A Topography of Climate Change Research

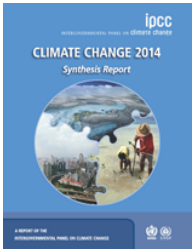
Max Callaghan

with Jan Minx, Piers Forster

January 27, 2019

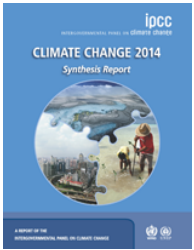
Figure: Portrait of map-makers, Gerardus Mercator and Jodocus Hondius (Jodocus Hondius) source: Wikipedia Commons

Figure: Portrait of map-makers, Gerardus Mercator and Jodocus Hondius (Jodocus Hondius) source: Wikipedia Commons



- The IPCC has a *cartographic* role at the science policy interface (Edenhofer and Minx, 2014; Edenhofer and Kowarsch, 2015)

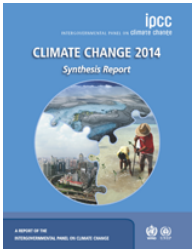
Figure: Portrait of map-makers, Gerardus Mercator and Jodocus Hondius (Jodocus Hondius) source: Wikipedia Commons



- The IPCC has a *cartographic* role at the science policy interface (Edenhofer and Minx, 2014; Edenhofer and Kowarsch, 2015)

- A topography is a description of a landscape

Figure: Portrait of map-makers, Gerardus Mercator and Jodocus Hondius (Jodocus Hondius) source: Wikipedia Commons



- The IPCC has a *cartographic* role at the science policy interface (Edenhofer and Minx, 2014; Edenhofer and Kowarsch, 2015)

- A topography is a description of a landscape
- Topics (from the Greek "topos", place) can describe the features of a body of text

Outline

1. Motivation
2. Methods
3. Results

1 Motivation

2 Methods

3 Results

Context - Big Literature

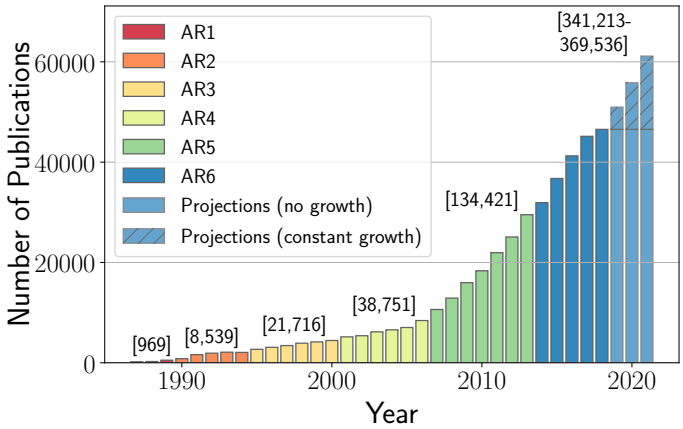


Figure: Articles on climate change in the Web of Science

A challenge for

- Global environmental assessments
- Our understanding of global environmental assessments
- Evidence synthesis more generally

The IPCC in the age of Big Literature

- We entrust the IPCC with providing a *comprehensive* and *transparent* assessment of the literature

The IPCC in the age of Big Literature

- We entrust the IPCC with providing a *comprehensive* and *transparent* assessment of the literature
- Although IPCC reports cite ever greater numbers of papers, this number decreases in proportion to the number of papers in literature

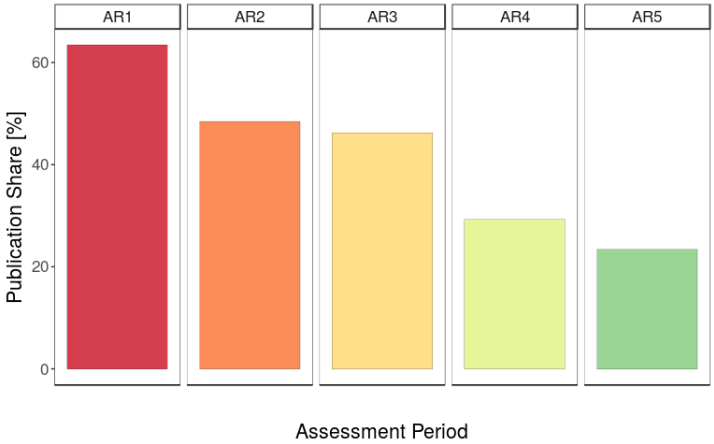


Figure: (Minx et al., 2017)

What Do We Know About the IPCC?

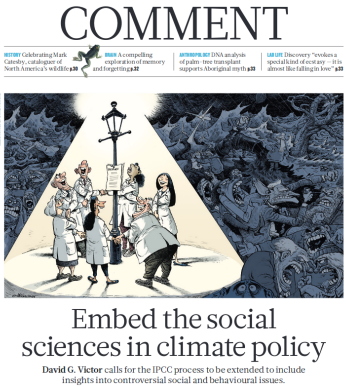
What Do We Know About the IPCC?

- The social sciences are seen as under-represented in IPCC reports

Figure: (David G. Victor, 2015)

What Do We Know About the IPCC?

- The social sciences are seen as under-represented in IPCC reports
- (Bjurström and Polk, 2011) are often cited as evidence of this under-representation, e.g. as demonstrating "a powerful bias to the natural sicences in the construction of 'IPCC knowledge' " Hulme and Mahony (2010)


COMMENT

HISTORY Celebrating Mark Catesby, cataloguer of North America's wildlife p30 | VISION A compelling exploration of memory and forgetting p32 | ANTHROPOLOGY DNA analysis of palm-tree transplant supports Aboriginal myth p33 | LAB LIFE Discovery "evokes a special kind of ecstasy — it is almost like falling in love" p33

Embed the social sciences in climate policy

David G. Victor calls for the IPCC process to be extended to include insights into controversial social and behavioural issues.

Figure: (David G. Victor, 2015)

What Do We Know About the IPCC?

- The social sciences are seen as under-represented in IPCC reports
- (Bjurström and Polk, 2011) are often cited as evidence of this under-representation, e.g. as demonstrating "a powerful bias to the natural sicences in the construction of 'IPCC knowledge' " Hulme and Mahony (2010)
- The evidence is simply the relative shares of the different disciplines in IPCC citations


COMMENT

HISTORY Celebrating Mark Catesby, cataloguer of North America's wildlife p.30 | BOOKS A compelling exploration of memory and forgetting p.32 | ANTHROPOLOGY DNA analysis of palm-tree transplant supports Aboriginal myth p.33 | LAB LIFE Discovery "evokes a special kind of ecstasy — it is almost like falling in love" p.33

Embed the social sciences in climate policy

David G. Victor calls for the IPCC process to be extended to include insights into controversial social and behavioural issues.

Figure: (David G. Victor, 2015)

The Age of Climate Solutions?

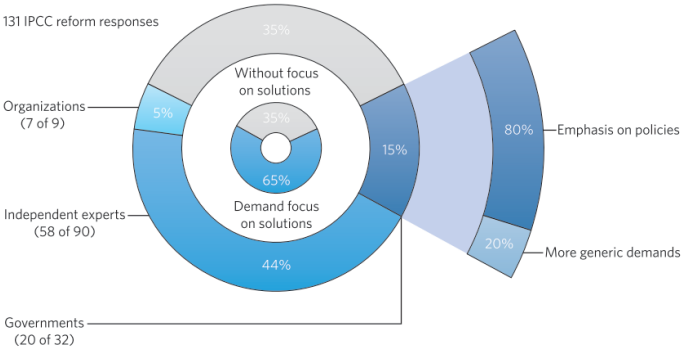


Figure: (Kowarsch et al., 2017)

- Demand for solutions is increasing

The Age of Climate Solutions?

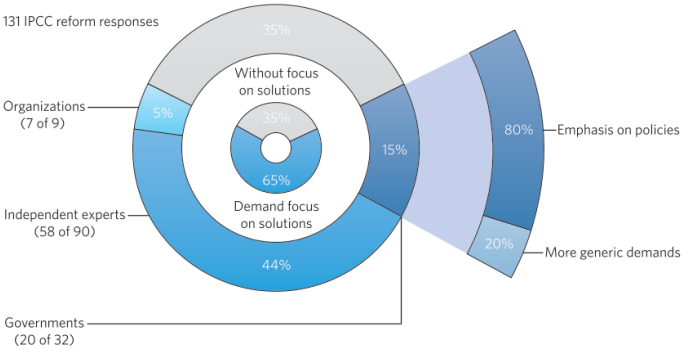


Figure: (Kowarsch et al., 2017)

- Demand for solutions is increasing
- We know little about the supply of solutions in the literature

1 Motivation

2 Methods

3 Results

Investigating bibliometric and text data

To understand the representation of social science and solutions relevant knowledge in IPCC reports, we look at journal classification and document abstracts

Data:
400,000 papers on climate change from the Web of Science (query following Grieneisen and Zhang (2011)), matched with 70,000 IPCC citations (Using Doc2Vec)

Topic modelling:
We use topic modelling (with NMF (Lee and Seung, 1999)) to understand the thematic content of papers

Topographic mapping:
We project the documents' topical locations into 2 dimensions using t-SNE (van der Maaten and Hinton, 2008)

Measuring representation:
We compare the proportions of categories of documents in the whole of the literature with the subset of the literature that is cited by the IPCC

Words, words, words

	AR1	AR2	AR3	AR4	AR5	AR6
Years	1986-1989	1990-1994	1995-2000	2001-2006	2007-2013	2014-
Documents	1,167	8,539	21,716	38,750	134,413	201,606
Unique words	2,000	12,480	23,346	34,637	71,867	94,746
New words	change (560)	oil (287)	downscaling (217)	sres (234)	biochar (1,791)	mmms (313)
	climate (428)	deltac (283)	degreesc (187)	petm (95)	redd (1,113)	cop21 (234)
	co2 (318)	whole (256)	ncep (130)	amf (88)	cmip5 (679)	c3n4 (214)
	climatic (289)	tax (254)	fco (107)	sf5cf3 (86)	cmip3 (587)	sdg (187)
	model (288)	landscape (249)	pfc (98)	clc (81)	mofs (299)	zika (182)
	atmospheric (281)	alternative (243)	otcs (98)	embankment (81)	sdm (297)	ndcs (168)
	effect (280)	availability (242)	dtr (95)	cwd (79)	mof (275)	indc (164)
	global (224)	life (239)	nee (89)	etm (75)	biochars (252)	indcs (134)

Table: Growth in climate change literature

Data from WoS Core Collection, query following Grieneisen and Zhang (2011)

Approach - What is the matter?

- Topic modelling (Blei et al., 2012) describes a suite of algorithms to discover the latent semantic content of documents

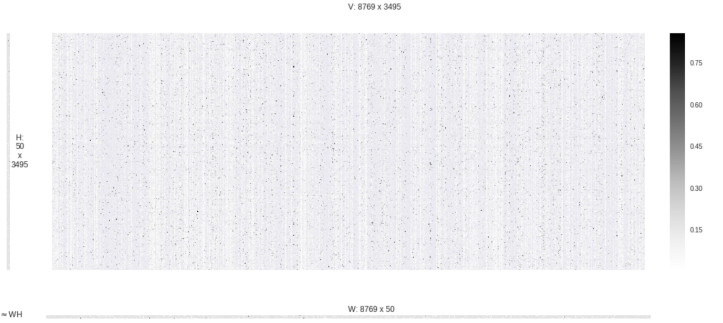
Approach - What is the matter?

- Topic modelling (Blei et al., 2012) describes a suite of algorithms to discover the latent semantic content of documents
- NMF (Lee and Seung, 1999) is a dimensionality reduction technique that can be used for topic modelling

$V_{i\mu}$ is a term frequency-inverse document frequency matrix of *stemmed* terms

$$V_{i\mu} \approx (WH)_{i\mu} = \sum_{a=1}^{r} W_{ia} H_{a\mu}$$

V is approximated by the product of W and H



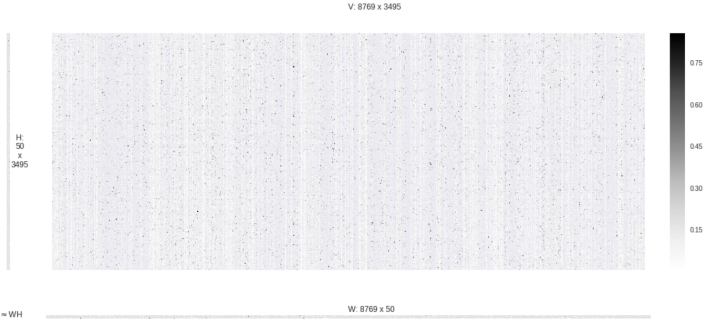
Approach - What is the matter?

- Topic modelling (Blei et al., 2012) describes a suite of algorithms to discover the latent semantic content of documents
- NMF (Lee and Seung, 1999) is a dimensionality reduction technique that can be used for topic modelling
- Topics are distributions of words. They describe documents.

$V_{i\mu}$ is a term frequency-inverse document frequency matrix of *stemmed* terms

$$V_{i\mu} \approx (WH)_{i\mu} = \sum_{a=1}^{r} W_{ia} H_{a\mu}$$

V is approximated by the product of W and H

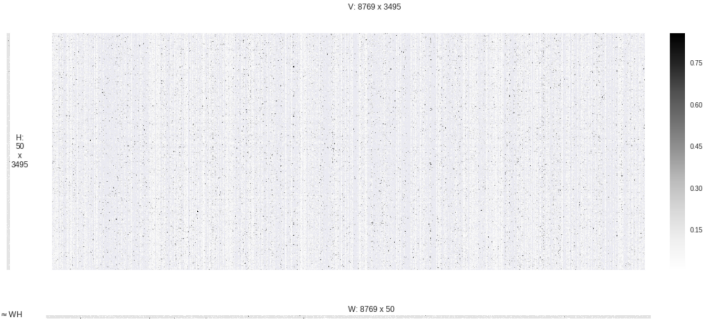


- Topic modelling (Blei et al., 2012) describes a suite of algorithms to discover the latent semantic content of documents
- NMF (Lee and Seung, 1999) is a dimensionality reduction technique that can be used for topic modelling
- Topics are distributions of words. They describe documents.
- A document's topic scores describe its association with each topic

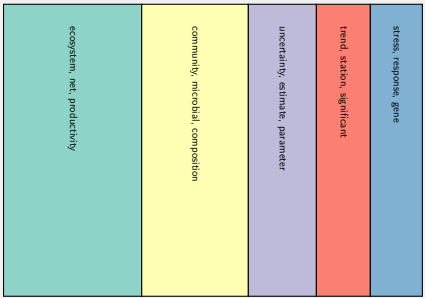
$V_{i\mu}$ is a term frequency-inverse document frequency matrix of *stemmed* terms

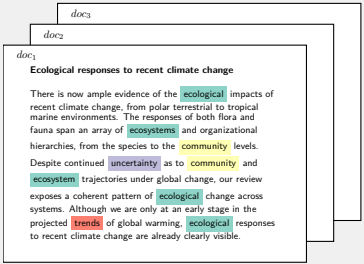
$$V_{i\mu} \approx (WH)_{i\mu} = \sum_{a=1}^{r} W_{ia} H_{a\mu}$$

V is approximated by the product of W and H



Doc Topic Example





Topic / Doc	ecosystem net productivity	community microbial composition	uncertainty estimate parameter	trend station significant
doc_1	0.022	0.017	0.011	0.009
doc_2	...	...	...	...
doc_3	...	...	...	...

Doc Topic Matrix

×

Term / Topic	ecological	ecosystem	recent	community
ecosystem net productivity	1.08	9.18	0	0
community microbial composition	0.19	0	0	9.76
uncertainty estimate parameter	0	0	0.01	0

Topic Term Matrix

≈

Term / Doc	ecological	ecosystem	recent	community
doc_1	3	2	2	2
doc_2	...	...	...	...
doc_3	...	...	...	...
doc_4	...	...	...	...

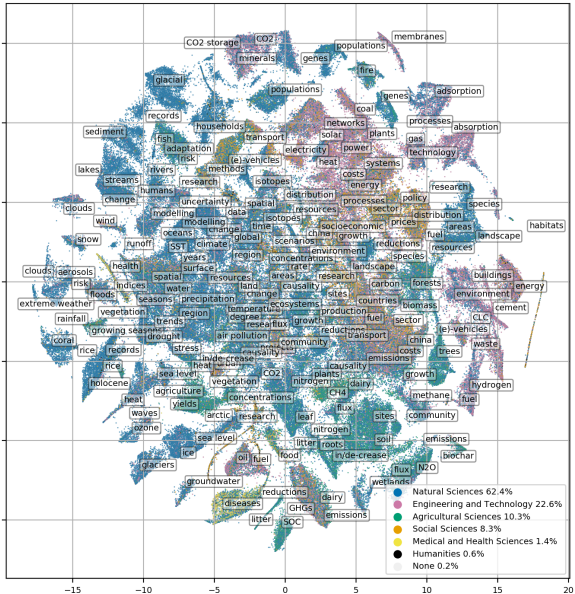
Doc Term Matrix

1 Motivation

2 Methods

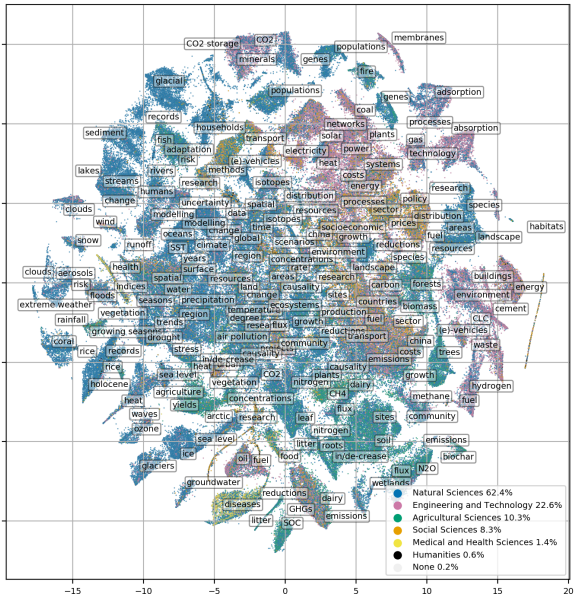
3 Results

A Topography of Climate Change Literature



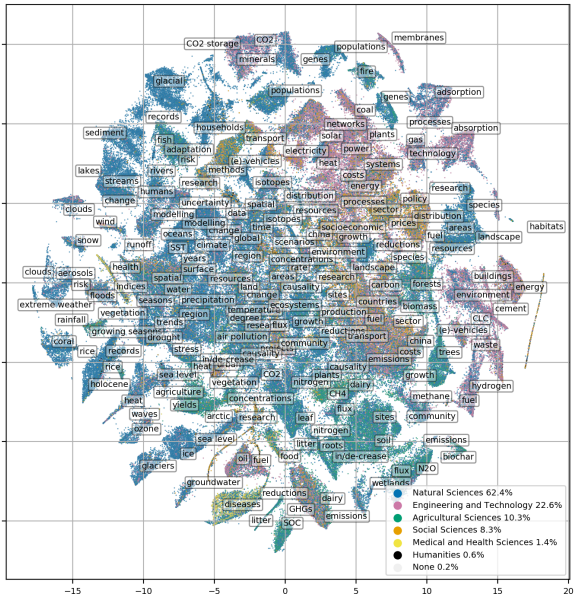
- We use t-distributed stochastic neighbour embedding (van der Maaten and Hinton, 2008) to reduce documents' topic vectors to 2 dimensions

A Topography of Climate Change Literature

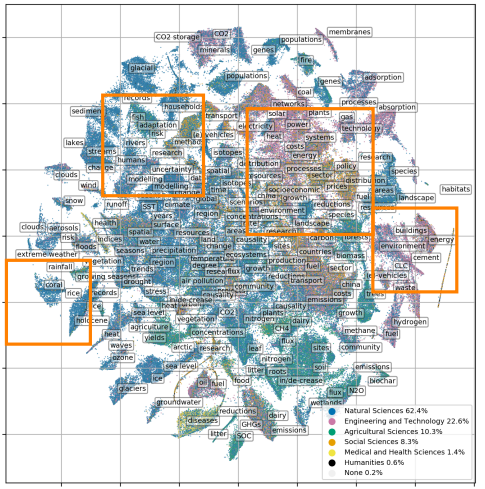


- We use t-distributed stochastic neighbour embedding (van der Maaten and Hinton, 2008) to reduce documents' topic vectors to 2 dimensions
- Each dot is a document, and documents with similar topic vectors are close together in the 2-dimensional space

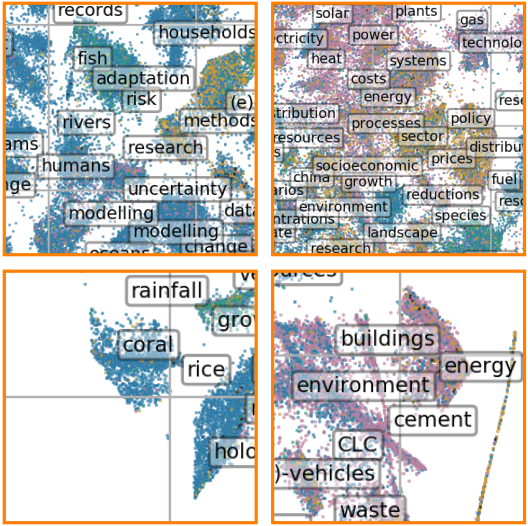
A Topography of Climate Change Literature



- We use t-distributed stochastic neighbour embedding (van der Maaten and Hinton, 2008) to reduce documents' topic vectors to 2 dimensions
- Each dot is a document, and documents with similar topic vectors are close together in the 2-dimensional space
- We can see the preponderance of natural sciences, and the greater or lesser clustering of disciplines in certain topic areas


Natural Sciences 62.4%
Engineering and Technology 22.6%
Agricultural Sciences 10.3%
Social Sciences 8.3%
Medical and Health Sciences 1.4%
Humanities 0.6%
None 0.2%

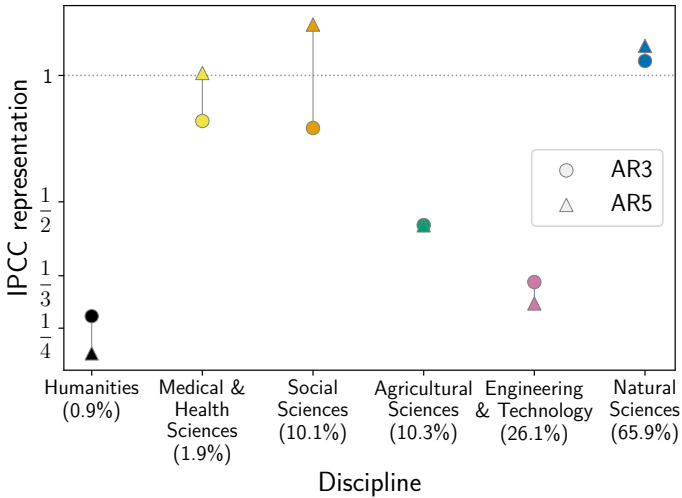
(a)


records
households
fish
adaptation
risk
rivers
methods
research
humans
uncertainty
modelling
solar
plants
gas
power
heat
systems
costs
energy
policy
processes
resources
sector
socioeconomic
prices
growth
reductions
environment
species
landscape
rainfall
coral
rice
buildings
energy
environment
cement
CLC
waste

(b)

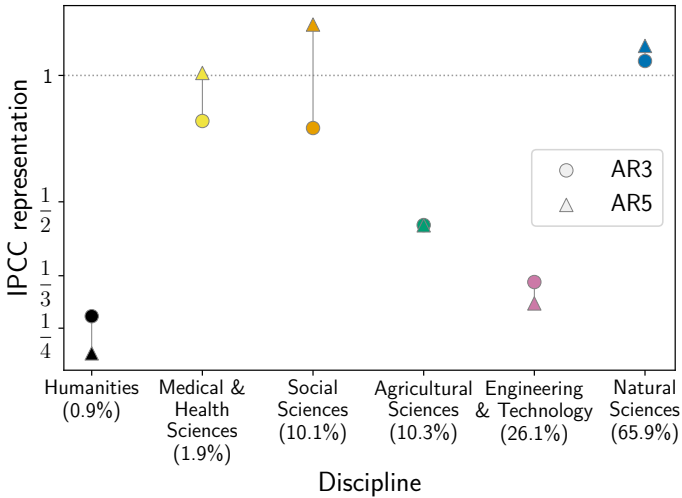
The Social Sciences are No Longer Under-Represented in IPCC reports

- The social sciences were under-represented in AR3, but by AR5 are over-represented



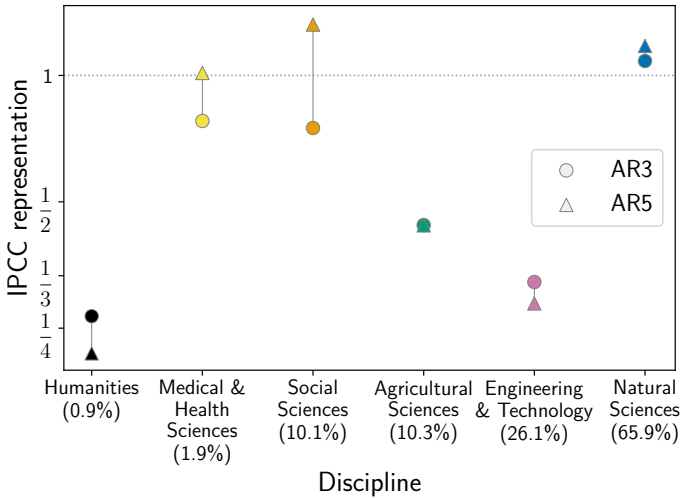
The Social Sciences are No Longer Under-Represented in IPCC reports

- The social sciences were under-represented in AR3, but by AR5 are over-represented
- The share of natural science documents in IPCC citations is similar to the share in the wider literature

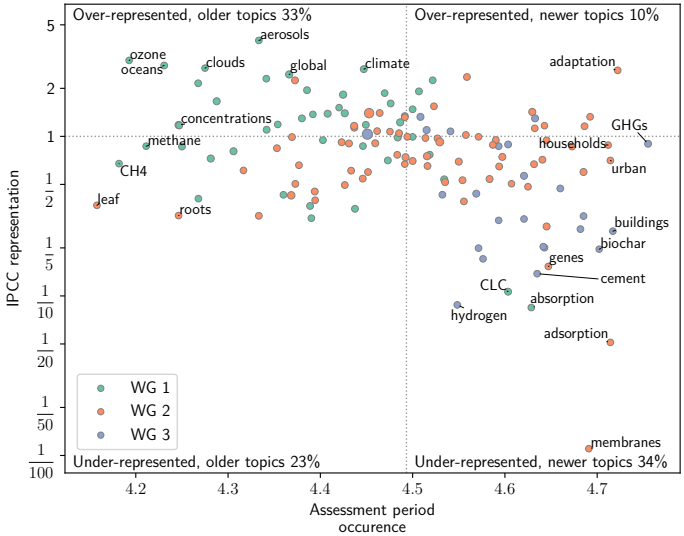


The Social Sciences are No Longer Under-Represented in IPCC reports

- The social sciences were under-represented in AR3, but by AR5 are over-represented
- The share of natural science documents in IPCC citations is similar to the share in the wider literature
- Agricultural sciences and engineering & technology are under-represented

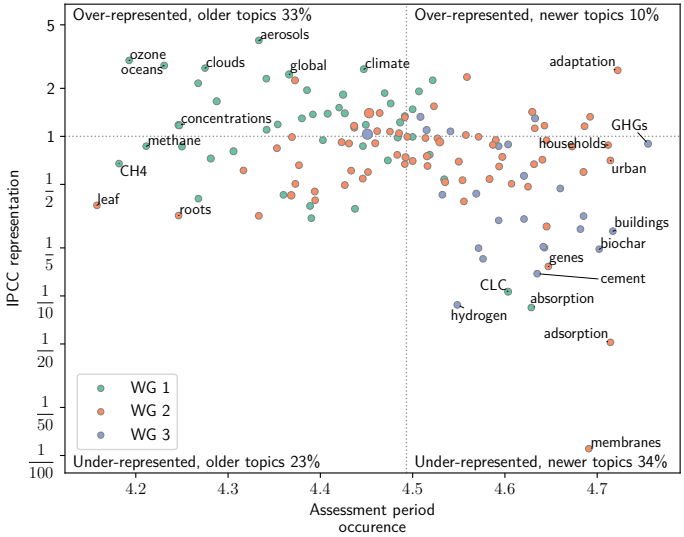


Topics on solutions are newer and under-represented



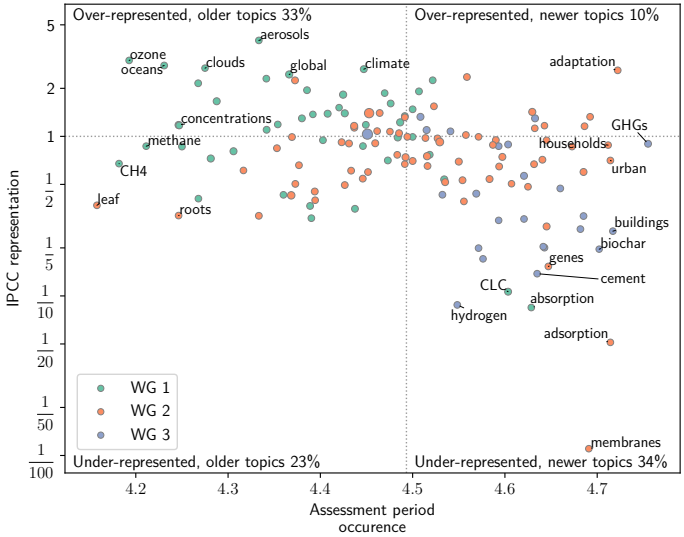
- The physical science of climate change is older and better covered

Topics on solutions are newer and under-represented



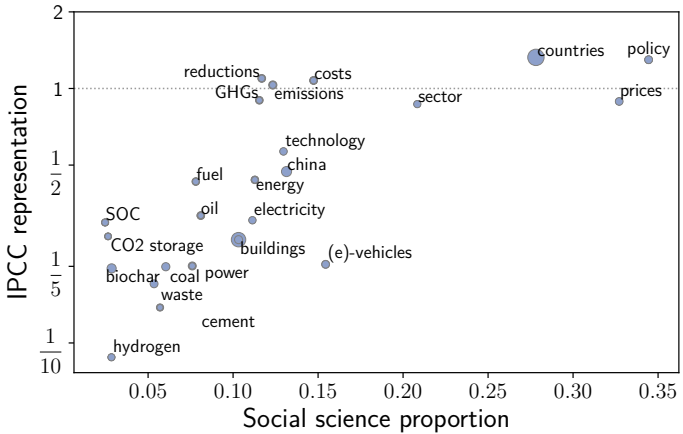
- The physical science of climate change is older and better covered
- Topics on "solutions" (although rather technical than policy) are newer and under-represented

Topics on solutions are newer and under-represented



- The physical science of climate change is older and better covered
- Topics on "solutions" (although rather technical than policy) are newer and under-represented
- Newer WGII topics are better covered than newer WGIII topics

WGIII topics with little social science are under-represented



- Technical solutions topics in WGIII contain little social science research and are under-represented

Conclusions

- Lack of social science knowledge in IPCC $\neq$ IPCC natural science bias

Conclusions

- Lack of social science knowledge in IPCC $\neq$ IPCC natural science bias
- If there is a lack, we need to produce more social science knowledge, particularly on technical solutions

Conclusions

- Lack of social science knowledge in IPCC $\neq$ IPCC natural science bias
- If there is a lack, we need to produce more social science knowledge, particularly on technical solutions
- There is a vast technical literature on solutions which is relatively untapped

Conclusions

- Lack of social science knowledge in IPCC $\neq$ IPCC natural science bias
- If there is a lack, we need to produce more social science knowledge, particularly on technical solutions
- There is a vast technical literature on solutions which is relatively untapped

But,

- Literature here is not all relevant climate knowledge (only WoS, only directly on climate change)

Conclusions

- Lack of social science knowledge in IPCC $\neq$ IPCC natural science bias
- If there is a lack, we need to produce more social science knowledge, particularly on technical solutions
- There is a vast technical literature on solutions which is relatively untapped

But,

- Literature here is not all relevant climate knowledge (only WoS, only directly on climate change)
- Relative representation of different parts of the literature is job of IPCC to decide, not a computer

Conclusions

- Lack of social science knowledge in IPCC $\neq$ IPCC natural science bias
- If there is a lack, we need to produce more social science knowledge, particularly on technical solutions
- There is a vast technical literature on solutions which is relatively untapped

But,

- Literature here is not all relevant climate knowledge (only WoS, only directly on climate change)
- Relative representation of different parts of the literature is job of IPCC to decide, not a computer

Nevertheless,

- Computer assisted methods can help the IPCC make its decisions on how to represent the literature from a more solid basis, and efficiently point to areas of growth or of under-representation

Conclusions

- Lack of social science knowledge in IPCC $\neq$ IPCC natural science bias
- If there is a lack, we need to produce more social science knowledge, particularly on technical solutions
- There is a vast technical literature on solutions which is relatively untapped

---

A Topography of Climate Change Research

`https://dx.doi.org/10.1038/s41558-019-0684-5`

`callaghan@mcc-berlin.net`

But,

- Literature here is not all relevant climate knowledge (only WoS, only directly on climate change)
- Relative representation of different parts of the literature is job of IPCC to decide, not a computer

Nevertheless,

- Computer assisted methods can help the IPCC make its decisions on how to represent the literature from a more solid basis, and efficiently point to areas of growth or of under-representation

Bibliography

Bjurström, A. and Polk, M. (2011). Physical and economic bias in climate change research: A scientometric study of IPCC Third Assessment Report. *Climatic Change*, 108(1):1–22.

Blei, D., Carin, L., and Dunson, D. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4):77–84.

David G. Victor (2015). Embed the social sciences in climate policy - David Victor. *Nature*, 520:7–9.

Edenhofer, O. and Kowarsch, M. (2015). Cartography of pathways: A new model for environmental policy assessments. *Environmental Science and Policy*, 51:56–64.

Edenhofer, O. and Minx, J. (2014). Mapmakers and navigators, facts and values. *Science*, 345(6192):37–38.

Grieneisen, M. and Zhang, M. (2011). The Current Status of Climate Change Research. *Nature Climate Change*, 1:72–73.

Hulme, M. and Mahony, M. (2010). Climate change: What do we know about the IPCC? *Progress in Physical Geography*, 34(5):705–718.

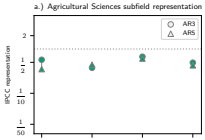
Kowarsch, M., Jabbour, J., Flachsland, C., Kok, M. T. J., Watson, R., Haas, P. M., Minx, J. C., Alcamo, J., Garard, J., Riousset, P., Pintér, L., Langford, C., Yamineva, Y., von Stechow, C., O'Reilly, J., and Edenhofer, O. (2017). A road map for global environmental assessments. *Nature Climate Change*, 7(6):379–382.

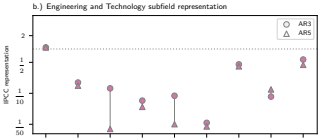
Lee, D. D. and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–91.

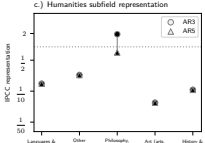
Minx, J. C., Callaghan, M., Lamb, W. F., Garard, J., and Edenhofer, O. (2017). Learning about climate change solutions in the IPCC and beyond. *Environmental Science & Policy*.

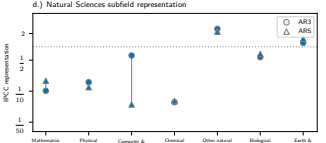
van der Maaten, L. and Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605.

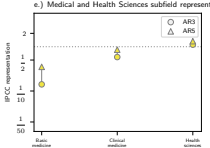
Topic disciplinary entropy and subdiscipline representation

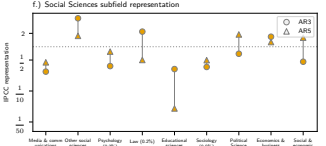




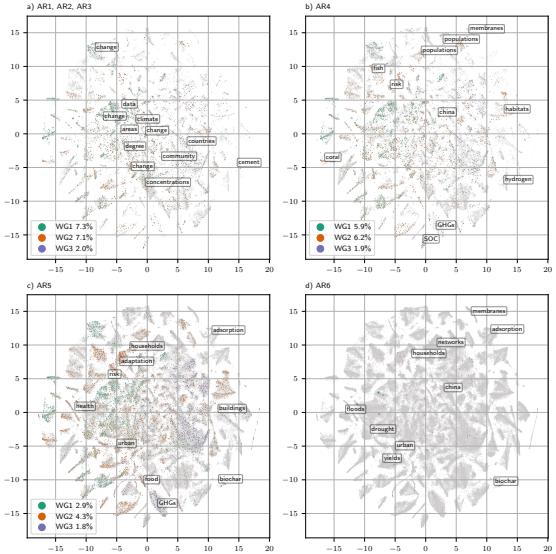




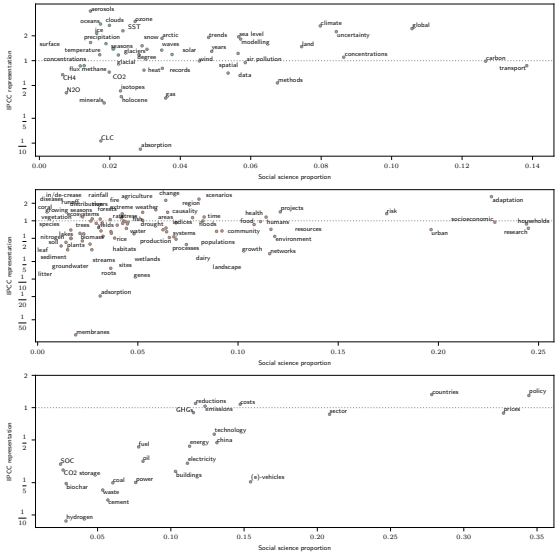




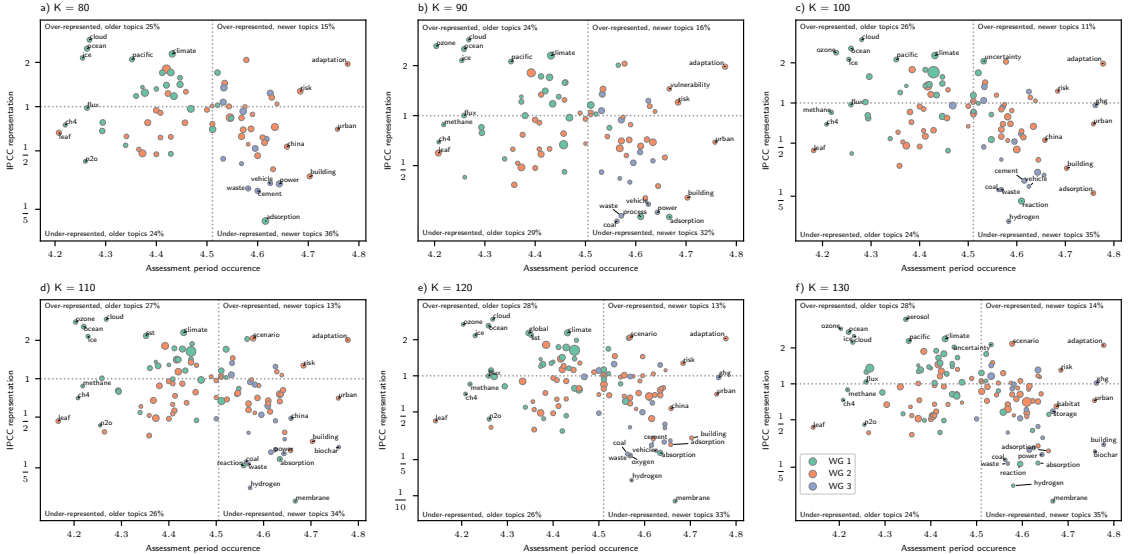
Topic growth



Social sciences and topic representation



n Topics


a) K = 80
Over-represented, older topics 25%
Over-represented, newer topics 15%
Under-represented, older topics 24%
Under-represented, newer topics 36%
b) K = 90
Over-represented, older topics 24%
Over-represented, newer topics 16%
Under-represented, older topics 29%
Under-represented, newer topics 32%
c) K = 100
Over-represented, older topics 26%
Over-represented, newer topics 11%
Under-represented, older topics 24%
Under-represented, newer topics 35%
d) K = 110
Over-represented, older topics 27%
Over-represented, newer topics 13%
Under-represented, older topics 26%
Under-represented, newer topics 34%
e) K = 120
Over-represented, older topics 28%
Over-represented, newer topics 13%
Under-represented, older topics 26%
Under-represented, newer topics 33%
f) K = 130
Over-represented, older topics 28%
Over-represented, newer topics 14%
Under-represented, older topics 24%
Under-represented, newer topics 35%
WG 1
WG 2
WG 3
IPCC representation
Assessment period occurence