Received: 31 July 2024 | Accepted: 6 January 2025

DOI: 10.1111/1365-2664.70010

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology

PERSPECTIVE

A protocol for error prevention and quality control in camera trap datasets

Eduardo A. Silva-Rodríguez[1,2] | Esteban I. Cortés[1,2] | Viviana Vasquez-Ibarra[1,2] | Nicolás Gálvez[3] | Jeremy Cusack[4] | Omar Ohrens[5] | Darío Moreira-Arce[6,7] | Ariel A. Farías[8] | José Infante-Varela[1,9]

[1]Laboratorio de Fauna Silvestre, Instituto de Conservación, Biodiversidad y Territorio, Facultad de Ciencias Forestales y Recursos Naturales, Universidad Austral de Chile, Valdivia, Chile; [2]Programa Austral Patagonia, Universidad Austral de Chile, Valdivia, Chile; [3]Pontificia Universidad Católica de Chile, Wildlife Ecology and Coexistence Lab and Centre for Local Development (CEDEL), Villarrica, Chile; [4]Okala, Stirling, UK; [5]Panthera, New York, New York, USA; [6]Departamento de Gestión Agraria, Universidad de Santiago de Chile (USACH), Santiago, Chile; [7]Institute of Ecology and Biodiversity (IEB), Santiago, Chile; [8]Departamento de Ecología & Gestión Ambiental, Centro Universitario Regional del Este (CURE-Maldonado), Universidad de la República, Maldonado, Uruguay and [9]Doctorado en Ecosistemas Forestales y Recursos Naturales, Escuela de Graduados, Facultad de Ciencias Forestales y Recursos Naturales, Universidad Austral de Chile, Valdivia, Chile

Correspondence
Eduardo A. Silva-Rodríguez
Email: eduardo.silva@uach.cl; eduardosilvar@gmail.com

José Infante-Varela
Email: jose.infantevarela@gmail.com

Funding information
Fondo Nacional de Desarrollo Científico y Tecnológico (ANID Fondecyt), Grant/Award Number: 11220713, 1221528 and 1231261; Agencia Nacional de Investigación y Desarrollo, Grant/Award Number: BECAS/DOCTORADO NACIONAL 21212206 and FB210006

Handling Editor: Jérémy Froidevaux

Abstract

1. Camera traps are a mainstream methodology in applied ecology, but surprisingly there are no widely accepted protocols to ensure the quality of the data obtained from these devices.
2. We reviewed a sample of 147 articles from the recent camera-trapping literature and found that only 4.8% report a measure of quality control.
3. We propose a framework to process media files obtained from camera traps that minimises errors by adopting a series of systematic procedures. Before classification, the focus is on detecting camera malfunctions, correcting storage and programming errors and establishing clear exclusion criteria. Classification can follow different approaches, including single or double human-eye review, which can be supported by artificial intelligence.
4. The protocol is followed by quality control procedures that enable users to determine whether a dataset meets quality standards and is ready to be analysed, or if further revision is needed.
5. *Synthesis and applications*: The proposed protocol introduces quality control as a key component of camera trap data processing, thus reducing error rates and making the reporting process more transparent. These principles also apply to other methods, such as autonomous sound-recording units. We suggest that by adopting formal quality control procedures, applied ecology will be able to capitalise the many advantages brought by new technologies and data processing tools.

© 2025 The Author(s). Journal of Applied Ecology © 2025 British Ecological Society.

KEYWORDS
artificial intelligence, camera trap, data standards, protocols, quality control, reproducibility, wildlife sensors

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010. By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

1 | INTRODUCTION

Estimating key population parameters for elusive vertebrates–such as nocturnal mammals–has historically been a major challenge. For some rare and elusive target species, approaches based on direct observation of animals or their signs (e.g. tracks and scats) pose additional logistical challenges (Thompson, 2004). In other cases, sympatric species (or their signs) share morphological characteristics that can make field identification difficult (e.g. Davison et al., 2002; Potter et al., 2019). Thus, the quality of resulting datasets heavily depends on the ability of field personnel to correctly identify target taxa (Davison et al., 2002). Indeed, in these cases, it is generally not possible to conduct quality control on the actual recording and identification processes (e.g. was the bird correctly identified?), implying that the basics behind the data are a matter of trust. The increasing use of wildlife sensors–such as camera traps–offer an opportunity to overcome this problem by allowing researchers and practitioners to collect a plethora of data through media files (e.g. photo, video or audio files) that can be revised multiple times (Caravaggi et al., 2017).

Over the last few decades, camera traps have emerged as a mainstream methodology for studying wildlife ecology (Steenweg et al., 2017). Their use is nearly non-invasive, allowing for the collection of large amounts of data—across multiple species simultaneously— and at relatively low cost. Importantly, the information obtained can be verified (Caravaggi et al., 2017). Their widespread use has led to a significant increase in our understanding of the ecology and conservation of wild vertebrates, including the discovery of new populations and species (Farias et al., 2014; Rovero et al., 2008), the evaluation of wildlife responses to human activity (e.g. Burton et al., 2024) and the implementation of large-scale monitoring programs at national (e.g. Chile; www.fotomonitoreo.cl, United States of America, Cove et al., 2021) and international scales (e.g. Rovero & Ahumada, 2017).

In this perspective, we first provide a brief overview of the multiple types of mistakes that can be made during the classification process of camera trap data. Then, based on a sample of recent journal articles, we present the status of classification and quality control procedures as reported by authors. Finally, based on the gaps detected, we present a general framework to both minimise errors and conduct quality control on datasets before using them in any analysis. We expect that this explicit framework to address quality control in camera trap datasets will contribute to improving the quality, credibility and reproducibility of data used in ecological science and practice.

2 | ERRORS IN CAMERA-TRAPPING DATASETS

Typical camera trap studies can yield tens of thousands of images, and in the case of extensive collaborations, these numbers can easily scale to hundreds of thousands of records (e.g. Cove et al., 2021). Handling such enormous amounts of media files and associated metadata is susceptible to different types of errors throughout the data generation process. Such errors will determine the quality of the resulting dataset and bear weight on the ecological inferences made (Johansson et al., 2020).

Camera trap data typically consist of spatial or deployment data (i.e. corresponding site, camera trap unit and coordinates for each record), temporal data (i.e. date and time when the record was produced) and the assigned classification (i.e. what species/individual/behaviour/etc. is assigned to each record; Bubnicki et al., 2023; Meek et al., 2014). The quality of camera trap datasets will depend on errors in each of these levels of information (Table S1). For example, errors in the spatial data can occur if the coordinate data is wrongly or inconsistently recorded (e.g. Scotson et al., 2017) or when media files from a given camera are assigned to another camera during the back-up process; errors in temporal data can occur when the date and time on a camera trap are set wrongly (Sanderson & Harris, 2013); and classification errors can happen when the assigned category (individual, species or behaviour depending on the focus of the study) does not correspond to what was actually recorded (e.g. Johansson et al., 2020; Potter et al., 2019; Zett et al., 2022).

Classification errors, in particular, can arise during manual (a person reviews the complete raw database and classifies each media file according to pre-established classes of interest), automatic (same as before but the review is performed completely by an artificial intelligence) or semi-automatic (a combination of manual and automatic classification) classification processes performed using any of the available classification softwares (e.g. Camera Base, Camelot, Timelapse, DigiKam, DataOrganize, Wildlife Insights, etc., Ahumada et al., 2020; Niedballa et al., 2016; Sanderson & Harris, 2013; Vélez et al., 2023; Young et al., 2018). Classification errors occur for several reasons, including confounding species of similar morphology (Gooliaff & Hodges, 2019), failing to visually detect an animal in a photo or video, forcing the classification of species of interest in doubtful records, incorrectly identifying species occurring in the study area (Zett et al., 2022), and inadvertently selecting wrong categories in classification software (Sundaresan et al., 2011; Table S1). Also, this can happen for different reasons such as low accuracy of automated classification

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology

software (Vélez et al., 2023), insufficient experience of classification personnel, fatigue of personnel during classification, low image quality, among others (Zett et al., 2022). Since different classification mistakes can occur and accumulate in the same dataset, the resulting influence on posterior analyses can be considerable. For instance, Johansson et al. (2020) found that a 12.5% misclassification of individuals in capture events led to a 35% overestimation in the abundance of snow leopards (*Panthera uncia*).

Errors will occur frequently in camera-trapping data if specific quality control measures are not implemented. Protocols aimed at reducing errors and enhancing transparency in camera trap data reporting have been developed for camera deployment (Wearn & Glover-Kapfer, 2017), reporting (Bubnicki et al., 2023; Forrester et al., 2016; Meek et al., 2014), individual animal recognition (Choo et al., 2020), and for automated and semi-automated classification (Böhner et al., 2023; Celis et al., 2022). However, although the risk of different types of errors–including false positives and false negatives–is well known and recognised (e.g. Meek et al., 2014), there is a lack of widely accepted protocols ensuring data quality controls in camera trap datasets.

3 | CURRENT QUALITY CONTROL PRACTICES

To determine the current prevalence of quality control reporting practices in the scientific literature we conducted a search of the recent camera trap literature in the Web of Science database (https://www.webofscience.com/wos/woscc/basic-search). We limited our search to articles published between 2021 and 2023 to ensure that our analyses were focused on up-to-date practices in the field and prevent biases toward practices that could already be outdated. The search was conducted on February 24th, 2024, using the following terms and years: 'camera trap*' OR 'trail camera*' (All Fields) AND 2021-01-01/2023-12-31 (Publication Date) AND Review Article OR Correction OR Editorial Material OR Meeting Abstract or News Item (Exclude—Document Types). The search yielded 2134 results. From the results we randomly selected a sample of 175 articles. This sample size was estimated using Epitools (Sergeant, ESG, 2018) and the following parameters: expected prevalence of quality control lower than 10%, desired confidence levels of 95% and precision of 5%. To the resulting sample size ($n=139$) we added 36 additional articles to account for the fact that we expected that up to 20% of the articles could be excluded.

Each selected article was reviewed independently by two researchers (EC and ES). First, articles were assessed to determine if they met inclusion criteria. We included all articles that reported camera trap surveys, except those that were based on published datasets (and therefore did not report methods) or on anecdotal records for which the actual survey details were not reported. A final set of 147 articles (84% of the initial 175) were further examined to determine if they met–at least partially–nine reporting criteria (see Table 1). Once completed, datasets obtained by each reviewer were compared and a consensus reached in case of discrepancies. We estimated the proportion of articles that reported–at least partially–each criterion and their associated Wilson's score confidence interval (Agresti, 2007) using Epitools (Sergeant, ESG, 2018).

The analysis of recent literature shows that image classification and quality control procedures are rarely reported in camera trap studies. For example, 48.3% of studies report the criteria used to classify records (at least partially) (e.g. species, individual, behaviour) and 28.6% state the use of specialised software to aid in classification or alternatively that the procedure was manual. Furthermore, only 10.2% of analysed studies reported the number of reviewers that classified records. Sampling effort (e.g. the number of cameras deployed) is reported in most articles (94.6%), but not the criteria used to exclude cameras from further analysis (only 23.1% did so) and procedures implemented to correct mistakes (0.7%). Lastly, quality control is very rarely reported (4.8%), if at all (Table 1).

4 | A FRAMEWORK FOR QUALITY CONTROL

Protocols are fundamental to secure the quality, credibility, traceability and reproducibility standard of scientific data (Munafò et al., 2017; Treves, 2022). However, as our review shows, it is not common practice to report on quality control procedures carried out on the datasets resulting from camera trap sampling. The fact that error rates are not reported leads to the implicit assumption that camera trap databases are error free. This assumption is, however, unlikely to be met in the vast majority of cases (e.g. Choo et al., 2020; Zett et al., 2022). To address this problem, we detail below a three-stage framework for quality assurance in the classification of camera trap datasets (Figure 1). The aim of this framework is to ensure that datasets meet all standards mentioned before data analysis. An example of the application of the proposed quality control is provided in the Supporting Information. The camera-trapping study reported in the Supporting Information was approved by the Comité Institucional de Cuidado y Uso de Animales (CICUA) at Universidad Austral de Chile (approval number 458/2022).

4.1 | Checks for media files from the field

Before classifying images, it is important to check for three types of errors (Figure 1, Stage a):

Storage mistakes are those associated with the backing up of memory cards. In our experience, one of the most common storage mistakes is to upload data from a memory card to a folder labelled with the wrong camera name or sampling unit. For this reason, it is important that each memory card has a unique ID associated with the camera site, sampling unit and deployment period. This is of particular importance if the project encompasses many sites

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010 By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Proportion of articles meeting nine criteria defined to evaluate classification procedures reported in recent research articles that used camera traps. The proportions are expressed as percentages, including 95% Wilson's score confidence intervals.

Criteria	Description	Proportion that met criteria	
		%	95% CI
Exclusion criteria	Criteria for the partial or total exclusion of cameras from further analysis reported	23.1	17.0–30.6
Sampling effort	The article reports the sampling effort included in the analyses (number of cameras and/or trap days)	94.6	89.6–97.2
Programming mistakes	Procedures to correct programming mistakes are reported	0.7	0.1–3.8
Classification criteria	The article indicates the criteria by which photos were classified to the categories needed to meet its objective (species, individual, behaviour, etc.). Articles that partially met this criterion are counted as meeting the criteria for the purpose of this article	48.3	40.4–56.3
Software	The article reports the software used or, alternatively, indicates that the classification was manual	28.6	21.9–36.3
Number of Reviewers	The article indicates the number of reviewers that classified the record files	11.6	7.3–17.7
Experience of reviewers	The article provides information on the experience of the reviewer (e.g. trained reviewers, undergraduate or graduate students, volunteers, etc.)	7.5	4.2–12.9
Number of revisions	The article indicates how many revisions of the dataset were conducted	10.2	6.3–16.2
Quality control	The article shows metrics regarding the quality of the data set (e.g. error rate, precision, recall, etc.)	4.8	2.3–9.5

and repeated memory card swaps. We recommend that, whenever possible, cameras be programmed to stamp the code of the site or camera on the images or videos themselves.

Programming mistakes refer primarily to errors in the setting of the date and/or time. Time errors are common because cameras often employ the 24-h format, which can cause confusion in areas where the use of the 12-h clock is more common (3 pm instead of 15:00). Date errors are common both due to lack of attention when programming and also because the most frequent date format used by cameras (mm-dd-yyyy) is not customary in most of the world. To ensure that mistakes in camera programming can be corrected, pictures or videos of the setting and removal need to be taken by the camera trap, and the time of the camera set up and removal needs to be recorded independently. The first and last media files recorded, representing deployment and retrieval events, should be later reviewed to detect errors. This approach is also useful to detect storage mistakes. Corrective measures used to address programming errors should be reported, and if errors cannot be corrected then exclusion criteria should be stated.

Camera malfunction refers to all cases where cameras did not work as expected. These include–but are not limited to–cameras whose placement in the field did not allow them to obtain analysable records (e.g. the camera was pointing too high), cameras that had reduced activation periods (e.g. ran out of batteries), or whose batteries were removed (with the time reset to the manufacture date of the device), devices that were moved by animals or people, cameras that were not turned on, blocked or corrupt memory cards and flash or infrared failure. Camera malfunction checks require the actual media files to be compared to the field dataset, and the manual inspection of at least a portion of the dataset. As for programming errors, camera malfunction can lead to data exclusion, for which the criteria used need to be stated explicitly.

4.2 | Checks for database classification

Classification categories and criteria should be defined first (Figure 1, Stage b). For some taxa, classification will be conducted to species

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010. By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology

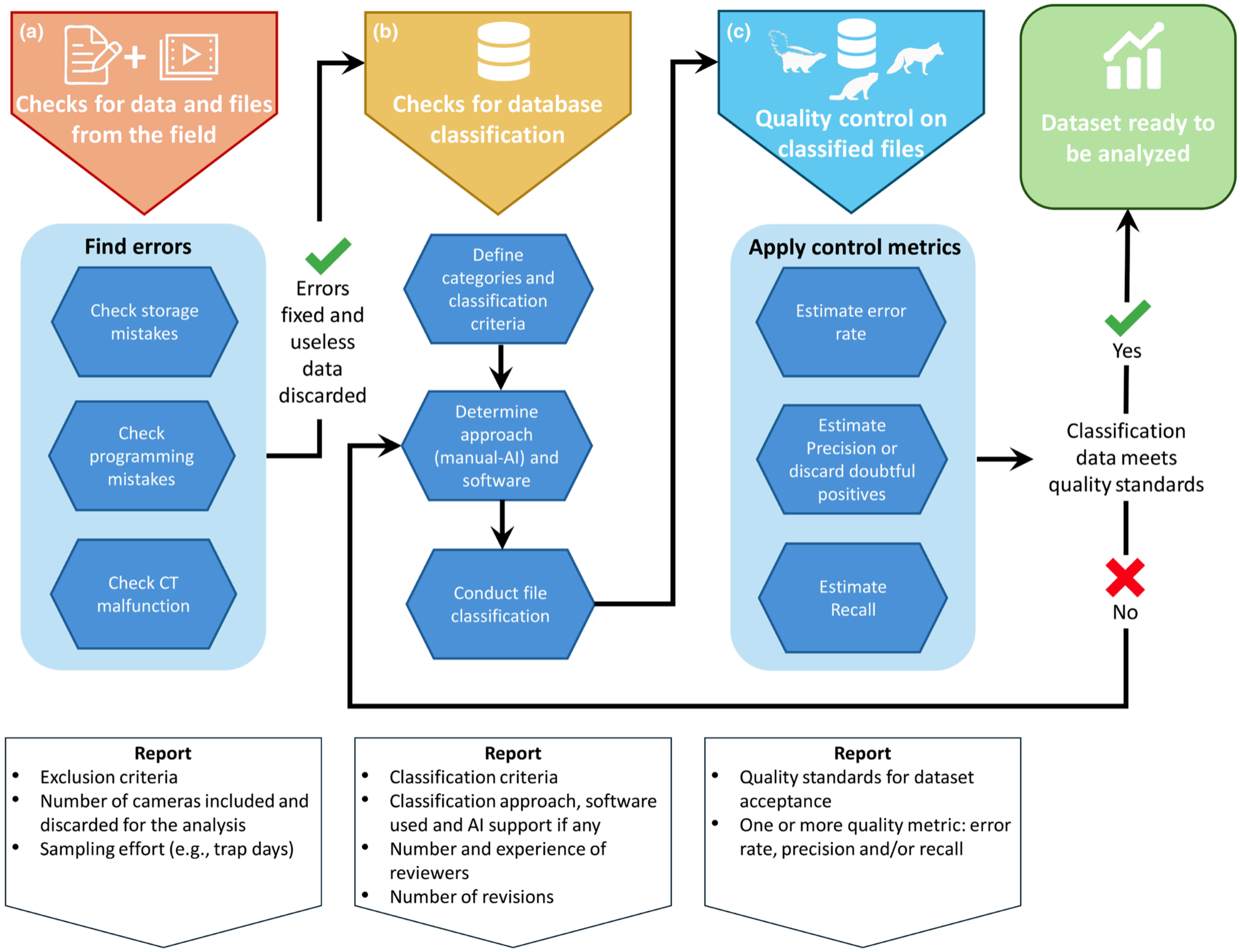

FIGURE 1 Framework for quality assurance and control in camera trap (CT) datasets and aspects that need to be reported for transparency of the dataset preparation process. The framework includes (a) checks for data and files from the field; (b) checks for database classification; and (c) quality control on the classified files. Precision is the proportion of records correctly assigned to a given category, whereas recall is the probability that a true record of a given species (or behaviour, individuals, etc.) is correctly classified as belonging to that species.

or genus level (see Kays et al., 2022). However, depending on the objective of the study, species can be grouped (e.g. birds, small mammals). The definition of categories requires clear classification criteria, as also recommended for the recognition of individual animals (Choo et al., 2020). In cases where species are similar and thus possibly indistinguishable due to poor image or video quality, or bad camera angles that do not permit unique characteristics to be discerned (e.g. Kays et al., 2022; Murray et al., 2023; Zett et al., 2022), there should be clear criteria by which individuals are assigned to a species (see Kays et al., 2022). Considering that some analyses are very sensitive to false positives (e.g. occupancy estimation, MacKenzie et al., 2017), we recommend not forcing species level identification for doubtful record files, and instead to adhere to the most precise classification possible (e.g. 'unknown canids', Murray et al., 2023). Such an approach also enables uncertain records to be filtered and reviewed at a later stage.

Classification criteria also need to be explicit if multiple individuals appearing in a single image or video will be counted. Two frequent approaches involve counting individuals in a given media file (i.e. a single image) and counting individuals in a sequence of media files (sensu Meek et al., 2014). From a biological point of view, it is often preferable to consider the entire sequence when counting individuals, so as to minimise double counting of gregarious animals that cannot be differentiated individually. In general terms, the criterion for any classification to be made needs to be explicitly reported as part of the protocol. This includes species and number of individuals, but also sex, age class and association to people (e.g. in the case of pets).

The classification approach (*manual, fully automated or semi-automated*) should be described in detail. In particular, the use of a given software (Vélez et al., 2023; Young et al., 2018) may imply different types of error. Ideally, the level of experience of reviewers

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010. By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

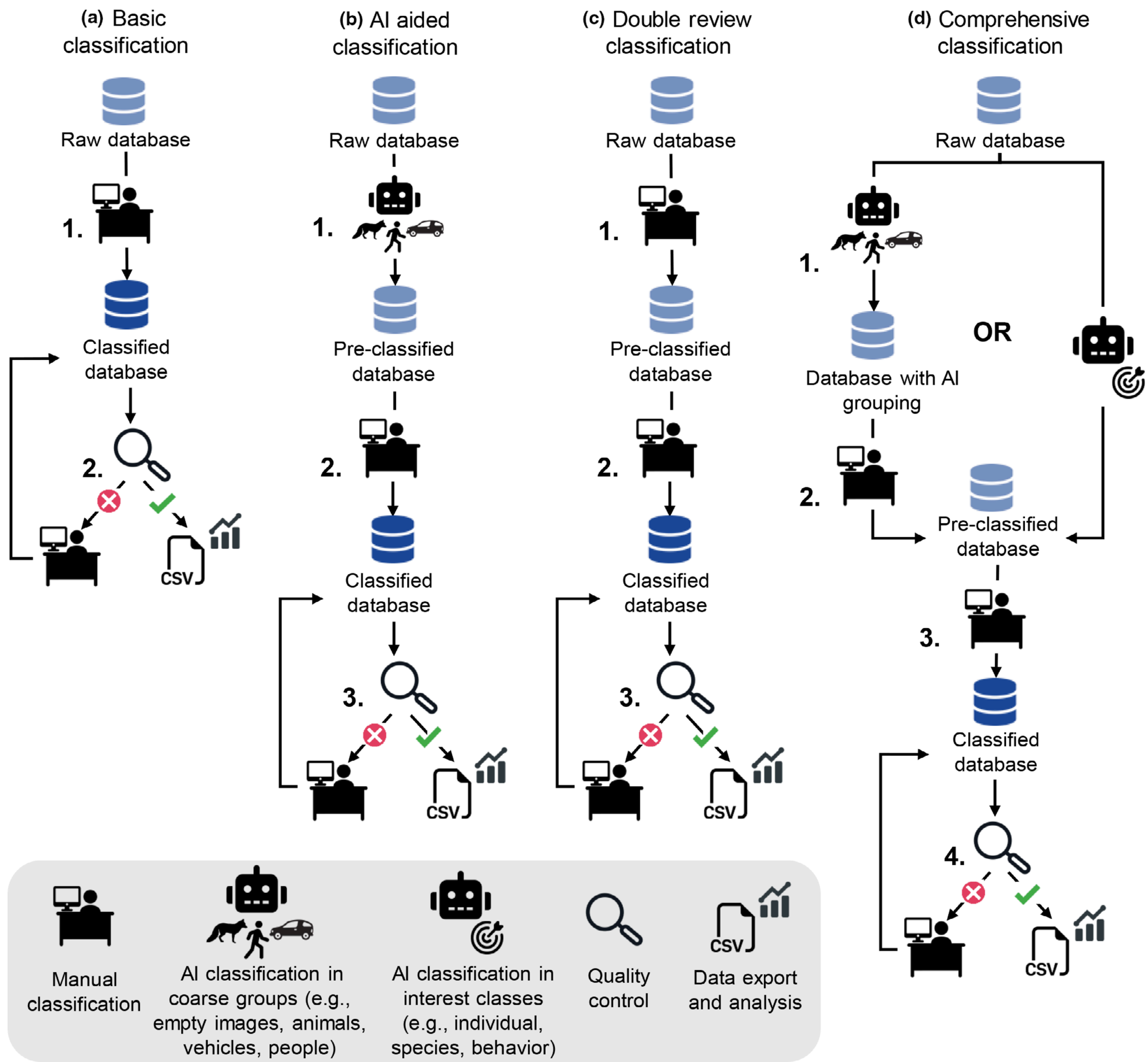

FIGURE 2 Four different classification protocols for camera trap datasets: (a) Basic classification; (b) artificial intelligence (AI) aided classification; (c) double review classification; and (d) comprehensive classification.

should be stated. In addition, automated or semi-automated image classification through artificial intelligence (AI), if used, needs to be declared. Furthermore, the metrics (e.g. accuracy, precision, recall, sensitivity, specificity) that underlie the performance of the classification model used need to be reported (Vélez et al., 2023) or, if using well established models (e.g. Megadetector, Beery et al., 2019), the classification model and model version.

Data classification can follow different protocols that differ in the number and order of manual and AI-assisted steps but share the fact that each of them culminates in a quality control procedure before data analysis (Figure 2; Section 4.3). The 'Basic classification' protocol relies on a single manual classification step followed by quality control (Figure 2a). Artificial intelligence support can be included in the first revision to help differentiate coarse groups (e.g. Beery et al., 2019, Figure 2b). The double review (Figure 2c) and comprehensive (Figure 2d) classification protocols include two human-eye revisions for all media files and–as for all protocols–a quality control procedure at the end. In the comprehensive protocol, one of the human-eye revisions could be replaced by automated classification as species level classification reaches higher levels of precision.

The inclusion of a second review in some of the protocols (Figure 2c,d) serves to detect and correct mistakes that could have occurred during the first classification. At this stage there are two options: to conduct an independent classification of all media files, or a revision of the first classification. If an independent evaluation is chosen, discrepancies can be found using functions available in

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010. By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

BRITISH ECOLOGICAL SOCIETY Journal of Applied Ecology

camtrapR (Niedballa et al., 2016). In general, the second option is more efficient, especially if the first round involves classifying the entire dataset by less experienced reviewers or volunteers (e.g. Arandjelovic et al., 2024), followed by a second round of revision and correction by more experienced personnel. Ideally, the two-stage classification should be performed by different people, although we acknowledge that this is not always feasible. During this stage, each record file will be validated or, alternatively, flagged as discrepant. Media files that are flagged should be reviewed to resolve discrepancies. If the discrepancies arise from classification errors (e.g. there was an animal in a file labelled as empty), these errors should be corrected. In cases where reviewers disagree on the assigned category, the available options are to label record files as unidentified, classify up to the lowest taxonomic level possible (e.g. unknown canid, Murray et al., 2023) or revise after discussion and agreement. Criteria for managing discrepancies relating to the identification of individual animals include exclusion from subsequent analysis, revision after discussion, modification of classification criteria after comparing results and analysing data separately (Choo et al., 2020). In all cases, the procedure used to solve discrepancies should be reported.

Although we favour protocols that involve double-reviews as an option to minimise errors in the classification process, alternatives that involve less steps (Figure 2a,b) or that involve a double review only for species that are difficult to identify can also be chosen. These alternatives can be adopted if the personnel conducting the classification is experienced (but see Choo et al., 2020) or if automated image classification with very high sensitivity and specificity are available.

4.3 | Quality control on classified files

Independent of the classification protocol chosen, once media files have been classified and reviewed, the database should undergo a quality control to determine if the error rate is low enough to accept the resulting dataset, or if a new revision of the database is needed. We suggest that the quality control should be conducted by the person(s) that holds responsibility for the database. At this stage, the purpose of the revision is not to rectify misclassified records, but to estimate the database's error rate (or other metrics, see below) as a criterion to determine if the database is acceptable *as is*. For this purpose, a random sample of the dataset must be drawn and reviewed to determine the error rate, which is the proportion of total record files in the sample that are incorrectly classified.

We suggest that a sample size of around 5000 record files should be large enough to determine precisely the error rate of most datasets (see details in the Supporting Information). Confidence intervals should be included in the error rate estimation. Given the behaviour of proportion data, we suggest using the Wilson's score confidence interval rather than the normal approximation (Agresti, 2007). Confidence interval estimators are available online (e.g. https://epitools.ausvet.com.au/ciproportion, Sergeant, ESG, 2018) even for camera trap users who are not familiar with more advanced statistical software. If error rates are higher than acceptable (e.g. >0.5% or any other predefined value), the camera trap dataset needs to go through a new classification process.

The error rate of a classified camera trap database is an important descriptor of the resulting dataset. However, the proportion of correctly classified records may be misleading, because the indicator will be influenced by the most frequent categories (e.g. empty record files). Therefore, we suggest that this indicator can be complemented or replaced by additional measures. We propose the use of two additional indicators either as a complement or alternative to error rate: recall (also known as sensitivity) and precision. Recall is the probability that a true record of a given species (or behaviour, individuals, etc.) is correctly classified as belonging to that species (Vélez et al., 2023). Precision is the proportion of records assigned to a given species (or focal behaviour, individuals, etc.) that are truly of that species (Vélez et al., 2023). Recall and precision can be estimated as a single-class indicator (i.e. for each species of interest in studies focused on a single or a few species) or as a macro-averaged multi-class indicator (i.e. as a descriptor of the whole dataset on community level studies; Table 2).

Recall can be estimated from the same sample used to estimate the error rate, regardless of whether it is to be used as a single-class or multi-class indicator. In the case of single-class precision, if the number of positive classifications is high, then it would be adequate to estimate it from a sample. However, if the analyses to be conducted are sensitive to false positives (e.g. occupancy models, MacKenzie et al., 2017), or when the focal species is rare or very difficult to detect (thus producing too few records), instead of estimating precision, we recommend reviewing all positive classifications and discarding doubtful cases. In the case of multi-class precision, the evaluation of a systematic sample of each of the classes (e.g. 200 media files per class), including empty and unidentified, would be more efficient. We note that macro-averaged precision and recall could be used as an alternative to error rate, especially in unbalanced datasets.

5 | FUTURE DIRECTIONS

The protocol described here not only aims to minimise errors at every stage of the camera trap data processing sequence, but also seeks to make error rates in resulting datasets more transparent. Although it provides a useful way for researchers and practitioners to standardise efforts, yet, we caution that it should not be adopted blindly. The reality of different projects, such as their objectives or the number of records, may require modifications to be made (see Choo et al., 2020 for individual organism identification). Furthermore, constraints on financial resources and/or staff may limit the adoption of some of the protocol steps proposed here, which is the reason why we propose different data processing and reviewing pipelines (Figure 2). Nevertheless, our aim is to promote a standardised way of minimising errors in the processing of camera trap datasets and to

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010. By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Measures available to conduct quality control on classified camera trap datasets (modified from Sokolova & Lapalme, 2009; Vélez et al., 2023).

Metric	Formula	Description	Advantages and disadvantages
Error rate	$\frac{fp+fn}{tp+fp+tn+fn}$	Proportion of misclassified media files across the whole dataset	Good to detect poorly classified datasets. Highly influenced by dominant classes (e.g. empty)
Precision	$\frac{tp}{tp+fp}$	Proportion of records assigned to a given species (or behaviour, individuals, etc.) that are truly of that species	Useful when the focus is on one or a few species. Can be labour intensive if many species are included
Recall (sensitivity)	$\frac{tp}{tp+fn}$	Proportion of all true records of a given species (or behaviour, individuals, etc.) that are correctly classified as belonging to that species	Useful when the focus is on one or a few species. Can be labour intensive if many species are included
Macro-averaged precision	$\frac{\sum_{i=1}^{c}\frac{tp_i}{tp_i+fp_i}}{c}$	Average of the precision across all classes (c), giving equal weight to each class	Useful when datasets are dominated by one or a few classes and the interest is in many different classes. Low classification quality in one or a few classes of interests may not be detected
Macro-averaged recall	$\frac{\sum_{i=1}^{c}\frac{tp_i}{tp_i+fn_i}}{c}$	Average of the recall across all classes (c), giving equal weight to each class	Useful when datasets are dominated by one or a few classes and the interest is in many different classes. Low classification quality in one or a few classes of interests may not be detected

Note: True positive (*tp*) is the number of times a species, individual or behaviour was correctly identified as present in a media file. False positive (*fp*) is the number of times a species, individual or behaviour was wrongly identified as present in a media file. True negative (*tn*) is the number of times a species, individual or behaviour was correctly identified as absent in a media file. False negative (*fn*) is the number of times a species, individual or behaviour was wrongly identified as absent in a media file (modified from Vélez et al., 2023).

increase transparency regarding the quality of said datasets for future camera trap studies. The key point is that procedures to ensure data quality need to be implemented and reported.

The present protocol also serves as a tool to distribute responsibilities within collaborative projects. Camera trap data are often processed by many people with different levels of experience (Steenweg et al., 2017; Zett et al., 2022). Our proposal provides a feasible solution by identifying specific points at which to conduct quality control. This can support decisions regarding who in a team might be better placed to carry out quality control at different stages, and thresholds above which a dataset is considered ready to be shared. This is particularly relevant for large-scale collaborations involving millions of records (e.g. Burton et al., 2024; Cove et al., 2021), for which it is rarely feasible for project leaders to check all data or media files. In these cases, we suggest quality control could involve multiple levels. First, by including quality control as part of the protocols for collaboration and requesting quality indicators for each contributed dataset (e.g. error rate, precision and recall). Second, by conducting a quality control on a sample of the whole dataset, using the protocol presented here. In any case, it is fundamental to acknowledge who conducted each stage of the review (coauthors, field assistants, volunteers, etc.) and who was responsible for the quality control. This latter role should be acknowledged in the author contribution section of published outputs.

Although our work focuses on the processing of camera trap data, the underlying principles are applicable to data collected using other technologies, such as audio files from autonomous recording units (e.g. Darras et al., 2019; Priyadarshani et al., 2018). Indeed, these often generate large volumes of data that are very demanding in terms of processing, making it more likely that errors will occur. The key message of the present work is that, unlike traditional wildlife survey methods based on human observers, recent technological methods give us the opportunity to check the quality of the resulting data and avoid biases in subsequent analyses (Caravaggi et al., 2017; Darras et al., 2019). We suggest that by adopting formal quality control procedures, it will be possible to capitalise on the many advantages brought about by these new technologies and data processing tools.

AUTHOR CONTRIBUTIONS

Eduardo A. Silva-Rodríguez conceived the idea. Eduardo A. Silva-Rodríguez, José Infante-Varela, Esteban I. Cortés and Viviana Vásquez-Ibarra developed the early versions of the protocol. Darío Moreira-Arce, Nicolás Gálvez, Jeremy Cusack, Ariel A. Farías and Omar Ohrens contributed to improve the protocol. Eduardo A. Silva-Rodríguez and José Infante-Varela drafted the first version of the manuscript, analyzed the data, and conducted quality control on the camera trap dataset. Eduardo A. Silva-Rodríguez, José Infante-Varela and Esteban I. Cortés, conducted the analysis of the recent literature. All authors contributed to review and improve the manuscript.

ACKNOWLEDGEMENTS

This study was funded by FONDECYT-ANID 1221528. D.M.-A. also thanks grant FONDECYT-ANID N°1231261 and Grant FON

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010. By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

DECYT-ANID PIA/BASAL FB210006. J.I.-V. acknowledges the support from ANID BECAS/DOCTORADO NACIONAL 21212206. J.C. acknowledges grant FONDECYT-ANID 11220713. The authors thank two anonymous reviewers for their insightful comments. We declare that during the preparation of our work, we used ChatGPT 3.5 and 4.0 (OpenAI, 2023, 2024) to detect potential errors in grammar, spelling and English style in our texts, as we are not native English speakers. Additionally, we used these tools to support code writing in R. We have reviewed and edited the content as needed and take full responsibility for the content of our manuscript. The camera-trapping study reported in the Supporting Information was approved by the Comité Institucional de Cuidado y Uso de Animales (CICUA) at Universidad Austral de Chile (approval number 458/2022). We thank Nadja Aranda and Fernanda Zenteno for their support in classifying camera trap data.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

Data available from the Figshare Repository https://doi.org/10.6084/m9.figshare.28330655 (Silva-Rodriguez et al., 2025).

ORCID

Eduardo A. Silva-Rodríguez https://orcid.org/0000-0001-9416-8653
Esteban I. Cortés https://orcid.org/0000-0001-6170-232X
Viviana Vasquez-Ibarra https://orcid.org/0009-0009-8944-4719
Nicolás Gálvez https://orcid.org/0000-0002-5129-0090
Jeremy Cusack https://orcid.org/0000-0003-3004-1586
Omar Ohrens https://orcid.org/0000-0001-9047-0804
Darío Moreira-Arce https://orcid.org/0000-0002-1188-496X
Ariel A. Farías https://orcid.org/0000-0001-7510-2253
José Infante-Varela https://orcid.org/0000-0003-2296-9616

REFERENCES

Agresti, A. (2007). *An Introduction to Categorical Data Analysis* (2nd ed.). John Wiley & Sons, Inc.

Ahumada, J. A., Fegraus, E., Birch, T., Flores, N., Kays, R., O'Brien, T. G., Palmer, J., Schuttler, S., Zhao, J. Y., Jetz, W., Kinnaird, M., Kulkarni, S., Lyet, A., Thau, D., Duong, M., Oliver, R., & Dancer, A. (2020). Wildlife insights: A platform to maximize the potential of camera trap and other passive sensor wildlife data for the planet. *Environmental Conservation*, *47*(1), 1–6.

Arandjelovic, M., Stephens, C. R., Dieguez, P., Maldonado, N., Bocksberger, G., Després-Einspenner, M. L., & Kühl, H. S. (2024). Highly precise community science annotations of video camera-trapped fauna in challenging environments. *Remote Sensing in Ecology and Conservation*, *10*(6), 702–724. https://doi.org/10.1002/rse2.402

Beery, S., Morris, D., & Yang, S. (2019). Efficient pipeline for camera trap image review. arXiv preprint arXiv:1907.06772. 10.48550/arXiv.1907.06772.

Böhner, H., Kleiven, E. F., Ims, R. A., & Soininen, E. M. (2023). A semi-automatic workflow to process images from small mammal camera traps. *Ecological Informatics*, *76*, 102150. https://doi.org/10.1016/j.ecoinf.2023.102150

Bubnicki, J. W., Norton, B., Baskauf, S. J., Bruce, T., Cagnacci, F., Casaer, J., & Desmet, P. (2023). Camtrap DP: An open standard for the FAIR exchange and archiving of camera trap data. *Remote Sensing in Ecology and Conservation*, *10*(3), 283–295. https://doi.org/10.1002/rse2.374

Burton, A. C., Beirne, C., Gaynor, K. M., Sun, C., Granados, A., Allen, M. L., & Kays, R. (2024). Mammal responses to global changes in human activity vary by trophic group and landscape. *Nature Ecology & Evolution*, *8*(5), 924–935.

Caravaggi, A., Banks, P. B., Burton, A. C., Finlay, C. M., Haswell, P. M., Hayward, M. W., & Wood, M. D. (2017). A review of camera trapping for conservation behaviour research. *Remote Sensing in Ecology and Conservation*, *3*(3), 109–122.

Celis, G., Ungar, P., Sokolov, A., Sokolova, N., Böhner, H., Liu, D., Gilg, O., Fufachev, I., Pokrovskaya, O., Ims, R. A., Zhou, W., Morris, D., & Ehrich, D. (2022). A versatile semiautomated image analysis workflow for time-lapsed camera trap image classification. *bioRxiv*, 2022–12. https://doi.org/10.1101/2022.12.28.522027

Choo, Y. R., Kudavidanage, E. P., Amarasinghe, T. R., Nimalrathna, T., Chua, M. A., & Webb, E. L. (2020). Best practices for reporting individual identification using camera trap photographs. *Global Ecology and Conservation*, *24*, e01294.

Cove, M. V., Kays, R., Bontrager, H., Bresnan, C., Lasky, M., Frerichs, T., Klann, R., Lee, T. E., Jr., Crockett, S. C., Crupi, A. P., Weiss, K. C. B., Rowe, H., Sprague, T., Schipper, J., Tellez, C., Lepczyk, C. A., Fantle-Lepczyk, J. E., LaPoint, S., Williamson, J., … Jordan, M. J. (2021). SNAPSHOT USA 2019: A coordinated national camera trap survey of the United States. *Ecology*, *102*(6), e03353. https://doi.org/10.1002/ecy.3353

Darras, K., Batáry, P., Furnas, B. J., Grass, I., Mulyani, Y. A., & Tscharntke, T. (2019). Autonomous sound recording outperforms human observation for sampling birds: A systematic map and user guide. *Ecological Applications*, *29*(6), e01954.

Davison, A., Birks, J. D., Brookes, R. C., Braithwaite, T. C., & Messenger, J. E. (2002). On the origin of faeces: Morphological versus molecular methods for surveying rare carnivores from their scats. *Journal of Zoology*, *257*(2), 141–143.

Farias, A. A., Sepúlveda, M. A., Silva-Rodríguez, E. A., Eguren, A., González, D., Jordán, N. I., & Svensson, G. L. (2014). A new population of Darwin's fox (*Lycalopex fulvipes*) in the Valdivian Coastal Range. *Revista Chilena de Historia Natural*, *87*, 1–3.

Forrester, T., O'Brien, T., Fegraus, E., Jansen, P. A., Palmer, J., Kays, R., & McShea, W. (2016). An open standard for camera trap data. *Biodiversity Data Journal*, *4*, e10197. https://doi.org/10.3897/BDJ.4.e10197

Gooliaff, T. J., & Hodges, K. E. (2019). Error rates in wildlife image classification. *Ecology and Evolution*, *9*(11), 6738–6740.

Johansson, Ö., Samelius, G., Wikberg, E., Chapron, G., Mishra, C., & Low, M. (2020). Identification errors in camera-trap studies result in systematic population overestimation. *Scientific Reports*, *10*(1), 6393.

Kays, R., Lasky, M., Allen, M. L., Dowler, R. C., Hawkins, M. T., Hope, A. G., & Cove, M. V. (2022). Which mammals can be identified from camera traps and crowdsourced photographs? *Journal of Mammalogy*, *103*(4), 767–775.

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K., Bailey, L. L., & Hines, J. E. (2017). *Occupancy estimation and modeling* (2nd ed., p. 641). Elsevier Academic Press. https://doi.org/10.1016/C2012-0-01164-7

Meek, P. D., Ballard, G., Claridge, A., Kays, R., Moseby, K., O'brien, T., & Townsend, S. (2014). Recommended guiding principles for reporting on camera trapping research. *Biodiversity and Conservation*, *23*, 2321–2343.

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E. J., Ware, J. J.,

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010. By Pontificia Universidad Catolica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

& Ioannidis, J. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9.

Murray, A., Sutherland, R., & Kays, R. (2023). Ecological effects of a declining red wolf population. *Animal Conservation*, *27*(3), 396–407. https://doi.org/10.1111/acv.12919

Niedballa, J., Sollmann, R., Courtiol, A., & Wilting, A. (2016). camtrapR: An R package for efficient camera trap data management. *Methods in Ecology and Evolution*, *7*(12), 1457–1462.

Potter, L. C., Brady, C. J., & Murphy, B. P. (2019). Accuracy of identifications of mammal species from camera trap images: A northern Australian case study. *Austral Ecology*, *44*(3), 473–483.

Priyadarshani, N., Marsland, S., & Castro, I. (2018). Automated birdsong recognition in complex acoustic environments: A review. *Journal of Avian Biology*, *49*(5), jav-01447.

Rovero, F., & Ahumada, J. (2017). The Tropical Ecology, Assessment and Monitoring (TEAM) Network: An early warning system for tropical rain forests. *Science of the Total Environment*, *574*, 914–923.

Rovero, F., Rathbun, G. B., Perkin, A., Jones, T., Ribble, D. O., Leonard, C., & Doggart, N. (2008). A new species of giant sengi or elephant-shrew (genus Rhynchocyon) highlights the exceptional biodiversity of the Udzungwa Mountains of Tanzania. *Journal of Zoology*, *274*(2), 126–133.

Sanderson, J., & Harris, G. (2013). Automatic data organization, storage, and analysis of camera trap pictures. *Journal of Indonesian Natural History*, *1*(1), 11–19.

Scotson, L., Johnston, L. R., Iannarilli, F., Wearn, O. R., Mohd-Azlan, J., Wong, W. M., & Fieberg, J. (2017). Best practices and software for the management and sharing of camera trap data for small and large scales studies. *Remote Sensing in Ecology and Conservation*, *3*(3), 158–172.

Sergeant, ESG. (2018). Epitools Epidemiological Calculators. Ausvet. http://epitools.ausvet.com.au

Silva-Rodriguez, E., Cortés, E., Vasquez-Ibarra, V., Gálvez, N., Cusack, J., Ohrens, O., Moreira-Arce, D., Farías, A., & Infante-Varela, J. (2025). Data from: A protocol for error prevention and quality control in camera trap datasets. *Figshare*. https://doi.org/10.6084/m9.figshare.28330655.v1

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437.

Steenweg, R., Hebblewhite, M., Kays, R., Ahumada, J., Fisher, J. T., Burton, C., & Rich, L. N. (2017). Scaling-up camera traps: Monitoring the planet's biodiversity with networks of remote sensors. *Frontiers in Ecology and the Environment*, *15*(1), 26–34.

Sundaresan, S. R., Riginos, C., & Abelson, E. S. (2011). Management and analysis of camera trap data: Alternative approaches (response to Harris et al. 2010). *Bulletin of the Ecological Society of America*, *92*(2), 188–195.

Thompson, W. (Ed.). (2004). *Sampling rare or elusive species: Concepts, designs, and techniques for estimating population parameters*. Island Press.

Treves, A. (2022). "Best available science" and the reproducibility crisis. *Frontiers in Ecology and the Environment*, *20*(9), 495.

Vélez, J., McShea, W., Shamon, H., Castiblanco-Camacho, P. J., Tabak, M. A., Chalmers, C., & Fieberg, J. (2023). An evaluation of platforms for processing camera-trap data using artificial intelligence. *Methods in Ecology and Evolution*, *14*(2), 459–477.

Wearn, O. R., & Glover-Kapfer, P. (2017). Camera-trapping for conservation: A guide to best-practices. WWF Conservation Technology Series 1(1). WWF-UK, Woking, United Kingdom.

Young, S., Rode-Margono, J., & Amin, R. (2018). Software to facilitate and streamline camera trap data management: A review. *Ecology and Evolution*, *8*(19), 9947–9957.

Zett, T., Stratford, K. J., & Weise, F. J. (2022). Inter-observer variance and agreement of wildlife information extracted from camera trap images. *Biodiversity and Conservation*, *31*(12), 3019–3037.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Table S1. Framework to categorise potential error types in camera trap (CT) datasets, their causes, and recommended precautions.

Figure S1. Estimated sample size required (i.e. number of media files to be revised during quality control) as a function of the expected proportion of misclassified pictures.

How to cite this article: Silva-Rodríguez, E. A., Cortés, E. I., Vasquez-Ibarra, V., Gálvez, N., Cusack, J., Ohrens, O., Moreira-Arce, D., Farías, A. A., & Infante-Varela, J. (2025). A protocol for error prevention and quality control in camera trap datasets. *Journal of Applied Ecology*, *62*, 773–782. https://doi.org/10.1111/1365-2664.70010

13652664, 2025, 4, Downloaded from https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2664.70010 By Pontificia Universidad Católica De Chile, Wiley Online Library on [03/04/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License