Supplementary Materials

Section I: Identification of topics through keyword clustering and topic titles

Three manual inputs are required to perform the automatic classification of the literature in topics. The first is the decision upon the number of topics. The number depends on the aim of the article and on the type of research question to be addressed. As our intent is to place and distinguish the literature in large – yet coherent – topics, we decided to reduce the number of topics to 20 (*T=20*), a number that suffices to distinguish branches of scientific literature, with topics that characterize DRR, and others that are typically associated to the CCA research stream. The second is the tuning of the classification tool. We carried out the topic modelling adopting the Gibbs algorithm following a rule of thumb proposed by Steyvers and Griffiths ((2007)) to obtain enough granularity of results and clear proportion of topics. Thus, we set the variable of topic smoothing equal to 50/*T* keeping the term smoothing equal to 0.01. One last manual setup before the run of topic modelling was the exclusion of the 20 most frequent words[1] and of irrelevant words, such as articles, prepositions and verbs like to show, exhibit, prove, and others, that are too pervasive to characterize topics, resulting in additional noise to the classification task. The classification task of articles into topic proportion requires also a minimum amount of words available that we set to 4. This procedure let to the exclusion of 4 articles.

Words (in decreasing order of probability)					
Topic 00	***Topic 01***	***Topic 02***	***Topic 03***	***Topic 04***	***Topic 05***
food	drought	model	events	economic	data
variability	resources	scenarios	extreme	future	analysis
agriculture	river	future	weather	sector	spatial
security	changes	impact	during	tourism	index
agricultural	region	models	population	energy	area
farmers	basin	under	over	potential	indicators
strategies	system	global	years	within	using
systems	resource	regional	past	sectors	method
africa	regions	using	time	increasing	information
production	western	potential	frequency	can	gis
land	fisheries	uncertainty	losses	current	china
south	land	scenario	increase	industry	hazard
rainfall	availability	projections	extremes	effects	exposure
climatic	conditions	changes	hurricane	importance	areas
changes	hydrological	current	intensity	take	each
livestock	potential	modelling	last	strategies	maps
practices	north	different	number	strategy	geographic

	Topic Titles
Topic 00	food security
Topic 01	drought & water resources
Topic 02	models & scenarios
Topic 03	extreme events
Topic 04	social (& economic) issues
Topic 05	data analysis (indexes, indicators, maps)
Topic 06	society & people (rural)
Topic 07	health & people
Topic 08	coastal vulnerability & sea level rise
Topic 09	policy & governance
Topic 10	ecology (forestry & biodiversity)
Topic 11	methods & assessment
Topic 12	national & global studies (mitigation)
Topic 13	resilience of socio-ecosystems
Topic 14	local communities & knowledge
Topic 15	climate science
Topic 16	planning (urban & local)
Topic 17	floods

[1] Such an operation is usual in machine learning approaches in computational linguistics when the idea is to discover the latent semantics and not articles' stylistic nuances.

farm	droughts	data	tropical	regional	different
farming	variability	modeling	increasing	direct	mapping
important	four	projected	exposure	supply	applied

Topic 06	***Topic 07***	***Topic 08***	***Topic 09***	***Topic 10***	***Topic 11***
social	health	coastal	policy	species	framework
rural	human	level	research	forest	approach
strategies	heat	rise	challenges	conservation	assessments
people	population	sea	governance	ecosystems	can
livelihoods	mortality	areas	sustainable	forests	integrated
migration	effects	sea-level	institutional	biodiversity	system
communities	populations	land	key	ecological	decision
poverty	stress	island	review	marine	analysis
factors	between	vulnerable	literature	warming	different
coping	vulnerable	storm	between	vegetation	approaches
households	exposure	erosion	action	ecosystem	identify
livelihood	factors	zone	issues	growth	process
vulnerable	associated	flooding	need	climatic	making
poor	disease	coast	studies	range	provide
environmental	evidence	pacific	institutions	tree	methods
household	including	islands	barriers	distribution	systems
access	risks	including	processes	habitat	application
economic	among	inundation	practice	populations	decision-making
groups	diseases	small	focus	diversity	provides
capacities	increased	along	opportunities	fire	tool

Topic 12	***Topic 13***	***Topic 14***	***Topic 15***	***Topic 16***	***Topic 17***
countries	resilience	community	temperature	planning	flood
mitigation	social	communities	increase	urban	measures
national	environmental	knowledge	crop	local	floods
developing	systems	local	precipitation	information	services
policies	human	research	conditions	risks	analysis
global	sustainability	perceptions	production	infrastructure	risks
efforts	political	through	changes	areas	potential
reduce	global	findings	climatic	can	should
world	social-ecological	perception	soil	cities	flooding
states	climate-change	canada	higher	city	case
emissions	dimensions	understanding	regions	government	costs
projects	between	case	growth	strategies	impact
effective	concept	participatory	yield	vulnerabilities	ecosystem
country	science	environmental	degrees	design	european
united	ecological	arctic	increased	plans	important
level	perspective	indigenous	rainfall	environment	protection
greenhouse	processes	conducted	crops	process	economic
there	complex	interviews	trends	through	part

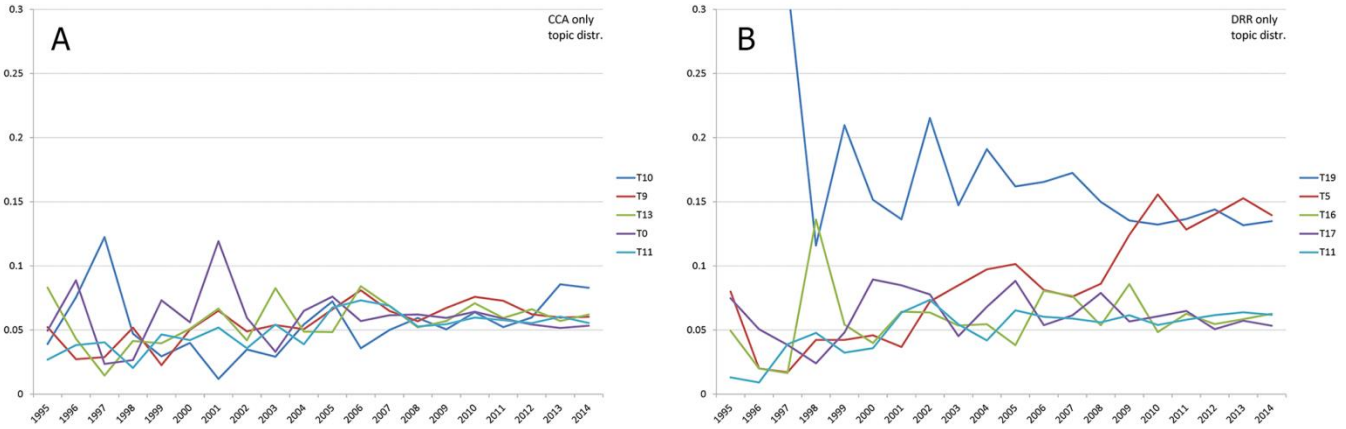
Topic 18	adaptation responses
Topic 19	hazards & disasters

need	within	changes	increases	effective	benefits
address	through	northern	carbon	existing	reduce

Topic 18	***Topic 19***
response	natural
responses	hazards
our	disasters
adapt	hazard
can	reduction
however	damage
important	building
australia	earthquake
there	mitigation
ability	preparedness
across	emergency
factors	loss
those	seismic
where	response
place	buildings
individual	prevention
nature	recovery
while	physical
limited	structural
likely	earthquakes

Section II. Evolution of top-5 topics in the CCA and DRR streams.

The dynamics of the average topic-proportion of articles for each community (CCA and DRR literatures in graph A and B, respectively) for the five topics with highest absolute topic-proportion per each community, with the exclusion of the information coming from those articles that appear in both dataset.

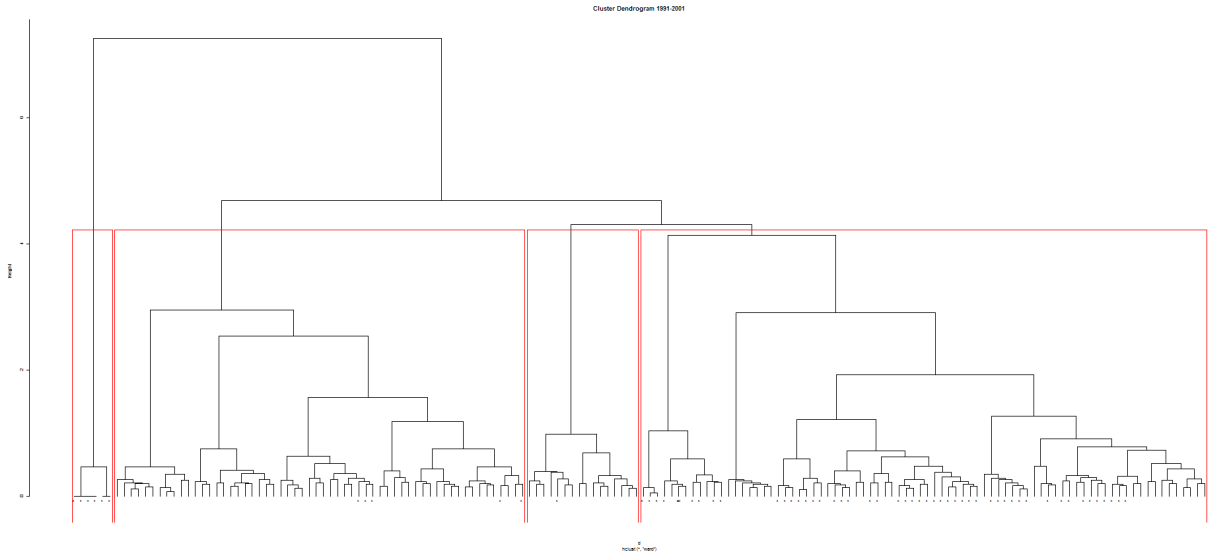


Section III. Clusters of articles based on their content: dendrograms and centroids

Having transformed each article in a vector of document topic distribution with values between 0 and 1, it is possible to generate a distance matrix and compute a classic clustering analysis. We first calculated the Euclidean distance for the articles in each time slice and used the aggregating Ward method (Murtagh and Contreras, 2012). Then we plot and compute centroids to characterize the different clusters. In the following figures and tables we show and analyze the evolution of the literature. Dendrograms were plotted to show a synthetic representation of the results.

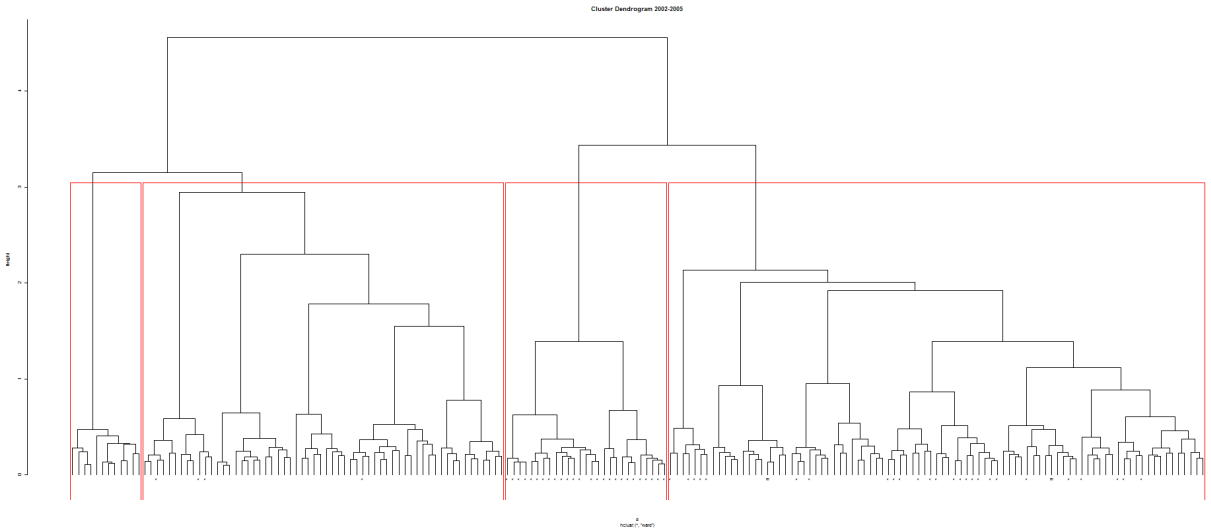
Identification of clusters through topic weighs of cluster centroids in four time slices (clusters are in the same visual order)
In dendrograms (*) display DRR articles, and (!!) highlight CCA&DRR articles.

	topic weight				
TS1 (1991-2001)	*cl. 1 (far left)*	*cl. 2*	*cl. 3*	*cl.4 (far right)*	***topic titles***
T.0	0.00	0.08	0.05	0.02	food security
T.1	0.00	0.09	0.04	0.02	drought & water resources
T.2	0.00	0.11	0.04	0.04	models & scenarios
T.3	0.00	0.05	0.07	0.06	extreme events
T.4	0.00	0.04	0.03	0.04	social (&economic) issues
T.5	0.00	0.03	0.02	0.03	data analysis (indexes, indicators, maps)
T.6	0.04	0.02	0.03	0.04	society & people (rural)
T.7	0.00	0.02	0.02	0.08	health & people
T.8	0.00	0.02	0.37	0.03	coastal vuln & sea level rise
T.9	0.00	0.02	0.03	0.06	policy & govenrnace
T.10	0.00	0.09	0.02	0.02	ecology (& forestry)
T.11	0.00	0.03	0.03	0.05	methods & assessment
T.12	0.00	0.04	0.04	0.10	countries and the world
T.13	0.00	0.03	0.02	0.07	resilence
T.14	0.08	0.02	0.01	0.03	(local) communities
T.15	0.00	0.15	0.02	0.02	climate science
T.16	0.00	0.03	0.05	0.05	planning (urban&local)
T.17	0.00	0.04	0.07	0.07	floods
T.18	0.00	0.05	0.02	0.05	adaptation responses
T.19	0.88	0.02	0.03	0.12	hazards & disasters
papers	6	58	16	80	



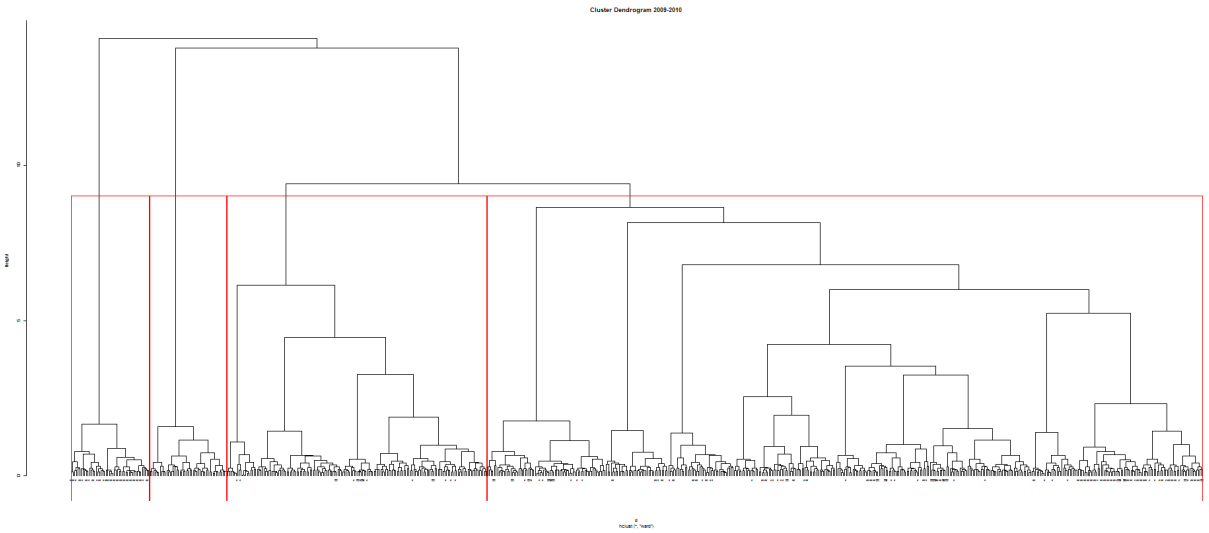
TS2 (2002-05)

T.0	0.04	0.09	0.01	0.03	food security
T.1	0.03	0.07	0.03	0.04	drought & water resources
T.2	0.05	0.12	0.04	0.03	models & scenarios
T.3	0.05	0.04	0.05	0.06	extreme events
T.4	0.04	0.04	0.03	0.05	social (&economic) issues
T.5	0.03	0.04	0.11	0.04	data analysis (indexes, indicators, maps)
T.6	0.02	0.03	0.03	0.06	society & people (rural)
T.7	0.02	0.03	0.04	0.05	health & people
T.8	0.02	0.10	0.03	0.03	coastal vuln & sea level rise
T.9	0.02	0.03	0.03	0.07	policy & govenrnace
T.10	0.34	0.03	0.01	0.02	ecology (& forestry)
T.11	0.03	0.04	0.06	0.07	methods & assessment
T.12	0.03	0.04	0.06	0.08	countries and the world
T.13	0.01	0.02	0.04	0.08	resilence
T.14	0.04	0.02	0.03	0.05	(local) communities
T.15	0.08	0.10	0.01	0.02	climate science
T.16	0.02	0.03	0.04	0.05	planning (urban&local)
T.17	0.02	0.05	0.06	0.06	floods
T.18	0.07	0.05	0.04	0.07	adaptation responses
T.19	0.02	0.02	0.26	0.05	hazards & disasters
papers	12	60	27	89	



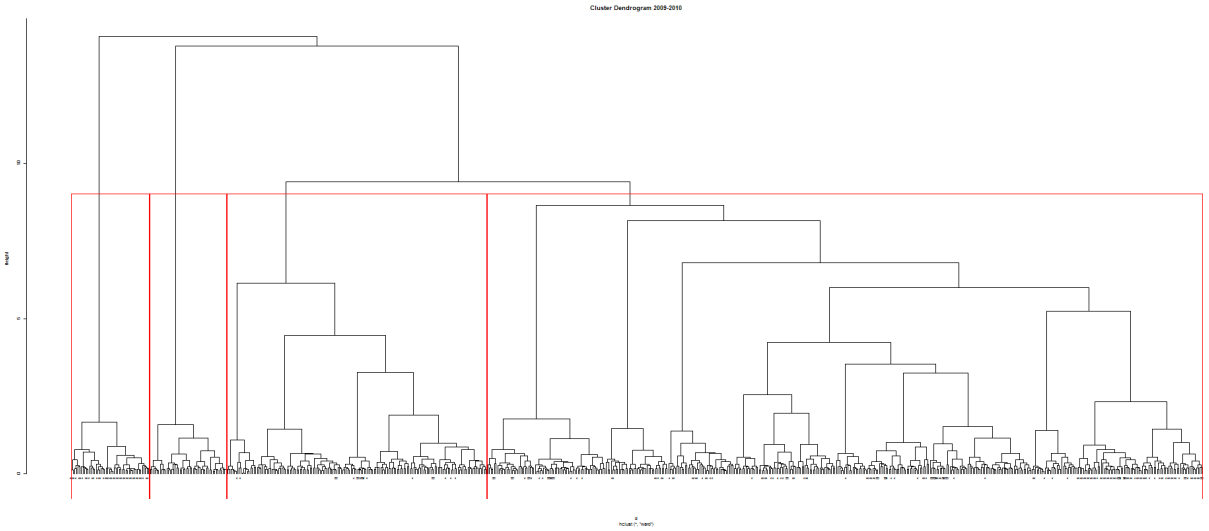
TS4 (2009-10)

T.0	0.01	0.02	0.13	0.03	food security
T.1	0.05	0.06	0.08	0.03	drought & water resources
T.2	0.04	0.05	0.04	0.04	models & scenarios
T.3	0.04	0.05	0.04	0.04	extreme events
T.4	0.02	0.03	0.04	0.05	social (&economic) issues
T.5	0.38	0.03	0.03	0.04	data analysis (indexes, indicators, maps)
T.6	0.02	0.01	0.10	0.04	society & people (rural)
T.7	0.01	0.03	0.03	0.05	health & people
T.8	0.04	0.03	0.02	0.05	coastal vuln & sea level rise
T.9	0.02	0.02	0.05	0.08	policy & govenrnace
T.10	0.02	0.35	0.03	0.02	ecology (& forestry)
T.11	0.06	0.02	0.03	0.07	methods & assessment
T.12	0.02	0.02	0.05	0.06	countries and the world
T.13	0.03	0.02	0.04	0.08	resilence
T.14	0.02	0.02	0.04	0.07	(local) communities
T.15	0.02	0.11	0.09	0.02	climate science
T.16	0.04	0.02	0.03	0.07	planning (urban&local)
T.17	0.05	0.02	0.04	0.05	floods
T.18	0.02	0.05	0.05	0.05	adaptation responses
T.19	0.09	0.01	0.02	0.07	hazards & disasters
papers	48	47	158	435	



TS6 (2013-14)

T.0	0.02	0.01	0.12	0.03	food security
T.1	0.05	0.03	0.09	0.03	drought & water resources
T.2	0.05	0.05	0.05	0.05	models & scenarios
T.3	0.04	0.05	0.04	0.05	extreme events
T.4	0.03	0.03	0.03	0.06	social (&economic) issues
T.5	0.04	0.21	0.05	0.04	data analysis (indexes, indicators, maps)
T.6	0.02	0.03	0.12	0.04	society & people (rural)
T.7	0.03	0.02	0.03	0.05	health & people
T.8	0.02	0.17	0.03	0.03	coastal vuln & sea level rise
T.9	0.03	0.02	0.04	0.08	policy & govenrnace
T.10	0.35	0.02	0.02	0.02	ecology (& forestry)
T.11	0.04	0.07	0.04	0.07	methods & assessment
T.12	0.03	0.02	0.03	0.05	countries and the world
T.13	0.03	0.04	0.03	0.08	resilence
T.14	0.03	0.03	0.04	0.07	(local) communities
T.15	0.08	0.03	0.10	0.03	climate science
T.16	0.02	0.04	0.03	0.06	planning (urban&local)
T.17	0.03	0.05	0.04	0.05	floods
T.18	0.05	0.03	0.05	0.06	adaptation responses
T.19	0.01	0.05	0.02	0.05	hazards & disasters
papers	169	120	230	620	



The cluster analysis segmented into six time slices aimed at understanding whether the topics could characterized the evolution of the literature and allow for the identification of research sub-streams characterised by peculiar topic distributions. In order to have comparable results, we focused at the hierarchical level of the classification tree at which the set of paper is subdivided into four branches (clusters). The analysis of the first time period (1991-2001) shows two well identified groups of papers: the first is characterised by the prevalence of T19 (hazards and disasters), with 6 papers only, belonging to the DRR set. The second cluster is also clearly identified by T2 and T15, identifying a subset of 58 papers (almost entirely belonging to CCA) focused on climatic science, scenarios and models. The third group of 16 papers (almost all CCA) is characterised by T8 and also T3 and T17, thus identifying papers dealing with costal vulnerability, sea level rise, floods and other extreme events. The remaining 80 papers (mainly belonging to DRR) are rather varied in topics. The following time period (2002-05) shows again a cluster of 27 DRR papers characterised by T19, but with

T5 (indices, indicators and maps) as a second most important topic. Again, a second cluster of almost only CCA papers (60 in total) is identifiable, while a new small cluster of 12 CCA papers appears, focused on ecology (T10) and climate change adaptation (T15 and T18). Later on, the DRR cluster with focus on disaster continues its evolution, with T19 still relevant, but with T5 (indices, indicators and maps) as the most relevant. Similarly, an evolving cluster with almost only CCA with focus on ecology can be identified as one of the four in the following periods.

References

Murtagh, F., Contreras, P. (2012) Algorithms for hierarchical clustering: an overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 2, 86-97.
Steyvers, M., Griffiths, T., (2007) Probabilistic topic models, in: Landauer, T.K., McNamara, D.S., Dennis, S., Kintsch, W. (Eds.), Handbook of latent semantic analysis. Routledge, New York, NY, pp. 427-448.