The data reliability in ecological research: a proposal for a quick self-assessment tool

Corrado Battisti[1*], Giuseppe Dodaro[2], Daniel Franco[3]

Abstract - Researchers and practitioners in basic and applied ecology provide to private or public clients assessment documents on various concern topics, such as the state of ecosystem components, the type of threats and their regime or the level of pressure and impact on biodiversity. These assessments, carried out by ecological field studies, may be strategic in addressing conservation research, plans and actions. Therefore, data provided in these documents should be characterized by a high reliability, that is, they should be based on standard methods and protocols, independence of data samples, absence of pseudo-replication, control of different levels of detectability among sampled individuals or species, high level of precision and accuracy etc. In this paper we propose a simple two-data-sheet format for a data reliability assessment of a professional study that may facilitate a rapid check of the more important requirements of a correct ecological field research. This format may be useful to students, technicians, professionals and researchers as well as public or private commissioning agencies (e.g. to evaluate the suitability of the study, possibly suggesting additions or modifications).

Key-words: ecological field study, data reliability, self-assessment.

Riassunto - Ricercatori, tecnici, professionisti, studenti di formazione naturalistico-ambientale sono spesso chiamati a redigere ricerche, tesi o consulenze per enti di ricerca e società private nei settori della pianificazione, gestione e conservazione. Tali documenti sono estremamente eterogenei e vanno dalle analisi di base (es., sullo stato delle componenti ambientali in un'area) ad indagini più complesse comprendenti valutazioni sulle relazioni causa-effetto tra attività antropiche e componenti ambientali. Le indagini inserite in queste consulenze costituiscono pertanto un momento importante del processo decisionale, consentendo di effettuare diagnosi e/o previsioni su specifici aspetti ambientali e indirizzando le scelte di piani, progetti, programmi. A fronte di questa responsabilità, è importante fornire relazioni che contengano dati di elevata attendibilità. In questo lavoro viene proposta una scheda sintetica di autovalutazione della attendibilità (*reliability*) dei dati inseriti

in una ricerca effettuata da tecnici, ricercatori e professionisti del settore ambientale, con particolare riferimento agli studi di campo nel settore naturalistico. Essa consente al tecnico di verificare se la ricerca ha rispettato i requisiti che garantiscono un'elevata attendibilità dei dati (standardizzazione, replicazione, indipendenza dei dati, controllo della *detectability*, ecc.). Al tempo stesso, questa scheda può fornire al soggetto committente (pubblico o privato) una base per verificare l'adeguatezza degli studi presentati e richiedere eventuali modifiche e integrazioni.

Parole chiave: studi di campo ecologici, attendibilità dei dati, autovalutazione.

INTRODUCTION

Researchers and professionals in environmental and ecological sectors provide a wide range of documental consultancies (reviews, field or experimental research, advices), spanning from the site environmental samplings, monitoring and assessment to the complex estimation of the variations of the ecosystem components due to anthropogenic activities, pressures and threats induced, for example, by large planning strategies or local single measures.

The research building up these documents come from reviews, metadata analyses or from field sampling (manipulative or mensurative ecological field research *sensu* Hurlbert, 1984), and outputs can be originally elaborated or processed by specific softwares (e.g. Band *et al.,* 2005). The research and the related documents are critical, allowing estimations and forecasting on specific ecological/environmental aspects and the dimension of possible cause-effect relationships among anthropogenic activities, pressures, and impacts on ecosystems components and biodiversity (Morrison, 2002; Salafsky *et al.,* 2008).

Therefore, researchers and professionals providing these documents have a great responsibility and their outputs – based on their research – can be crucial in basic and applied research (e.g. orienting decision-making processes impacting on public goods, influencing the authorization process of programs/plans/projects by the definition of different solutions or suggesting the prescription of mitigation and compensation measures). Given that these research data could be crucial to determine the state and the impact of/on populations, communities and ecosystems they must be reliable.

A shallow baseline survey of a site of interest, the use of unsuitable methods or protocols not well-structured (i.e. not representative) in time and space and the lack of standards can bring to inaccurate or unreliable results with strong consequences on the significance of pressure and impact relationships with the ecological components (Sutherland, 2006). Considering, for instance, rare and/or elusive species which are occasionally and/or seasonally present in a study site a trivial un-standardized survey can bring to their presence and abundance underestimation that, in turn, can bring to a weak or wrong environmental impact assessment (e.g. large migrant raptors moving across turbines in a wind farm) and a subsequent serious threat for the species of conservation concern. From these trial and error history a growing literature pointed out the requirement for ecological sampling reliability (Best, 1975; Cochran, 1977; Bibby *et al.*, 1992; Sutherland, 2000; Morrison, 2002; Thompson, 2002; Sutherland, 2006).

The aim of this short communication is to introduce a self-assessment synthetic sheet on the reliability of professional consultancies on ecological field research.

METHODS

We based the implementation sheet on the field research requirements developed for animal and plant ecology (i.e. Eberhardt & Thomas, 1991; Bibby *et al.*, 1992; Morrison, 2002; Sutherland *et al.*, 2004; Gotelli & Ellison, 2004; Greenwood & Robinson, 2006; Sutherland, 2006).

In the first part a quick assessment of the methodological design correctness is presented. It is divided in three sections: preparation, method, protocol. The queries are in an expert based format with yes/no qualitative judgements and an annotation box (Tab. 1).

In the preparation section the assessment was designed asking to verify if the hypothesis, the aims and the objectives were clearly and correctly stated and if a full review of the specific literature on the topic was done.

In the method section we ask for checking if the selected method is appropriate for the stated hypothesis, and if a fitting standard method exists.

In the protocol section we ask to verify:

1) if the sampling is representative of the numeric, spatial and time scale and if the grain (study frame) covers the studied processes. When compared to the entire phenomenon to assess, a weakness or a lack in representativeness occurs when the sample was too limited (weak or lacking numerical representativeness; e.g. too few specimens were collected), a too small study area was defined (weak or lacking spatial representativeness; e.g. too few sample points for check target species were located) or the temporal range for sampling was too limited or not coincident with the timing of the phenomenon under study (weak or lacking time representativeness; e.g. few days defined to sample an anthropogenic disturbance with an annual duration or a sampling carried out outside the period of presence of target species of conservation concern: for example, a sampling on the impact of water stress on migrant birds in a wetland must match the spring or autumn period when these species perform their migration).

2) If the sampling is adequately spatially and temporally replicated; since phenomena are variable in space and time, replication provide a measure of this variability (Eberhardt & Thomas, 1991). Consequently, data from ecological field studies that are based on a limited number of replicated samples are not reliable. Examples may be: for birds, mist-net sessions replicated only for a few days or weeks and then used to study the annual temporal pattern of dispersion by ringing techniques or, for bats, using a limited number of bat-detector based point transects to compile a species checklist for a wide geographical area. In this last case we have at least two design flaws: i) species identification by ultrasound is rarely reliable (in particular with bat faunas accounting for Italian species), and ii) small sample size.

3) If the sampling design follow a regular, random or stratified distribution of sample points. For example, when spatial heterogeneity occur at the study scale a stratification of sampling points is needed: point counts detecting animals or plants in a patchy Mediterranean landscape ("arlequin landscape" *sensu* Blondel & Aronson, 1999) should be stratified and not simply random located since a large part of the information, related to different habitat types of the landscape, may be lost with a randomization lacking in stratification.

4) If data independence and pseudo-replication have been accounted for (see Hurlbert, 1984). Many vagile animals (e.g. butterflies, birds, bats, fishes) actively move during the sampling period and, without paying attention to the independence of the data, the same individuals can be sampled multiple times from multiple sampling points: this induces data pseudo replication and as a consequence a less powerful – if unreliable at all – analysis.

5) If the individual detectability has been accounted for. Detectability of a sampling unit (e.g. an individual of a species) depends both by extrinsic and intrinsic factors depending on weather conditions, vegetation composition, structure and dynamic, geomorphology, behavioural differences of the studied species, expertise, fitness and behaviour of the researchers that carried out the study (O'Connor & Hicks, 1980; Sutherland, 2006).

6) If accuracy and precision are controlled.

The second part (Tab. 2) summarizes the results of the first part with the aim to point out the presence of possible bias. The assessment is based on a qualitative judgement (high, sufficient, low reliability) but a quantitative approach using scores (integer values, e.g. from 1, low, to 3, high) may be also suggested. The last row gives an overall judgement on the reliability level of field research. When using scores, the assignment of the final judgement can derive: i) calculating the mode from the values in column, or ii) summing up the values along the column and then match that total with a predefined interval which give an overall rank.

Tab.1 - First part of the self assessment sheet for the reliability of a data sampling in basic and applied ecology research.

Sampling reliability		yes/no	note
Preparation			
Work hypothesis, aims	Are the hypotheses, aims and objectives of the research correctly and clearly formulated?		
Literature review	Has been a full review on the specific literature on the topic carried out?		
Methods			
Methods used	Are the methods used appropriate to the hypotheses and aims of the research? Are they sufficient to give answers to the research questions?		
	Are there intrinsic weaknesses in the methods used? If so, can they mislead data interpretation and result discussion?		
Standard methods	Does a standard method exist?		
Protocol			
Representativeness *space*	Is the adopted sampling spatially representative? Is the sampling protocol representative of the studied process? Are the sampling scale and grain appropriately covering the studied processes?		
time	Is the sampling period representative of the studied process?		
numeric	Is the sample size statistically representative?		
Repetition			
space	Has the sampling been replicated in space?		
time	Has the sampling been replicated in time?		
Design (regular, random or stratified sampling distribution)	Has the sampling been distributed considering the site heterogeneity characteristics (stratified or regular or distributed)?		
Sample independence and pseudo-repetition	Are the data independent?		
	Does the possibility of pseudo-repetition exist?		
Detectability	Do species-intrinsic factors influencing species detectability exist?		
	If so, does the adopted method/protocol account for these extrinsic factors?		
	Do species-extrinsic factors (whether or vegetation shielding effects influencing the observer) influencing species detectability exist?		
	If so, does the adopted method/protocol account for these extrinsic factors?		
Precision and accuracy	Do precision and accuracy are controlled?		

Tab. 2 - Second part of the self assessment sheet for the reliability of a data sampling in basic and applied ecology research. The nominal ranks may be converted in integer values (e.g. scores from 1, low, to 3, high) to allow a quantitative approach (see text).

Work hypothesis, aims		Clearly defined				
Literature review		deep	sufficient	superficial (preliminary)		
Methods		appropriate	improvable	insufficient	*bias*	
Standard method		high	good -improvable	insufficient	absent	possible	certain
Representativeness	Space	high	good –improvable	insufficient	absent	possible	certain
	Time	high	good –improvable	insufficient	absent	possible	certain
	Numeric	high	good -improvable	insufficient	absent	possible	certain
Repetition	Space	high	good -improvable	insufficient	absent	possible	certain
	Time	high	good -improvable	insufficient	absent	possible	certain
Sampling distribution		optimum	good -improvable	insufficient (not defined)	absent	possible	certain
Sample data independence		high	good -improvable	insufficient (possible data pseudo-repetition)	absent (data pseudo-repetition)	possible	certain
Detectability	intrinsic	high	good -improvable	insufficient (sampling bias)	absent (sampling bias)	possible	certain
	extrinsic	high	good -improvable	insufficient (sampling bias)	absent (sampling bias)	possible	certain
Precision and accuracy		high	good -improvable	insufficient (sampling bias)	absent (sampling bias)	possible	certain
				Overall research reliability	High	Sufficient	Low

DISCUSSION AND CONCLUSION

The proposed sheet (Tab. 1, Tab. 2) can be used by different subjects (students, researchers, environmental practitioners and professionals), and for different aims. In particular the sheet can be used as a structured self-assessment tool, to find out possible weaknesses in an ecological field research design, and correct them adaptively. It can also be applied as a structured framework to communicate in an effective way the quality of the work performed by a professional advisor, pointing out the strengths and weaknesses.

In a similar way, the sheet can objectively define whether to use a research survey results conservatively or not, as well as if it is needed to deepen some specific aspects of the research to come to a reliable assessment. This is the case when the constrains (as time and budget) given by the client appear to be insufficient to representatively describe the ecological processes or phenomenon and to provide robust conclusions. Finally, the sheet can help decision makers (in the case of projects, plans, designs assessments) to quickly check the robustness and reliability of the available information and, possibly, to request possible changes or addictions.

In conclusion, the sheet can be used to check the reliability of methods and protocols used for field data sampling and analyses, overcoming the need for a right collection of preliminary information. Since there is a plethora of different types and regimes of field studies that researchers and professionals in basic ecology, environmental planning, management and conservation carried out, we are aware that there can be no one 'catch-all' assessment tool, included the one proposed. For example, this sheet can be insufficiently effective in case of studies based on expert judgements, which are often intrinsically uncertain (Linstone & Turoff, 1975). Therefore, we recommend the proposed evaluation framework only for original sampling protocols in ecological field studies which provide a collection of data suitable to obtain reliable information on specific components, factors, ecological processes and phenomena in a specific study area.

However, depending of the different type of study, it is possible to integrate or substitute some of the requirements since that the sheet has been designed as a modular tool, to be adapted to specific needs, yet maintaining its logical setting and validity.

Acknowledgements

REFERENCES

Band W., Madders M. & Whitfield D.P., 2005 – Developing field and analytical methods to assess avian collision risk at wind farms. In: Birds and Wind Power. Janss G., Ferrer M. & de Lucas M. (eds.). *Lynx Edicions*, Barcelona: 259-275.

Best L.B., 1975 – Interpretational errors in the mapping method as a census technique. *Auk*, 92: 452-460.

Bibby C.J., Burgess N.D., Hill D.A. & Mustoe S.H., 1992 – Bird Census Techniques. *Academic Press*, London.

Blondel J. & Aronson J., 1999 – Biology and wildlife of the Mediterranean region. *Oxford University Press*, Oxford.

Cochran W.G., 1977 – Sampling Techniques. 3rd edition. *John Wiley & Sons*, New York.

Eberhardt L.L. & Thomas J.M., 1991 – Designing environmental field study. *Ecological Monographs*, 61: 53-73.

Gotelli N.J. & Ellison A.M., 2004 – A primer of ecological statistics. *Sinauer*, New York.

Greenwood J.J.D. & Robinson R.A., 2006 – Principles of sampling. In: Ecological census techniques. Sutherland W.J. (ed.). *Cambridge University Press*, Cambridge: 11-86.

Hurlbert S.H., 1984 – Pseudoreplication and the design of ecological field experiments. *Ecological. Monographs*, 54: 187-211.

Linstone H.A. & Turoff M., 1975 – The Delphi Method: techniques and applications. *Addison-Wesley*, New York.

Morrison M.L., 2002 – Wildlife Restoration. Techniques for habitat analysis and animal monitoring. *Island Press*, Washington.

O'Connor R.J. & Hicks R.K., 1980 – The influence of weather conditions on the detection of birds during Common Birds Census fieldwork. *Bird Study,* 27: 137-151.

Salafsky N., Salzer D., Stattersfield A.J., Hilton-Taylor C., Neugarten R., Butchart S.H.M., Collen B., Cox N., Master L.L., O'Connor S. & Wilkie D., 2008 – Standard Lexicon for Biodiversity Conservation: Unified Classifications of Threats and Actions. *Conservation Biology*, 22: 897-911.

Sutherland W.J., 2000 – The Conservation Handbook. *Wiley Online Library*, New York.

Sutherland W.J., 2006 – Ecological census techniques. *Cambridge University Press*, Cambridge.

Sutherland W.J., Newton I. & Green R.E., 2004 – Bird Ecology and Conservation. *Oxford University Press*, Oxford.

Thompson S.K., 2002 – Sampling. 2nd edition. *John Wiley & Sons*, New York.