



Tracking digital footprints in Bonaire's landscapes

Spatial distribution and characterization of tourists on Bonaire using social media

Author(s): Diana Slijkerman, Jan Tjalling van der Wal, Pepijn de Vries, Peter Verweij

Wageningen University &
Research report C052/20

Tracking digital footprints in Bonaire's landscapes

Spatial distribution and characterization of tourists on Bonaire using social media

Author(s): Diana Slijkerman (1), Jan Tjalling van der Wal (1), Pepijn de Vries (1), Peter Verweij (2)

1: Wageningen Marine Research

2: Wageningen Environmental Research

The authors would like to acknowledge funding from the Wageningen University Knowledge Base programme: KB36 Biodiversity in a Nature Inclusive Society (project number KB-36-005-002) - that is supported by finance from the Dutch Ministry of Agriculture, Nature and Food Quality

Wageningen Marine Research
Den Helder, June 2020

CONFIDENTIAL No

Wageningen Marine Research report C052/20

Keywords: digital footprint, Bonaire, Tourism, photo's FLICKR

This report can be downloaded for free from <https://doi.org/10.18174/523643>
Wageningen Marine Research provides no printed copies of reports

Wageningen Marine Research is ISO 9001:2015 certified.

© Wageningen Marine Research

Wageningen Marine Research, an institute within the legal entity Stichting Wageningen Research (a foundation under Dutch private law) represented by Dr. M.C.Th. Scholten, Managing Director

KvK nr. 09098104,
WMR BTW nr. NL 8113.83.696.B16.
Code BIC/SWIFT address: RABONL2U
IBAN code: NL 73 RABO 0373599285

Wageningen Marine Research accepts no liability for consequential damage, nor for damage resulting from applications of the results of work or other data obtained from Wageningen Marine Research. Client indemnifies Wageningen Marine Research from claims of third parties in connection with this application. All rights reserved. No part of this publication may be reproduced and / or published, photocopied or used in any other way without the written permission of the publisher or author.

A_4_3_2 V29 (2019)

Contents

| | |
|--|-----------|
| Tracking digital footprints in Bonaire's landscapes | 1 |
| Summary | 4 |
| 1 Introduction | 6 |
| 1.1 Social media and the applicability in spatial -temporal research | 6 |
| 1.2 Bonaire studies with big data | 6 |
| 1.3 Aim of this study | 6 |
| 2 Methods | 8 |
| 2.1 Website and data targeting | 8 |
| 2.2 Data collection | 8 |
| 2.2.1 Search | 8 |
| 2.2.2 Photo categories | 8 |
| 2.2.3 Spatial grid | 9 |
| 2.3 Data processing | 9 |
| 2.3.1 Photo User Day intensity | 9 |
| 2.3.2 Spatial characteristics | 9 |
| 2.3.3 User groups identification | 10 |
| 2.3.4 Data analysis | 10 |
| 3 Results and discussion | 11 |
| 3.1 Photo User Days: How many, what and when | 11 |
| 3.1.1 How many | 11 |
| 3.1.2 What | 11 |
| 3.1.3 When | 12 |
| 3.2 Where: Overall distribution and intensity maps | 14 |
| 3.3 By whom: Type of photographer and origin | 18 |
| 3.3.1 Origin of photographers and their interest | 18 |
| 3.3.2 Type of photographer | 21 |
| 3.4 Comparison with earlier study | 22 |
| 4 Conclusions and recommendations | 24 |
| 4.1 Research aims | 24 |
| 4.1.1 Evaluation of the reproducibility | 24 |
| 4.1.2 Distribution patterns: What, Where, When, by Whom | 24 |
| 4.1.3 Proxy for behaviour and changed patterns | 25 |
| 4.2 Methodological aspects | 25 |
| 4.2.1 Tagging photos and machine learning | 25 |
| 4.2.2 Future of this type of research | 26 |
| 5 Quality Assurance | 27 |
| References | 28 |
| Justification | 29 |
| Annex 1 Assigning categories | 30 |

Summary

Introduction and aims

With the introduction of smart phones that also take photos combined with GPS tracking applications, more tourists are able to take more geo-tagged photo's during their travels. In combination with the options to upload these photos to online platforms, new ways of collecting data from the Internet provide new research opportunities. These digital footprints, combined with their specific meta-data regarding geo-location the data provide much information to be of use in spatial-temporal research.

A previous study on Bonaire showed the potential of this kind of research. It showed some understanding of the spatial movement of tourists and the number of tourists that visit different parts of Bonaire. In order to use these kinds of data in e.g. potential impact studies, we aimed to give a follow up.

In the present study the approach from Schep *et al.* (2016) was revisited with the following goals:

- Update distribution maps with the latest data (2016-2020) and evaluate the reproducibility of the maps.
- Detect whether distribution patterns and densities has changed following the recent developed trails and thus tourist spreading.
- Study if densities at specific locations can be related to local characteristics, such as:
 - Spatial characteristics such as distance roads
 - Landscape characteristics (landscapes)
 - Tourist type (cruise versus stay-over)

Methods

For this study, FLICKR was the only online platform to collect photographs and their metadata. Others were no longer available or unsuitable.

All photographs taken between November 2002 and October 2019, within a Bonaire surrounding bounding box were collected, including their meta-info. This resulted in 13026 photos, coming from 421 photographers.

Using a self-built Python application "PhotoCategoriser" each photo was assigned to a category (coastal, seascape, wildlife, underwater, terrestrial, other). Metadata of each photo and the assigned category allowed to analyse on origin of the photographer, to estimate the type of tourist (cruise or stay-over), their interest, and differences in spatial and temporal distribution. The specified resolution to aggregate the data was set at grid cells with a mean surface area of 0.301 km².

Photographer intensity is determined by condensing photographs into Photo User Days (PUD). One PUD stands for one (or more) photographs taken on a given day by specific photographer for a category in a grid cell.

Results and conclusions

Overall results

- The report provides various figures and maps presenting the spatial distribution of PUD as a proxy for tourist distribution. Temporal aspects in PUDs reflect the annual dynamics in tourist numbers.
- Trends in tourist numbers are not equally reflected in the numbers of PUDs. PUDs are therefore a proxy of tourist distribution, but not a strong indicator for trends in absolute numbers and intensity.

Reproducibility

- The additional ~ 4000 photos on top of the estimated ~ 10.000 FLICKR photos that were analyzed by Schep did not add much extra information. Also, the applied resolution did not refine the possibilities of performing risk assessments on habitats or species due to the limited number of data in those areas. Distribution patterns and intensity trends were similar. Category distribution however differed slightly. This can be explained by the differences in used datasets, and by the boundary criteria for assigning categories.

Detection of (changed) distribution and relation to local characteristics

- The overall distribution of PUDs shows higher intensity along the west coast, near Kralendijk and its tourist area. In addition, some higher intensity spots are visible near Sorobon in the east, and Seru

Largu in the middle of the island. The hotspots such as Goto and Washington Slagbaai in the north are clearly highlighted, as well as several scenic spots along the southern flats (Salt pans, Slavery houses, Lighthouse).

- Less frequented regions mainly include landscapes on the eastern part of the island. The low numbers of PUDs in these regions did not allow additional analysis on changes in distribution of PUD between years. It was considered not to be of added value. Hence, the effectiveness of the recent established trails could not be assessed any further, and an additional preliminary risk assessment for habitats or species was left out of the study.
- Distribution of PUDs reflect mainly the accessibility of regions: hence the roads and hotspots are clearly visible, and only limited PUDs were plotted further away. Analysis of distribution of tourists in specific habitats or nearby certain living areas of species were therefore considered not to be of added value.
- The interests (reflected by the categories) of the photographers slightly vary over the years and within a year, both by origin and by tourist type (also reflected by the cruise season). Also, the distribution and intensity of tourist types and origins seems to slightly vary. Details are provided in the report.

Future application and methodological issues

- We suggest that studies that use these data sources first look into the generic distribution and intensity of photo's PUDs collected (data coverage) before taking the effort of categorizing. Based on the general overview, following analysis steps such as categorizing and environmental risk assessment could be added.
- Manual assignment of categories to photos is a subjective exercise. Assigning categories requires strict criteria and midterm evaluation of results.
- Online platforms are variable in their existence and terms of use, leading to an uncertain accessibility and application of these kinds of data in future studies.

1 Introduction

1.1 Social media and the applicability in spatial -temporal research

With the introduction of smart phones that can take high-quality photos, and with increasing application of GPS and Wi-Fi on digital cameras, more tourists are able to take more photo's during their travels without the needs of large equipment. In combination with the recent development in social media and upload possibilities of these photos, new ways of collecting data from the Internet provide new research and analysis opportunities. These digital footprints include for example uploaded photos and combined with specific meta-data regarding geo-location the data provide much information to be of use in spatial-temporal research. Research using geo-tagged photos e.g. focus on the application in the mapping of tourist densities, taken routes, places visited, duration of stay, and the attractiveness of the places visited (what was on the photo) (Önder *et al.*, 2017; Li *et al.*, 2018). Hence, the spatial- temporal dynamics of tourism can be studied in more detail. These analyses can aid to e.g. adjust marketing strategies and implement crowd control measures. We are interested in the use of such data and its derived information to estimate tourist's behavior, possible impact on nature, prevent impact and propose measures accordingly.

1.2 Bonaire studies with big data

For Bonaire, Schep *et al.* (2016) was the first to demonstrate how data from social media could be used to map the spatial distribution of tourists and characterise their interest. The aim of that pilot study was to gain an understanding of the spatial movement of tourists and the number of tourists that visit different parts of Bonaire. The focus was to identify the most popular ecosystems for locals and tourists, and which aspects of the environment were of interest.

The data and results provided an overview of the usability of these kind of data and results from it. Based on the results, the local government can get insights into the relative importance of locations for nature-based tourism and incorporate this knowledge in their nature policy, their economic policy related to tourism and their Spatial Development Plan (Schep *et al.*, 2016). The study of Schep *et al.* (2016) comprises the data collection within social media platforms between summer 2005 and February 2016. They made spatial analysis on the square meter, grouped photos into 6 categories, and evaluated amongst other things the spatial and temporal distribution of photos and relative differences in photo densities among various types of areas on the island. The study clearly illustrated spatial and temporal variation in photo/tourist density, illustrating touristic and natural hotspots on the island. This pilot study provided valuable input to support a sustainable tourism development in Bonaire. Another step forward is to use these kind of data and analyses to guide the management of human impacts on the natural environment.

1.3 Aim of this study

This study aimed to identify possibilities and limitations of using social media data to estimate the spatial and temporal distribution of tourists, intended to support sustainable tourism development. Since 2016 more photos should have been available, and cruise tourism has increased (see figures in chapter 3). Furthermore, development projects on the island (via Plattelands Ontwikkeling Programma's) added touristic routes and trails on the island to revitalise outback regions on Bonaire, such as the east side of the island. Changes in distribution might be detected via geo-tagged photo's.

In the present study the approach from Schep *et al.* (2016) was revisited with the following goals:

-
- Update distribution maps with the latest data (2016-2020) and evaluate the reproducibility of the maps.
 - Detect whether distribution patterns and densities have changed following the recently developed trails and thus tourist spreading.
 - Study if photographer densities at specific locations can be related to characteristics, such as:
 - Spatial characteristics (autocorrelation) such as distance to roads
 - Landscape characteristics (landscapes)
 - Tourist type (cruise versus stay-over)

2 Methods

In short, the methodology of using geo-tagged data comprises few basic steps. First the selection and targeting of websites and available data, second the collection of data, and thereafter a data check and cleaning before the intended analysis can start (based on Li *et al.*, 2018). In this study we follow the set-up from Schep *et al.* (2016) in order to be able to compare results.

2.1 Website and data targeting

Schep *et al.* (2016) used several different social media platforms to collect their data in 2016, namely: Panoramio, Flickr and Instagram. Unfortunately, Panoramio has been shut down and Instagram has changed their terms of use making it virtually impossible to collect photographs and their metadata on a large scale. Data from these sources can therefore no longer be collected in the same manner as has been done by Schep *et al.* (2016). In addition, the popularity and usage intensity of social media are variable over time. Also, the target audience may differ along each platform and may also change over time. All of these traits make these platforms unsuitable for structural analyses over longer periods of time and are limited to opportunistic studies, such as the present study.

To compensate for the loss of two data sources (Panoramio, Instagram) a search was done to identify possible replacement sources. None were found that offered a suitably large user group, sufficient geographical detail, and accommodating (legal) terms. Some options came close, but i.e. due to cost involved (paid service, access) could not be included within this study.

2.2 Data collection

2.2.1 Search

In the present study Flickr (<https://www.flickr.com/>) was the only platform used, because it provides an Application Programming Interface (API) that allows for the extraction of detailed information. The R language was used to set up a script to extract data from Flickr. As Flickr's information query only allows for a limited data amount per request, a separate search per upload year was performed. For each upload year, all photographs taken within a Bonaire surrounding bounding box for Bonaire (between -68.56 and -68.03 longitude and 11.92 and 12.43 latitude) were collected, including their meta-info (such as photo owner). The data collection covered a period starting November 2002 and ending October 2019.

2.2.2 Photo categories

A Python-application 'PhotoCategoriser' ¹was built, based on a photo-viewer example, to assign a category to each photograph. The specifications of the 6 categories are described in Annex 1. The application locates the first uncategorised photograph from a SQLite-database holding the collected meta-data, downloads a temporary copy and shows it on screen for assignment.

¹ DOI: <https://zenodo.org/record/3873067#.XtZJODPiuCg>

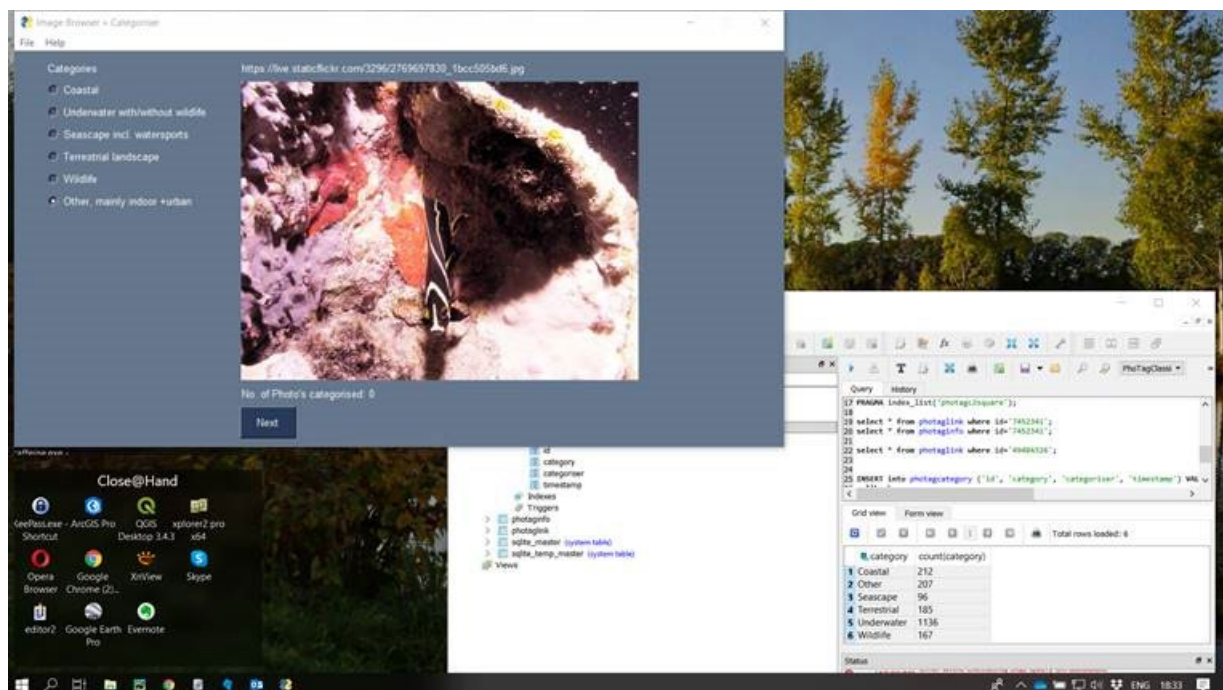


Figure 1. Screenshot of a photograph being categorised using the app.

With the PhotoCategoriser, in place, it took about two person days to categorise all pictures. It was noticed that often a photographer made more than one picture at a given location but that these frequently were put in different categories: 'I was here' (Other), picture of the beach (Coastal), picture of yacht sailing by (Seascape). Based on this, the perceived intention of the photographer clearly influenced the category assigned to the photograph. This may complicate things with respect to different approach that was considered: artificial intelligence/computer vision/machine learning. Such subtle shifts in intention may not be suitable for automated processing.

The category results allow for additional analyses on main interests by photographer types, origin, or season for example.

2.2.3 Spatial grid

For spatial analyses a C-squares (concise spatial query and representation system) grid covering the island is defined with a 0.005 degree resolution. This grid easily allows for aggregation at coarser resolutions at a later stage. At the specified resolution, grid cells have a mean surface area of 0.301 km². This spatial resolution is thus more accurate than the square kilometre used in the analysis of Schep *et al.* (2016).

2.3 Data processing

2.3.1 Photo User Day intensity

Photographer intensity is determined by condensing photographs into Photo User Days (PUD): one PUD stands for one (or more) photographs taken on a given day by specific photographer for a category in a grid cell. So, with six possible categories there remain at most six PUD. The record does contain data on how many original photographs were available.

The intensity maps show the PUD count across photographers and years for a given category per grid cell, or across all categories combined.

2.3.2 Spatial characteristics

Several spatial characteristics can be determined for each grid cell.

One option that was tested is: Distance to road. This has been calculated for each photograph position using a Bonaire Roads dataset dated 2013 (sourced from DCBD.nl). The cartesian distances are calculated in meters

using the UTM N19 WGS84 projection. And was summarised to min/max/avg at the same time the data was condensed to PUD.

Additionally, PUD intensity has been related to landscape types (Verweij *et al.*, 2020).

Other unattempted options include distance to shore, presence of types facilities etc.

Spatial accuracy in relation to assigned photo categories, have proved to be a potential problem of limited size. Once assigned a category (see also Figure 9) a number of photos shows up in locations that are at odds with the category. A data cleaning options has been investigated and the results are visible as hollow squares in the middle of the island, which clearly is a problematic place to take underwater photographs. In the end the number found was low and not used beyond identifying the potential problem (Figure 9).

2.3.3 User groups identification

Photographers were categorized in two ways: Region of Origin and Photographer Type.

Region of Origin was determined based on available metadata on owner location, and where possible an assignment was made to either: Local (only Bonaire), Latin America (including other Caribbean islands), North America, Europe or Other. Other also comprises photographers of which the origin could not be determined.

To distinguish the type of photographer (focusing on cruise tourist, stay-over, or local resident) we set criteria to try to assign each photographer to a group. The temporal and spatial variation within the uploaded photos can provide a proxy of type of tourists and can help to identify differences in behavior between tourist groups. The length of the period from first to last photo taken (per year) for a given photographer was taken as a measure for the length the stay and from that to assign a category (Table 1).

Table 1. Criteria to assign photographer type

| Length of Stay | Photographer Type | Reasoning |
|-----------------------|-------------------|--|
| <= 0.5 day | Cruise | Cruise ships visits last for ca 12 hours (0.5 day) within which passengers also must debark and re-embark. They may however also take a few pictures on arrival or at departure. |
| >0.5 day to <=30 days | Stay-over | Most stay-over tourist spend a week on the island, but a few may stay for a couple of weeks longer. |
| >30 days | Local | |

2.3.4 Data analysis

Statistical analysis on data such as correlation between time, photographer type, and categories could be modeled via statistical models were not included in this phase of the project. The project was time limited and thus contained a fist exploration.

3 Results and discussion

3.1 Photo User Days: How many, what and when

3.1.1 How many

The number of photos downloaded from Flickr were 13026 individual photos. The photos taken by 421 individual photographers resulted in 4062 photo user days (PUD) to describe the distribution of people-, and the likely interests they have. This selection was thus an ~ 4000 extra photos taken in the years 2016-2019 on top of the selection by Schep *et al.* (2016) which covered photo's till February 2016. Schep *et al.*, (2016) obtained a total of 2797 PUDs. The extra years covered in this study resulted in an addition of 1265 PUDs.

3.1.2 What

Each photo was assigned to 1 category which resulted in the total numbers and relative comparison presented in Table 2. The (estimated) contribution of each category in the study of Schep *et al.* (2016) differs from our findings (Table 2). Largest difference is seen in the allocation of photos to the category "Landscape". A clear explanation for this difference is not easy to provide. Similar differences in tagged categories were also observed during mid-term evaluations during the photo tagging phase in this project among two team members. It became clear that previously agreed criteria could be interpreted differently (subjective aspects). After evaluation, these could be restored. This observation shows that assigning a category to a photo could be subjective and should be evaluated and discussed among a team in order to prevent these variances. The difference between Schep *et al.* (2016) results and this study could also be a result of a slight difference in the assessment criteria for a given category, or the interpretation of these. Another explanation could be the photo collection: its origin, size and time span differs largely and could influence outcomes too.

Table 2. Category sizes among the two studies.

| Category | Schep et al (2016) | This study |
|--------------------|--------------------|------------|
| Seascape + coastal | 26% | 31% |
| Underwater | 15% | 20% |
| Terrestrial | 25% | 12% |
| Wildlife | 10% | 16% |
| Other | 24% | 21% |

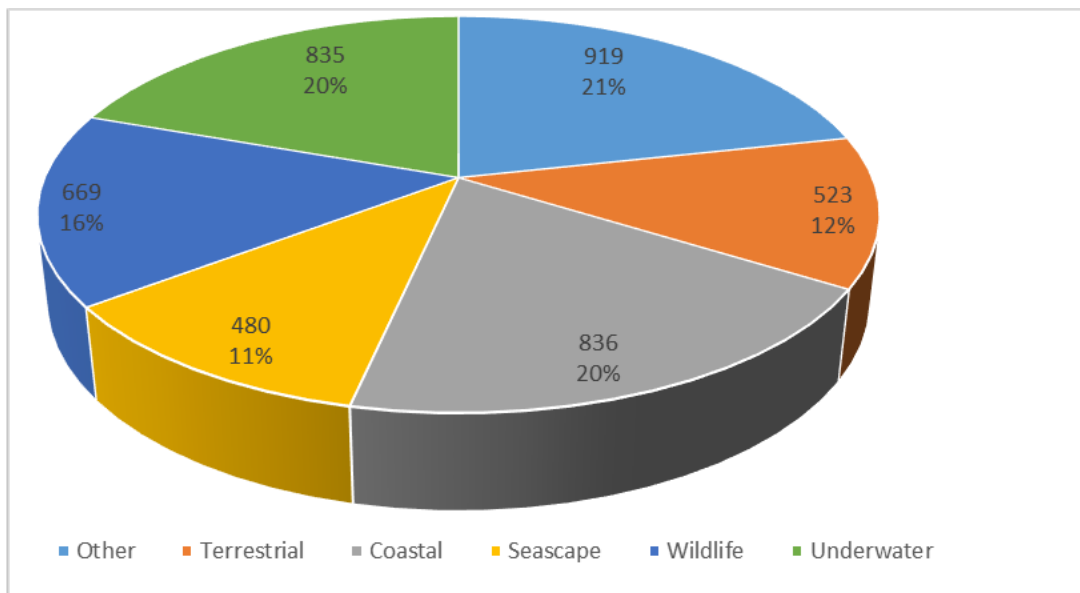


Figure 2. Total photo user days (PUD) per category (relative and absolute numbers given in chart).

Most photos taken are nature related (48%) divided in wildlife, underwater and terrestrial landscapes. Seascape and coastal photos are the next most popular category, and account for 31% of the photos taken. The other category consists of street views, buildings, restaurant meals and gatherings, and cultural happenings such as carnival. Depending on the interest of the study, the “other” category could be subdivided into more distinct categories.

3.1.3 When

The uploaded photos at Flickr vary over time (years as well as months) (Figure 3, Figure 4). This variation could likely be explained by the varying numbers of tourists over the years, but is however not the case when looking at CBS data on incoming tourism (Figure 5). Based on tourism data available for 2015-2019, no clear coinciding pattern becomes apparent with tourist numbers over time (Figure 5). The increase in tourist numbers, with roughly double the number of cruise tourists from 2017 onwards, is not reflected by an increase in PUDs. Flickr (and other photo-posting social media) may have become less popular. Or with changes to terms and conditions recently (privacy terms) e.g. less emphasis on free and more on paid accounts may also play a part.

Within a year, the number of PUDs also varies over the month, with July having the lowest numbers. Overall, May-September are months in which lesser cruise tourists visit Bonaire which might explain the relatively lower number of PUDs in this period. This generic “dip” is also seen in the number of PUDs between May-September. In case of equal contribution of all PUDs over each month, May-September should have contributed with 42% of the number of PUDs (5/12). Instead, they contribute with (on average) 36% of the PUDs, slightly less.

The interests of the photographers vary over the years (Figure 3). Within a year (Figure 4), the relative interest is also variable. Comparing the relatively even distribution across the year of stay-over tourist (arriving by air) and the seasonal character of the cruise arrivals, a stronger seasonality in the PUD data might be expected (not further analyzed yet).

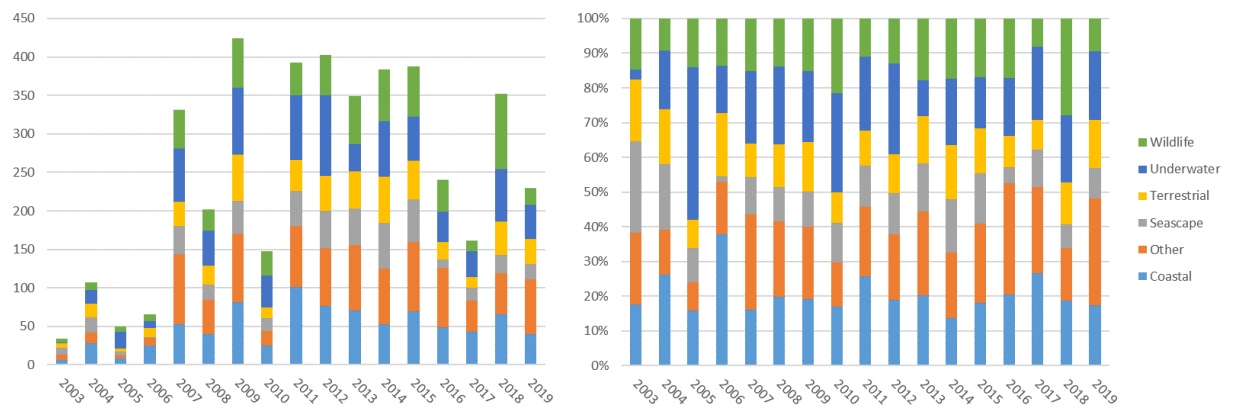


Figure 3. Overview of number of PUDs over the last 16 years, divided by photo-category. Left: absolute PUDs over time per category. Y - axes : number of PUD. Right: relative PUDs over time per category.

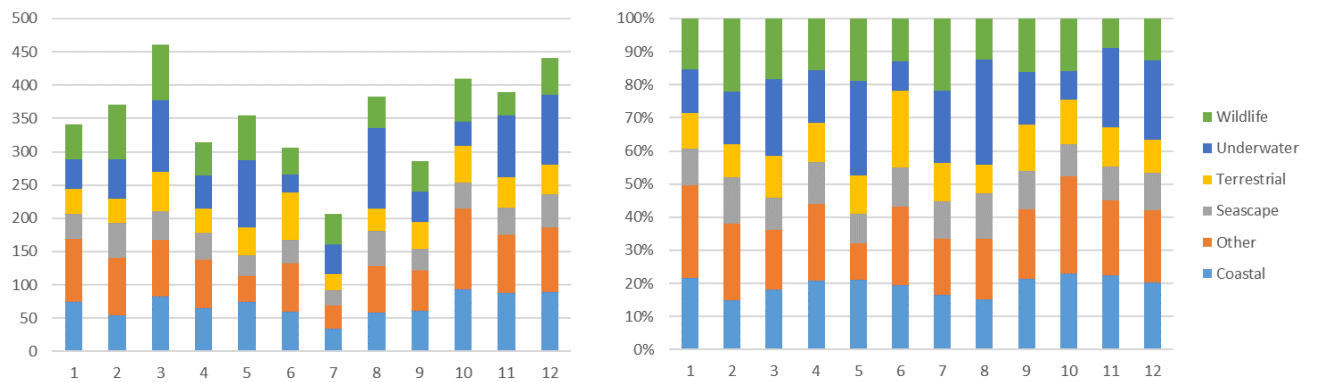


Figure 4. Overview of number of PUDs over a year (x-axes= months), divided by photo-category. Left: absolute number of PUDs, right: relative distribution of PUDs.

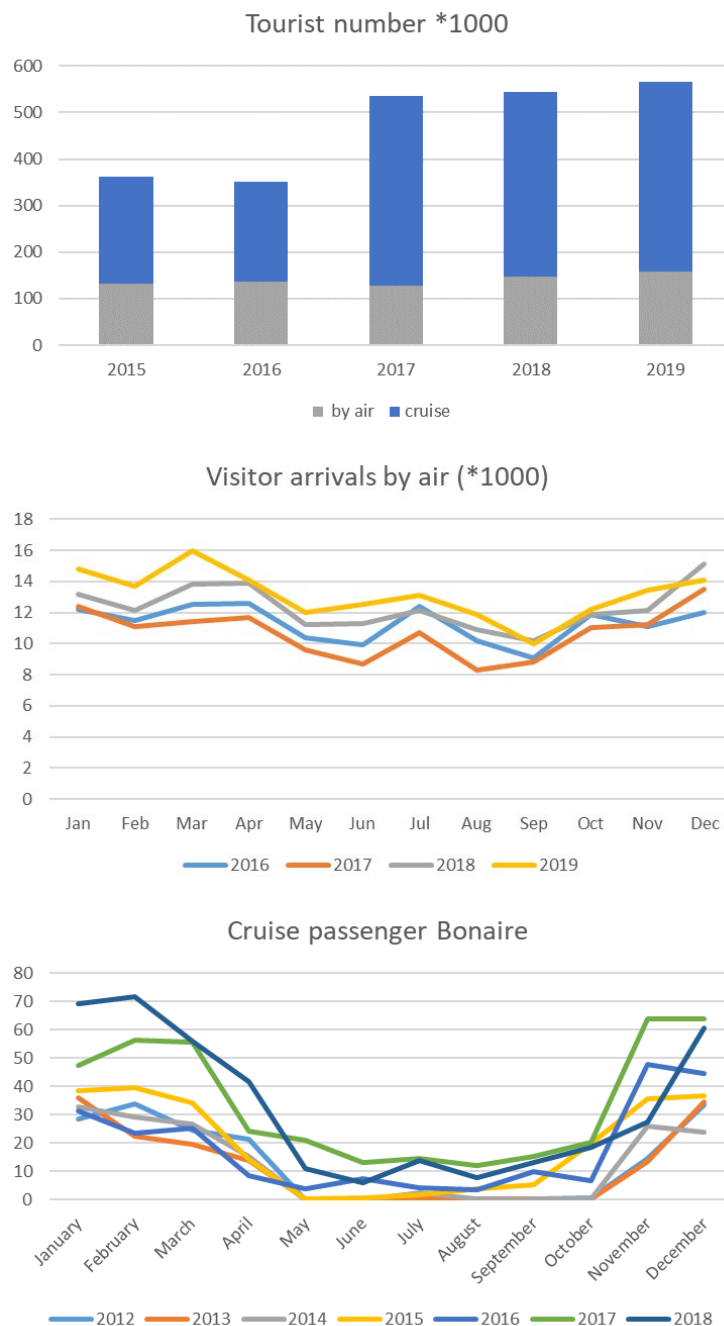


Figure 5. Top: Number of tourists visiting Bonaire over 2015 – 2019, (*1000) based on CBS data (CBS 2019), presenting cruise tourist and tourists arriving by air (stay-over tourist). Middle: Annual variation of visitors by air: Bottom: Annual variation cruise visitors (CBS data various year reports)

3.2 Where: Overall distribution and intensity maps

The overall distribution of PUDs (and thereby a proxy for tourists) (Figure 6) shows a higher intensity of PUD's along the west coast, near Kralendijk and its tourist area (Figure 7). In addition, some higher intensity spots are visible near Sorobon in the east, Seru Largu in the middle of the island. The hotspots such as Goto and Washington Slagbaai in the north are clearly highlighted, as well as several scenic spots along the southern flats (Salt pans, Slavery houses, Lighthouse). A map for specific land marks is presented in annex 2. The table in Figure 7 also shows that limited total number of PUDs is available for the various landscapes. Up to ~ 300 PUDs for the Northern Hills can be plotted, but for all other landscapes, the number is even less. Note that these totals account for all years (17 years) included in the database, resulting in a highest PUD availability of ~ 20/year, but for most landscapes far less. Based on this observation, additional analysis on changes in distribution of PUD between

years was considered not to be of added value. Hence, the effectiveness of the recent established trails was not assessed any further.

The intensity of all photo categories is plotted in Figure 8, and it shows the positions of the PUDs per category. Both figures (6 and 7) show a clear pattern:

- Each category seems to follow a patchy pattern, including the roads and Kralendijk. The latter can be explained to the fact that most hotels are situated there, and all types of photos can be taken from that location (wildlife e.g. in hotel gardens).
- Per category:
 - o Other: Kralendijk and Rincon, explained by the presence of roads and by the cultural aspects of these places.
 - o Seascape: Kralendijk and Sorobon, explained by the scenery at site.
 - o Coastal: besides Kralendijk, the roads along the west coast, the north and south are highlighted. Not many photos are taken along the east coast. This could be explained by the fact that it's a dirt track and very likely more daunting in appearance than most drivers would feel comfortable with. Especially in a rental.
 - o Underwater photos are taken near the coast, obviously. Some dive sites of interest are clearly highlighted such as the Hilma Hooker and Salt Pier in the south, and Buddies, Karpata and Andrea in the middle.
 - o Terrestrial: Mostly taken along the west coast at the salt pans, Goto, and Slagbaai park.
 - o Wildlife: PUDs are scattered over the island, concentrated by the roads and places of interest.

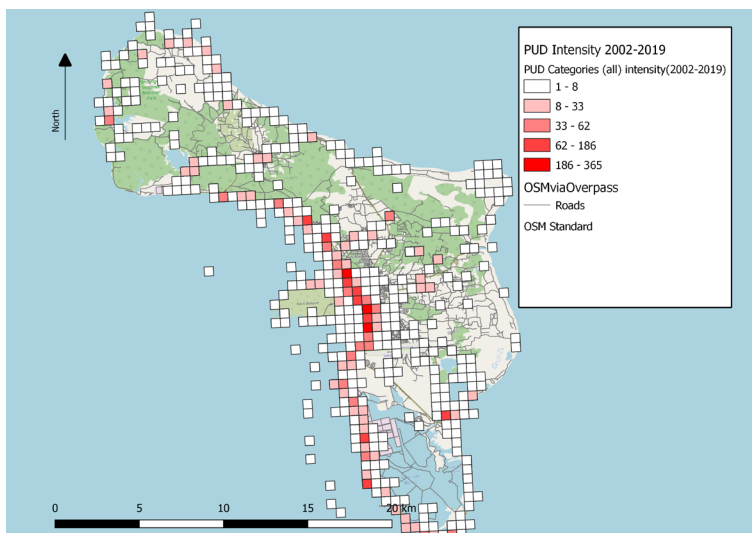


Figure 6. Spatial distribution and intensity of all PUD based on Flickr downloads (2002-2019).

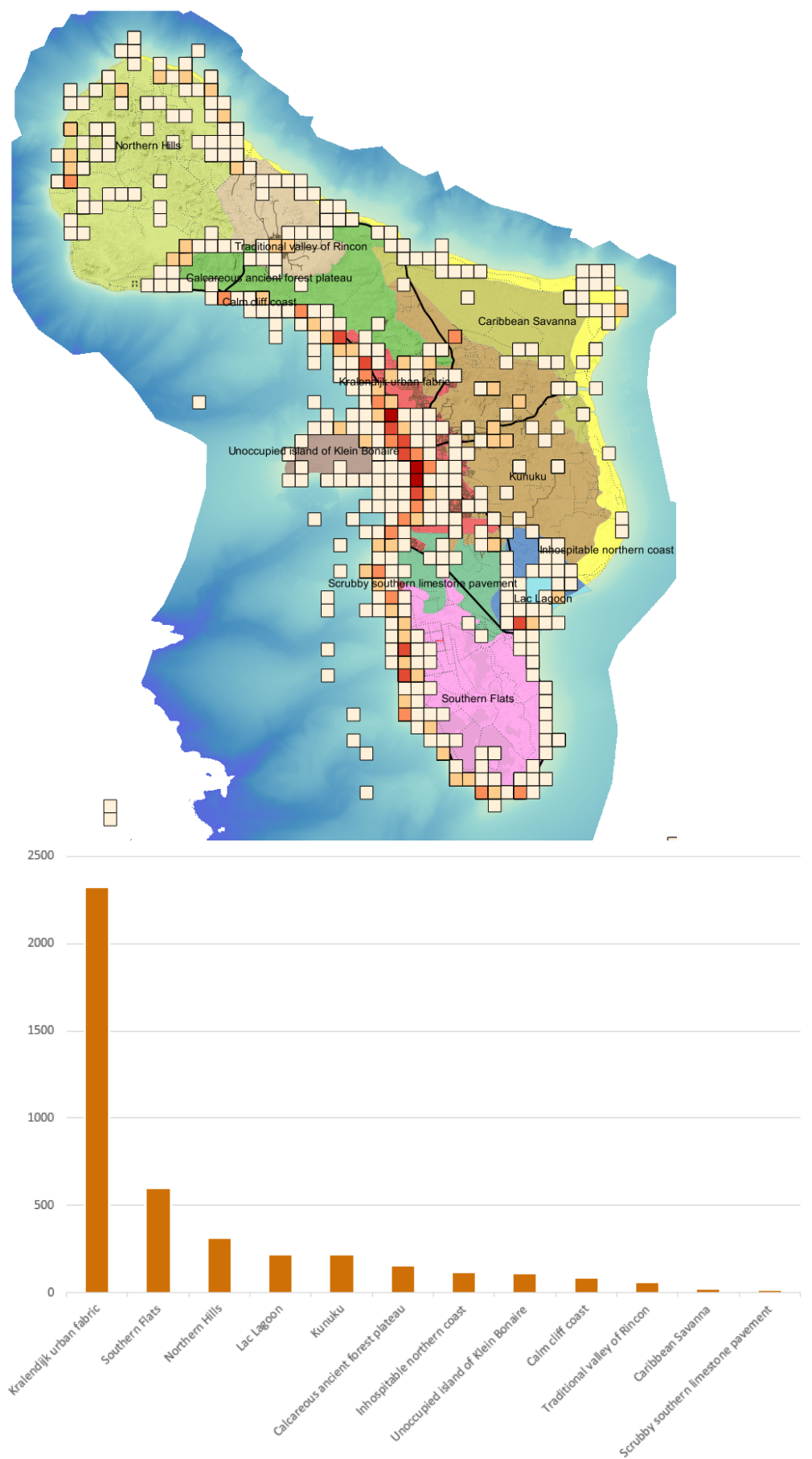


Figure 7. Above: Spatial distribution and intensity of PUDs in combination with a selection of Bonairean landscapes (Verweij et al., 2020). Below: Total PUDs per landscape.

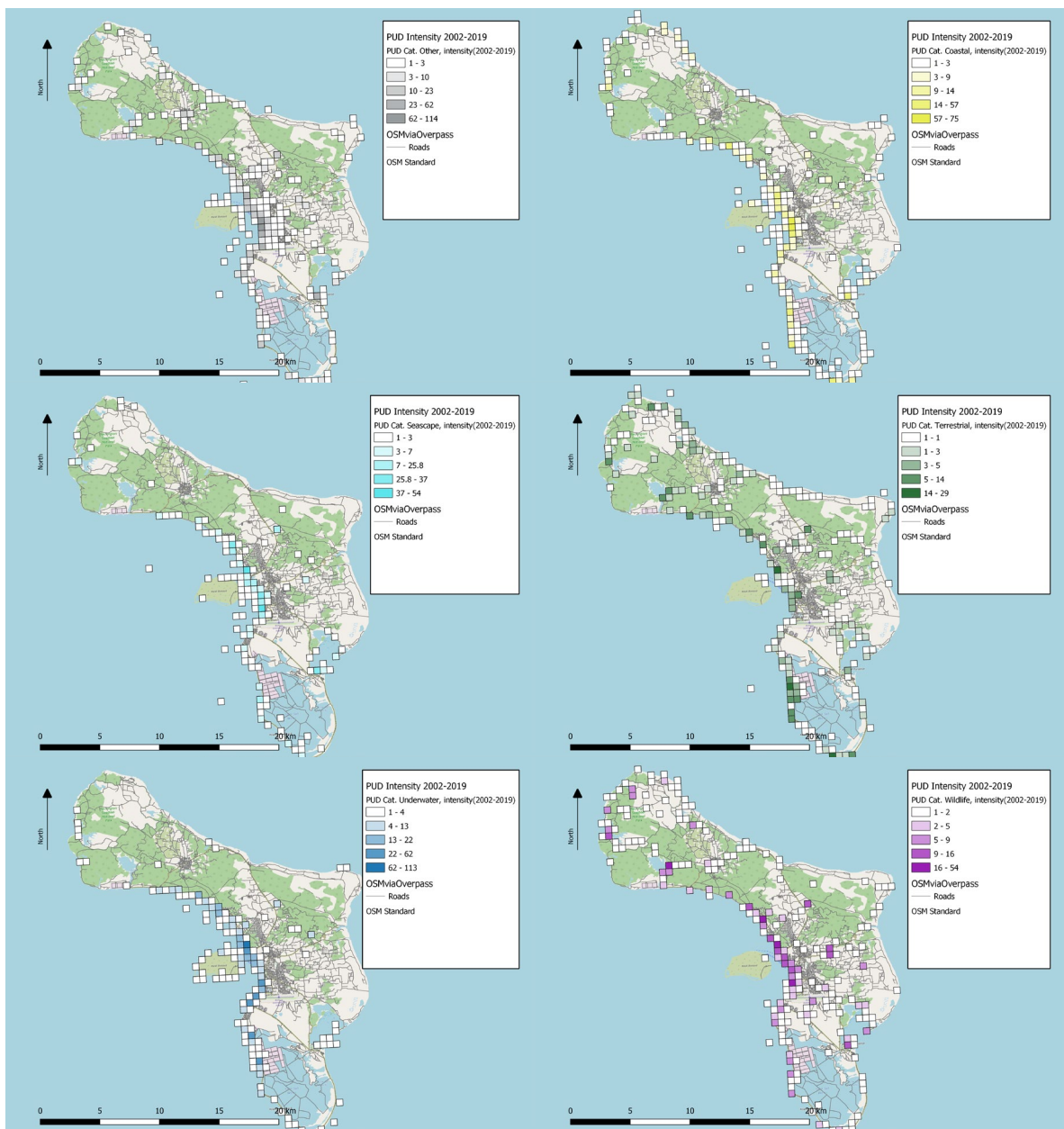


Figure 8. PUD intensity per photo category (Other, Coastal, Seascape, Terrestrial, Underwater, Wildlife).

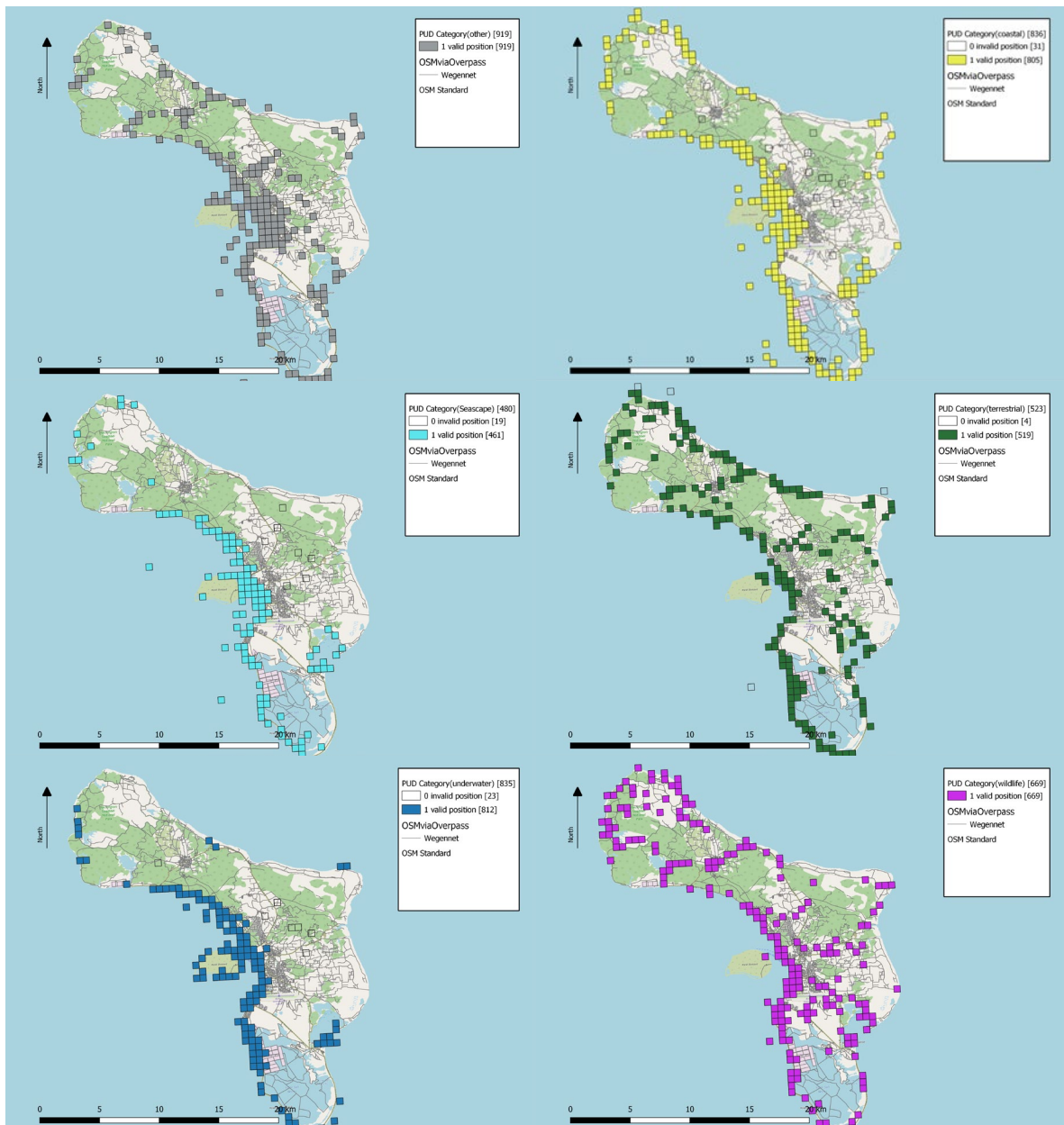


Figure 9. Spatial distribution of PUDs per category (Other, Coastal, Seascape, Terrestrial, Underwater, Wildlife). Invalid position marks the position of a photographer that is likely to be false combination (see method section).

3.3 By whom: Type of photographer and origin

3.3.1 Origin of photographers and their interest

Most photographers (that upload to Flickr) come from North America (USA and Canada) (39%), followed by "other" (mostly unknown) (28%), and Europe (26%). The places they visit and photograph do not differ much, it is the intensity that differs (Figure 9, and Figure 11), which can also be a result from their lower number in total. Locals contribute with only a few PUDs, and this contribution patterns cannot be used for a definite evaluation. The same is true for Latin American PUDs, which contribute with much smaller numbers compared to the other origins.

The interest of photo categories per photographer origin varies slightly: Photographers from Latin America seem to take relatively less wildlife and underwater photos compared to other groups, but

more coastal photos. Underwater photos are taken relatively more by Europeans and North Americans. A hypothetical option for Latin Americans taking fewer wildlife pictures is that local wild life on Bonaire is more familiar to them and thus less interesting as a topic for a photograph.

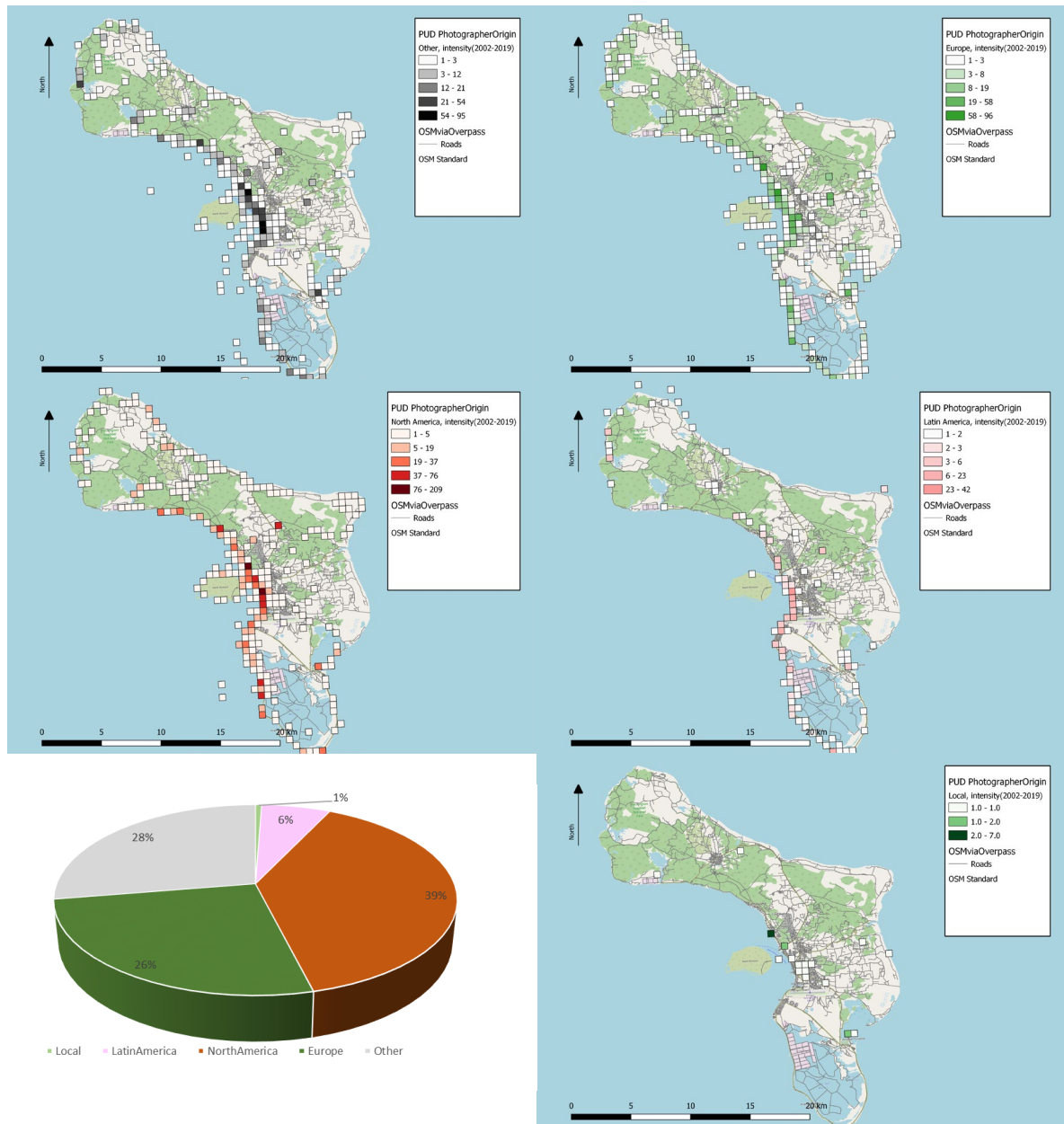


Figure 10. Spatial distribution and intensity of PUD divided by origin of the photographer (Local, Latin American (including Caribbean), North America (including Canada), Europe and other (also including unknown origin)).

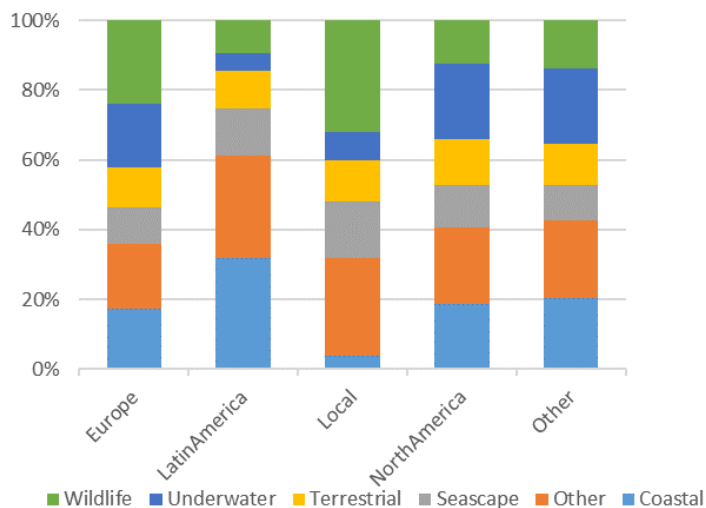


Figure 11. Relative interest of photo categories per photographer's origin. (Local, Latin American (including Caribbean), North America (including Canada), Europe and other (also including unknown origin)).

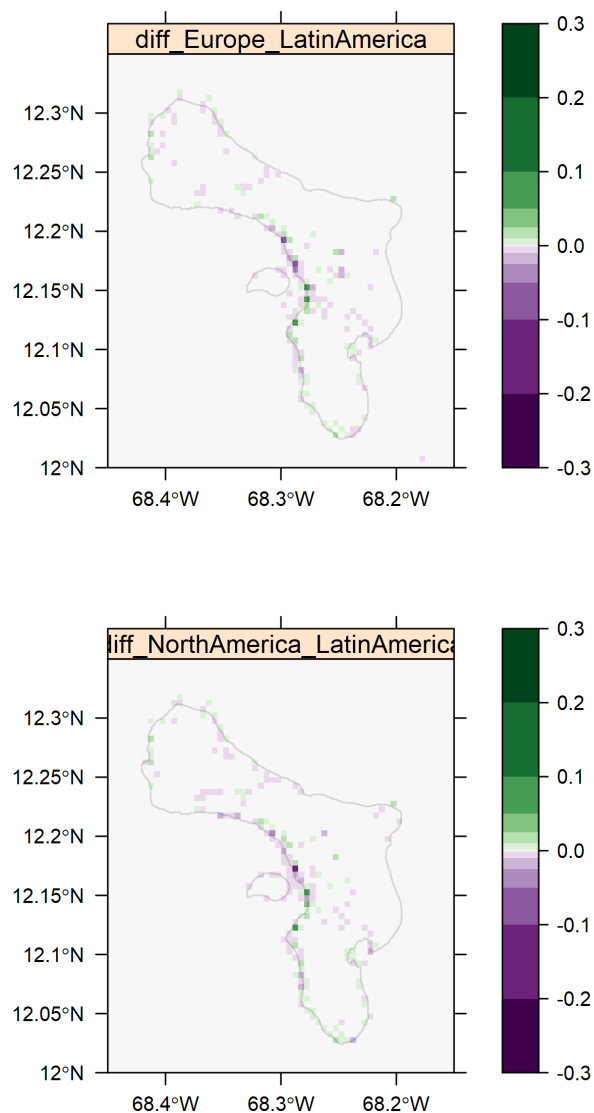


Figure 12 Difference in PUDS for European vs Latin American and North American vs Latin American photographers. Green marks relatively more Latin American tourists, purple marks relatively more European (upper figure) or North American (lower figure) tourists.

3.3.2 Type of photographer

Stay-over tourists account for 50% PUDs, and cruise tourist for 40%, PUDs from local photographers only account for 10% (Figure 12, Figure 14). And although it seems that the spatial distribution of each type of photographer is the same (Figure 12), an additional analysis shows slight differences (Figure 13).

Stay-over tourists seem to visit relatively more locations that require more time to reach. Typical places such as Slagbaai Park, dive sites as Karpata at the west, Sorobon, and Klein Bonaire are photographed more often (an indication of more visits) by stay-over tourists than by cruise tourists. Also, typical accommodation sites as Belnem- south of Kralendijk - are highlighted by stay-over tourists instead of cruise tourist. Cruise tourists seems to be more localized around Kralendijk itself, which can be explained that most cruise tourists stay near the boat or at least pass by after leaving the ship. The intensity of stay-over related PUDs seems to spread more inland and to the Northern Hills of Washington Slagbaai. This can be explained by the fact that a tour to Slagbaai park often takes a full day. Cruise tourist often visit the island in a shorter time frame (leaving the boat earliest around 8AM, and returning before 15AM).

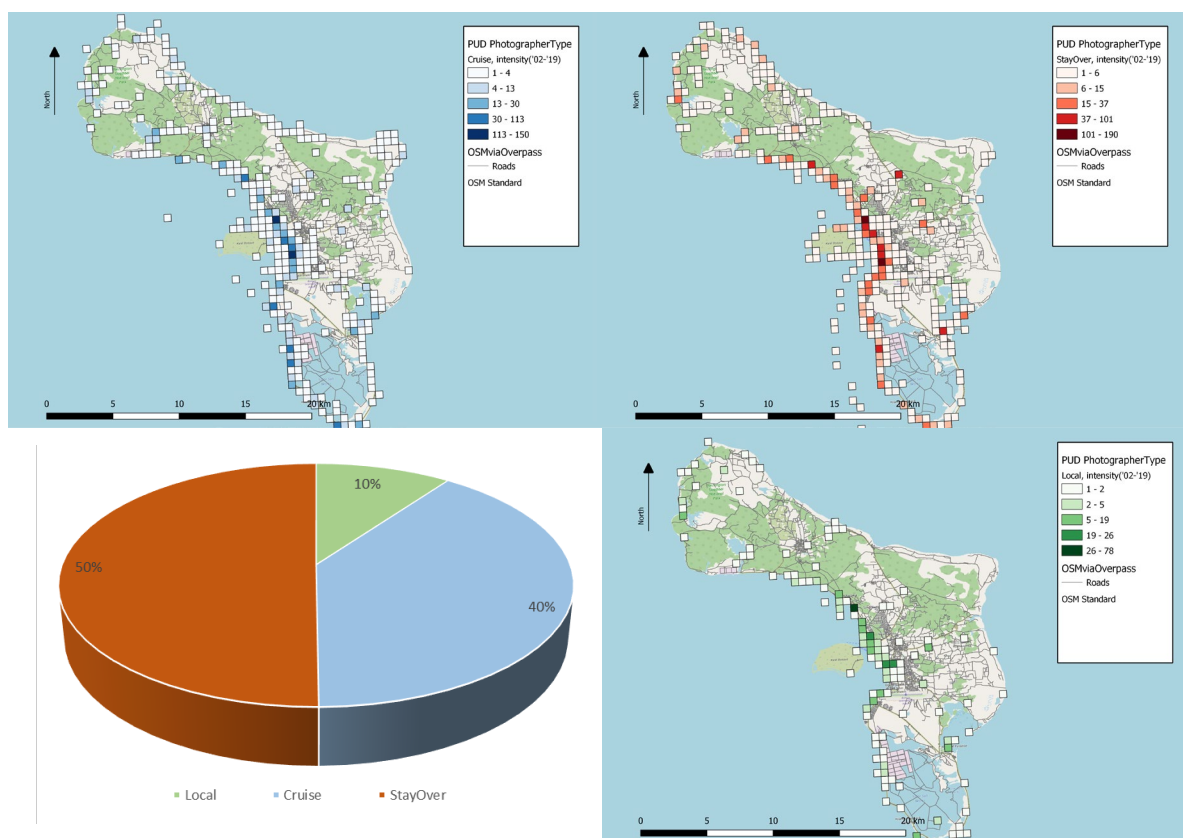


Figure 13. Distribution and intensity of PUDs by photographer type.

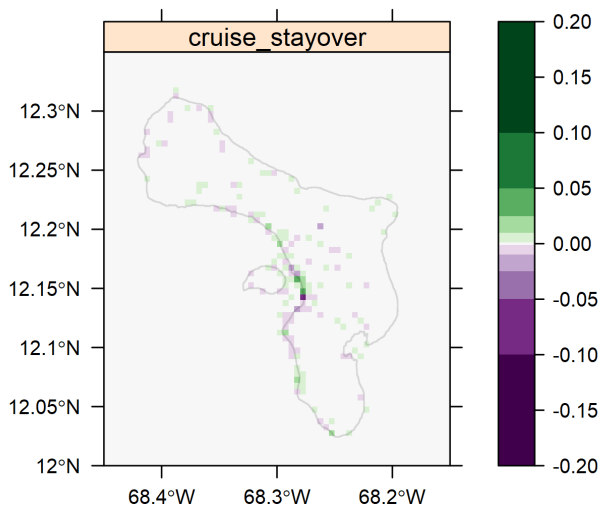


Figure 14. Difference in PUDs between cruise and stay-over photographers. Green marks relatively more cruise tourists, purple marks relatively more stay-over tourists.

Photographer interests per photographer type is presented in Figure 15. Interest of photo-objects do not seem to vary much between photographer type. Stay-over tourists and locals seem to make relatively more underwater photos compared to cruise tourists which can be explained by the time needed to go diving. Cruise tourists also seem to take less wildlife photos. Cruise tourists take relatively more “other” photos, which consisted of e.g. street views, selfies, buildings.

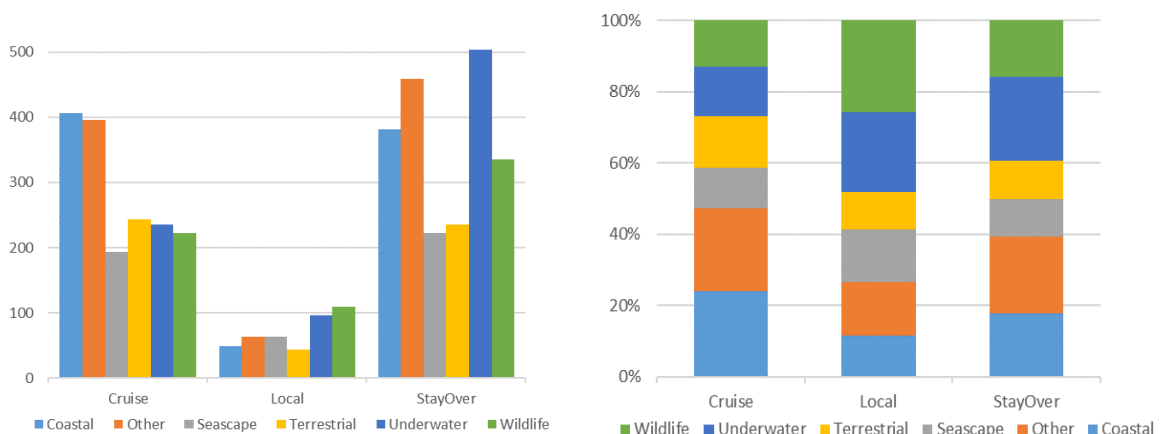


Figure 15. Contribution of photographer type by photo category Left: in absolute PUDs, right: relative contribution.

3.4 Comparison with earlier study

In Figure 15, data points from Schep *et al.* (2016) are presented (darker colour) in combination with data used within this study (lighter colour). For most categories there is a clear overlap for the majority of the locations. Especially for Flickr pictures, Panoramio users seem to have posted more inland terrestrial landscapes (mid, right). The current study has a somewhat wider coverage for categories Seascape and Wildlife with new locations mainly along the west coast (for Seascape) or more centrally on the island (Wildlife).

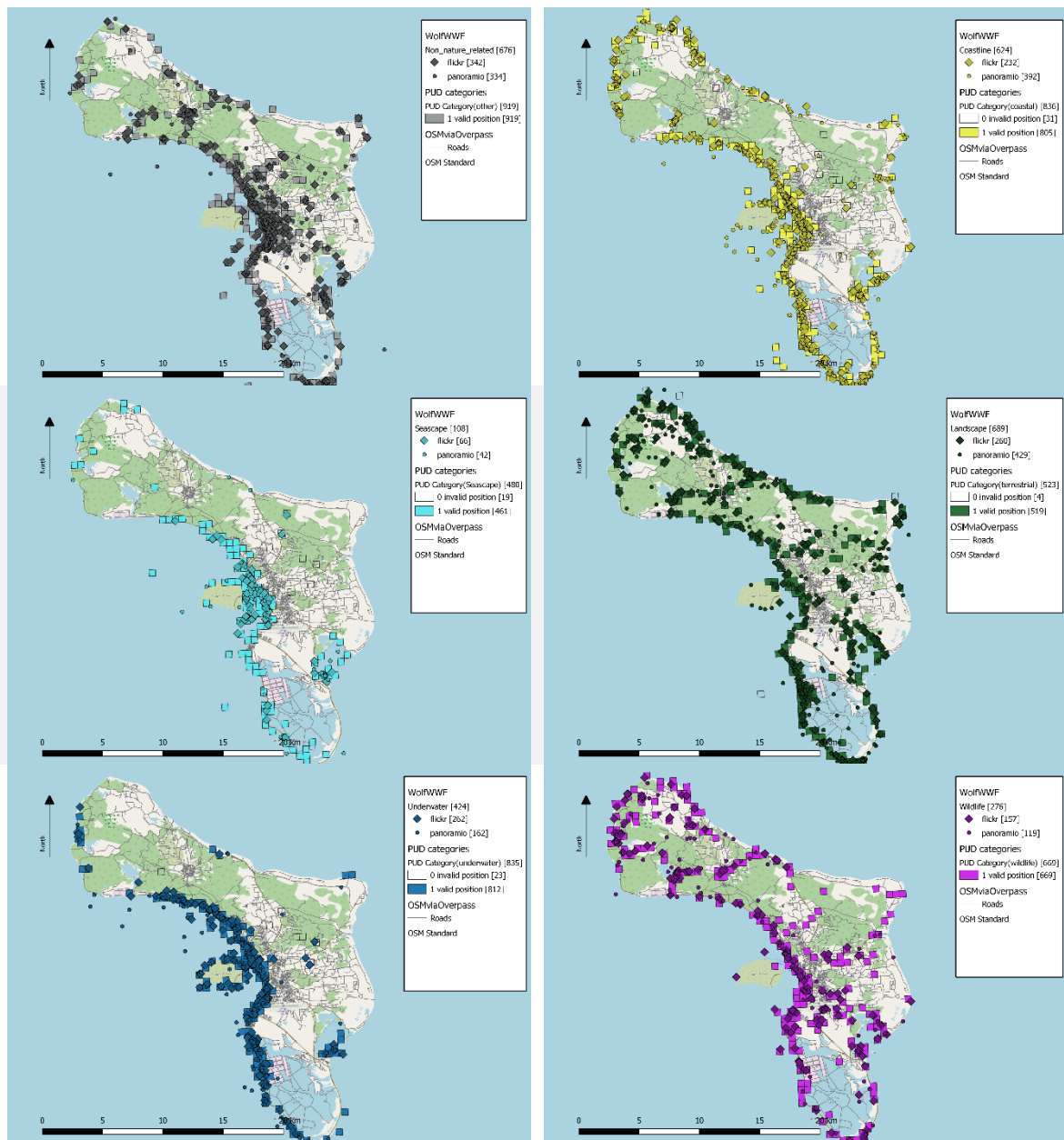


Figure 16. Comparison of the spatial distribution PUD current study (lighter colour) and the previous study (darker colour, (Schep *et al.*, 2016)). For the previous study the shape of the symbol identifies the data source (diamond for Flickr, dot for Panoramio and Instagram).

4 Conclusions and recommendations

In this project we aimed to identify possibilities and limitations of using social media data in order to estimate spatial and temporal distribution of tourists, intended to support sustainable tourism development. Since the study of Schep *et al.* in 2016 many more photos were available, and cruise tourism has increased.

We aimed to

- Update distribution maps with the latest data (2016-2020), and evaluate the reproducibility of the maps.
- Detect whether distribution patterns and densities has changed following the recent developed trails and thus tourist spreading.
- Study if photographer densities at specific locations can be related to local characteristics.

In this chapter we will conclude on the research aims, but also reflect on the methodological aspects of the study.

4.1 Research aims

4.1.1 Evaluation of the reproducibility

Although we have used more photos, covering more recent years, the distribution and density maps of Schep *et al.*, and those produced in this study do not differ much. Even the loss of two photo-sharing platforms did not considerably affect the outcome

While we have used a finer grid to present our data, this does not affect the overall distribution pattern or density per landscape either. An explanation could be that the density of PUD in low density areas will not discriminate differences using a finer grid (resolution). From this exercise, we can conclude that before applying a finer grid, it is useful to get a general idea of the overall spreading of PUDs.

4.1.2 Distribution patterns: Where, What, When, by Whom

Where

The spatial distribution and intensity maps showed clear patterns and hotpots of photographers which can be used as proxy for tourist patterns. The distribution maps showed a clear heterogeneity in the spreading of tourists. West coast, Kralendijk urban area (including underwater photo's), and southern flats were the most photographed areas, and all other landscapes far less, reflecting a low visitation of tourists, or less likable scenery to take pictures from.

What

Typical differences in interests were seen in the dataset used. Most photographed category is Seascape and coastal views, followed by underwater scenery or animals. "Other" represents a large number of photo's, and could have been refined more by adding e.g. a category "cultural aspects", reflecting typical Bonairean streetviews, historic buildings and people. We found the relatively low number of PUD in the category "terrestrial" remarkable, since Bonaire has much to offer on the landscape level. As this category differs largely from the outcome of Schep *et al.* this observation can also be a result of a bias in the criteria setting and steps in the assignment process.

When

The PUD numbers vary over time both in terms of years and months, explained by the variation in tourist arrivals and its trends in time. But although tourist arrivals (especially cruise tourism) has doubled, this is not reflected in uploaded photo's. Hence, the PUDs give an indication of tourist spreading over the island, but certainly not the overall trend in numbers (and its increase).

By Whom

The origin and type of photographer was used to achieve more insights into cultural interests. This can be of use for marketing, but also to educate or inform specific groups on basis of their interest and distribution over the islands. In case certain groups would concentrate e.g. near certain landscapes, targeted information can be send, and provided on site. A group, reflecting 28%, could not be given an "origin". This nearly one third of the data is a gap in the available information, that steers the analysis.

North Americans represent the largest group of visitors based on the remaining data.

Europeans and North Americans show the same distribution, and whereas some differences occur in interests (e.g. related to wildlife photo's) they have the same interest for underwater photography.

The data of type of photographer (cruise versus stay-over) indicated some slight pattern differences.

The differences seem to reflect the time needed to visit certain places (e.g. Slagbaai) or the activities to undertake. The overall number of underwater photos by cruise tourists was much lower while the absolute number of cruise tourists is much higher than that of stay-over tourists. Cruise tourists seem to take less photos of wildlife too, but relatively more coastal pictures. These observations could have value for targeted island marketing.

4.1.3 Proxy for behaviour and changed patterns

Intensity and spatial distribution patterns can help targeting potential risks for vulnerable habitats and or species. In this study, an attempt was made to mark photos taken further away from the road as a proxy for people scattering around "off-road" and potentially disturbing species or trampling plants.

The finer displayed grid and way of analysing, compared to Schep *et al.* (2016) was intended to be of added value to this exercise.

The actual numbers of photos taken further away from roads was however too low (data not shown), and not useful at this stage to analyse any further. If additional modelling detects patterns near vulnerable habitats and or species, educational programs can be targeted towards the photographer type (if specific groups are causing specific risks).

Also, no clear patterns in the east coast were detected (data not shown). The newly established routes and trails were looked into in order to potentially observe an increase of photos taken before and after measures taken (e.g. the spread of tourists). However, the PUD intensity in this region (e.g. reflected by the low numbers in the Caribbean savanna) was too low to detect any changes of importance.

4.2 Methodological aspects

4.2.1 Tagging photos and machine learning

Assigning categories requires strict criteria and midterm evaluation of results. Human assignment of categories to photos is a subjective exercise. Mid-term evaluation is therefore needed to asses potential differences in interpretation of criteria and to correct for these. We put more attention to the "perceived intention of the photographer" as a criterium. This "intention" can exist as subtle characteristics within a photo. For example, when a photo representing a landscape on land, would have been automatically tagged as "terrestrial" in most of the cases. It is as tagged as "wildlife" instead, in case we believed the photo was taken because of the purpose of photographing an animal such as a lizard or bird. However, these wildlife elements could be very small, and depending the photographer skill these elements are not always obvious. Before starting our human assignments, machine learning (ML) was considered a future better option. However, based on the subjective nature, and nuances within the picture that could make a difference in tagging outcome, ML requires a good background knowledge of photos posted, and compromise on the quality of the outcome, as we are not sure whether ML could easily be instructed to cover for these human interpretations.

Knowledge of the photo dataset provides potential refinement of categories to include in the tagging process. This asks for a first screening of many photo's, and to align those with the scope of the study. E.g.: within wildlife, a sub category of birds, iguanas, and life stock could have been added. Or within the "Bird" category, typical flag ship species as flamingos could have been tagged. In case of cultural

interests, photos from the category “other” could have been tagged with “cultural” or historical buildings, street views, or “people”. Depending the scope of the study, stakeholders (e.g. local governments, NGOs) could also be consulted for determining a suitable list of categories, as some subjects may be of specific interest.

4.2.2 Future of this type of research

Since the study by Schep *et al.* (2016), the free location-based photo-service Panoramio has been shut down. Flickr has been acquired by a different company (SmugMug) that has kept the service going for the last three years or so. Given these recent changes, it is unsure how for long these platforms will provide their (free) services and thus if this research will be possible in future. We believe that given the experience in this project, that although more and more people will upload data, the retrieval of these data will be limited, or has to be paid for. Hence, these kinds of studies might be restricted by that development.

Other sources did not offer sufficient geographic detail to go with the pictures to be useful for this type of study (Instagram, Facebook, Pinterest to name a few). From these sources pictures could potentially be extracted (scraped) based on other available mechanisms (i.e. tags). Enriching these pictures with a more accurate position could be attempted based on Artificial Intelligence/Machine Learning/Computer Vision-techniques. The training set would consist of pictures with a known location and a recognisable landmark and from this, similar pictures could be enriched with a more precise location. Pictures on Bonaire of landmarks such as the Cruise Ship pier, Seru Largu, Lighthouse, Salt Pier, Slave huts etc. are good candidates for such an approach. As would the landscape of Goto, the yellow marker rocks of dive spots, which also have a name on them, and the beaches of Slagbaai. Alternatively, the focus could be placed more on what (and/or when) rather than the where, easing the constraints on what social media can be included.

In addition to the above, with just one suitable (social media) source remaining, the initial number of photographs categorised appears high, but once boiled down to PUD there remain only just over four thousand observations from roughly four hundred people. Furthermore, there are only PUD on under 1500 separate days (n=1374). The dataset spans nearly seventeen years, more than six thousand days. Coverage is about 25% of the days. Effectively, the available sample in this study area is probably too small to enable answering all the questions that the study was aiming for. We haven't performed an analysis to assess the needed minimum number of PUDs. Other studies should be aware of these kinds of limitations before setting goals. We suggest that studies that use these data sources first look into the overall number of photos, the number of photographers, the generic distribution and intensity of photo's PUDs collected (data coverage also over the years) BEFORE taking additional steps such as too detailed aims or taking the effort of categorizing. Based on the general overview, following analysis steps such as categorizing and environmental risk assessment could be added.

5 Quality Assurance

Wageningen Marine Research utilises an ISO 9001:2015 certified quality management system. This certificate is valid until 15 December 2021. The organisation has been certified since 27 February 2001. The certification was issued by DNV GL.

References

- Li, J. L. Xu, L. Tang, S. Wang, L. Li (2018). Big data in tourism research: A literature review. *Tourism Management*, 68 (2018), pp. 301-323
- Önder I., W. Koerbitz, A. Hubmann-Haidvogel (2014). Tracing tourists by their digital footprints. The case of Austria. *Journal of Travel Research*, 55 (5) (2014), pp. 566-573
- Schep et al (2016). Nature on Bonaire and social media behaviour. Assessing the spatial distribution of tourism and recreation on Bonaire through social media. URL:
<https://www.dcbd.nl/document/nature-bonaire-and-social-media-behavior-assessing-spatial-distribution-tourism-and>
- CBS (2019). Data retrieved from <https://www.cbs.nl/en-gb/news/2020/17/tourism-on-bonaire-and-saba-up-in-2019> and previous studies related to this site.
- Verweij, P., Cormont, A., Nel, J., de Rooij, B., Jones-Walters, L., Slijkerman, D., Soma, K., van Eupen, M., 2020, A nature inclusive vision for Bonaire in 2050

Justification

Report C052/20

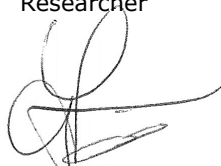
Project Number: 4318300121

The scientific quality of this report has been peer reviewed by a colleague scientist and a member of the Management Team of Wageningen Marine Research

Approved: Drs. J. Tamis
Researcher

Signature:

Date:



8 June 2020

Approved: Drs. J. Asjes
Manager

Signature:

Date:



8 June 2020

Annex 1 Assigning categories

We kept the same categories as Schep et al. (2016) has used. Without pictures/photos with categories of Schep et al. to calibrate our assignment, we have stayed as close as we could to their categories. All photos were assigned to a category, keeping the assumed intention of the photographer in mind. E.g.: in case a terrestrial photo includes wildlife, then the photo is assigned to wildlife instead of terrestrial landscape.

Coastal: Coastal views, including coastline

Seascape: Photo's with views of the sea itself, including watersports. Sunsets. No or very limited coastlines. Watersport including a.o. windsurfing, kayaking, sailing. Moored yachts are not sporty (inactive).

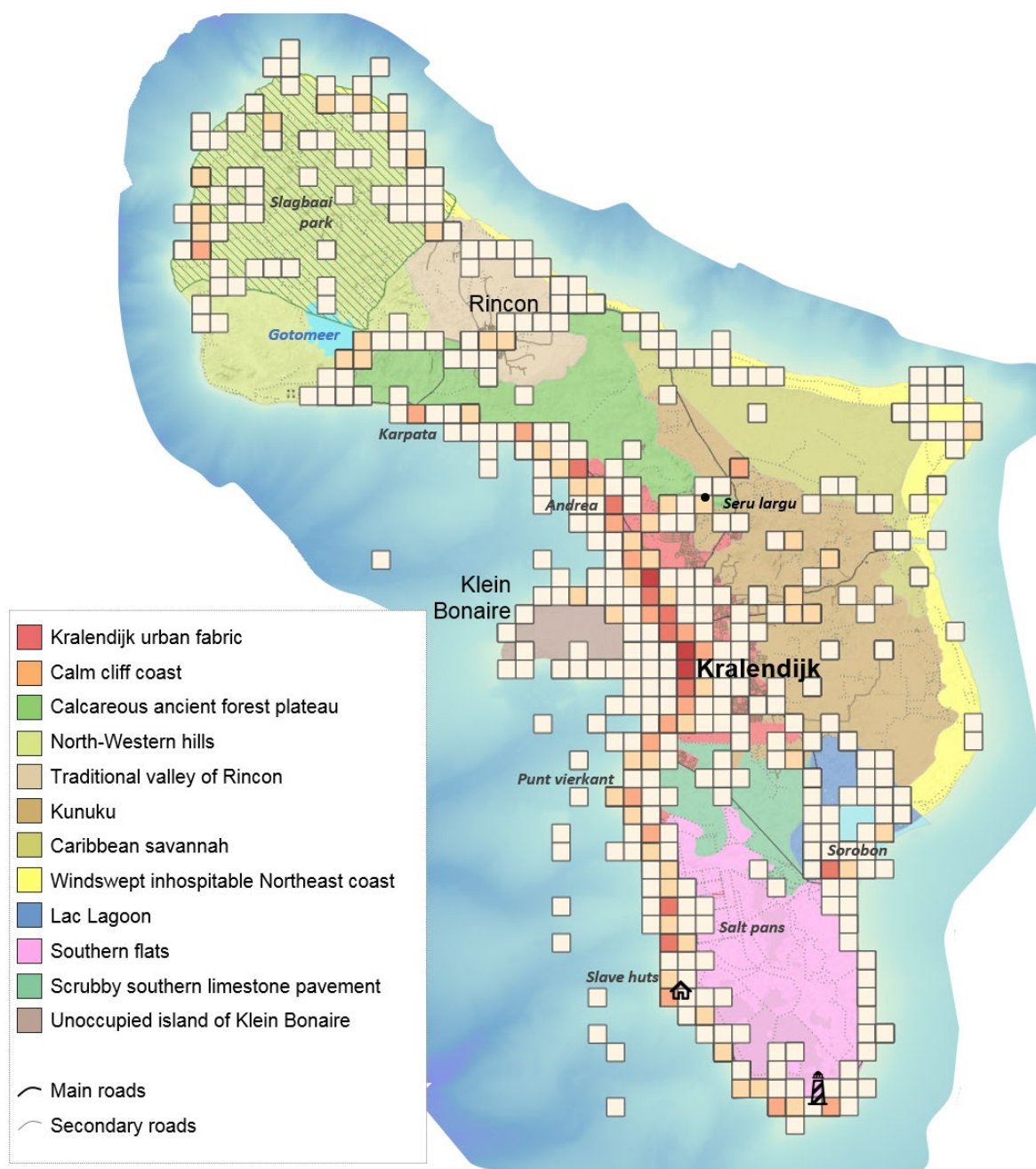
Terrestrial: Landscapes, views on land, can include smaller objects such as man-made structures, but these seemed less important than the landscape itself

Underwater: all underwater pictures (with or without wildlife). When taken above water, before or after a dive, then photos are assigned to other categories, depending the context. (e.g. seascape, coastal, other)

Wildlife: photo's intended to capture an animal on camera. Other than underwater fauna, so focusing on terrestrial wildlife and birds.

Other: indoor photo's, restaurants, streetviews, cultural aspects (market, carnival, etc.) historical buildings, selfies.

Annex 2 Landmarks



Wageningen Marine Research
T +31 (0)317 48 09 00
E: marine-research@wur.nl
www.wur.eu/marine-research

Visitors' address

- Ankerpark 27 1781 AG Den Helder
- Korringaweg 7, 4401 NT Yerseke
- Haringkade 1, 1976 CP IJmuiden

With knowledge, independent scientific research and advice, **Wageningen Marine Research** substantially contributes to more sustainable and more careful management, use and protection of natural riches in marine, coastal and freshwater areas.



Wageningen Marine Research is part of Wageningen University & Research. Wageningen University & Research is the collaboration between Wageningen University and the Wageningen Research Foundation and its mission is: 'To explore the potential for improving the quality of life'
