Machine learning-based evidence and attribution mapping of 100,000 climate impact studies

Max Callaghan, Carl-Friedrich Schleussner, Shruti Nath, Gerrit Hansen, Quentin Lejeune, Tom Knutson, Markus Reichstein, Emily Theokritoff, Marina Andrijevic, Robert Brecha, Michael Hegarty, Chelsea Jones, Kaylin Lee, Agathe Lucas, Nicole van Maanen, Inga Menke, Peter Pfleiderer, Burcu Yesil, Jan Minx

Context

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

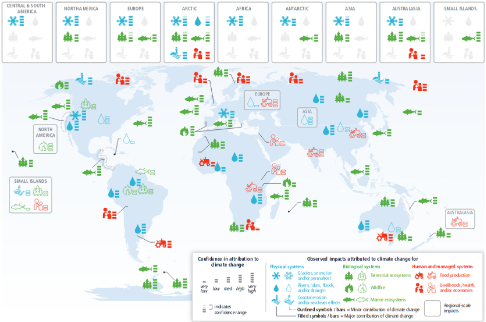


Figure 18-3 | Global patterns of observed climate change impacts reported since AR4. Each filled symbol in the top panels indicates a class of systems for which climate change has played a major role in observed changes in at least one system within that class across the respective region, with the range 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.

1011

Context

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

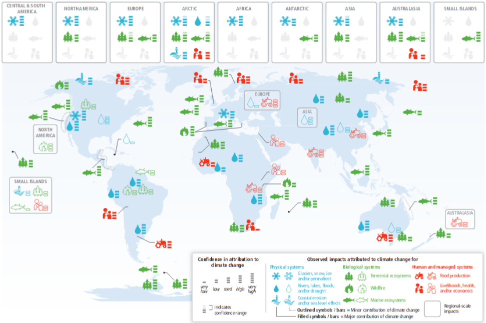
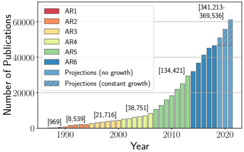


Figure 18-3 | Global patterns of observed climate change impacts reported since AR4. Each filled symbol in the top panels indicates a class of systems for which climate change has played a major role in observed changes in at least one system within that class across the respective region, with the range 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.

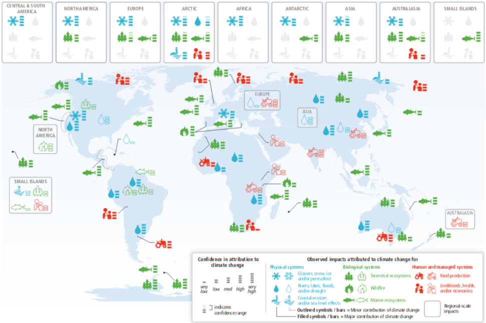
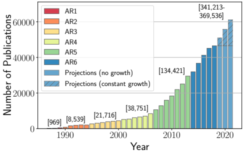


Figure 18-3 | Global patterns of observed climate change impacts reported since AR4. Each filled symbol in the top panels indicates a class of systems for which climate change has played a major role in observed changes in at least one system within that class across the respective region, with the range 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.

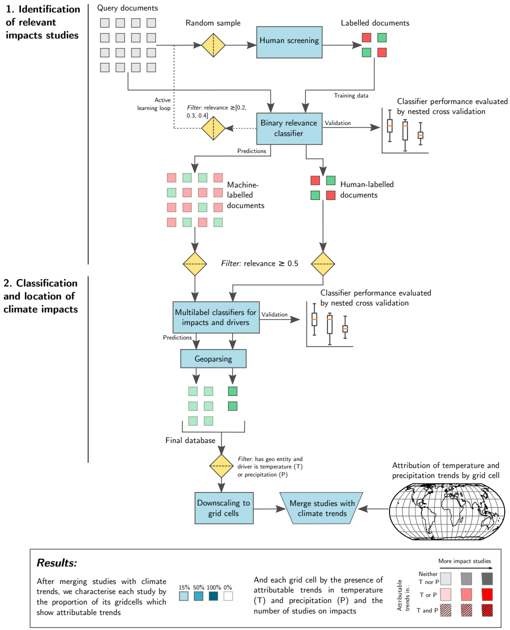


- These are challenged by big literature **?**
- They do not account for uncertainty about what literature is available

Process

1. Broad **search** in literature databases (Web of Science & Scopus) for literature on climate impacts
2. Hand **screen and code** documents to include only documents on observed climate impacts, and code the type of impacts and type of evidence
3. Combine this training data with the categorisation of documents in **AR5**
4. Use supervised **machine learning** to predict the inclusion and impact type of 100s of thousands or remaining documents
5. Use named entity recognition to extract **geographical locations** from titles and abstracts
6. Map entities to grid cells and combine with **WGI style D&A** of temperature and precipitation trends at the grid cell level
7. Describe **evidence gluts and gaps** at a grid cell level

Process

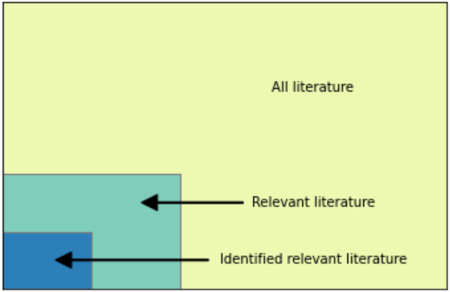

1. Identification of relevant impacts studies
Query documents
Random sample
Human screening
Labelled documents
Active learning loop
Filter: relevance ≥[0.2, 0.3, 0.4]
Training data
Binary relevance classifier
Validation
Classifier performance evaluated by nested cross validation
Predictions
Machine-labelled documents
Human-labelled documents
Filter: relevance ≥ 0.5
2. Classification and location of climate impacts
Multilabel classifiers for impacts and drivers
Validation
Classifier performance evaluated by nested cross validation
Predictions
Geoparsing
Final database
Filter: has geo entity and driver is temperature (T) or precipitation (P)
Downscaling to grid cells
Merge studies with climate trends
Attribution of temperature and precipitation trends by grid cell
Results:
After merging studies with climate trends, we characterise each study by the proportion of its gridcells which show attributable trends
15% 50% 100% 0%
And each grid cell by the presence of attributable trends in temperature (T) and precipitation (P) and the number of studies on impacts
More impact studies
Attributable trends in..
Neither T nor P
T or P
T and P

Questions for CVM

- How to aggregate, how to merge (w.r.t space and time and content) with other data?
- What documents to include?
- At what level do we want to categorise impact type?
- Can we screen more documents to improve predictions?

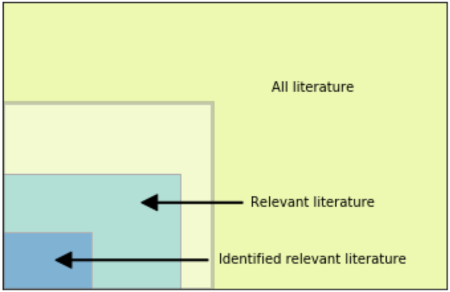
Search query I

We built a query using the documents from the tables in AR5 WGII Chapter 18. The ideal query should contain *all* documents included in those tables, along with *all* additional relevant documents (untestable) and a hopefully minimal amount of irrelevant documents



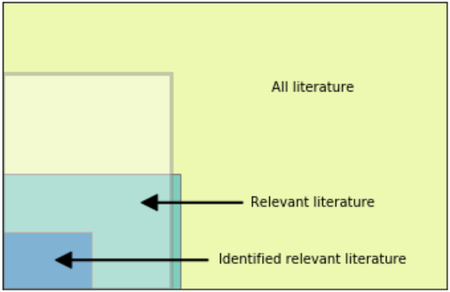
Search query I

We built a query using the documents from the tables in AR5 WGII Chapter 18. The ideal query should contain *all* documents included in those tables, along with *all* additional relevant documents (untestable) and a hopefully minimal amount of irrelevant documents



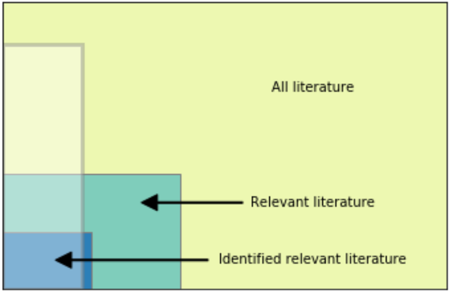
Search query I

We built a query using the documents from the tables in AR5 WGII Chapter 18. The ideal query should contain *all* documents included in those tables, along with *all* additional relevant documents (untestable) and a hopefully minimal amount of irrelevant documents



Search query I

We built a query using the documents from the tables in AR5 WGII Chapter 18. The ideal query should contain *all* documents included in those tables, along with *all* additional relevant documents (untestable) and a hopefully minimal amount of irrelevant documents



Search query II

Our query contained a list of climate variables, a list of impacts, and a list of words narrowing down the literature on observed impacts

Climate

TS=("climate model" OR "elevated* temperatur" OR "ocean* warming" OR "saline* intrusion" OR "chang* climat" OR "environment* change" OR "climat* change" OR "climat* warm" OR "warming* climat" OR "climat* varia" OR "global* warming" OR "global* change" OR "greenhouse* effect" OR "snow cover" OR "extreme temperature" OR "cyclone" OR "ocean acidification" OR "anthropogen*" OR "sea* level" OR "precipitation variabil*" OR "precipitation change*" OR "temperature* impact" OR "environmental* variab" OR "weather* pattern" OR "weather* factor*" OR "climat*") OR TS=("change* NEAR/5 cryosphere" OR "increase* NEAR/3 temperatur*"))

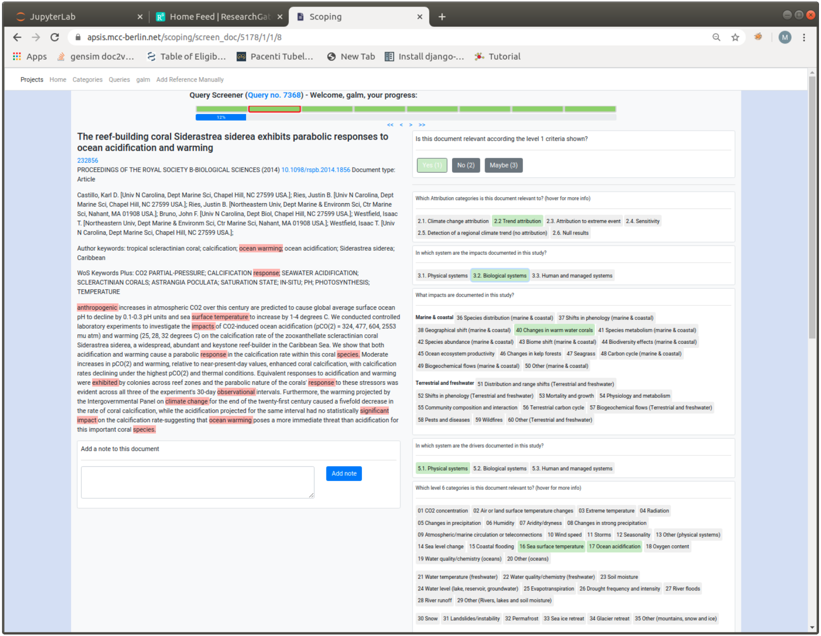
Impacts

AND (TS=("migration" OR "impact*" OR "specie*" OR "mortality*" OR "health" OR "disease*" OR "ecosystem*" OR "mass balance" OR "flood*" OR "drought" OR "disease*" OR "adaptation" OR "malaria" OR "fire" OR "water scarcity" OR "water supply" OR "permafrost" OR "biological response" OR "food availability" OR "food security" OR "vegetation dynamic*" OR "cyclone*" OR "yield*" OR "gender" OR "indigenous" OR "conflict" OR "inequality" OR "snow water equival*" OR "surface temp*") OR TS=("glacier* NEAR/3 melt*" OR "glacier* NEAR/3 mass*" OR "erosion* NEAR/5 coast*" OR "glacier* NEAR/5 retreat*" OR "rainfall* NEAR/5 reduc*" OR "coral* NEAR/5 stress*" OR "precip* NEAR/5 *crease*" OR "river NEAR/5 flow"))

Observed

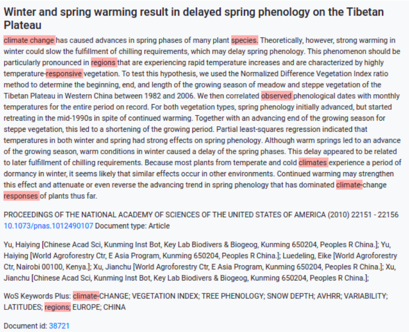
AND (TS=("recent" OR "current" OR "modern" OR "observ*" OR "evidence*" OR "past" OR "local" OR "region*" OR "significant" OR "driver*" OR "driving OR "respon*" OR "were responsible" OR "was responsible" OR "exhibited" OR "witnessed" OR "attribut*" OR "has increased" OR "has decreased" OR "histor*" OR "correlation" OR "evaluation"))

Screening & Labelling



Basic ML I - text as data

We can build a set of features is from a TFIDF weighted set of unigrams and bigrams from the documents' abstracts



Winter and spring warming result in delayed spring phenology on the Tibetan Plateau

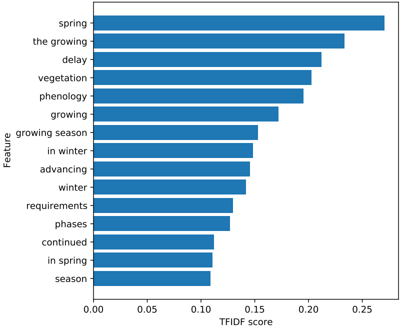
climate change has caused advances in spring phases of many plant species. Theoretically, however, strong warming in winter could slow the fulfillment of chilling requirements, which may delay spring phenology. This phenomenon should be particularly pronounced in regions that are experiencing rapid temperature increases and are characterized by highly temperature-responsive vegetation. To test this hypothesis, we used the Normalized Difference Vegetation Index ratio method to determine the beginning, end, and length of the growing season of meadow and steppe vegetation of the Tibetan Plateau in Western China between 1982 and 2006. We then correlated observed phenological dates with monthly temperatures for the entire period on record. For both vegetation types, spring phenology initially advanced, but started retreating in the mid-1990s in spite of continued warming. Together with an advancing end of the growing season for steppe vegetation, this led to a shortening of the growing period. Partial least-squares regression indicated that temperatures in both winter and spring had strong effects on spring phenology. Although warm springs led to an advance of the growing season, warm conditions in winter caused a delay of the spring phases. This delay appeared to be related to later fulfillment of chilling requirements. Because most plants from temperate and cold climates experience a period of dormancy in winter, it seems likely that similar effects occur in other environments. Continued warming may strengthen this effect and attenuate or even reverse the advancing trend in spring phenology that has dominated climate-change responses of plants thus far.

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA (2010) 22151 - 22156
10.1073/pnas.1012490107 Document type: Article

Yu, Haiying [Chinese Acad Sci, Kunming Inst Bot, Key Lab Biodivers & Biogeog, Kunming 650204, Peoples R China.]; Yu, Haiying [World Agroforestry Ctr, E Asia Program, Kunming 650204, Peoples R China.]; Luedeling, Eike [World Agroforestry Ctr, Nairobi 00100, Kenya.]; Xu, Jianchu [World Agroforestry Ctr, E Asia Program, Kunming 650204, Peoples R China.]; Xu, Jianchu [Chinese Acad Sci, Kunming Inst Bot, Key Lab Biodivers & Biogeog, Kunming 650204, Peoples R China.];

WoS Keywords Plus: climate-CHANGE; VEGETATION INDEX; TREE PHENOLOGY; SNOW DEPTH; AVHRR; VARIABILITY; LATITUDES; regions; EUROPE; CHINA

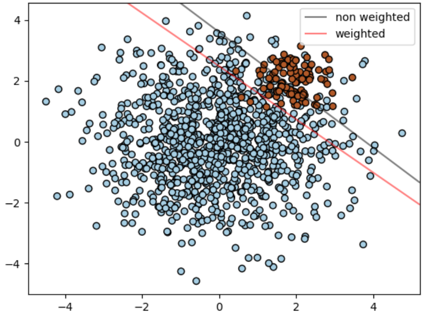
Document id: 38721



We discard very uncommon and very common features, leaving us with a vocabulary of 7,394 unique features.

Basic ML II - Support Vector Machines

SVMs try to fit a hyperplane through the multidimensional feature space (represented below in 2D) that best separates the classes in the training data.



With SVMs we ignore word order and context (bag of words assumption).

Fancy ML I - BERT

BERT (Bidirectional Representations from Transformers) is trained (by Google) on huge text corpora, and can be **"fine tuned"** on custom tasks.

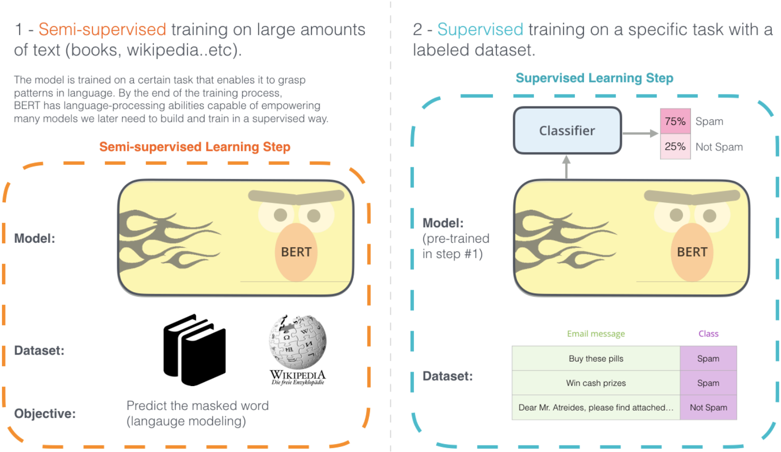


Figure: Source: `https://jalammar.github.io/illustrated-bert/`

Fancy ML II - Masked Language Modelling

BERT is trained by feeding it millions of sentences from Wikipedia etc. and asking it to predict missing words.

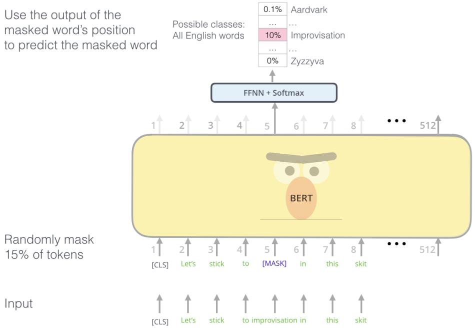


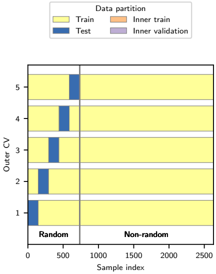
Figure: Source: `https://jalammar.github.io/illustrated-bert/`

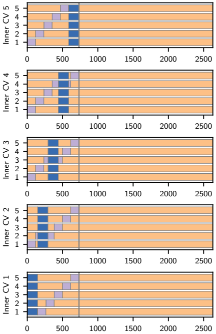
The model is trained on this task in two directions (left to right and right to left) - hence Bidirectional.

This generates embeddings (representations of words as vectors) which are contextually aware.

ML I - Training and Validation

You can assess a model by **training** it on one subset of data, and **testing** it on another. Models have different settings (hyperparameters) that affect how well they perform on a given dataset. Separating hyperparameter optimisation from performance estimation prevents overfitting, hence **nested cross-validation**.



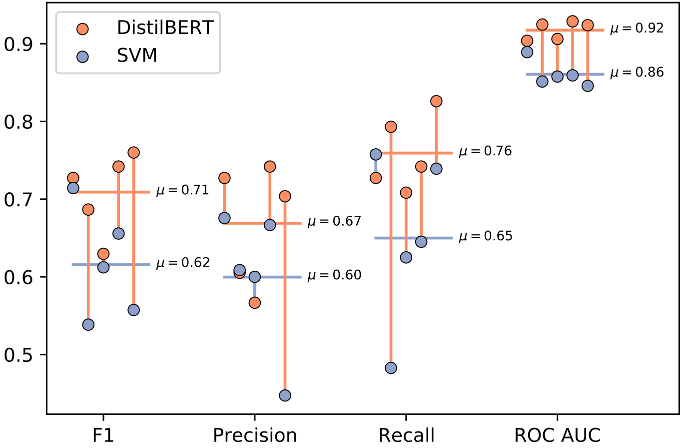


- In the inner loop, a model is trained on **Inner train** and tested on **Inner validation** with each combination of hyperparameters.
- The best performing model is selected and then trained on the outer **Train** and tested on the outer **Test**.
- For the final model used to make predictions, each candidate is tested in each outer loop and the best performing is selected.

This process validates the **hyperparameter optimisation procedure** itself, rather than a specific set of hyperparameters. The estimation of performance is more generalisable.

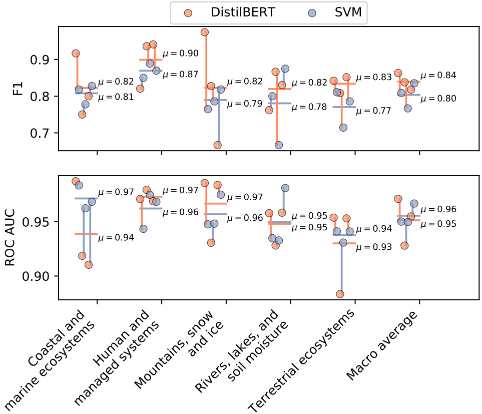
ML II - results

BERT outperformed SVM by all performance metrics



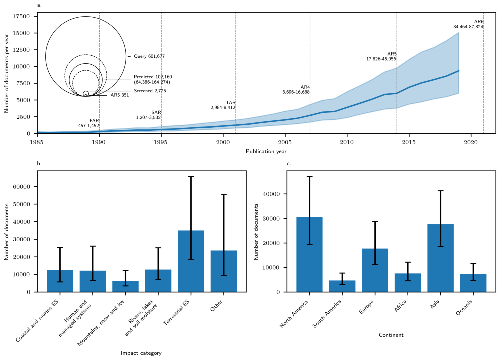
ML II - results II

BERT outperformed SVM by all performance metrics



Performance of impact type classifier higher than classifier for inclusion (provides evidence of observed impacts of climate change)

Results

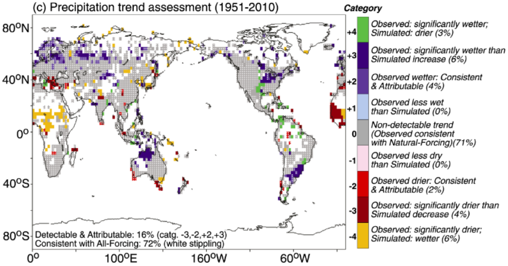


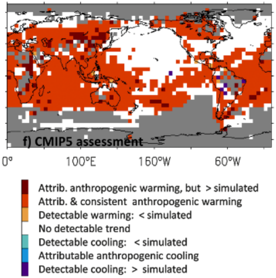
- We identify nearly 100,000 documents likely to be relevant
- We predict impact type, and extract locations

Synthesizing impacts evidence with quantitative detection and attribution evidence

We know from detection and attribution studies whether observed trends in temperature and precipitation are attributable to human influence on the climate.

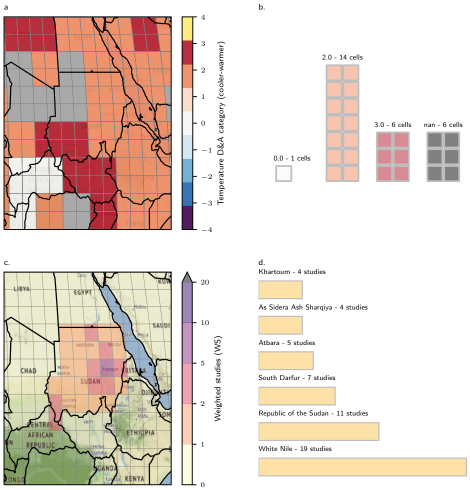
?? show this on a grid cell level





We update these calculations and combine with information from our database of impacts evidence, in which the locations, and the climate drivers have been predicted

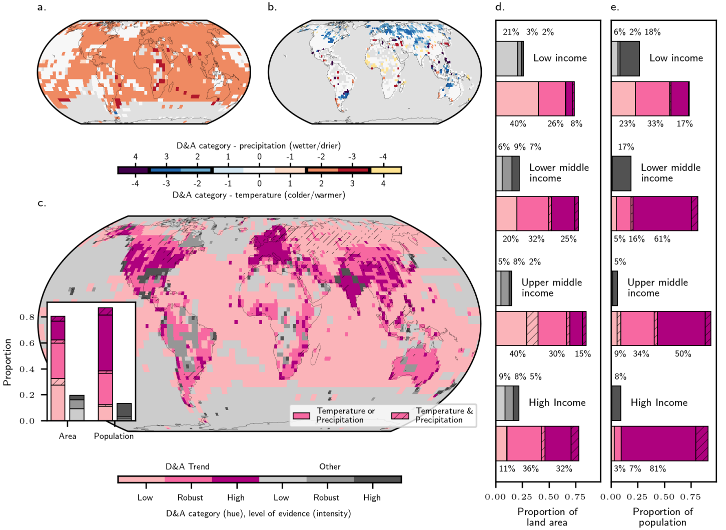
Synthesising impacts with D&A evidence



- 20 out of 27 gridcells in Sudan display an attributable increase in temperature
- So each study referring to Sudan refers to a place where around 74% of the gridcells display attributable increases in temp.

- 11 studies refer to Sudan (as the smallest identifiable geographical entity), and Sudan has 27 gridcells
- We apportion these studies to the relevant gridcells, calculating that each gridcell in Sudan has $\frac{11}{27}$ studies referring to it
- We do the same for each further geographical entity

We combine all this information to show evidence gaps and gluts



Conclusions

- We identify a large body of evidence about climate impacts, emphasising what we have seen recently: we are already feeling the effects of climate change
- What we know about the effects of a changing climate on human and natural systems does not always match with what we know about how (and where) humans are driving changes in climate variables:
  - In high income countries, 88%5 population live in areas with attributable climate changes and high evidence of the impacts of those changes in human and natural systems
  - In low income countries, 74% of population live in areas with attributable climate changes, but for almost a third of that population, there low evidence of the impacts of those changes -> **attribution gap**

But,

- Current results only show studies in Web of Science and scopus, so definitely do not show all relevant studies
- Although our query returned all papers in the relevant AR5 section, it may still miss potentially relevant literature.
- Study identification is approximate and uncertain, trends in studies may not correspond to trends attributed to human influence
- Geoparsing is also inexact, and is unable to grasp fuzzy geographical content e.g. "Western China"

Summary - Machine learning-based evidence and attribution mapping of 100,000 climate impact studies

- In a large collaborative coding exercise, we examined thousands of papers *potentially* relevant to understanding observed impacts of climate change
- We used machine learning to identify tens $>$ 100,000 studies *likely* to be relevant.
- We predicted the sector, climate driver, and location for each of these studies
- We used the location and predicted climate driver to synthesise this information with existing quantitative Detection and Attribution knowledge.

Takeaways

- Machine learning can inform and support global environmental assessments
- We have lots of evidence of observed impacts of climate change, on all continents and in all systems.
- What we know about the effects of a changing climate on human and natural systems does not always match with what we know about how (and where) humans are driving changes in climate variables

---

Thanks!

Contact: `callaghan@mcc-berlin.net`

Bibliography