabids are pinned and sent to expert taxonomists for verification. In the initial phases of the program, NEON allocated funds for secondary verification of 36% of parataxonomist-identified specimens by expert taxonomists.

Expert taxonomists are individuals who hold a PhD in entomology, trained in taxonomy/systematics of ground beetles, and have robust publishing records demonstrating their expertise in the field of carabid systematics. Once verified by expert taxonomists, a subset of confirmed specimens for each species are returned annually to parataxonomists to expand local reference collections.

Data treatment

We downloaded 841 zipped data files from the NEON data portal (http://data.neonscience.org) on 1 April 2019. These zip files contained monthly ground beetle collection data from each NEON site sampled from 2013 to 2017. Each.zip file contains multiple files. For our analysis, we used only.csv files pertaining to parataxonomist identifications which were subjected to expert validation. We then merged.csv files into two consolidated files. As of 1 April 2019, the parataxonomist files contained 48898 records of parataxonomist-identified, pinned specimens. Expert taxonomist data initially contained 33,567 records of verified pinned specimens; however, 41 duplicate pairs of records were present in the data. We were able to resolve 31 duplicate pairs by deletion. The remaining 10 duplicate pairs could not be resolved because of multiple, conflicting identifications by experts and were removed from the dataset. The final corrected expert dataset contained 33,516 records. It is important to note that these data represent a subset of the entire NEON ground beetle dataset. At the time of writing this paper, there were an additional 50,000+ carabid specimens identified by parataxonomists stored in ethanol and not subject to secondary review by experts. For this study however, we used only the 33,516 records of individuals validated by taxonomic experts to assess the accuracy of parataxonomists.

Quantifying parataxonomist error rate and relative contribution of expert taxonomists

We compared inconsistencies in taxonomic identifications made by parataxonomists and experts on the same specimens. Inconsistencies can derive from either limitation in the taxonomic skill of the parataxonomist or wholly erroneous taxonomic assignments made by the parataxonomist. We defined taxonomic limitation as any case where a parataxonomist made an identification at a broader taxonomic resolution (such as genus), which was later resolved to a finer level by the expert (such as species). We defined taxonomic error as cases where parataxonomists assigned a specimen the incorrect family, genus, species, or subspecies, which was later corrected by experts.

Evaluating implication of parataxonomist errors on biodiversity monitoring objectives

We evaluated the implications of parataxonomist error rates by comparing unverified identifications (parataxonomist data) with verified identifications (expert data) to address two common goals of biodiversity monitoring: (1) comparisons of regional differences in species richness and (2) capacity to detect changes in taxa of interest such as changes in the occurrence of introduced species. We compared rank order of species richness among NEON domains using unverified and verified identifications to detect discrepancies or reversals in domain ranks and to determine the level at which parataxonomic error rate would be sufficiently high to obscure actual trends in richness among regions. We also compared the capacity to detect differences in the number of non-indigenous ground beetles using NEON data by comparing verified and unverified identifications across ecological domains. Non-indigenous species of ground beetles were determined by Bousquet (2012). We consider comparisons of the number of non-indigenous individuals to be a finer scale application of the NEON data where the relative abundance of a non-indigenous species may dictate policy actions (i.e., quarantine or control measures). It is important to note that at present we are unaware of any quarantine/control measures prescribed for adventive ground beetles.

Rather, we use this as a test case to relate parataxonomist error rate to a possible application of large-scale monitoring data.

Statistical analysis

Parataxonomist error rate (ratio of erroneous parataxonomist identifications to total specimens collected) and the limitation rate (ratio of partially identified specimens to total specimens collected) were analyzed using a beta regression model. Beta regression has been recommended for analyzing percentage data (Ferrari and Cribari-Neto 2004). Our analysis was carried out in R 3.5.1 and RStudio 1.1.456 using the package *betareg* (Cribari-Neto and Zeileis 2010, R Core Team 2018). To describe implications of parataxonomist errors for comparisons of species richness, we compared differences in rank-order correlation between the number of species identified by parataxonomists and experts across domains using Kendall's tau (Noether 1981, R Core Team 2018).

We observed three domains which relied heavily on expert taxonomists to identify specimens to the species level (D15, Utah; D18 and D19, Alaska) and two domains with extreme error rates in identification (D04, Puerto Rico; D17, California). In these cases, we judged parataxonomic limitations and error rates to be caused by factors other than species richness observed within a domain. In the case of D04, high error rates in identifications were attributed to repeated misidentifications of *Discoderus*, a particularly difficult genus to identify due to an absence of recent, revised taxonomic keys to species or any recent descriptions (Bousquet 2012, P. Messer, *personal communication*). High error rates at D17 were attributed to the short timeframe of sampling, and the extremely small quantity of specimens was collected. It is likely that D17 had not yet been provided comparable identification resources, such as a local reference collection. At D15 (Utah), D18 and D19 (Alaska) unusually high numbers of specimens were identified to family only (D15) or genus only (D18/19), indicating significantly less energy spent on specimen identifications in these regions as compared to others. For these reasons, these domains were removed from our beta regression model as outliers.

Results

After removing outliers, parataxonomist error rate did not significantly increase with greater overall species richness within a domain (Fig. 2). Overall errors in identification at family, genus, species, or subspecies level were 11.1% (±9.2% standard error [SE]). Errors occurred most frequently within the genera *Pterostichus, Cyclotrachelus,* and *Harpalus*. Taxonomic limitations of parataxonomist at family, genus, or species level affected an additional 16% (±10.8% SE). However, when subspecies identifications were excluded, limitations accounted for only an additional 11.6% (±8.7% SE) of specimen identifications. Limitations in parataxonomist identifications occurred most frequently within *Pterostichus*, *Amara,* and *Harpalus*.

A total of 33,516 specimens identified by parataxonomists were verified by experts (Table 1). In 1536 instances (4.6% of identifications), expert taxonomists were unable to provide species names for specimens. Only five of these were listed as unidentifiable due to damaged specimens. Species-level identifications of 32 individuals, identified as *Pterostichus, Loxandrus*, *Amara,* and *Brachinus,* were hindered by life stage of the specimens (teneral or immature). Four-hundred twenty-eight individuals (410 *Pterostichus*, 6 *Rhadine,* and 12 *Selenophorus*) were described as new taxa to be included in forthcoming manuscripts by taxonomic experts contracted by the NEON project. Remaining specimens, in some instances, were characterized with reference to other species or by distinctive morphological characters. The majority of specimens that were unable to be resolved to species by the expert taxonomists (85%) came from four domains (D04, Puerto Rico; D02, Maryland/Virginia; D14, Arizona/New Mexico; and D13, Colorado/Utah).

Species richness determined by parataxonomists generally mirrored richness determined by experts (Kendall's tau = 0.88; Fig. 3). However, a comparatively low error rate did not necessarily indicate an accurate estimation of species richness. In D18 (Alaska), average error rate in identification was relatively low (6.3%) but species richness was underestimated by 50%. In contrast, higher average error rates found at D14 (Arizona/New Mexico, 28.8%) underestimated

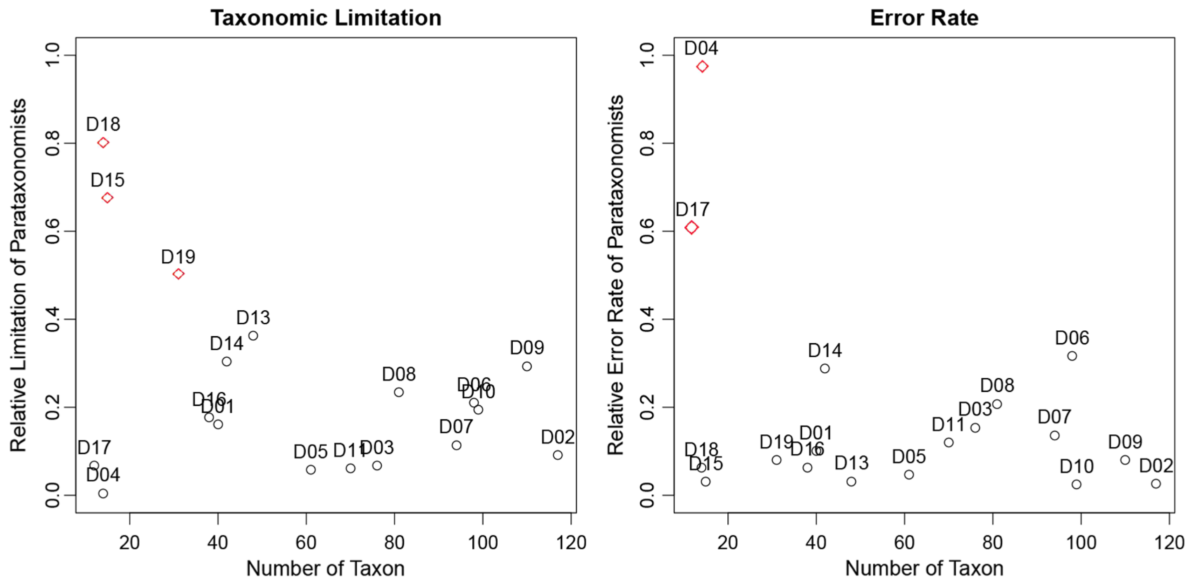

Fig. 2. Taxonomic limitations and error rates of parataxonomists by species richness. Domains which were removed from the beta regression model as outliers are denoted with a red diamond. After removal of outliers, the relationship between relative limitations of parataxonomist identifications and number of taxon had a slope of 0.002 (±0.006) with a nonsignificant *P*-value of 0.766. The relationship between relative error rate of parataxonomist identifications and number of taxon also had a slope of 0.008 (±0.006) and nonsignificant *P*-value of 0.149.

Table 1. Domain site locations by state, trapping effort, specimen abundance, richness, limitations in parataxonomist identifications at family, genus, and species-level, and incorrect parataxonomist identifications separated by errors at family, genus, species, and subspecies level by domain.

					Limitations				Error				
Domain	Location	Trap days†	Total specimens	Total taxon	Family	Genus	Species	Total	Family	Genus	Species	Sub-species	Total
D01	MA, NH	41,977	2991	40	144	3	336	483	4	0	300	0	304
D02	MD, VA	59,425	5289	117	170	302	12	484	0	0	134	6	140
D03	FL, GA	71,158	1562	76	0	101	4	105	0	66	175	0	241
D04	PR	27,874	600	14	0	3	0	3	0	3	582	0	585
D05	MI, WI	39,594	3729	61	209	5	3	217	0	0	167	10	177
D06	KS	39,018	2059	98	145	49	240	434	3	41	587	21	652
D07	VA, TN	41,258	2694	94	89	188	32	309	1	80	213	75	369
D08	AL	45,752	1575	81	272	5	92	369	1	0	296	31	328
D09	ND	32,455	1827	110	527	9	1	537	0	0	118	28	146
D10	CO	45,940	5188	99	557	263	190	1010	0	1	128	0	129
D11	OK, TX	32,187	750	70	0	39	7	46	0	6	84	0	90
D13	CO, UT	26,420	2250	48	279	120	417	816	0	1	69	2	72
D14	AZ, NM	34,829	733	42	25	196	2	223	0	17	194	0	211
D15	UT	16,758	608	15	345	37	29	411	0	0	19	0	19
D16	OR, WA	6966	503	38	3	26	60	89	0	2	27	3	32
D17	CA	9238	253	12	0	17	0	17	0	20	134	0	154
D18	AK	2226	522	14	0	419	0	419	0	11	1	21	33
D19	AK	16,818	383	31	0	192	0	192	0	1	29	1	31
Total		589,893	33,516	1060	2765	1974	1425	6164	9	249	3257	198	3713

† Total number of days sampled by each trap within a domain.

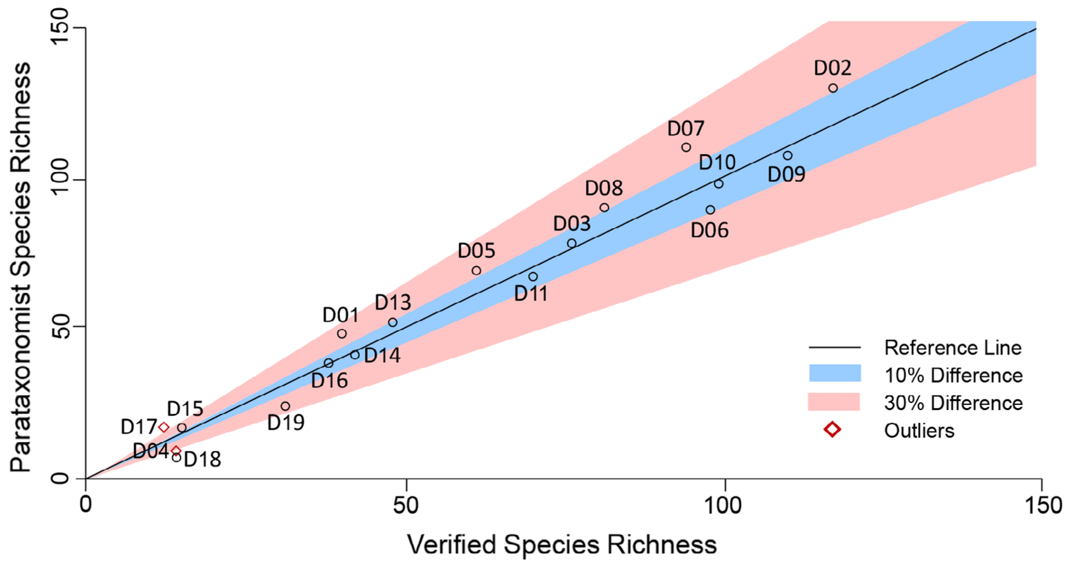

Fig. 3. Comparison of species richness identified by parataxonomists and expert taxonomist across 18 ecological domains at the National Ecological Observatory Network project. Solid line is a reference line which indicates whether parataxonomists over- or underestimated total species richness reported by expert taxonomists. Blue and red regions correspond with 10% and 30% difference between parataxonomist and expert species richness. Domains which were removed from the beta regression model as outliers in parataxonomist error rates are denoted with a red diamond.

species richness by only 2.4%. In several cases, parataxonomist errors affected the order of species richness among regions. In the most severe case, overestimation of species richness by parataxonomists at D07 (California; 110 parataxonomist-identified species, 94 species confirmed by experts) incorrectly indicated D07 to have a higher species richness than three other domains (D06 [Kansas; 98 confirmed species], D10 [Colorado; 99 confirmed species], and D09 [North Dakota; 110 confirmed species]). However, a parataxonomist error rate as low as 4.8% was also sufficient to misrepresent larger gradients in species richness, falsely concluding that D05 (Michigan/Wisconsin) had higher richness than D11 (Oklahoma/Texas).

The distribution of non-indigenous species was fairly consistent between parataxonomist and expert taxonomists. Non-indigenous species were correctly identified by parataxonomists at all sites in which they were present with the exception of one instance of *Pterostichus melanarius melanarius* (Illiger) at D10 (Colorado) and single specimens of *Bembidion lampros* (Herbst) at D16 (Oregon/Washington) and D17 (California). In these three instances, non-indigenous species were misidentified as native species, incorrectly indicating that invasive species were absent in these locations. Parataxonomists also erroneously identified three individuals of *Amara lindrothi* Hieke (a native species) as *Amara apricaria* (Paykull) (an invasive species) at D13 (Colorado/Utah) and incorrectly identified four instances of *Amara flebilis* (Casey; a native species) as *Amara aenea* (DeGeer) and *Amara familiaris* (Duftschmid; invasive species) at D06 (Kansas). In these cases, parataxonomists indicated a presence of non-indigenous species where they were actually none. In addition, in two instances where parataxonomists correctly identified the presence of non-indigenous species at a site, the abundance of the non-indigenous species population was significantly underestimated. Abundance was over 400% higher in *P. melanarius melanarius* at D09 (North Dakota) and D05 (Michigan/Wisconsin) when expert taxonomist identifications were compared to the identifications made by parataxonomists.

Discussion

Accuracy and limitations of parataxonomist identifications

Contrary to our initial hypothesis, parataxonomist error rate for ground beetle identification was consistently 11% across regions and did not increase with increasing ground beetle species richness. Other studies that have relied on parataxonomists for arthropod identifications have reported similar error rates; however, formal verifications of parataxonomist error rates are limited. For example, in a study of parataxonomist identification capabilities in Australia, Oliver and Beattie (1993) reported error rates of identification of topical ants at 6% and spiders at 13%. In Oliver and Beattie's study, parataxonomists received only three hours of initial training from Oliver, who was previously trained by an experienced taxonomist, and were provided with a reference collection of specimens. Parataxonomists identified 1850 spider and 1233 ant specimens to recognizable taxonomic units (RTUs) rather than formal taxonomic classifications. Parataxonomists identified specimens of these taxa into 136 RTUs, which were later reverified by professional taxonomists as 126 different species. Parataxonomists at NEON are provided with similar resources as in Oliver and Beattie's study, often receiving initial training from fellow parataxonomists located at the same domain, and are provided a local reference collection. Unlike Oliver and Beattie however, NEON's parataxonomists are instructed to identify specimens to the lowest taxonomic rank possible given their capabilities and are provided with any taxonomic keys available for the region. Considering the comparable error rates between the two methods, identifying specimens directly to recognized taxonomic classifications, where taxonomic keys are available, rather than RTUs may speed accessibility of the data and decrease reliance on expert taxonomists to translate RTU identifications into formally recognized classifications.

Inability to resolve specimens to the lowest taxonomic rank including subspecies was a significant restriction across the NEON parataxonomist data and affected 16% of the identifications. However, limitations for species identifications affected only 11.6%. We expect the number of individuals left at a higher taxonomic rank to decrease with time as the local teaching collections and identification references grow. As this report covers the earliest NEON data available, we expect these numbers should improve over time as identification resources accumulate and taxonomist feedback corrects identification errors. It is important to note that expert taxonomists also suffered from limitations which are inherent to large-scale biodiversity monitoring projects, such as the collection of little studied or undescribed species. The inability of professional taxonomist to identify specimens to species affected 4.6% of the specimens sent for verification. As a result, it is unlikely that parataxonomists will achieve perfect accuracy.

Limitations of data derived by parataxonomists

Parataxonomist error rates, if sufficiently large, could impact conclusions drawn from biomonitoring programs. Correct assessment of species richness among ecoregions is arguably one of the most basic applications of NEON data. With an average identification error rate of 11%, parataxonomist and expert assessments of species richness among domains were relatively consistent. However, in some cases, conclusions regarding species richness among domains based solely on parataxonomist data would misrepresent species richness among regions. In these cases, nuanced questions and comparisons, such as which community has a higher proportion of rare species, may prove unreliable. In addition, these misrepresentations could affect efforts to accurately identify regions with the highest biodiversity, a common conservation strategy (Myers et al. 2000). However, error rates in species identification do not necessarily have to affect estimations of species richness if parataxonomists can accurately identify specimens to morphospecies. In cases like D04 (Puerto Rico), where identification resources are limited, parataxonomists may still produce accurate estimations of the species richness regardless of errors in species-level identifications through the use of RTUs rather than attempting to identify individuals directly to species (Basset et al. 2000).

Differences in species richness among ecoregions were likely affected by differences in sampling intensities. The three regions with the highest verified species richness (D02, Maryland/

Virginia; D09, North Dakota; and D10, Colorado) also had relatively high sampling intensity. In contrast, ecologically diverse regions such as D16 (Oregon/Washington) and D17 (California) had relatively limited trapping effort and low species richness. We anticipate that with further sampling, species richness will increase in ecoregions with longer sampling seasons and those with increased landscape heterogeneity. For example, D02, which had one of the most intense sampling efforts and the highest species richness, also contained a wide variety of sampled habitat types including cultivated crops, pasture lands, evergreen and deciduous forests, and an area bordering a coastal inlet.

Finer scale applications of NEON biomonitoring data to detect non-indigenous species were also affected by parataxonomist error rates. Parataxonomists produced a reasonably accurate distribution map of invasive ground beetle species across the United States which was sufficient to gather a coarse picture of non-indigenous carabid distributions. However, the numbers of invasive carabid populations were underestimated in multiple locations and missed entirely in several others. In these cases, conclusions drawn from this data could mislead strategies related to the monitoring and/or control of invasive species. In many cases, the timely identification of invasive species is paramount in preventing these individuals from forming established populations which become increasingly challenging to eliminate (Anderson 2005). By failing to enact necessary control procedures in areas which are invaded by non-indigenous species, but do not appear so according to parataxonomist identifications, unchecked populations of invasive species could negatively affect native biodiversity (Crowder and Snyder 2010, Caffrey et al. 2014).

Optimizing oversight by expert taxonomists

Our study made use of a relatively large number of identifications made by NEON parataxonomists which were verified by experts. Intensive verification is prudent in the initial phases of a project but is likely infeasible over the long term. Establishment of local resources (species keys, teaching collections, species lists, etc.) may reduce the need for secondary verification by experts. In order to optimize oversite of expert taxonomists, verification of a small subsample of specimens with a focus on particularly cryptic or specious groups (e.g., *Pterostichus, Harpalus, Amara*) may be sufficient to assess parataxonomist error rates. Oliver and Beattie (1993) examined a subset of 10 random specimens per RTU identified by parataxonomist to analyze whether a subset of the data could estimate with reasonable precision parataxonomist accuracy. This subset, 16% of their total spider and ant collections, was found to capture 98.6% of the species collected in the full study. Nevertheless, some degree of taxonomic feedback should continue throughout even the latter stages of a large-scale study to address rare or cryptic species which are collected infrequently. Continuing taxonomic feedback will also be particularly essential in the case of new hires who will require additional support to minimize their error rates.

Reducing parataxonomist error rate

Parataxonomist accuracy is heavily reliant on the amount of training and feedback received. In situations without input from professional taxonomists, error rates increase. For example, Baraloto et al. (2007) found 87% error rates of parataxonomist identifications of woody tree species in Bolivia where only common names were utilized and no consultation with professional taxonomists occurred. In contrast, in studies where parataxonomists were initially provided with simplified determination keys for plants, parataxonomist error rates were reduced to 19.5% (Abadie et al. 2008). The training of parataxonomists should not occur in a vacuum, and continued feedback from professional taxonomists on the accuracy rates, and where weaknesses lie, is paramount to creating robust data (Basset et al. 2004). Error rates of NEON parataxonomists could likely be further reduced with additional initial training by experts, which would improve parataxonomist abilities to interpret taxonomic keys (Derraik et al. 2002). In addition, in regions with abundant populations of particularly homogenous taxa (i.e., *Pterostichus, Cyclotrachelus,* and *Harpalus*), parataxonomists would benefit from advanced taxonomic training, such as in the dissection of specimen genitalia, which has been shown to resolved identification issues in these cases (Basset et al. 2000, 2001). Supplementary tools such as taxonomic keys specific to a given region or locale, high-quality digital photographs, and a robust verified reference collection could also help further reduce parataxonomist error rates (Gotelli 2004). These tools will also maintain continuity in biomonitoring programs with significant personnel turnover, as positions for parataxonomists are rarely permanent (Janzen 2004).

Conclusions

Parataxonomists can identify specimens with consistent accuracy across regions with increasing species richness; however, a constant error rate will likely be a persistent challenge for large-scale biodiversity monitoring. Whether these errors will have significant negative effects on the interpretation of large-scale biomonitoring data will depend on the goals of monitoring programs and the research questions being investigated. Nevertheless, to ensure robust data a subset of parataxonomist identifications should be regularly verified throughout their employment and increased taxonomic support should be directed toward regions demonstrating particular difficulties. To minimize weaknesses, resources such as taxonomic keys, regular reports of errors in identifications, local reference collections, and initial training should be provided.

Acknowledgments

The National Ecological Observatory Network is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle Memorial Institute. This material is based in part upon work supported by the National Science Foundation through the NEON Program. Funding provided through the Natural Sciences and Engineering Research Council of Canada NSERC Discovery Grant RGPIN-2015-03826. Lauren Egli and Timothy Work conceived the ideas, designed methodology, analyzed the data, and led the writing of the manuscript. Katherine LeVan revised the manuscript critically for important intellectual content. All authors have no conflicts of interest.

Literature Cited

Abadie, J. C., C. Andrade, N. Machon, and E. Porcher. 2008. On the use of parataxonomy in biodiversity monitoring: a case study on wild flora. Biodiversity and Conservation 17:3485–3500.

Anderson, L. W. J. 2005. California's reaction to *Caulerpa taxifolia*: a model for invasive species rapid response. Biological Invasions 7:1003–1016.

Baraloto, C., E. Ferreira, C. Rockwell, and F. Walthier. 2007. Limitations and applications of parataxonomy for community forest management in Southwestern Amazonia. Ethnobotany Research and Applications 5:77–84.

Basset, Y., E. Charles, D. S. Hammond, and V. K. Brown. 2001. Short-term effects of canopy openness on insect herbivores in a rain forest in Guyana. Journal of Applied Ecology 38:1045–1058.

Basset, Y., V. Novotny, S. Miller, and R. Pyle. 2000. Quantifying biodiversity: experience with parataxonomists and digital photography in Papua New Guinea and Guyana. BioScience 50:899–908.

Basset, Y., V. Novotny, S. Miller, G. D. Weiblen, O. Missa, and A. J. A. Stewart. 2004. Conservation and biological monitoring of tropical forests: the role of parataxonomists. Journal of Applied Ecology 41:163–174.

Bousquet, Y. 2012. Catalogue of Geadephaga (Coleoptera, Adephaga) of America, North of Mexico. ZooKeys 245:1–1722.

Caffrey, J. M., et al. 2014. Tackling invasive alien species in Europe: the top 20 issues. Management of Biological Invasions 5:1–20.

Cameron, K. H., and S. R. Leather. 2012. How good are carabid beetles (Coleoptera, Carabidae) as indicators of invertebrate abundance and order richness? Biodiversity and Conservation 21:763–779.

Cranston, P. S., and T. Hillman. 1992. Rapid assessment of biodiversity using biological diversity technicians. Department of Arts, Sport, Environment and Territories, Murray Darling Freshwater Research Centre, Albury, New South Wales, Australia.

Cribari-Neto, F., and A. Zeileis. 2010. Beta regression in R. Journal of Statistical Software 34:1–24.

Crowder, D. W., and W. E. Snyder. 2010. Eating their way to the top? Mechanisms underlying the success of invasive insect generalist predators. Biological Invasions 12:2857–2876.

Derraik, J. G. B., G. P. Closs, K. J. M. Dickinson, P. Sirvid, B. I. P. Barratt, and B. H. Patrick. 2002. Arthropod morphospecies versus taxonomic species: a case study with Araneae, Coleoptera, and Lepidoptera. Conservation Biology 16:1015–1023.

Didham, R. K., O. R. Edwards, S. R. Leather, and Y. Basset. 2013. Arthropod diversity and the future of all-taxa inventories. Insect Conservation and Diversity 6:1–4.

Ferrari, S. L. P., and F. Cribari-Neto. 2004. Beta regression for modelling rates and proportions. Journal of Applied Statistics 31:799–815.

Fore, L. S., K. Paulsen, and K. O'Laughlin. 2001. Assessing the performance of volunteers in monitoring streams. Freshwater Biology 46:109–123.

Friberg, N., A. Baattrup-Pedersen, M. L. Pedersen, and J. Skriver. 2005. The new Danish stream monitoring programme (Novana)-Preparing monitoring activities for the water framework directive era. Environmental Monitoring and Assessment 111:27–42.

Gewin, V. 2002. All living things, online. Nature 418:362–363.

Gotelli, N. J. 2004. A taxonomic wish-list for community ecology. Philosophical Transactions of the Royal Society B: Biological Sciences 359:585–597.

Hoekman, D., et al. 2017. Design for ground beetle abundance and diversity sampling within the National Ecological Observatory Network. Ecosphere 8:1–17.

Janzen, D. H. 1988. Ecological characterization of a Costa Rican dry forest caterpillar fauna. Biotropica 20:120–135.

Janzen, D. H. 1991. How to save tropical biodiversity: the national biodiversity institute of Costa Rica. National Biodiversity Institute of Costa Rica 3:159–171.

Janzen, D. H. 2000. Costa Rica's Area de Conservacion Guanacaste: a long march to survival through non-damaging biodiversity and ecosystem development. Biodiversity 1:7–20.

Janzen, D. H. 2004. Setting up tropical biodiversity for conservation through non-damaging use: participation by parataxonomists. Journal of Applied Ecology 41:181–187.

Janzen, D. H., and W. Hallwachs. 2011. Joining Inventory by parataxonomists with DNA barcoding of a large complex tropical conserved wildland in northwestern Costa Rica. PLOS ONE 6:1–13.

Janzen, D. H., W. Hallwachs, J. Jimenez, R. Gámez. 1993. The role of the parataxonomists, inventory managers and taxonomists in Costa Rica's national biodiversity inventory. Pages 223–254 *in* W. V. Reid, et al., editors. Biodiversity prospecting. World Resources Institute, Washington, D.C., USA.

Janzen, D. H., M. J. Sharkey, and J. M. Burns. 1998. Parasitization biology of a new species of Braconidae (Hymenoptera) feeding on larvae of Costa Rican dry forest skippers (Lepidoptera: Hesperiidae: Pyrginae). Tropical Lepidoptera 9:33–41.

Kao, R. H., et al. 2012. NEON terrestrial field observations: designing continental-scale, standardized sampling. Ecosphere 3:1–17.

Karatayev, A. Y., L. E. Burlakova, D. K. Padilla, S. E. Mastitsky, and S. Olenin. 2009. Invaders are not a random selection of species. Biological Invasions 11:2009–2019.

Kenis, M., et al. 2009. Ecological effects of invasive alien insects. Biological Invasions 11:21–45.

Kotze, D. J., et al. 2011. Forty years of carabid beetle research in Europe-from taxonomy, biology, ecology and population studies to bioindication, habitat assessment and conservation. ZooKeys 100:55–148.

LeVan, K., N. Robinson, D. Hoekman, and K. Blevins. 2018. TOS protocol and procedure: ground beetle sampling. National Ecological Observatory Network, Boulder, Colorado, USA.

Longino, J. T., and R. K. Colwell. 1997. Biodiversity assessment using structured inventory: capturing the ant fauna of a tropical rain forest. Ecological Applications 7:1263–1277.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.

National Ecological Observatory Network. 2018. Data Products: NOEN.DP1.10022.001. http://data.neonscience.org

Noether, G. 1981. Why Kendall Tau? Teaching Statistics 3:41–43.

Novotny, V., and Y. Basset. 1998. Seasonality of sap-sucking insects (Auchenorrhyncha, Hemiptera) feeding on *Ficus* (Moraceae) in a lowland rain forest in Papua New Guinea. Oecologia 115:514–522.

Novotny, V., and Y. Basset. 1999. Body size and host plant specialization: a relationship from a community of herbivorous insects on *Ficus* from Papua New Guinea. Journal of Tropical Ecology 15:315–328.

Novotny, V., Y. Basset, S. E. Miller, P. Drozd, and L. Cizek. 2002. Host specialization of leaf-chewing insects in a New Guinea rainforest. Journal of Animal Ecology 71:400–412.

Oliver, I., and A. J. Beattie. 1993. A possible method for the rapid assessment of biodiversity. Conservation Biology 7:562–568.

Pereira, H. M., B. Reyers, M. Watanabe, E. Bohensky, S. Foale, and C. Palm. 2005. Condition and trends of ecosystem services and biodiversity. Pages 171–201 *in* J. Plesnik, V. Neronov, and B. O'Regan, editors. Ecosystem and human well-being: multiscale assessments. Island Press, Washington, D.C., USA.

R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Schauff, M. E., and D. H. Janzen. 2001. Taxonomy and ecology of Costa Rican Euplectrus (Hymenoptera: Eulophidae), parasitoids of caterpillars (Lepidoptera). Journal of Hymenoptera Research 10: 181–230.

Schimel, D.. 2013. NEON Observatory Design. National Ecological Observatory Network, Boulder, Colorado, USA. https://data.neonscience.org/documents/10179/11206/NEON.DOC.000001vD

Ståhl, G., et al. 2011. National Inventory of Landscapes in Sweden (NILS) – scope, design, and experiences from establishing a multiscale biodiversity monitoring system. Environmental Monitoring and Assessment 173:579–595.

Work, T. T., M. Koivula, J. Klimaszewski, D. Langor, J. Spence, J. Sweeney, and C. Hébert. 2008. Evaluation of carabid beetles as indicators of forest change in Canada. Canadian Entomologist 140: 393–414.

Data Accessibility

Data available from the NEON data portal at http://data.neonscience.org