



HAL
open science

Widespread imprecision in estimates of the economic costs of invasive alien species worldwide

Philip E Hulme, Danish A Ahmed, Phillip J Haubrock, Brooks A Kaiser, Melina Kourantidou, Boris Leroy, Shana M Mcdermott

► **To cite this version:**

Philip E Hulme, Danish A Ahmed, Phillip J Haubrock, Brooks A Kaiser, Melina Kourantidou, et al.. Widespread imprecision in estimates of the economic costs of invasive alien species worldwide. *Science of the Total Environment*, 2024, 909, 10.1016/j.scitotenv.2023.167997 . hal-04633043

HAL Id: hal-04633043

<https://hal.science/hal-04633043v1>

Submitted on 3 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Widespread imprecision in estimates of the economic costs of invasive alien species worldwide

Philip E. Hulme^{a,*}, Danish A. Ahmed^b, Phillip J. Haubrock^{b,c,d}, Brooks A. Kaiser^e,
Melina Kourantidou^{e,f}, Boris Leroy^g, Shana M. McDermott^h

^a *Bioprotection Aotearoa, Lincoln University, PO Box 85084, Christchurch, Canterbury, New Zealand*

^b *CAMB, Center for Applied Mathematics and Bioinformatics, Gulf University for Science and Technology, Kuwait*

^c *Department of River Ecology and Conservation, Senckenberg Research Institute and Natural History Museum Frankfurt, Clamecystr. 12, 63571 Gelnhausen, Germany*

^d *Faculty of Fisheries and Protection of Waters, South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, University of South Bohemia in České Budějovice, Zátisí 728/II, 389 25 Vodňany, Czech Republic*

^e *MERE, SEBE, University of Southern Denmark, Degnevej 14a, 6705 Esbjerg Ø, Denmark*

^f *Université de Bretagne Occidentale, UMR 6308 AMURE, IUEM, 29280, Plouzané, France*

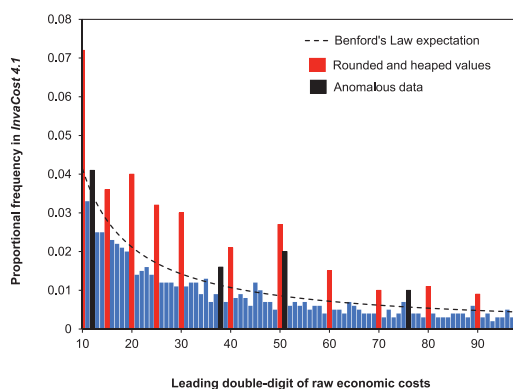
^g *UMR 8067, Biologie Des Organismes Et Écosystèmes Aquatiques (BOREA), Sorbonne Université, Muséum National d'Histoire Naturelle, Université de Caen Normandie, Université Des Antilles, CNRS, IRD, CP26, 43 Rue Cuvier, 75005 Paris, France*

^h *Department of Economics, Trinity University, San Antonio, TX 78216, USA*

HIGHLIGHTS

- The precision of global monetary cost estimates of biological invasions is assessed quantitatively for the first time.
- The distributions of the leading digits of reported cost estimates were compared against expectations from Benford's Law.
- Strong evidence of irregularities in reported costs was found irrespective of the type of cost or perceived reliability.
- The prevalence of round numbers and numerical heaping suggests many cost estimates are imprecise and possibly inflated.
- Tests of Benford's Law indicate a need for greater transparency and rigour when estimating costs of biological invasions.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Sergi Sabater

Keywords:

Auditing

Bayes

Biological invasions

Data quality

ABSTRACT

Several hundred studies have attempted to estimate the monetary cost arising from the management and/or impacts of invasive alien species. However, the diversity of methods used to estimate the monetary costs of invasive alien species, the types of costs that have been reported, and the spatial scales at which they have been assessed raise important questions as to the precision of these reported monetary costs. Benford's Law has been increasingly used as a diagnostic tool to assess the accuracy and reliability of estimates reported in financial accounts but has rarely been applied to audit data on environmental costs. Therefore, the distributions of first,

* Corresponding author.

E-mail address: philip.hulme@lincoln.ac.nz (P.E. Hulme).

<https://doi.org/10.1016/j.scitotenv.2023.167997>

Received 6 April 2023; Received in revised form 12 October 2023; Accepted 19 October 2023

Available online 30 October 2023

0048-9697/© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Non-native species
Scale-invariance

second- and leading double-digits of the monetary costs arising from biological invasions, as reported in the *InvaCost* database, were compared with the null expectations under Benford’s Law. There was strong evidence that the reported monetary costs of biological invasions departed considerably from Benford’s Law and the departures were of a scale equal to that found in global macroeconomic data. The rounding upwards of costs appears to be widespread. Furthermore, numerical heaping, where values cluster around specific numbers was evident with only 901 unique cost values accounting for half of the 13,553 cost estimates within the *InvaCost* database. Irrespective of the currency, the value of 1,000,000 was the most common cost estimate. An investigation of anomalous data entries concluded that non-peer reviewed official government reports need to provide greater detail regarding how costs are estimated. Despite the undeniably high economic cost of biological invasions worldwide, individual records of costs were often found to be imprecise and possibly inflated and this emphasises the need for greater transparency and rigour when reporting the costs of biological invasions. Identifying whether the irregularities found for the costs of biological invasions are general for other types of environmental costs should be a research priority.

1. Introduction

It has been well established that the frequency distribution of the first or leading digit in a set of numbers, rather than being uniform, generally follows a logarithmically decaying distribution (Berger and Hill, 2021; Miller, 2015). This first-digit law, often referred to as Benford’s Law, has been shown to hold across a diverse range of environmental datasets including the travel distances of tropical cyclones (Joannes-Boyau et al., 2015), daily pollen counts in cities (Docampo et al., 2009), annual streamflow of rivers (Nigrini and Miller, 2007), the duration of volcanic eruptions (Geyer and Marti, 2012), the emissions of greenhouse gases per country (Sambridge et al., 2010), as well as phytoplankton abundance in lakes (Ozkundakci and Pingram, 2019).

Despite Benford’s Law being widely applicable, it is not generally known, and thus it is assumed that individuals manipulating data would be unlikely to replicate a realistic digit distribution (Moul and Nye, 2015; Nigrini, 2012; Tam Cho and Gaines, 2007). For this reason Benford’s Law has been applied to detect voting fraud (Breunig and Goerres, 2011), the manipulation of international trade statistics (Cerioli et al., 2019), and the trustworthiness of COVID-19 case reports (Silva and Figueiredo, 2021). More recently, Benford’s Law has been applied to detect data manipulation and anomalies in environmental datasets including air quality information (Stoerk, 2016), the integrity of emission reduction data (Cole et al., 2020), and the reliability of fish-catch reporting (Noletto et al., 2022).

While the majority of environmental studies have focused on the frequency distribution of the first-digit, Benford’s Law involves a more general framework based on the joint distribution of the whole set of significant digits of each number (Barabesi et al., 2021). Integrating the analysis of the first-, second- and the leading double-digit in a set of numbers (e.g., in the number \$6543.21, they are 6, 5 and 65,

respectively) has been shown to be an effective means to reveal potential data irregularities (Nigrini, 2012). As a result, Benford’s Law can identify varying levels of imprecision in cost data (Fig. 1).

Nevertheless, not all datasets conform to Benford’s Law and it is generally applicable only where there is a large sample of individual estimates ($N > 1000$) that range across more than three orders of magnitude and exhibit positive skew such that the mean is greater than the median (Melita and Miraglia, 2021; Tam Cho and Gaines, 2007). Furthermore, Benford’s Law does not apply to sets of sequential numbers, assigned numbers (such as postcodes, car license plates or telephone numbers) or numerical series that are restricted to certain maximum or minimum bounds (Clippe and Ausloos, 2012; Pinilla et al., 2018). Although shaped by human activity, both actual and forecast monetary values (where not deliberately manipulated or subject to excessive rounding) have generally been expected to follow Benford’s Law (Cunha and Bugarin, 2015; Günnel and Tödter, 2009). Understanding the nature of data that can be expected to follow Benford’s law is clearly of critical importance.

However, since Benford’s Law is an empirical observation, any significant departures from expectations should be interpreted carefully and cannot provide irrefutable evidence of intentional data manipulation (Balashov et al., 2021). Departures from Benford’s Law can occur for legitimate reasons or inadvertently without any need to imply deceitful data manipulation (Nigrini, 2017). Therefore, tests of Benford’s Law should be viewed as an initial filtering step in data analysis providing an indication of whether more detailed investigation is warranted (Druică et al., 2018).

Although Benford’s Law has been applied to financial data in order to audit irregularities in fiscal accounts (Moul and Nye, 2015; Nigrini, 2012) there has been no attempt thus far to audit estimates of the monetary costs of environmental degradation (Arendt et al., 2022; Feng

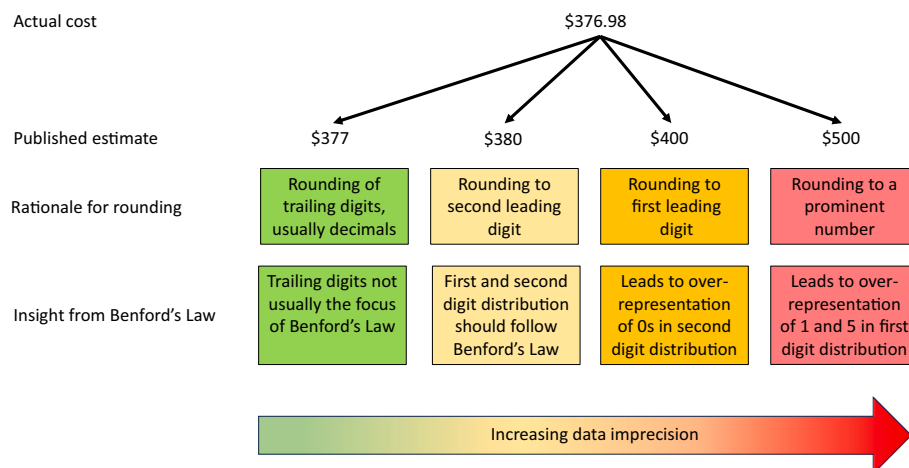


Fig. 1. Schematic illustration of the different forms of numerical rounding, the appropriate test using Benford’s law and the interpretation of such tests.

et al., 2021). One of the primary threats to economic and environmental wellbeing globally is the impact of invasive alien species on human health, food security, and ecosystem function (Pysek et al., 2020). The increasing application of global data to address the monetary costs of invasive alien species has provided some of the most striking evidence as to why biological invasions need to be prevented or rapidly contained (Cuthbert et al., 2021; Diagne et al., 2021; Haubrock et al., 2022; Kouba et al., 2022; Renault et al., 2022; Wang et al., 2023). Nevertheless, the diversity of sources from which costs of invasive species have been drawn, the types of costs that have been estimated, and the spatial scales at which they have been assessed raise important questions as to the precision of reported monetary cost estimates pertaining to biological invasions.

For example, should estimates of monetary costs derived from peer reviewed publications be assumed to be more reliable than those obtained from grey literature (Augusto et al., 2010)? Similarly, are costs arising from the management of invasive species more verifiable than less easily quantified measures of ecosystem damage (Emerton and Howard, 2008; Vaissiere et al., 2022)? Scaled-costs that result from mathematical combinations of numbers, such as when costs estimated at a single location are extrapolated to a much larger area, often conform to Benford’s Law (Tam Cho and Gaines, 2007). Thus, would data that have been scaled-up from observations at individual sites to provide country-level cost estimates or that are projected from a subset of data be expected to show fewer deviations from null expectations? To find answers to these questions, the distributions of first-, second- and leading double-digits of costs reported as arising from the impact and management of biological invasions and collated in the extensive global *Invacost* database (Diagne et al., 2020) were examined with reference to the null expectations under Benford’s Law.

2. Materials and methods

2.1. Monetary cost estimates of invasive alien species

The most recently updated version of the *InvaCost* database (v 4.1) includes 13,553 entries that describe the estimated or reported monetary costs from invasive alien species retrieved from both peer-reviewed as well as other literature sources worldwide (Balzani et al., 2022; Jiang et al., 2022; Kourantidou et al., 2022; Turbelin et al., 2023; Wang et al., 2023). All studies to date using the *InvaCost 4.1* database have standardised these original cost records using appropriate currency conversion and inflation correction factors to enable global scale comparisons (Diagne et al., 2023; Jiang et al., 2022; Kourantidou et al., 2022; Renault et al., 2022; Wang et al., 2023). However, standardisation of economic data to a common currency benchmark often involves sufficient compression or transformation of the data to account for currency conversions or inflation that the base assumptions for conformity with Benford’s Law will likely not hold (Nye and Moul, 2007). For this

reason, only the original raw cost estimates in the local currency, even if accumulated over several years, were used in this analysis since it is these values that have been drawn from reports and publications and most likely to exhibit any underlying departure from Benford’s Law.

In addition to reporting the estimated monetary costs in local currency at the approximate time of expenditure, the *InvaCost 4.1* database includes multiple additional fields that facilitate the refining of analyses and in particular four fields are the subject of the present analysis (Leroy et al., 2022): a) source reliability (high or low); b) cost realisation (observed or potential), c) type of cost (management or damage), and d) spatial scale (continent, country, region, site etc.). The perceived reliability of a data entry was classed as either high or low with the former class being assigned only if the entry was associated with well documented, reproducible, and traceable methods. A distinction was also made as to whether data entries were costs that were realised as having been observed or were only based on predictions of projected expected costs (potential). Costs were classed as either arising from damage incurred because of the invasive alien species (e.g., yield losses of crops, human health costs) or instead reflected the management expenditure arising from dealing with the species (e.g., labour and pesticide used). Although costs have been reported at multiple spatial scales, most estimates have been drawn from studies at either a country or a site scale, with the latter being defined as an area at intra-country level, including USA states, thus comparisons were made at only these two spatial scales.

2.2. Null model expectations of digit frequencies under Benford’s Law

If a set of numbers follows Benford’s Law, then the expected frequency of digit d occurring as the first or second digit is given by (Miller, 2015):

$$\text{First digit probability } P(d) = \log_{10} \left(\frac{1+d}{d} \right) \text{ for } d \in \{1, \dots, 9\} \quad (1)$$

$$\text{Second digit probability } P(d) = \sum_{k=1}^9 \log_{10} \left(1 + \frac{1}{10k+d} \right) \text{ for } d \in \{0, \dots, 9\} \quad (2)$$

The first-digit test is the most widely used approach for detecting non-conformity in monetary data and is useful for picking-up obvious anomalies (Moul and Nye, 2015). However, the second-digit test is recommended as a more robust indicator of data anomalies than the first-digit-test because the deviations between the observed and expected frequencies are more pronounced when there are systematic irregularities in the data (Breunig and Goerres, 2011; Diekmann, 2007; Todter, 2009). When examined together, an overrepresentation of ones (1 s) in the first-digit test combined with an over-representation of zeroes (0 s) in the second-digit test, with both digit tests revealing an underrepresentation of nines (9 s) is indicative of a tendency to round

Table 1

Sample size (N), numerical range and coefficient of skew for the overall *InvaCost 4.1* database and subsets examined in this study. Bayes factors (B_{10}) are presented for the analyses testing whether the first-, second- and leading double-digit distributions are consistent (H_0) or depart (H_1) from a null expectation based on Benford’s Law. Key to interpreting Bayes Factors: >100 extreme evidence for H_1 ; 30–100 very strong evidence for H_1 ; 10–30 strong evidence for H_1 ; 3–10 moderate evidence for H_1 ; 1–3 anecdotal evidence for H_1 ; 1 no evidence.

	N	Range	Skew	First-digit	Second-digit	Double-digit
				$B_{F_{10}}$	$B_{F_{10}}$	$B_{F_{10}}$
Overall database	13,553	1.71×10^{13}	106.85	1.04×10^9	∞	∞
High reliability	12,410	1.79×10^{12}	52.85	1.79×10^6	1.17×10^{276}	∞
Low reliability	1143	1.71×10^{13}	31.94	1.22×10^{-5}	8.92×10^{71}	6.44×10^{55}
Observed costs	11,009	1.71×10^{13}	98.07	9.10×10^7	∞	∞
Potential costs	2544	1.79×10^{12}	38.27	2.50×10^{-6}	5.42×10^{20}	9.48×10^{-11}
Damage	3484	1.71×10^{13}	55.02	1.89×10^{-7}	2.82×10^{39}	0.793
Management	9392	1.79×10^{12}	94.72	3.65×10^5	9.07×10^{283}	∞
Country scale	2956	1.71×10^{13}	50.07	1.59×10^{-9}	1.04×10^{83}	2.29×10^{43}
Site scale	9027	2.00×10^{11}	34.40	2.68×10^{14}	2.26×10^{214}	∞

values upwards (Charles, 1988; Lacina et al., 2018). Rounding of numerical values to a smaller number of significant leading digits is common practice in both environmental science (e.g., 67,835 rounded to 70,000) as well as in financial audits (e.g., values expressed to the nearest \$1000) yet will alter expectations arising from Benford’s Law. Under such circumstances, comparison with the rounded Benford distribution for both the first and second leading digit is often more appropriate (Günnel and Tödter, 2009; Slepkov et al., 2015).

Benford’s Law can be extended to two digits with the expected frequency of digit d_1d_2 occurring as the leading first two-digits is given by (Miller, 2015):

$$\text{First two-digit probability } P(d_1d_2) = \log_{10} \left(1 + \frac{1}{d_1d_2} \right) \text{ for } d_1d_2 \in \{10, \dots, 99\} \tag{3}$$

In contrast to the first- and second-digit tests, the double-digit test is primarily used to provide greater granularity in the analysis by pinpointing abnormal duplications of double digits and possible biases in the data. It is based on the observation that pairs of digits in many datasets are not uniformly distributed and follow a predictable pattern. Analyses of the first-, second-, and leading double-digits were undertaken for the entire dataset as well as separately for each of the paired comparison examining method reliability, cost realisation, type of cost, and spatial scale. Additional analyses of the first- and second-digit distribution against the rounded Benford distribution were undertaken to gauge the extent to which rounding might explain deviations from Benford’s Law.

2.3. Statistical analysis of goodness of fit to Benford’s Law

Both the first and second non-zero digit left of the decimal point were extracted from each raw cost estimate in the local currency presented in the *InvaCost 4.1* database (accessed on 11th November 2022). The observed frequency of first-, second- and leading double-digits in the raw economic cost data were compared with the expected null (H_0) frequency under Benford’s Law. Statistical comparisons between the observed and expected leading digit distributions have most often used Pearson χ^2 goodness of fit tests, however, the interpretation of χ^2 is markedly influenced by sample size such that for large samples size ($N > 500$) any deviation from the expected, no matter how small, can lead to Type I errors (Cleary and Thibodeau, 2005; Nye and Moul, 2007; Tam Cho and Gaines, 2007). Where sample sizes are large and risk of Type I errors is high, metrics that quantify how much more likely data occurred under the alternative (H_1) compared to null hypothesis (H_0) are more appropriate (Pericchi and Torres, 2011). The Bayes factor (BF_{10}) quantifies the relative predictive performance of two rival hypotheses, such that a $BF_{10} = 2$ means that the data are twice as likely under H_1 than the H_0 , values >10 are considered strong evidence in support of H_1 while a $BF_{10} > 100$ indicates extreme evidence (Stefan et al., 2019; van Doorn et al., 2021). Bayes factors have previously been shown to outperform standard Z-tests and χ^2 tests of Benford’s Law (Derks and Wagenmakers, 2021; Geyer and Williamson, 2004; Pericchi and Torres, 2011). Analyses therefore calculated BF_{10} under a default uniform prior as implemented in the Bayesian multinomial package in JASP v 0.17.1 (JASP Team, 2023).

3. Results

The raw cost estimates in the entire *InvaCost 4.1* database as well as in the different subsets satisfied the basic requirements for testing using Benford’s Law, being large enough sample sizes of positively skewed cost data ranging across at least twelve orders of magnitude (Table 1).

3.1. First-digit distribution

There was strong evidence that the first-digit distribution of the overall costs in the *InvaCost 4.1* database did not conform to Benford’s Law, with an overrepresentation of the digits 1 and 5 and an underrepresentation of the digits 8 and 9 (Fig. 2a, Table 1). This trend was also evident among the different subsets of the dataset but there was stronger evidence for a departure from Benford’s Law for data that were deemed more reliable (Fig. 3a), estimated management rather than damage costs (Fig. 3b), were observed instead of projected (Fig. 3c), and that were assessed at a site as opposed to a country spatial scale (Fig. 3d). First-digit distributions of estimates from sources deemed of low reliability, that estimated damage, or were assessed at a country scale appeared to

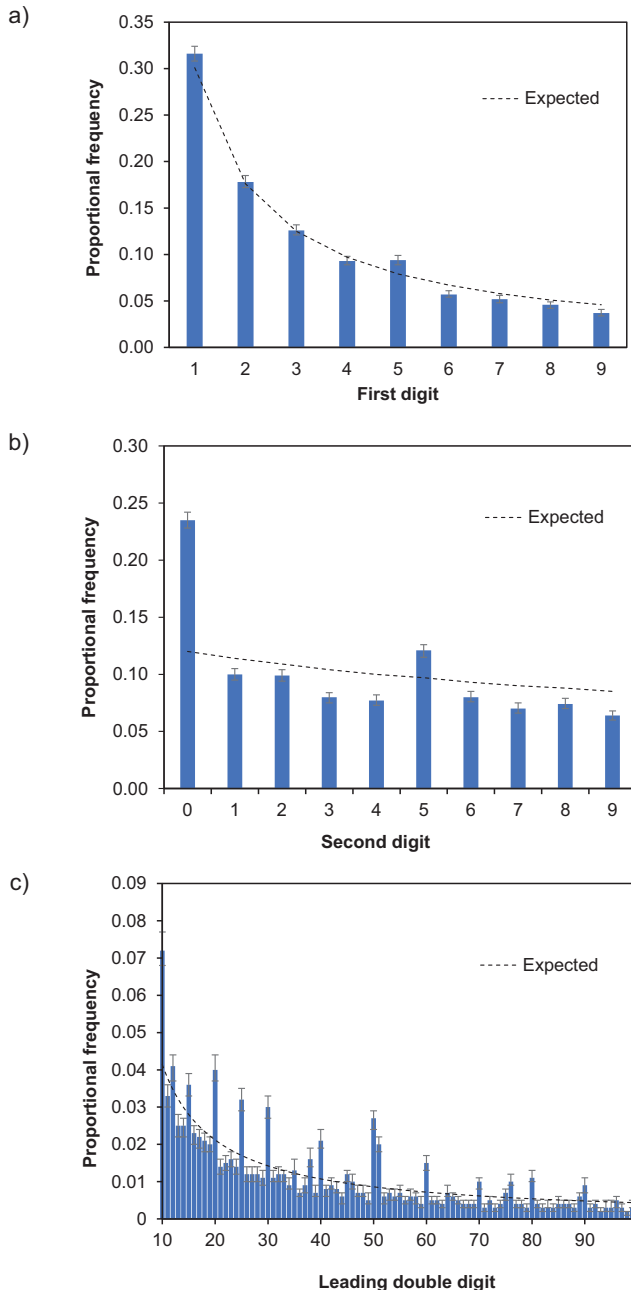


Fig. 2. The frequency of a) first-, b) second- and c) leading double-digits drawn from the economic costs of invasive alien species (columns) with the expected frequency derived from Benford’s Law (line). The 95 % confidence intervals are presented and are calculated based on the independent binomial distributions of each digit.

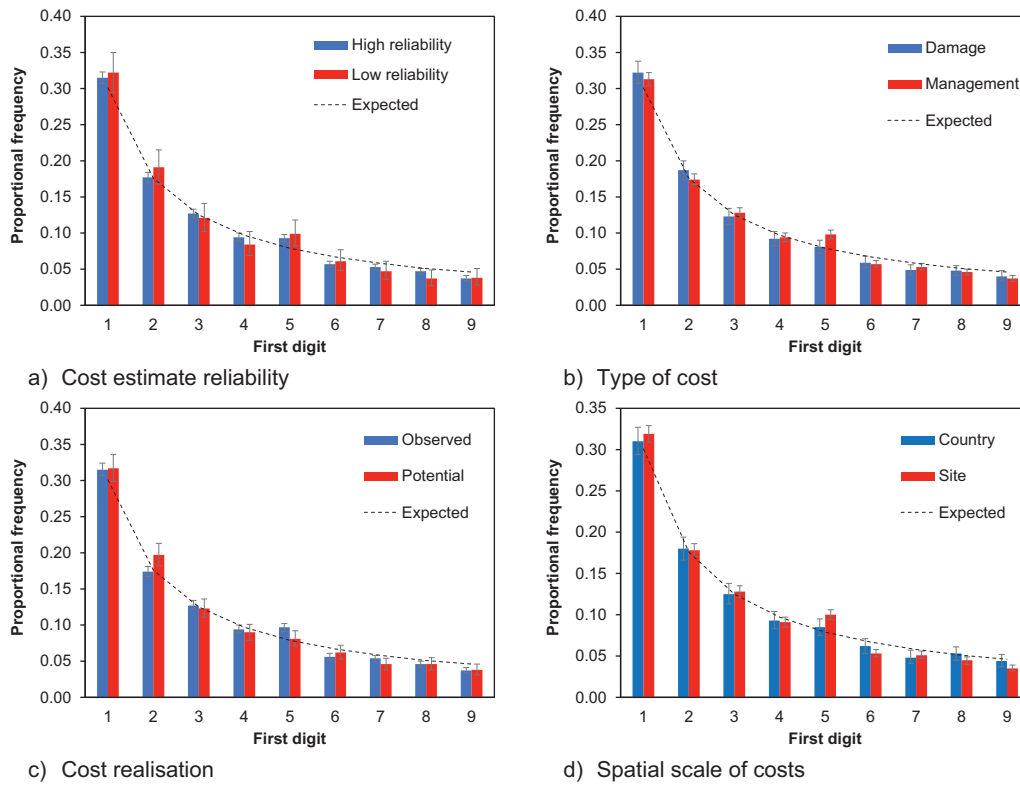


Fig. 3. Comparison of the frequency of first digits drawn from the economic costs of invasive alien species (columns) with the expected frequency derived from Benford's Law (line) in relation to whether the costs were: a) perceived as having high or low reliability, b) damage or management costs, c) derived from observed or projected data, and d) assessed at a country or site scale. The 95 % confidence intervals are presented and are calculated based on the independent binomial distributions of each digit.

fit Benford's Law. The BF_{10} was not correlated with sample size (Pearson $r = 0.14$, $df = 7$, $P = 0.719$) and most first-digit distributions exhibiting a $BF_{10} > 100$ providing extreme evidence for departure from Benford's Law while the four values with $BF_{10} < 0.01$ indicated extreme support for the null expectation.

3.2. Second-digit distribution

The evidence for a departure from Benford's Law was much stronger for second- than first-digit distributions, as indicated by considerably larger BF_{10} values (Fig. 2b, Table 1). In every comparison the BF_{10} highlighted extreme evidence for departures from Benford's Law (Table 1) but the evidence was stronger for data from sources that were classed as of low reliability (Fig. 4a), quantified management rather than damage costs (Fig. 4b) and were observed instead of projected (Fig. 4c) at site rather than country spatial scale (Fig. 4d).

3.3. Evidence for rounding

There was no evidence that either the first- or second-digit distribution was shaped by simple rounding to the nearest significant figure and in all cases the fit to the rounded Benford was worse than to the standard Benford distribution (Table 2). Evidence of an overrepresentation of ones (1 s) in the first-digit test combined with an overrepresentation of zeroes (0 s) in the second-digit test, with both digit tests revealing an underrepresentation of nines (9 s) was used to test for evidence that values may have been rounded upwards. Significant departures from either the standard or rounded Benford expected distribution were tested using the confidence intervals derived from an independent binomial distribution of each digit (0, 1 or 9). Results were dependent on whether the rounded or standard Benford distribution was chosen but the most parsimonious interpretation was that there was

support for rounding up of values for the entire dataset, but the evidence was stronger for data from sources that were classed as of high reliability, quantified management rather than damage costs, and were observed instead of projected at site rather than country spatial scale (Table 2).

3.4. Double-digit distribution

Given the evidence for both the first- and second-digit distribution analyses, the finding that the leading double-digit distribution also failed to conform to Benford's Law was not surprising (Table 1). Across all data entries, marked departures from expected frequencies were found for the following double digits: 10, 20, 25, 30 and 50 all of which were strongly overrepresented. The complimentary underrepresentation of the double digits 49 and 99 is consistent with cost estimates being regularly rounded upwards (Fig. 2c). However, other double digits that could not be associated with rounding were also overrepresented including: 12, 13, 21 and 51. These trends were also apparent in the different subsets of data (Fig. 5, Table 1). In general, the double digits that tended to be overrepresented were similar for each pair among the four different comparisons, but the deviation from expected values appeared more marked for sources that were deemed reliable (Fig. 5a), estimated management rather than damage costs (Fig. 5b), were observed instead of projected (Fig. 5c), and quantified at a site as opposed to a country scale (Fig. 5d). The BF_{10} highlighted that there was support for costs that were projected or quantified damage to conform to Benford's Law but for all other comparisons, highlighted extreme evidence for departures from Benford's Law (Table 1).

3.5. Trends in overrepresented values

While there were 13,553 entries in the *InvaCost* 4.1 database, these

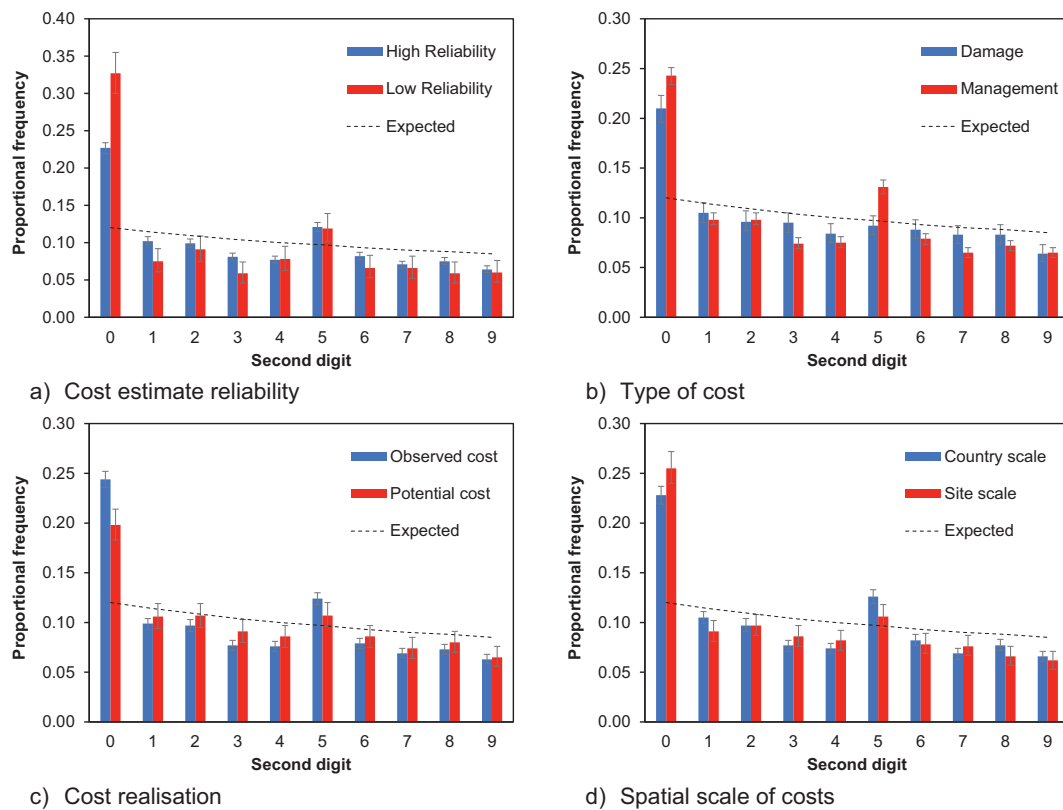


Fig. 4. Comparison of the frequency of second digits drawn from the economic costs of invasive alien species (columns) with the expected frequency derived from Benford's Law (line) in relation to whether the costs were: a) perceived as having high or low reliability, b) damage or management costs, c) derived from observed or projected data, and d) assessed at a country or site scale. The 95 % confidence intervals are presented and are calculated based on the independent binomial distributions of each digit.

Table 2

Bayes factors (B_{10}) are presented for the analyses testing whether the first- or second-digit distributions are consistent (H_0) or depart (H_1) from a null expectation based on the rounded Benford distribution. Key to interpreting Bayes Factors as in Table 1. Evidence of rounding up was tested against a null model that used the standard or rounded Benford distribution in relation to an overrepresentation of ones (1 s) in the first-digit test combined with an over-representation of zeroes (0 s) in the second-digit test, with both digit tests revealing an underrepresentation of nines (9 s).

	Rounded first-digit		Rounded second-digit		Evidence for rounding up	
	$B_{F_{10}}$		$B_{F_{10}}$		Standard Benford	Rounded Benford
Overall database	1.41×10^{230}		∞		Yes	Yes
High reliability	7.20×10^{207}		∞		Yes	Yes
Low reliability	4.09×10^{14}		1.83×10^{90}		No	No
Observed costs	1.96×10^{189}		∞		Yes	Yes
Potential costs	1.57×10^{34}		9.25×10^{36}		No	Yes
Damage	9.26×10^{53}		3.02×10^{64}		No	Yes
Management	7.29×10^{154}		∞		Yes	Yes
Country scale	1.99×10^{36}		2.54×10^{114}		No	No
Site scale	1.66×10^{167}		2.92×10^{292}		Yes	Yes

amounted to only 7,442 distinct values and this points to clustering of costs around certain numbers. A total of 901 different values accounted for 50 % of the 13,553 entries in the *InvaCost 4.1* database (Fig. 6a). The 20 most frequent values accounted for over 11 % of all records with the number 1,000,000 being the most common, accounting for 1.57 % of all values (Fig. 6b). Other multiples of one million were also frequent, including 2,000,000, 3,000,000, and 5,000,000, and there was also an overrepresentation of much smaller numbers, all of which were simple multiples of 25.6 (e.g., 512, 256, 128 etc., Fig. 6b). These values account for the spikes seen at values 12, and 51 in the double-digit distributions (Fig. 2c).

4. Discussion

Estimating the costs of invasive alien species is challenging and a wide variety of methods, each with its own limitations, has been used to derive such estimates (Emerton and Howard, 2008; Vaissiere et al., 2022). The researchers developing *InvaCost* have undertaken a major task to bring together these heterogeneous data sources and have done their best to screen all records in terms of objective measures of reliability (Diagne et al., 2020). Nevertheless, despite these quality assurance steps, there was strong evidence that the costs of biological invasions, as reported in the current *InvaCost 4.1* database, do not conform to Benford's Law indicating a problem with reporting and estimating costs. The magnitude of deviations from null expectations was dependent on the particular digit distribution examined

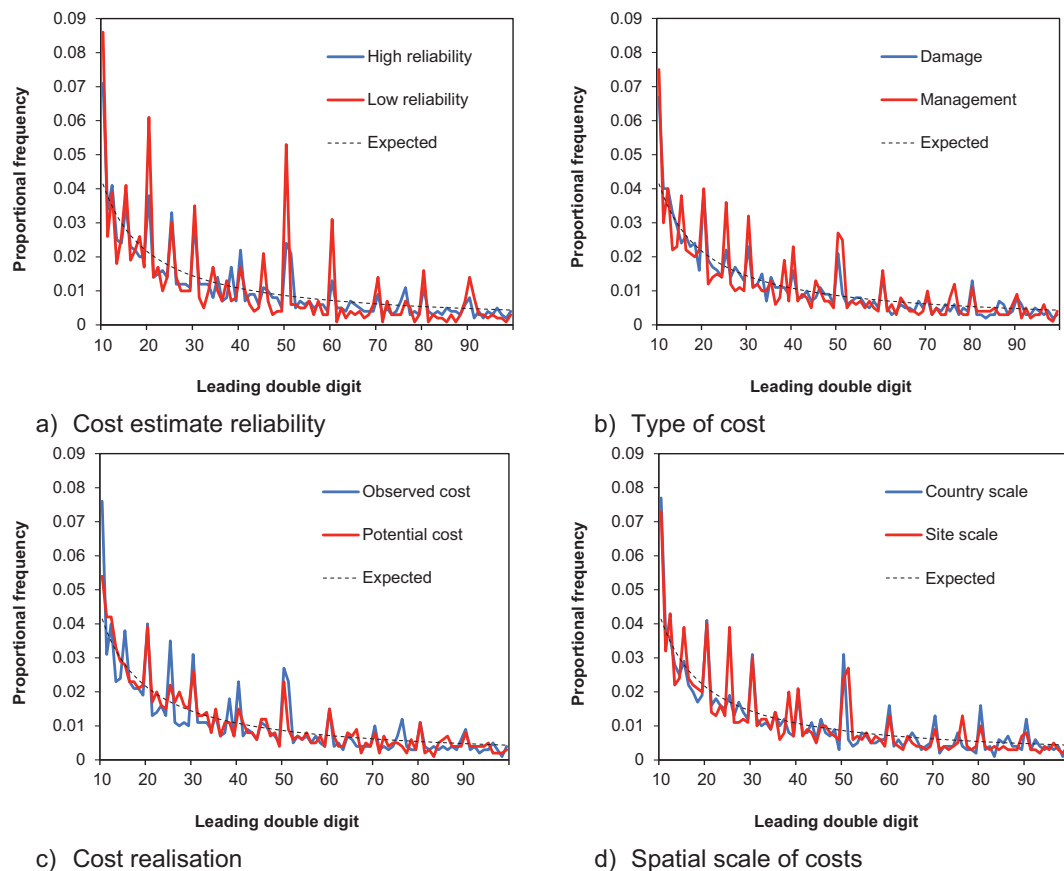


Fig. 5. Comparison of the frequency of the leading double digits drawn from the economic costs of invasive alien species (columns) with the expected frequency derived from Benford's Law (line) in relation to whether the costs were: a) perceived as having high or low reliability, b) damage or management costs, c) derived from observed or projected data, and d) assessed at a country or site scale.

emphasising the need for a broad approach simultaneously examining first-, second-, and leading double-digits. However, the marked departures from null expectation observed for the second-digit distribution across the entire dataset and all its subsets is of most concern since it is recognised that the second-digit test is the most robust indicator of systematic irregularities in reported values. The scale of these irregularities was greater for sources classed as having high reliability, that quantified management expenditure, were observed rather than projected, and were estimated at a site scale. Apart from the site-scale results, these findings are all contrary to the initial expectations. This suggests that the current screening criteria for data reliability currently employed in *InvaCost* may require revision and that the reports of cost values by the various authors (government agencies, scientists, or other actors) requires more detailed examination.

Benford's Law has previously been used to detect the possible deceitful manipulation of data in financial accounts (Rauch et al., 2011; Todter, 2009), electoral returns (Breunig and Goerres, 2011; Pericchi and Torres, 2011), and pollution records (de Marchi and Hamilton, 2006; Stoerk, 2016). However, it is important to remember that tests of Benford's law are only a first step in the identification of data irregularities after which more detailed investigation is required to attribute the cause and assess whether there was misreporting or deliberate data manipulation. Thus, the analyses presented here cannot be interpreted as evidence of any misconduct. In contrast to situations where fraud has been detected in the financial accounts of a single individual or organisation, the data collated in the *InvaCost 4.1* database stem from hundreds of different published sources and authors drawn across multiple years, regions, and taxa. A strong conspiracy mentality would therefore be required to interpret the departures from Benford's law found in the

InvaCost 4.1 database as a deliberate global collusion to manipulate the estimated monetary costs of biological invasions. Nevertheless, the marked departures, particularly of the second-digit distribution, point to irregularities that are worthy of further investigation.

It was evident that in many cases the reported monetary costs of biological invasions are presented as round numbers, defined as an integer which ends with one or more zeroes (0 s). The human tendency to use round numbers is commonly found in financial estimates and forecasts, often termed the round number bias (Lin and Pursiainen, 2021). Round numbers are generally used under conditions of high uncertainty (Hervé and Schwiendbacher, 2018) and this finding highlights that monetary cost estimates of biological invasions are often imprecise and that reporting data to many significant digits is rarely appropriate given the limited accuracy and repeatability of the methods (Emerton and Howard, 2008; Vaissiere et al., 2022).

Round numbers could be the result of rounding up or rounding down of uncertain estimates. The data suggest that estimates are frequently rounded up as shown by the overrepresentation of ones (1 s) in the first-digit test combined with an over-representation of zeroes (0 s) in the second-digit test, with both digit tests revealing an underrepresentation of nines (9 s). Humans often use the leading digit of a number to rapidly estimate its magnitude and as a result the rounding up of the leading first digit of reported monetary values is often undertaken to improve the perceptions of financial statements (Lin and Hwang, 2010; Van Cane-ghem, 2004). Nevertheless, whether due to using imprecise methods or the human tendency to use round numbers, the rounding upwards of the estimated costs of biological invasions appears to be widespread and a potentially significant issue. Although rounding upwards was evident in all comparisons, it appeared to be more of a problem for studies

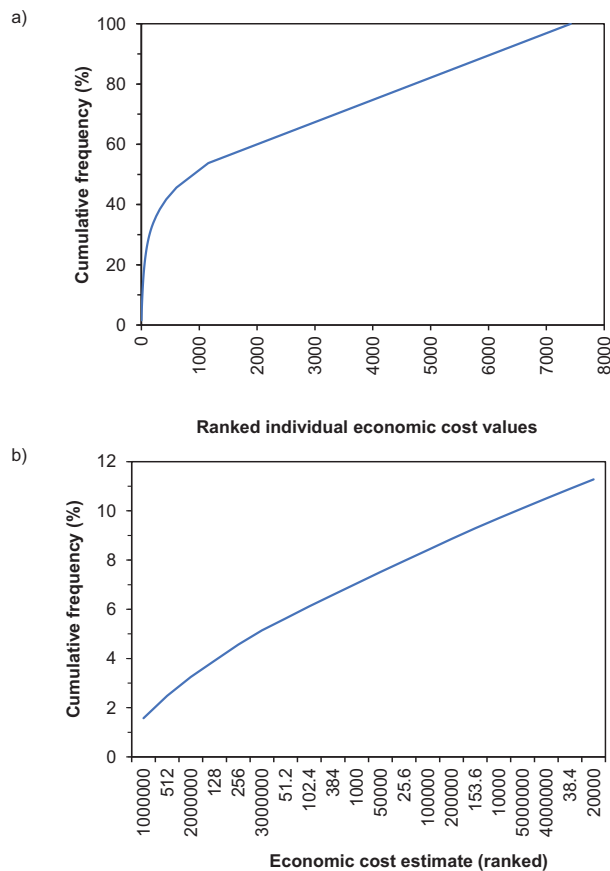


Fig. 6. Cumulative frequency of economic cost estimates when ranked by their prevalence in the *InvaCost 4.1* database for a) all data ranked from most frequent value to least and b) the top 20 most frequently occurring economic cost values.

identified as having high reliability and describing observed costs, those related to management or those obtained at a site scale.

One possible explanation for the high frequency of round numbers is that the expenditure on management is predetermined at a set value that is necessarily approximate and rounded up to only a few significant figures. Such values are often imprecise and may be shaped more by budgetary constraints of the commissioning organisation rather than by the actual scale of the problem to be managed. Under such circumstance, management expenditure is a poor estimate of the true cost of dealing with an invasive alien species. It is important to note that even when rounded up, any cost estimate subject to budgetary constraints could still be an underestimate of the true management cost of an invasive alien species. From an analytical perspective, using round data in place of the true values often results in biased estimates of regression coefficients unless an adjustment, such as Sheppard's correction, is applied (Schneeweiss and Komlos, 2009). Therefore, making projections of future costs based on round data should only be attempted with considerable caution since uncertainty can propagate in estimates drawn from extrapolations (Henry et al., 2023).

In addition to evidence of round numbers, the *InvaCost 4.1* database shows considerable support for numerical heaping, where values (often round numbers) cluster around certain values (Roberts and Brewer, 2001). For example, the value of 1,000,000 appeared repeatedly irrespective of whether the currency was United States, Singapore, New Zealand or Australian dollars, Japanese yen, European euros, or British pounds. Heaping appears to be marked for economic cost values that belong to the set of "Prominent Numbers" which includes the powers of ten (10; 100; 1,000 etc.), their doubles (20; 200; 2000 etc.) and their

halves (5; 50; 500 etc.) that can be formally represented as $n10^i$ where i is any integer and n is $\frac{1}{2}$, 1, or 2 (Converse and Dennis, 2018). Preferential use of prominent numbers would further contribute to the observed marked overrepresentation of 0 as a second digit. The frequency distributions of double digits highlighted multiple spikes associated with numerical heaping and while evident in all subsets of the *InvaCost 4.1* data, appeared more marked in sources classed as being of high reliability and associated with observed or management costs. Heaping is common in self-reported financial data, and where information retrieval and integration are particularly challenging, which is often the case for calculations of environmental costs (Gideon et al., 2017). Heaping indicates imprecision in data and is of concern because it distorts the distribution of the underlying data and may have substantial effects on model fit, leading to biased estimates and inaccurate predictions (Wang and Heitjan, 2008). Nevertheless, heaping around prominent numbers is a common cross-cultural behaviour related to the way humans express numerical values that are uncertain (Jansen and Pollmann, 2001) and while this undoubtedly limits the precision of the cost estimates of biological invasions it does not on its own indicate misconduct.

Benford's Law has been widely used in forensic accounting to identify specific records that depart markedly from expectations (Nigrini, 2012). A good illustration of this is the management cost dataset stemming from ten annual reports published by the regional government of Valencia, Spain between 2010 and 2019 describing the number of days (and fractions thereof) spent controlling different invasive alien species. These working days were converted to monetary values by multiplying them by an average constant daily rate of €128 across all reports (Angulo et al., 2021). However, the reported estimates of management expenditure do not conform to expectations under Benford's Law and out of the 123 records of €512 (equivalent to four workdays) in the *InvaCost 4.1* database, 48 (38.7 %) stem from a single annual report published in 2014 (Generalitat Valenciana, 2014). While it is easy to understand why the reported costs are multiples or fractions of the average constant daily rate, Benford's Law is scale invariant so the multiplication of worked hours by a constant hourly rate should not affect conformity with Benford's Law. In this case, part of the departure from Benford's Law is likely due to the hours worked being rounded to the nearest 0.1 days, which is logical for tracking staff time, but such a process will necessarily lead to values being imprecise. However, the heaping around certain values is less easy to explain given that effort would undoubtedly vary depending on the taxon targeted for control and both the year as well as location of any management activity. For example, why was the most frequent duration of management exactly four days, but rarely five or three days? A possible explanation is that heaping could indicate that effort is determined more by work-allocation models (e.g., having one day in the office a week) rather than the actual effort required to eradicate species. These types of factors known to influence imprecision, rounding, and heaping of management expenditure are unlikely to be unique to this dataset but it is in part the quantity and level of detail in these reports that have helped provide some insights into how management costs are sometimes estimated.

Another example is the frequency with which the value 1,000,000 was reported in the *InvaCost 4.1* database. Of the 212 reported cases of costs valued at exactly 1,000,000, the majority (146) were from Japan. These data all stem from government statistics published in official reports or websites by the Japanese Ministry of Environment and the Ministry of Agriculture, Forestry and Fisheries (JUDGIT!, 2019; Watari et al., 2021). These costs are associated with estimates of both damage and management at country as well as site scales and have been classed as reliable. However, the preponderance of values amounting to exactly ¥1,000,000 is not consistent with expectations under Benford's Law. While no claims of inappropriate behaviour are being made regarding the official government sources used in the *InvaCost 4.1* database, a deeper analysis on how such information is derived appears warranted.

The sources of management costs in *InvaCost* are highly

heterogeneous and include actual expenditures, estimates arising from research extrapolations, values that are estimated or planned budgets from government reports, and costs that are clearly coarse estimates (e.g., with cost values often being preceded by “up to” or “at least”). Understanding the degree to which different types of management costs depart from Benford’s Law expectations and which are the main sources of imprecision is necessary. For example, planned expenditures for invasive species management can be rounded or reworked to fit budget restrictions. In those cases, deviations from Benford’s Law can be expected.

The foregoing illustrates the value of applying Benford’s Law to environmental cost data. The analytical approach, while simple, provides an efficient means to interrogate large datasets to identify potentially problematic data sources or statistical challenges (such as use of round numbers and heaping). So how much confidence should be placed in the current assessments of the costs of biological invasions? It is important to restate that there is no suggestion that the researchers developing *InvaCost* have deliberately manipulated cost data, since the analyses are all based on previously published sources. Furthermore, publications arising from *InvaCost* have implemented quality assurance measures and have generally only used the most reliable data sources (Cuthbert et al., 2022; Fantle-Lepczyk et al., 2022; Renault et al., 2022). However, departures from Benford’s Law were more marked for sources classed as having high rather than low reliability. The Benford analyses present strong support for systematic use of round numbers, as well as heaping and this points to greater imprecision in estimates than previously recognised and likely result in somewhat inflated estimates of costs. Even with the *InvaCost* quality assurance procedures, it is evident that sources of monetary costs currently classed as reliable might not meet this criterion. *InvaCost* reflects the state of the literature on the monetary costs of biological invasions and is continually working towards improvement by evaluating and standardising the reporting and comparison of cost values (Leroy et al., 2022). Benford’s Law could become a new component of data quality assurance which would complement other ongoing efforts to provide assessments of the quality of cost estimates. Nevertheless, the key message that the costs of biological invasions arising from damage and management are considerable yet are widely underreported and thus underestimated likely remains true (Diagne et al., 2021). Thus, even with evidence of irregularities, a lack of precision in estimates reported in the literature, and possible inflation of some of the cost estimates, the scale of biological invasions and their impacts on the economy and ecosystems cannot be understated.

The Benford’s Law analyses point to several recommendations regarding how the monetary costs of biological invasions should be reported. First, there needs to be more transparent and thorough documentation of how cost estimates are derived, including the models and assumptions used, the types of costs included, the spatial scales at which costs are assessed, and where rounding or other adjustments have been made. This would allow others to evaluate and compare estimates and replicate or refine the analyses more critically as well as assess reliability more objectively. Second, it is essential that authors provide estimates of uncertainty around their cost values that could take the form of confidence intervals, sensitivity analyses, or scenario analyses that consider different assumptions in calculations. This would help data users understand the range of potential costs and the factors that contribute to uncertainty. Third, where possible, audited financial accounts of actual expenditures on invasive alien species management should be used to help increase the precision of cost estimates but also give greater confidence in the derivation of these data themselves. This seems particularly important for non-peer-reviewed official government reports. These recommendations likely hold more widely since here is no reason to believe that other large compilations of ecological data on monetary costs, such as those quantifying ecosystem services, would be any less likely to show similar departures from Benford’s Law. Forensic accounting, once the domain of financial auditors, would appear to be equally relevant to assessing the reliability of environmental data.

5. Conclusions

An integrated approach testing first-, second- and leading double-digit distributions against null expectations from Benford’s Law provides an objective and repeatable approach to assessing irregularities in cost estimates of environmental degradation. Using estimates of global costs of biological invasions as an example, it is evident that there can be quite large departures from null expectations. Detecting these irregularities using Benford’s Law should become a standard basis for thorough auditing and quality assurance of large environmental datasets. These results also point to the absence of agreed data standards for the estimation of costs of the management or damage arising from invasive alien species and this limits the traceability of estimates and their comparability. Benford’s Law can be used to pinpoint questionable sources that can then be examined in further detail and if necessary, rejected. The considerable efforts to date in collating reported estimates of costs of invasive alien species highlight the importance of obtaining accurate values to better inform policymakers and other stakeholders of the risk posed by biological invasions and the need to implement approaches that reduce these threats and mitigate their impacts.

CRediT authorship contribution statement

Philip E. Hulme: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Danish A. Ahmed:** Writing – review & editing. **Phillip J. Haubrock:** Writing – review & editing. **Brooks A. Kaiser:** Writing – review & editing. **Melina Kourantidou:** Writing – review & editing. **Boris Leroy:** Writing – review & editing. **Shana M. McDermott:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Underlying data are publicly available in an online repository (<https://doi.org/10.6084/m9.figshare.12668570>).

Acknowledgments

The authors wish to thank all researchers who contributed to the development of the *InvaCost* database and in particular Franck Courchamp for valuable comments on a previous draft of the manuscript and for coordinating a workshop in Marrakech where these ideas were first discussed.

References

- Angulo, E., Ballesteros-Mejia, L., Novoa, A., Duboscq-Carra, V.G., Diagne, C., Courchamp, F., 2021. Economic costs of invasive alien species in Spain. *NeoBiota* 67, 267–297.
- Arendt, R., Bach, V., Finkbeiner, M., 2022. The global environmental costs of mining and processing abiotic raw materials and their geographic distribution. *J. Clean. Prod.* 361, 132232.
- Augusto, L., Bakker, M.R., Morel, C., Meredieu, C., Trichet, P., Badaeu, V., et al., 2010. Is ‘grey literature’ a reliable source of data to characterize soils at the scale of a region? A case study in a maritime pine forest in southwestern France. *Eur. J. Soil Sci.* 61, 807–822.
- Balashov, V.S., Yan, Y., Zhu, X., 2021. Using the Newcomb–Benford law to study the association between a country’s COVID-19 reporting accuracy and its development. *Sci. Rep.* 11, 22914.
- Balzani, P., Cuthbert, R.N., Briski, E., Galil, B., Castellanos-Galindo, G.A., Kouba, A., et al., 2022. Knowledge needs in economic costs of invasive species facilitated by canalisation. *NeoBiota* 78, 207–223.
- Barabesi, L., Cerioli, A., Perrotta, D., 2021. Forum on Benford’s law and statistical methods for the detection of frauds. *Stat. Methods Appl.* 30, 767–778.

- Berger, A., Hill, T.P., 2021. The mathematics of Benford's law: a primer. *JISS* 30, 779–795.
- Breunig, C., Goerres, A., 2011. Searching for electoral irregularities in an established democracy: applying Benford's Law tests to Bundestag elections in Unified Germany. *Elect. Stud.* 30, 534–545.
- Cerlioli, A., Barabesi, L., Cerasa, A., Menegatti, M., Perrotta, D., 2019. Newcomb-Benford law and the detection of frauds in international trade. *Proc. Natl. Acad. Sci. U. S. A.* 116, 106–115.
- Charles, A.P.N.C., 1988. Anomalies in income numbers: evidence of goal oriented behavior. *Account. Rev.* 63, 321–327.
- Cleary, R., Thibodeau, J.C., 2005. Applying digital analysis using Benford's Law to detect fraud: the dangers of Type I errors. *Auditing J. Pract. Theory* 24, 77–81.
- Clippe, P., Ausloos, M., 2012. Benford's law and Theil transform of financial data. *Physica A: Statistical Mechanics and its Applications* 391, 6556–6567.
- Cole, M.A., Maddison, D.J., Zhang, L.Y., 2020. Testing the emission reduction claims of CDM projects using the Benford's Law. *Clim. Chang.* 160, 407–426.
- Converse, B.A., Dennis, P.J., 2018. The role of "Prominent Numbers" in open numerical judgment: strained decision makers choose from a limited set of accessible numbers. *Organ. Behav. Hum. Decis. Process.* 147, 94–107.
- Cuthbert, R.N., Pattison, Z., Taylor, N.G., Verbrugge, L., Diagne, C., Ahmed, D.A., et al., 2021. Global economic costs of aquatic invasive alien species. *Sci. Total Environ.* 775, 145238.
- Cunha, F., Bugarin, M.S., 2015. Benford's law for audit of public works: An analysis of overpricing in Maracanã soccer arena's renovation. *Econ. Bull.* 35, 1168–1176.
- Cuthbert, R.N., Diagne, C., Hudgins, E.J., Turbelin, A., Ahmed, D.A., Albert, C., et al., 2022. Biological invasion costs reveal insufficient proactive management worldwide. *Sci. Total Environ.* 819, 153404.
- de Marchi, S., Hamilton, J.T., 2006. Assessing the accuracy of self-reported data: an evaluation of the toxics release inventory. *J. Risk Uncertain.* 32, 57–76.
- Derks, K., Wagenmakers, E.-J., 2021. The Bayesian approach to audit evidence: quantifying statistical evidence using the Bayes Factor. Preprint PsyArXiv. <https://doi.org/10.31234/osf.io/kzqp5> (last accessed March 20, 2023).
- Diagne, C., Leroy, B., Gozlan, R.E., Vaissiere, A.C., Assailly, C., Nuninger, L., et al., 2020. InvaCost, a public database of the economic costs of biological invasions worldwide. *Sci. Data* 7.
- Diagne, C., Leroy, B., Vaissiere, A.C., Gozlan, R.E., Roiz, D., Jaric, I., et al., 2021. High and rising economic costs of biological invasions worldwide. *Nature* 592, 571–576.
- Diagne, C., Ballesteros-Mejia, L., Cuthbert, R.N., Bodey, T.W., Fantle-Lepczyk, J., Angulo, E., et al., 2023. Economic costs of invasive rodents worldwide: the tip of the iceberg. *PeerJ* 11, e14935.
- Diekmann, A., 2007. Not the first digit! Using Benford's law to detect fraudulent scientific data. *J. Appl. Stat.* 34, 321–329.
- Docampo, S., Trigo, M.D., Aira, M.J., Cabezedo, B., Flores-Moya, A., 2009. Benford's law applied to aerobiological data and its potential as a quality control tool. *Aerobiologia* 25, 275–283.
- Druică, E., Oancea, B., Vălsan, C., 2018. Benford's law and the limits of digit analysis. *Int. J. Account. Inf. Syst.* 31, 75–82.
- Emerton, L., Howard, G., 2008. A Toolkit for the Economic Analysis of Invasive Species. Global Invasive Species Programme, Nairobi.
- Fantle-Lepczyk, J.E., Haubrock, P.J., Kramer, A.M., Cuthbert, R.N., Turbelin, A.J., Crystal-Ornelas, R., et al., 2022. Economic costs of biological invasions in the United States. *Sci. Total Environ.* 806, 151318.
- Feng, Y.S., Liu, G.Y., Zhang, L.X., Casazza, M., 2021. Review on pollution damage costs accounting. *Sci. Total Environ.* 783, 147074.
- Generalitat Valenciana, 2014. Actuaciones de Control de Especies Invasoras. Memoria Anual 2013. Generalitat Valenciana-Conselleria d'Infraestructures-Territori i Medi Ambient-Direcció General el Medio Natural, Valencia.
- Geyer, A., Marti, J., 2012. Applying Benford's law to volcanology. *Geology* 40, 327–330.
- Geyer, C.L., Williamson, P.P., 2004. Detecting fraud in data sets using Benford's Law. *Commun. Stat. Simul. Comput.* 33, 229–246.
- Gideon, M., Helppie-McFall, B., Hsu, J.W., 2017. Heaping at round numbers on financial questions: the role of satisficing. *Survey Res. Methods* 11, 189–214.
- Günnel, S., Tödter, K.-H., 2009. Does Benford's Law hold in economic research and forecasting? *Empirica* 36, 273–292.
- Haubrock, P.J., Cuthbert, R.N., Hudgins, E.J., Crystal-Ornelas, R., Kourantidou, M., Moodley, D., et al., 2022. Geographic and taxonomic trends of rising biological invasion costs. *Sci. Total Environ.* 817, 152948.
- Henry, M., Leung, B., Cuthbert, R.N., Bodey, T.W., Ahmed, D.A., Angulo, E., et al., 2023. Unveiling the hidden economic toll of biological invasions in the European Union. *Environ. Sci. Eur.* 35, 43.
- Hervé, F., Schwienbacher, A., 2018. Round-number bias in investment: evidence from equity crowdfunding. *Finance* 39, 71–105.
- Jansen, C.J.M., Pollmann, M.M.W., 2001. On round numbers: pragmatic aspects of numerical expressions. *J. Quant. Linguist.* 8, 187–201.
- JASP Team, 2023. JASP (version 0.17.1). <https://jasp-stats.org>.
- Jiang, X., Zheng, P., Soto, I., Haubrock, P.J., Chen, J., Ji, L., 2022. Global economic costs and knowledge gaps of invasive gastropods. *Ecol. Indic.* 145, 109614.
- Joannes-Boyau, R., Bodin, T., Scheffers, A., Sambridge, M., May, S.M., 2015. Using Benford's law to investigate Natural Hazard dataset homogeneity. *Sci. Rep.* 5, 12046.
- JUDGIT!, 2019. JUDGIT!. <https://judgit.net/about>.
- Kouba, A., Oficialdegui, F.J., Cuthbert, R.N., Kourantidou, M., South, J., Tricarico, E., et al., 2022. Identifying economic costs and knowledge gaps of invasive aquatic crustaceans. *Sci. Total Environ.* 813, 152325W.
- Kourantidou, M., Verbrugge, L.N., Haubrock, P.J., Cuthbert, R.N., Angulo, E., Ahonen, I., et al., 2022. The economic costs, management and regulation of biological invasions in the Nordic countries. *J. Environ. Manag.* 324, 116374.
- Lacina, M., Lee, B.B., Kim, D.W., 2018. Management of revenue and earnings in Korean firms influenced by cognitive reference points. *Rev. Pac. Basin Financ. Mark. Policies* 21, 18500121.
- Leroy, B., Kramer, A.M., Vaissiere, A.C., Kourantidou, M., Courchamp, F., Diagne, C., 2022. Analysing economic costs of invasive alien species with the invacost r package. *Methods Ecol. Evol.* 13, 1930–1937.
- Lin, J.W., Hwang, M.I., 2010. Audit quality, corporate governance, and earnings management: a meta-analysis. *Int. J. Audit.* 14, 57–77.
- Lin, T.-C., Pursiainen, V., 2021. The round number heuristic and entrepreneur crowdfunding performance. *Finance* 68, 101894.
- Melita, M.D., Miraglia, J.E., 2021. On the applicability of Benford law to exoplanetary and asteroid data. *New Astron.* 89, 101654.
- Miller, S.J., 2015. A quick introduction to Benford's Law. In: Miller, S.J. (Ed.), *Benford's Law: Theory and Applications*. Princeton University Press, Princeton, pp. 3–22.
- Moul, C.C., Nye, J.V.C., 2015. Testing for strategic manipulation of economic and financial data. In: Miller, S.J. (Ed.), *Benford's Law: Theory and Applications*. Princeton University Press, Princeton, pp. 257–264.
- Nigrini, M.J., 2012. *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection*. John Wiley & Sons, Hoboken, NJ.
- Nigrini, M.J., 2017. Audit sampling using Benford's Law: a review of the literature with some new perspectives. *J. Emerg. Technol. Account.* 14, 29–46.
- Nigrini, M.J., Miller, S.J., 2007. Benford's law applied to hydrology data - results and relevance to other geophysical data. *Math. Geol.* 39, 469–490.
- Noletto, E.M., Carvalho, A.R., Thome-Souza, M.J.F., Angelini, R., 2022. Reporting the accuracy of small-scale fishing data by simply applying Benford's law. *Front. Mar. Sci.* 9, 947503.
- Nye, J., Moul, C., 2007. The political economy of numbers: on the application of Benford's Law to international macroeconomic statistics. *B E J. Macroecon.* 7, 1–12.
- Ozkundakci, D., Pingram, M.A., 2019. Nature favours "one" as the leading digit in phytoplankton abundance data. *Limnologica* 78, 125707.
- Pericchi, L., Torres, D., 2011. Quick anomaly detection by the Newcomb—Benford Law, with applications to electoral processes data from the USA, Puerto Rico and Venezuela. *Stat. Sci.* 26, 502–516.
- Pinilla, J., López-Valcárcel, B.G., González-Martel, C., Peiro, S., 2018. Pinocchio testing in the forensic analysis of waiting lists: Using public waiting list data from Finland and Spain for testing Newcomb-Benford's law. *BMJ Open* 8, e022079.
- Pyssek, P., Hulme, P.E., Simberloff, D., Bacher, S., Blackburn, T.M., Carlton, J.T., et al., 2020. Scientists' warning on invasive alien species. *Biol. Rev.* 95, 1511–1534.
- Rauch, B., Gottsche, M., Brahler, G., Engel, S., 2011. Fact and fiction in EU-Governmental economic data. *Ger. Econ. Rev.* 12, 243–255.
- Renault, D., Angulo, E., Cuthbert, R.N., Haubrock, P.J., Capinha, C., Bang, A., et al., 2022. The magnitude, diversity, and distribution of the economic costs of invasive terrestrial invertebrates worldwide. *Sci. Total Environ.* 835, 155391.
- Roberts, J.M., Brewer, D.D., 2001. Measures and tests of heaping in discrete quantitative distributions. *J. Appl. Stat.* 28, 887–896.
- Sambridge, M., Tkalcic, H., Jackson, A., 2010. Benford's law in the natural sciences. *Geophys. Res. Lett.* 37, L22301.
- Schneeweiss, H., Komlos, J., 2009. Probabilistic rounding and Sheppard's correction. *Stat. Methodol.* 6, 577–593.
- Silva, L., Figueiredo, D., 2021. Using Benford's law to assess the quality of COVID-19 register data in Brazil. *J. Public Health* 43, 107–110.
- Slepkov, A.D., Ironside, K.B., DiBattista, D., 2015. Benford's law: textbook exercises and multiple-choice testbanks. *PLoS One* 10, e0117972.
- Stefan, A.M., Gronau, Q.F., Schönbrodt, F.D., Wagenmakers, E.J., 2019. A tutorial on Bayes Factor Design Analysis using an informed prior. *Behav. Res. Methods* 51, 1042–1058.
- Stoerk, T., 2016. Statistical corruption in Beijing's air quality data has likely ended in 2012. *Atmos. Environ.* 127, 365–371.
- Tam Cho, W.K., Gaines, B.J., 2007. Breaking the (Benford) law: statistical fraud detection in campaign finance. *Am. Stat.* 61, 218–223.
- Todter, K.H., 2009. Benford's Law as an indicator of fraud in economics. *Ger. Econ. Rev.* 10, 339–351.
- Turbelin, A.J., Cuthbert, R.N., Essl, F., Haubrock, P.J., Ricciardi, A., Courchamp, F., 2023. Biological invasions are as costly as natural hazards. *Perspect. Ecol. Conserv.* 21, 143–150.
- Vaissiere, A.C., Courtois, P., Courchamp, F., Kourantidou, M., Diagne, C., Essl, F., et al., 2022. The nature of economic costs of biological invasions. *Biol. Invasions* 24, 2081–2101.
- Van Caneghem, T., 2004. The impact of audit quality on earnings rounding-up behaviour: some UK evidence. *Eur. Account. Rev.* 13, 771–786.
- van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., et al., 2021. The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychon. Bull. Rev.* 28, 813–826.
- Wang, H., Heitjan, D.F., 2008. Modeling heaping in self-reported cigarette counts. *Stat. Method.* 27, 3789–3804.
- Wang, S.Q., Deng, T., Zhang, J.Q., Li, Y.M., 2023. Global economic costs of mammal invasions. *Sci. Total Environ.* 857, 159479.
- Watari, Y., Komine, H., Angulo, E., Diagne, C., Ballesteros-Mejia, L., Courchamp, F., 2021. First synthesis of the economic costs of biological invasions in Japan. *NeoBiota* 67, 79–101.