

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

Max Callaghan^{1,2}, Carl-Friedrich Schleussner^{3,5}, Shruti Nath^{3,4}, Quentin Lejeune³, Thomas R. Knutson⁶, Markus Reichstein^{7,8}, Gerrit Hansen¹, Emily Theokritoff^{3,5}, Marina Andrijevic^{3,5}, Robert Brecha^{3,9}, Michael Hegarty³, Chelsea Jones³, Kaylin Lee³, Agathe Lucas³, Nicole van Maanen^{3,5}, Inga Menke³, Peter Pflieger^{3,5}, Burcu Yesil³, and Jan Minx^{1,2}

¹Mercator Research Institute on Global Commons and Climate Change, Torgauer StraSse, 10829 Berlin, Germany

²Priestley International Centre for Climate, University of Leeds, Leeds LS2 9JT, United Kingdom

³Climate Analytics, Berlin, Germany

⁴Institute of Atmospheric and Climate Sciences, ETH Zürich, Switzerland

⁵Integrative Research Institute on Transformations of Human-Environment Systems, Humboldt University, Berlin, Germany

⁶NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 08540, USA

⁷Max Planck Institute for Biogeochemistry, Department Biogeochemical Integration, D-07701 Jena, Germany

⁸Michael Stifel Center Jena for Data-driven and Simulation Science, Jena, Germany

⁹Hanley Sustainability Institute, Renewable and Clean Energy Program and Physics Dept., University of Dayton, Dayton, Ohio, USA

1 Search query

```
(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*")
)
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")
)
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")
)

```

2 Supplementary Data

2.1 Category Scheme

The file **category_aggregation.csv** lists the categories used to code relevant documents. Each category could be used as an impact or a driver. To make the classification problem tractable, the categories were merged into broad categories resembling those used in IPCC AR5.

2.2 Document metadata

The file **0c_doc_info.csv** contains basic metadata about each of the studies considered. For rights reasons the abstracts are not included, although these are used in the analysis.

2.3 Predictions

Predicted scores are given in the following files for relevance **1_document_relevance.csv**, impact category **1_impact_predictions.csv**, and climate drivers **1_driver_predictions.csv**.

2.4 Grid cells

The file **2_merged_da_data.csv** shows the attribution category and weighted studies score (for temperature, precipitation and for all studies) for each grid cell.

2.5 Study grid cell matching

The file **study_gridcell_2.5.csv** has a row for each grid cell study combination, where the column 'ndf_id' refers to the 'index' column of the aforementioned grid cell dataset.

2.6 Study D&A

The file **2_study_da.csv** shows the number of gridcells with each detection and attribution category that each study refers to.

3 Supplementary Tables

		Gridcell sums
South America	D&A	3674 (2400-5760)
	Other	1061 (605-1995)
North America	D&A	21745 (14364-31884)
	Other	8868 (5002-15196)
Africa	D&A	5323 (3391-8104)
	Other	2251 (1220-4105)
Europe	D&A	13991 (9105-21466)
	Other	3762 (2089-7232)
Asia	D&A	20885 (14705-29783)
	Other	6764 (3999-11548)
Oceania	D&A	5482 (3579-8202)
	Other	1922 (1047-3441)
Global	D&A	57366 (38371-85227)
	Other	20419 (11697-35705)
Without location	D&A	0 (0-0)
	Other	23954 (13897-42921)

Table 1: The number of studies in each impact category and each continent

		Coastal and marine Ecosystems (WS>1)		Human and managed systems (WS>1)		Mountains, snow and ice (WS>1)		Rivers, lakes, and soil moisture (WS>1)		Terrestrial ecosystems (WS>1)		Other systems (WS>1)		Total (WS>5)	
		area	population	area	population	area	population	area	population	area	population	area	population	area	population
South America	D&A	8% (5%-31%)	26% (18%-51%)	26% (13%-62%)	55% (38%-79%)	11% (8%-15%)	19% (14%-32%)	16% (6%-45%)	33% (14%-68%)	63% (38%-64%)	77% (67%-81%)	45% (11%-64%)	63% (27%-81%)	52% (18%-63%)	75% (46%-81%)
	Other	3% (1%-27%)	6% (2%-15%)	24% (5%-34%)	14% (9%-17%)	13% (6%-14%)	6% (3%-10%)	19% (3%-34%)	15% (5%-19%)	36% (29%-36%)	19% (17%-19%)	28% (8%-36%)	17% (7%-19%)	33% (12%-36%)	19% (13%-19%)
North America	D&A	33% (19%-50%)	58% (43%-70%)	40% (25%-57%)	64% (48%-72%)	34% (18%-54%)	43% (27%-56%)	52% (24%-64%)	60% (41%-69%)	70% (61%-70%)	70% (64%-72%)	58% (32%-71%)	72% (58%-72%)	62% (55%-70%)	70% (69%-72%)
	Other	11% (5%-12%)	27% (9%-27%)	9% (7%-15%)	28% (26%-28%)	8% (5%-19%)	19% (8%-24%)	15% (9%-18%)	28% (28%-28%)	21% (16%-21%)	28% (28%-28%)	16% (11%-21%)	28% (27%-28%)	17% (16%-21%)	28% (28%-28%)
Africa	D&A	12% (4%-23%)	22% (12%-37%)	32% (25%-46%)	68% (58%-79%)	3% (1%-9%)	7% (6%-14%)	17% (5%-35%)	44% (18%-68%)	40% (22%-51%)	71% (46%-79%)	35% (15%-52%)	64% (26%-79%)	32% (24%-41%)	66% (51%-75%)
	Other	2% (0%-4%)	1% (0%-4%)	7% (3%-14%)	3% (3%-5%)	0% (0%-1%)	0% (0%-1%)	9% (1%-13%)	4% (2%-7%)	7% (3%-20%)	4% (2%-9%)	9% (2%-23%)	6% (2%-8%)	6% (2%-12%)	4% (2%-6%)
Europe	D&A	16% (11%-21%)	59% (41%-71%)	25% (20%-33%)	80% (73%-88%)	21% (8%-30%)	58% (25%-79%)	24% (18%-31%)	77% (64%-86%)	35% (28%-56%)	88% (83%-95%)	29% (21%-45%)	85% (73%-93%)	28% (25%-34%)	83% (78%-90%)
	Other	0% (0%-0%)	0% (0%-0%)	1% (0%-1%)	0% (0%-0%)	0% (0%-1%)	0% (0%-0%)	0% (0%-0%)	0% (0%-0%)	1% (1%-3%)	0% (0%-0%)	0% (0%-1%)	0% (0%-0%)	0% (0%-1%)	0% (0%-0%)
Asia	D&A	19% (6%-29%)	47% (18%-58%)	61% (44%-78%)	80% (75%-82%)	27% (14%-43%)	39% (12%-56%)	61% (34%-72%)	75% (50%-81%)	71% (60%-81%)	80% (72%-83%)	73% (51%-83%)	80% (73%-84%)	70% (60%-76%)	80% (78%-82%)
	Other	3% (1%-4%)	9% (4%-13%)	5% (5%-7%)	15% (15%-15%)	4% (2%-4%)	14% (6%-14%)	5% (5%-5%)	15% (15%-15%)	6% (4%-7%)	15% (14%-15%)	6% (5%-7%)	15% (15%-15%)	5% (5%-6%)	15% (15%-15%)
Oceania	D&A	76% (72%-76%)	91% (59%-91%)	74% (53%-84%)	70% (54%-86%)	3% (2%-8%)	37% (15%-53%)	19% (4%-83%)	53% (43%-82%)	84% (79%-86%)	77% (74%-92%)	86% (73%-86%)	89% (61%-91%)	84% (74%-86%)	84% (75%-92%)
	Other	14% (14%-14%)	5% (5%-5%)	14% (14%-14%)	5% (5%-5%)	1% (1%-3%)	4% (4%-4%)	14% (1%-14%)	5% (4%-5%)	14% (14%-14%)	5% (5%-5%)	14% (14%-14%)	5% (5%-5%)	14% (14%-14%)	5% (5%-5%)
Global	D&A	22% (13%-33%)	44% (23%-57%)	38% (27%-53%)	73% (64%-79%)	17% (9%-27%)	32% (13%-46%)	32% (16%-49%)	62% (40%-75%)	53% (41%-62%)	75% (63%-81%)	49% (29%-63%)	75% (59%-82%)	48% (37%-56%)	74% (67%-80%)
	Other	5% (2%-8%)	8% (3%-11%)	8% (4%-12%)	11% (11%-12%)	4% (2%-7%)	9% (4%-10%)	9% (3%-12%)	11% (10%-12%)	12% (9%-15%)	12% (11%-13%)	11% (5%-16%)	12% (11%-13%)	10% (6%-13%)	12% (11%-12%)

Table 2: The percentage of land area and population with robust evidence by impact category and continent

		Coastal and marine Ecosystems (WS>1)		Human and managed systems (WS>1)		Mountains, snow and ice (WS>1)		Rivers, lakes, and soil moisture (WS>1)		Terrestrial ecosystems (WS>1)		Other systems (WS>1)		Total (WS>5)	
		area	population	area	population	area	population	area	population	area	population	area	population	area	population
Low income	D&A	8% (2%-20%)	17% (5%-27%)	35% (29%-45%)	53% (46%-61%)	6% (3%-11%)	6% (3%-17%)	22% (7%-37%)	36% (11%-54%)	37% (25%-48%)	51% (41%-61%)	38% (17%-54%)	51% (27%-64%)	34% (27%-43%)	50% (43%-59%)
	Other	3% (0%-5%)	12% (11%-17%)	5% (4%-11%)	20% (20%-21%)	1% (1%-2%)	16% (5%-17%)	4% (2%-8%)	20% (18%-23%)	7% (3%-19%)	21% (18%-26%)	8% (4%-20%)	22% (19%-24%)	5% (3%-7%)	20% (19%-22%)
Lower middle income	D&A	26% (9%-37%)	59% (19%-65%)	50% (42%-74%)	77% (71%-80%)	22% (6%-29%)	49% (12%-58%)	38% (19%-59%)	67% (37%-77%)	64% (41%-71%)	78% (58%-80%)	58% (31%-72%)	76% (62%-80%)	58% (40%-65%)	77% (70%-79%)
	Other	5% (1%-8%)	10% (2%-16%)	14% (9%-19%)	18% (18%-18%)	12% (3%-12%)	17% (7%-17%)	13% (8%-20%)	18% (17%-19%)	19% (14%-21%)	19% (17%-19%)	15% (7%-22%)	18% (17%-19%)	16% (9%-21%)	18% (18%-19%)
Upper middle income	D&A	11% (5%-23%)	29% (17%-49%)	34% (21%-51%)	80% (69%-88%)	13% (9%-24%)	19% (11%-42%)	31% (16%-47%)	73% (52%-85%)	53% (39%-64%)	87% (81%-91%)	45% (25%-59%)	83% (67%-90%)	45% (33%-53%)	85% (77%-89%)
	Other	2% (1%-9%)	4% (1%-6%)	9% (3%-11%)	6% (4%-6%)	2% (2%-3%)	2% (0%-4%)	9% (2%-12%)	6% (4%-6%)	10% (8%-13%)	6% (5%-6%)	10% (4%-13%)	6% (5%-6%)	10% (5%-12%)	6% (5%-6%)
High Income	D&A	45% (34%-56%)	77% (60%-85%)	50% (37%-66%)	85% (80%-90%)	32% (16%-48%)	68% (37%-82%)	46% (24%-68%)	81% (71%-86%)	73% (66%-76%)	87% (86%-89%)	67% (47%-78%)	89% (81%-91%)	67% (60%-74%)	88% (85%-89%)
	Other	9% (6%-10%)	9% (6%-9%)	8% (6%-13%)	9% (8%-9%)	5% (3%-13%)	7% (6%-8%)	12% (5%-14%)	9% (8%-9%)	15% (12%-17%)	9% (9%-9%)	13% (9%-17%)	9% (9%-9%)	13% (12%-16%)	9% (9%-9%)
Global	D&A	22% (13%-33%)	44% (23%-57%)	38% (27%-53%)	73% (64%-79%)	17% (9%-27%)	32% (13%-46%)	32% (16%-49%)	62% (40%-75%)	53% (41%-62%)	75% (63%-81%)	49% (29%-63%)	75% (59%-82%)	48% (37%-56%)	74% (67%-80%)
	Other	5% (2%-8%)	8% (3%-11%)	8% (4%-12%)	11% (11%-12%)	4% (2%-7%)	9% (4%-10%)	9% (3%-12%)	11% (10%-12%)	12% (9%-15%)	12% (11%-13%)	11% (5%-16%)	12% (11%-13%)	10% (6%-13%)	12% (11%-12%)

Table 3: The percentage of land area and population with robust evidence by impact category and income category

		Gridcell sums	
		area	population
South America	D&A	64%	81%
	Other	36%	19%
North America	D&A	71%	72%
	Other	29%	28%
Africa	D&A	73%	89%
	Other	27%	11%
Europe	D&A	95%	100%
	Other	5%	0%
Asia	D&A	93%	85%
	Other	7%	15%
Oceania	D&A	86%	95%
	Other	14%	5%
Global	D&A	80%	87%
	Other	20%	13%

Table 4: The percentage of land area and population with attributable impacts in temperature and/or precipitation by continent

		Gridcell sums	
		area	population
5. Low income	D&A	74%	73%
	Other	26%	27%
4. Lower middle income	D&A	78%	81%
	Other	22%	19%
3. Upper middle income	D&A	85%	94%
	Other	15%	6%
1. High Income	D&A	78%	91%
	Other	22%	9%
Global	D&A	80%	87%
	Other	20%	13%

Table 5: The percentage of land area and population with attributable impacts in temperature and/or precipitation by income category