A Rapid Computer-assisted Systematic Map of Regional Climate Impacts

June 12, 2020

Outline

Context

Systematic assessments of the evidence on Climate Change like those conducted by the IPCC are vital.

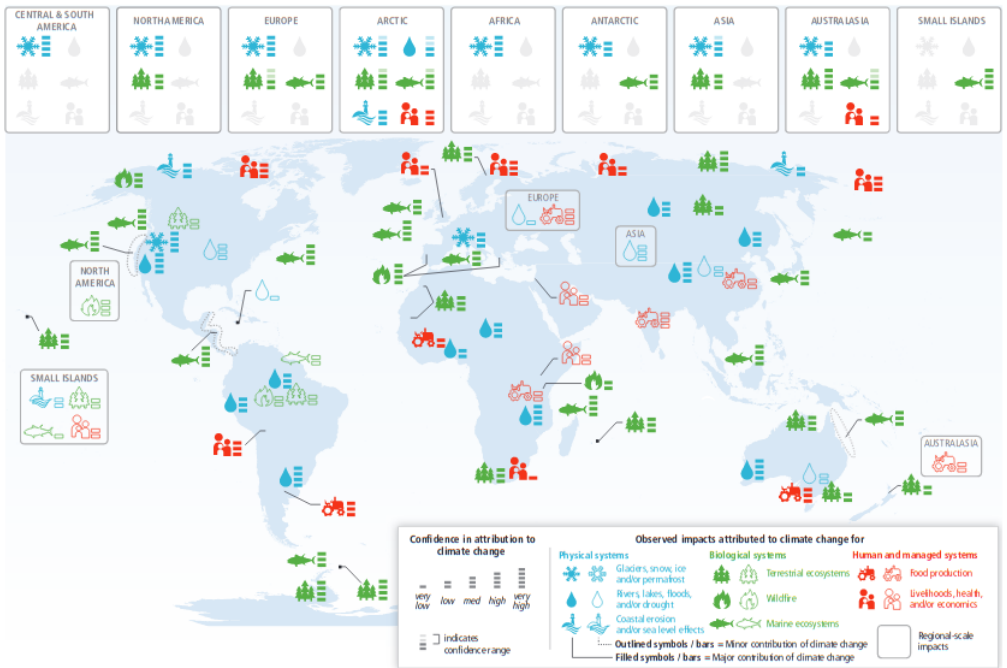


Figure 18-3 | Global patterns of observed climate change impacts reported since AR4. Each filled symbol in the top panels indicates a class of systems for which climate change has played a major role in observed changes in at least one system within that class across the respective region, with the range of confidence in attribution for those region-wide impacts indicated by the bars. Regional-scale impacts where climate change has played a minor role are shown by outlined symbols in a box in the respective region. Sub-regional impacts are indicated with symbols on the map, placed in the approximate area of their occurrence. The impacted area can vary from specific locations to broad areas such as a major river basin. Impacts on physical (blue), biological (green), and human (red) systems are differentiated by color. This map represents a graphical synthesis of Tables 18-5, 18-6, 18-7, 18-8, and 18-9. Absence of climate change impacts from this figure does not imply that such impacts have not occurred.

Context

Systematic assessments of the evidence on Climate Change like those conducted by the IPCC are vital.

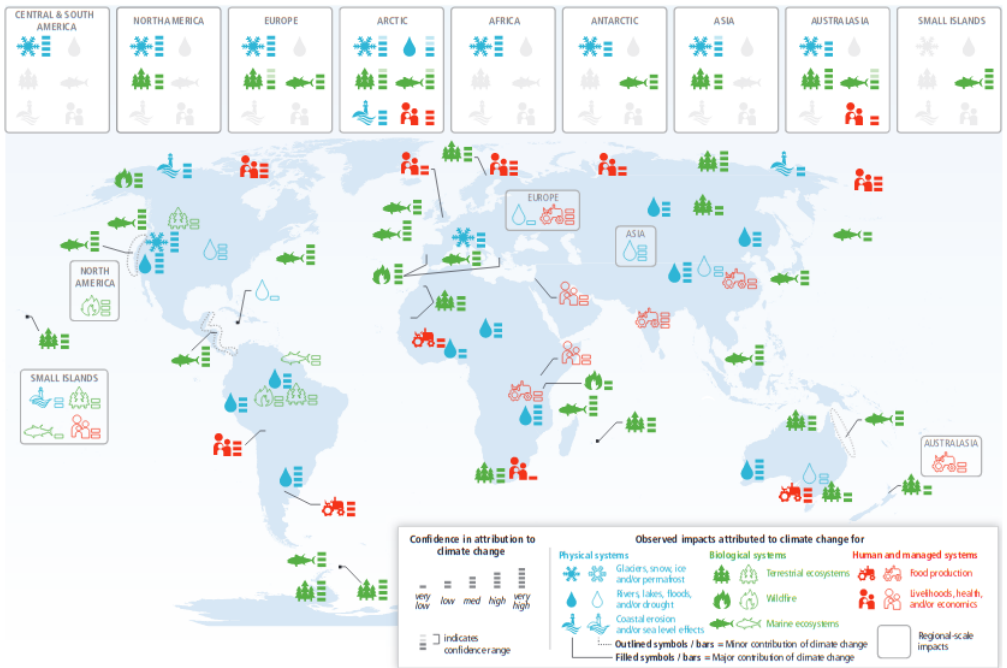
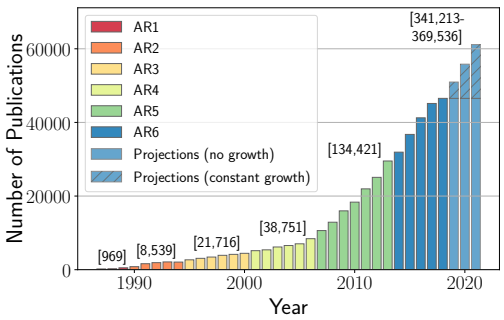


Figure 18-3 | Global patterns of observed climate change impacts reported since AR4. Each filled symbol in the top panels indicates a class of systems for which climate change has played a major role in observed changes in at least one system within that class across the respective region, with the range of confidence in attribution for those region-wide impacts indicated by the bars. Regional-scale impacts where climate change has played a minor role are shown by outlined symbols in a box in the respective region. Sub-regional impacts are indicated with symbols on the map, placed in the approximate area of their occurrence. The impacted area can vary from specific locations to broad areas such as a major river basin. Impacts on physical (blue), biological (green), and human (red) systems are differentiated by color. This map represents a graphical synthesis of Tables 18-5, 18-6, 18-7, 18-8, and 18-9. Absence of climate change impacts from this figure does not imply that such impacts have not occurred.



- These are challenged by big literature

Context

Systematic assessments of the evidence on Climate Change like those conducted by the IPCC are vital.

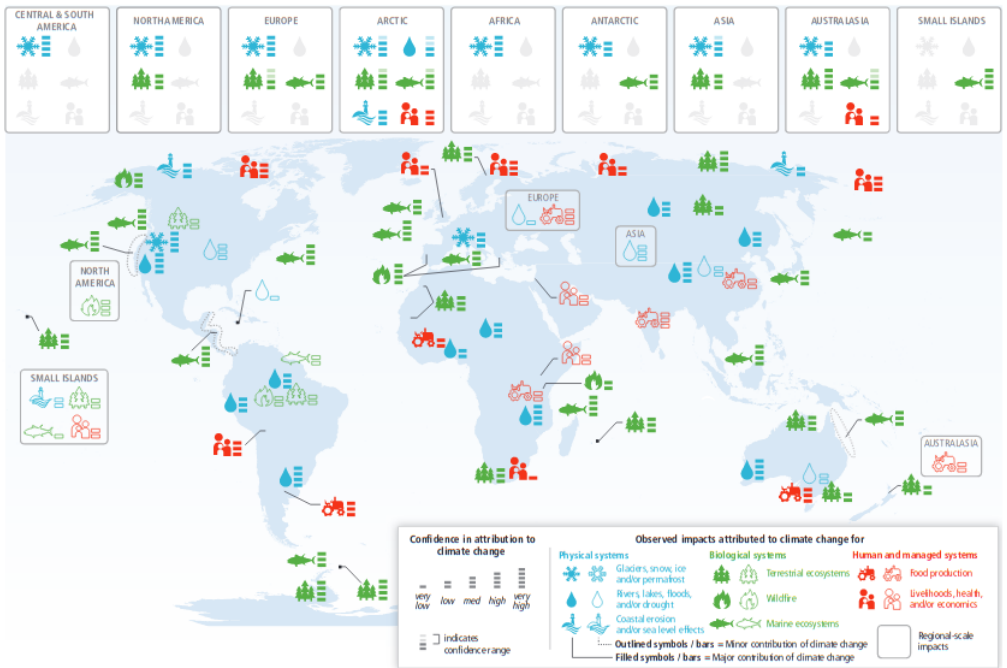
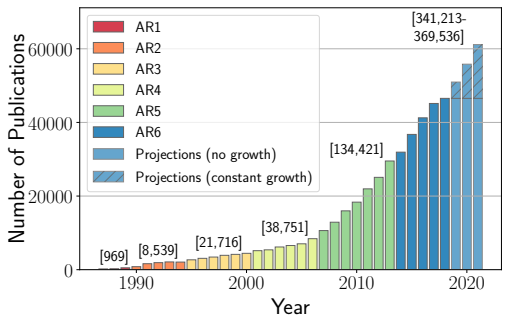


Figure 18-3 | Global patterns of observed climate change impacts reported since AR4. Each filled symbol in the top panels indicates a class of systems for which climate change has played a major role in observed changes in at least one system within that class across the respective region, with the range of confidence in attribution for those region-wide impacts indicated by the bars. Regional-scale impacts where climate change has played a minor role are shown by outlined symbols in a box in the respective region. Sub-regional impacts are indicated with symbols on the map, placed in the approximate area of their occurrence. The impacted area can vary from specific locations to broad areas such as a major river basin. Impacts on physical (blue), biological (green), and human (red) systems are differentiated by color. This map represents a graphical synthesis of Tables 18-5, 18-6, 18-7, 18-8, and 18-9. Absence of climate change impacts from this figure does not imply that such impacts have not occurred.



- These are challenged by big literature
- They do not account for uncertainty about what literature is available

Goal

There are hundreds of thousands of documents potentially relevant to observed climate impacts. We want to be able to do two things:

Goal

There are hundreds of thousands of documents potentially relevant to observed climate impacts. We want to be able to do two things:

- Separate those documents which *are* relevant from those that are not

Goal

There are hundreds of thousands of documents potentially relevant to observed climate impacts. We want to be able to do two things:

- Separate those documents which *are* relevant from those that are not
- Predict in what way relevant documents are relevant:
  - What impacts do they document?
  - What type of evidence do they provide?
  - In which locations is there evidence

Goal

There are hundreds of thousands of documents potentially relevant to observed climate impacts. We want to be able to do two things:

- Separate those documents which *are* relevant from those that are not
- Predict in what way relevant documents are relevant:
  - What impacts do they document?
  - What type of evidence do they provide?
  - In which locations is there evidence

Once we can do that, we can draw a rough map of the available evidence, and/or aid the production of an *assessment* of the available evidence

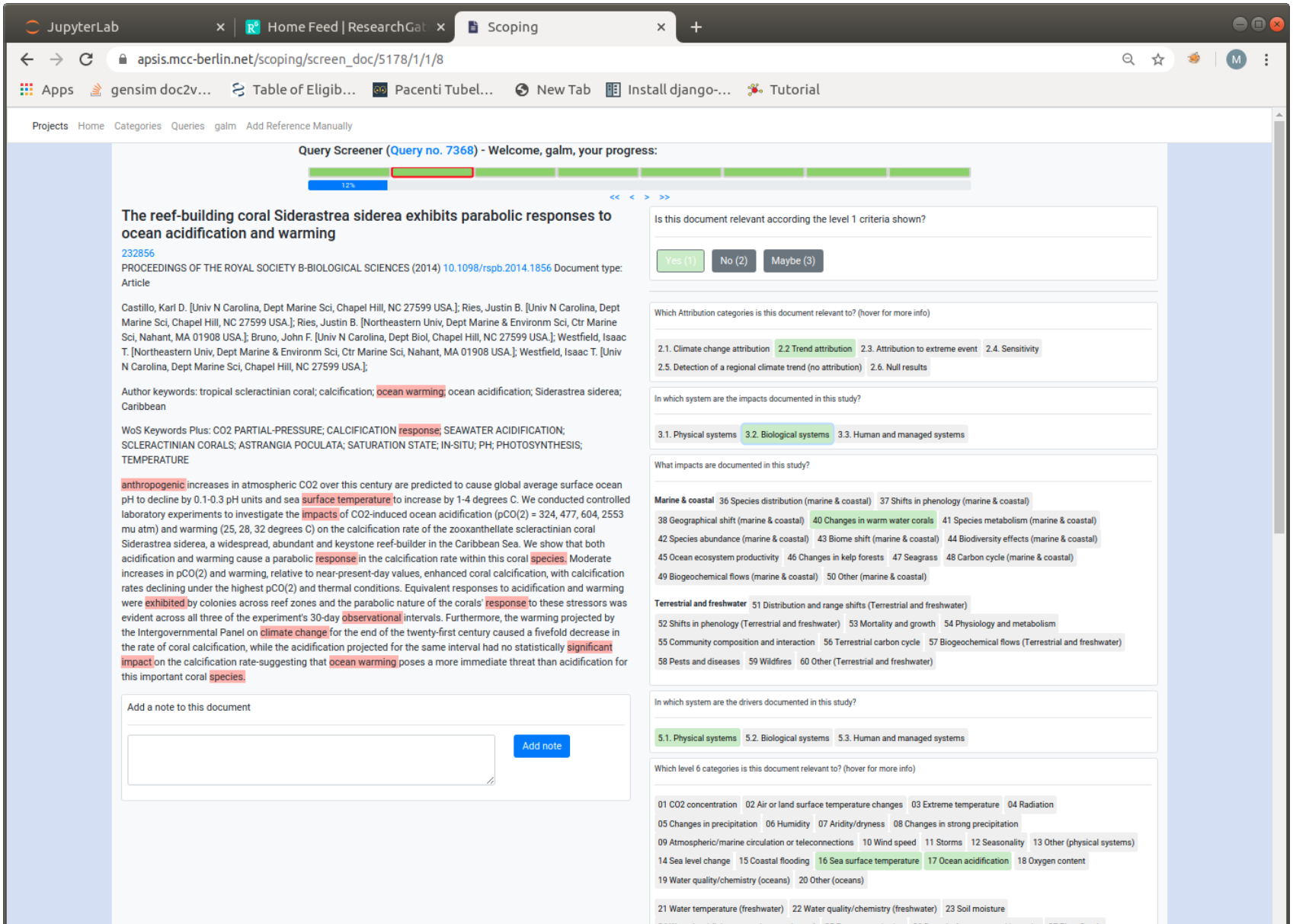
Distribution of labour between humans and machines

A human expert or a team of human experts is best placed to answer those questions for any single document, but they can't look at all potentially relevant documents

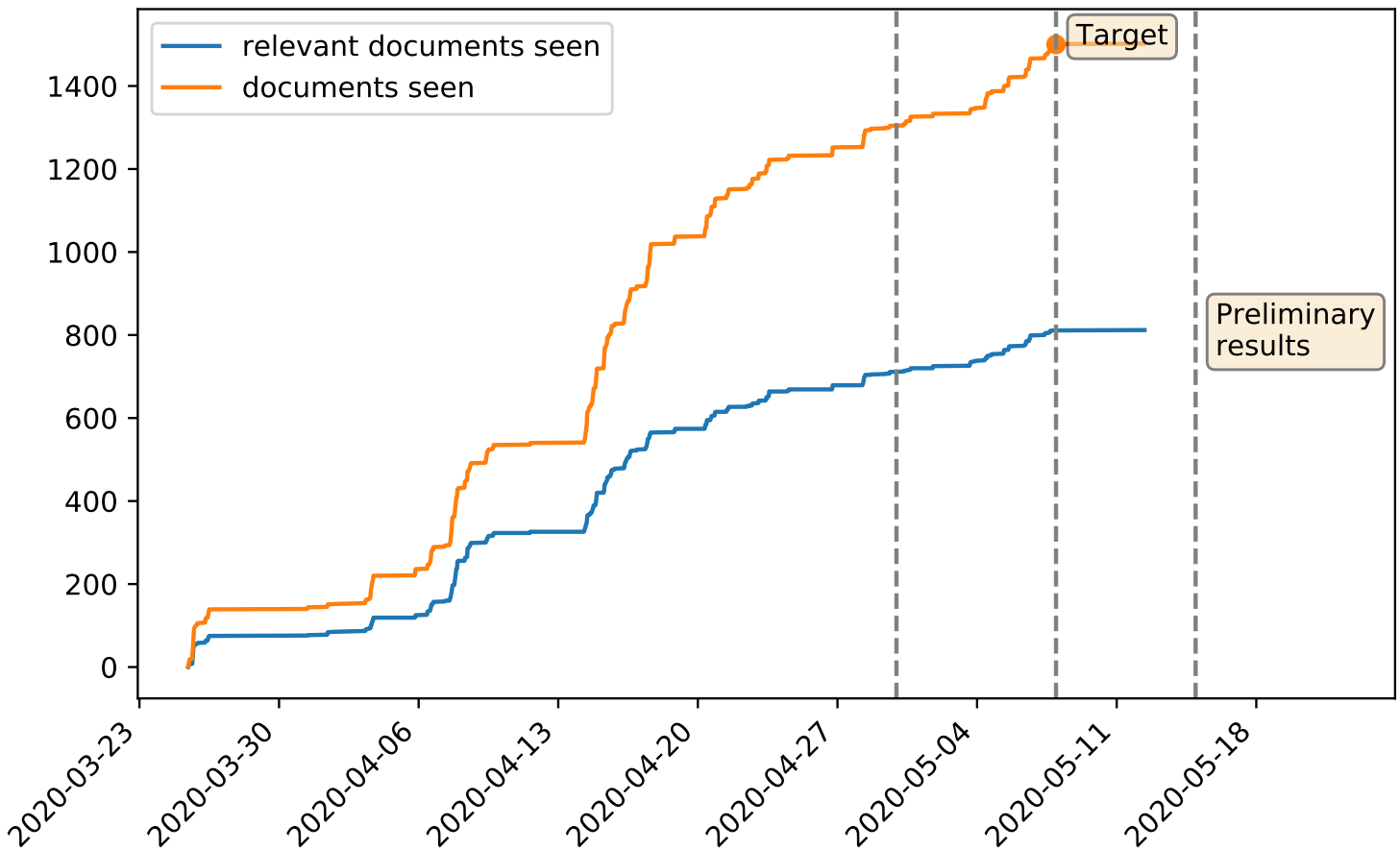
We can use labels generated by humans to try to teach a computer what a relevant document looks like, and how to decide in what way it is relevant.

If this works well, we can predict, with some uncertainty, how much evidence there is, and where and on what topic it is.

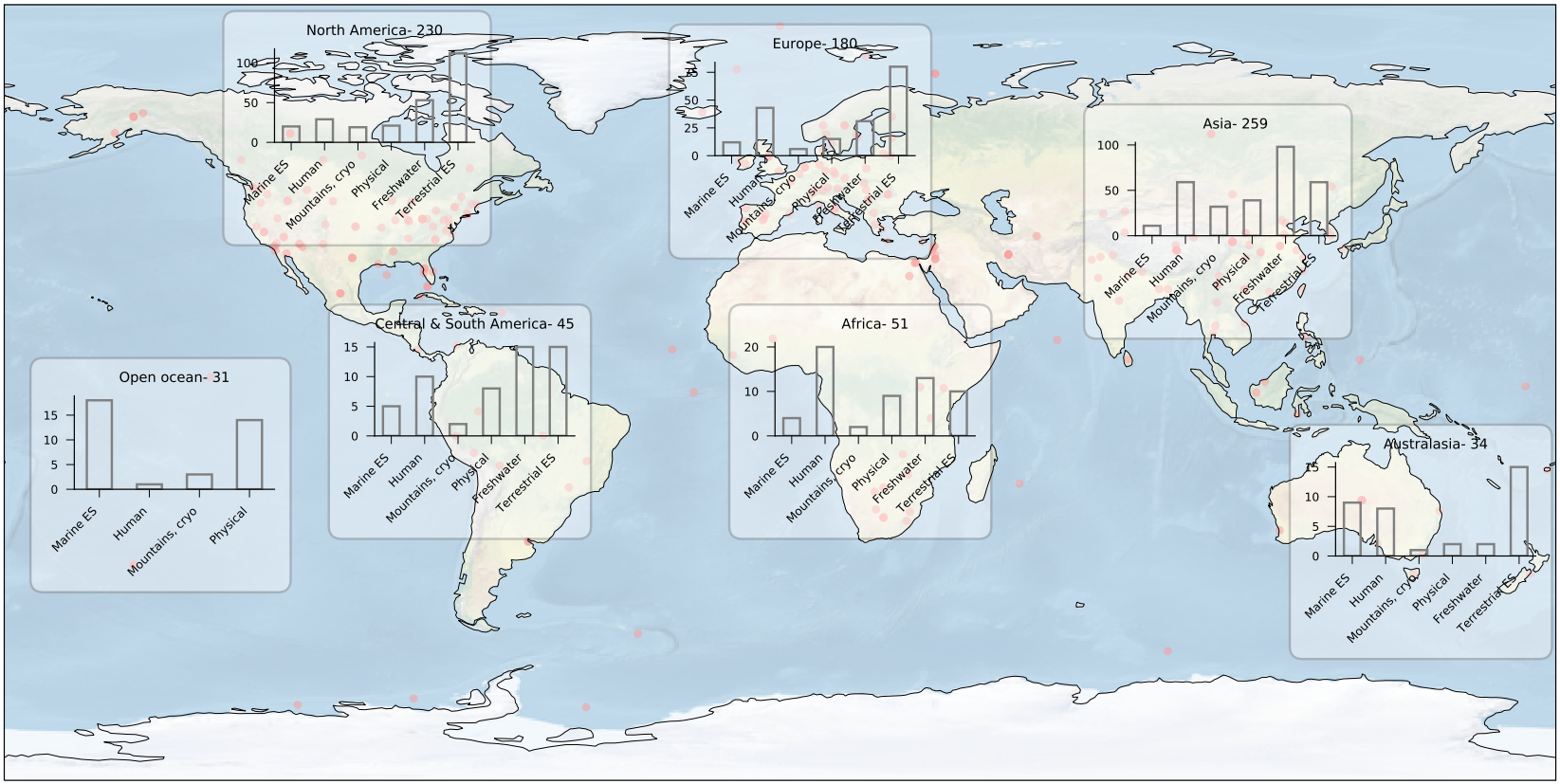
We set up our platform to record the relevance and lots of other information about each document



During the first couple of months of lockdown we screened 1500 documents



This already constitutes a useful information gathering exercise



Support Vector Machines

- We train SVMs to predict binary outcomes for relevance, and for each category
- SVMs build a hyperplane which best separates the features of data points of different classes
- Our features are 1-2 word ngrams taken from the document titles and abstracts

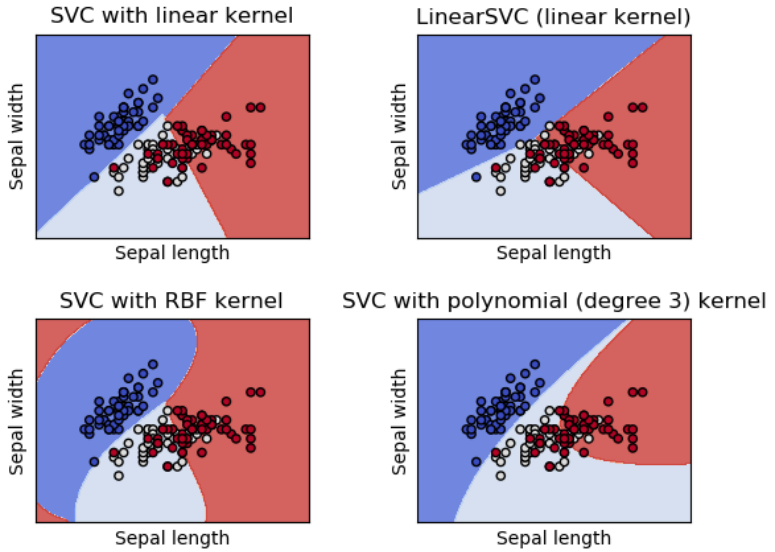
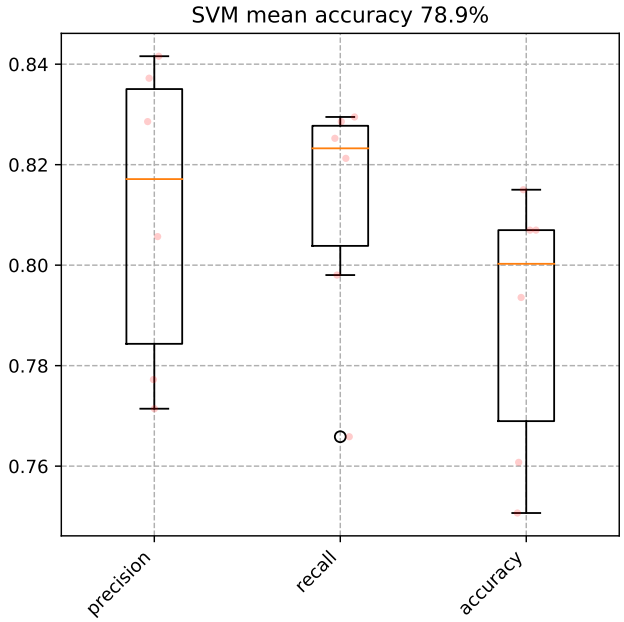
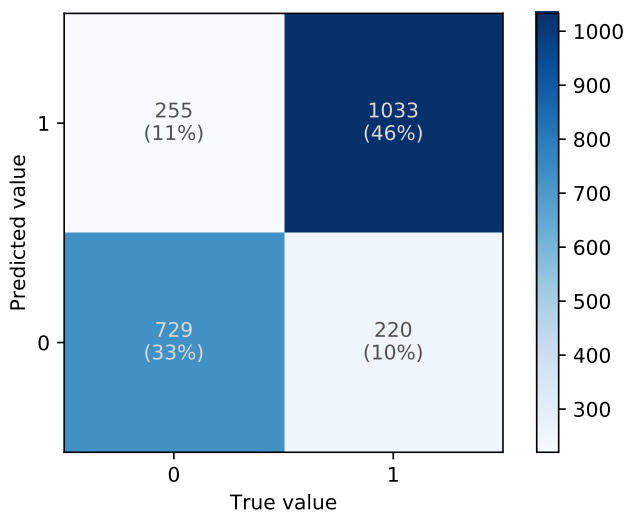


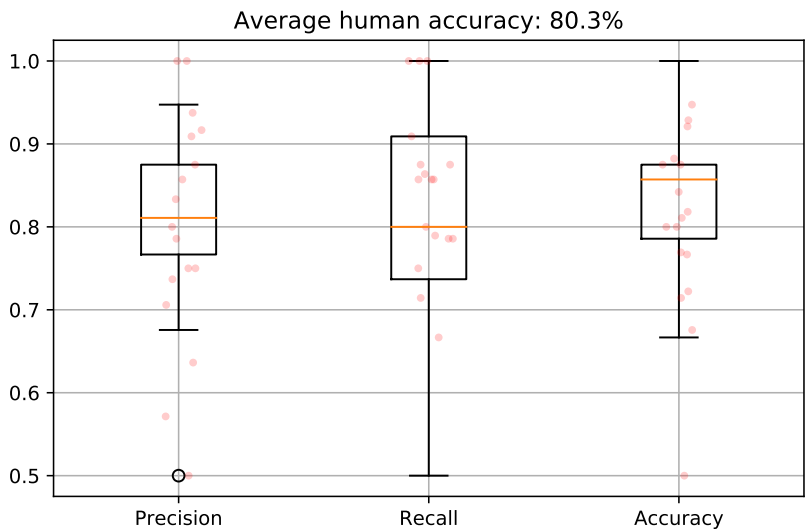
Figure: Source: `https://scikit-learn.org/stable/auto_examples/svm/plot_iris_svc.html`

We predict the relevance of a document most of the time

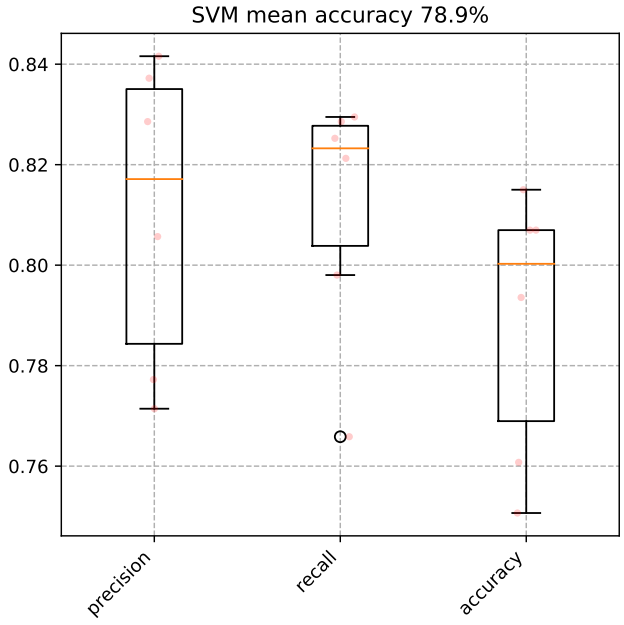


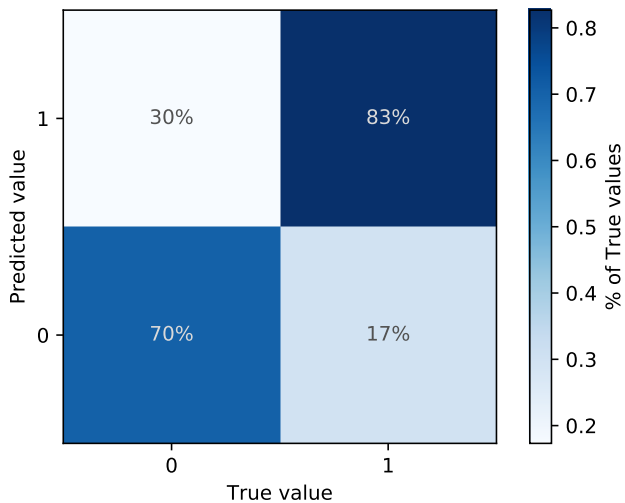


This has been steadily increasing by using the model itself as a "second pair of eyes" to check for errors, and I expect it to increase further (partly due to different criteria for inclusion at different stages of the project)



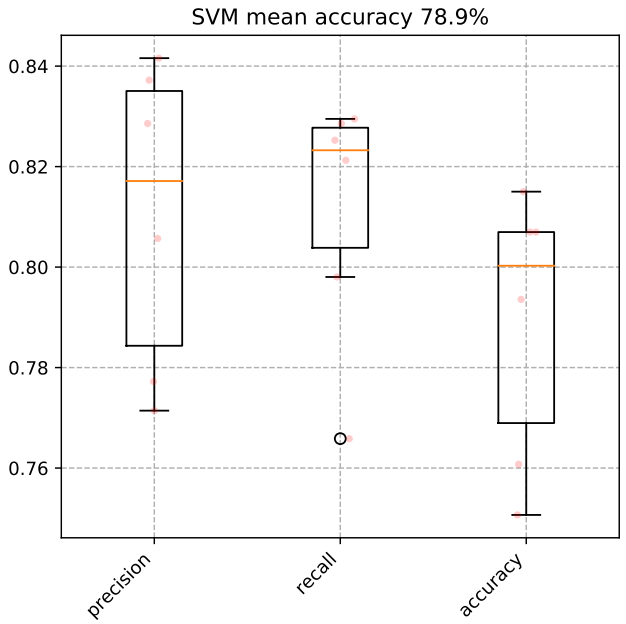
We predict the relevance of a document most of the time

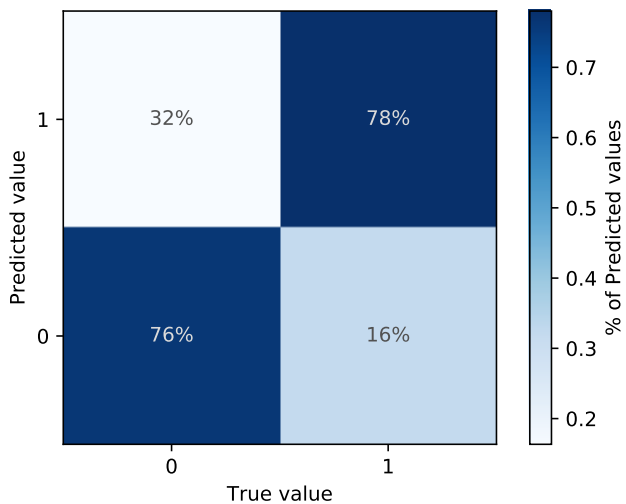




This has been steadily increasing by using the model itself as a "second pair of eyes" to check for errors, and I expect it to increase further (partly due to different criteria for inclusion at different stages of the project)

We predict the relevance of a document most of the time





This has been steadily increasing by using the model itself as a "second pair of eyes" to check for errors, and I expect it to increase further (partly due to different criteria for inclusion at different stages of the project)

We can clearly identify what impact category a document is related to

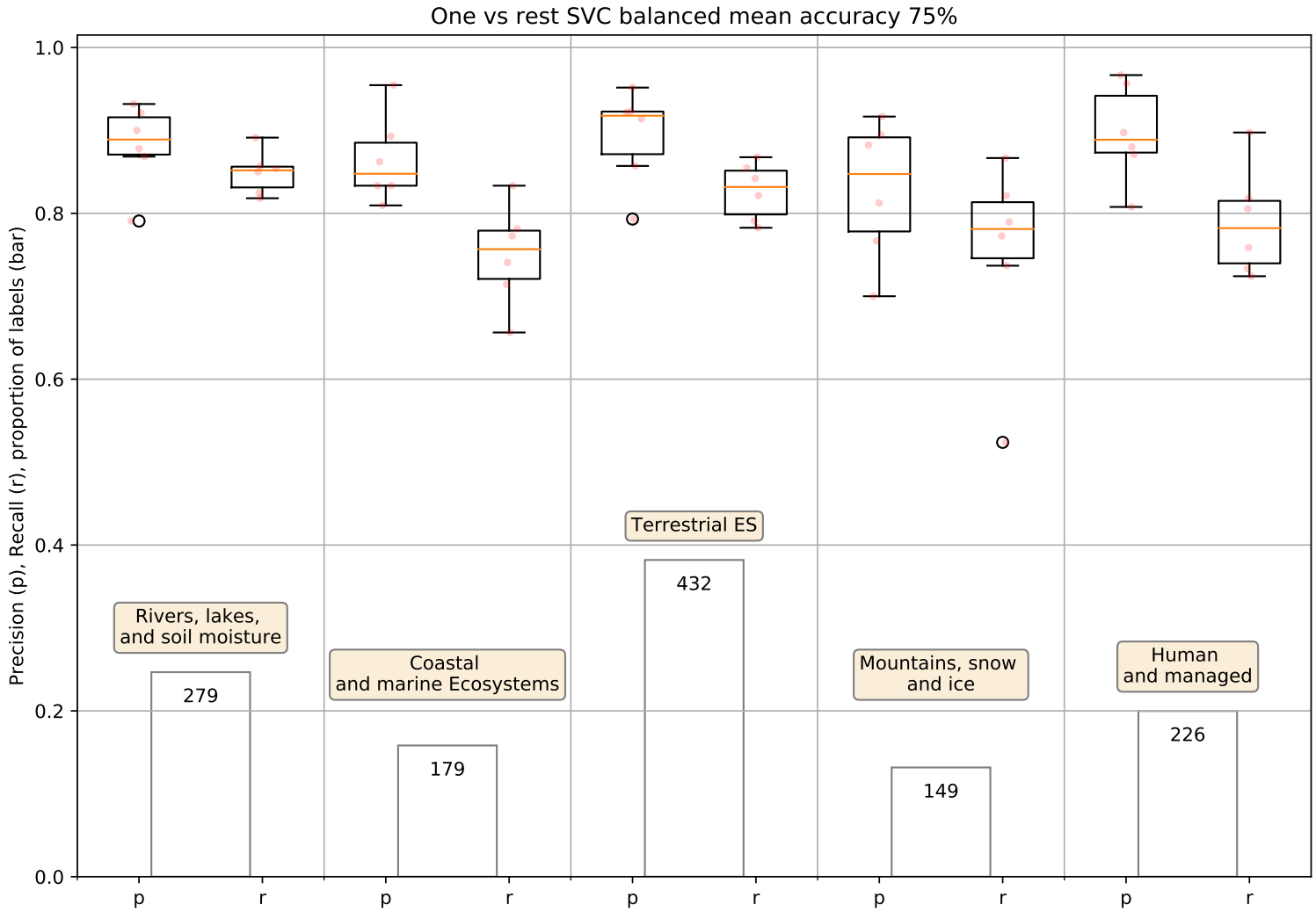
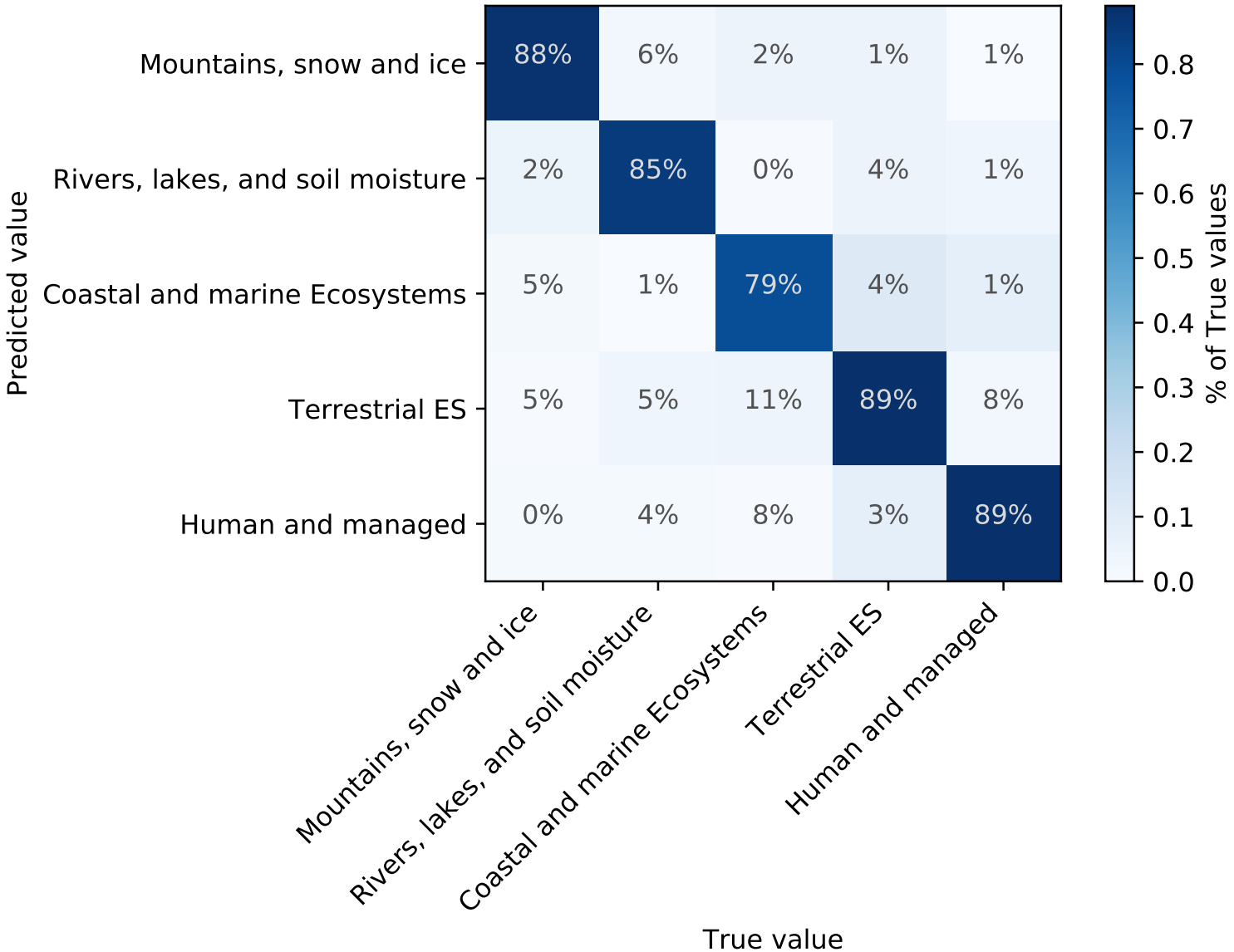
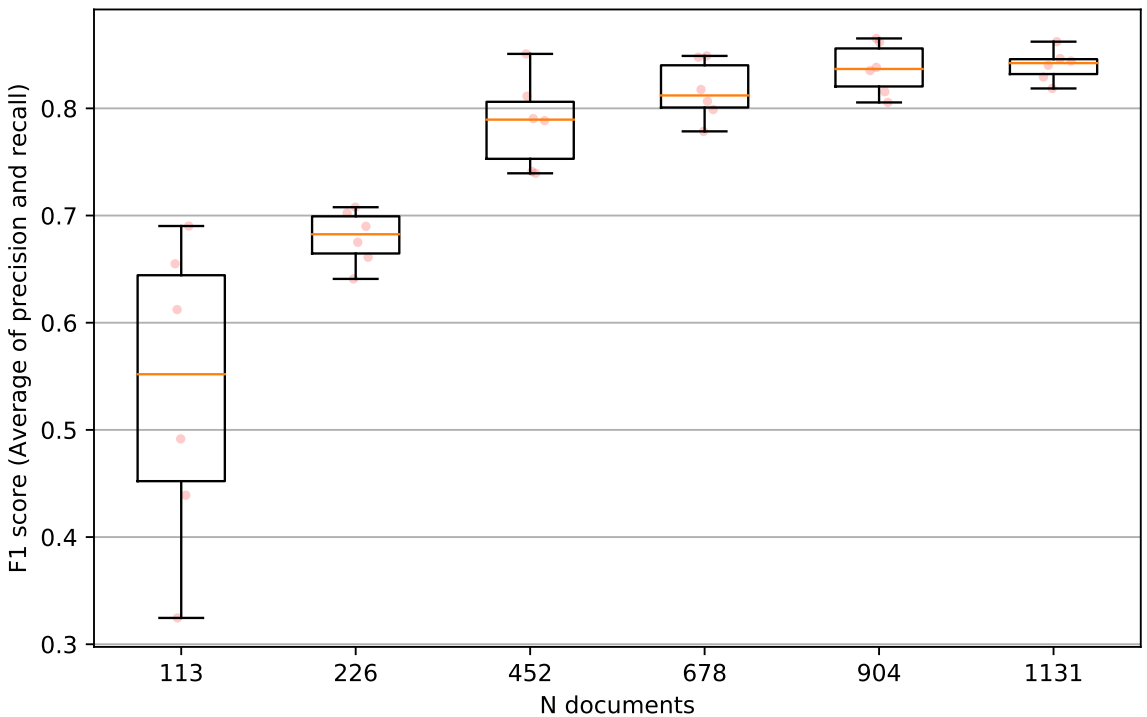


Figure: Precision (how many documents predicted to be in a category actually had that label) and Recall (how many documents with a label were predicted to be in that category) for each broad impact category

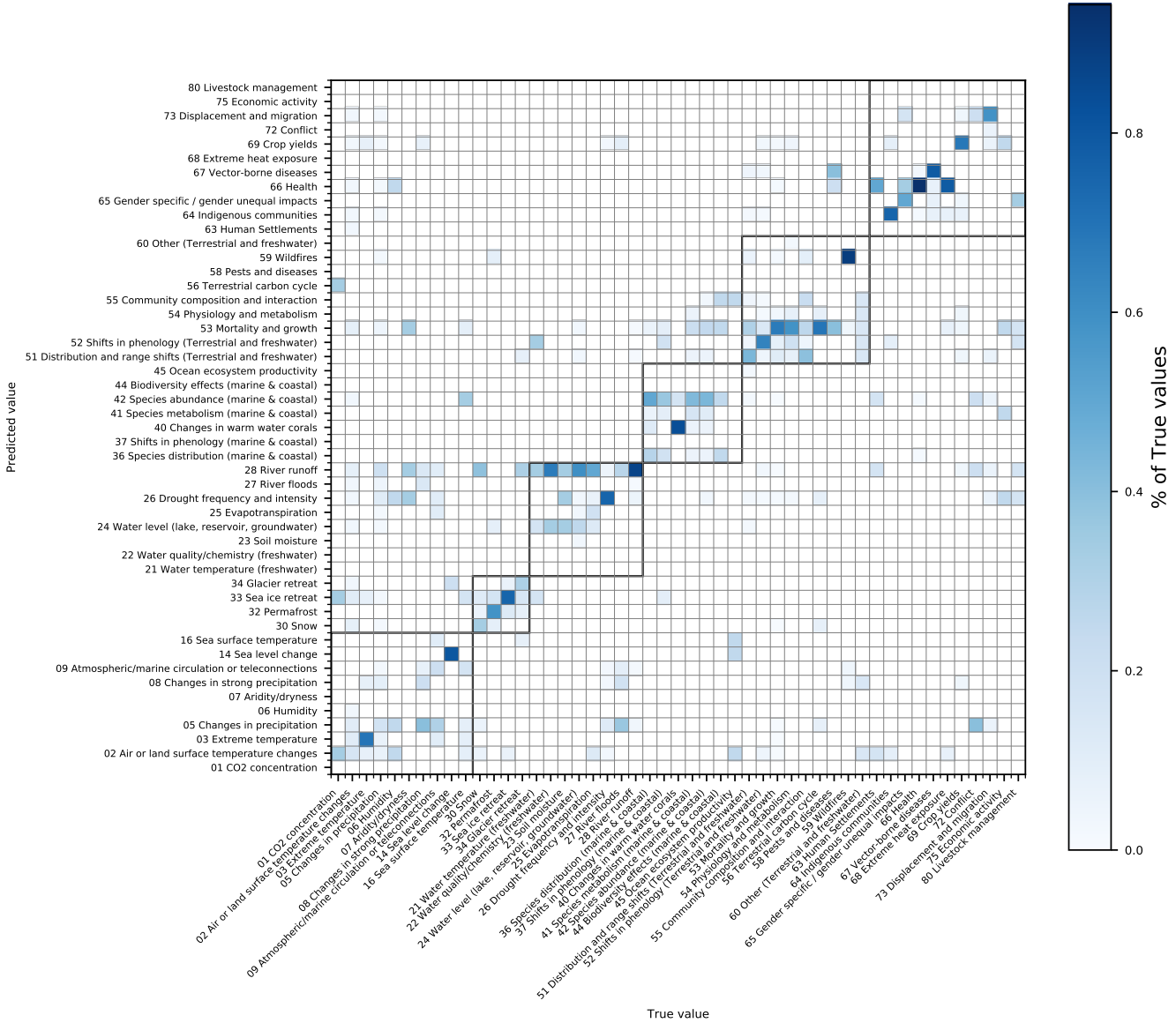
We can clearly identify what impact category a document is related to



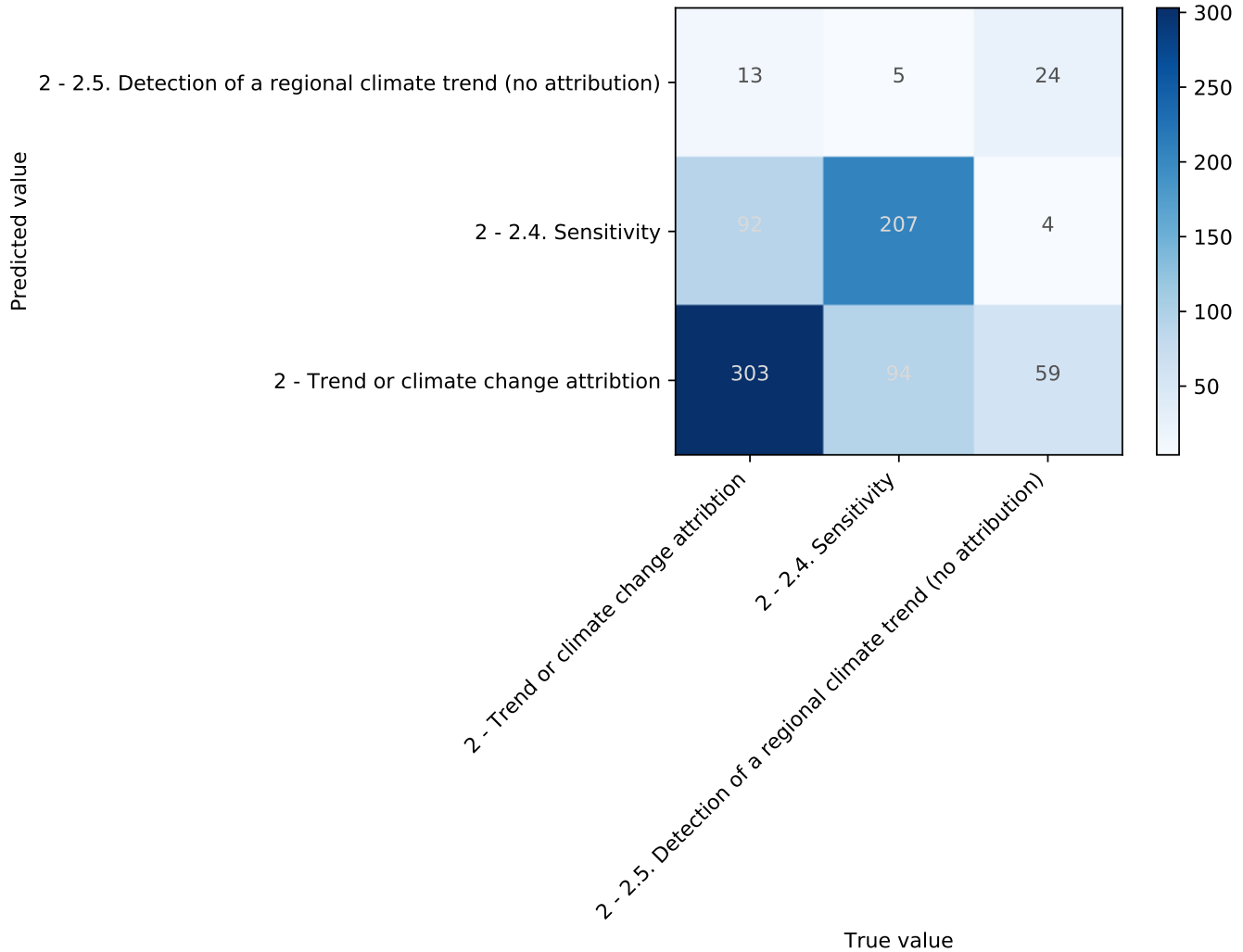
Accuracy increased and uncertainty decreased with the number of labels available



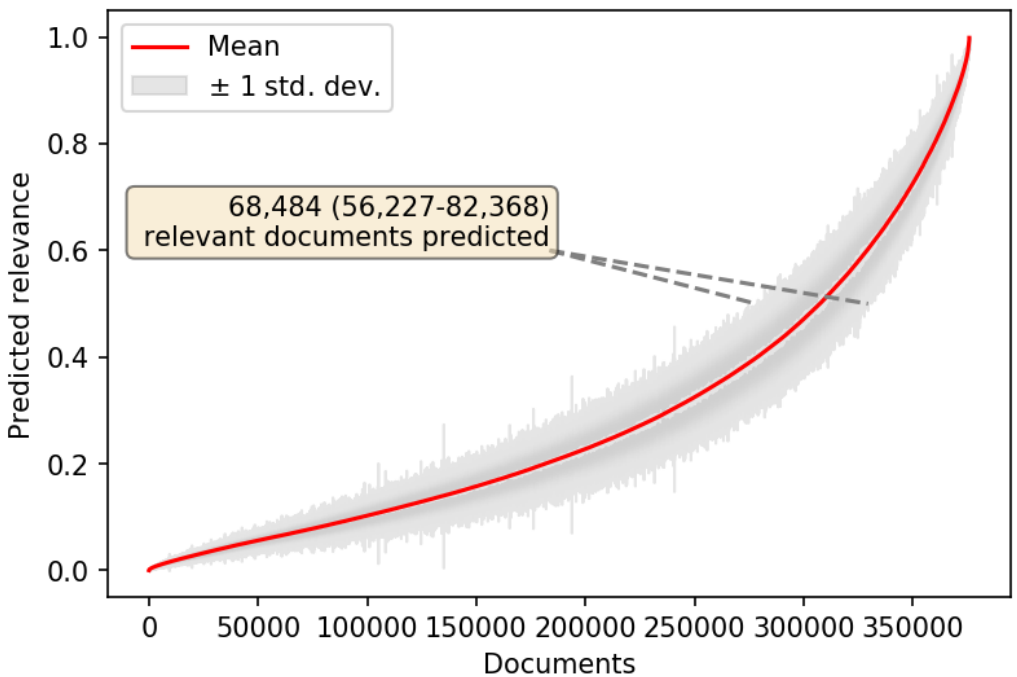
We are also broadly correct on subcategories - impressive given the amount of data and the complexity of the coding scheme



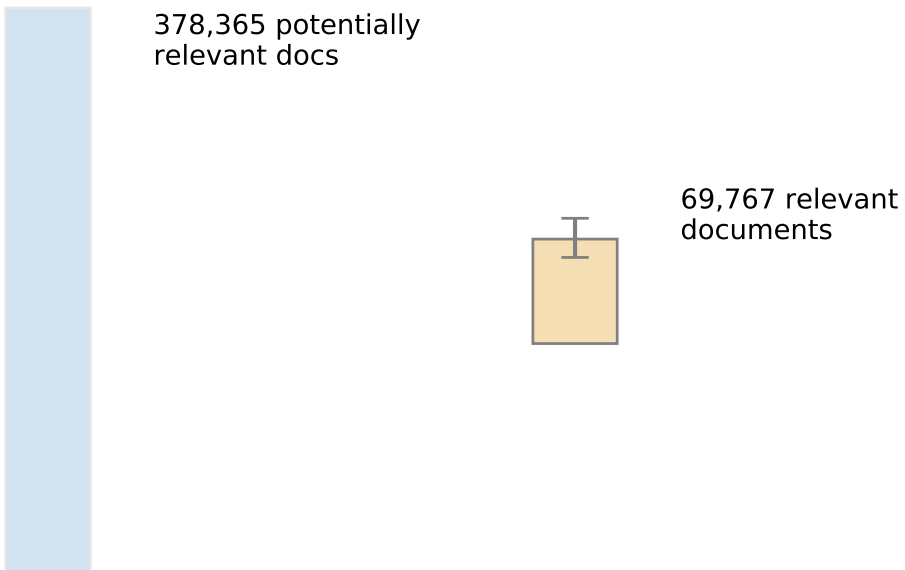
Getting attribution correct is harder, but we have fewer labels

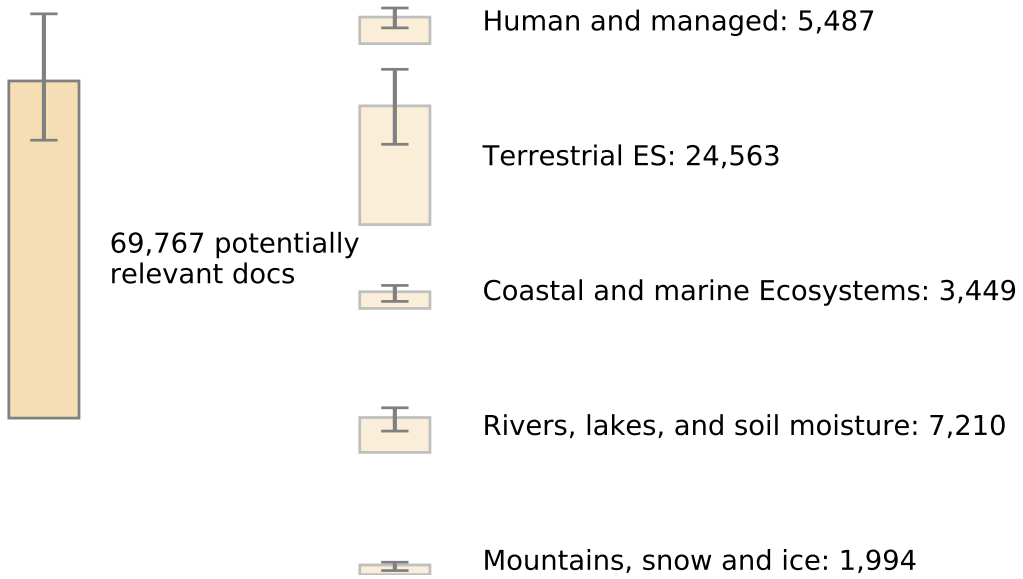


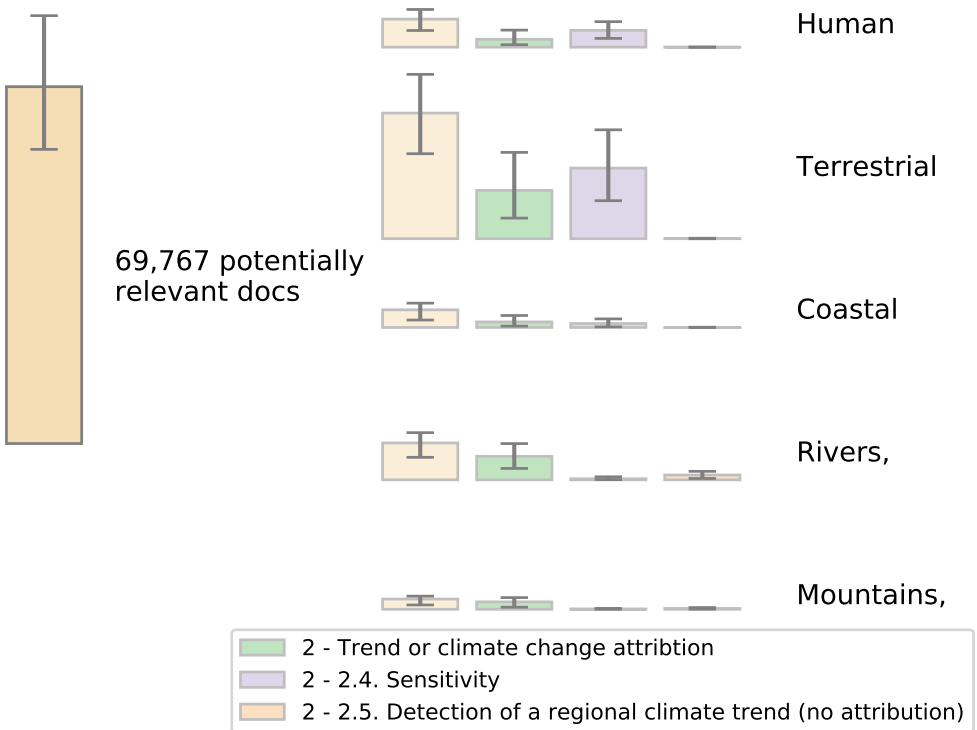
We predict tens of thousands of additional documents relevant according to the criteria we defined



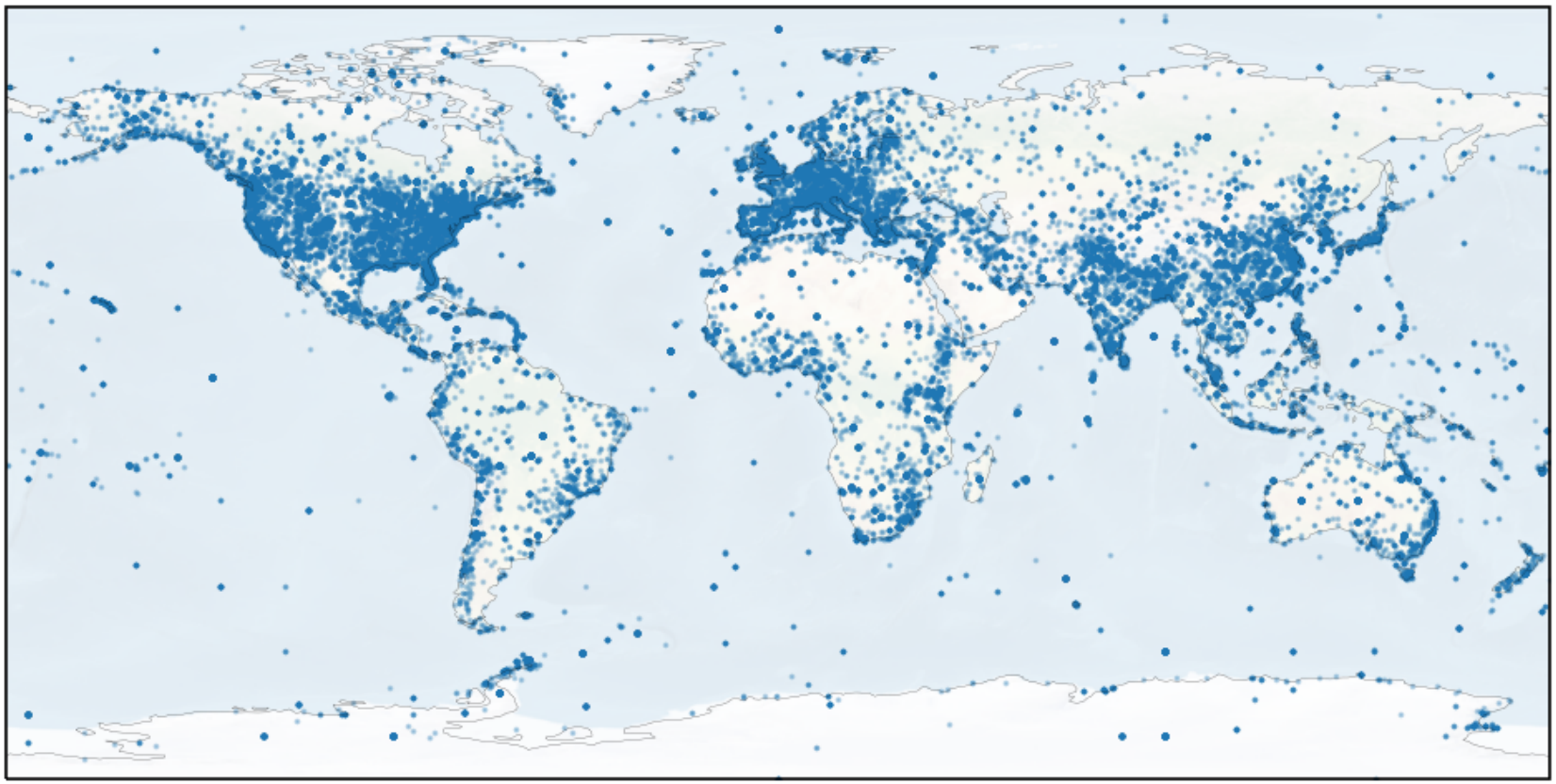
- We train 6 classifiers on random partitions of the labelled dataset
- This gives us 6 estimates for each unseen document
- The mean and standard deviation of these estimates give us an idea, with some uncertainty, of how many documents are in each category
- There must be a better way to incorporate what we know from our test statistics into our uncertainty ranges, but I can't figure it out


378,365 potentially relevant docs
69,767 relevant documents

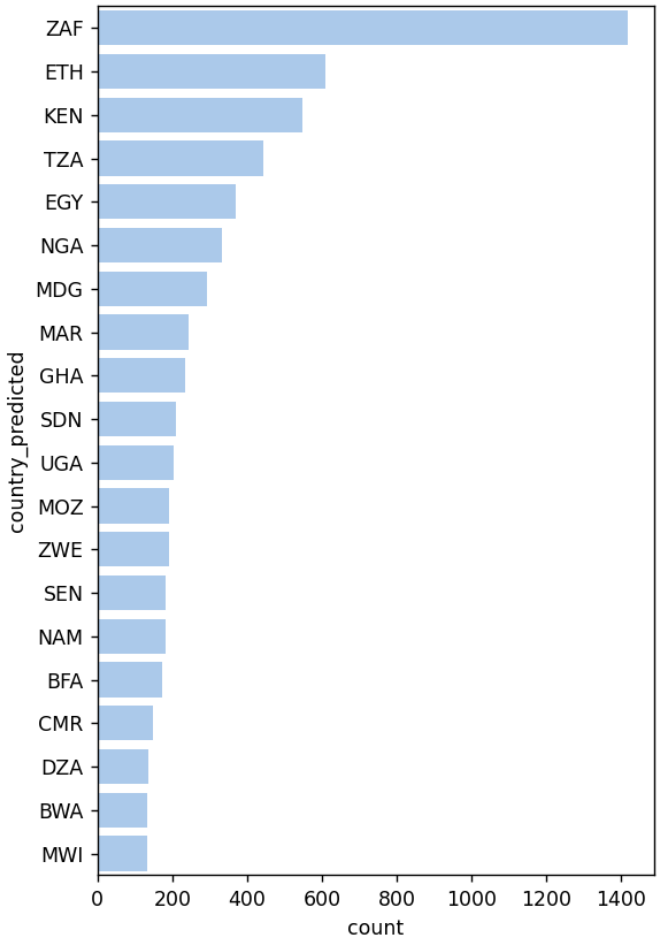

69,767 potentially relevant docs
Human and managed: 5,487
Terrestrial ES: 24,563
Coastal and marine Ecosystems: 3,449
Rivers, lakes, and soil moisture: 7,210
Mountains, snow and ice: 1,994

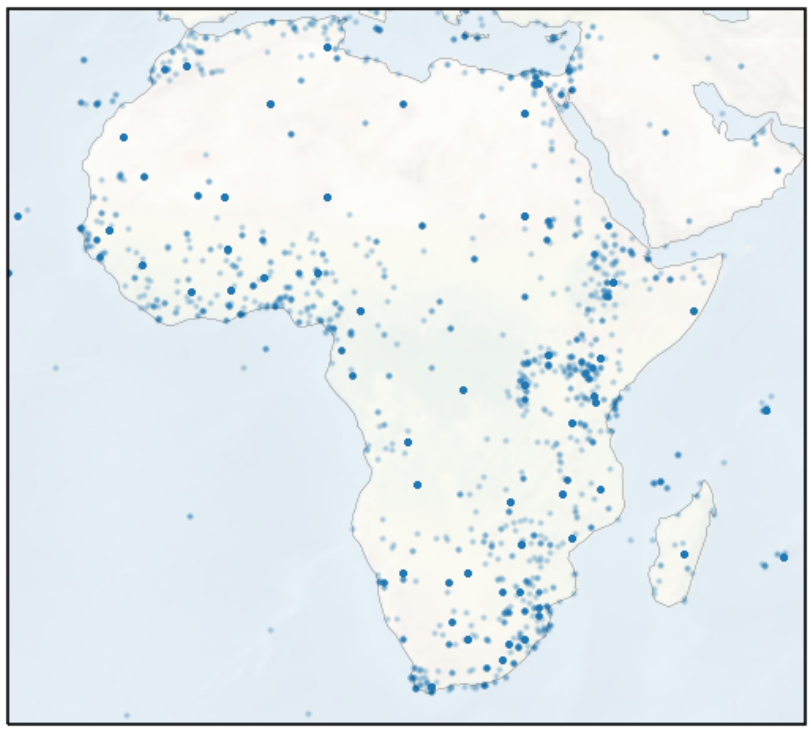

69,767 potentially relevant docs
Human
Terrestrial
Coastal
Rivers,
Mountains,
2 - Trend or climate change attribtion
2 - 2.4. Sensitivity
2 - 2.5. Detection of a regional climate trend (no attribution)

The studies predicted to be relevant cover a much broader array of places, but geographic imbalances persist

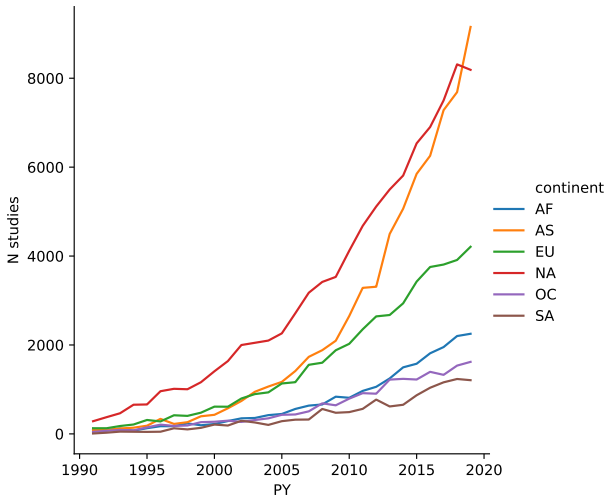


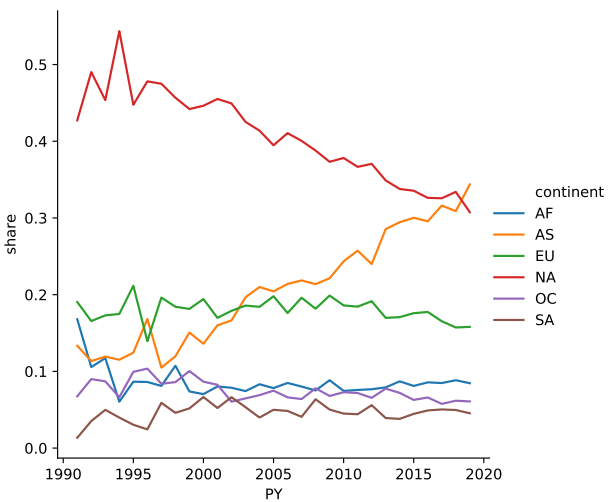
We have around 7,500 place names (from 3,250 unique documents predicted to have mentioned a location in Africa). These are also not evenly distributed.





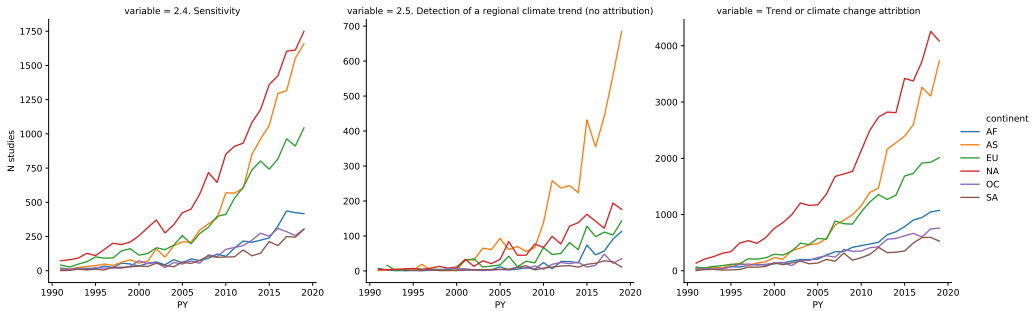
In which locations is there evidence? What impacts does it document? Since when has there been evidence?



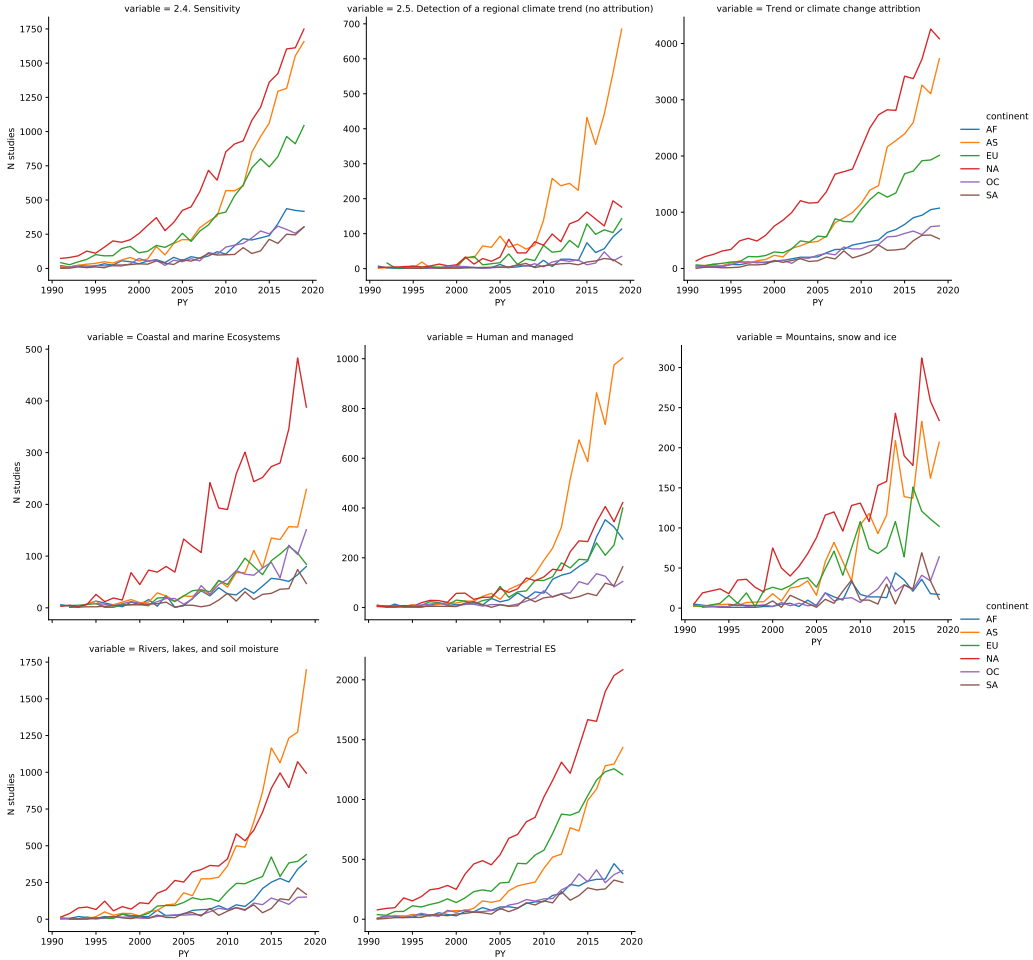


- More studies on Asia than North America since 2018
- Africa now more frequently studied than South America and Oceania

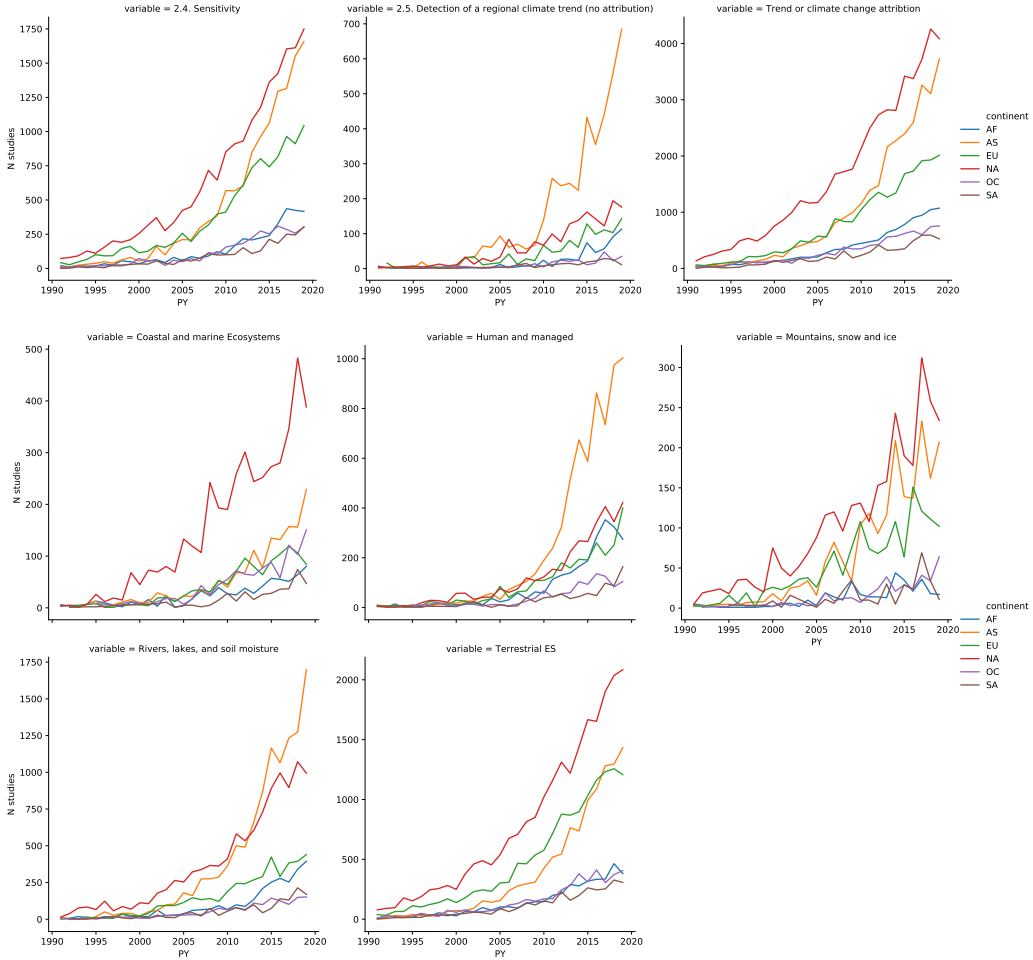
- Lots of the new studies on Asia have been about Detection (is there a regional climate trend)



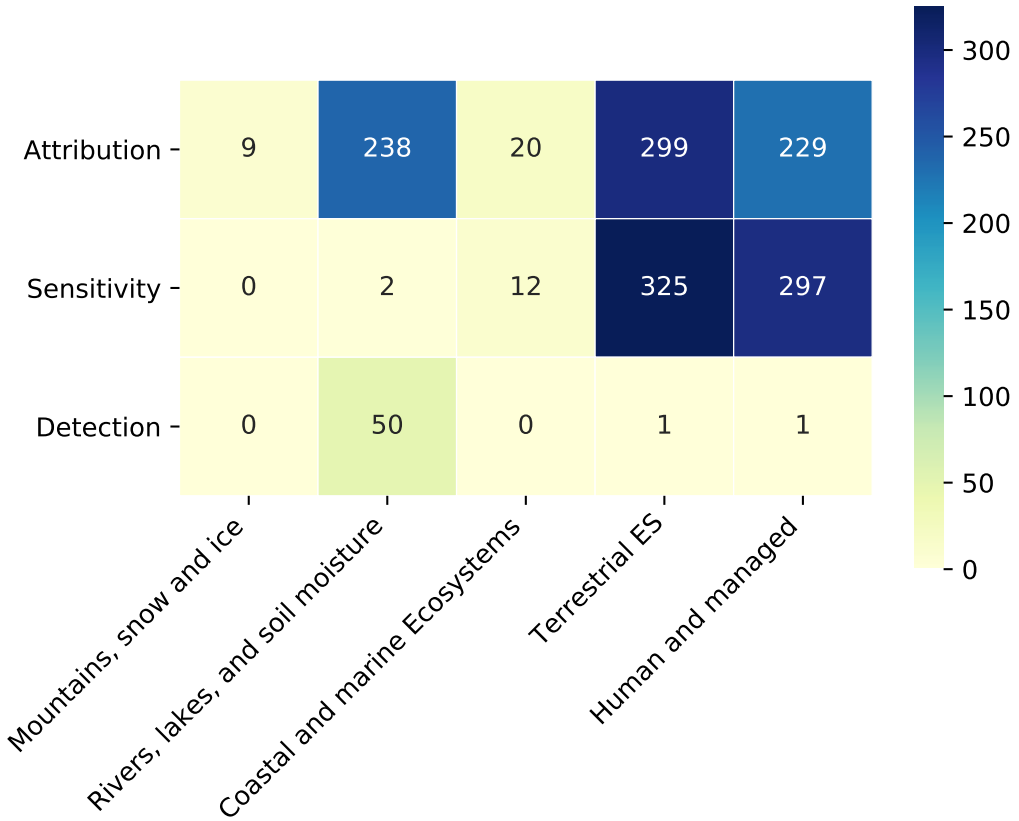
- Lots of the new studies on Asia have been about Detection (is there a regional climate trend)
- There is also more literature on human impacts and the water cycle in Asia



- Lots of the new studies on Asia have been about Detection (is there a regional climate trend)
- There is also more literature on human impacts and the water cycle in Asia
- On human impacts, Africa is as much studied as anywhere apart from Asia

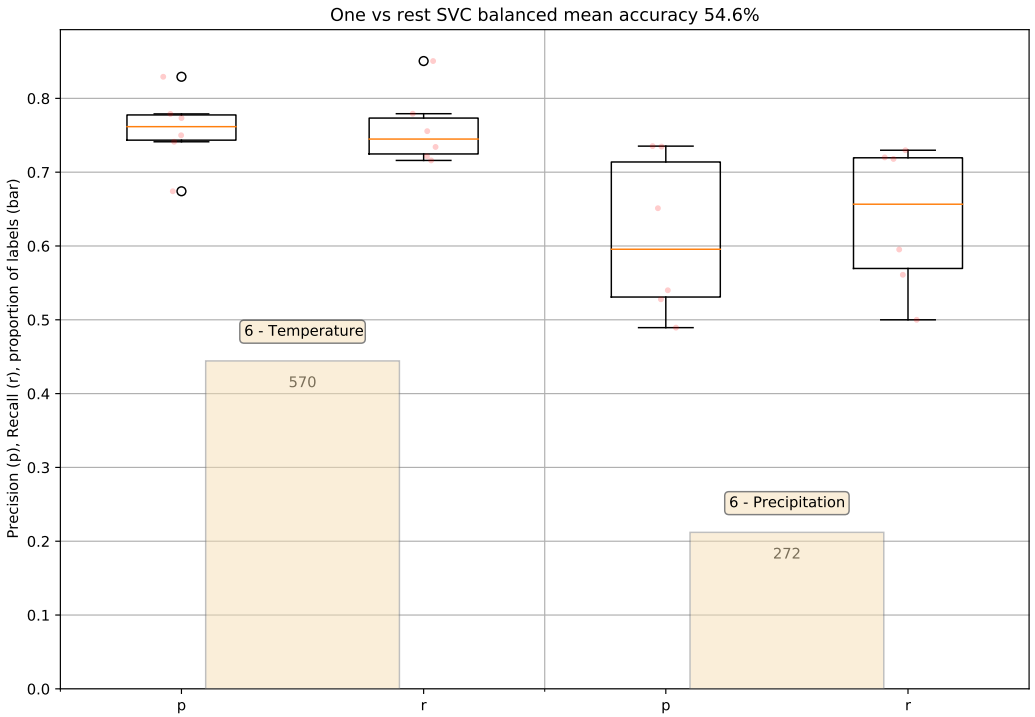


Study types



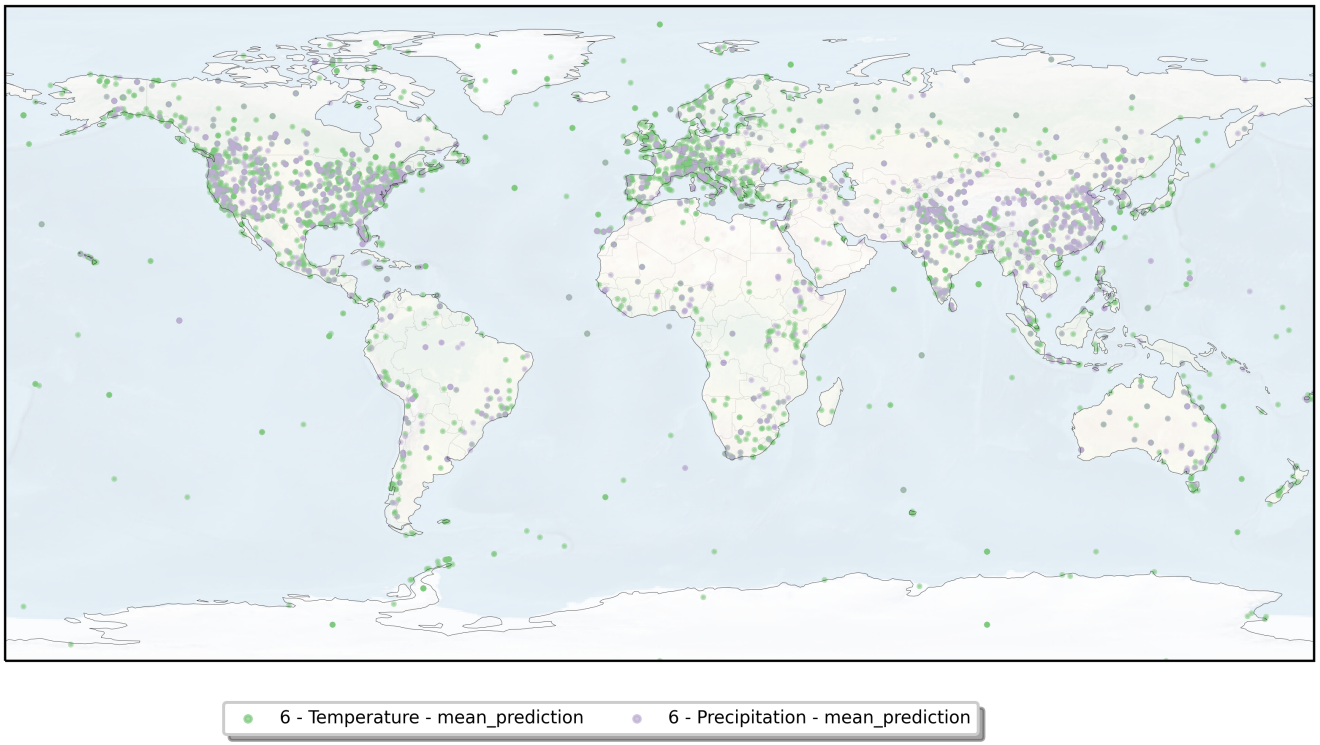
- Most literature is on terrestrial ecosystems
- There's also a lot of attribution, as well as detection, literature on rivers, lakes and soil.
- There's a large chunk on human and managed systems, although the majority examines sensitivity rather than attribution

Combining data with climate models



- Not many papers do *Full Attribution*: Human influence -> changes in regional climate variables -> regional impacts
- But human fingerprints on the climate are now everywhere (Sippel et al., 2020)
- By combining our results with climate model data, we can show where impacts might be driven by anthropogenic climate change.

Combining data with climate models



Outlook

So far

- Data collection
- Coding scheme
- Coding
- Learning and predictions
- Collation of results

Still to come

- Further data checking
- Investigating distribution of evidence and comparing with IPCC
- Predicting drivers and mapping driver-impact pathways
- Write up
- Interactive map

References

Sippel, S., Meinshausen, N., Fischer, E. M., Székely, E., and Knutti, R. (2020). Climate change now detectable from any single day of weather at global scale. *Nature Climate Change*, 10(January).