A social media-based framework for quantifying temporal changes to wildlife viewing intensity

Kostas Papafitsoros [a,b,*], Lukáš Adam [c], Gail Schofield [d]

[a] School of Mathematical Sciences, Queen Mary University of London, London E14NS, UK
[b] ARCHELON, the Sea Turtle Protection Society of Greece, Solomou 57, 10432 Athens, Greece
[c] Artificial Intelligence Center, Czech Technical University in Prague, Karlovo Namesti 13, 120 00 Praha 2, Czech Republic
[d] School of Biological and Behavioural Sciences, Queen Mary University of London, London E14NS, UK

ARTICLE INFO

ABSTRACT

Documenting how human pressure on wildlife changes over time is important to minimise potential adverse effects through implementing appropriate management and policy actions; however, obtaining objective measures of these changes and their potential impacts is often logistically challenging, particularly in the natural environment. Here, we developed a modular stochastic model that infers the ratio of actual viewing pressure on wildlife in consecutive time periods (years) using social media, as this medium is widespread and easily accessible. Pressure was calculated from the number of times individual animals appeared in social media in pre-defined time windows, accounting for time-dependent variables that influence them (e.g. number of people with access to social media). Formulas for the confidence intervals of viewing pressure ratios were rigorously developed and validated, and corresponding uncertainty was quantified. We applied the developed framework to calculate changes to wildlife viewing pressure on loggerhead sea turtles (*Caretta caretta*) at Zakynthos island (Greece) before and during the COVID-19 pandemic (2019–2021) based on 2646 social media entries. Our model ensured temporal comparability across years of social media data grouped in time window sizes, by correcting for the interannual increase of social media use. Optimal sizes for these windows were delineated, reducing uncertainty while maintaining high time-scale resolution. The optimal time window was around 7-days during the peak tourist season when more data were available in all three years, and >15 days during the low season. In contrast, raw social media data exhibited clear bias when quantifying changes to viewing pressure, with unknown uncertainty. The framework developed here allows widely-available social media data to be used objectively when quantifying temporal changes to wildlife viewing pressure. Its modularity allowed viewing pressure to be quantified for all data combined, or subsets of data (different groups, situations or locations), and could be applied to any site supporting wildlife exposed to tourism.

1. Introduction

As awareness of the ecological, economical, and intrinsic value of wildlife has grown in recent decades (König et al., 2020), demand to observe animals in their natural environment and consequently wildlife viewing has also risen (Moorhouse et al., 2015). To address any potential adverse effects that this activity has on wild animals (i.e., elevated risk of injury from boat collisions during viewing or behavioural changes), documentation of its impacts are widespread in both terrestrial and marine environments (Burgin and Hardiman, 2015; Larson et al., 2016; König et al., 2020). However, wildlife viewing pressure is not consistent within or across years (Moorhouse et al., 2015), particularly in areas supporting seasonal migrants, such as breeding

humpback whales and sea turtles. Even within the same population, animals are not randomly distributed in time or space, with certain individuals (in hotspots or residents) being disproportionately targeted (Semeniuk et al., 2009; Christiansen and Lusseau, 2014). Another source of temporal fluctuation in viewing pressure is due to changes on the numbers of actual human observers. This issue which has been exemplified by the recent COVID-19 pandemic via the unprecedented absence of ecotourism in hotspots globally (Rutz et al., 2020; Bates et al., 2021; March et al., 2021; Schofield et al., 2021). As a result, analytical approaches that capture temporal variation in both animal presence and viewing records are required to quantify viewing impacts on at multiple scales, from individuals to groups and populations

(Moorhouse et al., 2015; Birk et al., 2020; Western et al., 2020; Marion et al., 2020). By understanding how these pressures change within and across years, relevant actions that promote conservation efforts and mitigate disturbance could be implemented.

The rise of social media, image sharing and the widespread use of mobile phone cameras has generated an extra level of pressure on wildlife, through huge demand for "selfies" and "close-up" images with animals, resulting in close encounters between humans and wild animals that have potential negative impacts (Semeniuk et al., 2009; Christiansen and Lusseau, 2014; Lenzi et al., 2020; Papafitsoros et al., 2021; Molyneaux et al., 2021; Van Hamme et al., 2021). Yet, because viewing wild animals is intrinsically linked with taking photographs and videos, social media is also being explored as a useful tool to inform conservation science (Dickinson et al., 2012; Di Minin et al., 2015; Toivonen et al., 2019). In particular, wildlife imagery initially uploaded online for purposes other than facilitating conservation studies (e.g. to social media to share personal experiences) is being increasingly applied for science under various names such as *conservation culturomics* (Ladle et al., 2016), *passive crowdsourcing* (Ghermandi and Sinclair, 2019), *iEcology* (Jarić et al., 2020), and *passive citizen science* (Edwards et al., 2021). Such data are being used to quantitatively evaluate interactions between humans and wildlife in both terrestrial (e.g. mountain gorillas (*Gorilla beringei*) Van Hamme et al., 2021, orangutans (*Pongo abelii*) Molyneaux et al., 2021) and marine environments (e.g. sea turtles (*Caretta caretta*) Papafitsoros et al., 2021, monk seals (*Neomonachus schauinslandi*) Sullivan et al., 2019).

This application of social media has several advantages over typical researcher-based approaches. For instance, to measure how ecotourism places pressure on wildlife, simple counts of tourist numbers frequenting focal sites (or subset of tourists participating on organised activities) do not necessarily reflect the actual (true) viewing pressure that animals (individuals or groups) are subjected to within a larger population, as operator strategies and animal behaviour and movement patterns also have an effect (Semeniuk et al., 2009; Christiansen and Lusseau, 2014). This issue becomes even more challenging when considering nonorganised and nonregulated activities where encounters are often incidental (Papafitsoros et al., 2021). In comparison, counting the number of times that animals appear on social media in given time frames could be used to quantify actual viewing pressure on animals more objectively, particularly when exploring how pressure changes over time (seasons/years).

However, systematic approaches to make such comparisons over time based on social media are lacking (Barros et al., 2019; Rice and Pan, 2021). Such approaches are needed, because the frequency at which animals are recorded is influenced by the location and number of people accessing social media platforms, plus the availability of cameras and smartphones (Toivonen et al., 2019; Ghermandi and Sinclair, 2019). For instance, the number of Instagram (a popular sharing social media platform) users globally has risen from around two million in 2010 to more than one billion in 2020 (https://www.statista.com/statistics/183585/instagram-number-of-global-users/). Consequently, any change in pressure based on comparing the number of animal images uploaded to this platform must be corrected to account for this increase. Challenges also exist because the flow of information from a given human–animal interaction to it appearing in social media is governed by many factors (Tenkanen et al., 2017; Jarić et al., 2020; Edwards et al., 2021). Not all observed animals are captured on camera, nor do all those captured on camera actually appear on social media (Tenkanen et al., 2017). Ultimately, the number of human–animal encounters appearing on social media tend to be several orders of magnitude lower than the actual number originally observed (Wood et al., 2013; Papafitsoros et al., 2021). Even when social media data can be compared temporally (i.e. records are consistent), certainty varies with the number of data samples (observations), with confidence being higher when sightings are higher (consistency of an estimator, Lehmann and Casella, 2006). Thus, it is important to model the uncertainty that exists in this flow of information, and establish the minimum number of social media records required to ensure interpretations have high confidence. Such models are currently absent from the literature for social media, but parallels exist for in other fields (Huang et al., 2017).

Here, we developed a rigorous and mathematically consistent framework to quantify temporal changes in the number of human–animal encounters using social media data. We focused on quantifying uncertainty in these changes. We then applied the framework to quantify changes to wildlife viewing pressure on loggerhead sea turtles (Zakynthos Island, Greece) during the course of COVID-19 global travel disruption (2019 to 2021). Zakynthos Island supports both a major population of breeding loggerheads, which are seasonally present from May to August, and a small population of year-round resident turtles (immature and adults). The turtles frequenting the island are subjected to intense ecotourism viewing pressure from May to October (Papafitsoros et al., 2021). This sea turtle population forms part of the Mediterranean loggerhead sea turtle population, which is considered conservation dependent, based on data assimilated from long-term conservation efforts (Casale et al., 2018). However, because our framework only requires targeted social media mining for photographs and videos of animals in a given focal area, it could be widely applied to other sites where humans interact with and photograph wildlife, provided these interactions are accompanied by a substantial, consistent and timely social media activity. The information provided by our model could be used, for instance, by management agencies of wildlife habitats. We anticipate that as the data mining of social media becomes automated, this information could be made available in real time, informing managers of noticeable changes to viewing pressure of specific threatened animal groups, prompting action to mitigate disturbance (i.e. dynamic management). Through increasing the reliability of using social media-based methods to quantify wildlife tourism pressure, our model facilitates the integration of this global citizen-based phenomenon as a science tool to identify and mitigate adverse effects of human–wildlife interactions.

2. Methods: Model development and validation

2.1. Real and detected viewing pressures

We developed a rigorous method to compare human viewing pressure on wildlife in a given year against other years, based on detected observations from social media. The following parameters were, first, defined:

- **Real viewing pressure** V_{real}: number of times that an animal at a focal site is observed by a human. It is difficult to obtain this value because it requires continuous observation of the same animal(s).
- **Photographs of individuals** V_{photo}: number of times that a photograph/video of an animal at a focal site is recorded. By definition, this number is smaller than or equal to V_{real}.
- **Detected viewing pressure** $V_{detected}$: number of times for which a person recorded an image (photograph/video) of an animal and uploaded that image on a public social media account together with detectable identifiers, e.g. "hashtags" ("#"). We used the term *entry* for such images. By definition this number is smaller than or equal to V_{photo}.

These parameters were evaluated for all animals that were observed across all situations, locations and years. However, certain animal groups, observation conditions, locations or years could also be separated to evaluate different types of pressures at different temporal scales. We focused on comparing V_{real} across years using the corresponding $V_{detected}$, which was assumed to be known, following a targeted search in social media.
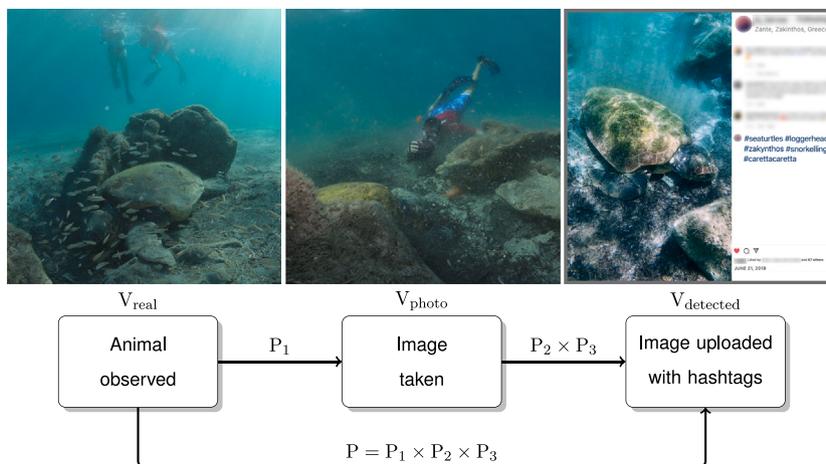
Fig. 1. Relationship of V_{real}, V_{photo}, $V_{detected}$ with P_1, P_2, P_3. Left: actual observation event of an animal, contributing to V_{real}. Middle: observer takes a photograph of the animal with probability P_1, contributing to V_{photo}. Right: Given this, the photograph appears on a public social media account with detectable hashtags and probability $P_2 \times P_3$, contributing to $V_{detected}$.

2.2. Transition probabilities

To link $V_{detected}$ and V_{real}, we delineated three *transition probabilities*. When a single observation of an individual animal is made (i.e. one person observes one animal):

- P_1 is the probability that the person took at least one photograph/ video of this animal.
- P_2 is the probability that the person has a public social media account.
- P_3 is the probability that given that the tourist has a public social media account, they uploaded a photograph/video of the animal to it with detectable identifiers.

The relationship of V_{real}, V_{photo}, $V_{detected}$ with P_1, P_2, P_3 is shown in Fig. 1. Furthermore, Fig. 2 presents a schematic of the relationship between the real and detected viewing pressure, as well as the role of the confidence interval formula (4) described in the next sections.

From the definitions above it follows that the approximate value of V_{photo} is equal to:

$$V_{photo} \simeq P_1 \times V_{real}. \tag{1}$$

Thus, P_1 is used to transfer from V_{real} to V_{photo}. To transfer from V_{photo} to $V_{detected}$, we separated the transition probability into whether the photographer (1) has a public social media account and (2) uploaded the photograph/video with detectable identifiers. The probability that both components are satisfied is equal to $P_2 \times P_3$. Thus, $P_2 \times P_3$ is used to pass from V_{photo} to $V_{detected}$, meaning that $V_{detected}$ is approximately equal to:

$$V_{detected} \simeq P_2 \times P_3 \times V_{photo}. \tag{2}$$

By combining Eqs. (1) and (2) we end up with $V_{detected}$ being approximately equal to

$$V_{detected} \simeq P_1 \times P_2 \times P_3 \times V_{real} = P \times V_{real}. \tag{3}$$

where P denotes the product of the three probabilities. In other words, P is the probability that all the three following events occur, given that an animal is observed by a human: (i) the observer took at least one photograph/video of the animal (ii) the observer has a public social media account (iii) the observer uploaded the photograph/video to that public social media account with detectable identifiers (given that the observer has a public social account). These three events will occur with probabilities P_1, P_2, P_3 respectively and they occur simultaneously with probability $P := P_1 \times P_2 \times P_3$. Some assumptions must be made on the transition probabilities P_1, P_2, P_3:

ID-independence assumption: We assumed that P_1, P_2, P_3 do not depend on the identity of the observed animal. For instance, certain individuals are not expected to have higher or lower probabilities of being photographed/videoed than other individuals or for these images to be uploaded on social media.

Time-independence assumption: We assumed P_1, P_3 would be constant across time. For instance, given that a person observes an animal, this person is as likely to record the viewed animal on a device with a camera in all years. The assumption of time independence for P_1 implies that the availability of photographic equipment (mobile phones, cameras, underwater cameras) remains (approximately) stable across the selected timeframe.

In contrast, P_2 is considered time dependent because social media users increase annually, and so the probability that a person has a social media account increases, thus we used the notation $P_2^{year_1}$, $P_2^{year_2}$ etc.

Condition of observation-independence assumption: We assumed that P_2 and P_3 do not depend on the *condition* in which the observation took place. For instance, a person would upload photographs/videos with the same probability, regardless of the condition (where or how) in which the observation took place at the focal site. Different conditions could be locations (underwater vs. boat based for marine life), means of recording (normal cameras vs. drones), or time of the day (daytime vs. nighttime). However, in certain situations (i.e. when viewing marine wildlife), P_1 is likely condition-dependent. This is because, more images are obtained from boat viewings compared to underwater viewings, as devices with cameras that can be used underwater tend to be more expensive and so used less. Thus, we used the notation $P_1^{condition_1}$, $P_1^{condition_2}$ etc.

Outcome of viewings-independence assumption: We assumed that the viewing of an animal resulting in a photograph/video that is uploaded to social media would not depend on the corresponding event of a different viewing (i.e. each viewing is independent). For instance, when two viewings are done by two different persons, it can be assumed that they act independently. However, if the same person observes two different animals, this independence is lost. For instance, if that person uploads a photograph/video of the first animal, it is more likely that they upload a photograph/video for the second animal as well. Hence, for this assumption to be true, a person would have to observe just one animal, which is not necessarily the case. This assumption can be validated by quantifying the number of entries uploaded to unique social media accounts. At sites where this is not the case (i.e. where a single person can view multiple animals), the model must be interpreted with caution.
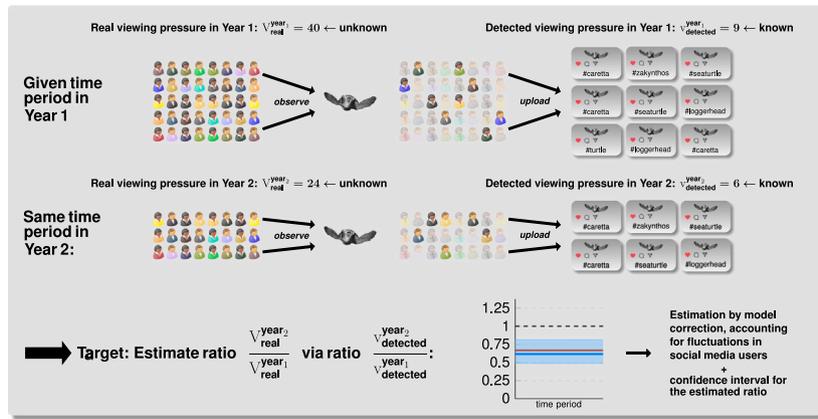
Fig. 2. Schematic illustrating the real and detected viewing pressures for an individual animal, in a specific 2-year time period. During this time period in Year 1, the animal is observed 40 times but only 9 of these encounters are uploaded on social media. In comparison, in the same time period in Year 2, the same individual is observed 24 times of which 6 are uploaded on social media. Only the detected viewing pressures in the two years ($v_{detected}^{year_1} = 9$ and $v_{detected}^{year_2} = 6$ respectively) are known. Our model uses the ratio of the detected viewing pressure $\frac{v_{detected}^{year_2}}{v_{detected}^{year_1}}$ to estimate the ratio of the real viewing pressure $\frac{V_{real}^{year_2}}{V_{real}^{year_1}}$. The estimate is corrected accounting for fluctuations in social media users and a confidence interval of this estimation is calculated via formula (4).

2.3. Linking detected viewing pressure to real viewing pressure and related uncertainty quantification

The previous section argued that $V_{detected}$ equals approximately $P \times V_{real}$, where $P := P_1 \times P_2 \times P_3$. To be more precise from a modelling perspective, V_{real} is a deterministic quantity, whereas $V_{detected}$ is a *random* variable that follows the *binomial* distribution with parameters V_{real} and P, denoted as $Bi(V_{real}, P)$. This means that $V_{detected}$ is the number of *successes* of V_{real} independent *experiments* each of which has success probability P. Here, *success* is defined as the event in which a photograph/video was taken and uploaded with hashtags, and *experiments* are defined as animal viewings. We stated that $V_{detected}$ is *approximately* equal to $P \times V_{real}$. This is because $P \times V_{real}$ is the *expected value* of $V_{detected}$; however because $V_{detected}$ is distributed around its expected value (with nonzero variance), its true value is not known. We can only observe its realisation, denoted by $v_{detected}$, which is the number of detected images in social media. This distinction is important from a modelling point of view. The conceptual difference between $V_{detected}$ and $v_{detected}$ can be understood, if a hypothetical experiment is "repeated" by "going back in time". Even though the probability that $V_{detected}$ belongs to a certain interval remains the same, $v_{detected}$ (its realisation) would differ due to randomness (in practice also due to unquantified, difficult to track, factors).

Given some detected entries in two years, $v_{detected}^{year_1}$ and $v_{detected}^{year_2}$, referring to a given time window, condition of observation and group of animals, we wanted to quantify the uncertainty of the corresponding real viewing pressure change, i.e. how much larger/smaller $V_{real}^{year_2}$ is compared to $V_{real}^{year_1}$. Given an acceptable level of error $0 < \alpha < 1$, we aimed to find an "interval I", such that:

$\frac{V_{real}^{year_2}}{V_{real}^{year_1}}$ belongs to the interval I with probability $1 - \alpha$.

Specifically, we showed that the confidence interval I is defined by the following inequalities

$$\frac{v_{detected}^{year_2}}{v_{detected}^{year_1}} \frac{P^{year_1}}{P^{year_2}} e^{-k\sqrt{\frac{1}{v_{detected}^{year_2}} + \frac{1}{v_{detected}^{year_1}}}} \leq \frac{V_{real}^{year_2}}{V_{real}^{year_1}} \leq \frac{v_{detected}^{year_2}}{v_{detected}^{year_1}} \frac{P^{year_1}}{P^{year_2}}$$

$$\times \, e^{k\sqrt{\frac{1}{v_{detected}^{year_2}} + \frac{1}{v_{detected}^{year_1}}}}, \tag{4}$$

where, k is the value of the quantile function of the standard normal distribution at level $1 - \frac{\alpha}{2}$. For instance for $\alpha = 0.05$, which we fix henceforth, we have $k = 1.96$. Supplementary material A shows how

the confidence interval (4) was derived, and associated justification. In brief, the confidence interval was derived using three assumptions: (i) relatively large values for $v_{real}^{year_1}$, $v_{real}^{year_2}$, (ii) relatively small values for P^{year_1}, P^{year_2} and (iii) the observed values $v_{detected}^{year_1}$, $v_{detected}^{year_2}$ being close to their respective means. Supplementary material B explains these assumptions.

The confidence interval (4) implies the point estimate (most probable value) for real pressure ratio $V_{real}^{year_2}/V_{real}^{year_1}$ is equal to

$$\frac{v_{detected}^{year_2}}{v_{detected}^{year_1}} \frac{P^{year_1}}{P^{year_2}}, \tag{5}$$

which is natural, as it is simply the ratio of the detected pressure $v_{detected}^{year_2}/v_{detected}^{year_1}$ adjusted by the ratio of probabilities P^{year_2}/P^{year_1}. Even though the true probabilities P^{year_1} and P^{year_2} are not known, due to our assumptions, we estimate their ratio as

$$\frac{P^{year_2}}{P^{year_1}} = \frac{P_1^{year_2} P_2^{year_2} P_3^{year_2}}{P_1^{year_1} P_2^{year_1} P_3^{year_1}} = \frac{P_1 P_2^{year_2} P_3}{P_1 P_2^{year_1} P_3} = \frac{P_2^{year_2}}{P_2^{year_1}} \approx \frac{N_{platform}^{year_2}}{N_{platform}^{year_1}}, \tag{6}$$

that is, as the ratio of the global number of users of the fixed social media platform in consecutive time frames, i.e. number of users in year$_1$ and year$_2$, denoted by $N_{platform}^{year_1}$ and $N_{platform}^{year_2}$, respectively.

2.4. A simple illustration of the confidence interval

To illustrate the confidence interval (4), we provide an example with $v_{detected}^{2021} = 200$ and $v_{detected}^{2020} = 100$ as the number of detected entries. If we assume that the number of social media users did not change, (i.e. $N_{platform}^{year_1} = N_{platform}^{year_2}$), then the confidence interval based on formula (4) is calculated as

$$1.57 \leq \frac{V_{real}^{2021}}{V_{real}^{2020}} \leq 2.54.$$

If the number of observations increases (e.g. $v_{detected}^{2021} = 2000$ and $v_{detected}^{2020} = 1000$) then the new confidence interval shrinks to

$$1.85 \leq \frac{V_{real}^{2021}}{V_{real}^{2020}} \leq 2.16.$$

In both cases, the confidence interval lies around the point estimate $v_{detected}^{2021}/v_{detected}^{2020} = 200/100 = 2$ due to (5). As the number of entries increases, the confidence interval shrinks, increasing its robustness and
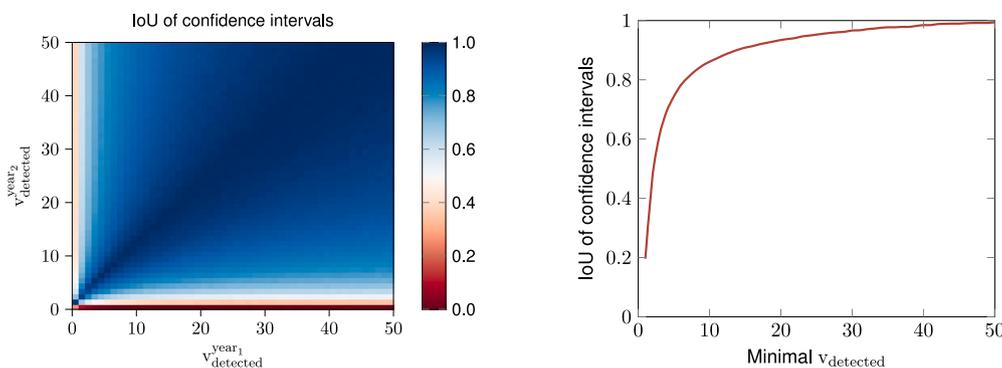
Fig. 3. Left: Comparison of our estimated confidence interval (4) and the true confidence interval computed via simulations for $v_{detected}^{year_1}$ and $v_{detected}^{year_2}$ between 1 and 50, using IoU (intersection over union as the similarity metric). Right: IoU of estimated and true confidence intervals as a function of the minimum number of detected entries, with Minimal $v_{detected} = \text{minimum}(v_{detected}^{year_1}, v_{detected}^{year_2})$.

confidence we have in it. However, if the number of social media users doubled, then the confidence interval would shrink by a factor of two to

$$0.93 \leq \frac{V_{real}^{2021}}{V_{real}^{2020}} \leq 1.08.$$

This is logical as the same real pressure with two times the number of Instagram users should result in twice as many entries and consequently double the observed pressure.

2.5. Confidence interval validation

Because three additional assumptions were required to derive the confidence interval (4) we verified its accuracy and validated its use. We computed the confidence interval (4) for a series of $v_{detected}^{year_1}$ and $v_{detected}^{year_2}$ and compared the outputs with the true confidence interval, computed via simulations. Supplementary material D presents the procedure used to compute the true confidence interval. In brief, it was based on the knowledge of transition probabilities, which are not known for real data, and were chosen as $P^{year_1} = P^{year_2}$ for the simulation. Once we computed these two confidence intervals, we computed their IoU (intersection over union) as a comparison metric. IoU = 0 represented no intersection (our estimated confidence interval is not accurate), and IoU = 1 represented a perfect alignment (our estimated confidence interval is perfectly accurate).

Fig. 3 presents the outputs of this procedure based on $v_{detected}^{year_1}$ and $v_{detected}^{year_2}$ for all possible combinations of numbers between 1 and 50. If either number of detected entries $v_{detected}^{year_1}$ and $v_{detected}^{year_2}$ is small, then our estimated and true confidence interval had small IoU, and the estimated uncertainty is not trustworthy. However, this discrepancy is rectified when the number of detected entries rises. Fig. 3 (right) quantifies this and shows the minimum number of detected entries $v_{detected}^{year_1}$ and $v_{detected}^{year_2}$ so that IoU reaches a certain threshold. For example, 5 entries are sufficient to obtain an IoU of 75%, whereas 10 entries produce an IoU of 86%. Thus, the results are trustworthy once the number of entries for both are greater than or equal to 5. Of note, we also performed simulations with $P^{year_1} \neq P^{year_2}$ with similar conclusions, and for the sake of simplicity (the resulting graph is symmetric) we presented here the case $P^{year_1} = P^{year_2}$.

2.6. Empirical application of the model

2.6.1. Study site

We tested the model by applying it to a site (Laganas Bay, Zakynthos Island, Greece) supporting large numbers of loggerhead sea turtles (*Caretta caretta*) subject to intense ecotourism (Fig. 4). The site supports around 300 breeding adult females seasonally (April–August)

(Margaritoulis, 2005; Margaritoulis et al., 2011; Schofield et al., 2017) and around 40 year-round residents of mostly juveniles and adult males (Schofield et al., 2020; Papafitsoros et al., 2021). The island is a popular holiday destination in summer (May–October) with over 850,000 visitors (2018–2019 on average, Papafitsoros et al., 2021). There is a well established wildlife-watching industry, where tourists observe turtles on organised boat tours (Fig. 4). This industry is estimated to service 180,000 tourists over 9000 trips per year, generating an annual revenue of more than 2.7 million euros (Schofield et al., 2015; Papafitsoros et al., 2021).

2.6.2. Collection of social media records

Social media records were collected for 2019, 2020 and 2021 following the methods described in Papafitsoros et al. (2021). The first year represented a typical tourism year (2019) and the next two years (2020 and 2021) were heavily and mildly impacted by COVID-19, respectively, when tourism levels were limited by global travel restrictions (Schofield et al., 2021). Instagram was selected over other social media types because of its popularity and convenient search framework via "hashtags" and "places". We searched for entries using specific hashtags and places related to the study site and species. The list of hashtags used is provided in the supplementary material of Papafitsoros et al. (2021). For an entry to be detected, the user must have a public profile on Instagram and the entry must be uploaded using one of the specified hashtags. The search was performed on at least once weekly from 1 May to 31 October each year, with regular retrospective searches to minimise missed entries. We did not record entries from previous seasons or those used as advertisements by tour operators. "Throwback" posts and re-posts from a different account (duplicate posts) were detected by inspecting the captions, and were excluded from our dataset. For each entry we recorded the date that it was uploaded. We previously showed that more than 80% of entries are uploaded within 2 days of being captured (Papafitsoros et al., 2021). This information was obtained because a proportion of Instagram entries (74 and 97 entries in 2018 and 2019 respectively) showed both the date they had been uploaded and the actual dates that they had been recorded in the caption of the entry. Observation conditions that could be detected from social media were considered, allowing the uploaded entries to be distinguished based on them (i.e., underwater, boat, type of device, location). Here, we only selected entries taken from a boat, i.e. organised tours, and private hire boats, which defines the "condition of the observation". Of note, at this site, tourists remain on board boats during organised boat tours, and take photographs of sea turtles when they surface to breathe, using standard commercial digital cameras or mobile phones.

Fig. 4. Left: Southern part of Zakynthos Island (Greece) showing Laganas Bay and the maritime zoning of the National Marine Park of Zakynthos (Zones A-B-C). Inset: arrow indicates location of Zakynthos Island in Greece. Right: Wildlife watching vessels observing sea turtles. Photograph credits: Kostas Papafitsoros.

2.6.3. Number of instagram users and visitors to Zakynthos

The number of annual Instagram users was obtained from the Statista website: https://www.statista.com/statistics/183585/instagram-number-of-global-users, *(accessed on 14.01.2022)*, see also Supplementary table 3.

The number of arrivals to Zakynthos airport were obtained for each month (May–October 2019–2021) from the Hellenic Civil Aviation Authority (CAA, http://www.ypa.gr). For 2019–2020, the number of daily airport arrivals was also available. Because daily airport arrivals in 2021 were not available, we employed a mass conserving interpolation approach to generate a series of simulated daily arrivals, the sum of which (over all days) is the same with the original sum (over months) (https://uk.mathworks.com/matlabcentral/fileexchange/99404-conservative-regridding). This approach was validated using 2019 and 2020 airport arrivals for which the real number of daily arrivals was available.

2.6.4. Comparing real viewing pressure ratios with baseline values

We investigated how our developed framework provided more reliable information on changes to viewing pressure from $year_1$ to another year, (i.e. $year_2$). We compared the estimated ratio of V_{real} with the corresponding confidence intervals to two baseline values (constant 1 and ratio of visitor arrivals). Constant 1 represented equal viewing pressure in $year_1$ and $year_2$. When the ratio of V_{real} was above (below) value 1, then viewing pressure increased (decreased) from $year_1$ to $year_2$. The second baseline value was the ratio of visitor arrivals. We considered the ratio of the airport arrivals to be representative of the ratio of the total number of people present at this focal site and observing sea turtles. The comparison of the V_{real} ratio to arrivals tested to what degree V_{real} is directly proportional to the number of visitors. When the ratio of V_{real} was not equal to the arrival ratio, changes to V_{real} might have also been driven by additional factors, not just the change in the number of tourist arrivals (e.g. more organised tours operating between years). When the ratio of V_{real} was above (resp. below) the arrival ratio, this meant that changes to V_{real} were higher (resp. lower) than that predicted by changes to tourist numbers when assuming a simple linear relationship between the number of visitors and number of animal viewings. This could happen due to other factors that affect real viewing pressure (apart from the number of visitors), which differ in a non-predictable way from $year_1$ to $year_2$. For instance, small fluctuations in ecotourism activity might arise due to sudden severe weather events or disruptions in the normal functioning of social media platforms.

3. Results

3.1. Overview of instagram records and tourist numbers

We recorded 2646 Instagram entries from boats for 2019, 2020 and 2021 (n = 1382, n = 387 and n = 877 respectively; Fig. 5;
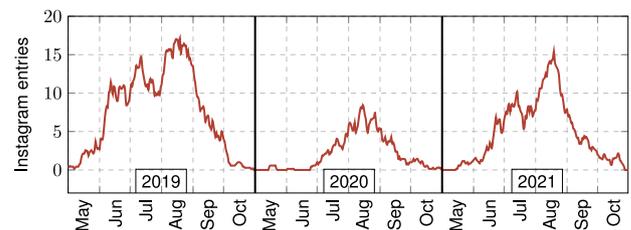

Fig. 5. Number of recorded daily Instagram entries taken from boats for May–October for 2019–2021 (moving mean of 7 days).

Supplementary table 1). The total number of visitors in 2019 was 1,288,651 (airport and port combined) versus 386,756 (70% lower) in 2020, due to COVID-19 pandemic travel restrictions. In 2021, there were 803,868 visitors (107.8% higher than 2020, but still 37.6% lower than 2019; Supplementary table 2).

3.2. Model versus raw data

We investigated three quantities as estimates for the ratio of the ratio $V_{real}^{year_2}/V_{real}^{year_1}$:

- raw data ratio $v_{detected}^{year_2}/v_{detected}^{year_1}$;
- point estimate (5) ratio $\frac{v_{detected}^{year_2}}{v_{detected}^{year_1}}\frac{N_{insta}^{year_2}}{N_{insta}^{year_1}}$;
- associated confidence interval (4).

The raw data ratio was a natural candidate for estimating the real pressure ratio. However, unlike the other two estimates, changes to the number of Instagram users and uncertainty of that estimate were not considered. Therefore, confidence interval (4) might be more reliable than the raw data ratio estimate.

For example, we compared boat viewing pressure before (2019) and during the COVID-19 pandemic (2021) for these three quantities when aggregating viewing images at 3 and 7 day intervals (Fig. 6). For the 3 day aggregation window, the raw data ratios $v_{detected}^{2021}/v_{detected}^{2019}$ were primarily below 1, apart from three time windows (mid-May, late June, mid-July; see three arrows in Fig. 6), suggesting that boat viewing pressure rose from 2019 to 2021 for these specific 3-day windows. However, when adjusting for changes in the number of Instagram users, the corresponding point estimate value of these 3-day windows decreased close to 1, with the confidence interval being above and below 1. Thus, the number of data (detected entries) was insufficient to make objective inferences for this narrow time interval. In contrast, when aggregating viewing images to 7 day intervals, the confidence intervals for June–September remained below 1, indicating that boat pressure change could be inferred with high certainty, and that it was lower in 2020 than 2019.
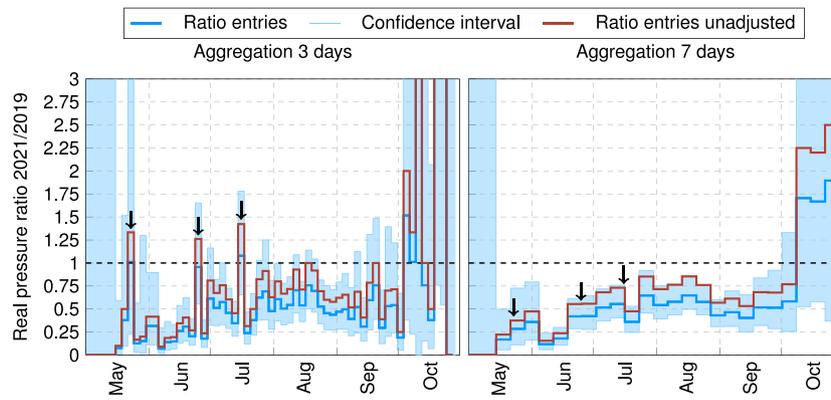
Fig. 6. Confidence intervals (light blue) for 2021/2019 ratios of the real boat pressure $V_{real}^{2021}/V_{real}^{2019}$ for loggerhead turtles (aggregating viewings in 3 and 7-days intervals). Dark blue line represents the best point estimate, $\frac{v_{detected}^{2021}}{v_{detected}^{2019}}\frac{N_{insta}^{2021}}{N_{insta}^{2019}}$; red line represents the simple ratio of detected entries $v_{detected}^{2021}/v_{detected}^{2019}$ i.e., not accounting for changes in the number of Instagram users. Horizontal dashed line represents equal real viewing pressures for the two years (ratio equal to 1). The three arrows point at time windows where the unadjusted ratio was above one (with low confidence) and dropped below one (with high confidence) when adjusting for changes in the number of Instagram users and using a larger aggregation window. See https://github.com/sadda/Turtles_Covid for interactive plot for all year ratios 2020/2019, 2021/2019, 2021/2020.
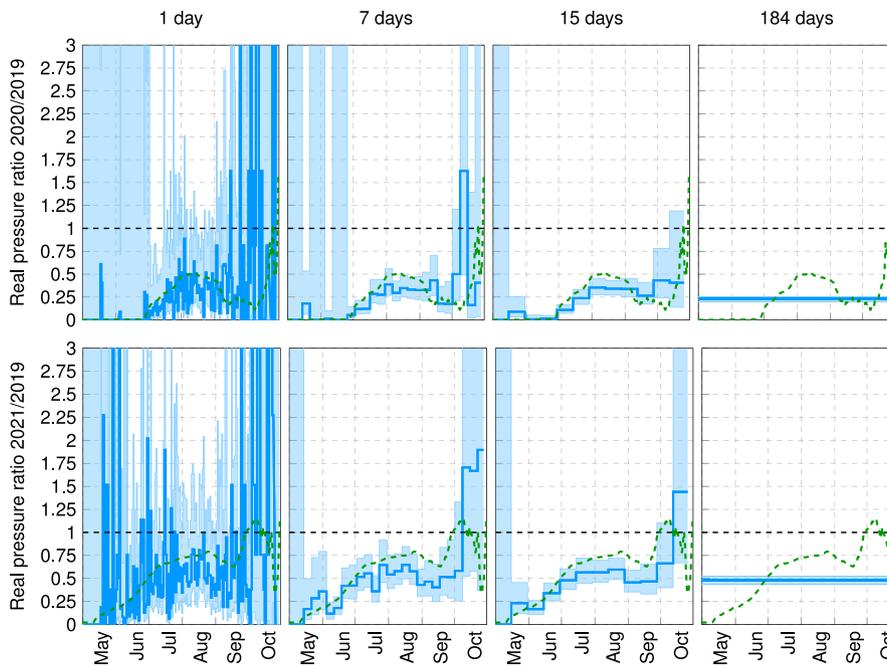
Fig. 7. Confidence intervals (light blue) for 2020/2019 (top) and 2021/2019 (bottom) ratios of real boat pressure on loggerhead turtles, $V_{real}^{2020}/V_{real}^{2019}$, $V_{real}^{2021}/V_{real}^{2019}$ for a series of 1, 7, 15 and 184-day aggregation windows (the latter refers to the full period of May–October considered separately for each year). Dark blue line represents the best point estimate. Green line represents the corresponding ratio of visitor arrivals (moving mean of 7 days). Horizontal dashed line represents equal real viewing pressure for the two years (ratio equal to 1). Arrows are referred to in the text. See https://github.com/sadda/Turtles_Covid for interactive plot of all year ratios 2020/2019, 2021/2019, 2021/2020.

We note that for visualisation purposes we selected the 2020/2019 and 2021/2019 ratios, i.e., keeping the year 2019 as the denominator, and we did not work with the inverse ratios, in order for the ratios to mostly fall between zero and one.

### 3.3. Effect of time aggregations on the uncertainty of real pressure ratio

To evaluate how aggregating image records in different time intervals impacted uncertainty in the real pressure ratio, we evaluated different time windows (1, 7, 15, 184 days (entire season); Fig. 7) for 2020/2019 and 2021/2019. Data based on daily records led to extremely high uncertainty for the estimated real pressure ratio (very large confidence intervals), and should not be used (Fig. 7). Certainty increased as the data were aggregated into larger time intervals, with 7 and 15-day intervals providing higher confidence for July–September, as the corresponding detected entries were higher (Fig. 8). The real

pressure ratio during this period was below 1, with high certainty, as all confidence intervals were below this number. The arrival ratio closely followed the real pressure ratio, indicating proportionality between real boat pressure and the number of tourists frequenting the site. Based on the entire season (184 days), compared to 2019, boat observation pressure was 25% and 50% lower in 2020 and 2021, respectively, with very high certainty, due to the small confidence interval.

Fig. 9 provides a further visualisation on how the size of the confidence intervals varied with the aggregation window (number of days). We aggregated the days of each month based on this window. For example, for a 4-day aggregation window, a month with 30 days was split into 7 windows (excluding the last 2 days). Then, for each month (May–October), we computed the average size of the confidence interval for each window, defined as the ratio upper over the lower bound of the interval, with 1 corresponding to the smallest possible confidence interval (Supplementary material C). During July–September (and June
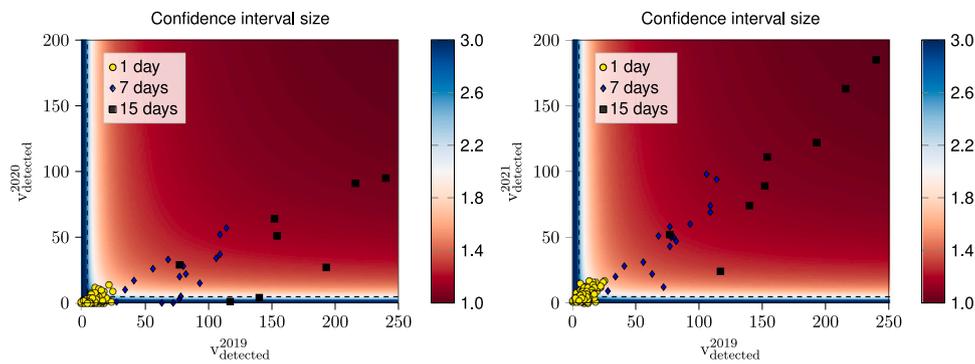
Fig. 8. Scatter plots of the values for pairs of detected entries ($v^{2019}_{detected}$, $v^{2020}_{detected}$) and ($v^{2019}_{detected}$, $v^{2021}_{detected}$) when considering 1-, 7- and 15-day aggregation windows, corresponding to the plots in Fig. 6. The quadrants are colour-coded based on the size of the confidence interval for the corresponding real viewing pressure ratios $V^{2020}_{real}/V^{2019}_{real}$, $V^{2021}_{real}/V^{2019}_{real}$, to show the degree of uncertainty as a function of the magnitude of the values for the detected entries. Larger aggregation windows, lead to a larger number of detected entries and thus smaller confidence intervals (higher certainty) for the real viewing pressure ratio.
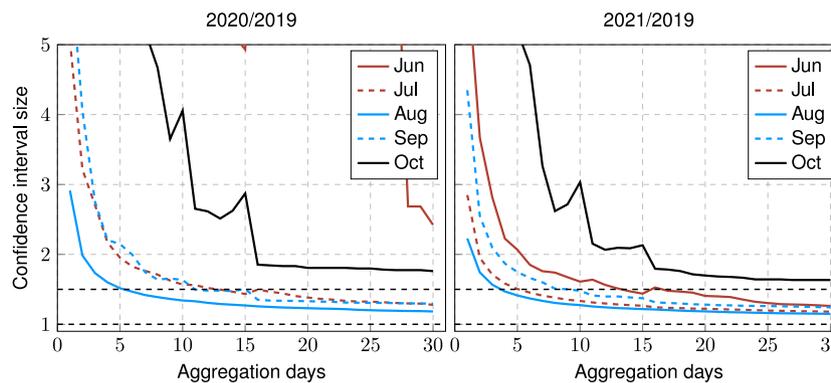
Fig. 9. Average confidence interval size (defined as upper bound/lower bound ratio) for real boat pressure ratio in 2020/2019 and 2021/2019 versus different day aggregation windows. Aggregation windows were plotted separately for each month (June–October). May was excluded due to insufficient data.

2021/2019), the confidence interval had a smaller average width than October, due to the larger number of detected entries. In parallel, the width decreased in all months as the aggregation window increased. For July–September (high visitor season) of 2021/2019, a 7-day aggregation window represents an optimal balance between achieving high certainty (confidence interval size < 1.5) with sufficiently fine time scale resolution, since aggregation windows of >7-days do not cause the width of the confidence interval to significantly decrease further. In contrast, a 15-day window was needed for July–September for 2020/2019 ratio (and June 2021/2019), due to fewer entries. For October 2020/2019 and 2021/2019, the same certainty was achieved by setting an aggregation window of >30 days.

4. Discussion

A key issue of passive crowdsourcing/iEcology raw data and associated analysis is determining uncertainty to improve objective interpretation (Isaac et al., 2014; Jarić et al., 2020). As such, the present study provided a rigorous framework for using social media imagery with confidence to infer temporal changes (within and across years) in wildlife watching pressure. We showed that confidence increased when integrating multiple days. This framework modelled the flow of information from a human–animal interaction event to that event appearing in social media in a detectable way. This was achieved by introducing the notions of real and detected viewing pressure, linking them via transition probabilities and showing how temporal changes in the former can be estimated by temporal changes in the latter. Detected viewing pressure was modelled as a random variable, allowing uncertainty in real viewing pressure change to be quantified by rigorously deriving confidence intervals. Through applying the model at a site supporting

large scale sea turtle ecotourism, we demonstrated its advantages over simply using raw data (i.e., simple ratio of detected animal entries in given time periods). Through increasing the reliability of using social media-based methods to quantify wildlife tourism pressure, our model facilitates the use of social media as a scientific tool providing evidence based information on human pressure.

Our study delineated appropriate time-windows to analyse social media data temporally, allowing robust comparison across years, despite highly different visitor levels due to COVID-19. The optimal aggregation window (balance of high confidence and high temporal resolution) for images at our study site was 7 days during the peak period, with more days being needed during the low periods, due to fewer images. Our results supported those of Tenkanen et al. (2017), who showed that social media captured temporal variation in national park visitation rates at a monthly scale, but was challenging on a daily basis, due to fewer social media records. Other studies arbitrarily grouped information into monthly periods. For instance, Molyneaux et al. (2021) compared the monthly number of photographs posted on Instagram before and after the onset of the COVID-19 pandemic to quantify variation in interactions between tourists and orangutans. Barros et al. (2019) observed that Flickr geotagged data have enough information to capture daily, weekly and monthly distribution patterns of visitors of Spanish national parks, but did not perform temporal comparisons of visitor numbers across years. By implementing our approach, studies could maximise social media data by selecting appropriate aggregation windows of sufficient temporal granularity, guaranteeing certainty in interpretations.

This modularity of our model allows more refined versions to be developed as knowledge of its key constituents becomes available. For instance, estimates of the ratio $P^{year_2}_2/P^{year_1}_2$ could be improved by

directly incorporating the number of users that belong to the main demographic characteristics (nationality, age group, culture) of visitors to a focal site (Väisänen et al., 2021) or if available, the total number of social media posts made from that site. The time-independence assumption of transition probability P_1 could also be refined (i.e. availability of photographic equipment across years) by incorporating information on the yearly ratios of global (or more refined demographically) sales of equipment (e.g. smartphones/underwater cameras associated to marine wildlife viewing). Our model could easily incorporate such refinements by adjusting the corresponding ratio formulas. Our model also allowed temporal changes to viewing pressure of specific animal individuals/groups to be focused on by combining social media images with photo-identification records. Examples of this include analysis of specific African wild dog dens (Cloutier et al., 2021), gorilla family groups (Molyneux et al., 2021), and resident foraging sea turtles versus migratory turtles (Papafitsoros et al., 2021). This flexibility is particularly important because wildlife viewing pressure is not equally distributed across all animals present at a given site or time period. Certain individuals are often subjected to disproportionly higher viewing pressure, due to ecotourism activities incidentally or deliberately targeting these groups, particularly resident animals (Semeniuk et al., 2009; Christiansen and Lusseau, 2014; Schofield et al., 2015; Papafitsoros et al., 2021). Thus, social media could be used to tease out this information quantitatively, and to introduce more appropriate watching practices and conservation measures.

The selected social media platform also influenced interpretation (Ghermandi et al., 2020), particularly as the demographics of visitors and social media use change over time. Ideally, social media data should reflect actual human–animal interactions, while minimising user-induced biases, in parallel to revealing temporal variability in these interactions. For this reason, we used Instagram, because it captures real life human activity effectively (Tenkanen et al., 2017; Hausmann et al., 2018). This attribute allowed us to objectively identify temporal variation in viewing pressure. Alternatively, Flickr has been widely used to infer spatial information on national park visitations, due to its easily accessible geotagged photographs (Wood et al., 2013; Barros et al., 2019; Ghermandi et al., 2020; Edwards et al., 2021); yet, its temporal correlation with ground-truthed data is lower than of Instagram (Tenkanen et al., 2017). Other platforms could also be used in our model, like YouTube (Otsuka and Yamakoshi, 2020; Taklis et al., 2020); however, larger time aggregation windows might be required to account for lower temporal correlation between the time of viewing and time of video uploading. In contrast, the demographics of Instagram users, as well as those of more recently popular platforms (e.g. TikTok), are not always representative of visitors to a focal site. For instance, generally younger people use other social media forms (Heikinheimo et al., 2017; Hausmann et al., 2018). Furthermore, the country of origin might also affect social media use (Ghermandi and Sinclair, 2019). Biases might also be self-diminishing, such as if, hypothetically, smartphone availability increased across years whereas Instagram (or any social media) use decreases. In this case, the net output of social media content might remain constant, even though the processes that govern information flow from the focal site to social media platforms are time dependent.

As with other technologies involving the remote collection of data, passive crowdsourcing/social media data should be validated (or ground-truthed) using robust field data to guarantee conservation policies are informed appropriately (Jarić et al., 2020). Since information flow from the human–animal encounter to it being detectable in social media changes across sites, validation should be site-specific. The current study was methods-based, and so ground truthing was not the primary focus; however, our outputs closely aligned with previous studies at our site (Schofield et al., 2015; Papafitsoros et al., 2021). The ground-truthing of our model would involve estimating P, which would allow V_{real} to be directly estimated at a given time interval, rather than only the ratio in two time periods. This approach could provide a useful comparison of the ratio estimation we provide here; however, determining P remains challenging. P could be estimated by using questionnaires targeting visitors at focal sites, investigating whether visitors observed wildlife and uploaded any images on social media, or by quantifying actual viewing pressure using direct observations over time. Regular repetition of such surveys is necessary to determine the time dependence of P, particularly in the long term, i.e. across years.

Finally, for our model to be applied successfully, three minimal parameters are required at a given site; high levels of interactions and high levels of social media activity, timely social media activity, and extra information to validate the model. First, the activity of humans observing, interacting, and photographing wildlife must be accompanied by substantial social media activity. This is required since as we showed, high numbers of social media entries increase the confidence we have on the estimated temporal change of the real viewing pressure. Thus, while our framework is suitable at sites where large scale tourism and large numbers of wildlife coexist (>20 individuals), it is unlikely to be useful in areas low levels of where human–wildlife interactions are low (e.g. scientific expeditions in remote areas). Second, social media activity should be timely. Our model works best in areas where people upload images of their encounters with animals within 1–2 days of making observations. If this is not the case, a correction for time delay should be incorporated in the model. This phenomenon depends on the social media platform used and site characteristics (good available internet connection). For instance, people on vacation are more likely to share personal moments almost immediately (Bayer et al., 2016). Thus, sites that support tourist activities are more likely to be suitable for our framework. Thirdly, the site should allow for some validation of the model in any of the ways described in the previous paragraph. Sites where numbers of visitors are recorded, either via park visitation ticketing schemes or simply by recording the number of arrivals in main entrance points (as it was the case for the site we considered here), are more suitable than sites where there is no information on visitation numbers is available whatsoever.

5. Conclusion

Most protected areas globally receive visitors that produce social media content related to them (Tenkanen et al., 2017), leading to the emergence of conservation culturomics/iEcology/passive crowdsourcing. These records are typically used to study spatiotemporal visitation patterns and interaction of humans with the natural environment, and identify potential threats to wildlife (Sullivan et al., 2019; Jarić et al., 2020; Edwards et al., 2021; Papafitsoros et al., 2021; Molyneaux et al., 2021; Van Hamme et al., 2021; Cloutier et al., 2021). However, to accomplish this, social media-related biases and uncertainties must be overcome, with the current study making an important step towards this. We modelled the flow of information from human–animal encounter to its appearance in social media, and inferred temporal changes in the number of encounters from temporal changes to corresponding social media imagery. We focused on quantifying uncertainty underlying such inferences, and identifying aggregation windows that combined increased temporal granularity to reduce uncertainty. We expect that continued advances in automating the mining of social data, and machine learning, will facilitate the creation of well-curated and meaningful datasets (Väisänen et al., 2021; Tuia et al., 2022). Combined with our framework, such advances would increase the number of studies using social media data to infer human impacts on wildlife in different situations.

CRediT authorship contribution statement

Kostas Papafitsoros: Conceived the study, Assimilated the social media data, Led the writing of the manuscript with critical contributions from Lukas Adam and Gail Schofield. **Lukáš Adam:** Led the model development, Data analysis, Critical contributions to the writing of the manuscript. **Gail Schofield:** Critical contributions to the writing of the manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Code and data used are publicly available in the following github repository: https://github.com/sadda/Turtles_Covid.

Acknowledgements

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ecolmodel.2022.110223.

References

Barros, C., Moya-Gómez, B., García-Palomares, J.C., 2019. Identifying temporal patterns of visitors to national parks through geotagged photographs. Sustainability 11 (24), http://dx.doi.org/10.1016/j.tree.2020.03.003.

Bates, A.E., Primack, R.B., Biggar, B.S., et al., 2021. Global COVID-19 lockdown highlights humans as both threats and custodians of the environment. Biol. Cons. 263, 109175. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Bayer, J., Ellison, N., Schoenebeck, S., Falk, E., 2016. Sharing the small moments: ephemeral social interaction on Snapchat. Inf. Commun. Soc. 19 (7), 956–977. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Birk, S., Chapman, D., Carvalho, L., et al., 2020. Impacts of multiple stressors on freshwater biota across spatial scales and ecosystems. Nat. Ecol. Evol. 4, 1060–1068. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Burgin, S., Hardiman, N., 2015. Effects of non-consumptive wildlife-oriented tourism on marine species and prospects for their sustainable management. J. Environ. Manag. 151, 210–220. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Casale, P., Broderick, A., Camiñas, J., et al., 2018. Mediterranean sea turtles: current knowledge and priorities for conservation and research. Endanger. Species Res. 229–267. http://dx.doi.org/10.3354/esr00901.

Christiansen, F., Lusseau, D., 2014. Understanding the ecological effects of whale-watching on cetaceans. In: Higham, J., Bejder, L., Williams, R. (Eds.), Whale-Watching. Cambridge University Press, pp. 177–192. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Cloutier, T.L., Rasmussen, G.S.A., Giordano, A.J., Kaplin, B.A., Willey, L., 2021. Digital conservation: using social media to investigate the scope of african painted dog den disturbance by humans. Hum. Dimens. Wildl. 26 (5), 481–491. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Di Minin, E., Tenkanen, H., Toivonen, T., 2015. Prospects and challenges for social media data in conservation science. Front. Environ. Sci. 3, 63. http://dx.doi.org/10.3389/fenvs.2015.00063.

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J., Phillips, T., Purcell, K., 2012. The current state of citizen science as a tool for ecological research and public engagement. Front. Ecol. Environ. 10 (6), 291–297. http://dx.doi.org/10.1890/110236.

Edwards, T., Jones, C.B., Perkins, S.E., Corcoran, P., 2021. Passive citizen science: The role of social media in wildlife observations. PLOS ONE 16 (8), 1–22. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Ghermandi, A., Sinclair, M., 2019. Passive crowdsourcing of social media in environmental research: a systematic map. Global Environmental Change 55, 36–47. http://dx.doi.org/10.1016/j.gloenvcha.2019.02.003.

Ghermandi, A., Sinclair, M., Fichtman, E., Gish, M., 2020. Novel insights on intensity and typology of direct human-nature interactions in protected areas through passive crowdsourcing. Global Environ. Change 65, 102189. http://dx.doi.org/10.1016/j.gloenvcha.2020.102189.

Hausmann, A., Toivonen, T., Slotow, R., Tenkanen, H., Moilanen, A., Heikinheimo, V., Di Minin, E., 2018. Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas. Conserv. Lett. 11 (1), e12343. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Heikinheimo, V., Di Minin, E., Tenkanen, H., Hausmann, A., Erkkonen, J., Toivonen, T., 2017. User-generated geographic information for visitor monitoring in a national park: A comparison of social media data and visitor survey. ISPRS Int. J. Geo-Inf. 6 (3), http://dx.doi.org/10.1016/j.tree.2020.03.003.

Huang, T., Elghafari, A., Relia, K., Chunara, R., 2017. High-resolution temporal representations of alcohol and tobacco behaviors from social media data. Proc. ACM Human-Comput. Interact. 1 (54), 1–26. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Isaac, N.J.B., van Strien, A.J., August, T.A., de Zeeuw, M.P., Roy, D.B., 2014. Statistics for citizen science: extracting signals of change from noisy ecological data. Methods Ecol. Evol. 5 (10), 1052–1060. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Jarić, I., Correia, R.A., Brook, B.W., Buettel, J.C., Courchamp, F., Di Minin, E., Firth, J.A., Gaston, K.J., Jepson, P., Kalinkat, G., Ladle, R., Soriano-Redondo, A., Souza, A.T., Roll, U., 2020. Iecology: Harnessing large online resources to generate ecological insights. Trends Ecol. Evol. 35 (7), 630–639. http://dx.doi.org/10.1016/j.tree.2020.03.003.

König, H.J., Kiffner, C., Kramer-Schadt, S., Fürst, C., Keuling, O., Ford, A.T., 2020. Human–wildlife coexistence in a changing world. Conserv. Biol. 34 (4), 786–794. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Ladle, R.J., Correia, R.A., Do, Y., Joo, G.-J., Malhado, A.C.M., Proulx, R., Roberge, J.-M., Jepson, P., 2016. Conservation culturomics. Front. Ecol. Environ. 14 (5), 269–275. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Larson, C.L., Reed, S.E., Merenlender, A.M., Crooks, K.R., 2016. Effects of recreation on animals revealed as widespread through a global systematic review. PLOS ONE 11 (12), 1–21. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Lehmann, E.L., Casella, G., 2006. Theory of Point Estimation. Springer Science & Business Media.

Lenzi, C., Speiran, S., Grasso, C., 2020. "Let me take a selfie": Implications of social media for public perceptions of wild animals. Soc. Animals 1–20. http://dx.doi.org/10.1016/j.tree.2020.03.003.

March, D., Metcalfe, K., Tintoré, J., Godley, B.J., 2021. Tracking the global reduction of marine traffic during the COVID-19 pandemic. Nature Commun. 12, 2415. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Margaritoulis, D., 2005. Nesting activity and reproductive output of loggerhead sea turtles, caretta caretta, over 19 seasons (1984–2002) at Laganas Bay, Zakynthos, Greece: The largest rookery in the Mediterranean. Chelonian Conserv. Biol. 4 (4), 916–929.

Margaritoulis, D., Rees, A.F., Riggal, T., 2011. Reproductive data of loggerhead turtles in Laganas Bay, Zakynthos island, Greece, 2003–2009. Mar. Turt. Newsl. (131), 2–6.

Marion, S., Davies, A., Demšar, U., Irvine, R.J., Stephens, P.A., Long, J., 2020. A systematic review of methods for studying the impacts of outdoor recreation on terrestrial wildlife. Global Ecol. Conserv. 22, e00917. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Molyneaux, A., Hankinson, E., Kaban, M., Svensson, M.S., Cheyne, S., Nijman, V., 2021. Primate selfies and anthropozoonotic diseases: Lack of rule compliance and poor risk perception threatens orangutans. Folia Primatol. 92, 296–305. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Moorhouse, T.P., Dahlsjö, C.A.L., Baker, S.E., D'Cruze, N.C., Macdonald, D.W., 2015. The customer isn't always right–conservation and animal welfare implications of the increasing demand for wildlife tourism. PLOS ONE 10 (10), 1–15. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Otsuka, R., Yamakoshi, G., 2020. Analyzing the popularity of YouTube videos that violate mountain gorilla tourism regulations. PLOS ONE 15 (5), 1–20. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Papafitsoros, K., Panagopoulou, A., Schofield, G., 2021. Social media reveals consistently disproportionate tourism pressure on a threatened marine vertebrate. Animal Conserv. 24 (4), 568–579. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Rice, W.L., Pan, B., 2021. Understanding changes in park visitation during the COVID-19 pandemic: A spatial application of big data. Wellbeing, Space and Society 2, 100037. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Rutz, C., Loretto, M.C., Bates, A.E., Davidson, S.C., Duarte, C.C., Jetz, W., Johnson, M., Kato, A., Kays, R., Mueller, T., Primack, R.B., Ropert-Coudert, Y., Tucker, M.A., Wikelski, M., Cagnacci, F., 2020. COVID-19 lockdown allows researchers to quantify the effects of human activity on wildlife. Nat. Ecol. Evol. 4, 1156–1159. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Schofield, G., Dickson, L.C.D., Westover, L., Dujon, A.M., Katselidis, K.A., 2021. COVID-19 disruption reveals mass-tourism pressure on nearshore sea turtle distributions and access to optimal breeding habitat. Evol. Appl. 14 (10), 2516–2526. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Schofield, G., Katselidis, K.A., Lilley, M.K.S., Reina, R.D., Hays, G.C., 2017. Detecting elusive aspects of wildlife ecology using drones: New insights on the mating dynamics and operational sex ratios of sea turtles. Funct. Ecol. 31 (12), 2310–2319.

Schofield, G., Klaassen, M., Papafitsoros, K., Lilley, M., Katselidis, K.A., Hays, G.C., 2020. Long-term photo-id and satellite tracking reveal sex-biased survival linked to movements in an endangered species. Ecology 11, e03027. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Schofield, G., Scott, R., Katselidis, K.A., Mazaris, A.D., Hays, G.C., 2015. Quantifying wildlife-watching ecotourism intensity on an endangered marine vertebrate. Animal Conserv. 18 (6), 517–528. http://dx.doi.org/10.1111/acv.12202.

Semeniuk, C..A.D., Bourgeon, S., Smith, S.L., Rothley, K.D., 2009. Hematological differences between stingrays at tourist and non-visited sites suggest physiological costs of wildlife tourism. Biol. Cons. 142 (8), 1818–1829. http://dx.doi.org/10.1016/j.biocon.2009.03.022.

Sullivan, M., Robinson, S., Littnan, C., 2019. Social media as a data resource for #monkseal conservation. PLOS ONE 14 (10), 1–11. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Taklis, C., Giovos, I., Karamanlidis, A.A., 2020. Social media: a valuable tool to inform shark conservation in Greece. Mediterr. Mar. Sci. 21 (3), 493–498. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., Toivonen, T., 2017. Instagram, flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. Sci. Rep. 7, 17615. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., Tenkanen, H., Di Minin, E., 2019. Social media data for conservation science: A methodological overview. Biol. Cons. 233, 298–315. http://dx.doi.org/10.1016/j.biocon.2019.01.023.

Tuia, D., Kellenberger, B., Beery, S., Costelloe, B.R., Zuffi, S., Risse, B., Mathis, A., Mathis, M.W., van Langevelde, F., Burghardt, T., et al., 2022. Perspectives in machine learning for wildlife conservation. Nature Commun. 13 (1), 1–15. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Väisänen, T., Heikinheimo, V., Hiippala, T., Toivonen, T., 2021. Exploring human–nature interactions in national parks with social media photographs and computer vision. Conserv. Biol. 35 (2), 424–436. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Van Hamme, G., Svensson, M.S., Morcatty, T.Q., Nekaris, K.A.I., Nijman, V., 2021. Keep your distance: Using instagram posts to evaluate the risk of anthroponotic disease transmission in gorilla ecotourism. People Nat. 3 (2), 325–334. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Western, D., Tyrrell, P., Brehony, P., Russell, S., Western, G., Kamanga, J., 2020. Conservation from the inside-out: Winning space and a place for wildlife in working landscapes. People Nat. 2 (2), 279–291. http://dx.doi.org/10.1016/j.tree.2020.03.003.

Wood, S.A., Guerry, A.D., Silver, J.M., Lacayo, M., 2013. Using social media to quantify nature-based tourism and recreation. Sci. Rep. 3, 2976. http://dx.doi.org/10.1016/j.tree.2020.03.003.