



JASMIN (STFC/Stephen Kill)



# CEDA, JASMIN and Managing User Communities

UK Tier 0 Collaboration Meeting  
14-16 March 2018

**Victoria Bennett**

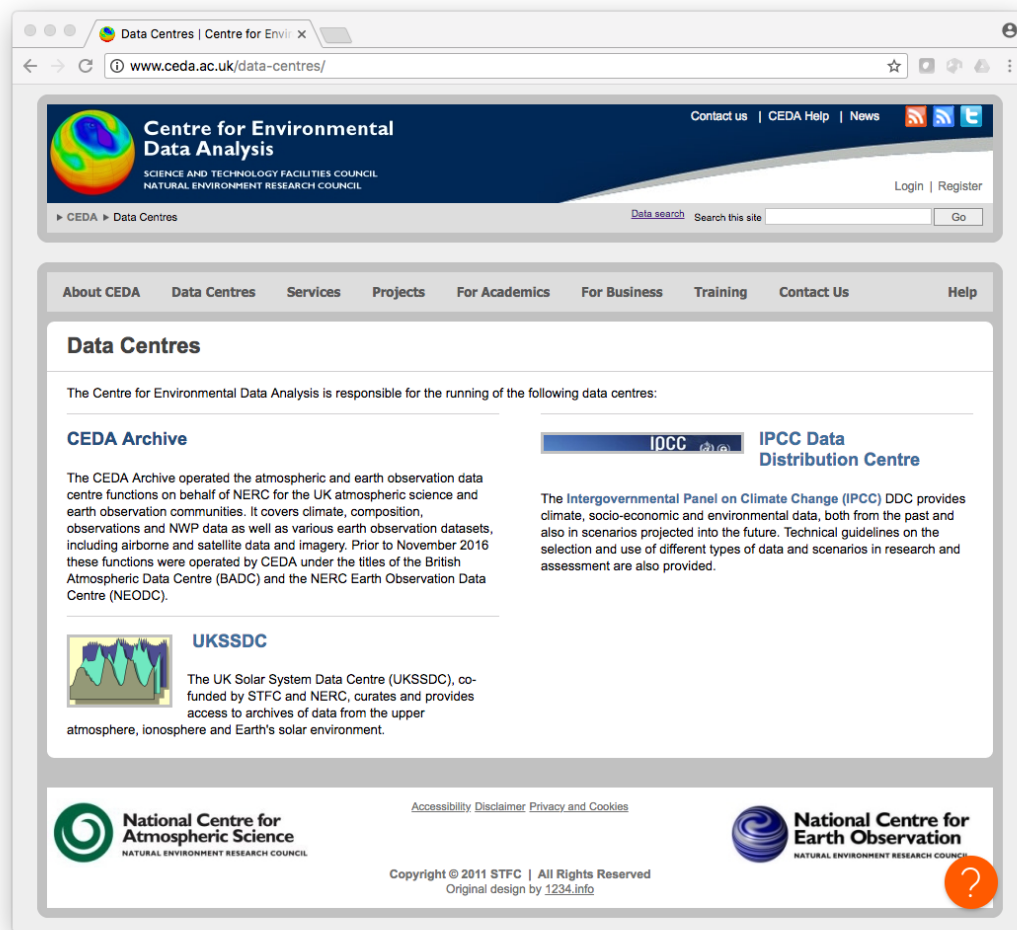
NCAS/NCEO, Centre for Environmental Data Analysis, RAL Space



# CEDA Context

“to support environmental science, further environmental data archival practices, and develop and deploy new technologies to enhance access to data”

➔ Curation & Facilitation





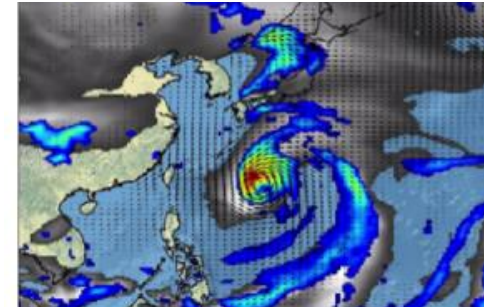
# JASMIN Context

- JASMIN provides a Data Commons for environmental sciences community
- Hosted & operated by STFC for NERC
- *An STFC cross-department collaboration*
  - STFC Scientific Computing Department
    - Hardware architecture, cloud software architecture & operation
  - CEDA (RAL Space)
    - User services & applications development, cloud software architecture
- Both partners involved from the outset in architectural design and implementation

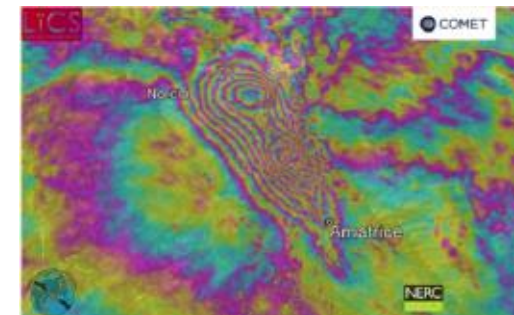


# Science Goals

- Support the science & impact objectives of major stakeholders
  - NERC : NCAS & NCEO
- Help community deal with demands of 21<sup>st</sup> century environmental science
  - Growth in numerical simulation ...& simulated data
  - Increasing complexity of algorithms needed to exploit data
  - Managing & preserving increasing volumes of NERC community data
  - Address digital skills gap
  - Curating digital information products of national significance



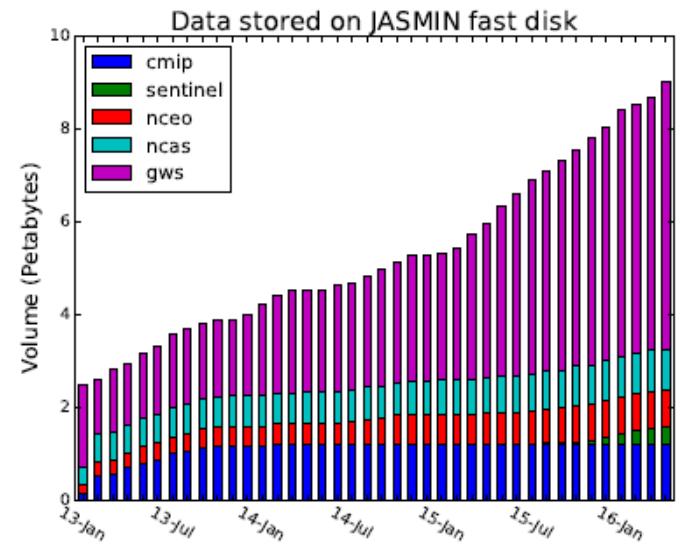
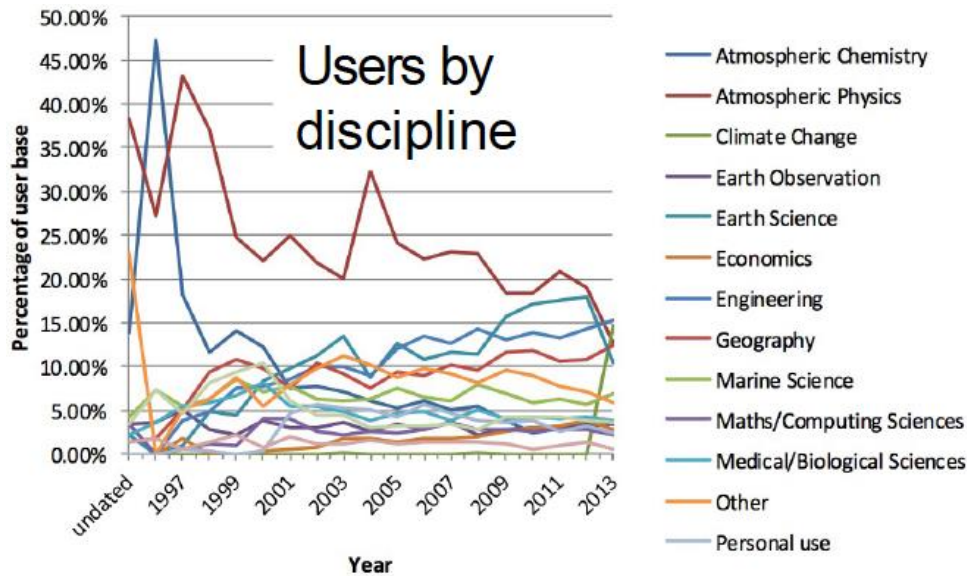
High Res Climate Model



Seismic Fault analysis

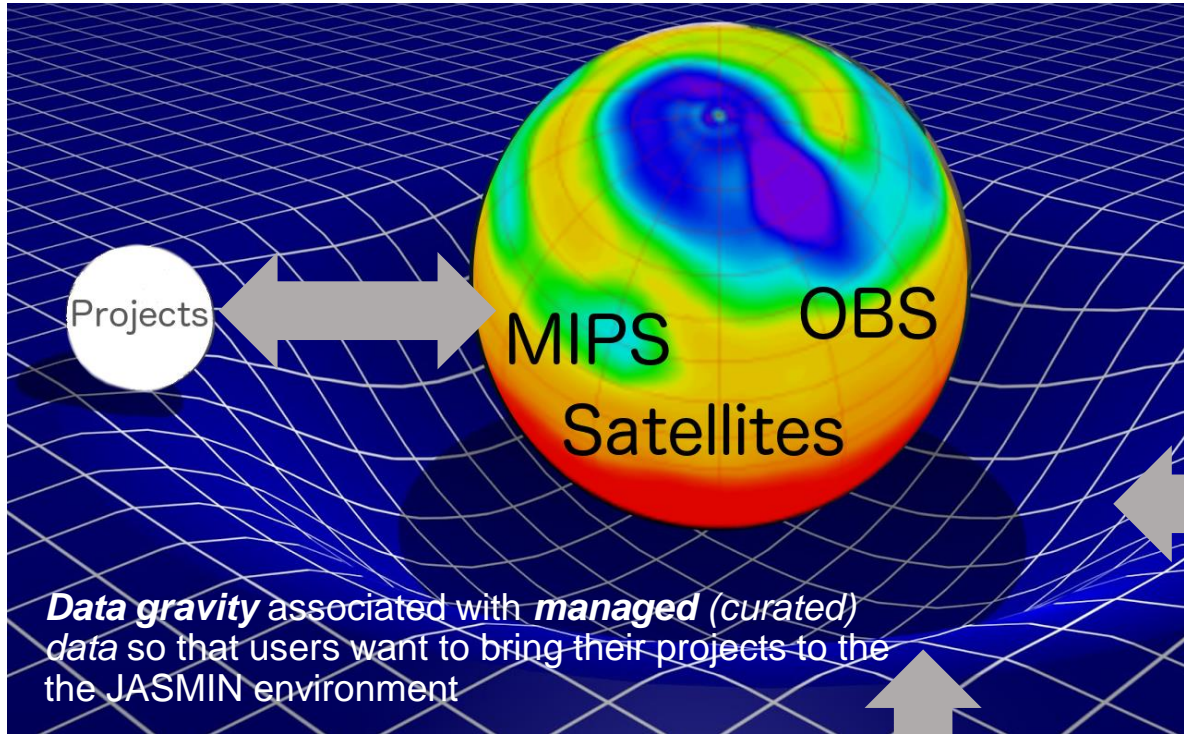


# Data Growth and Diversification of user community



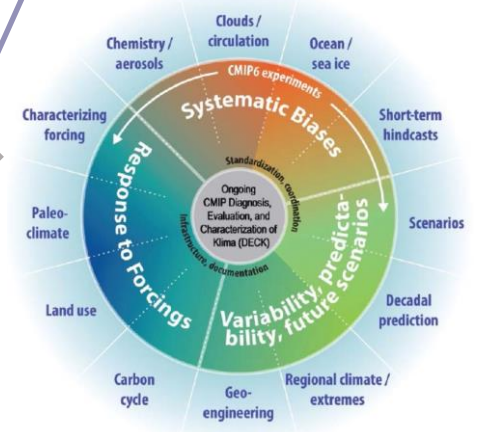
2013-2016 increasing data storage on JASMIN, in Group Workspaces (GWS) and archive

# JASMIN as a Data Commons



CEDA CREPP system to ingest from MetOffice Hadley Centre

## CMIP6



Climate Models  
Couple-Model Intercomparison Projects

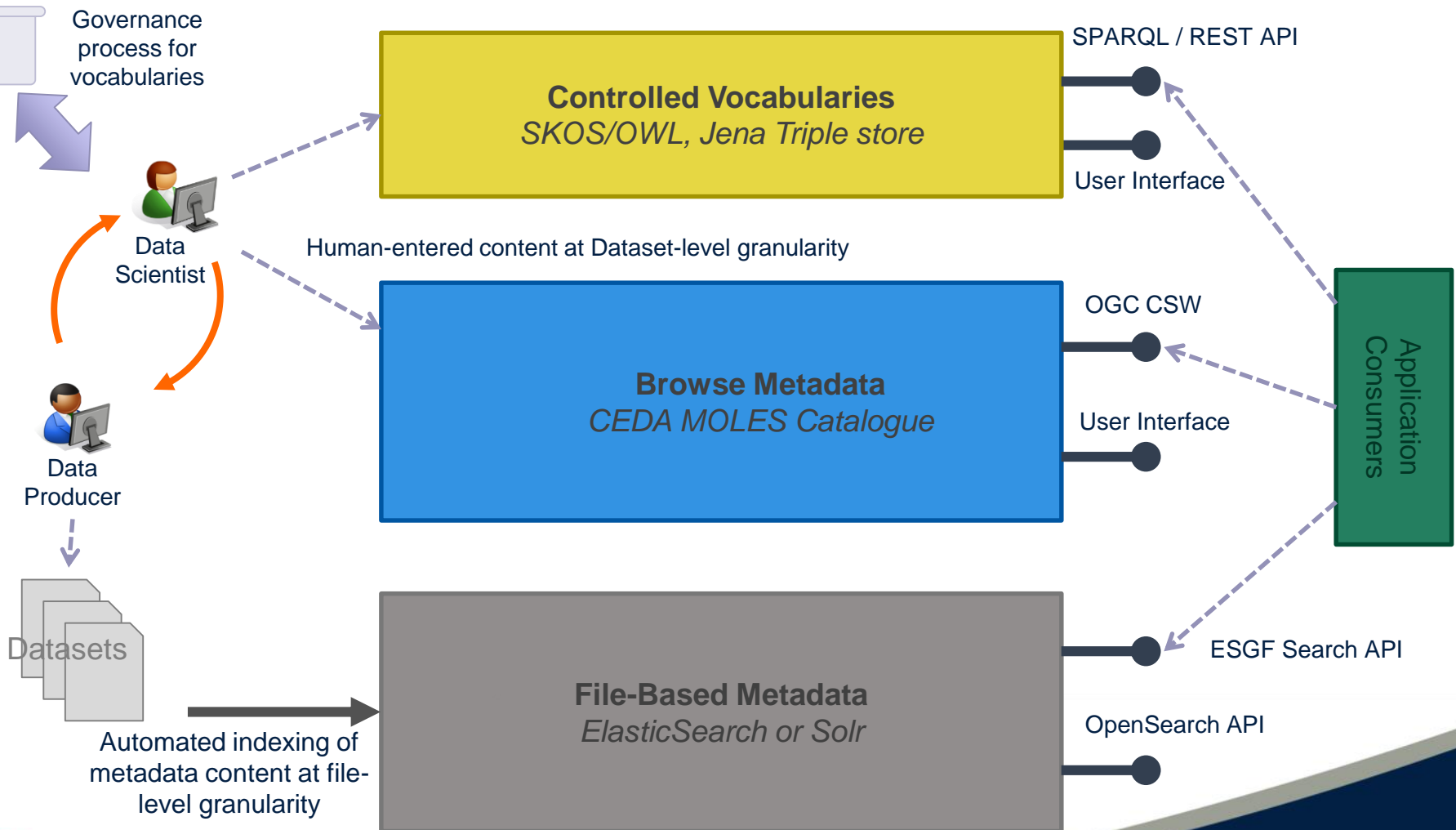
Sentinel Earth  
Observation Data



Sentinel missions data rate: ~6PB/year

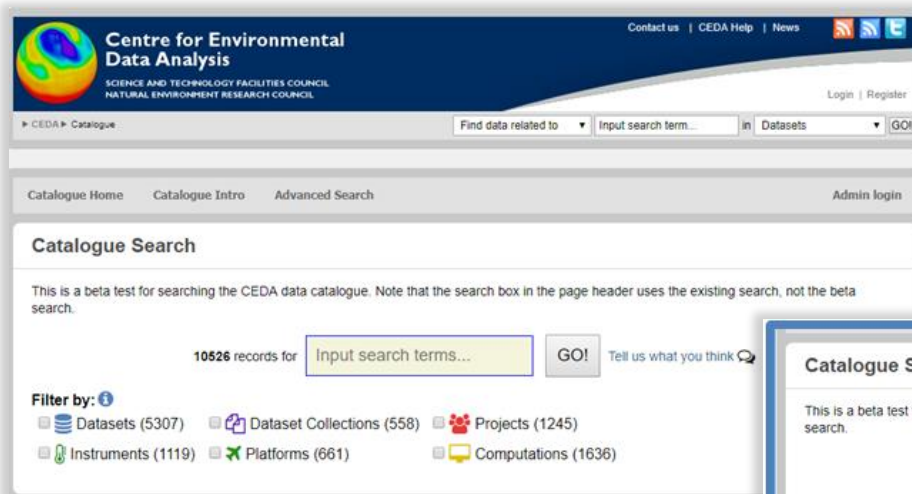
# Managed Data:

## Sourcing information for Data discovery





# MOLES – Catalogue Search



**Centre for Environmental Data Analysis**  
 SCIENCE AND TECHNOLOGY FACILITIES COUNCIL  
 NATURAL ENVIRONMENT RESEARCH COUNCIL

Find data related to  in **Datasets**

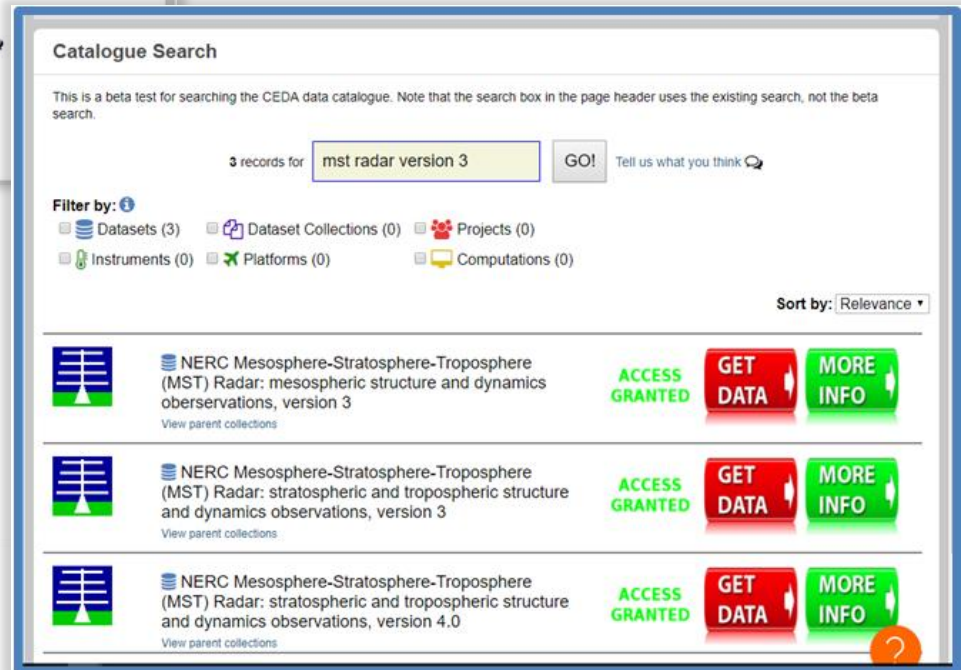
Catalogue Home | Catalogue Intro | Advanced Search | Admin login

### Catalogue Search

This is a beta test for searching the CEDA data catalogue. Note that the search box in the page header uses the existing search, not the beta search.

10526 records for   Tell us what you think

**Filter by:**  
 Datasets (5307)  
  Dataset Collections (558)  
  Projects (1245)  
 Instruments (1119)  
  Platforms (661)  
  Computations (1636)






### Catalogue Search

This is a beta test for searching the CEDA data catalogue. Note that the search box in the page header uses the existing search, not the beta search.

3 records for   Tell us what you think

**Filter by:**  
 Datasets (3)  
  Dataset Collections (0)  
  Projects (0)  
 Instruments (0)  
  Platforms (0)  
  Computations (0)

Sort by: **Relevance**

	NERC Mesosphere-Stratosphere-Troposphere (MST) Radar: mesospheric structure and dynamics observations, version 3 <small>View parent collections</small>	ACCESS GRANTED	GET DATA	MORE INFO
	NERC Mesosphere-Stratosphere-Troposphere (MST) Radar: stratospheric and tropospheric structure and dynamics observations, version 3 <small>View parent collections</small>	ACCESS GRANTED	GET DATA	MORE INFO
	NERC Mesosphere-Stratosphere-Troposphere (MST) Radar: stratospheric and tropospheric structure and dynamics observations, version 4.0 <small>View parent collections</small>	ACCESS GRANTED	GET DATA	MORE INFO





flight-finder.ceda.ac.uk

League Kibana Intranet SBS SIMS Helpdesk League - old accounts-admin PRISM Clarity travel eelog J/M Dashboard Ganglia Acclim

**EUFAR Flight Finder**

Search for flights within the EUFAR archive using the parameters below. Refine your search with geographical, temporal, and text search terms. Flights from FAAM, NERC-ARSF, SAFIRE, AAI-Polars, K&E-Enduro, and INTA-CASA aircraft and the APEX instrument flown on the DLR aircraft are now included.

Help (Documentation)  
Tutorial Video

Choose an index:

Click an item to expand the panel.

Geographical Search

Temporal Filter

Refine Search via Time Range

2017-01-01 to 2018-02-01

1k  
100  
10

**EUFAR Flight Finder & CEDA Satellite Data Finder**  
Wendy Garland, Richard Smith and Ag Stephens  
NCAS/NERC, Centre for Environmental Data Analysis, RAL Space, STFC, UK

**Introduction**  
Searching for data collected on moving platforms such as research aircraft can be difficult – unless you were involved in the flight planning you are not likely to know where or when a flight took place, or what instruments were operated. The EUFAR Flight Finder (EFF) is a geospatial/temporal/keyword search tool developed during the EUFAR2+ FP7 project to maximise discovery, access and reuse of data collected by the many disparate environmental research aircraft – both atmospheric in-situ and remote-sensing measurements – during EUFAR funded flights whose data are stored in the EUFAR data archive at CEDA.

The scope of the EFF tool was widened to include the full catalogue of flights from the FAAM Bae-146 aircraft and the NERC-ARSF Research Facility (formerly ARSF), also stored in the CEDA archive, and now covers 1278 of date and nearly 2000 flights.

Following the success of the EFF, the concept was refined and reconfigured to display satellite data in the CEDA archive in the CEDA Satellite Data Finder (SDFinder), also stored in the CEDA archive, and now covers 1.2 million Sentinel 1, 2, 3 and Landsat 5, 7, 8 images.

**How the EFF and Satellite Finder works**  
 - Location information: flight tracks, name, aircraft, sensor information, along with parameter and raw search data to be extracted from the data files, for each flight scenario. This is a complex task that requires a detailed understanding of the flight scenarios and the data files to be searched. It is a complex task that requires a detailed understanding of the flight scenarios and the data files to be searched.

**What is EUFAR?**  
 The European Centre for Earth Observation Research in Environmental and Geosciences (EUFAR) is an acronym for the European Research Project of European Commission, which is a project under Horizon 2020, FP7, FP6, FP5, FP4, FP3, FP2, FP1, FP0, FP-1, FP-2, FP-3, FP-4, FP-5, FP-6, FP-7, FP-8, FP-9, FP-10, FP-11, FP-12, FP-13, FP-14, FP-15, FP-16, FP-17, FP-18, FP-19, FP-20, FP-21, FP-22, FP-23, FP-24, FP-25, FP-26, FP-27, FP-28, FP-29, FP-30, FP-31, FP-32, FP-33, FP-34, FP-35, FP-36, FP-37, FP-38, FP-39, FP-40, FP-41, FP-42, FP-43, FP-44, FP-45, FP-46, FP-47, FP-48, FP-49, FP-50, FP-51, FP-52, FP-53, FP-54, FP-55, FP-56, FP-57, FP-58, FP-59, FP-60, FP-61, FP-62, FP-63, FP-64, FP-65, FP-66, FP-67, FP-68, FP-69, FP-70, FP-71, FP-72, FP-73, FP-74, FP-75, FP-76, FP-77, FP-78, FP-79, FP-80, FP-81, FP-82, FP-83, FP-84, FP-85, FP-86, FP-87, FP-88, FP-89, FP-90, FP-91, FP-92, FP-93, FP-94, FP-95, FP-96, FP-97, FP-98, FP-99, FP-100, FP-101, FP-102, FP-103, FP-104, FP-105, FP-106, FP-107, FP-108, FP-109, FP-110, FP-111, FP-112, FP-113, FP-114, FP-115, FP-116, FP-117, FP-118, FP-119, FP-120, FP-121, FP-122, FP-123, FP-124, FP-125, FP-126, FP-127, FP-128, FP-129, FP-130, FP-131, FP-132, FP-133, FP-134, FP-135, FP-136, FP-137, FP-138, FP-139, FP-140, FP-141, FP-142, FP-143, FP-144, FP-145, FP-146, FP-147, FP-148, FP-149, FP-150, FP-151, FP-152, FP-153, FP-154, FP-155, FP-156, FP-157, FP-158, FP-159, FP-160, FP-161, FP-162, FP-163, FP-164, FP-165, FP-166, FP-167, FP-168, FP-169, FP-170, FP-171, FP-172, FP-173, FP-174, FP-175, FP-176, FP-177, FP-178, FP-179, FP-180, FP-181, FP-182, FP-183, FP-184, FP-185, FP-186, FP-187, FP-188, FP-189, FP-190, FP-191, FP-192, FP-193, FP-194, FP-195, FP-196, FP-197, FP-198, FP-199, FP-200, FP-201, FP-202, FP-203, FP-204, FP-205, FP-206, FP-207, FP-208, FP-209, FP-210, FP-211, FP-212, FP-213, FP-214, FP-215, FP-216, FP-217, FP-218, FP-219, FP-220, FP-221, FP-222, FP-223, FP-224, FP-225, FP-226, FP-227, FP-228, FP-229, FP-230, FP-231, FP-232, FP-233, FP-234, FP-235, FP-236, FP-237, FP-238, FP-239, FP-240, FP-241, FP-242, FP-243, FP-244, FP-245, FP-246, FP-247, FP-248, FP-249, FP-250, FP-251, FP-252, FP-253, FP-254, FP-255, FP-256, FP-257, FP-258, FP-259, FP-260, FP-261, FP-262, FP-263, FP-264, FP-265, FP-266, FP-267, FP-268, FP-269, FP-270, FP-271, FP-272, FP-273, FP-274, FP-275, FP-276, FP-277, FP-278, FP-279, FP-280, FP-281, FP-282, FP-283, FP-284, FP-285, FP-286, FP-287, FP-288, FP-289, FP-290, FP-291, FP-292, FP-293, FP-294, FP-295, FP-296, FP-297, FP-298, FP-299, FP-300, FP-301, FP-302, FP-303, FP-304, FP-305, FP-306, FP-307, FP-308, FP-309, FP-310, FP-311, FP-312, FP-313, FP-314, FP-315, FP-316, FP-317, FP-318, FP-319, FP-320, FP-321, FP-322, FP-323, FP-324, FP-325, FP-326, FP-327, FP-328, FP-329, FP-330, FP-331, FP-332, FP-333, FP-334, FP-335, FP-336, FP-337, FP-338, FP-339, FP-340, FP-341, FP-342, FP-343, FP-344, FP-345, FP-346, FP-347, FP-348, FP-349, FP-350, FP-351, FP-352, FP-353, FP-354, FP-355, FP-356, FP-357, FP-358, FP-359, FP-360, FP-361, FP-362, FP-363, FP-364, FP-365, FP-366, FP-367, FP-368, FP-369, FP-370, FP-371, FP-372, FP-373, FP-374, FP-375, FP-376, FP-377, FP-378, FP-379, FP-380, FP-381, FP-382, FP-383, FP-384, FP-385, FP-386, FP-387, FP-388, FP-389, FP-390, FP-391, FP-392, FP-393, FP-394, FP-395, FP-396, FP-397, FP-398, FP-399, FP-400, FP-401, FP-402, FP-403, FP-404, FP-405, FP-406, FP-407, FP-408, FP-409, FP-410, FP-411, FP-412, FP-413, FP-414, FP-415, FP-416, FP-417, FP-418, FP-419, FP-420, FP-421, FP-422, FP-423, FP-424, FP-425, FP-426, FP-427, FP-428, FP-429, FP-430, FP-431, FP-432, FP-433, FP-434, FP-435, FP-436, FP-437, FP-438, FP-439, FP-440, FP-441, FP-442, FP-443, FP-444, FP-445, FP-446, FP-447, FP-448, FP-449, FP-450, FP-451, FP-452, FP-453, FP-454, FP-455, FP-456, FP-457, FP-458, FP-459, FP-460, FP-461, FP-462, FP-463, FP-464, FP-465, FP-466, FP-467, FP-468, FP-469, FP-470, FP-471, FP-472, FP-473, FP-474, FP-475, FP-476, FP-477, FP-478, FP-479, FP-480, FP-481, FP-482, FP-483, FP-484, FP-485, FP-486, FP-487, FP-488, FP-489, FP-490, FP-491, FP-492, FP-493, FP-494, FP-495, FP-496, FP-497, FP-498, FP-499, FP-500, FP-501, FP-502, FP-503, FP-504, FP-505, FP-506, FP-507, FP-508, FP-509, FP-510, FP-511, FP-512, FP-513, FP-514, FP-515, FP-516, FP-517, FP-518, FP-519, FP-520, FP-521, FP-522, FP-523, FP-524, FP-525, FP-526, FP-527, FP-528, FP-529, FP-530, FP-531, FP-532, FP-533, FP-534, FP-535, FP-536, FP-537, FP-538, FP-539, FP-540, FP-541, FP-542, FP-543, FP-544, FP-545, FP-546, FP-547, FP-548, FP-549, FP-550, FP-551, FP-552, FP-553, FP-554, FP-555, FP-556, FP-557, FP-558, FP-559, FP-560, FP-561, FP-562, FP-563, FP-564, FP-565, FP-566, FP-567, FP-568, FP-569, FP-570, FP-571, FP-572, FP-573, FP-574, FP-575, FP-576, FP-577, FP-578, FP-579, FP-580, FP-581, FP-582, FP-583, FP-584, FP-585, FP-586, FP-587, FP-588, FP-589, FP-590, FP-591, FP-592, FP-593, FP-594, FP-595, FP-596, FP-597, FP-598, FP-599, FP-600, FP-601, FP-602, FP-603, FP-604, FP-605, FP-606, FP-607, FP-608, FP-609, FP-610, FP-611, FP-612, FP-613, FP-614, FP-615, FP-616, FP-617, FP-618, FP-619, FP-620, FP-621, FP-622, FP-623, FP-624, FP-625, FP-626, FP-627, FP-628, FP-629, FP-630, FP-631, FP-632, FP-633, FP-634, FP-635, FP-636, FP-637, FP-638, FP-639, FP-640, FP-641, FP-642, FP-643, FP-644, FP-645, FP-646, FP-647, FP-648, FP-649, FP-650, FP-651, FP-652, FP-653, FP-654, FP-655, FP-656, FP-657, FP-658, FP-659, FP-660, FP-661, FP-662, FP-663, FP-664, FP-665, FP-666, FP-667, FP-668, FP-669, FP-670, FP-671, FP-672, FP-673, FP-674, FP-675, FP-676, FP-677, FP-678, FP-679, FP-680, FP-681, FP-682, FP-683, FP-684, FP-685, FP-686, FP-687, FP-688, FP-689, FP-690, FP-691, FP-692, FP-693, FP-694, FP-695, FP-696, FP-697, FP-698, FP-699, FP-700, FP-701, FP-702, FP-703, FP-704, FP-705, FP-706, FP-707, FP-708, FP-709, FP-710, FP-711, FP-712, FP-713, FP-714, FP-715, FP-716, FP-717, FP-718, FP-719, FP-720, FP-721, FP-722, FP-723, FP-724, FP-725, FP-726, FP-727, FP-728, FP-729, FP-730, FP-731, FP-732, FP-733, FP-734, FP-735, FP-736, FP-737, FP-738, FP-739, FP-740, FP-741, FP-742, FP-743, FP-744, FP-745, FP-746, FP-747, FP-748, FP-749, FP-750, FP-751, FP-752, FP-753, FP-754, FP-755, FP-756, FP-757, FP-758, FP-759, FP-760, FP-761, FP-762, FP-763, FP-764, FP-765, FP-766, FP-767, FP-768, FP-769, FP-770, FP-771, FP-772, FP-773, FP-774, FP-775, FP-776, FP-777, FP-778, FP-779, FP-780, FP-781, FP-782, FP-783, FP-784, FP-785, FP-786, FP-787, FP-788, FP-789, FP-790, FP-791, FP-792, FP-793, FP-794, FP-795, FP-796, FP-797, FP-798, FP-799, FP-800, FP-801, FP-802, FP-803, FP-804, FP-805, FP-806, FP-807, FP-808, FP-809, FP-810, FP-811, FP-812, FP-813, FP-814, FP-815, FP-816, FP-817, FP-818, FP-819, FP-820, FP-821, FP-822, FP-823, FP-824, FP-825, FP-826, FP-827, FP-828, FP-829, FP-830, FP-831, FP-832, FP-833, FP-834, FP-835, FP-836, FP-837, FP-838, FP-839, FP-840, FP-841, FP-842, FP-843, FP-844, FP-845, FP-846, FP-847, FP-848, FP-849, FP-850, FP-851, FP-852, FP-853, FP-854, FP-855, FP-856, FP-857, FP-858, FP-859, FP-860, FP-861, FP-862, FP-863, FP-864, FP-865, FP-866, FP-867, FP-868, FP-869, FP-870, FP-871, FP-872, FP-873, FP-874, FP-875, FP-876, FP-877, FP-878, FP-879, FP-880, FP-881, FP-882, FP-883, FP-884, FP-885, FP-886, FP-887, FP-888, FP-889, FP-890, FP-891, FP-892, FP-893, FP-894, FP-895, FP-896, FP-897, FP-898, FP-899, FP-900, FP-901, FP-902, FP-903, FP-904, FP-905, FP-906, FP-907, FP-908, FP-909, FP-910, FP-911, FP-912, FP-913, FP-914, FP-915, FP-916, FP-917, FP-918, FP-919, FP-920, FP-921, FP-922, FP-923, FP-924, FP-925, FP-926, FP-927, FP-928, FP-929, FP-930, FP-931, FP-932, FP-933, FP-934, FP-935, FP-936, FP-937, FP-938, FP-939, FP-940, FP-941, FP-942, FP-943, FP-944, FP-945, FP-946, FP-947, FP-948, FP-949, FP-950, FP-951, FP-952, FP-953, FP-954, FP-955, FP-956, FP-957, FP-958, FP-959, FP-960, FP-961, FP-962, FP-963, FP-964, FP-965, FP-966, FP-967, FP-968, FP-969, FP-970, FP-971, FP-972, FP-973, FP-974, FP-975, FP-976, FP-977, FP-978, FP-979, FP-980, FP-981, FP-982, FP-983, FP-984, FP-985, FP-986, FP-987, FP-988, FP-989, FP-990, FP-991, FP-992, FP-993, FP-994, FP-995, FP-996, FP-997, FP-998, FP-999, FP-1000.

**Conclusions and future plans**  
 The EUFAR Flight Finder and CEDA Satellite Finder tools are now available for use and make it easier to discover and access data from 2000 flights and 2 million satellite scenes.  
 The development of the EFF and similar tools highlights the need for well-organised, comprehensive, consistent datasets to maximise the use and re-use of the data.  
 Future plans under consideration include:  
 - Extending the EFF to display regular (non-EUFAR) flight paths of the wider EUFAR fleet into operational data.  
 - Including ground-based NCAS observations.  
 - Adding further satellite missions as they are launched by ESA and adding AATOS scenes.  
 - Including the search capabilities, metadata, data processing, data management, and data access capabilities.  
 - Finding a way to integrate satellite and flight data to validation studies.

http://flight-finder.ceda.ac.uk  
http://geo-search.ceda.ac.uk  
Contact: Wendy Garland  
wendy.garland@ncas.ac.uk

Only the first 1000 results will be plotted on the map with most recent on top.

Help Tutorial Feedback

Click an item to expand the panel.

Temporal Filter

Change Map Centre

Rectangle Search

Satellite Filter

Apply Filters Clear Filters

Export Results

814 hits with current selection.  
123 milliseconds for response.  
ouse: Lat: 77.43, Lng: -18.83

Map Satellite

Rectangle search: On

Key:  
 - Sentinel 1  
 - Sentinel 2  
 - Sentinel 3  
 - Landsat  
 - Other

# Logical View



## CEDA Archive Services

Data Centres, Curation, DB systems  
User management, External Helpdesk

## Analysis Environment

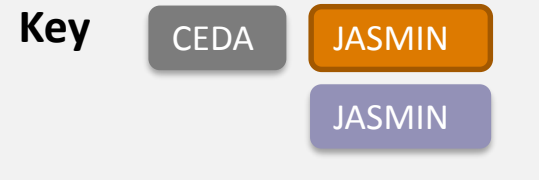
Compute:  
Science VMs + User Management, PaaS and IaaS Cloud,  
Group Workspaces: Fast Disk & Elastic Tape  
External Helpdesk



## JASMIN Compute and Storage

Managed Compute (Lotus and GWSs, Tape Store + Data Transfer Zone),  
Community Cloud + Internal Helpdesk

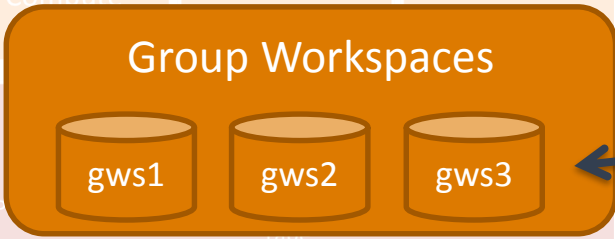




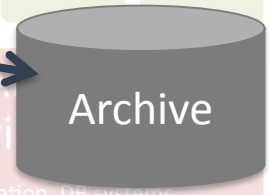
**Long Term Archive Storage**



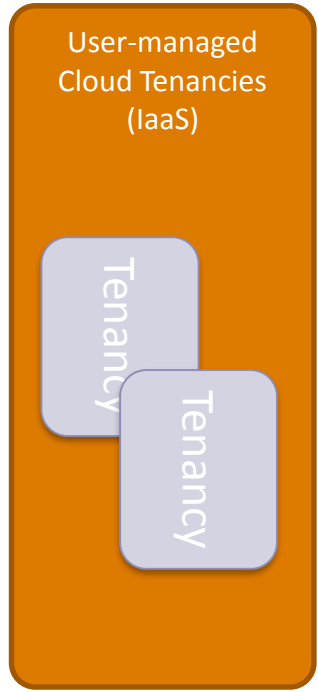
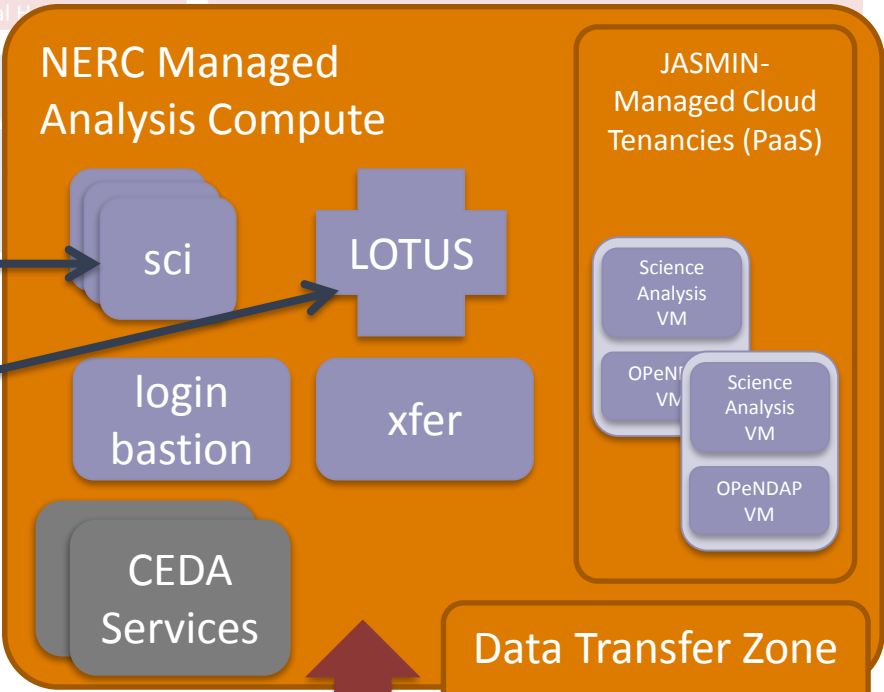
**Short-Term Project Storage**



Phase 4: 17PB



Phase 4: 26PB



Phase 4: 11500 cores – all types of compute

**Interactive Compute**

**Batch Compute**



**Functional View**

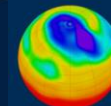






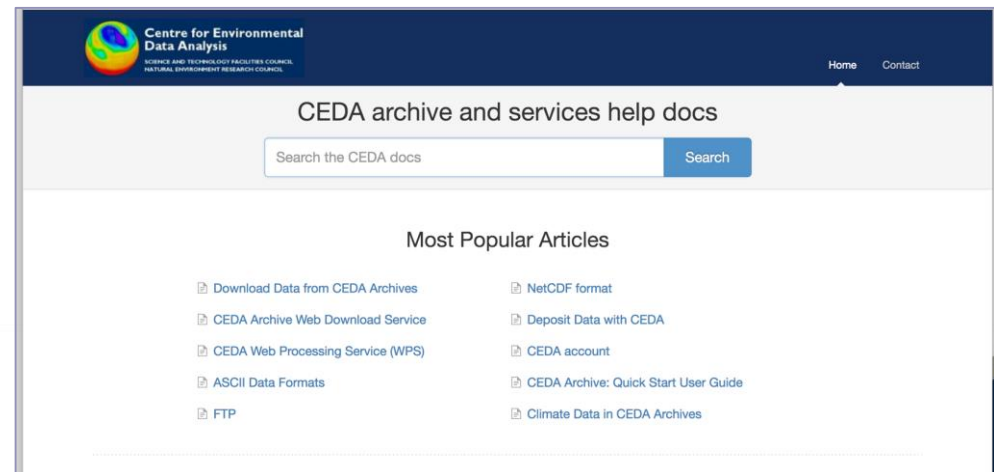
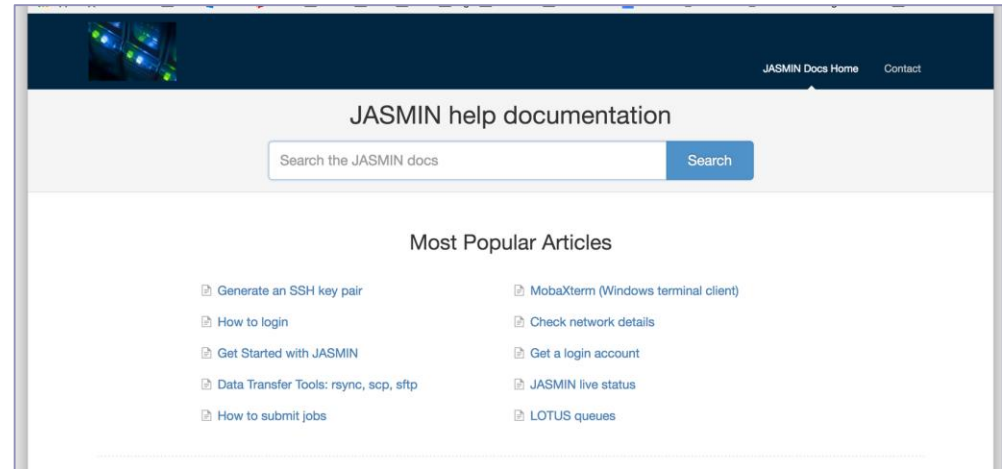
# Main service options for JASMIN users

- Direct POSIX-based access (majority of JASMIN use)
  - Group workspaces
  - Lotus batch compute
  - Direct access to the archive
- Cloud
  - Platform-as-a-Service:
    - tenants deploy pre-defined VM templates
    - VMs have direct POSIX access to group workspace and data archive
    - VMs are patched by the JASMIN team
  - Infrastructure-as-a-Service:
    - tenants deploy pre-defined VM templates
    - But have full root access to VMs
    - Can allocate their own storage
    - VMs managed by the tenancy owner
    - HTTP and FTP access only to GWS and data archive
- *Object storage and dual S3/POSIX interface storage being deployed this year for JASMIN Phase 4*

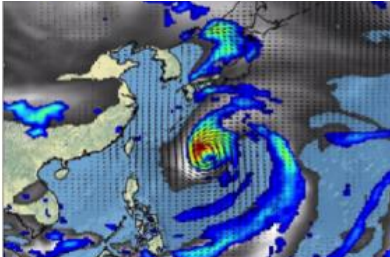


# User support

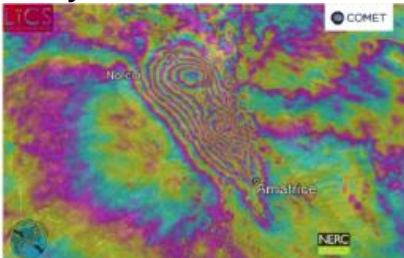
- Helpdesk service for CEDA and JASMIN users
  - Approx. 50 queries/week
  - Initial response within 1 working day
- Documentation – continuously evolving
- Data provider support via team of data scientists
  - Formats
  - Metadata
  - Data upload
  - Data policy



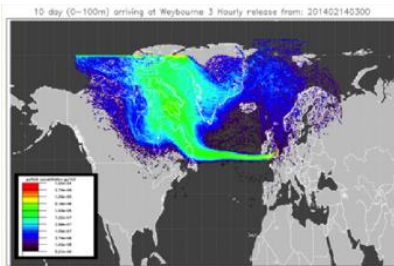
# ~150 Science projects on JASMIN to date



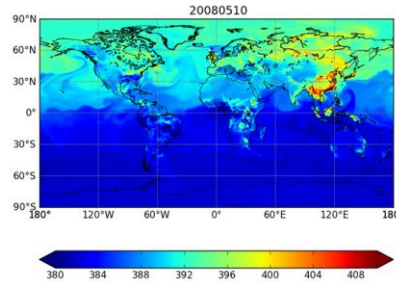
High Res Climate Model analysis



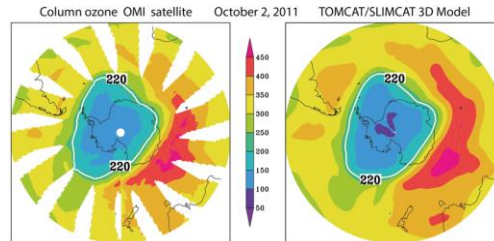
Fault analysis



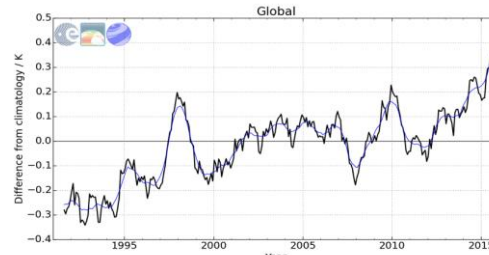
Atmospheric dispersion



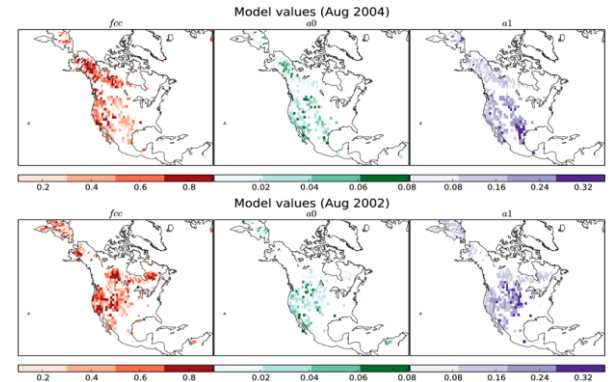
Regional carbon balance on a global scale



Antarctic Ozone hole: model vs. observations

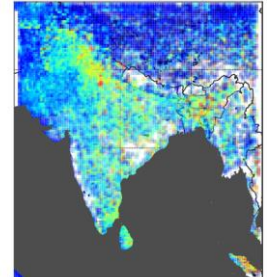


Sea Surface Temperature from satellite observations

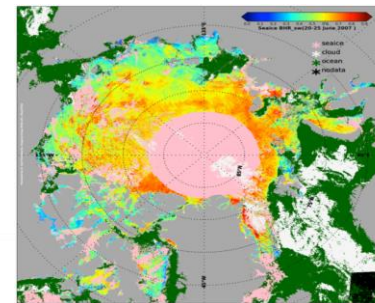


Deriving the impact of fire on vegetation from earth observation data

OMI Vertical Columns monthly average 0.5x0.5 grid



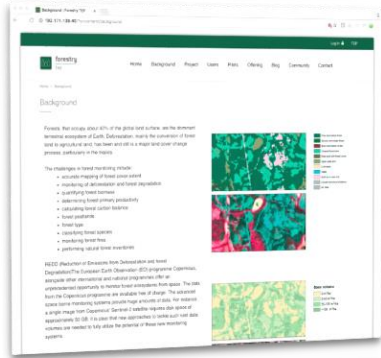
Understanding oxidant chemistry over the Indian subcontinent



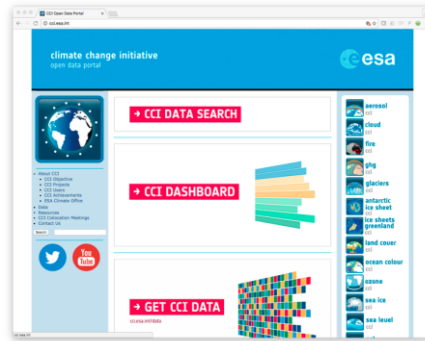
Climate variables from European and US instruments/satellites



# JASMIN usage: Cloud



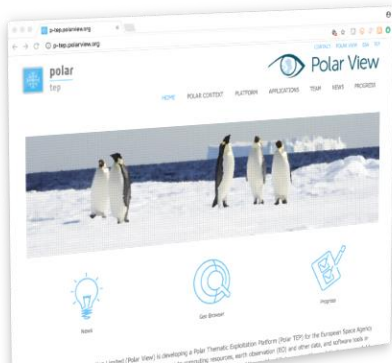
ESA Forestry Thematic Exploitation Platform



ESA Climate Change Initiative Open Data Portal



Majic interface to Jules Land-surface model on JASMIN

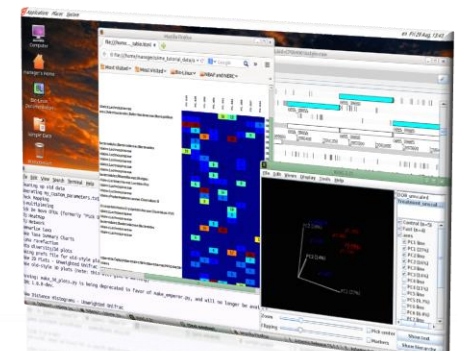


ESA Polar Thematic Exploitation Platform



Attendees at ESA Summer school, ESRIN used OPTIRAD Jupyter Notebook environment

– Credit ESA



EOS Cloud – Desktop-as-a-Service for Environmental Genomics

# Funding & Governance

- Funding
  - Capital
    - "Periodic" capital investment
      - Phase 1, 2, 3 : 2011 - 2015 (~£12M total)
      - Phase 4 currently underway (£8M)
  - Recurrent
    - £650K/year via NERC SLA
  - New governance arrangements April 2018 as part of NERC's National Capability commissioning / funding route
    - NERC, NCAS & NCEO have primary role in approving JASMIN strategy





# Issues & Gaps

- Data intensive computing facility
  - Storage is a major cost
  - Phase 4 introduces more cost effective storage technologies for future scaling: object storage and scale-out file system
- Balancing "bleeding edge" technology with demands of operational service
  - Research infrastructure, not 99.999% available
  - Small team (capital-heavy)
- Diverse user community
  - Wide range of workflow requirements
  - Wide range of user skill level
  - Training, user education, engagement





# Further Information

- CEDA and JASMIN:
  - <http://www.jasmin.ac.uk/>
  - <http://www.ceda.ac.uk/>
- JASMIN paper  
Lawrence, B.N. , V.L. Bennett, J. Churchill, M. Jukes, P. Kershaw, S. Pascoe, S. Pepler, M. Pritchard, and A. Stephens. **Storing and manipulating environmental big data with JASMIN.** *Proceedings of IEEE Big Data 2013*, p68-75, [doi:10.1109/BigData.2013.6691556](https://doi.org/10.1109/BigData.2013.6691556)