Center for International Earth Science Information Network
EARTH INSTITUTE | COLUMBIA UNIVERSITY

Assessing the Science Impact of Gridded Population Data: A Pilot Study

Dr. Robert R. Downs[1]

[1] rdowns@ciesin.columbia.edu

Center for International Earth Science Information Network (CIESIN)
The Earth Institute, Columbia University

Presented to
CODATA 2019 Beijing
Towards next-generation data-driven science: policies, practices and platforms
Beijing, China,
19-20 September 2019

Session: Metrics for Assessing the Use, Impact, and Value of Scientific Data (150)

Copyright 2019. The Trustees of Columbia University in the City of New York.

PopGrid Data Products – Available Recommended Citations

Gridded Data Products	Notes	Dataset	Paper
Gridded Population of the World (GPW), version 4	Collection: 9 datasets	9	
Global Human Settlement Layer – Population (GHS-POP)	1 dataset available in 4 different resolutions	1	1
Global Rural Urban Mapping Project (GRUMP)	Collection: 10 datasets	10	1
History Database of the Global Environment (HYDE) Population Grids, version 3.1	Series of 4 datasets		2
LandScan Global Population database	18 datasets, 1/year (2000-2017)	18	
World Population Estimate	9 products and services		1
WorldPop	“Global Per Country 2000-2020 Datasets” are described for ‘Population’		2

POPGRID Data Collaborative https://www.popgrid.org/

Sources of Recommended Citations:
GPWv4: https://sedac.ciesin.columbia.edu/data/collection/gpw-v4
GHS-Pop: https://ghsl.jrc.ec.europa.eu/datasets.php
GRUMP: https://sedac.ciesin.columbia.edu/data/collection/grump-v1
HYDE: https://themasites.pbl.nl/tridion/en/themasites/hyde/download/index-2.html
LandScan: https://landscan.ornl.gov/index.php/node/183
World Population Estimate: https://livingatlas.arcgis.com/en/browse/#d=2&cont=esriOnly&rgnCode=WO&q=World%20Population%20Estimate
WorldPop: https://www.worldpop.org/project/categories?id=3

PopGrid Data Products – Elements of Recommended Citations*

Gridded Data Products	Author or Creator	Public Release Date	Title	Version ID	Repository or Publisher	Resolvable Persistent Identifier (PID)
Gridded Population of the World (GPW), version 4	Y	Y	Y	Y	Y	Y
Global Human Settlement Layer – Population (GHS-POP)	Y	Y	Y	Y	Y	Y
Global Rural Urban Mapping Project (GRUMP)	Y	Y	Y	Y	Y	Y
History Database of the Global Environment (HYDE) Population Grids, version 3.1	Y	Y	Y	Y	Y	Y
LandScan Global Population database	Y	Y	Y	Y	Y	
World Population Estimate	Y	Y	Y	Y	Y	Y
WorldPop	Y	Y	Y	Y	Y	Y

Rows in yellow represent recommended citations for an original article in the absence of a recommended data citation.

* Based on: ESIP Data Preservation and Stewardship Committee. 2019. Data Citation Guidelines for Earth Science Data. Ver. 2. Earth Science Information Partners. https://doi.org/10.6084/m9.figshare.8441816

PopGrid Data Products – Disseminated Data Citations

Gridded Data Products	Online Access	Search Feature	Filter Feature	Number of References
Gridded Population of the World (GPW), version 4	Data citation database	Y	Y	452
Global Human Settlement Layer – Population (GHS-POP)	Selected references & link to bookshop	Y	Y	115
Global Rural Urban Mapping Project (GRUMP)	Data citation database	Y	Y	890
History Database of the Global Environment (HYDE) Population Grids, version 3.1	Selected references			19
LandScan Global Population database	Selected references			33
World Population Estimate	Original Publication			1
WorldPop	Original Publication		Y	51

Sources of Disseminated Data Citations (accessed on 12 August 2019):
GPWv4: https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/citations
GHS-Pop: https://ghsl.jrc.ec.europa.eu/documents.php
GRUMP https://sedac.ciesin.columbia.edu/data/collection/grump-v1/citations
HYDE: https://themasites.pbl.nl/tridion/en/themasites/hyde/publications/index-2.html
LandScan: https://landscan.ornl.gov/publications
World Population Estimate: https://datascience.codata.org/articles/10.5334/dsj-2018-020/ [Retrieved from Google Scholar from DOI query]
WorldPop: https://www.worldpop.org/publications

PopGrid Data Products – Disaggregated Disseminated Citations

Gridded Data Products	Reported Citations	Journal Articles	Books, Conference Proceedings, Reports, etc.
Gridded Population of the World (GPW), version 4	452	369	83
Global Human Settlement Layer – Population (GHS-POP)	115	11	104
Global Rural Urban Mapping Project (GRUMP)	890	681	209
History Database of the Global Environment (HYDE) Population Grids, version 3.1	19	12	7
LandScan Global Population database	32	16	16
World Population Estimate	1	1	0
WorldPop	51	50	1

Sources of Disseminated Data Citations (accessed on 19 August 2019):
GPWv4: https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/citations
GHS-Pop: https://ghsl.jrc.ec.europa.eu/documents.php
GRUMP https://sedac.ciesin.columbia.edu/data/collection/grump-v1/citations
HYDE: https://themasites.pbl.nl/tridion/en/themasites/hyde/publications/index-2.html
LandScan: https://landscan.ornl.gov/publications
World Population Estimate: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C33&q=10.5334%2Fdsj-2018-020&btnG=
WorldPop: http://citations.springer.com/item?doi=10.1038/sdata.2017.4

PopGrid Data Products – Disciplines of Reported Data Citations

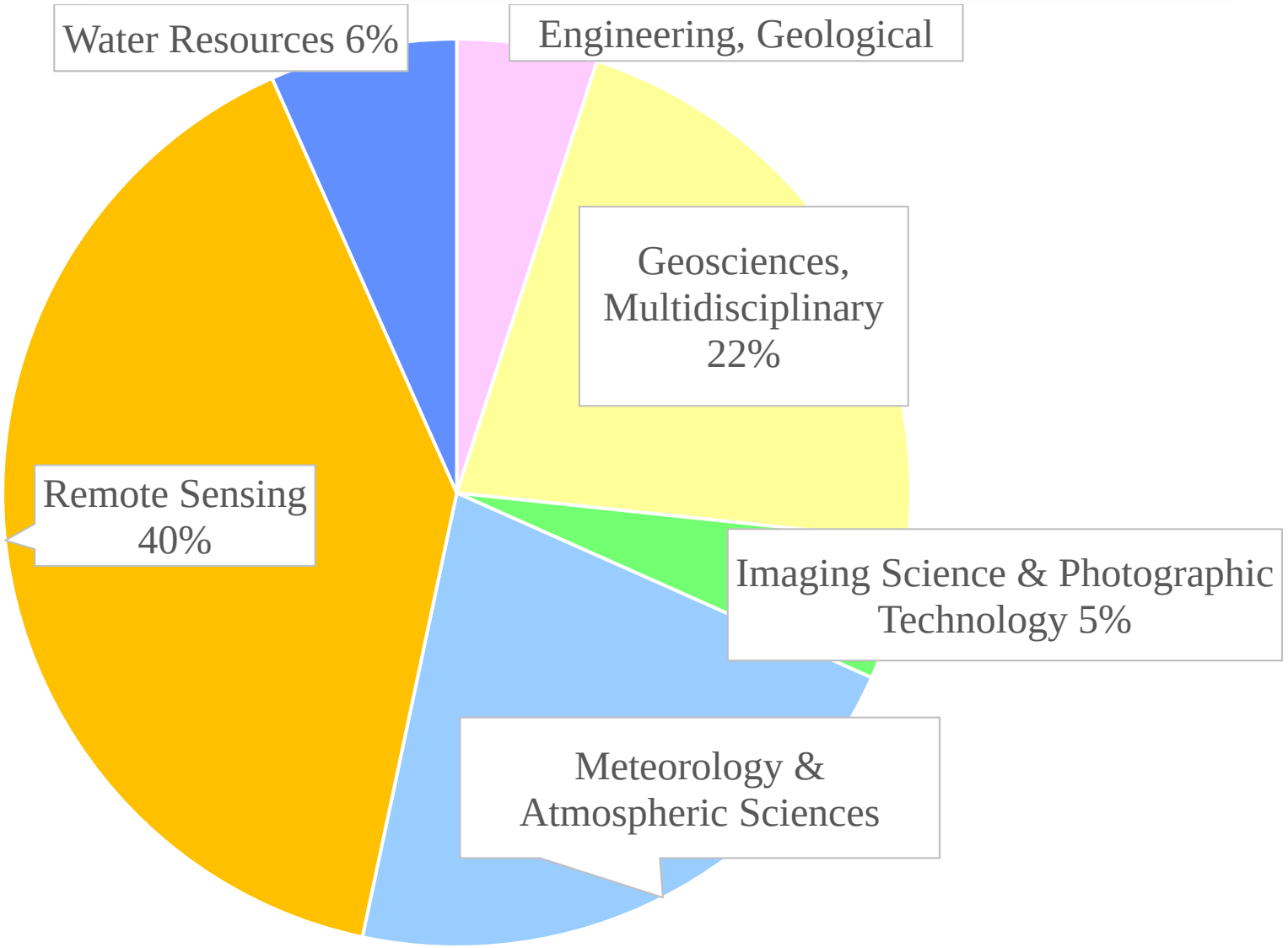
Gridded Data Products	Reported Citations	Journal Articles	Citing Journals	Discipline Categories
Gridded Population of the World (GPW), version 4	452	369	221	85
Global Human Settlement Layer – Population (GHS-POP)	115	11	6	6
Global Rural Urban Mapping Project (GRUMP)	890	681	289	95
History Database of the Global Environment (HYDE) Population Grids, version 3.1	19	12	9	12
LandScan Global Population database	32	16	11	16
World Population Estimate	1	1	1	2
WorldPop	51	50	34	49

Sources of Disseminated Data Citations (accessed on 19 August 2019):
GPWv4: https://sedac.ciesin.columbia.edu/data/collection/gpw-v4/citations
GHS-Pop: https://ghsl.jrc.ec.europa.eu/documents.php
GRUMP https://sedac.ciesin.columbia.edu/data/collection/grump-v1/citations
HYDE: https://themasites.pbl.nl/tridion/en/themasites/hyde/publications/index-2.html
LandScan: https://landscan.ornl.gov/publications
World Population Estimate: https://scholar.google.com/scholar?hl=en&as_sdt=0%2C33&q=10.5334%2Fdsj-2018-020&btnG=
WorldPop: http://citations.springer.com/item?doi=10.1038/sdata.2017.4

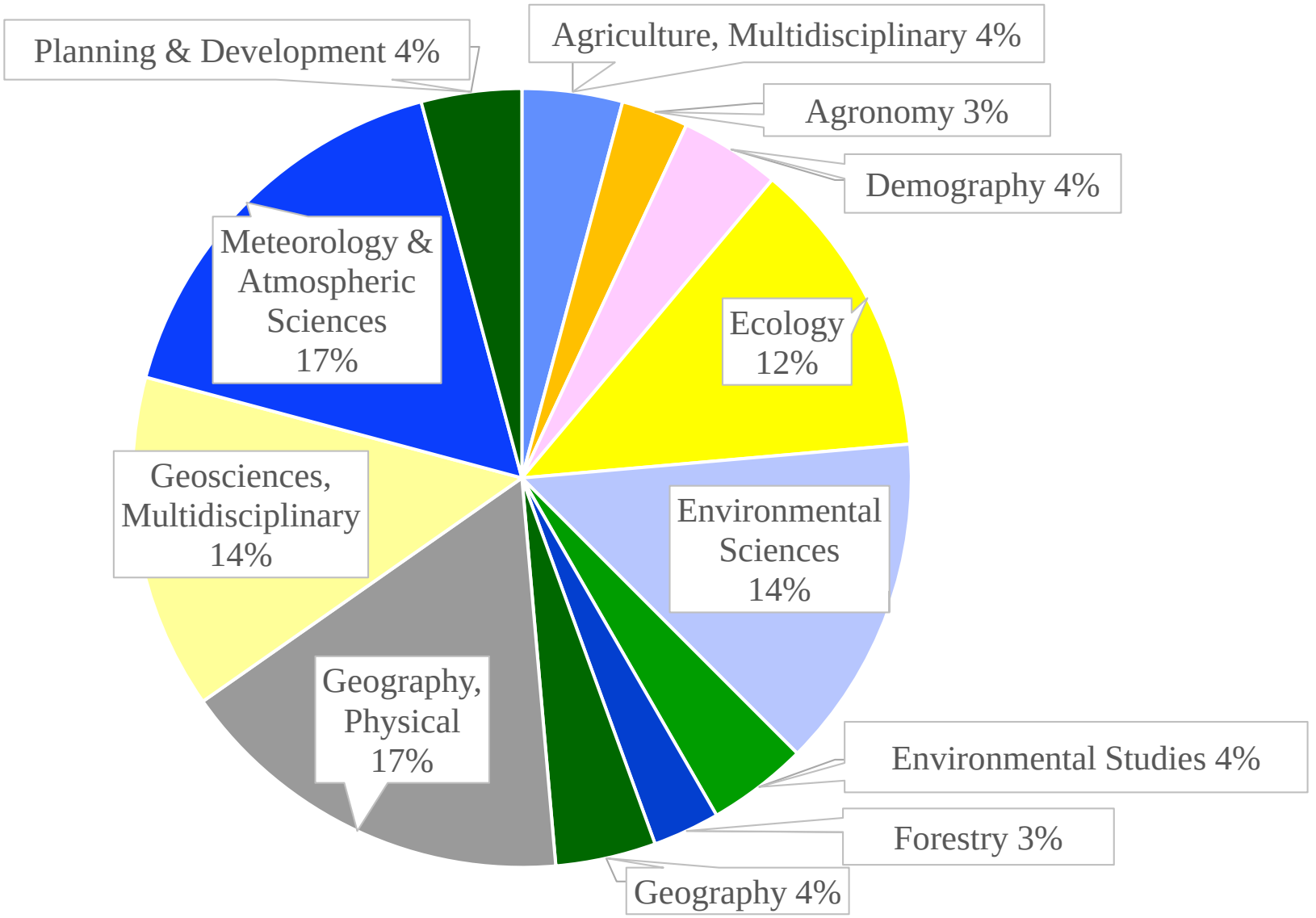
World Population Estimate

- One article cited the original article
  - Tuholske C, Caylor K, Evans T, Avery R. 2019. Environmental Research Letters, Volume 14(8). https://doi.org/10.1088/1748-9326/ab2432
- WoS Categories of citing journal (2 Categories):
  - Environmental Sciences;
  - Meteorology & Atmospheric Sciences

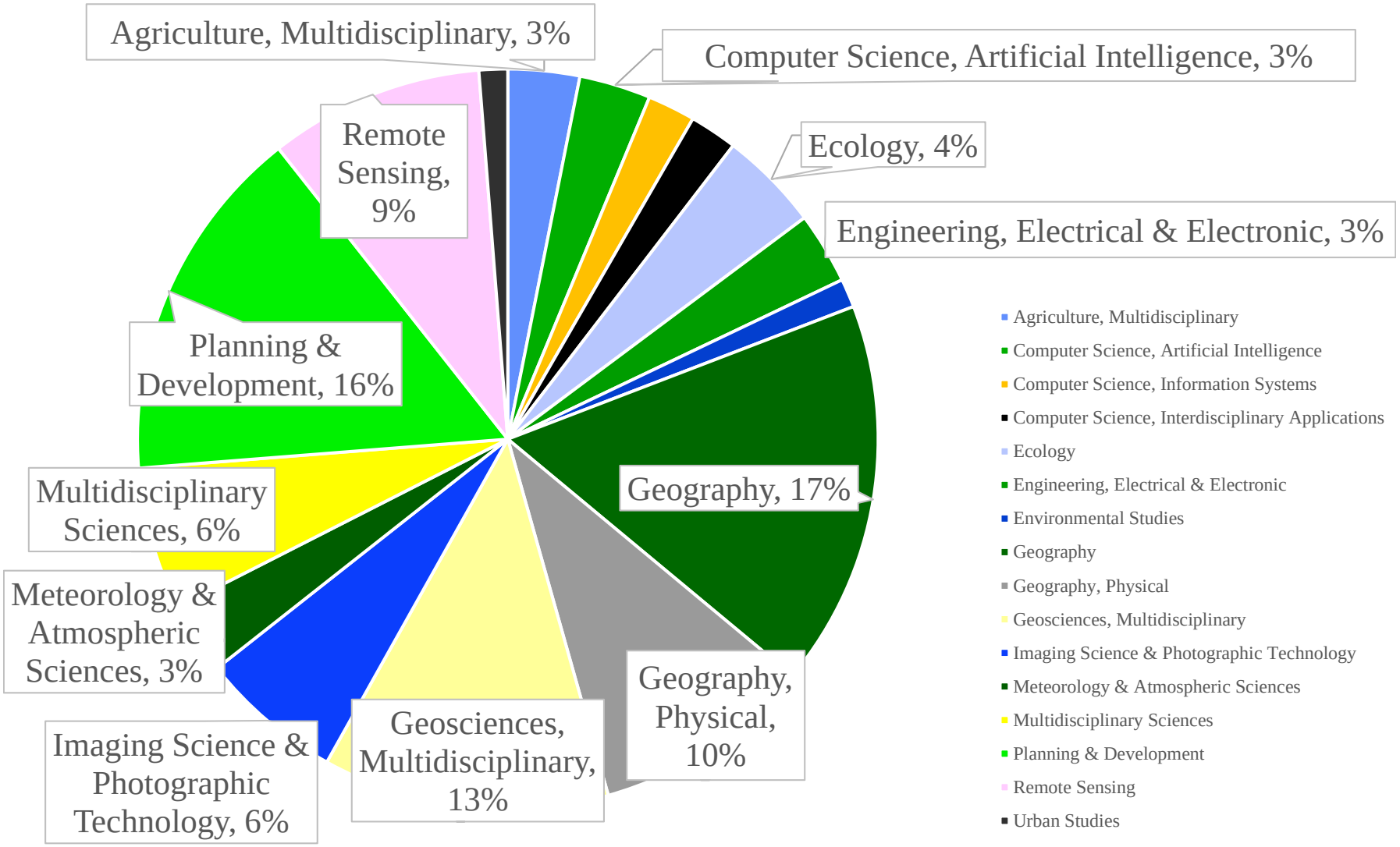
6 Discipline Categories of GHS-Pop Data Citations



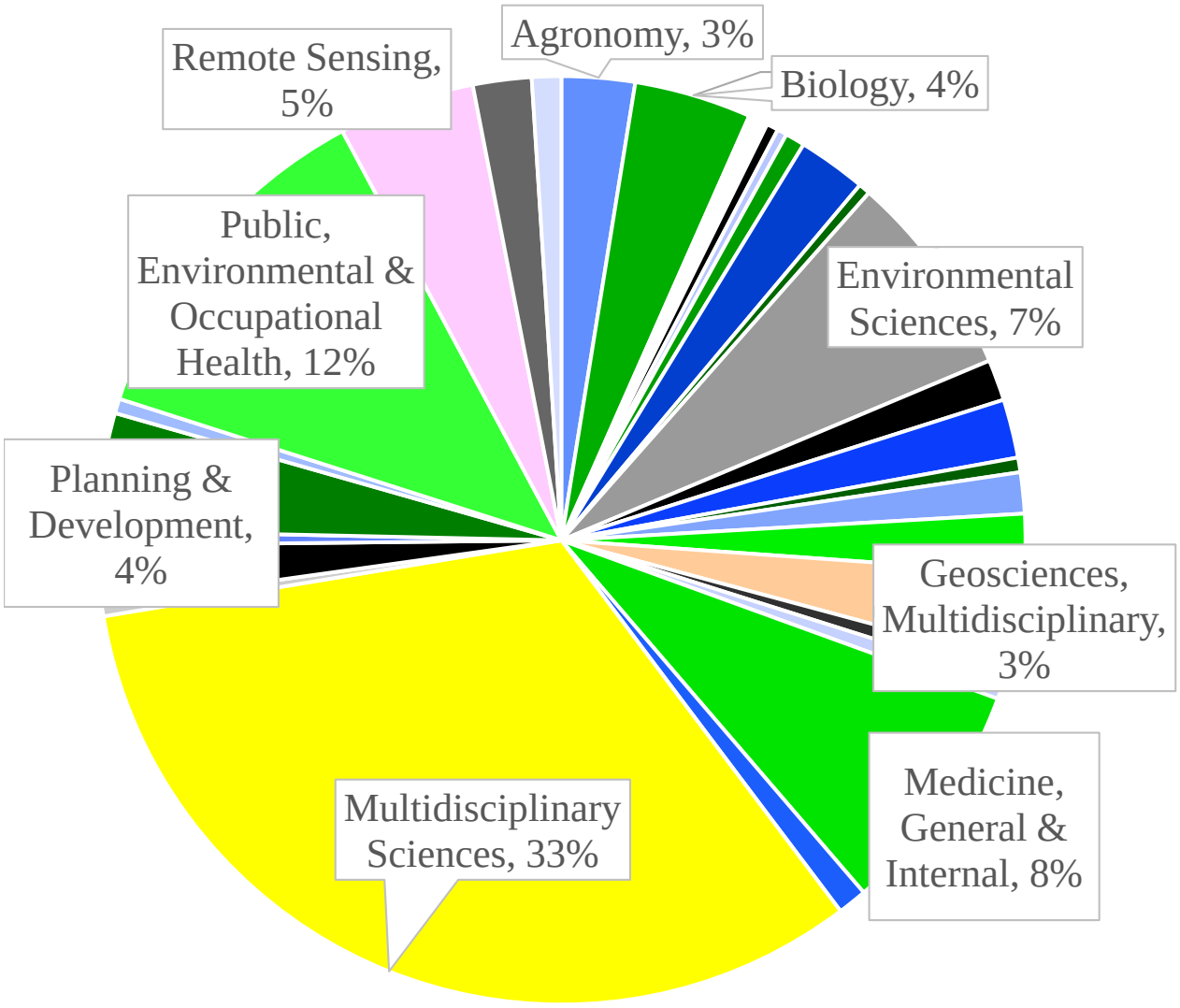
12 Discipline Categories of HYDE Citations



16 Discipline Categories of LandScan Data Citations

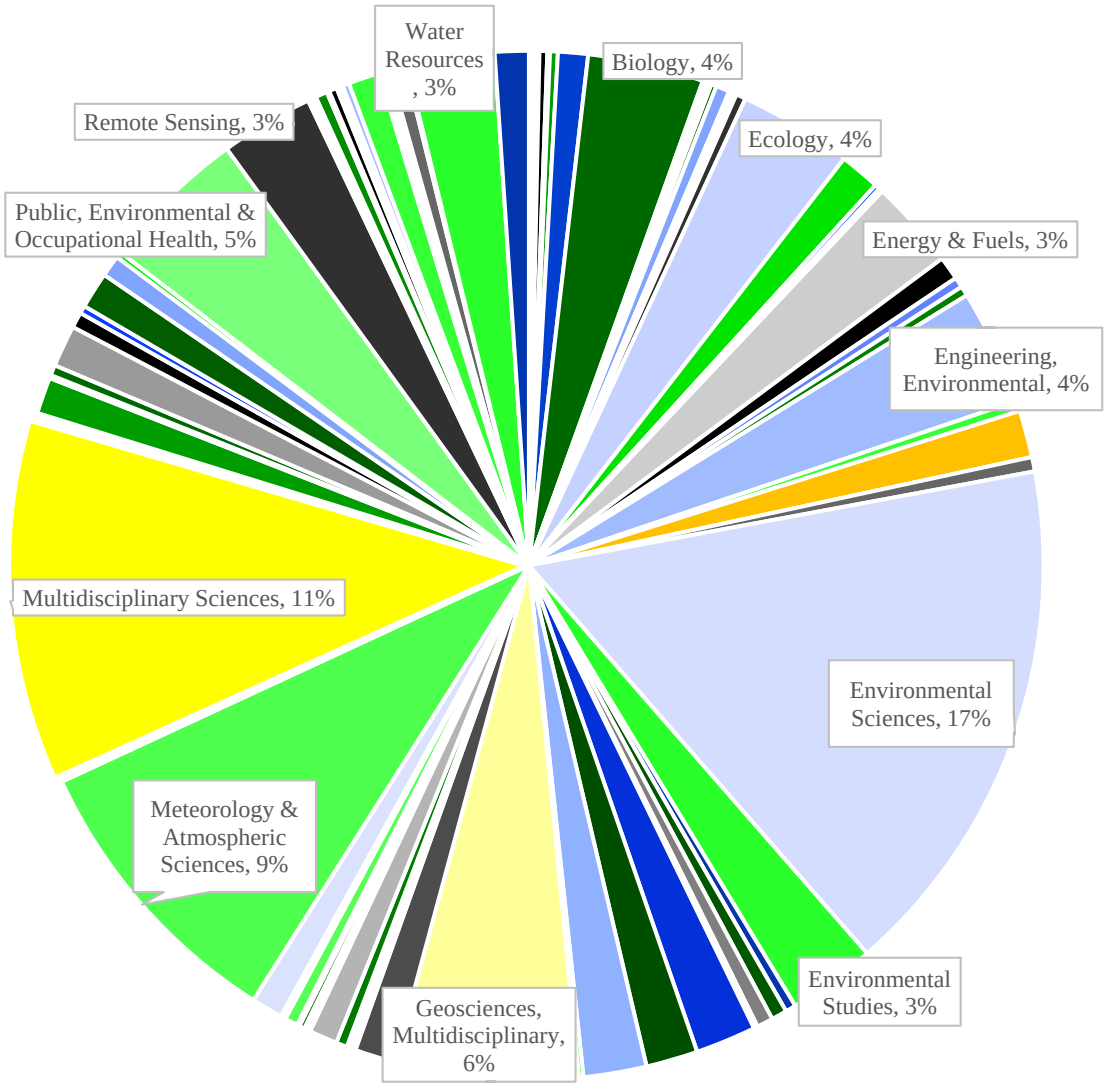


49 Discipline Categories of WorldPop Citations



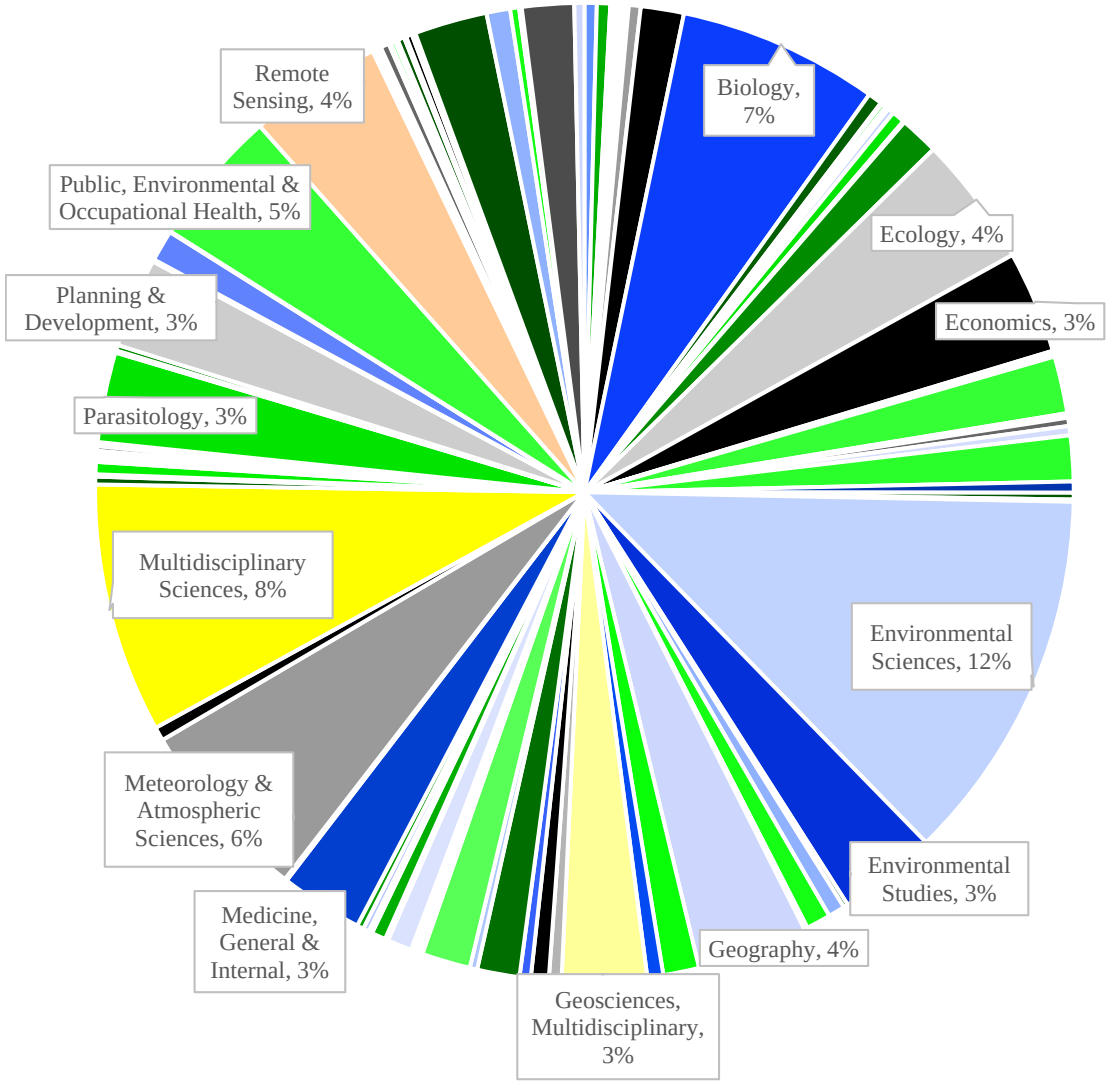
- Agronomy
- Biology
- Chemistry, Analytical
- Computer Science, Interdisciplinary Applications
- Education, Scientific Disciplines
- Electrochemistry
- Energy & Fuels
- Engineering, Environmental
- Environmental Sciences
- Environmental Studies
- Forestry
- Genetics & Heredity
- Geography
- Geography, Physical
- Geosciences, Multidisciplinary
- Imaging Science & Photographic Technology
- Instruments & Instrumentation
- Medicine, General & Internal
- Microbiology
- Multidisciplinary Sciences
- Operations Research & Management Science
- Physics, Multidisciplinary
- Physiology
- Planning & Development
- Plant Sciences
- Public, Environmental & Occupational Health
- Remote Sensing
- Statistics & Probability
- Tropical Medicine

85 Discipline Categories of GPWv4 Data Citations



- Agriculture, Dairy & Animal Science
- Agriculture, Multidisciplinary
- Agronomy
- Astronomy & Astrophysics
- Automation & Control Systems
- Biochemistry & Molecular Biology
- Biodiversity Conservation
- Biology
- Business
- Cell Biology
- Chemistry, Physical
- Computer Science, Information Systems
- Computer Science, Interdisciplinary Applications
- Computer Science, Software Engineering
- Construction & Building Technology
- Demography
- Ecology
- Economics
- Education, Scientific Disciplines
- Emergency Medicine
- Energy & Fuels
- Engineering, Aerospace
- Engineering, Civil
- Engineering, Electrical & Electronic
- Engineering, Environmental
- Engineering, Geological
- Engineering, Multidisciplinary
- Entomology
- Environmental Sciences
- Environmental Studies
- Evolutionary Biology
- Food Science & Technology
- Forestry
- Genetics & Heredity
- Geochemistry & Geophysics
- Geography
- Geography, Physical
- Geology
- Geosciences, Multidisciplinary
- Green & Sustainable Science & Technology
- Health Policy & Services
- History
- History of Social Sciences
- Imaging Science & Photographic Technology
- Infectious Diseases
- Information Science & Library Science

95 Discipline Categories of GRUMPv1 Data Citations



- Agricultural Economics & Policy
- Agriculture, Multidisciplinary
- Agronomy
- Area Studies
- Astronomy & Astrophysics
- Automation & Control Systems
- Behavioral Sciences
- Biochemical Research Methods
- Biochemistry & Molecular Biology
- Biodiversity Conservation
- Biology
- Business, Finance
- Chemistry, Analytical
- Chemistry, Physical
- Clinical Neurology
- Communication
- Computer Science, Information Systems
- Computer Science, Interdisciplinary Applications
- Construction & Building Technology
- Demography
- Ecology
- Economics
- Education, Scientific Disciplines

Recommendations to Improve Science Impact of PopGrid Data

- Provide online access to recommended data citation *
  - Easily discoverable recommended data citation
  - Recommended data citation for data as well as for original article
- Include elements of data citation in recommended data citation
  - Version number differentiates current version from previous versions
  - Persistent identifier, such as a DOI, is critical for identifying citations
- Provide access to publications that cite data (data citations)
  - Associate data citations with the collection or dataset cited
  - Provide capabilities for searching (full-text, faceted) data citations

* ESIP Data Preservation and Stewardship Committee. 2019. Data Citation Guidelines for Earth Science Data. Ver. 2. Earth Science Information Partners. https://doi.org/10.6084/m9.figshare.8441816

Limitations

- Investigator involved in production of gridded data products
  - GPWv4 and GRUMPv1 collections
- Data citations analyses based on online accessible data citations
  - Incomplete or selected identification and dissemination of data citations
  - Variation on categorization of journal articles versus conference proceedings
- Multiple sources used to identify categories of journals
  - WoS category, then Scopus, then journal scope statement

Thank you!

- Joe Schumacher, of CIESIN, provided formatted reports of data citations for GPWv4 and GRUMP data

- Support from sponsors of the PopGrid Data Collaborative project is appreciated, including the Columbia University Earth Institute's Cross-Cutting Initiatives, the Bill & Melinda Gates Foundation, and the Group on Earth Observations (GEO) Human Planet Initiative (HPI). https://www.popgrid.org

- Support from ESIP also is appreciated. http://esipfed.org