**Citation patterns of publications using unmanned aerial vehicles in ecology and conservation**

<table>
<thead>
<tr>
<th><strong>Journal:</strong></th>
<th><em>Journal of Unmanned Vehicle Systems</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>juvs-2019-0013.R1</td>
</tr>
<tr>
<td><strong>Manuscript Type:</strong></td>
<td>Review</td>
</tr>
<tr>
<td><strong>Date Submitted by the Author:</strong></td>
<td>05-Sep-2019</td>
</tr>
<tr>
<td><strong>Complete List of Authors:</strong></td>
<td>Dujon, Antoine; Deakin University Faculty of Science Engineering and Built Environment,</td>
</tr>
<tr>
<td><strong>Keyword:</strong></td>
<td>bibliometry, drone, meta-analysis, research impact, review</td>
</tr>
<tr>
<td><strong>Is the invited manuscript for consideration in a Special Issue? :</strong></td>
<td>Not applicable (regular submission)</td>
</tr>
</tbody>
</table>
Citation trajectory of publications using unmanned aerial vehicles in ecology and conservation

Graphical Abstract

109x94mm (150 x 150 DPI)
Citation patterns of publications using unmanned aerial vehicles in ecology and conservation.

Antoine M. Dujon

1Deakin University, Australia, School of Life and Environmental Sciences, Centre for Integrative Ecology, (Waurn Ponds Campus). Waurn Ponds, VIC 3216, Australia.

*Corresponding author: Antoine M. Dujon (antoine.dujon@yahoo.fr)
Abstract

Unmanned aerial vehicles (UAVs) are incorporated as an important part of the toolbox to complement existing methods used in publications in ecology. It is therefore useful to understand how those publications accumulate citations over time. In this study I used 213 articles in which UAVs have been used and I investigate for potential factor underlying how many citations they received. I used metrics that were already shown to be correlated with the number of citations in other fields, and tested more specific effects such as the ecosystem, habitat type or the International Union for Conservation of Nature (IUCN) Red List status of the study species. I found that the time elapsed since publication was the only variable explaining the number of citations a publication received. The average number of citations was 12.1 [95% credible intervals: 8.8,16.7] after two years and 41.8 [95% credible intervals: 27.1,63.7] after five years. In total, <6 % of publications had no citations after one year and <0.5% of publication after two years, which is less than for the field of biology as a whole. This study allows a baseline to be established, from which we can compare the evolution of the field in the future.

Key words: bibliometry, drone, meta-analysis, machine learning, research impact, review
Introduction

Over the past 15 years, unmanned aerial vehicles (UAVs) emerged as a promising technology in ecology (Koh and Wich 2012; Anderson and Gaston 2013; Chabot 2018). UAVs have been used as a tool to obtain a range of ecological insights such as establish species counts (Chabot et al. 2015; Hodgson et al. 2018), monitoring the health of species (Barasona et al. 2014; Christiansen et al. 2016), detecting elusive species (Schofield et al. 2017; Bonnin et al. 2018; Fu et al. 2018), monitoring pollution (Hirds et al. 2017; Martin et al. 2018), and mapping habitat features or engineering species distribution (Kaneko et al. 2014; Barrell and Grant 2015; Szantoi et al. 2017). A such, UAVs are becoming an important part of the ecologist’s toolbox used to complement a range of already existing methods (Gonzalez et al. 2016; Johnston 2018; Rees et al. 2018). As consequence an increasing number of articles using UAVs are published in the field of ecology (Dujon and Schofield 2019).

The number of citations an article receives is usually considered as a relatively reliable measurement of its impact and quality within the scientific community (Aksnes and Rip 2009). Citations acknowledge the impact an author has had on another's work and can be conceived of as “the currency by which we repay the intellectual debt we owe our predecessors” (Antonakis et al. 2014). In addition to advancing their field, highly cited publications increase a scientist’s prestige, and help in obtaining grants which then facilitate the production of new impactful studies and the obtention of additional grants in a positive feedback loop (Judge et al. 2007; Ali et al. 2016). Therefore, obtaining insights on the factors underlying how a publication will accumulate over time has been the focus of multiple studies (Aksnes 2003; Antonakis et al. 2014; Newman 2014; Fox et al. 2016). However, the number of citations a publication receives varies a great deal between different scientific fields, and while general trends have been established (i.e. because of the large amount of publications available when multiple fields are pooled), they may miss specificities in the
citation pattern of a particular field (Radicchi et al. 2008; Patience et al. 2017). As the number of publications using UAVs is increasing rapidly, it then becomes interesting to have a focus on ecological studies in which this relatively new technology is being used.

In this publication I investigated factors underlying the citation pattern of publications using UAVs in ecology. More specifically, I focused on the cases in which UAVs are used to detect and monitor individual or clusters of plant or animals, but also the presence of animal nests. I investigated and quantified the effect of a range of variables including the number of authors, the impact factor of the journal of publication, the IUCN red list status of the study species and whether machine learning was used. I expected the number of authors and the journal impact factor to have a positive impact of the number of citations (Lozano and Larivière 2012; Fox et al. 2016). In addition, I expected publications focusing on species of conservation concern (i.e. classified as vulnerable, endangered or critically endangered) to be more cited compared to the other status. Based on previous findings, and despite its growing popularity, I did not expected the use of machine learning to have an effect on the number of citations (Jordan and Mitchell 2015; Dujon and Schofield 2019). I had no a-priori for the other variables I investigated. With this publication, I expect to provide scientists using UAVs in ecology with an explanation of the factors underlying the citation patterns for publications in this field.

Material and methods

Literature review of publications using UAVs in ecology

For this study I adapted the protocol from Dujon and Schofield (2019). I assembled data on ecological studies using UAVs in ecology from published sources. I searched the Thomson Reuters ISI Web of Science™ database and Google Scholar for papers that included a combination of one of the following terms: ‘drone’ or ‘unmanned aerial vehicle’ or
‘unmanned aerial system’ or ‘unoccupied aircraft system’ or ‘remotely piloted aerial system’
or their abbreviations with any of the following search terms ‘wildlife’ or ‘ecology’ or
‘conservation’ or ‘behaviour’ or ‘behavior’ and ‘machine learning’ in the topic field, which
includes the title, abstract, keywords and Keywords Plus (i.e. words that frequently appear in
the titles of the articles cited within a publication). To locate additional articles that may not
have been identified by the initial search, I also checked the reference lists of relevant papers
based on the pre-defined keywords. In total, I identified 213 publications since 2004 up to the
end of the 2018 (i.e. 31 December) that met the criteria (see Dujon and Schofield 2019 for the
full list of publications). The year 2004 corresponded to the year of publication of the oldest
study that was found and that met the criteria.

For each publication I determined: (1) the number of citations as of the 31th of June
2019 using Google Scholar (because it has a broader coverage compared to the Web of
Science and Scopus databases, Harzing and Alakangas 2016; available at
https://scholar.google.com), (2) the ecosystem type (terrestrial, marine or freshwater), (3) the
continent where the study site(s) is/are located, (4) the habitat type(s) following the highest
classification level of the International Union for Conservation of Nature (IUCN)
Classification Scheme (which defines 17 main habitats; International Union for Conservation
of Nature 2012), (5) whether the authors used machine learning to determine the plant or
animals species, (6) the study species and their IUCN Red List category at the global level
when available (using the IUCN database at www.iucnredlist.org; see Dujon and Schofield
2019), and (7) if the university of affiliation of the first author (which is often considered to
be taking full responsibility for the publication, Duffy 2017) was located in a country with an
“advanced” or “developing” economy (see Dujon and Schofield 2019 and International
Monetary Fund 2016 for full detail on the economic classification).
Data consolidation

Using peer-reviewed articles that investigated patterns in why publications are cited in different fields of science, I identified a range of variables potentially associated with why publications using UAVs in ecology accumulate citations over time (Aksnes 2003; Leimu and Koricheva 2005a, 2005b; Lozano and Larivière 2012; Vanclay 2013; Fox et al. 2016).

Those variables included: (1) whether the article was published in a journal with an impact factor, (2) the impact factor of the journal of publication obtained using InCites Citation Reports website (https://clarivate.com/products/incites) on the year of publication when available, (3) the length of an article, expressed in number of pages, (4) the number of authors, (5) the number of institutions involved, (6) the number of references cited in the publication, and (7) the h-index of the first author (obtained from www.scopus.com) on the year of publication. The h-index is equal to the number of papers (N) in an author’s publication list that have N or more citations (Hirsch 2005).

Statistical analyses

Because the sample size of the different explanatory variables were not always equal, I used a series of three Bayesian multiple linear regression models to investigate how they affected the number of citations a the publications received. The first model (termed as model 1) included the time since publication, the number of authors, the number of references, the publication length, the number of institutions involved, the continent of location of the study site, the ecosystem, the economy type, whether the article was published in a journal with an impact factor, and whether machine learning as explanatory variables was used for a total of 212 publications. One publication was excluded because I was unable to determine its ecosystem type.

The second model (termed as model 2) included all of the variables used in model 1
plus the impact factor of the journal of publication as a continuous variable for a total of 159 publication. A total of 53 publications were excluded from this model because they were published in a journal with no impact factor at the time of publication.

The third model (termed as model 3) included all of the variables used in model 1 plus the IUCN Red List status of the study species as an additional variable. This model included 129 publications for which I was able to determine the species’ IUCN red list status.

Prior to fitting the three models, I added one to the publication’s citation count and log10 transformed them to ensure the response variable followed a normal distribution. In addition I log10 transformed the time since publication to ensure that the residual variance was homogeneous after fitting a model (Zuur et al. 2009). Homogeneity of variance was investigated from scatter plots of residuals vs. fitted values and residuals against each explanatory variable in the model. The normality of the response variable and of the residuals were visually checked using quantile–quantile plots and residual histograms. All assumptions of the models were fulfilled. In addition, I estimated the percentage of variance explained by each of the three models (i.e. using the adjusted R-squared coefficient, as per Zuur et al. 2009).

Throughout this manuscript, I report the estimated parameters followed by their 95% credible intervals between square brackets. The slopes of the regression models for the continuous variables and the mean differences estimated (estimated using contrasts) between two levels of a categorical variable were considered to be significant if their 95% credible interval did not include zero (Kruschke 2015).

All Bayesian models were computed using the MCMCglmm package (Hadfield 2010) within the R software version 3.3.2. (R Development Core Team 2013) and the models fitted using non-informative priors (Hadfield 2010).
Results

Time since publication and other bibliometric measurements

The average number of citations received by an article was significantly associated with the time since publication. The average number of citations for a publication was 12.1 [95%CI: 8.8, 16.7] after two years and 41.8 [95%CI: 27.1, 63.7] after five years (Figure 1). Out of the 213 publications, 11 (<6%) had zero citations after one year, and only one (<0.5%) had zero citations after two years. There was no significant effect of the number of authors, number of institutions involved, the article length, the number of references or the economy type on the average number of citations increase over time (Table 1). Similarly, the journal impact factor or the first author’s h-index did not significantly affect the number of citations. Articles published in a journal with an impact factor were not cited more when compared to articles published in a journal without an impact factor (Table 1).

Ecosystem type, habitat type, location of study site, species of interest and use of machine learning

There was no significant effect of the ecosystem or habitat type and location on the average number of citations a publication accumulated over time (Table 1). Similarly, publications in which machine learning was used to detect the species of interest were cited at a similar level compared to those which did not. In addition, there was no significant difference in the average number of citations for publications covering species with a different IUCN Red List status. Publications that focused on species of conservation concern (i.e. vulnerable, endangered or critically endangered) were not more cited compared to the other IUCN red list categories. There was no significant difference in the citation rate of publications when comparing the continent on which the study site was located on (Table 1).
Percentage of variance explained by the three models.

Model 1 explained 71% of the variance of the data. Excluding time since publication from this model while keeping all other variables (i.e. a non-parsimonious model) resulted in a decrease of the percentage of the variance explained by the model to 23%. Similarly, the percentage of variance explained dropped from 75% to 24% for model 2 and 81% to 30% for model 3 when the time since publication variable was excluded from the models. These results confirm that the time since publication is the main variable with explanatory power for the number of citations a publication using UAVs in ecology will receive.

Discussion

Main findings

I found that the time elapsed since publication was the main variable explaining the increase in the average number of citations a publication accumulated over time. No other variables I investigated had a significant association with the publication citation rate, which contradicts the findings from previous studies in other fields.

Citation rate and proportion of uncited publications

Most of the publications using UAVs in ecology were published in the past five years, and the use of this technology in ecology is still relatively new (Dujon and Schofield 2019). The pool of publications scientists can cite while using this tool is relatively small compared to the tens of thousands of studies published in ecology each year. As a consequence, the number of publications with no citations is relatively low (<6% after one years, <0.5% after two years) and articles quickly accumulate citations. These values are lower than the estimates of ~50% of publications with no citations after one year and ~20% after two years reported by Van Noorden (2017) for the field of biology. As the number of publications using
UAVs in ecology increases it is likely that this phenomenon will disappear and that the proportion of publications with a very low number of citations will increase.

*Journal impact factor and first author h-index*

I found no significant effect of the journal impact factor on the citation trajectories of publications using UAVs in ecology. A weak correlation between those two variables was reported in some fields (e.g. Callaham et al. 2002; Leimu & Koricheva 2005b). However, more broadly, the relationship between impact factor and number of citations has been weakening over the past 20 years. This is likely because most articles are now accessible through institution’s libraries subscriptions, or as open access articles due to the development of online tools, and are no longer bound to their paper copies (Lozano and Larivière 2012; Piwowar et al. 2018). It is, therefore, increasingly easier to access high quality publications regardless of the impact factor of the journal they are published in. As most publications using UAVs in ecology were published in the last five years, my finding is consistent with the broader pattern previously observed by Lozano and Larivière (2012) and it is unlikely a strong correlation would appears in the coming years. Similarly, I found that the h-index of the first author did not affect the number of citations a publication received. The h-index of authors was sometimes found to be correlated with the citation rate (Vanclay 2013). A potential explanation for this is that authors that publish a higher quantity of quality publications tend to see those publications quickly accumulate citations which, as a consequence, reflects on their h-index. This effect was not detected in this study.

*Publication length, number of authors, institutions involved and references*

I investigated whether the average number of citations would increase with the length of a publication, the number of authors, the number of institutions involved and the number of references cited as per Fox et al. (2016) and Leimu and Koricheva (2005a). However, none of these relationships were observed for publications using UAVs in ecology. Longer articles
may give more space for the author to develop key concepts and theories in its field which might explain the increase in citation rate (Leimu and Koricheva 2005b; Fox et al. 2016). A longer reference list may make a publication more visible when searched for in online databases (Didegah and Thelwall 2013). In addition, the existence of a reciprocal citation “game” (i.e. “I cite you and you cite me” scheme) might explain why a longer reference list results in an increased citation rate, and was suggested in other fields (Antonakis et al. 2014). In this field and at the time of this study, these trends were not apparent, likely due to the novelty of the use of UAVs in ecology, and that these effects have not had the time to appear yet, despite a large number of publications.

Location of the university of affiliation and study sites

I found that in the case of ecology studies in which UAVs were used, the economy type of the country of the university to which the first author was affiliated or of the study sites had no effect on the average number of citations a publication received, a finding similar to Antonakis et al. (2014). This was despite a bias towards a higher number of publications using UAVs in ecology being produced by countries with an advanced economy (>90% of publications, Dujon and Schofield 2019). In other fields, countries with an advanced economy produce more publications which receive more citations compared to countries with a developing economy (Leimu and Koricheva 2005a; Mazloumian et al. 2013). As the overall number of publications using UAVs in ecology produced by universities located in a country with a developing economy increases in the future, differences might appear with an overall lower average number of citations when compared to universities located in a country with an advanced economy.
Ecosystem type, habitat type, conservation status of study species and use of machine learning.

There was no difference in the citation rate between publications covering different ecosystem and habitats types. A similar result was observed when considering the study species’ IUCN Red List conservation. While there are biases in the coverage of species in conservation biology (e.g. temperate species are the subject of more studies when compared to tropical species, Titley et al. 2017; and vertebrates are more studied than invertebrates, Lawler et al. 2006), and this coverage does not always align with conservation priorities (Lawler et al. 2006, Meijaard et al. 2015), those biases do not translate to a difference in the number of citations. In addition, the use of machine learning to detect the species of interest did not increase the citation rate. This finding is consistent with Dujon & Schofield (2019) in which we found that an article using machine learning to process UAV data in ecology is not more likely to be published in a journal with a high impact factor compared to a publication in which machine learning was not used. With their rapid increase in popularity, the use of machine learning methods to process UAV imagery data, when considered alone, is unlikely to become correlated with the number of citations a publication will receive; rather it is more likely highly cited publications using machine learning will be those focusing on obtaining insights to answer key ecological questions (Wagstaff 2012; Jordan and Mitchell 2015; Dujon and Schofield 2019).

Conclusion

I found that, except the time elapsed since publication, none of the variables I explored explains the number of citations a publication using UAVs in ecology received. This suggests that when a relatively new technology is used, and when the overall pool of publications is
relatively small, most publications tend to rapidly accumulate citations regardless of their content. This effect may not persist as the total number of publications increases. Furthermore, this study allows the establishment of a baseline from which to compare the evolution of this field in the near future.

Acknowledgments.

I am grateful to Anna Miltiadous for her constructive comments on the manuscript.
Bibliography


Meijaard, E., Cardillo, M., Meijaard, E.M., and Possingham, H.P. 2015. Geographic bias in


Table 1 caption:

Results of the Bayesian multiple linear regression model linking the transformed number of citations received by a publication using UAVs in ecology to a range of continuous and categorical variables. The parameters reported for the variable marked with a (a) were estimated using the model 2, and with a (b) using the model 3. All other variables were investigated using the model 1. Estimates in bold are considered as being statistically significant.
**Table 1:**

### Continuous explanatory variable

<table>
<thead>
<tr>
<th></th>
<th>Slope and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log_{10}(time since publication)</td>
<td><strong>1.457 [0.935, 1.982]</strong></td>
</tr>
<tr>
<td>Number of references cited</td>
<td>0.004 [-0.005, 0.012]</td>
</tr>
<tr>
<td>Number of authors</td>
<td>0.022 [-0.069, 0.107]</td>
</tr>
<tr>
<td>Article length</td>
<td>-0.007 [-0.039, 0.028]</td>
</tr>
<tr>
<td>Number of institutions</td>
<td>-0.033 [-0.149, 0.104]</td>
</tr>
<tr>
<td>H-index of the first author</td>
<td>0.001 [-0.020, 0.023]</td>
</tr>
<tr>
<td>Impact factor(a)</td>
<td>0.077 [-0.044, 0.213]</td>
</tr>
</tbody>
</table>

### Categorical explanatory variable

<table>
<thead>
<tr>
<th>Reference group</th>
<th>Comparison group</th>
<th>Mean difference between groups and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>In a journal with an impact factor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>0.127 [-0.221, 0.511]</td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>0.045 [-0.294, 0.381]</td>
</tr>
<tr>
<td>Used machine learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Yes</td>
<td>0.285 [-0.628, 1.280]</td>
</tr>
<tr>
<td>Economic country affiliation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced</td>
<td>Developing</td>
<td>0.051 [-0.462, 0.598]</td>
</tr>
<tr>
<td>North-America</td>
<td>Oceania</td>
<td></td>
</tr>
<tr>
<td>South-America</td>
<td>Europe</td>
<td>0.040 [-0.409, 0.460]</td>
</tr>
<tr>
<td>Africa</td>
<td>Asia</td>
<td>-0.479 [-1.439, 0.454]</td>
</tr>
<tr>
<td>Antarctica</td>
<td>Europe</td>
<td>-0.237 [-1.420, 1.019]</td>
</tr>
<tr>
<td>Ecosystem</td>
<td>Marine</td>
<td>-0.598 [-2.743, 1.704]</td>
</tr>
<tr>
<td>Terrestrial</td>
<td>Freshwater</td>
<td>-0.064 [-0.967, 0.852]</td>
</tr>
<tr>
<td>Habitat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Terrestrial artificial</td>
<td>Aquatic artificial</td>
<td>0.206 [-2.093, 2.288]</td>
</tr>
<tr>
<td>Forest</td>
<td>0.314 [-1.722, 2.472]</td>
<td></td>
</tr>
<tr>
<td>Grassland</td>
<td>0.056 [-2.190, 1.983]</td>
<td></td>
</tr>
<tr>
<td>Introduced vegetation</td>
<td>0.276 [-1.970, 2.427]</td>
<td></td>
</tr>
<tr>
<td>Marine coastal and supratidal</td>
<td>0.197 [-2.070, 2.343]</td>
<td></td>
</tr>
<tr>
<td>Marine intertidal</td>
<td>-0.029 [-2.414, 2.386]</td>
<td></td>
</tr>
<tr>
<td>Marine neritic</td>
<td>0.193 [-1.911, 2.468]</td>
<td></td>
</tr>
<tr>
<td>Savanna</td>
<td>0.150 [-2.108, 2.518]</td>
<td></td>
</tr>
<tr>
<td>Shrubland</td>
<td>0.036 [-2.016, 2.286]</td>
<td></td>
</tr>
<tr>
<td>Wetland</td>
<td>0.182 [-1.997, 2.561]</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.100 [-2.280, 2.331]</td>
<td></td>
</tr>
<tr>
<td>IUCN Red List Status(b)</td>
<td>Not Evaluated</td>
<td></td>
</tr>
<tr>
<td>Data deficient</td>
<td>-0.087 [-1.279, 1.271]</td>
<td></td>
</tr>
<tr>
<td>Least concern</td>
<td>0.066 [-0.514, 0.609]</td>
<td></td>
</tr>
<tr>
<td>Near threatened</td>
<td>-0.031 [-0.926, 0.871]</td>
<td></td>
</tr>
<tr>
<td>Vulnerable</td>
<td>-0.116 [-0.875, 0.696]</td>
<td></td>
</tr>
<tr>
<td>Endangered</td>
<td>-0.117 [-1.034, 0.755]</td>
<td></td>
</tr>
</tbody>
</table>

(a) Impact factor is the impact factor of the journal where the article was published.

(b) IUCN Red List Status includes Not Evaluated, Data Deficient, Least Concern, Near Threatened, Vulnerable, and Endangered.
Figure caption:

Figure 1: Number of citations in function of the time elapsed since publication for 213 articles published in the field of ecology and in which UAVs were used. One grey dot represents one publication, the solid and dashed black lines represent the average number of citations and the associated 95% credible interval predicted using a Bayesian linear regression model.
Figure 1

![Graph showing the relationship between number of citations and time since publication. The X-axis represents time since publication in years, ranging from 0 to 15. The Y-axis represents the number of citations, ranging from 0 to 1000. The graph includes multiple data points and fitted curves.]