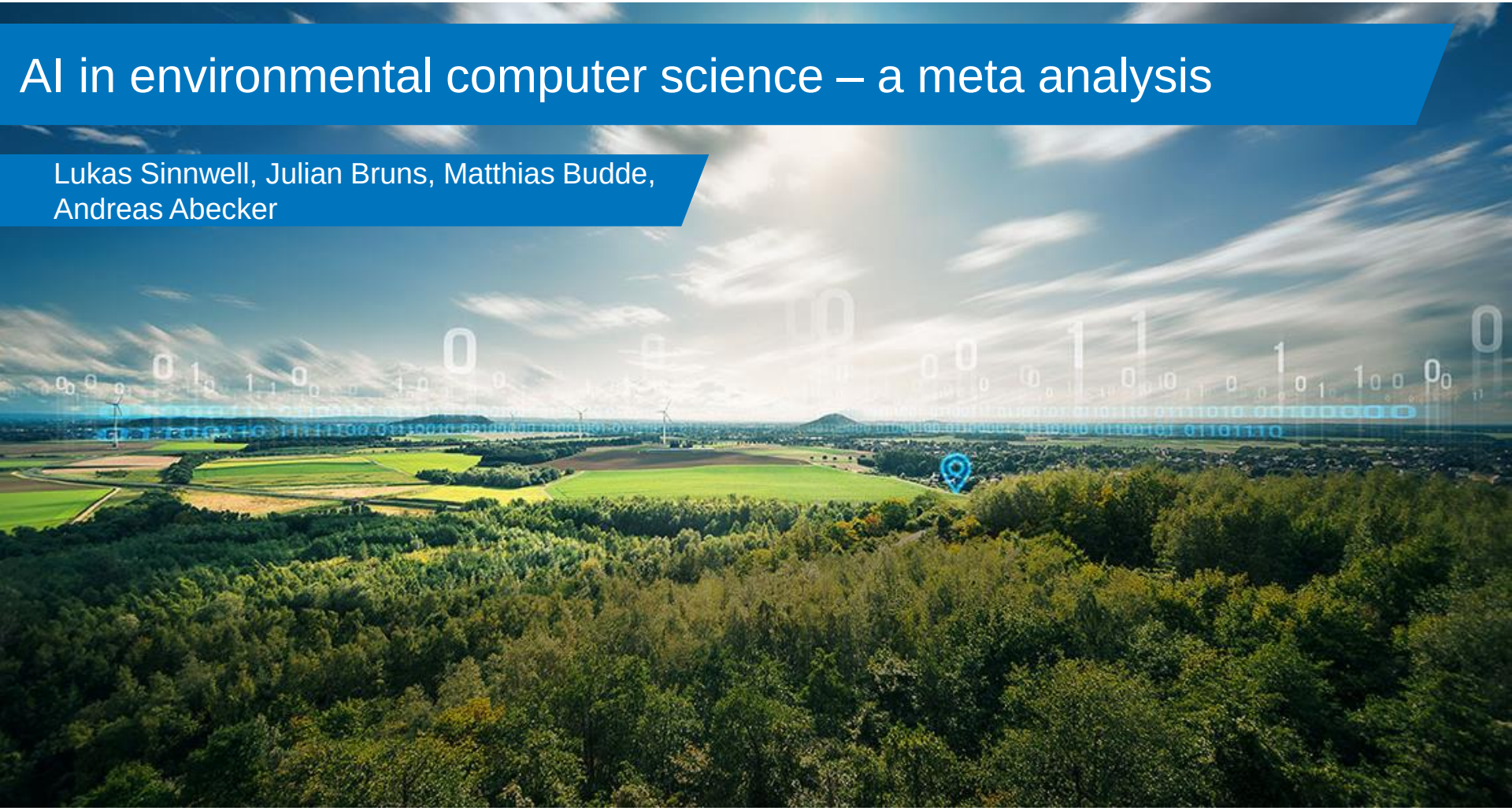
disy

Geo Intelligence for your Data

AI in environmental computer science – a meta analysis

Lukas Sinnwell, Julian Bruns, Matthias Budde,
Andreas Abecker



Starting point

“There is [. . .] no publication that allows a comprehensive interpretation of the research activities in the context of environment and sustainability […]” (BMU 2019)

disy

Goals and Content

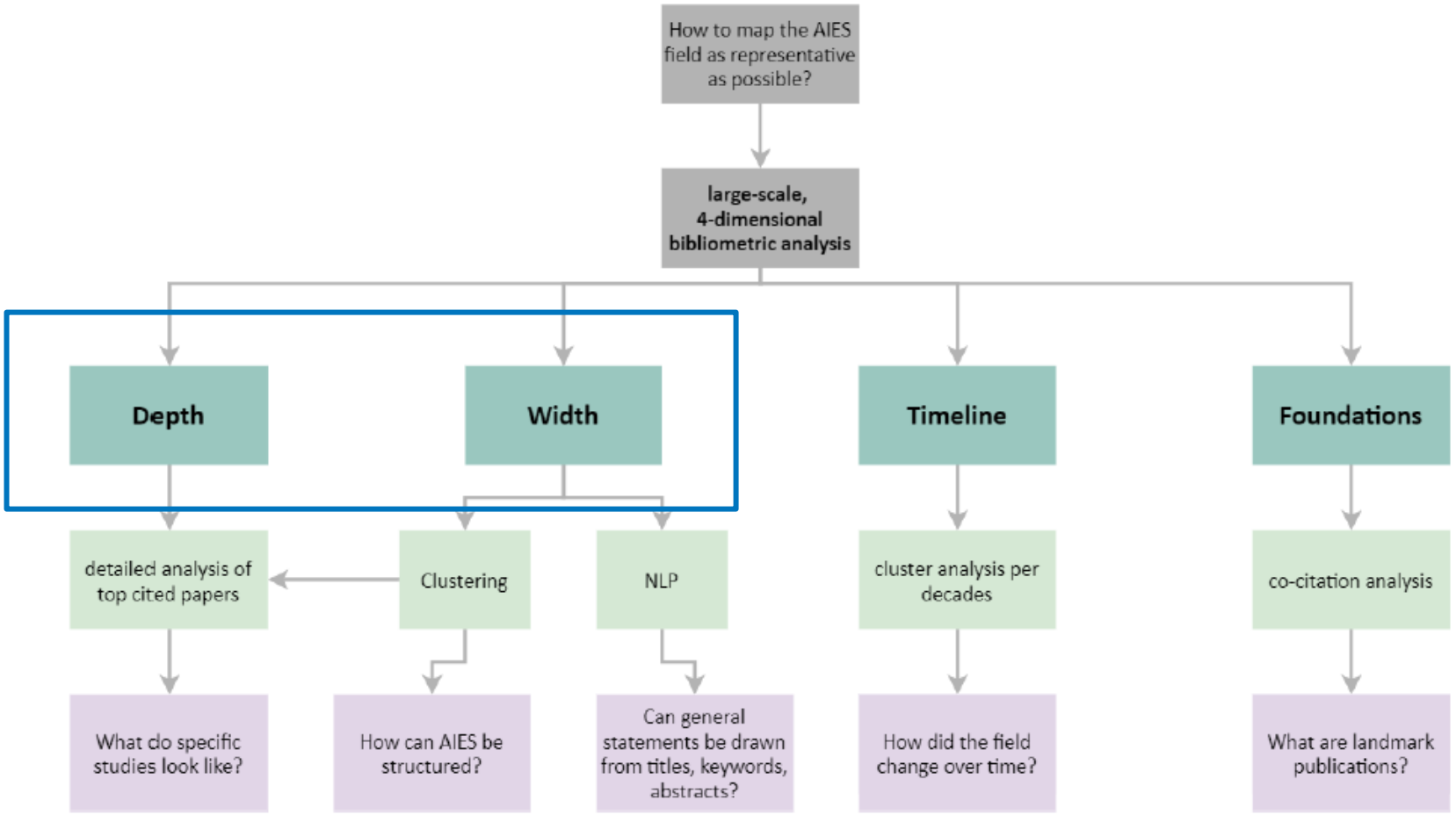
(1) Generate an overview of publications which apply to ECS and AI

(2) Detect subfields and map each publication to a subfield

(3) Provide a first evaluation of the the existing literature

(4) Discuss further steps

General approach



Criteria for in-depth comparison

- We chose the most common criteria in AI / Machine Learning
- It is based on the typical AI workflow and taken from existing literature
- We focused particulary on the ability to replicate the results of a given publication

General criteria
Research Area
Title
Authors
Times Cited
Year
Countries
Type
Journal
Conference
Keywords
AI methods in focus
Key notes
Outcome

Methodological criteria	
type of data	*Kind of data is used in study*
sample size	*size of dataset*
Input Selection	
Input Significance	*methods for deciding which inputs to use in model*
Input Independence	*methods for checking correlation among inputs*
Data Division	
Data Division	*method for dividing data into train / test set*
Training [%]	*Percentage of whole dataset used for training*
Test [%]	*Percentage of whole dataset used for testing*
Validation	*validation technique or percentage of dataset used for validation*
Model Building	
Hyperparameter tuning	*methods for setting HPs*
NN architecture	*if available, methods of determining NN architecture*
Evaluation & Reference	
Evaluation metrics	*metrics used for model evaluation*
Reference model	*reference for model developed in study*

disy

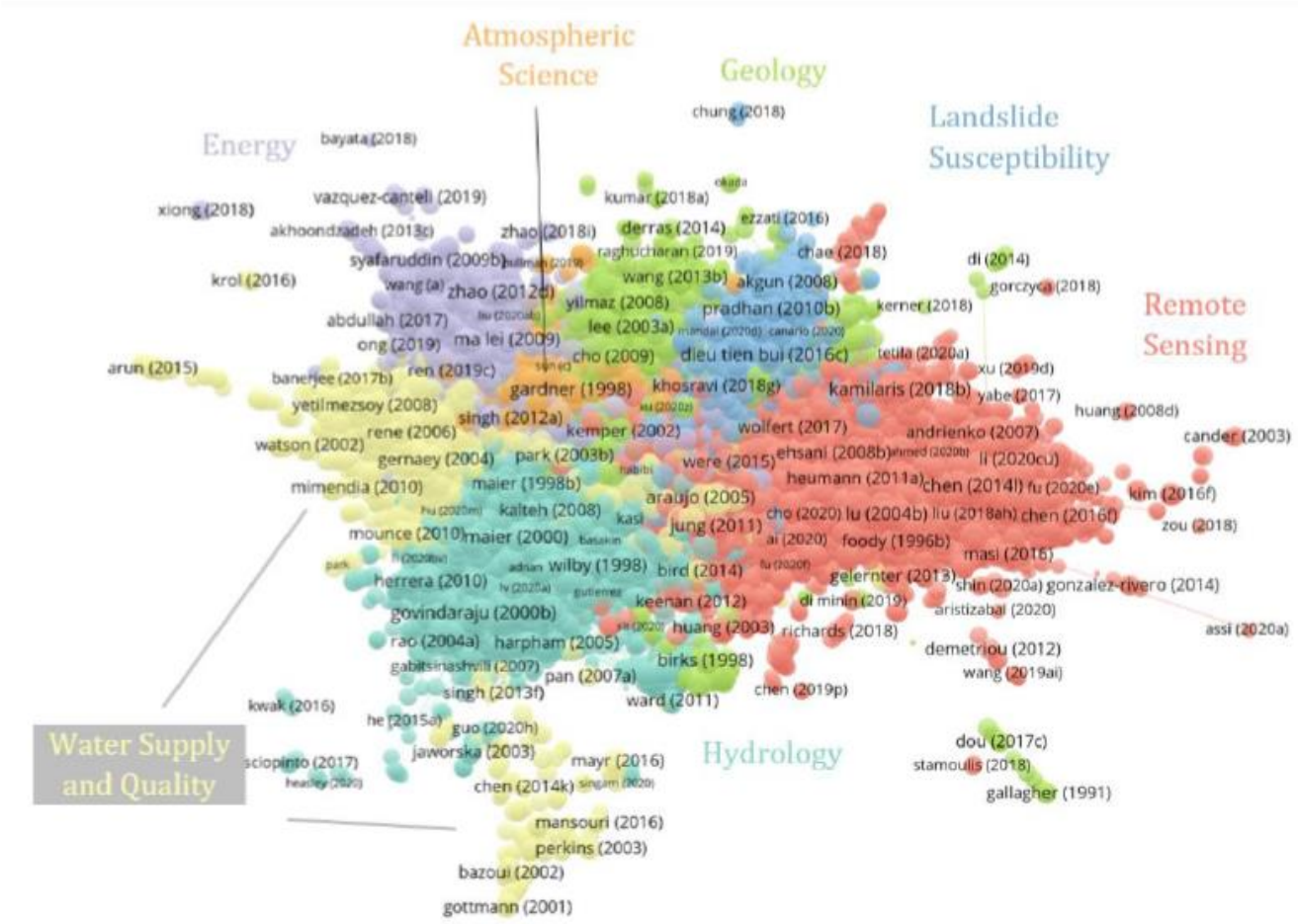
Gathering the basic corpus

```
(TS = ("Artificial Intelligence") OR TS = ("Machine Learning") OR
TS = ("Deep Learning") OR TS = ("Artificial Neural Network*"))
AND
(WC = ("ENVIRONMENTAL SCIENCES") OR WC = ("WATER RESOURCES") OR WC = ("GEO-
SCIENCES MULTIDISCIPLINARY") OR WC = ("REMOTE SENSING") OR
WC = ("GREEN SUSTAINABLE SCIENCE TECHNOLOGY") OR
WC = ("GEOGRAPHY PHYSICAL") OR WC = ("AGRICULTURE MULTIDISCIPLINARY") OR
WC = ("ENVIRONMENTAL STUDIES"))
```

- We use the Web of Science for our overall corpus
- We filter it down with the given query
  → This returns 18.254 records in November 2020
  → ~ 10% of the overall AI literature in WoS

- We apply a the Smart Local Moving Algorithm from VOSviewer to create suitable clusters for the domain and then compare selected results in-depth

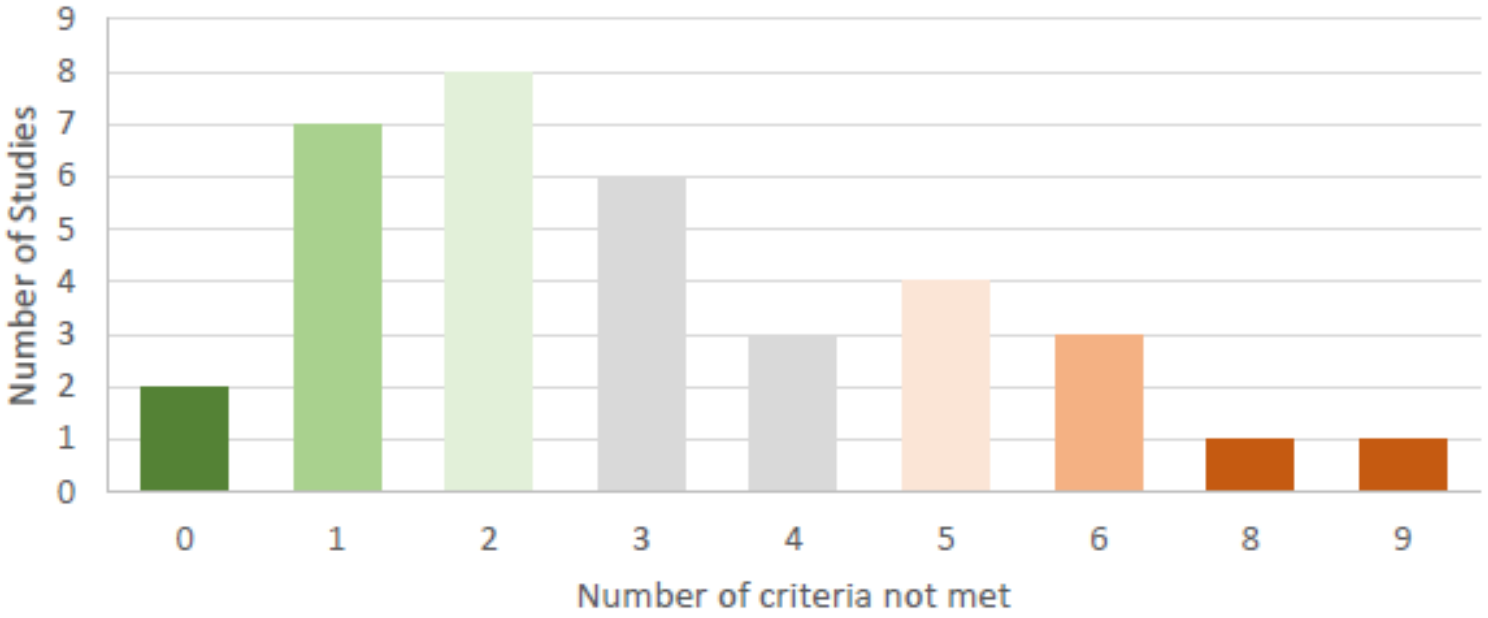
Cluster Results



Cluster Results

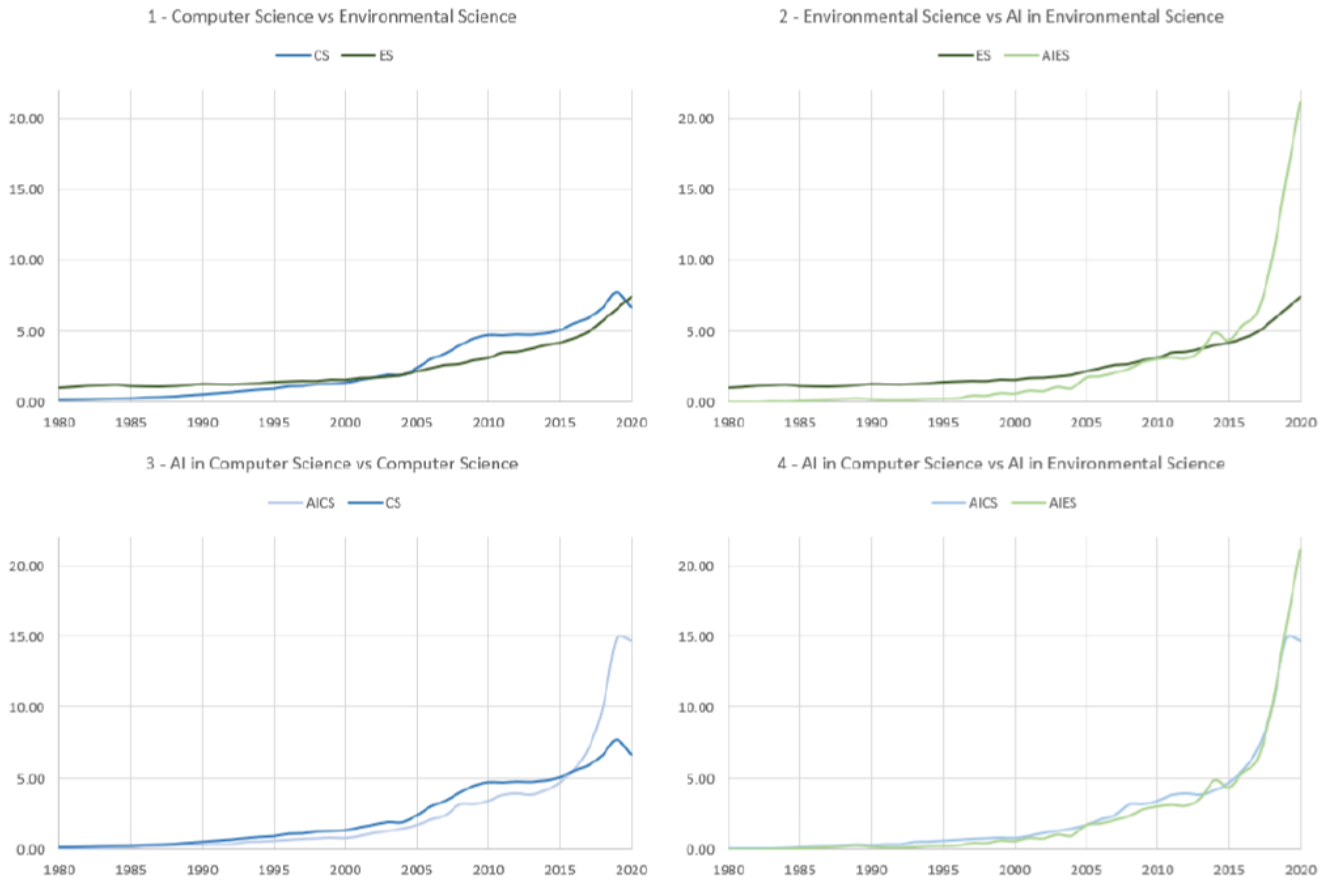
cluster 1	cluster 2	cluster 3	cluster 4	cluster 5	cluster 6	cluster 7
sentinel	river	solar radiation	adsorption	soil	emission	landslide susceptibility mapping
remote sensing	river basin	building	wastewater	rock	city	gis
forest	basin	energy	aqueous solution	blast	air quality	landslide
hyperspectral image classification	reservoir	energy consumption	treatment	earthquake	ozone concentration	landslide susceptibility assessment
remote sensing image	temperature	photovoltaic system	water	mine	ozone	landslide susceptibility
Remote Sensing	*Hydrology*	*Energy*	*Water Supply & Quality*	*Geology*	*Atmospheric Science*	*Landslide Susceptibility*

In-Depth



- Based on the 5 most cited studies in each cluster
- In total 35 publications were compared
- Most studies are older; they range from 1995 to 2016

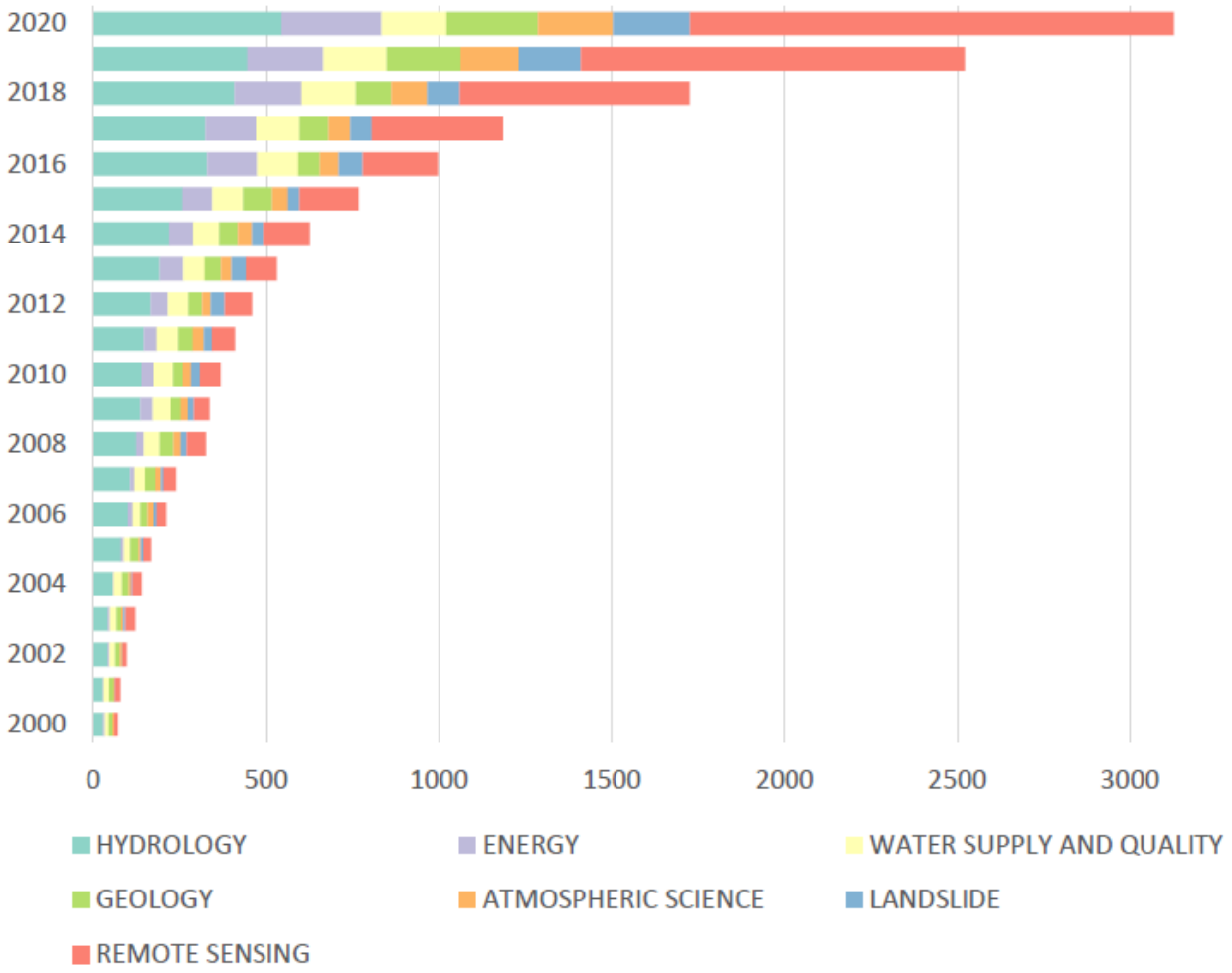
Publication history



Development of the percentage a year contributed to the whole of publications

disy

Publications per year and field in AI

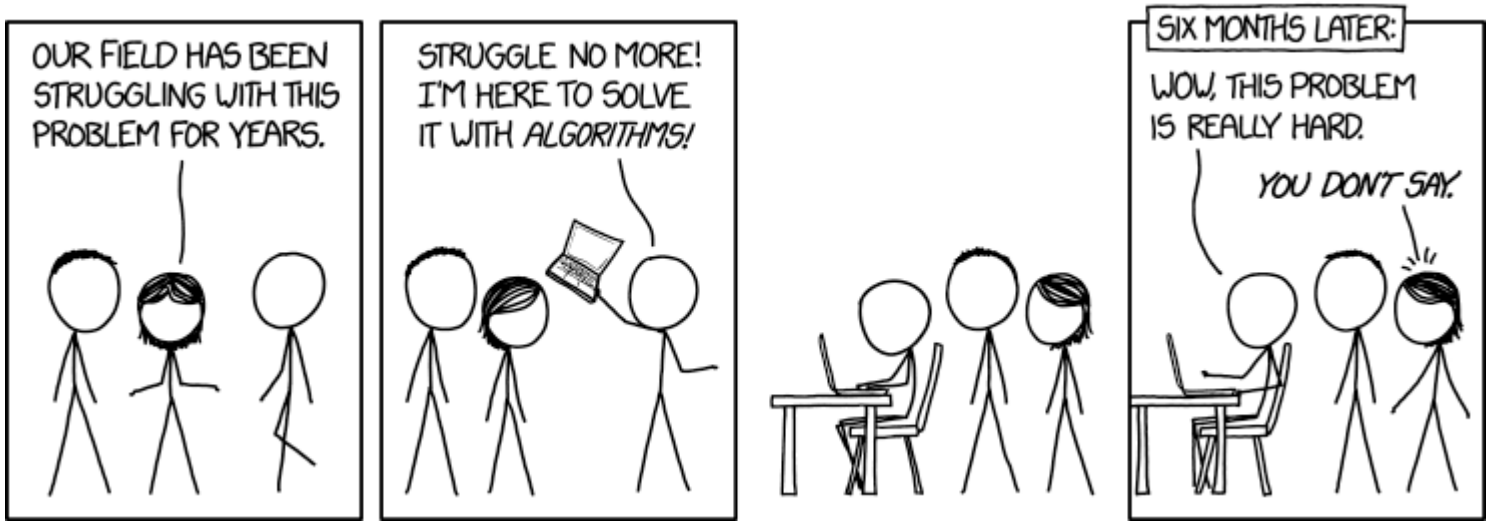


To summarize

- This work provides a first overview; it achieves its aim but more research is needed.
- Our results suggests that the field of AIECS is predominantly populated by domain researchers who use AI as a tool.
  → Models are highly specific, not general purpose
- Many publications can not be reproduced and are therefore difficult to analyze and correctly evaluate.
- The field is growing at a very high speed – older publications may not be truly representative of todays publications.
  - Almost a third of the examined publications, around 5600, were published in 2019 and 2020.

Thank you for your attention

Questions ?



https://xkcd.com/1831/

Bundesministerium für Umwelt, Naturschutz und Nukleare Sicherheit (BMU)

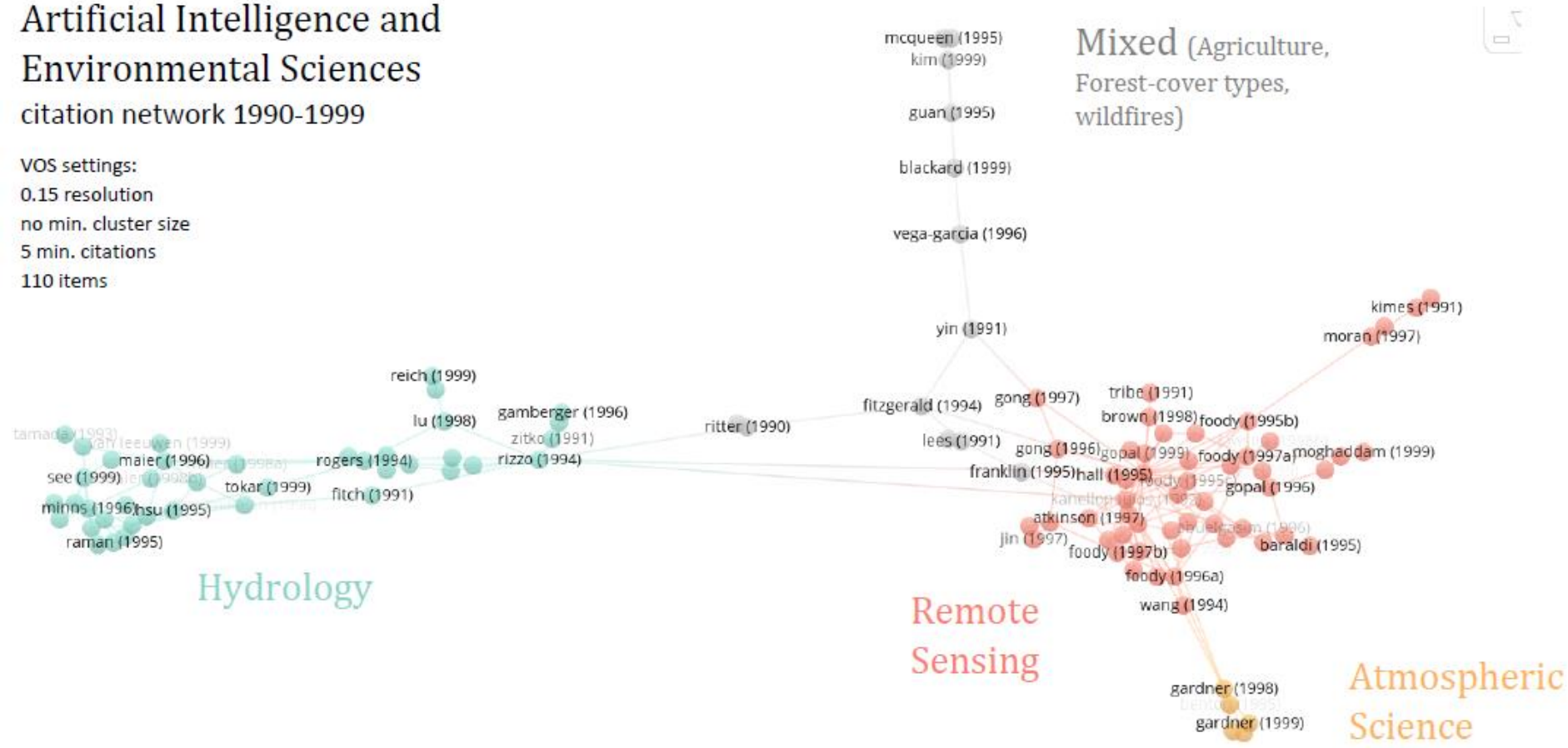
„KI-Leuchttürme für Umwelt, Klima, Natur und Ressourcen“

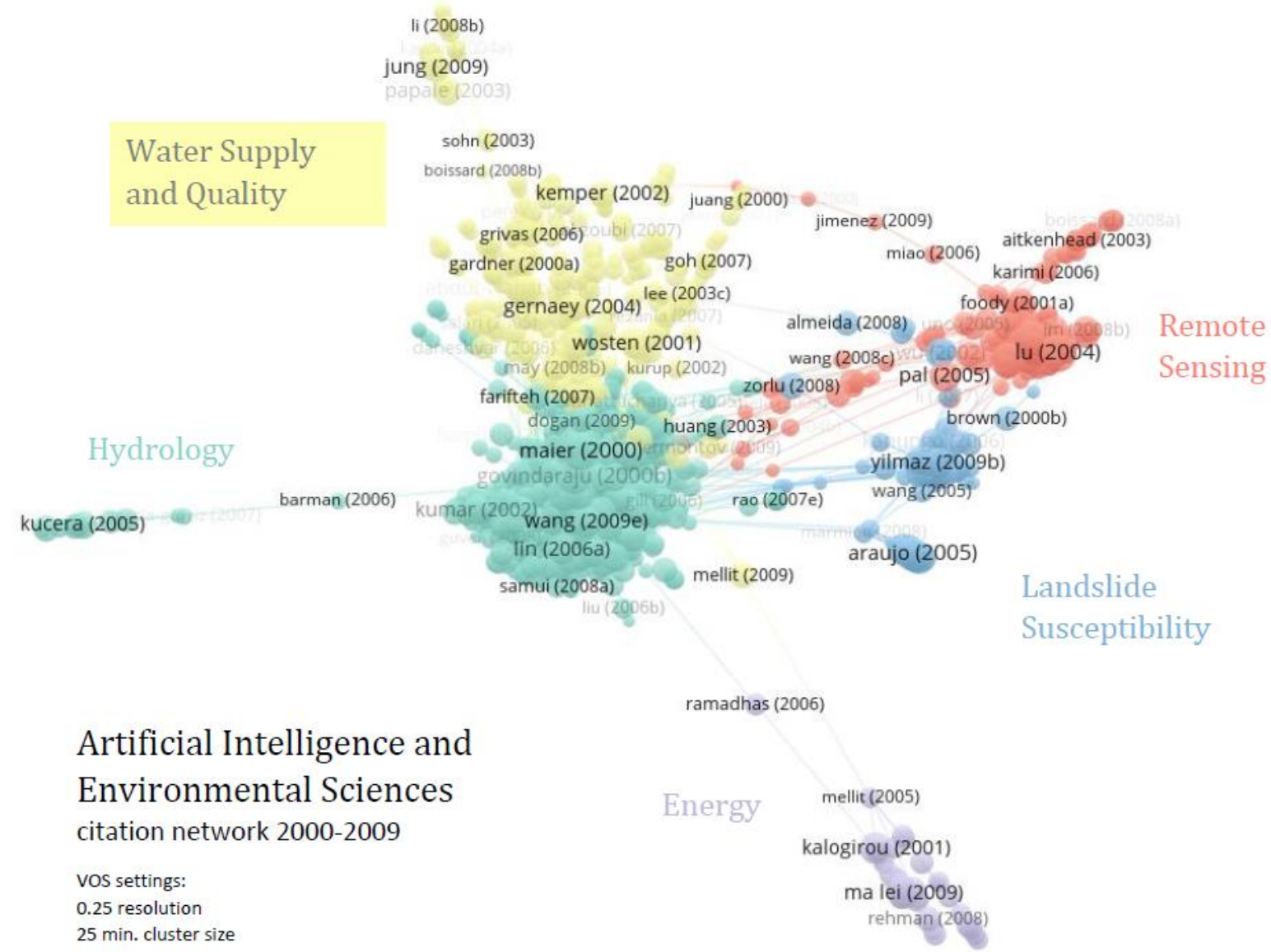
Förderkennzeichen FKZ 67KI2048.

Artificial Intelligence and Environmental Sciences

citation network 1990-1999

VOS settings:
0.15 resolution
no min. cluster size
5 min. citations
110 items

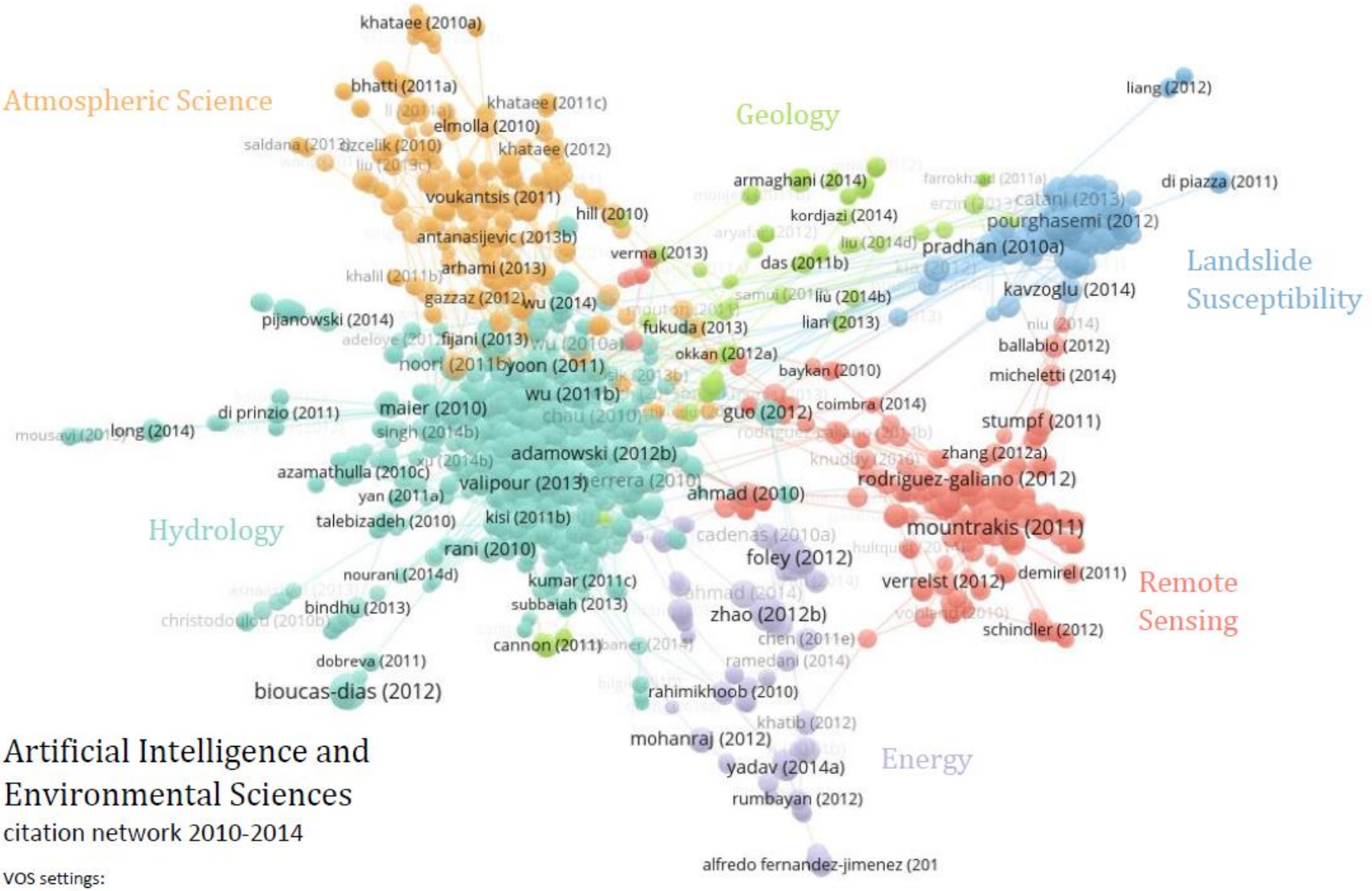




Artificial Intelligence and Environmental Sciences

citation network 2000-2009

VOS settings:
0.25 resolution
25 min. cluster size
min. 5 citations
1265 items – 5 clusters



Artificial Intelligence and Environmental Sciences

citation network 2010-2014

VOS settings:
0.10 resolution
25 min. cluster size
min. 5 citations
1117 items