



Trends and gaps in the last decade of movement ecology (Ideas for Manuscript)

- It's still a methodological paper on movement ecology, about where we are and where we are going (or should go)

Intro

An Intro on Nathan's paper, grow in the field due to technological advances for tagging, software for processing and analyzing data and methods. We can prepare a time line like in the poster with the number of publications + special features (Fig.). We could mention conferences related to movement ecology, biologging and statistical ecology. And then, introduction to a need to stop and see what we are going, where are we going and with what tools (... a movement ecology analysis of movement ecology??? sorry I may be thinking about this too much).

Two short paragraphs about methods: One about identifying the mov-eco papers. I don't like to spend much on this part but I guess that's important. Saying that with the combinations of words we use (we could only show them in the supp. mat.) we make sure that > 80% papers are mov-eco papers, but we may lose some others not included here. Could there be biases? We should certainly address this in one or two lines, but Idk how to tackle this exactly (i.e. finding out a type of paper we could be missing because of our word criterion). A second short paragraph about the actual analysis: examining title + keywords + abstract + materials and methods, asking about specific stuff (i.e. framework, software, devices, methods) with dictionaries and topic identification with LDA models. The idea would be to mention all this - along with the fact that we perform quality control at every stage of the analysis - but only detail these parts in Supp. Mat. In Supp. Mat, there will also be a list of caveats with our analyses; e.g. the fact that we can find a paper on birds and mammals with GLS data and our algorithms will not conclude that GLS were only used on birds (not mammals).

Main results:

Overall trends:

Increasing trend in the number of publications (Fig.1) and number of species studied (Fig.2), consistent with everything in the intro and the fact that there are more publications in science in general. Can someone get a reliable time series (or data to produce that curve) of the number of publications in science or in ecology? Or a publication showing it?

Framework:

Not much change in the studies in framework components respect to the decade before (Fig. 3 and 4). Still a lot more focus on external factors than the others. There should be some discussion about why this is happening. - I think some input from Tommy, Susana, Mathieu would be great for this part (and everybody else who'd like to lend a hand). Also, what can we say about the subcategories of the framework? What do you find interesting? (Tables 1-5) The link to the subcategories and their keywords is: https://docs.google.com/document/d/1F43--6wK_k54kwCyJezpZyiDI2rMnBAqEVmFz7kU0G0/edit?usp=sharing

A shift in the use of tagging devices

Increase use of video and accelerometer data in publications (Fig. 5), meaning that people are actually trying to measure movement explicitly, and not just by sampling locations in time (if the idea is not well expressed

A TIME LINE: PUBLICATIONS & EVENTS

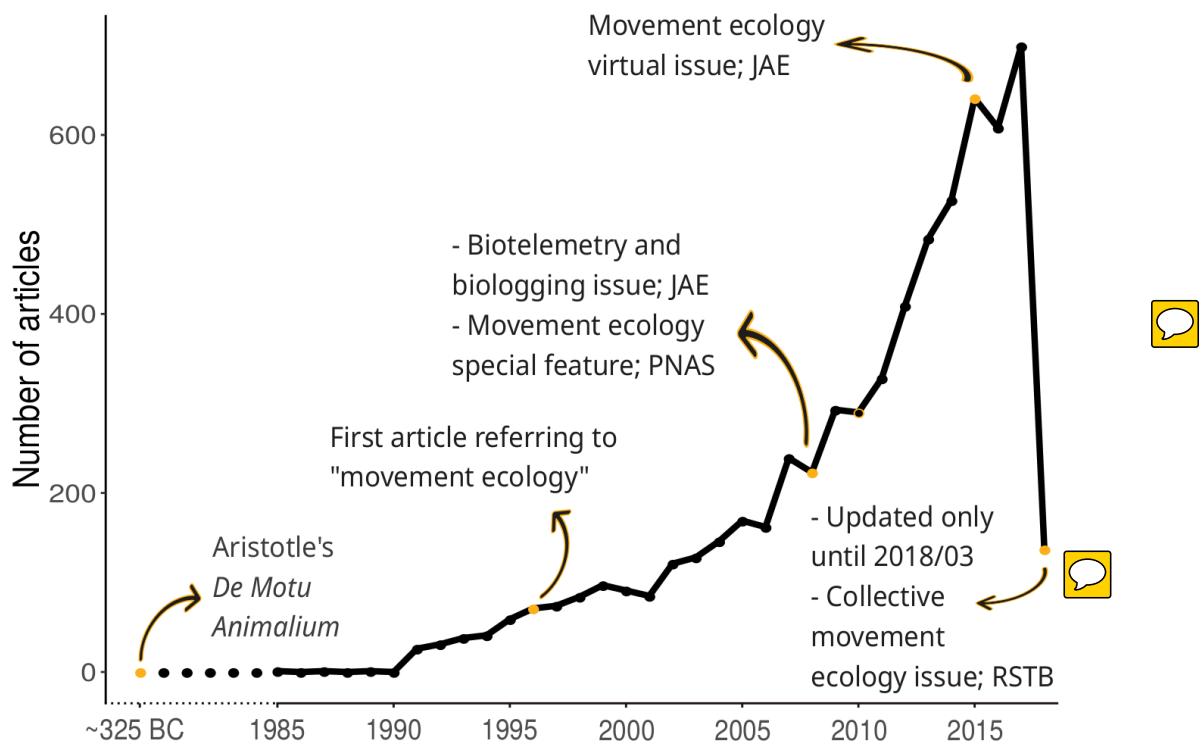


Figure 1: Timeline from poster (need to update 2018)

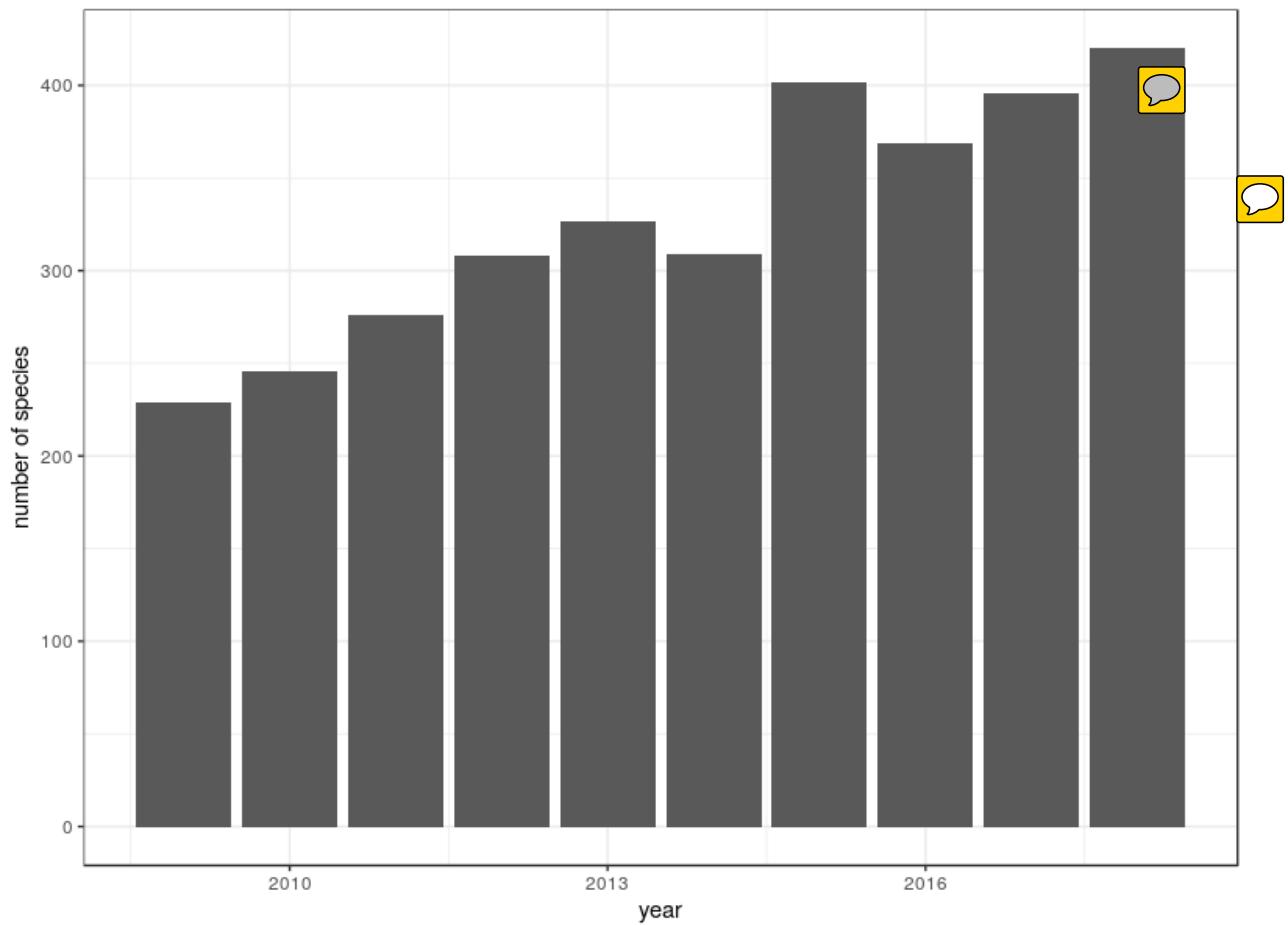


Figure 2: Time series for species (I'd do a line plot instead of bars)

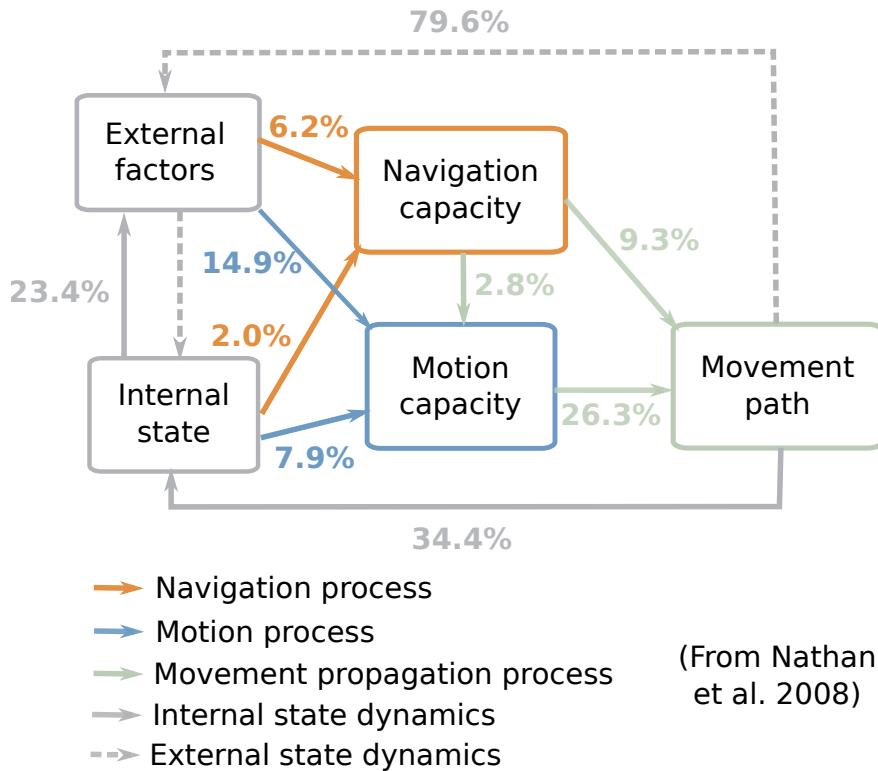


Figure 3: Framework between 2009-2018

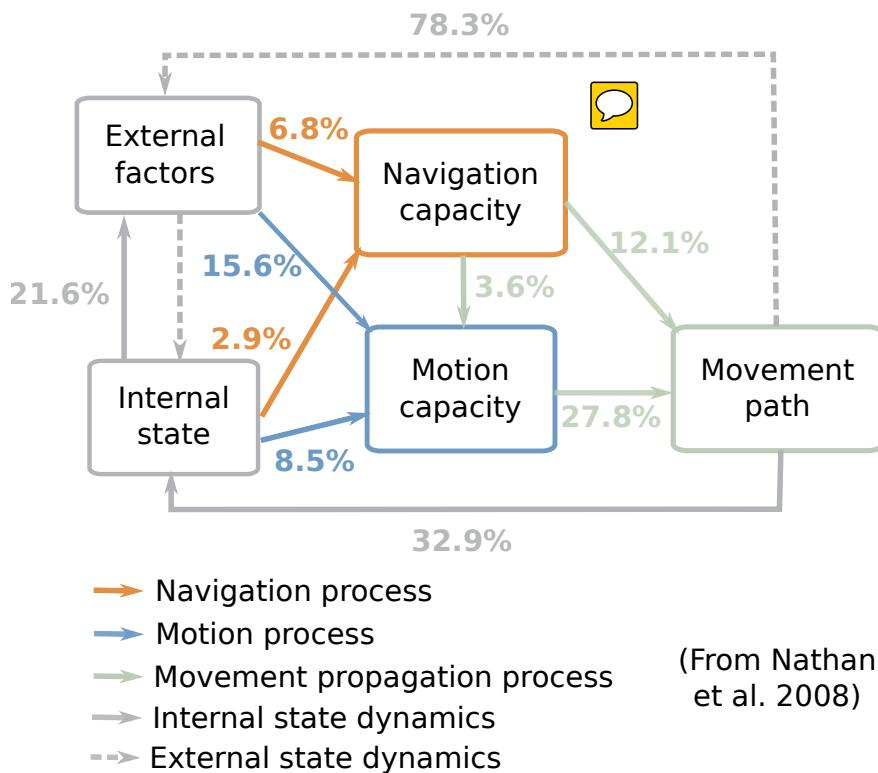


Figure 4: Framework between 1999-2008

External factor	Perc.
Environment	59.6%
Animal interaction	24.1%
Anthropogenic	7.1%
Other	0.5%
General	0.3%

Table 1: Percentage of external factor papers related to each category

Internal factor (TAC)	Perc.
Phenotypic trait	39.4%
Physiological state (shorter term)	30.9%
Individual state (longer-term)	4.8%
General	1.2%
Genotype	0.6%

Table 2: Percentage of internal state papers related to each category. Those are Tommy's categories.

*Phenotypic: social status, body mass, sex, personality; *Physiological: physiology, neurology; *Individual: breeding, age; *Genotype: genetic, telomere

please help me). This means that there are things that could be explored more deeply, like motion with accelerometer, or use of cues to navigate or change before with video. It also presents new challenges in terms of data processing and analysis.



Tommy's comment: I think the right terms are egocentric and exocentric. I.e. studies increasingly have an egocentric frame of reference (i.e. movement of animal) rather than just their x and y positions in the environment. See this paper which might be relevant here - <https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2656.13094>

Towards an R community

Software-wise, there is a homogenization process towards R (Fig. 6). The second most used software, ArcGIS, is steadily decreasing in use. It may be a sign that research is going from space occupancy or use, to studying actual movement. R doesn't only have a lot of packages to process and analyze movement (I'd mention the R-pkgs review), but it also offers a free software environment to program and create new methods. If the whole community is converging towards one software, there is potentially a lot that we can do together to rapidly improve processing and analysis. - I'm sure Mathieu could expand on this.

Mathieu's comment during the meeting was about ArcGIS being used mostly for data processing before (so analyses with other software), and now R is enough for them. He was surprised of both the poor use of R at the beginning of the decade and its huge popularity at the end.

Tommy's comment: Is use QGIS or similar programs increasing - this would tell you that spatial analysis/mapping is fairly constant still. I think you could mention that increasingly there are more sophisticated packages in R for creating maps/figures/data vis. Rocío: My answer to the QGIS part is that its declared use was very poor.

Methods for analysis: are we using movement specific methods?

Concerning methods, most papers are using general methods and not many "movement-specific" ones (e.g. step selection functions, state-space models, individual-based models, random walk models) [Check Table 6 and Fig. 7, and dictionary file in the repository, ./Data/Dictionary/Synonyms-Methods.csv, for methods words].

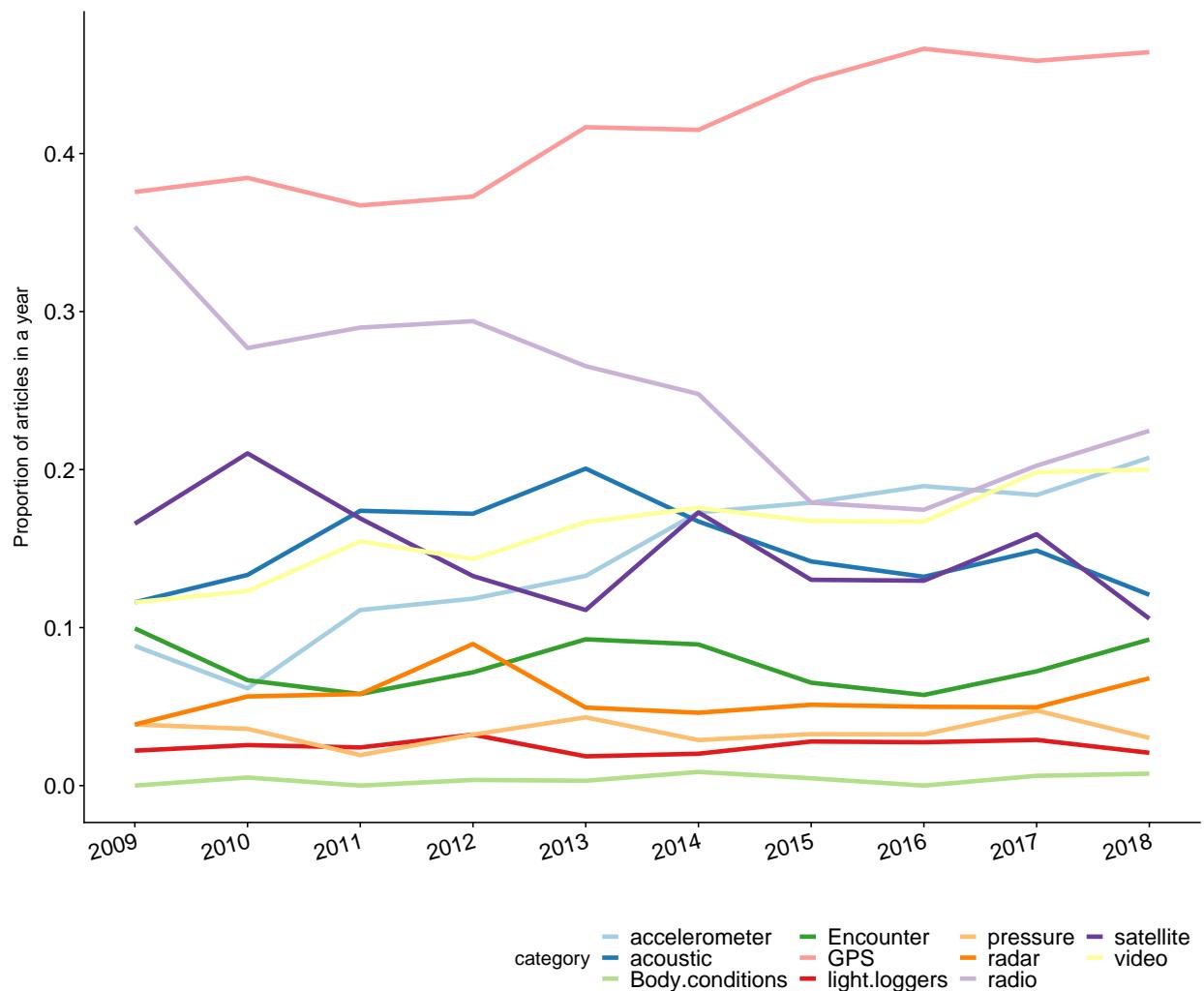


Figure 5: Proportion of papers of each year using each type of device

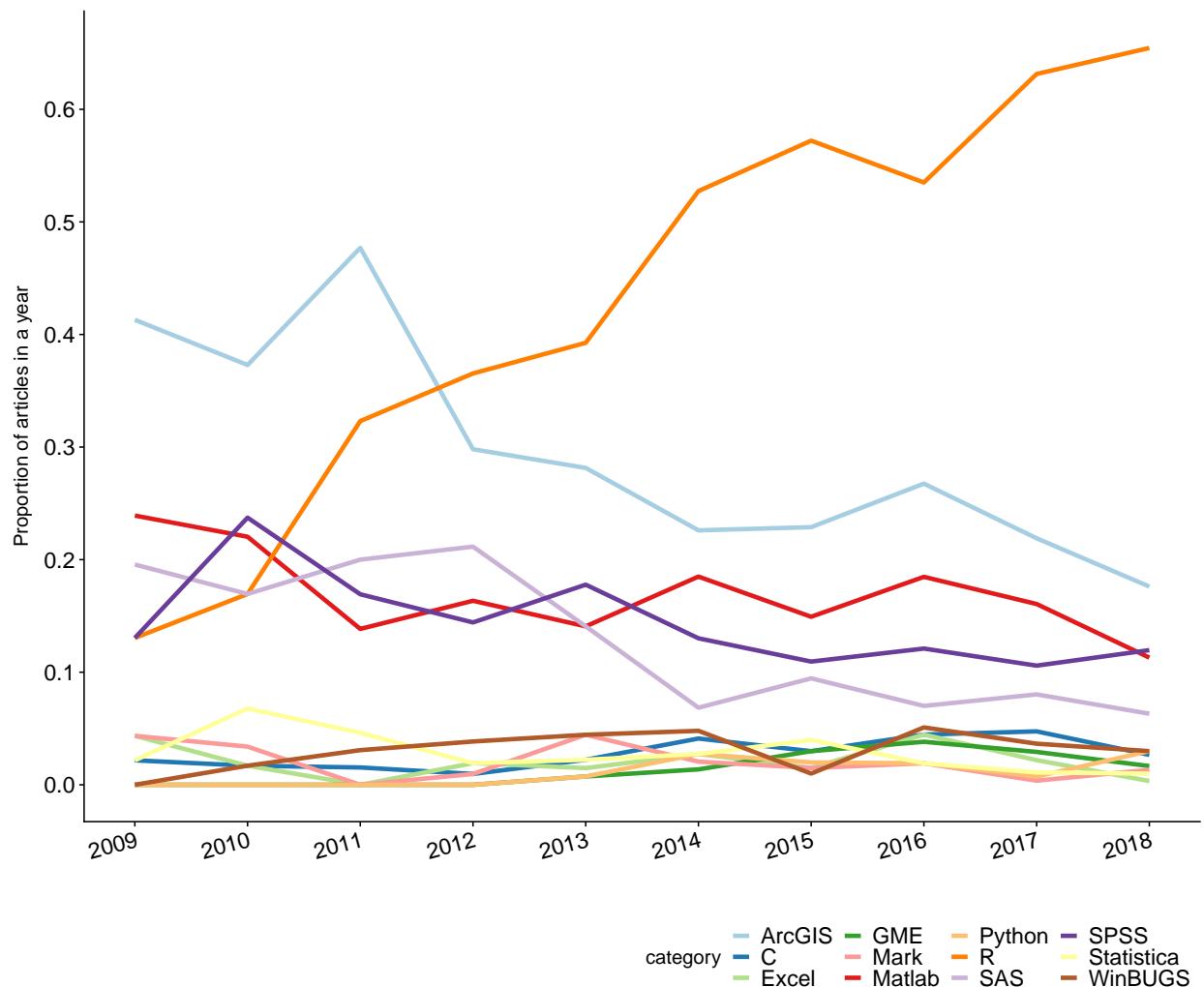


Figure 6: Proportion of papers of each year using each software (12 most mentioned software)

Internal factor (SCT)	Perc.
Life-history	37.4%
Physiological state	29.9%
Behavior	6.0%
General	1.6%
Cognition	1.5%
Genotype	0.6%

Table 3: Percentage of internal state papers related to each category. Those are Susana's categories. *Life-history: breeding, body mass, sex; *Physiological: physiology, neurology; *Behavior: social status, personality; *Genotype: genetic, telomere; *Cognition: psychology, past experience

Motion process	Perc.
Biophysical	34.5%
Empirical	26.7%
Theoretical	20.7%

Table 4: Percentage of motion papers related to each category. *Biophysical: biomechanic, locomotion; *Empirical: soaring, dive; *Theoretical: random walk, sinuosity;

Most papers use general methods like generalized linear (mixed and not mixed) models, ANOVAs, etc., when analyzing tracking data (mostly to analyze the effect of an environmental covariate, from what I can remember). Or, as Tommy said, there are many descriptive pieces. It is true that sometimes we don't require the use movement-specific methods, and the aim of the paper is not to judge, but we rather expect that in the context of increased number of publications, movement-specific devices and convergence towards the use of R, we expect that the use of movement-related methods will increase. The proportion of papers per year mentioning movement methods have actually decreased, and I wonder if it's because the use of these methods has not increased at the same speed as the number of publications [I could also have forgotten words in the dictionary; that's possible; but we had an acceptable quality control result. All QC results should go on Supp. Mat.; I only have to find where they are].

Main topics in movement ecology

For all coauthors: Topic analysis was performed using LDA. As I explained in an email, I could have used the standard way to choose topics which is the perplexity score, but that usually ends up with hundreds of topics. I found stuff online about using a coherence metric, but that may require a bit of time exploring. I personally think we could justify our choice of 13 as the minimum number that can provide interpretable topics. If you want to check if a lower/higher number would give interpretable topics, see word clouds in the repository: /Rocio/plots/ and check all files with a name like wordcloud_topics_X_alpha_1_method_VEM.pdf. You can share with us if you have a different choice [so, feel free to say if you think we absolutely need some time to use a coherence metric. I read a post online from a guy who is finishing his PhD on that, so he didn't have all results to show. But it is not impossible to do our own research on it.]

So this part would be about describing topics. Check Fig. 8.

Putting it all together (not a great title, I know)

The sankey graph below try to summarize the whole story. We show only the taxonomic groups that have the highest number of papers. (There are more graphs in /Matt/plots/Sankey. Feel free to check them and see if a plot I don't show here would be better for the manuscript.) The size of each column is proportional to the number of papers related to that category. For instance, there are way more papers focused on mammals than on the other groups. All animal groups are significantly more related to external factors. Humans, on the other hand, are almost equally focused on the analysis of external factors, the internal state and motion.

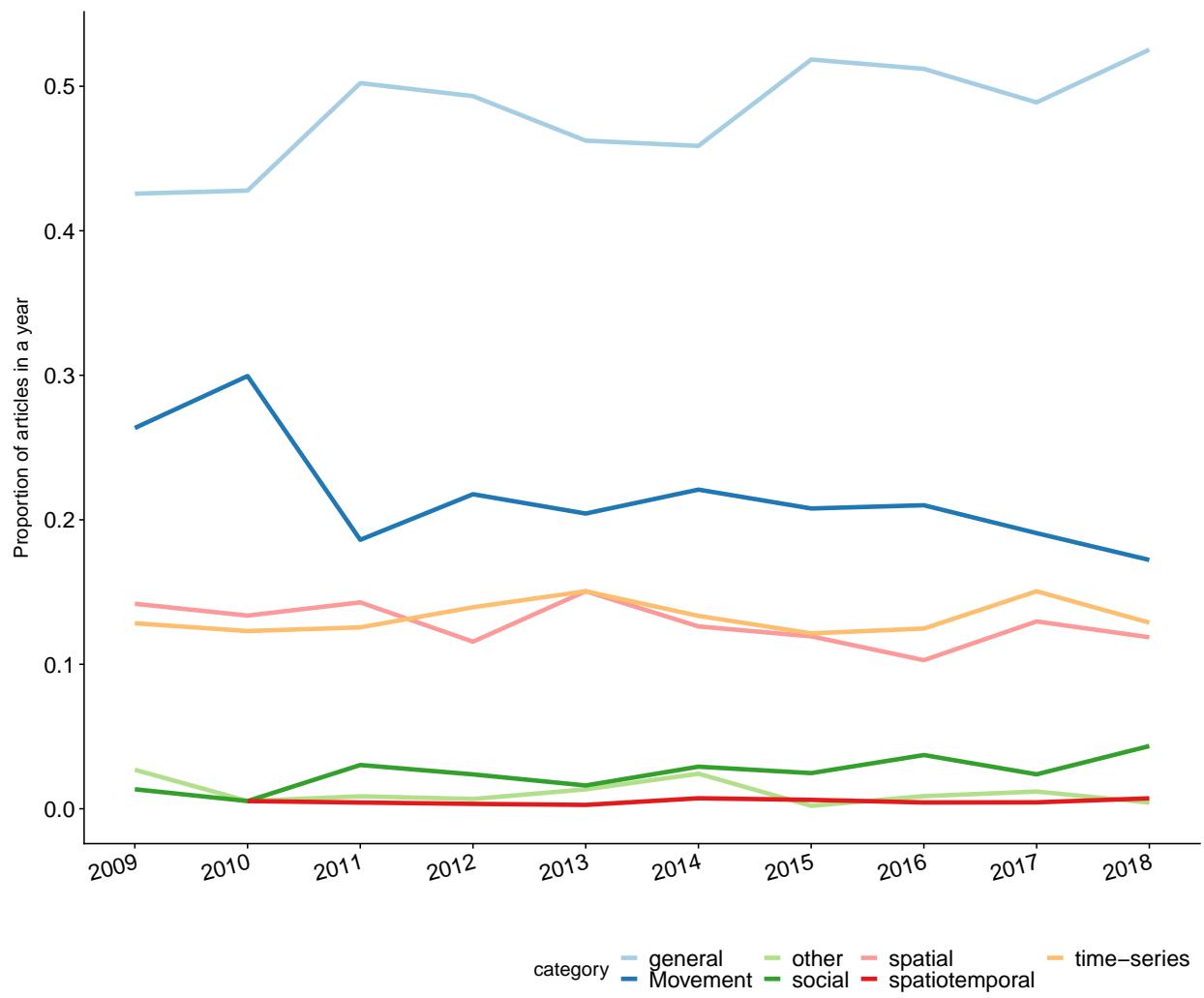


Figure 7: Proportion of papers of each year using each type of method

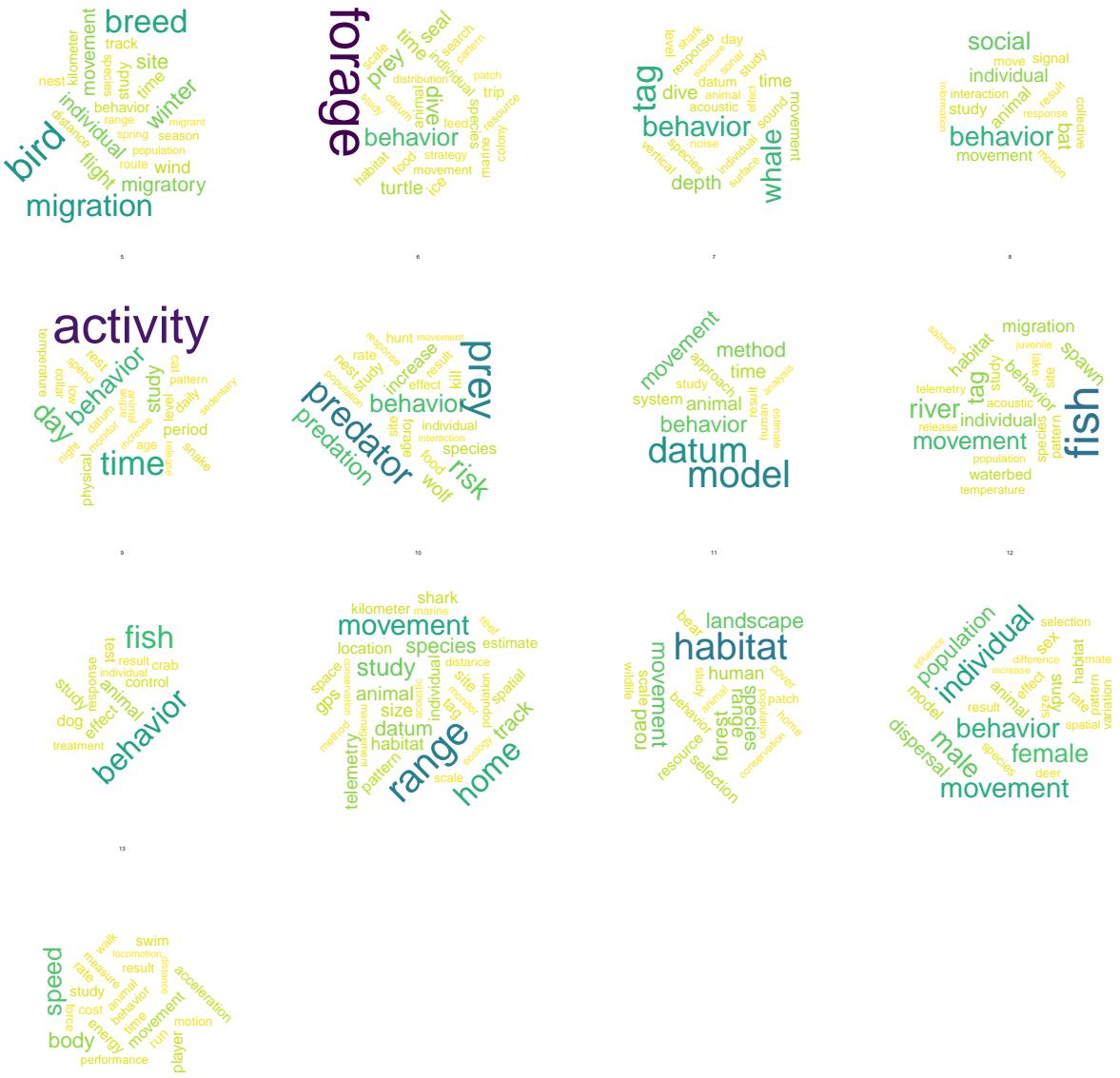


Figure 8: Word clouds of each topic

Navigation process	Perc.
Cognitive	21.4%
Orientation	14.5%
Sensory	5.0%

Table 5: Percentage of navigation papers related to each category. *Cognitive: memory, cognition; *Orientation: compass, homing, path integration; *Sensory: olfact, chemor, magnetor;

Method	Perc.
general	62.2%
Movement	20.5%
spatial	7.8%
time series	7.2%
social	2.3%
other	0.7%
spatiotemporal	0.4%

Table 6: Percentage of use of each type of statistical method in publications. A publication may be using several types of methods.

Navigation is the least studied in all groups. Some framework components are more linked to some topics than others. After analyzing the topics and labeling them, would you say that these links make sense? Is a component (e.g. navigation) been neglected in a topic where it should be important?

We also discussed the idea of shading some not so important topics. And actually the fact that some topics are seen in few papers should be analyzed. I don't want to go ahead on my ideas because I want you to first give it a try at labeling the topics, but after you do that, I'd love for you to share your thoughts. So please check the graphs per taxon in /Matt/plots/Sankey, add text to the document based on your interpretations (not just descriptions, but ideas of what to show in the paper) and say which graphs would be best.

Conclusions

The paper could close on the opportunities in working together, human and animal fields, using more sophisticated tools, in a united software community, and trying to tackle more the neglected components of the framework. We should discuss the challenges to overcome.



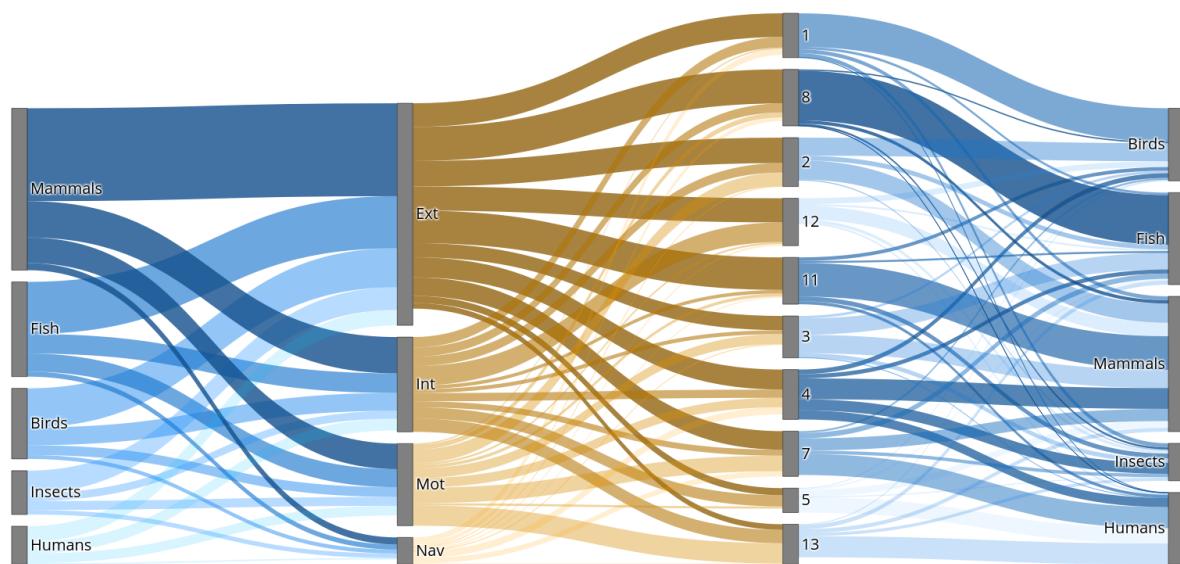


Figure 9: Taxa - Framework - Topics - Taxa