Online participatory GIS mapping of marine recreation in Denmark: contrasting crowdsourced and representative survey approaches

Hélène Draux, Department of Geosciences and Natural Resource Management, Copenhagen University, Denmark, hedr@ign.ku.dk,

Anton Stahl Olafsson, Department of Geosciences and Natural Resource Management, Copenhagen University, Denmark, asol@ign.ku.dk,

Berit Charlotte Kaae, Department of Geosciences and Natural Resource Management Copenhagen University, Denmark, bck@ign.ku.dk,

Hans Skov-Petersen, Department of Geosciences and Natural Resource Management Copenhagen University, Denmark, hsp@ign.ku.dk

Introduction

With more than 7,300 km of coastline, Denmark has a lot to offer in terms of marine recreational activities. A large part of the Danish coast and waters are indeed used all year round for organised and unorganised activities including swimming, kayaking, fishing, hunting, surfing and sailing. Some of these activities require formal facilities, such as marinas, piers, or ramps, while others,such as swimming, are informal and therefore take place anywhere at the coast. Although the entry points into the water are often known, butnot currently systematically collected,thespatial extentof these activitiesonce in the water is not known.

The lack of spatial information is a challenge to create a balanced and coherent planning and management of the sea, as framed in marine conservation planning andMarine Spatial Planning, MSP (Douvere & Ehler, 2009; Mazor et al 2014). The limited availability and quality of data on recreational marine uses is as an obstacle for implementing EU policy (EEA 2015).Innovative ways for documenting and mapping the missing 'social landscape' of the marine environment in terms of spatial attributes of recreational use are needed (Martin & Hall-Arber 2008).

To find out the extent of use of the Danish waters for marine recreation activities, our research aimed to collect the spatial extent, diversity, and intensityof these activities.Here, we discuss the quality of the data collected through two sampling strategies: an online crowdsourcedsurvey and a commercial representative panel survey.

Methodology

We focused on two populations groups to collect data representing marine recreational activities in Denmark: organised activity groups and the general population. The organised groupsalready actively take part in water-based activities and are part of a network sharing their specific recreation activity. The general populationhowever, generally takes part in marine recreation activities in aless frequent and less organised way. Thecombination of results from thesetwouser groups givesa comprehensive spatial representation of marine activities in Denmark; both studies complementing each others.

Geo-located information was collected using a custom-built webGIS, in which participants shared the location of the water activities they had taken part in. The interface allowed for the collection of points and lines (using drawing tools), and the uploading of GPX tracks (created by satnavs).

Crowdsourced strategy

Online sampling has been found effective in hard-to-reach populations (Baltar and Brunet, 2011).Since only a limited part of the Danish population takes part in marine recreation activities,sampling through organised channels such as social networking sites and federations aided in attaining arelevant small but important groups;Facebookrepresented 49% of the site visitsduringthe first sampling year.

In that period (November 2014 - November 2015), we collected 7344 points, 624 lines, and 54 GPX tracks that were shared by 2453 unique respondents.

Panel surveystrategy

Our crowdsourced sample effectively reached active marine recreationists. However, some marine activities areless organised; e.g. swimming or snorkelling. Therefore, our first sampling strategy did not fairly take into account such'laid-back' activities, and so we used an opt-in commercial panel survey sampling strategy which targeted the general population at large.

We recruited participants through Userneeds, a Danish company specialised in organising panel surveys. Userneedscontacted more than 10,000 pre-recruitedparticipants, representative of the Danish population in terms of home address, age, and gender.

We received data from 10,291 participants, of which 4054 (39%) added 7055 points, 1634 polylines, and 2 uploaded GPX track files.

Data collection evaluation

Our data quality evaluation is based on four main criteria: *representativeness* (geographically and demographically speaking), *completeness* (proportion of location added compared to other datasets), *accuracy* (scale at which the locations were added and their positional accuracy), and *quality* (relative amount of information added to each point).

Results

Representativeness

Our crowdsourced sample had an average age of 46.5, while the average Danish adult in a jetski, kayak, rowing, or sailing clubs was 50.4 (population size: 83,379). Therefore, our sample is consistent with age of the Danish adult population in a club.The sample had anover-representation of male (75% of respondents), gender unbalance that can also be found in clubs with 70% of males based on the same population as above.

Regarding the geographic representativeness, we found that a couple ofmunicipalitieswere overly represented (normalised with the population size; 4 to 8 times

more responses than the least represented) and that the region bordering Copenhagen was not well represented.

Completeness and accuracy

Kite-surfers in Denmark have a spatial database of used locations for kitesurfing (kitemekka.dk); it contains 214 locations (Figure 1(a)). During the crowdsourcing survey, one of the 32 activities was kite-surfing, for which we collected 612 kitesurfing points or lines; some locations being duplicated, since they were mapped by different respondents (Figure 1(b)). 60% of the locations from kitemekka were included in our dataset, while some locations from our dataset were not in the kitemekka datasets.The panel survey only resulted in 10 mapped kitesurfing locations; all were also included in the crowdsourceddataset.

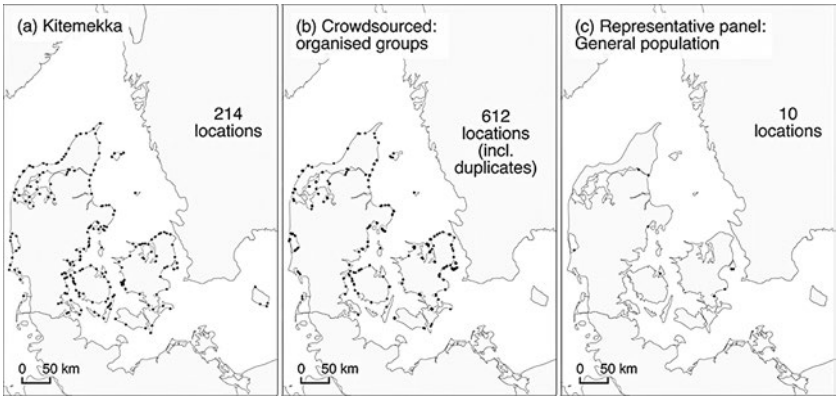

Figure 1. *Dataset comparison for kitesurfing locations.*

The maximum distance observed between two points referring to the same kitesurfing spot in the crowdsourced dataset was 3.2 km, with an average of 0.6 km. No restriction on the zoom was made when drawing points and lines; in both samples an average scale of 1:72,223 was used. This can be considered an accurate mapping scale in a national survey.

Quality

There were three types of fields to fill out in the survey: selection, simple data entry (1 word or number), and longer text fields. The selection fields were filled out for 86 to 95% of entries, data entries were filled out for 63 to 72% of entries, and longer text fields were filled out for 21 to 28% of the entries. The panel dataset gave similar numbers, and therefore the quality in terms of entries was comparable.

Conclusion

Our custom webGIS successfully gathered geo-located data from organised groups and the general population. Both strategies collected data from a representative sample of their respective population. The dataset produced, although of similar quality, wasdrastically different in terms of completeness for the kitesurfing activi-

ty; a niche activity that is therefore harder to reach through a representative sample of the population. More analysis needs to be carried out in order to explore further both datasets in terms of quality, accuracy and completeness for the 32 activities included.

Douvere, F., & Ehler, C. N. 2009. New perspectives on sea use management: initial findings from European experience with marine spatial planning. Journal of environmental management, 90(1), 77-88.

EEA, 2015. State of Europe's Seas. European Environmental Agency Report 2/2015.

Fabiola Baltar, Ignasi Brunet. 2012. Social research 2.0: virtual snowball sampling method using Facebook, Internet Research, Vol. 22 Iss: 1, pp.57 – 74

Martin, K. S., & Hall-Arber, M. (2008) The missing layer: Geo-technologies, communities, and implications for marine spatial planning. Marine Policy, 32(5), 779-786.

Mazor T, et al. (2014) The crowded sea: incorporating multiple marine activities in conservation plans can significantly alter spatial priorities. PLOS ONE 9(8):e104489.