Data mining on YouTube reveals fisher group-specific harvesting patterns and social engagement in recreational anglers and spearfishers

Valerio Sbragaglia^{1,2}, Ricardo A. Correia^{3,4}, Salvatore Coco⁵, Robert Arlinghaus^{2,6}

¹ Institute for Environmental Protection and Research (ISPRA), Livorno, Italy

² Department of Biology and Ecology of Fishes, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany.

³ DBIO & CESAM-Centre for Environmental and Marine Studies, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

⁴ Institute of Biological and Health Sciences, Federal University of Alagoas, Maceió, Alagoas, Brazil

⁵ Dipartimento di Scienze Biologiche, Geologiche e Ambientali, University of Bologna, Via S. Alberto 163, 48123 Ravenna (Italy)

⁶ Division of Integrative Fisheries Management, Department of Crop and Animal Sciences, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Philippstrasse 13, Haus 7, 10115 Berlin, Germany

* Corresponding author email: valeriosbra@gmail.com

ABSTRACT

Data about recreational fisheries are scarce in many areas of the world. In the absence of monitoring data collected *in situ*, alternative data sources, such as digital applications and social media platforms, have the potential to produce valuable insights. Yet, the potential of social media for drawing insights about recreational fisheries is still underexplored. In this study, we applied data mining on YouTube videos to better understand recreational fisheries targeting common dentex (*Dentex dentex*), an iconic species of Mediterranean recreational fisheries. We chose this model species because of ongoing controversies about the relative impact of recreational angling and recreational spearfishing on its conservation status. In Italy alone, from 2010 to 2016 recreational spearfishers posted 1051 videos compared to 692 videos posted by recreational anglers. Only the upload pattern of spearfishing videos followed a seasonal pattern with peaks in July, suggesting seasonality of spearfishing catches of *D. dentex* – a trend not found for anglers. The average mass of the fish declared in recreational angling videos (6.43 kg) was significantly larger than the one in spearfishing videos (4.50 kg). Videos posted by recreational spearfishers received significantly more likes and comments than those posted by recreational

anglers, suggesting that the social engagement among recreational spearfishers was stronger than in anglers. We also found that the mass of the fish positively predicted social engagement in recreational spearfishing videos, but not in videos posted by recreational anglers. This could be caused by the generally smaller odds of catching large *D. dentex* by spearfishing, possibly explaining why posting videos with particularly large specimen triggered larger social engagement by recreational spearfishers. Our case study demonstrates that data mining on YouTube can be a powerful tool to provide complementary data on controversial and data-poor aspects of recreational fisheries.

KEYWORDS: *Dentex dentex*, Mediterranean, mass, monitoring, social media

INTRODUCTION

Recreational fishers contribute a non-negligible proportion to the total fish harvested around the globe (Cooke and Cowx, 2004; World Bank, 2012; Hyder et al., 2018; Brownscombe et al., 2019). At the same time, recreational fishing generates manifold social and economic benefits to fishers and society at large (Rudd et al., 2002; Hyder et al., 2018; Brownscombe et al., 2019). However, many recreational fisheries are characterized by widespread lack of data regarding fishing effort, harvest and socio-economic aspects, making global monitoring and assessment difficult (FAO, 2012a; Arlinghaus et al., 2016; Elmer et al., 2017). Available data are often obtained with methods that are resource intensive (e.g., creel surveys; random telephone surveys; Pollock et al., 1994; Board and Council, 2006; Hartill et al., 2012; Rocklin et al., 2014; Hartill and Edwards, 2015; Zarauz et al., 2015; Bellanger and Levrel, 2017; Holdsworth et al., 2018; Hyder et al., 2018; van Poorten and Brydle, 2018). As a complement to more traditional methods, less resource-demanding data sourced from digital applications and social media platforms are attracting increasing attention (Martin et al., 2014; Carter et al., 2015; Venturelli et al., 2017). However, it remains unclear to what extent such data can produce reliable insights of use in management and monitoring of recreational fisheries.

The recent development and accessibility of video recording devices, including submergible ones, have increased the possibilities for recreational fishers to film their fishing activity and at the same time opened new possibilities for aquatic research (e.g., Bulleri and Benedetti-Cecchi, 2014; Struthers et al., 2015; Giovos et al., 2018). Once shared on social media, analysis of video footage posted by recreational fishers can potentially be a useful source of digital data, for example to compare the harvesting pattern of different types of recreational fishers (e.g., different groups of anglers or anglers and spearfishers). Furthermore, analyzing social media posting can allow comparing how users react to different forms of recreational fisheries. For example, it is possible that capture videos of the very same species of fish can trigger a different reactions and social engagement when captured by a recreational angler with hook and line or by a recreational spearfisher. Similar differences in responses may occur when different fisher types (e.g., harvest-oriented or catch-and-release types) post their captures, but scientific knowledge on this topic is not available.

Data mining refers to the extraction of information from a data set that is not readily apparent and not always easily obtainable (Barbier and Liu, 2011). Data mining from social media can play an important role in conservation science and in understanding and quantifying human-wildlife interactions (Di Minin et al., 2015; van der Wal and Arts, 2015; Ladle et al., 2016; Hausmann et al., 2018). For example, Monkman et al. (2018b) demonstrated that user knowledge retrieved by google search engines can be used as a proxy indicator of coastal utilization in marine recreational fisheries. Monkman et al. (2018c) employed text and data mining on social media to provide information on the spatio-temporal scales of species-specific marine recreational fisheries. Shiffman et al. (2017) used posts by recreational anglers on online forums to assess shark conservation issues, such as illegal catches, and Martin et al. (2014) showed that the activity on an online recreational fishing social network was related to the spatial distribution of fishing effort. These examples show that data mining of social media can provide relevant insights into the patterns and dynamics of selected recreational fisheries.

YouTube represents a worldwide dynamic cultural system where diverse groups interact in a participatory way around video footage (Burgess and Green, 2018). YouTube is also embedded into the recreational fishers' culture and appears to be one of the main web platforms used to share videos of catches or memorable fishing trips. YouTube videos have previously been used to document that recreational fisheries in the Mediterranean Sea are multispecies and exhibit spatially homogeneous patterns among different countries (Giovos et al., 2018). In another recent study, Belhabib et al. (2016) used YouTube videos to retrieve data for recreational catches in West Africa where such data are extremely scarce. It is worthwhile to conduct further studies using YouTube data because, in contrast to many other survey methods, one can also study how groups respond to the action of focal people shown on videos.

In the present study, we used YouTube to better understand recreational fisheries of the common dentex (*Dentex dentex*) in Italy. *Dentex dentex* constitutes an iconic target species for Mediterranean fisheries exploited by both recreational anglers and recreational spearfishers (Morales-Nin and Moranta, 1997; Morales-Nin et al., 2005; Font and Lloret, 2011; Marengo et al., 2014; Marengo et al., 2015). In the overview of the conservation status of the marine fishes of the Mediterranean Sea

published by the International Union for Conservation of Nature (IUCN), *D. dentex* was reported as a vulnerable species that experienced declines between 30-50% in the Mediterranean region (Malak et al., 2011). The establishment of restrictions to recreational spearfishing has been recommended as a possible solution for the recovery of the species (Malak et al., 2011; Carpenter and Russel, 2014). This recommendation implies that spearfishers harvesting of *D. dentex* is higher than recreational angling; however previous studies suggested the opposite in terms of harvesting rate and number of specimens caught (Marengo et al., 2015). In this context, the videos uploaded on YouTube can inform about differences in catch patterns between recreational anglers and recreational spearfishers. Moreover, as many recreational fishers declare the mass of the fish shown in video footage, analyzing YouTube videos from the same region provides an interesting opportunity to infer differences in the average mass of catches between groups. This could be particularly important for collecting complementary data to the ones already available on the vulnerability of large specimens of *D. dentex* to the different forms of recreational fishing (Morales-Nin et al., 2005; Marengo et al., 2015). Indeed, preserving large individuals could be one of the key for improving the conservation status of *D. dentex* (Grau et al., 2016).

In terms of social dimensions, the social engagement recorded on YouTube videos as well as the relationship of the mass of the reported catch and social engagement can provide insights on the online behavior of recreational fishers. Angling and spearfishing are extremely different fishing activities that differ in energy investment (e.g., into boats), required skills and possibly aesthetics. Recreational fisheries has a strong component of masculinity (Bull, 2009), and given the specific fishing style recreational spearfishers might be perceived differently than for example recreational angling from a motorized boat. Therefore, it is possible that the very same fish posted by either recreational anglers or recreational spearfishers induces different social engagement. As the size of the fish captured is of fundamental importance in the utility function of most recreational fishers (Arlinghaus et al., 2014) and is systematically related to fisher satisfaction (Beardmore et al., 2015), one would expect more social engagement to happen with videos showing larger sized specimens. Because spearfishers and anglers may differ in their general ability to catch large

specimens (e.g., Marengo et al., 2015), we expected that the social engagement for a given mass of fish reported in the video to be specific to the type of recreational fisheries.

Our main objective was to provide complementary knowledge on *D. dentex* recreational fisheries by using data mined from YouTube. We compare the harvesting pattern of *D. dentex* between recreational anglers and recreational spearfishers and provide first results on the potential of YouTube to characterize social aspects of recreational fisheries. Our research questions were:

(i) What is the upload pattern of *D. dentex* YouTube videos in recreational angling and recreational spearfishing and are there differences in the seasonal upload patterns among the two types of recreational fisheries?

(ii) Are there differences in the mass of *D. dentex* reported in the videos between recreational angling and recreational spearfishing?

(iii) Are there differences in the social engagement of videos between recreational angling and recreational spearfishing?

(iv) Does the mass declared in the video predict social engagement, and does this relationship differ between recreational angling and recreational spearfishing?

MATERIALS AND METHODS

Ethical aspects

Data mining on YouTube does not require specific ethical review because the data are publicly-available. However, we followed the framework presented by Monkman et al. (2018a) on the ethics of using social media in fisheries research. In particular, we anonymized all the data after performing the sorting of the videos according to their title and description. Moreover, we extended the anonymization to all possible direct and indirect identifiers of the people who uploaded the video footage (i.e., video id, date of publication, channel id, channel title, video duration).

The case study

We present a case study aiming at exploring the harvesting patterns of *D. dentex* in Italy using videos uploaded on YouTube from 2010 to 2016. We used a systematic analysis on videos to collect quantitative data on harvesting patterns and social

engagement. We automatically retrieved all the videos published concerning the species of interest and sorted them into two groups: one related to captures by recreational angling and the other one related to captures by recreational spearfishing. Spearfishing was defined as underwater fishing practiced by the exclusive use of free-diving techniques and a speargun (Sbragaglia et al., 2016; Sbragaglia et al., 2018); angling was defined as hook-and-line fishing from either the coastline or from a boat with natural baits or artificial lures.

Figure 1. Schematic representation of the data mining and filtering according to the keywords reported in Table 1. A first step is represented by the automatic identification of videos followed by a second step represented by the cross checked identification.

Data mining

We collected the data by using the YouTube Data API (v3), following a two-step process (Figure 1). First, we extracted the data from YouTube’s API in April 2018 using two keywords: the name of the species in Italian (“*Dentice*”) and the scientific name of the species (“*Dentex*”). Although mentions to vernacular and scientific names are usually highly associated in online databases (Jarić et al., 2016; Correia et al., 2017), both terms were employed to ensure all videos of potential interest were identified. We carried out searches with the Italian vernacular species name to limit data

collection to Italian recreational fishers. With regards to the scientific species name, we only used the genus name as a search term to ensure that videos citing occasionally used synonyms (e.g. *Dentex vulgaris*) were also captured (Correia et al., 2018). We compiled a raw dataset with the title and descriptions of videos along with social engagement metrics such as number of views, likes and comments.

Table 1. The keywords used in the automatic identification for recognizing whether a video was about the capture of a *D. dentex* by angling or spearfishing. Keywords are in Italian to limit the search to Italian recreational fishers and they represent the most popular techniques (e.g., “*traina*” meaning trolling), the most common techniques (e.g., “*aspetto*” meaning sit-and-wait hunting) and the most common tackle brands (e.g., *shimano*).

Angling	Spearfishing
Traina, spinning, jigging, surfcasting, calamaro, barca, canna, rockfishing, spiaggia, inchiku, shimano, kayak, strike, bolentino, scogliera, palamito	aspetto, agguato, caduta, pesca sub, apnea, spearfishing, pesca subacquea, dapiran, discesa, saber, bassofondo, seawolf, alemanni, michelangelo, tuffo, strappato, poca acqua, sea hawk, abisso, saber, roller, seawolf, gimansub, calibro

In a second step, we automatically searched the title and description of each video for specific keywords as reported in Table 1. We subdivided the keywords into two groups with the aim to sort the videos regarding recreational angling and recreational spearfishing. We choose keywords following the most popular fishing techniques, the most common gears used and the most common tackle brands. We stored the results of the second step in the automatic identified dataset that represented the output of the data mining without manual cross checking of the results.

Manual cross check of the automatic identification

Subsequently, we manually cross checked the raw dataset. We first excluded the videos that were: (i) not related to the target species; (ii) not showing the catch of the target species (i.e. catch and release or not shooting while diving underwater); (iii) not related to the target country; and (iv) repetitions of previously published videos. Then, we applied a manual cross check of the automatic classification to identify the occurrence of false negatives (i.e. target videos previously not recognized following the keywords), false positives (i.e. videos erroneously attributed to one of the two groups)

and mismatch recognitions (i.e. videos erroneously attributed to one fisher group instead of the other). Finally, during the manual cross check, we also annotated the mass of the fish reported in the title or description of the videos. We run all the analyses related to data mining in R (<https://www.r-project.org/>; version 3.5.0) with the additional package “jsonlite” (Ooms, 2014), “lubridate” (Grolemund and Wickham, 2011) and “curl” (<https://cran.r-project.org/web/packages/curl/index.html>).

Data analysis and statistical approach

We first used the cross-checked dataset to estimate the annual oscillation in the upload patterns of videos by group (question *i*) by using RAIN (Rhythmicity Analysis Incorporating Nonparametric methods). This method is a robust non-parametric method for the detection of rhythms in data that can detect arbitrary oscillations (Thaben and Westermark, 2014). Then, we used Generalized Linear Models (GLM; Nelder and Baker, 2004) to estimate differences between recreational angling and recreational spearfishing videos in the reported mass of the fish declared in the video footage (question *ii*) and in the social engagement parameters (number of views, likes and comments; question *iii*). Finally, we used GLMs to estimate the relationship between variables indexing social engagement and the reported mass of the fish (question *iv*). In order to reduce the skewness of the mass distribution between groups and have comparable regression results, we limited the comparison to reported masses between 2.5 and 10 kg. We employed a Gaussian distribution with GLM for modelling the data related to the mass of the fish, while we used a negative binomial distribution to account for overdispersion of the count data for social engagement variables (Bliss and Fisher, 1953; Gardner et al., 1995). We assessed model fit by checking the plot of the residuals vs. the fitted values. In all cases we used a 95% confidence interval. We run all analysis in R (<https://www.r-project.org/>; version 3.5.0) with the additional packages: “mass” (Ripley et al., 2013), and “rain” (Thaben and Westermark, 2014).

RESULTS

We identified a total of 21,441 videos published between 2010 and 2016. A first screening of the data revealed that 1680 videos (7.8 %) were: not related to the target

species, not related to the country of interest, repetitions of previously published videos or were videos where the fish was not captured. During the second cross check on the remaining 19,761 videos, we found that the automatic identification following the keyword in Table 1 overestimated (+18%) the number of videos related to recreational angling and underestimated (-17%) those related to recreational spearfishing (Table 2). Moreover, during the manual cross check we found that 42% of the videos showing the capture of *D. dentex* were wrongly assigned to the recreational angling group, and 38% wrongly assigned to the recreational spearfishing group. After the cross check, we correctly identified 692 videos related to recreational angling and 1051 videos related to recreational spearfishing (Table 2).

Table 2. The comparison of the results of the automatic identification and manual cross checked identification in terms of false positive, false negative and group mismatch.

Automatic identification	Checked identification	N
n.c.	n.c.	0
angling	n.c.	15
spearfishing	n.c.	13
TOTAL false positive		28
angling	angling	472
n.c.	angling	216
spearfishing	angling	4
TOTAL angling		692
spearfishing	spearfishing	540
n.c.	spearfishing	511
angling	spearfishing	0
TOTAL spearfishing		1051

Using the final cross-checked data set, we did not find significant ($p = 0.163$) annual oscillations (i.e., seasonality) in the upload pattern in the recreational angling group. In contrast, the recreational spearfishing group revealed significant ($p < 0.001$) seasonality with a peak of uploads of videos in July (Fig. 2).

Figure 2. The upload monthly patterns related to angling (red) and spearfishing (blue). The total number of identified videos is 692 for angling and 1051 for spearfishing.

During the cross check, we annotated the reported mass of the fish from either the title or description of 540 videos (214 for recreational angling and 326 for recreational spearfishing). The fish were significantly ($F_{1,538} = 76.64$; $p < 0.001$) lighter in recreational spearfishing videos compared to angling ones (-1.93 [$-2.36 - -1.49$] kg; Fig. 3). The extrapolation of the average mass to all the videos identified here indicated that recreational anglers and recreational spearfishers harvested similar biomass, 4.45 and 4.73 t, respectively. This is of course an absolute conservative estimate and only relates to the sample size of the video that we analysed.

Figure 3. Violin plots of the mass of the fish declared in the title and description of the videos for both angling (red) and spearfishing (blue). Numbers inside the violin plots indicate the mean mass. The total number of identified videos with the mass of the fish is 214 for angling and 326 for spearfishing. The black horizontal line represent significant differences between groups (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$).

Regarding social engagement, we did not find significant differences between recreational angling and recreational spearfishing in terms of views ($p = 0.200$; Table 3, Fig. 4). However, the number of likes was significantly greater in recreational spearfishing videos compared to videos posted about recreational angling of common dentex (Rate Ratio, RR = 2.15 [1.85 – 2.49]; $p < 0.001$; Table 3. Fig. 4). The same significant difference was found in relation to the number of comments (RR = 5.34 [4.55 – 6.25]; $p < 0.001$; Table 3, Fig. 4).

Table 3. Modelling results related to the differences in social engagement parameters between spearfishing and angling groups. Estimates (rate ratio) and confidence intervals ([2.5% – 97.5%]) are reported together with Z statistics and p values for spearfishing group with respect to angling one.

Variable	Group	Estimate	Z value	p value
Views	Spearfishing	1.10 [0.95 – 1.28]	1.283	0.200
Likes	Spearfishing	2.15 [1.85 – 2.49]	10.050	<0.001
Comments	Spearfishing	5.34 [4.55 – 6.25]	20.740	<0.001

Figure 4. Barplots of the average social engagement parameters for both angling (red) and spearfishing (blue). The total number of identified videos is 692 for angling and 1051 for spearfishing. Black lines represent significant differences between groups (*: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$).

Furthermore, the mass of the fish declared in the videos significantly ($p < 0.001$; Table 4) and positively predicted all social engagement variables (comments, views, likes) in recreational spearfishing; larger fish increased the degree of social engagement (views: $RR = 1.92 [1.73 - 2.12]$; likes: $RR = 1.57 [1.45 - 1.70]$; comments: $RR = 1.23 [1.13 - 1.35]$; Table 4, Fig. 5). The mass of the fish had a weak, yet significant impact on the number of views ($p < 0.001$; Table 4) and comments ($p < 0.05$; Table 4) for recreational angling videos; the lighter the fish the larger was the social engagement (views: $RR = 0.84 [0.74 - 0.95]$; comments: $RR = 0.85 [0.74 - 0.97]$; Table 4 and Fig. 5). Finally, all the relationships between the mass of the fish and social engagement variables obtained for videos of recreational spearfishing were significantly different (i.e. no overlap of confidence intervals) from the ones revealed for recreational angling (Table 4).

Table 4. Modelling results related to the correlation between mass of the fish declared in the video and social engagement parameters. Estimates (rate ratio) and confidence intervals ([2.5% – 97.5%]) are reported together with Z statistics and p values for both recreational angling and recreational spearfishing.

Variable	Group	Estimate	Z value	P value
Views:mass	Angling	0.84 [0.74 – 0.95]	-3.09	< 0.001
	Spearfishing	1.92 [1.73 – 2.12]	14.30	< 0.001
Likes:mass	Angling	0.90 [0.80 – 1.01]	-1.90	0.057
	Spearfishing	1.57 [1.45 – 1.70]	11.27	< 0.001
Comments:mass	Angling	0.85 [0.74 – 0.97]	-2.40	< 0.05
	Spearfishing	1.23 [1.13 – 1.35]	4.86	< 0.001

Figure 5. Scatter plots of the mass of the fish against social engagement parameters for both angling (red: N = 184) and spearfishing (blue; N = 234) groups. Lines represent a negative binomial smoothing function together with 95% confidence interval (grey area).

DISCUSSION

Our study demonstrated that data mining on YouTube is an effective method for providing complementary data on recreational fisheries and offers potential for understanding specific aspects of the social media engagement present in different types of recreational fisheries. We showed that during the monitoring period (2010-2016) recreational spearfishers and recreational anglers harvested similar *D. dentex* biomasses. Clearly this is an absolute conservative estimate only related to YouTube videos and it is not suitable to directly compare the two sectors in terms of their impacts on the stock. We also showed a peak of video posting in July only for spearfishing group, suggesting that the odds of catching *D. dentex* shows seasonality in recreational spearfishers, but not in anglers. Moreover, recreational spearfishing videos registered more social engagement (i.e., likes and comments) compared to recreational angling videos, supporting the idea that spearfishing videos enhanced social rewards, possibly because recreational spearfishing necessitates elevated

personal investments and skill for catching *D. dentex* compared to anglers. Finally, the mass of the fish positively predicted social engagement only in recreational spearfishing probably because they have less chances of catching large *D. dentex* than recreational anglers (Marengo et al., 2015), and consequently videos with bigger fish triggered more social engagement.

Our results represented only those recreational fishers that shared the video footage of their catches on YouTube and hence provided a measure of the minimum harvested biomass of *D. dentex* by recreational anglers and recreational spearfishing, respectively. The total harvest of *D. dentex* in Italy by both commercial and recreational fisheries during the study period (2010-2016) was at least 1254 t (FAO, 2012b). We estimated that the recreational anglers and recreational spearfishers posting videos on YouTube harvested a total biomass of 4.45 and 4.73 t, respectively. This means that the minimum recreational fishery harvest of *D. dentex* in Italy represented 0.73 % of the total harvested biomass. Because we lack an estimate of the total number of recreational fishers of both types, we are unable to scale our data to the total population of anglers/spearfishers. We also do not know the degree to which people posting videos are different from the average recreational fisher. However, Marengo et al. (2015) showed that recreational fishers can harvest up to 37% of the total biomass (artisanal and recreational fisheries) of *D. dentex* in the Bonifacio Strait (NW Mediterranean Sea). Based solely on the analysis of YouTube video footage, our results suggest that recreational anglers and recreational spearfishers harvest similar biomasses of *D. dentex*. This is in accordance with what has previously been reported in the Balearic archipelago where *D. dentex* is regularly caught by both boat anglers and spearfishers (Morales-Nin et al., 2005). By contrast, Marengo et al. (2015) showed that in the Bonifacio Strait the harvesting rate of angling (i.e., trolling; $359 \text{ g boat}^{-1} \text{ h}^{-1}$) was higher than spearfishing one ($68 \text{ g boat}^{-1} \text{ h}^{-1}$) and similarly the number of specimens caught (recreational angling: 94%; recreational spearfishing: 6% of catches; Marengo et al., 2015). Collectively, our results and data from other sources indicate that spearfishing cannot be the only form of recreational fishing responsible for the decline of *D. dentex* population in the Mediterranean Sea, in contrast to suggestions expressed previously (Malak et al., 2011; Carpenter and Russel, 2014).

The pattern of videos upload indicated seasonality only for recreational spearfishing with a peak of uploads in the month of July. There are several reasons that could explain these results. A first explanation - assuming that the videos are uploaded within few days after the fishing trip - could be related to a seasonal change of vulnerability of *D. dentex* to spearfishing with a peak of vulnerability in the month of July. Seasonal movements of *D. dentex* have been recently documented. Indeed, Aspillaga et al. (2017) demonstrated that *D. dentex* prefers the warm suprathermocline layer whose shallowest depths (at about 20-30 m) are usually reached in July and August. Furthermore, (Sbragaglia et al., *in press*) documented that *D. dentex* presence at an artificial reef at 22 m depth was higher during summer months and positively correlated with the water temperature. It is plausible that anglers can pursue *D. dentex* at all depths throughout the year; by contrast, spearfishers usually perform their freediving hunting actions at maximum depths of about 18-25 m (FIPSAS, 2002). Consequently, spearfishers have more chances of encountering *D. dentex* during summer months (e.g., July) than in other months, in line with the upload patterns we found. A second interpretation could be related to seasonality in the fishing activity of spearfishers. However, this explanation is not supported by previous results in the Mediterranean area that indicated angling has a more marked seasonality in fishing activity than spearfishing (Morales-Nin et al., 2005). Our results could also be related to other factors not related to fishing (e.g., seasonal activity on social media).

The *D. dentex* harvested in the videos posted by recreational anglers were significantly heavier than those harvested by recreational spearfishers. This is in accordance with the qualitative data reported by Marengo et al. (2015) for the Bonifacio strait showing that recreational angling (in particular trolling) harvested larger specimens relative to spearfishing. Moreover, the average mass reported in our study from YouTube for *D. dentex* harvested by spearfishers (4.50 kg) was larger than the average harvested mass of individual common dentex reported during an onsite survey in Cap de Creus, North-West Mediterranean Sea (1.77 kg; Lloret et al., 2008), suggesting that the average mass estimated here for spearfishing could be overestimated in YouTube videos compared to the true average in real fisheries. A possible explanation is that the upload of spearfishing videos may be biased toward

particularly memorable and hence large fish. Alternatively, spearfishers may inflate the mass of the fish declared in the videos for reasons of increasing prestige, possibly to generate more social engagement (see below). Our study, together with previous studies (Marengo et al., 2015), suggest that spearfishers have lower chances than anglers of capturing large *D. dentex*.

YouTube users left more likes and comments to spearfishing videos than to angling ones. Although we are unable to characterize what kind of public engaged with the videos (e.g., fishers or members of the non-fishing public), it is conceivable to assume that most of the social engagement with videos was triggered by either recreational spearfishers or recreational anglers. If this is the case, our results suggest that spearfishers either have a larger online community from which to draw social engagement or spearfishing videos truly trigger more online engagement than angling. The latter could be related to the fact that catching a *D. dentex* with spearfishing techniques requires more skills (e.g., long free diving usually at considerable depths) and personal investment of energy relative to angling, such that the posting of particularly remarkable individuals (large) stimulates increased feelings of excitement and arousal, in turn motivating more likes, views and comments. The intrinsic difficulty of catching *D. dentex* by spearfishing is reflected by the extremely low species abundance during spearfishing competitions (0.16-0.50 % of total number of fish; Coll et al., 2004). In fact, the first capture of a *D. dentex* is considered as the final graduation among spearfishers (FIPSAS, 2002). The differences observed in the number of likes and comments may also be related to a stronger masculine moral identity of spearfishers related to providing food for their families by challenging the adverse conditions of underwater hunting, as described for another hand-fishing technique with strong masculine traits targeting catfish (e.g., *Pylodictis olivaris*) in rivers in the USA (Grigsby, 2009; Grigsby, 2011). Finally, experiencing nature and mastering challenging situations constitute among the main motivations of recreational anglers and recreational spearfishers (Knopf et al., 1973; Holland and Ditton, 1992; Fedler and Ditton, 1994; Beardmore et al., 2011; Young et al., 2016). Spearfishing is a peculiar fishing modality that is more similar to hunting than fishing; it implies a particularly direct contact with the prey in its own habitat, and it has been reported that spearfishers reveal higher levels of satisfaction than anglers (Gordoa et

al., 2019). Therefore, spearfishing videos could be more effective in triggering enhanced appreciation among spearfishers and consequently they may receive more likes and comments than when the same-sized fish is posted by anglers, possibly captured from highly motorized boats.

The capture of large *D. dentex* triggered a particularly positive increase of all social engagement variables, but only in the spearfishing group. *Dentex dentex* may reach a maximum mass above 13 kg (Marengo et al., 2014), but fishes of 7 kg or larger may be already considered a unique trophy size by recreational fishers. As documented above, spearfishers could have lower chances of catching a trophy *D. dentex*; in particular, the percentage of the videos where the declared mass of the fish was above 7 kg represent 34% of the videos posted by recreational anglers and 14% of those posted by recreational spearfishers. A study conducted with online terrestrial hunting photographs similarly revealed that prey size influenced emotional signals (i.e., smiles) of hunters; true smiles were 1.5 times greater when posing with large prey compared with small prey (Child and Darimont, 2015). This can be related to an activity-specific achievement satisfaction of hunters for harvesting a large and rare prey (Potter et al., 1973; Child and Darimont, 2015). Our results regarding spearfishers are in agreement with this view, but relate to social appreciation, rather than individual satisfaction as shown by the fisher himself.

Although data mining produced interesting complementary information for characterizing recreational fisheries, there are several aspects that need to be addressed in the future before the technique can be more readily used. First, there is a need for a systematic study as to how recreational fishers posting on YouTube differ from the representative population of recreational fishers. Once these data are available, one could weigh YouTube data and infer insights about the population of users from YouTube data. Second, we do not know what kind of YouTube users contributed to the social engagement variables, and a proper identification is needed. Third, the automatic classification we applied was only based on sorting the videos of different forms of recreational fisheries according to specific keywords, which likely contributed to the large disparity of the automatic screening and the manual cross check; this is probably due to users tagging a broad range of keywords to attract more views. Future methodological refinements for implementing the sensitivity of

automatic sorting should also consider machine learning approaches (e.g., Roll et al., 2018). Finally, an interesting follow-up analysis could be related to the qualitative analysis of the content of the commentaries left after watching the videos. Such analysis could provide insights into more specific social aspects of recreational fishers, such as the exact type of social engagement, the content of the comments and the degree of user satisfaction with the footage (e.g., Ebeling-Schuld and Darimont, 2017). Moreover, one could specifically analyse discourses surrounding controversial footage to better understand what reactions are induced by specific technicalities of the hunt (e.g., whether a fish was captured from a motorized boat or from the shore, by which type of fisher etc.).

Our study opens a new avenue in using digital applications and social media platforms for producing valuable complementary data for characterizing catch patterns and social aspects of different forms of recreational fisheries. The approach we present is not overly time-consuming and can be implemented with limited resources. This will be particularly important for marine recreational fisheries in Europe where there are still significant gaps in knowledge (Hyder et al., 2018), and for developing countries where there are not many resources available for the monitoring of recreational fisheries. Data mining on YouTube could contribute to characterize multispecies harvest patterns at national level, contribute to understanding catch allocation and help understanding the human dimensions of recreational fishers (Arlinghaus et al., 2013; Hunt et al., 2013).

Authors' contribution

VS conceived the study; RC performed the data mining with inputs by VS; SC and VS analysed the data; VS statistically analysed the data, VS and RA interpreted the data, VS led the writing of the manuscript with inputs by all other co-authors. All authors gave final approval for publication.

Acknowledgments

VS was supported by a Leibniz-DAAD postdoctoral research fellowship (# 91632699). RAC is currently funded by a post-doctoral scholarship from FCT – Fundação para a Ciência e Tecnologia (SFRH/BPD/118635/2016).

REFERENCES

- Arlinghaus, R., Beardmore, B., Riepe, C., Meyerhoff, J., and Pagel, T. 2014. Species-specific preferences of German recreational anglers for freshwater fishing experiences, with emphasis on the intrinsic utilities of fish stocking and wild fishes. *Journal of Fish Biology*, 85: 1843-1867.
- Arlinghaus, R., Cooke, S. J., and Potts, W. 2013. Towards resilient recreational fisheries on a global scale through improved understanding of fish and fisher behaviour. *Fisheries Management and Ecology*, 20: 91-98.
- Arlinghaus, R., Cooke, S. J., Sutton, S. G., Danylchuk, A. J., Potts, W., Freire, K. d. M. F., Alós, J., et al. 2016. Recommendations for the future of recreational fisheries to prepare the social-ecological system to cope with change. *Fisheries Management and Ecology*, 23: 177-186.
- Aspillaga, E., Bartumeus, F., Starr, R. M., López-Sanz, À., Linares, C., Díaz, D., Garrabou, J., et al. 2017. Thermal stratification drives movement of a coastal apex predator. *Scientific Reports*, 7: 526.
- Barbier, G., and Liu, H. 2011. Data mining in social media. *In Social network data analytics*, pp. 327-352. Springer.
- Beardmore, B., Haider, W., Hunt, L. M., and Arlinghaus, R. 2011. The importance of trip context for determining primary angler motivations: Are more specialized anglers more catch-oriented than previously believed? *North American Journal of Fisheries Management*, 31: 861-879.
- Beardmore, B., Hunt, L. M., Haider, W., Dorow, M., and Arlinghaus, R. 2015. Effectively managing angler satisfaction in recreational fisheries requires understanding the fish species and the anglers. *Canadian Journal of Fisheries and Aquatic Sciences*, 72: 500-513.
- Belhabib, D., Campredon, P., Lazar, N., Sumaila, U. R., Baye, B. C., Kane, E. A., and Pauly, D. 2016. Best for pleasure, not for business: evaluating recreational marine fisheries in West Africa using unconventional sources of data. *Palgrave Communications*, 2: 15050.
- Bellanger, M., and Levrel, H. 2017. A cost-effectiveness analysis of alternative survey methods used for the monitoring of marine recreational fishing in France. *Ocean & Coastal Management*, 138: 19-28.
- Bliss, C. I., and Fisher, R. A. 1953. Fitting the negative binomial distribution to biological data. *Biometrics*, 9: 176-200.
- Board, O. S., and Council, N. R. 2006. Review of recreational fisheries survey methods, National Academies Press.
- Brownscombe, J. W., Hyder, K., Potts, W., Wilson, K. L., Pope, K. L., Danylchuk, A. J., Cooke, S. J., et al. 2019. The future of recreational fisheries: Advances in science, monitoring, management, and practice. *Fisheries Research*, 211: 247-255.
- Bull, J. 2009. Watery masculinities: fly-fishing and the angling male in the South West of England. *Gender, place and culture*, 16: 445-465.
- Bulleri, F., and Benedetti-Cecchi, L. 2014. Chasing fish and catching data: recreational spearfishing videos as a tool for assessing the structure of fish assemblages on shallow rocky reefs. *Marine Ecology Progress Series*, 506: 255-265.
- Burgess, J., and Green, J. 2018. YouTube: Online video and participatory culture, John Wiley & Sons.

- Carpenter, K. E., and Russel, B. 2014. *Dentex dentex*. The IUCN red list of threatened species.
- Carter, D. W., Crosson, S., and Liese, C. 2015. Nowcasting intraseasonal recreational fishing harvest with internet search volume. *PloS one*, 10: e0137752.
- Coll, J., Linde, M., García-Rubies, A., Riera, F., and Grau, A. M. 2004. Spear fishing in the Balearic Islands (west central Mediterranean): species affected and catch evolution during the period 1975–2001. *Fisheries Research*, 70: 97-111.
- Cooke, S. J., and Cowx, I. G. 2004. The role of recreational fishing in global fish crises. *AIBS Bulletin*, 54: 857-859.
- Correia, R. A., Jarić, I., Jepson, P., Malhado, A. C., Alves, J. A., and Ladle, R. J. 2018. Nomenclature instability in species culturomic assessments: Why synonyms matter. *Ecological Indicators*, 90: 74-78.
- Correia, R. A., Jepson, P., Malhado, A. C., and Ladle, R. J. 2017. Internet scientific name frequency as an indicator of cultural salience of biodiversity. *Ecological Indicators*, 78: 549-555.
- Child, K. R., and Darimont, C. T. 2015. Hunting for trophies: online hunting photographs reveal achievement satisfaction with large and dangerous prey. *Human Dimensions of Wildlife*, 20: 531-541.
- Di Minin, E., Tenkanen, H., and Toivonen, T. 2015. Prospects and challenges for social media data in conservation science. *Frontiers in Environmental Science*, 3: 63.
- Ebeling-Schuld, A. M., and Darimont, C. T. 2017. Online hunting forums identify achievement as prominent among multiple satisfactions. *Wildlife Society Bulletin*, 41: 523-529.
- Elmer, L. K., Kelly, L. A., Rivest, S., Steell, S. C., Twardek, W. M., Danylchuk, A. J., Arlinghaus, R., et al. 2017. Angling into the future: Ten commandments for recreational fisheries science, management, and stewardship in a good Anthropocene. *Environmental Management*, 60: 165-175.
- FAO 2012a. FAO Technical Guidelines for Responsible Fisheries.
- FAO 2012b. FISHSTAT J: FAO fishery and aquaculture global statistics. FAO fisheries department, fishery information, data and statistics unit.
- Fedler, A. J., and Ditton, R. B. 1994. Understanding angler motivations in fisheries management. *Fisheries*, 19: 6-13.
- FIPSAS 2002. Libro bianco; Le gare di pesca in apnea. Comitato di Settore Attività Subacquee.
- Font, T., and Lloret, J. 2011. Biological implications of recreational shore angling and harvest in a marine reserve: the case of Cape Creus. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 21: 210-217.
- Gardner, W., Mulvey, E. P., and Shaw, E. C. 1995. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological bulletin*, 118: 392.
- Giovos, I., Keramidas, I., Antoniou, C., Deidun, A., Font, T., Kleitou, P., Lloret, J., et al. 2018. Identifying recreational fisheries in the Mediterranean Sea through social media. *Fisheries Management and Ecology*, 25: 287-295.
- Gordoa, A., Dedeu, A. L., and Boada, J. 2019. Recreational fishing in Spain: First national estimates of fisher population size, fishing activity and fisher social profile. *Fisheries Research*, 211: 1-12.

- Grau, A., Saborido-Rey, F., Pastor, E., Palmer, M., Massuti-Pascual, E., Quetglas, A., Riera, I., et al. 2016. Reproductive strategy of common dentex *Dentex dentex*: management implications. *Mediterranean marine science*, 17: 552-566.
- Grigsby, M. 2009. Masculine identity work among Missouri noodlers: community providers, pleasure seeking comrades and family men, and tough courageous men. *Southern Rural Sociology*, 24.
- Grigsby, M. 2011. Subcultural masculine moral identity work among rural Missouri noodlers: "a special breed of men". *Anthropologica*: 159-171.
- Grolemund, G., and Wickham, H. 2011. Dates and times made easy with lubridate. *Journal of Statistical Software*, 40: 1-25.
- Hartill, B. W., Cryer, M., Lyle, J. M., Rees, E. B., Ryan, K. L., Steffe, A. S., Taylor, S. M., et al. 2012. Scale-and context-dependent selection of recreational harvest estimation methods: the Australasian experience. *North American Journal of Fisheries Management*, 32: 109-123.
- Hartill, B. W., and Edwards, C. T. T. 2015. Comparison of recreational harvest estimates provided by onsite and offsite surveys: detecting bias and corroborating estimates. *Canadian Journal of Fisheries and Aquatic Sciences*, 72: 1379-1389.
- Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., and Di Minin, E. 2018. Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. *Conservation Letters*, 11: e12343.
- Holdsworth, J. C., Hartill, B. W., Heinemann, A., and Wynne-Jones, J. 2018. Integrated survey methods to estimate harvest by marine recreational fishers in New Zealand. *Fisheries Research*, 204: 424-432.
- Holland, S. M., and Ditton, R. B. 1992. Fishing trip satisfaction: a typology of anglers. *North American Journal of Fisheries Management*, 12: 28-33.
- Hunt, L., Sutton, S., and Arlinghaus, R. 2013. Illustrating the critical role of human dimensions research for understanding and managing recreational fisheries within a social-ecological system framework. *Fisheries Management and Ecology*, 20: 111-124.
- Hyder, K., Weltersbach, M. S., Armstrong, M., Ferter, K., Townhill, B., Ahvonen, A., Arlinghaus, R., et al. 2018. Recreational sea fishing in Europe in a global context-Participation rates, fishing effort, expenditure, and implications for monitoring and assessment. *Fish and Fisheries*, 19: 225-243.
- Jarić, I., Courchamp, F., Gessner, J., and Roberts, D. L. 2016. Data mining in conservation research using Latin and vernacular species names. *PeerJ*, 4: e2202.
- Knopf, R. C., Driver, B., and Bassett, J. R. 1973. Motivations for fishing. *In Transactions of the North American Wildlife and Natural Resources Conference*, pp. 191-204.
- Ladle, R. J., Correia, R. A., Do, Y., Joo, G. J., Malhado, A. C., Proulx, R., Roberge, J. M., et al. 2016. Conservation culturomics. *Frontiers in Ecology and the Environment*, 14: 269-275.
- Lloret, J., Zaragoza, N., Caballero, D., Font, T., Casadevall, M., and Riera, V. 2008. Spearfishing pressure on fish communities in rocky coastal habitats in a Mediterranean marine protected area. *Fisheries Research*, 94: 84-91.

- Malak, D. A., Livingstone, S. R., Pollard, D., Polidoro, B. A., Cuttelod, A., Bariche, M., Bilecenoglu, M., et al. 2011. Overview of the conservation status of the marine fishes of the Mediterranean Sea. 61 pp.
- Marengo, M., Culioli, J.-M., Santoni, M.-C., Marchand, B., and Durieux, E. D. H. 2015. Comparative analysis of artisanal and recreational fisheries for *Dentex dentex* in a marine protected area. *Fisheries Management and Ecology*, 22: 249-260.
- Marengo, M., Durieux, E. D. H., Marchand, B., and Francour, P. 2014. A review of biology, fisheries and population structure of *Dentex dentex* (Sparidae). *Reviews in Fish Biology and Fisheries*, 24: 1065-1088.
- Martin, D. R., Chizinski, C. J., Eskridge, K. M., and Pope, K. L. 2014. Using posts to an online social network to assess fishing effort. *Fisheries Research*, 157: 24-27.
- Monkman, G. G., Kaiser, M., and Hyder, K. 2018a. The ethics of using social media in fisheries research. *Reviews in Fisheries Science & Aquaculture*, 26: 235-242.
- Monkman, G. G., Kaiser, M. J., and Hyder, K. 2018b. Heterogeneous public and local knowledge provides a qualitative indicator of coastal use by marine recreational fishers. *Journal of Environmental Management*, 228: 495-505.
- Monkman, G. G., Kaiser, M. J., and Hyder, K. 2018c. Text and data mining of social media to map wildlife recreation activity. *Biological Conservation*, 228: 89-99.
- Morales-Nin, B., and Moranta, J. 1997. Life history and fishery of the common dentex (*Dentex dentex*) in Mallorca (Balearic Islands, western Mediterranean). *Fisheries Research*, 30: 67-76.
- Morales-Nin, B., Moranta, J., García, C., Tugores, M. P., Grau, A. M., Riera, F., and Cerdà, M. 2005. The recreational fishery off Majorca Island (western Mediterranean): some implications for coastal resource management. *ICES Journal of Marine Science*, 62: 727-739.
- Nelder, J. A., and Baker, R. J. 2004. Generalized linear models. *Encyclopedia of statistical sciences*, 4.
- Ooms, J. 2014. The jsonlite package: A practical and consistent mapping between json data and r objects. *arXiv preprint arXiv:1403.2805*.
- Pollock, K. H., Jones, C. M., and T.L., B. 1994. Angler survey methods and their applications in fisheries management. *American Fisheries Society Special Publication*, 25.
- Potter, D. R., Hendee, J. C., and Clark, R. N. 1973. Hunting satisfaction: game, guns, or nature. *In Transactions of the North American Wildlife and Natural Resources Conference*, pp. 220-229.
- Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., and Ripley, M. B. 2013. Package 'mass'. CRAN Repos. [Httpcran R-Proj](http://cran.r-project.org).
- Rocklin, D., Levrel, H., Drogou, M., Herfaut, J., and Véron, G. 2014. Combining telephone surveys and fishing catches self-report: The French sea bass recreational fishery assessment. *PloS one*, 9: e87271.
- Roll, U., Correia, R. A., and Berger-Tal, O. 2018. Using machine learning to disentangle homonyms in large text corpora. *Conservation Biology*, 32: 716-724.
- Rudd, M. A., Folmer, H., and van Kooten, G. C. 2002. Economic evaluation of recreational fishery policies. *Recreational Fisheries: Ecological, Economic and Social Evaluation*, 3: 34-52.
- Sbragaglia, V., Morroni, L., Bramanti, L., Weitzmann, B., Arlinghaus, R., and Azzurro, E. 2018. Spearfishing modulates flight initiation distance of fishes: the effects of

- protection, individual size, and bearing a speargun. *ICES Journal of Marine Science*, 75: 1779-1789.
- Sbragaglia, V., Nuñez, J. D., Dominoni, D., Coco, S., Fanelli, E., Azzurro, E., Marini, S., et al. *in press*. Annual rhythms of temporal niche partitioning in the Sparidae family are correlated to different environmental variables. *Scientific Reports*.
- Sbragaglia, V., Pla, O. S., Gordo, A., Hernández, S. P., Elias, L. C., Pulido, M., Ruiz, R. T., et al. 2016. The Barcelona agreement: a manifesto towards the spearfishing of the future. *Scientia Marina*, 80: 423-426.
- Shiffman, D. S., Macdonald, C., Ganz, H. Y., and Hammerschlag, N. 2017. Fishing practices and representations of shark conservation issues among users of a land-based shark angling online forum. *Fisheries Research*, 196: 13-26.
- Struthers, D. P., Danylchuk, A. J., Wilson, A. D. M., and Cooke, S. J. 2015. Action cameras: bringing aquatic and fisheries research into view. *Fisheries*, 40: 502-512.
- Thaben, P. F., and Westermark, P. O. 2014. Detecting rhythms in time series with RAIN. *Journal of Biological Rhythms*, 29: 391-400.
- van der Wal, R., and Arts, K. 2015. Digital conservation: An introduction. *Ambio*, 44: 517-521.
- van Poorten, B. T., and Brydle, S. 2018. Estimating fishing effort from remote traffic counters: Opportunities and challenges. *Fisheries Research*, 204: 231-238.
- Venturelli, P. A., Hyder, K., and Skov, C. 2017. Angler apps as a source of recreational fisheries data: opportunities, challenges and proposed standards. *Fish and Fisheries*, 18: 578-595.
- World Bank 2012. Hidden harvest: The global contribution of capture fisheries. Worldbank; <https://openknowledge.worldbank.org/handle/10986/11873>
License: CC BY 3.0 IGO, Washington, DC.
- Young, M. A. L., Foale, S., and Bellwood, D. R. 2016. Why do fishers fish? A cross-cultural examination of the motivations for fishing. *Marine Policy*, 66: 114-123.
- Zarauz, L., Ruiz, J., Urtizberea, A., Andonegi, E., Mugerza, E., and Artetxe, I. 2015. Comparing different survey methods to estimate European sea bass recreational catches in the Basque Country. *ICES Journal of Marine Science*, 72: 1181-1191.