Biological Conservation 223 (2018) 82–96

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

The contribution of scientific research to conservation planning

Louise Mair[a,*], Aileen C. Mill[a], Peter A. Robertson[a], Steven P. Rushton[a], Mark D.F. Shirley[a], Jon Paul Rodriguez[b], Philip J.K. McGowan[a]

[a] School of Natural and Environmental Sciences, Ridley Building 2, Newcastle University, Newcastle upon Tyne NE1 7RU, United Kingdom
[b] IUCN Species Survival Commission, Provita, Center for Ecology, Venezuelan Institute for Scientific Investigation, Apdo. 20632, Caracas 1020-A, Venezuela

ARTICLE INFO

Keywords:
Conservation prioritisation
Literature review
Research-implementation gap
Topic modelling
Transdisciplinary research

ABSTRACT

Conservation planning plays an instrumental role in facilitating progress towards biodiversity targets by providing practitioners with the tools required to allocate resources and implement actions. However, the utility of a burgeoning scientific literature to on-the-ground conservation has been questioned. Given such criticisms, and the lack of progress towards the global Aichi Biodiversity Targets, we aim to assess the contribution of scientific research to the field of conservation planning. We applied topic modelling to a body of literature consisting of 4471 articles pertaining to conservation planning published between 2000 and 2016. We quantified changes in topic popularity, and assessed the extent to which different topics were addressed within the same articles. We found that research into the status of species and habitats was most prevalent, the process of action planning received considerably less attention, and implementation attracted the least research of all. The scientific literature was thus dominated by biological rather than socio-political research, and furthermore showed a general lack of inter-disciplinary research, which is problematic given that ultimately it is the socio-political context that will determine the success of conservation efforts. The number of publications on implementation and monitoring declined over time, suggesting a waning interest in publishing evidence of plan effectiveness, and that limited efforts have been made to address the "implementation crisis". We suggest that filling research gaps, through integration of the social sciences and placing greater value on evidence syntheses, would push scientific research towards greater applicability and help to provide the necessary information to achieve global biodiversity targets.

1. Introduction

Conservation planning is the process of "deciding where, when and how to allocate limited conservation resources" (Pressey and Bottrill, 2009). Planning provides practitioners with the information and direction required to allocate resources and implement actions, ranging from the recovery of endangered species (Clark et al., 2002) to the establishment of large-scale protected area networks (Margules and Pressey, 2000). As global conservation targets have evolved since the adoption of the Convention on Biological Diversity (CBD) at the Rio Earth Summit in 1992, the need for conservation planning has become increasingly evident, and planning is now considered essential for achieving the current global Aichi Biodiversity Targets (CBD, 2010, 2015). Each of the twenty Aichi targets was designed to contribute towards halting the global loss of biodiversity by 2020 (CBD, 2010), and conservation planning should play an instrumental role in facilitating progress towards these targets by providing the strategic framework for the implementation of connected, ecologically representative protected area networks (Aichi Target 11; e.g. Pollock et al., 2017; Venter et al., 2018) and the prevention of species extinctions (Aichi Target 12; e.g. Whitfield et al., 2006; Challender et al., 2014).

Research into conservation planning aims to assist progress towards such ambitious conservation targets, yet there are criticisms about the lack of applicability of much scientific work to practical conservation efforts such as habitat restoration or the designation of protected areas (Knight et al., 2008; Barmuta et al., 2011). Furthermore, current evidence indicates that the majority of the Aichi Biodiversity Targets are unlikely to be met (CBD, 2014); species extinctions and declines have not been halted (Tittensor et al., 2014), and while protected area networks are likely to meet the area coverage targets of 17% terrestrial and 10% marine, they do not adequately cover ecoregions or important areas for biodiversity (Butchart et al., 2015). In the context of current conservation shortcomings and deliberations over the utility of research, it is timely to assess the quantity and diversity of scientific research into conservation planning, and hence assess the availability and

* Corresponding author.
E-mail address: Louise.Mair@newcastle.ac.uk (L. Mair).

https://doi.org/10.1016/j.biocon.2018.04.037
Received 23 February 2018; Received in revised form 23 April 2018; Accepted 27 April 2018
0006-3207/

applicability of information and advice that can build towards achieving global biodiversity targets.

There is a broad range of different conservation planning frameworks outlined in both the scientific and grey literature (Redford et al., 2003; Pressey and Bottrill, 2009). These frameworks encompass many steps, with each step falling loosely into three broad categories: (i) assessing the current status of, and threats to, species or areas of conservation interest; (ii) determining what actions should be taken; and (iii) implementation and monitoring (Knight et al., 2006a). The specifics of each step can vary greatly among approaches, and different planning frameworks may diverge on issues such as the process of identifying explicit conservation objectives, and the incorporation of socio-economic considerations (Pressey and Bottrill, 2009). Furthermore, planning is a non-linear process, and adaptive management and the revision of plans in response to monitoring outcomes is required for success (Grantham et al., 2010).

The complete conservation planning process is complex, and scientific research projects often focus on in-depth examinations of individual steps or processes within the overall framework. Studies may, for example, assess data requirements (Boitani et al., 2011), incorporate costs estimates (Carwardine et al., 2010), or evaluate the suitability of taxonomic surrogates (Rodrigues and Brooks, 2007). This fragmentation of the overall process leads to a large and complex body of literature, and it has been argued that the consideration of individual aspects of the planning process in isolation can result in a disconnect between scientific advance and practical application (Knight et al., 2008). Recent research has suggested that landscape genetics has so far failed to make much impact on conservation planning (Keller et al., 2015) and species distribution models are used less often in planning than might be expected given the proliferation and sophistication of available methods (Tulloch et al., 2016). Furthermore, planning exercises are frequently carried out without the engagement of the end-user or relevant stakeholders, with one review finding that the majority of the publications considered had the aim of improving research techniques rather than achieving implementation (Knight et al., 2008). These issues bring into question the applicability of much of the research pertaining to conservation planning, and emphasise that research direction has different drivers to conservation needs. For example, funding availability has been shown to stimulate research priorities, and this is subject to politics and the changing popularity of research topics (Stroud et al., 2014).

Obtaining an overview of the availability of information in such a vast and complex body of literature is challenging, particularly when the aim is to capture the full extent of the publishing landscape. Topic modelling provides a statistical tool to assess the content of articles in a corpus (a large body of literature; Blei and Lafferty, 2009). The approach makes use of the co-occurrence patterns of words in article abstracts to identify a range of topics which represent the main ideas present in a corpus (Griffiths and Steyvers, 2004). Topic modelling provides quantitative rigour to summarising themes and allows synthesis across disparate information sources covering different biological, spatial and temporal scales (Westgate et al., 2015). The approach has recently been applied within ecological science to analyse publishing trends in arid ecology research (Greenville et al., 2017), and to compare the topics of conservation-prioritisation articles that did and did not apply species distribution models (Tulloch et al., 2016).

Here, we use topic modelling to assess the contribution of scientific research to the field of conservation planning. We quantify which aspects of the conservation planning process receive the most attention in the published literature, and how topic popularity has changed over time. We also assess the extent to which different aspects of conservation planning are either linked to the broader process or studied in isolation, in order to challenge the implicit assumption that research related to conservation planning is suitable for practical application. We aim to capture the full extent of the publishing landscape; the corpus we analyse consists of 4471 articles published from 2000 to 2016 pertaining to conservation planning. Consideration of this large body of literature allows us to determine potential gaps and neglected fields which could be addressed in order to aid progress towards global biodiversity targets.

2. Methods

2.1. Literature search

We searched Web of Science for articles published from 2000 to 2016 using the terms "conservation plan*" or "recovery plan*" and also "biodiversity", "species", "habitat*" or "ecosystem*". We included only articles published in English and which were categorised as articles or reviews according to document type, giving 4619 documents.

Citations and abstracts were downloaded and imported in to the program R (R Core Team, 2017) using the package *bibliometrix* (Aria and Cuccurullo, 2016). Articles which were categorised as "in proceedings" and articles that did not have abstracts were removed. This gave 4471 documents.

2.2. Abstract cleaning

A small number of documents included abstracts written in both English and either Spanish or French; the identification and removal of non-English text is detailed in Appendix A. Abstracts were then transformed into a corpus and processed using the R package *tm* (Feinerer et al., 2008). Search terms were removed as these words were common to all abstracts. Numbers written as words and digits were also removed (Grun and Hornik, 2011). The pre-defined list of English stop-words provided in the *tm* package (Feinerer et al., 2008) were removed and we expanded this list by removing the components of abbreviated words on the stop-words list, as well as "also" which was the most common synonym of the stop-word "and" (see Table A1 for list of words removed). Finally, terms added by the publishers for copyright reasons were removed, hyphens and forward slashes were changed to spaces, and all other punctuation was removed (sensu Grun and Hornik, 2011).

The suffixes of the abstract words were then removed to reduce words to their common root, and words that appeared in five or fewer articles were removed (following methods in Griffiths and Steyvers, 2004; Lu et al., 2017 demonstrated that removal of infrequent words had little impact on model performance). This gave a final corpus with a vocabulary of 4201 words.

2.3. Topic modelling

Topic modelling defines topics within a corpus based on sets of words that co-occur with unusual frequency (Griffiths and Steyvers, 2004; Grun and Hornik, 2011). Each topic can be understood as a meaningful combination of ideas within the corpus. Documents belong simultaneously to several topics, making topic modelling an appropriate tool to examine the cross-cutting nature of many research documents.

The inputs to the topic model are a matrix of document-word frequencies and the number of topics to be identified. The most appropriate number of topics for the corpus can determined a priori by carrying out block-cross validation and measuring perplexity (which is a measure of likelihood; Grun and Hornik, 2011). The model then provides the weight that each word contributes to a topic, allowing the main ideas of each topic to be inferred. Topic distributions vary over documents and the weight of each topic within a document is provided, which allows the main topic and diversity of topics within a document to be identified.

We identified 40 topics in the corpus by fitting a Latent Dirichlet Allocation (LDA) model with Gibbs sampling using the R package *topicmodels* (Grun and Hornik, 2011). Our results from block-cross validation (described in supporting information) indicated that model