Inferring public interest from search engine data requires caution

In a recent communication, Burivalova et al. (2018) analyzed Google Trends data with respect to interest in conservation and made two important claims: (1) that public interest in conservation is rising since 2004 and (2) that conservation and climate change–related topics have similar levels of public interest. Their assertions are based on a proposed new method to back-adjust Google Trends data from relative to absolute search volume. Their results contradict those of earlier studies using similar data to claim that public interest in conservation is waning (McCallum and Bury 2013; Troumbis 2017). However, after reproducing Burivalova et al.’s analysis and correcting an error in their algorithm, we find that their claims may not hold under scrutiny.

We applaud the effort by Burivalova and colleagues to develop new ways to explore Google Trends data, and we see great potential in their proposed method. However, their assertions rest on implicit assumptions that (1) an observed growth in absolute search volume reflects an increase in public interest and that (2) their method correctly reflects differences in public interest across topics. We reproduced Burivalova et al.’s Figure 3 using the same methodology (Figure 1a) and found that their metric was adjusted in relation to the maximum search volume (highest $a_i$) observed within each topic (Equation 8 and step 6 in their WebFigure 4). This is valid only when analyzing topics individually, because it does not preserve differences in search volume across topics. To preserve such differences using Equation 8, the data must be scaled using a value of $a_i$ that reflects the maximum search volume observed across all topics. Doing so, we find that search volume for climate-change topics is approximately double that observed for conservation topics in recent years (apart from “Extinction”; Figure 1b).

Furthermore, the rapid growth of search engine use over time means that the absolute number of searches is likely to have increased for any topic, independently of public interest. Given this, a more reasonable assessment of how public interest for conservation topics has changed might be achieved by comparing the rate of change across topics (Ngheim et al. 2016). Calculating the ratio between searches for conservation topics and searches for the topic “climate change”, we find results are not constant over time. The ratio decreases for all terms between 2004 and 2007, and again after 2014, indicating that searches for “climate change” took prominence in relation to conservation topics during these periods (Figure 1c). Repeating the analysis with “global warming” produces similar results, which concur with reports of an increase in news media attention toward climate-change topics during these periods (Legagneux et al. 2018). Overall, our results suggest that temporal changes in public interest toward conservation are different and more nuanced than those presented in Burivalova et al. (2018).

In our experience, temporal dynamics of search engine usage are complex, and inferences on changes in public interest derived from such data should be approached with caution. Accounting (or not) for multiple confounding factors such as the growth in internet access, search engine usage, time spent online, and the changing nature of internet usage (eg work versus leisure; see Ficetola [2013]) is likely to produce markedly different results. Furthermore, as Burivalova and colleagues rightly point out, Google

Figure 1. Changes in search volume for conservation and climate-change topics over time. (a) Reproduction of Figure 3 in Burivalova et al. (2018), where dashed lines represent original search volume obtained from Google Trends and solid lines represent adjusted search volume using the method proposed by Burivalova et al. (2018). (b) Search volume data adjusted according to our proposed method, which accounts for differences in search volume across topics. (c) The monthly proportion of absolute searches for conservation topics (color key same as in [a]) in relation to climate change, indicated by the thin solid lines, and underlying significant temporal trends, indicated by thick solid lines. The dashed line indicates a similar search volume for conservation topics and climate change; values above the dashed line indicate higher volume of searches for conservation topics and vice versa. Points of change in the temporal trends were identified using package “strucchange” and statistically non-zero trends (***, $P < 0.001$) were estimated using generalized linear models in R software v.3.4.2 (R Core Team 2017).
Trends data present an additional challenge in this context because the raw data are unavailable due to proprietary constraints. Moreover, considering only a single data source may produce a biased view of changes in public interest toward any topic. We believe that combining results from multiple sources (Veríssimo et al. 2016; Cooper et al. 2019; Jarić et al. 2019) – such as different search engines, social media platforms, online news media, Wikipedia page views, internet blogs and forums – is more likely to provide meaningful insights on social and cultural trends.

The exploration of conservation-related topics using digital data sources provides new opportunities for conservation science and practice (Sutherland et al. 2018). We are advocates of such potential (Di Minin et al. 2015; Ladle et al. 2016; Soriano-Redondo et al. 2017) and actively encourage efforts to positively engage with cultural trends for the benefit of conservation. We have also faced challenges and limitations associated with these methods including, for example, issues of semantic complexity, language dynamics, and data collection and curation (Ladle et al. 2016), and have been careful to elucidate them. Our research has aimed to offer practical solutions to some of these challenges (Jarić et al. 2016; Correia et al. 2018; Roll et al. 2018), and we have recently established a Conservation Culturomics working group within the Society for Conservation Biology. This group embodies our belief in open collaboration, and we hope it will facilitate knowledge-sharing and collaborative efforts to overcome challenges through a welcoming, supportive, and stimulating environment. We encourage all interested parties to join this endeavor toward advancing digital methods for conservation.

Ricardo A Correia1,2†, Enrico Di Minin3,5†, Ivan Jarić4,6,7†, Paul Jepson8†, Richard Ladle9, John Mittermeier8†, Minin3,4,5†, Ivan Jarić6,7†, Paul Jepson8†, 1DBIO & CESAM-Centre for Environmental and Marine Studies, University of Aveiro, Campus Universitário de Santiago, Aveiro, Portugal; 2Institute of Biological and Health Sciences, Federal University of Alagoas, Maceió, Brazil *(rahc85@gmail.com); 3Department of Geosciences and Geography, University of Helsinki, Helsinki, Finland; 4Helsinki Institute of Sustainability Science, University of Helsinki, Helsinki, Finland; 5School of Life Sciences, University of KwaZulu-Natal, Durban, South Africa; 6Biology Centre of the Czech Academy of Sciences, Institute of Hydrobiology, České Budějovice, Czech Republic; 7University of South Bohemia, Faculty of Science, Department of Ecosystem Biology, České Budějovice, Czech Republic; 8School of Geography and the Environment, University of Oxford, Oxford, UK; 9Mitrani Department of Desert Ecology, The Jacob Blaustein Institutes for Desert Research, Ben-Gurion University of the Negev, Midreshet Ben-Gurion, Israel; 10Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Cornwall Campus, UK; 11CIBIO/InBio, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Universidade do Porto, Portugal; 12CIBIO/InBio, Centro de Investigação em Biodiversidade e Recursos Genéticos, Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, Portugal; 13Department of Zoology, University of Oxford, Oxford, UK; 14Oxford Martin School, University of Oxford, Oxford, UK; 15San Diego Zoo Institute for Conservation Research, Escondido, CA; † all authors contributed equally to this work
