

Measuring the Interdisciplinary Value of SEDAC Data

Robert R. Downs¹

¹ rdowns@ciesin.columbia.edu

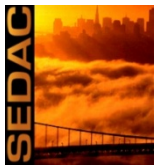
NASA Socioeconomic Data and Applications Center (SEDAC)
Center for International Earth Science Information Network (CIESIN)
The Earth Institute, Columbia University

2017 ESIP Winter Meeting
Bethesda, Maryland; 11-13 January 2017

Session: Measuring the Value of Data
Friday, 13 January 2017; 11:00 a.m. - 12:30 p.m.



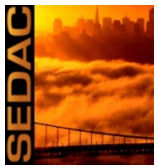
Why Measure the Value of Data?



- Inform planning for data collections
 - Maintenance of current data products and services
 - Development of future data products and services
- Identify value of data products, services, and collections
 - Recent and current value to communities served
 - Potential future value to extent or new user communities
- Determine if costs of data products and services are justified
 - Past, current, and future data collection costs
 - Past, current, and future data curation costs



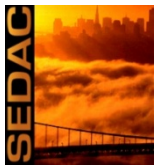
Stakeholder Interests in Measuring Data Value



- Data Producers
 - Want to collect data that will be valuable to themselves and others
- Data Centers
 - Want to focus efforts on managing data products and services of value
- Data Sponsors
 - Want to sponsor collection and curation of valuable data and services
- Data Users
 - Want access to data products and services that they can use



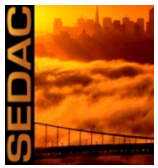
Which Data are Valued?



- Data products and services that are used are valuable
 - Use by the designated user community may be most valued
- Future value of data products and services may be anticipated
 - Current value of data products and services may indicate future value



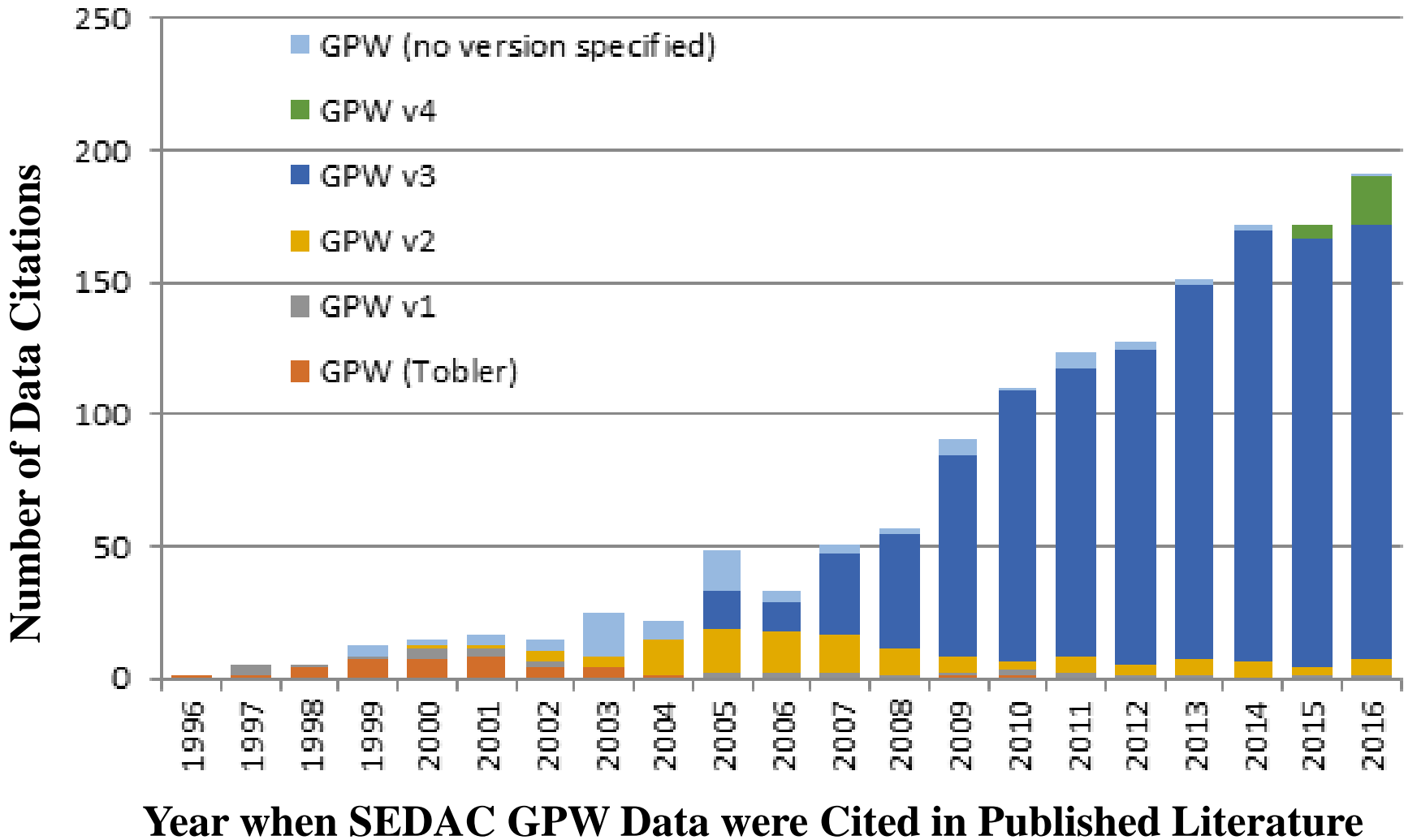
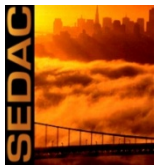
Measuring the Value of Science Data



- Formal use
 - Citations as references in scientific publications
 - Citations in reports, textbooks, popular media, social media
- Informal use
 - Downloads of data and documentation, landing page views
 - Questions and comments to reference (help desk) staff
- Unmeasured use
 - Not cited (omitted, incomplete, or incorrect) in scientific publications
 - Downloads from other sources (aggregators, colleagues)

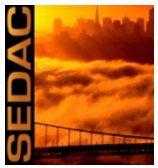


Gridded Population of the World (GPW) Data: Citations By Year





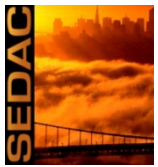
Value in Interdisciplinary Use (or Reuse) of Data



- Use within fields beyond disciplines of data collectors
- Use by teams representing multiple disciplines
- Use for purposes unintended by data collectors
- Combined with data from other disciplines to foster new uses



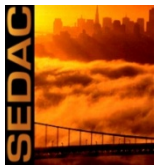
Interdisciplinary Aspects of SEDAC Mission



- Mission: “develop and operate applications that support the integration of socioeconomic and Earth science data and to serve as an "Information Gateway" between the Earth and social sciences”



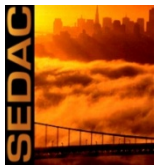
Methodology



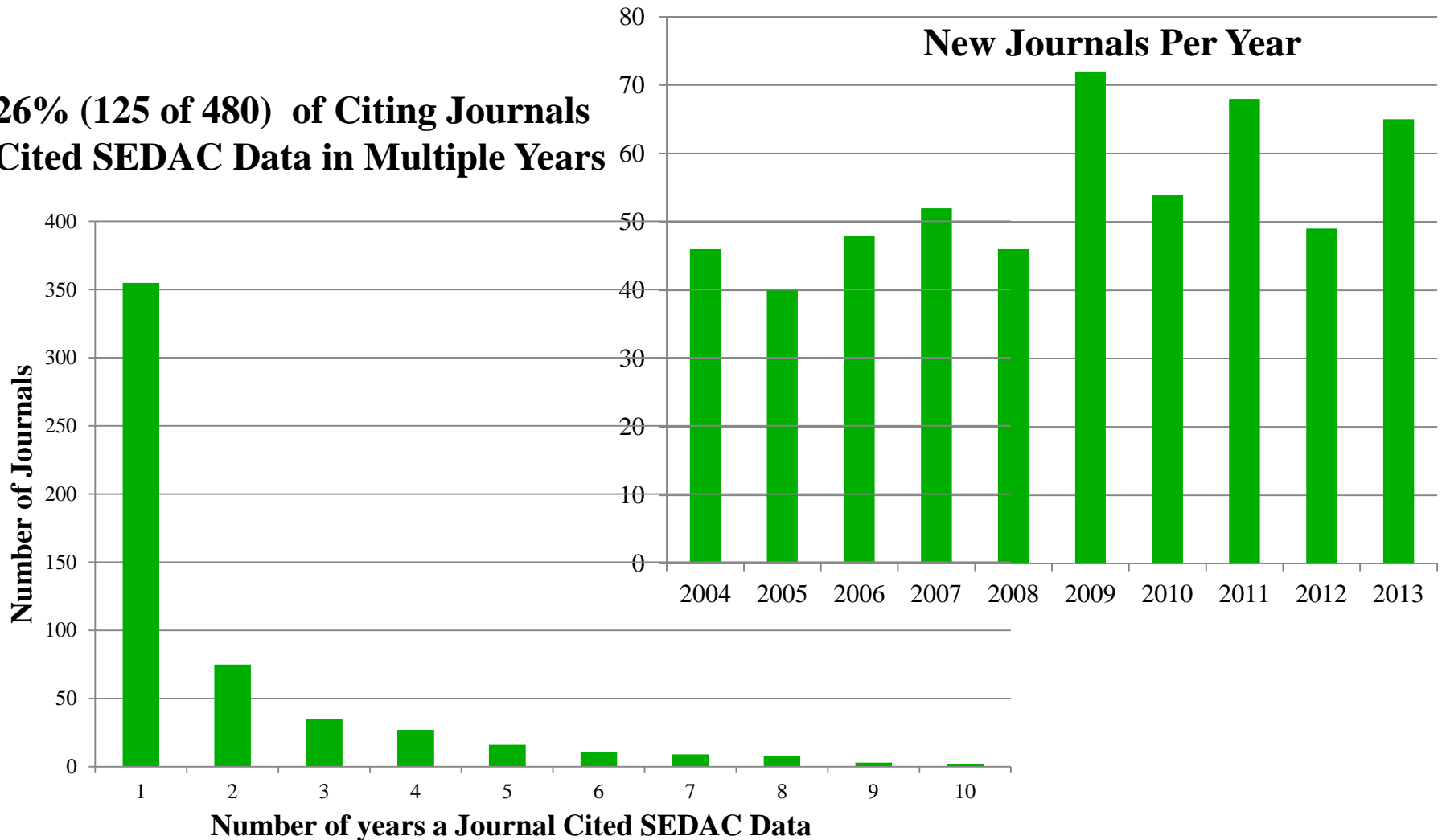
- **Collected citations of SEDAC data**
 - Obtained list of citing journal articles published during 2004 - 2013
- **Obtained classifications to identify disciplines of each citing journal**
 - Obtained Web of Science® (WoS) Category assignments
 - Obtained Web of Knowledge® Subject Classification of WoS Categories (WoK5.3) and Equivalent General Categories and Subject Areas
 - Obtained ScienceWatch® Field Definitions of Major Fields
- **Identified multidisciplinary use of SEDAC data**
 - Identified WoS Categories assigned to journals citing SEDAC data, and used Scopus®, journal titles, and publisher sites when Categories were not assigned
 - Paired assigned WoS Categories to Equivalent General Categories and Subjects
 - Identified Major Fields corresponding to assigned WoS Categories and Subjects
 - Normalized journals with WoS Categories, (WoK5.3) and Equivalent General Categories and Subject areas, and Field Definitions of Major Fields
 - Identified Categories, Subjects, and Major Fields of journals citing SEDAC data



540 Journals Published 1480 Articles Citing SEDAC Data During 2004-2013

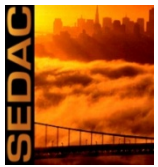


**26% (125 of 480) of Citing Journals
Cited SEDAC Data in Multiple Years**

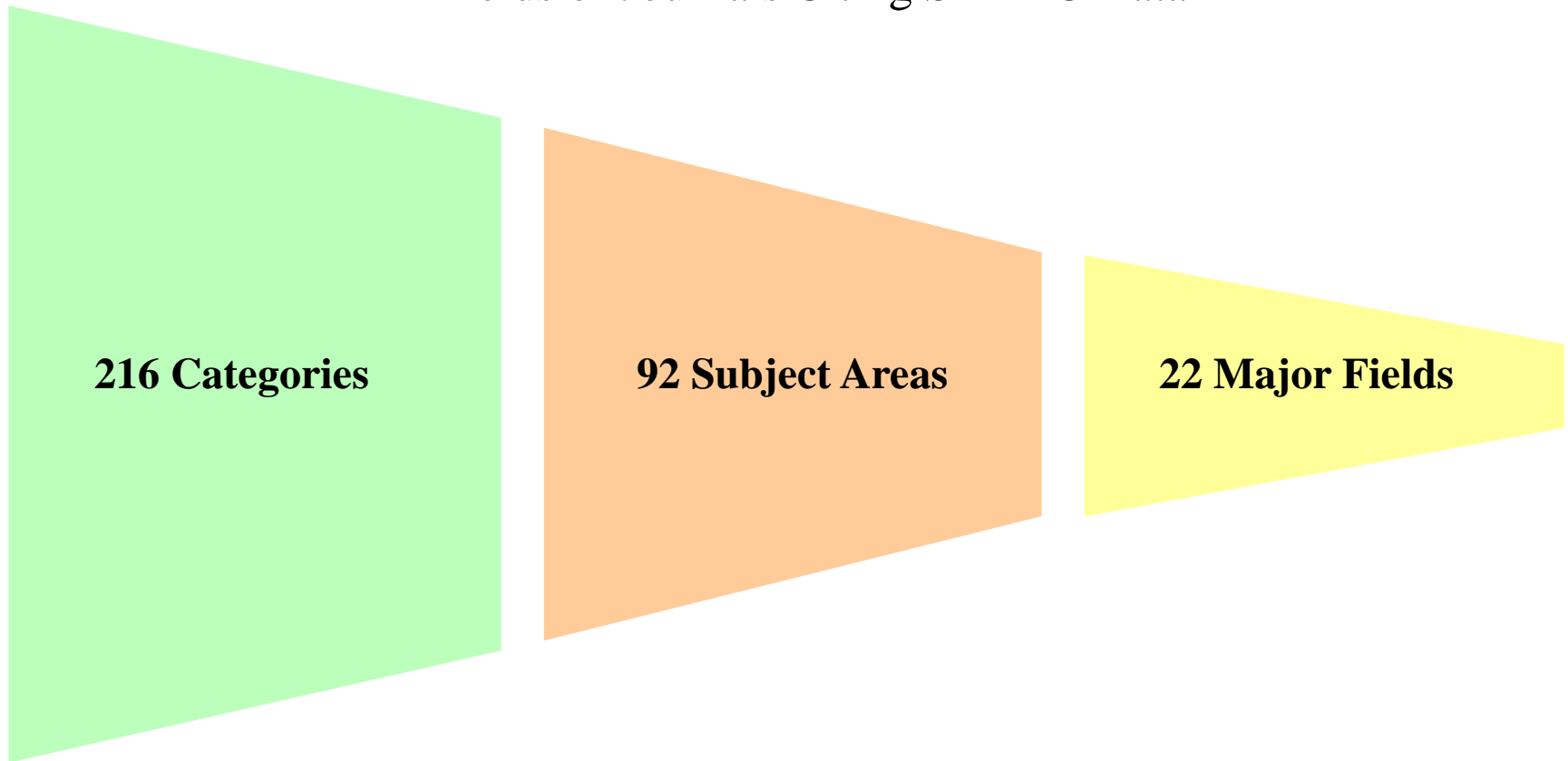




Multidisciplinary Impact of SEDAC Data: 2004-2013

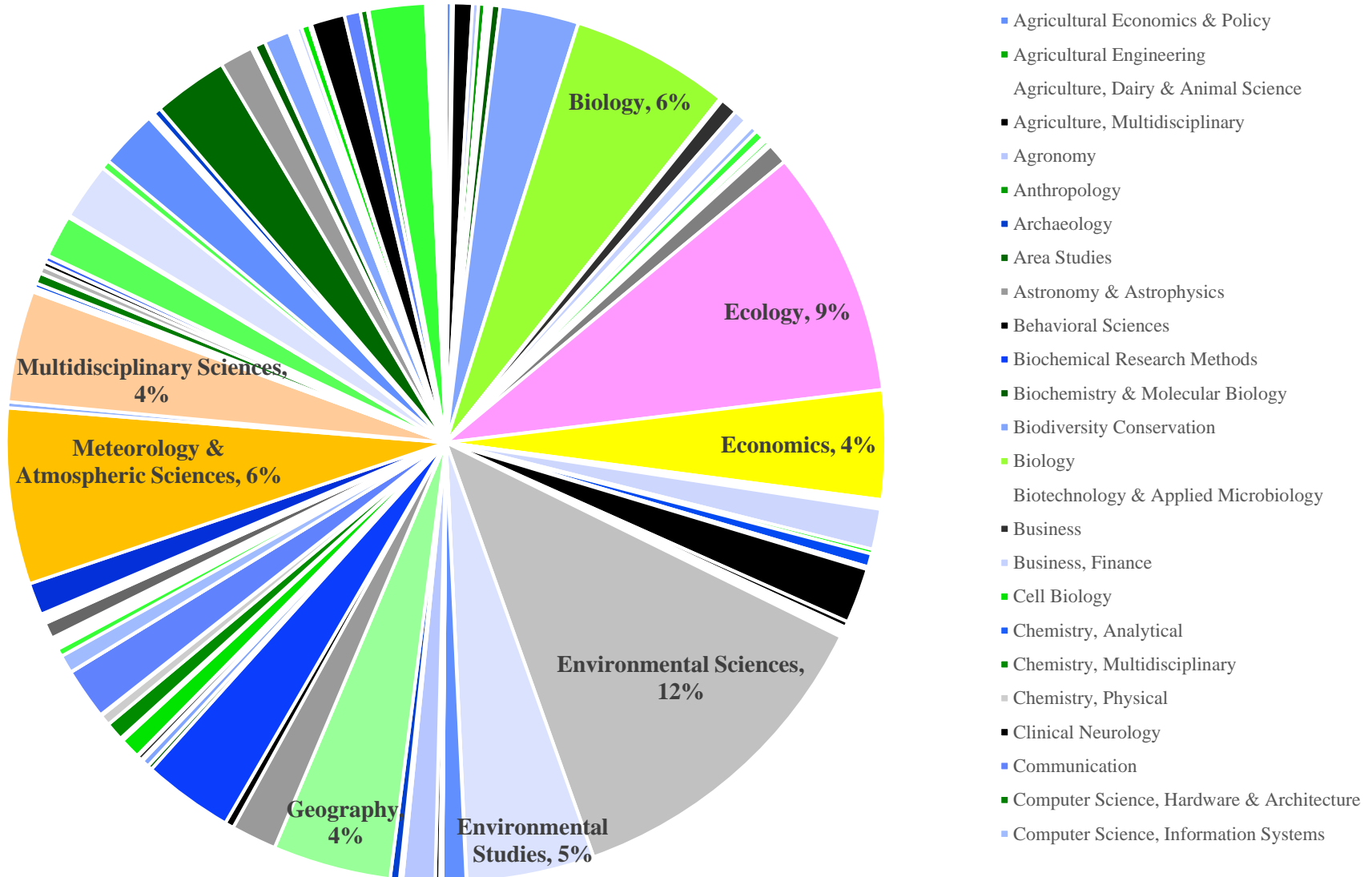
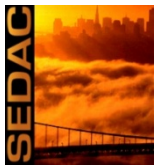


Web of Science Categories, Subject Areas, and Major
Fields of Journals Citing SEDAC Data



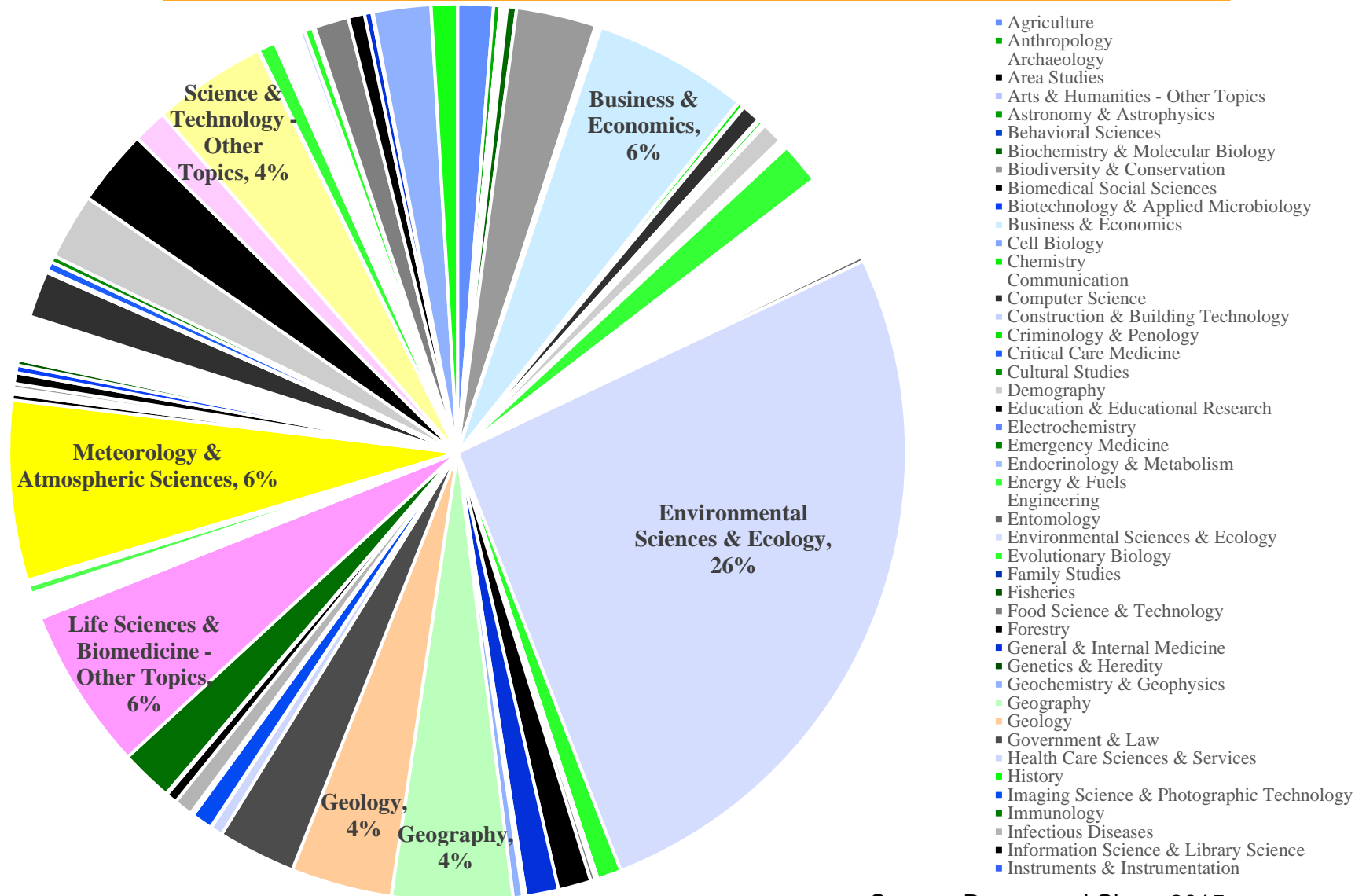
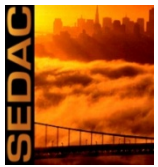


216 WoS Categories of Journals Containing Citations of SEDAC Data from 2004 through 2013





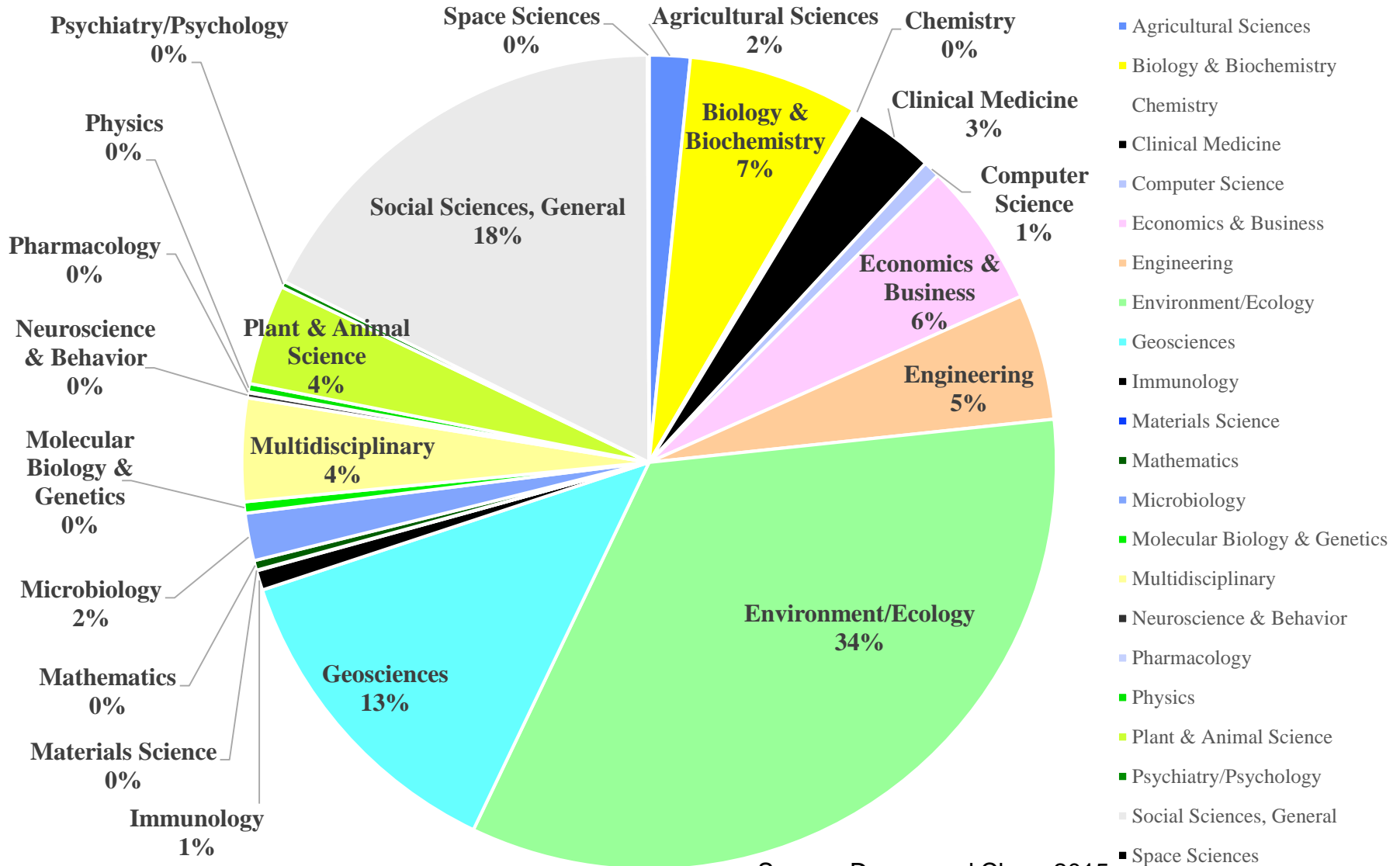
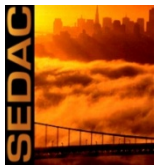
Proportional Representation of 92 WoK Subject Areas of Journals that Cited SEDAC Data 2004-2013



Source: Downs and Chen. 2015



22 Major Fields of Science Represented by Journals that Cited SEDAC Data 2004-2013



Source: Downs and Chen. 2015