



herschel



Status Report of, and plans for Herschel Data Processing/HIPE

Stephan Ott

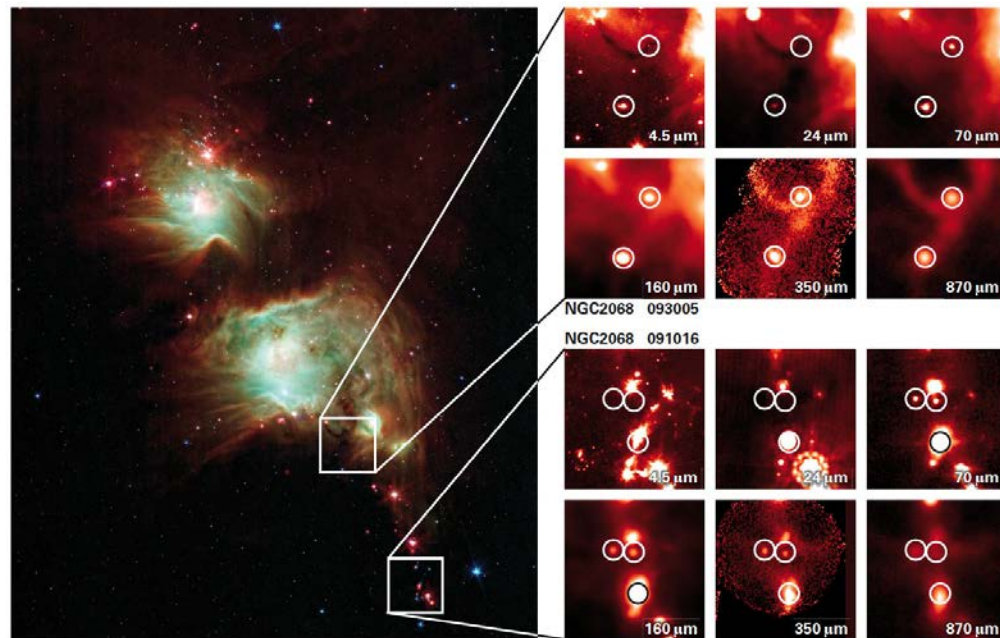
Herschel Science Data Processing Development Manager

Herschel Science Data Processing Coordinator

with many thanks to all colleagues who contributed and provided inputs

Outline

- **Data Processing Overview**
 - Events
 - Highlights of HCSS 10
 - Other Activities
 - HCSS release schedule
- **Post-Operations review**
- **Future activities and priorities for HCSS 11**
- **Project Development Statistics**
- **Payback time!**
- **On-demand Profiles**
- **Even more Questions and Wishes to the DPUG**



Herschel enables discovery of young protostars in Orion. Herschel/PACS imaging of an area in Orion, previously imaged by Spitzer, has led to a serendipitous discovery of young protostars ([Stutz et al. 2013](#))

Data Processing Overview – Event list

- **9/10/2012 – 11/10/2012** **HIPE Forum 2012 (ESAC)**
<http://herschel.esac.esa.int/twiki/bin/view/Hcss/HipeForum2012>
72 persons attended
- **17/10/2012 – 22/01/2013** **HCSS 10.0 branch-off , stabilisation, validation and operational deployment**
- **22/11/2012 – 27/2/2013** **Bulk reprocessing of Herschel observations with HCSS 9.1**
- **14/02/2013** **Branch-off point for HCSS 10.1. HCSS 10 bulk reprocessing will be performed with this version**
- **28/01/2013 – 31/01/2013** **Herschel PACS and SPIRE map-making workshop (ESAC)**
<http://herschel.esac.esa.int/2013Mapmaking.shtml>
57 persons attended

Data Processing Overview – Event list

- **25/03/2013 – 27/03/2013** **Herschel Calibration workshop (ESAC)**
<http://herschel.esac.esa.int/CalibrationWorkshop5.shtml>
63 persons attended
- **11/04/2013** **HCSS 11.0 branch-off point**
- **28/05/2013 – 30/05/2013** **Lessons Learned “Operations from cradle to grave” (ESAC)**
- **24/06/2013 – 27/06/2013** **Herschel Data Processing Workshop for Newcomers (ESAC)**
- **Summer 2013** **Herschel Data Processing Workshop for Archive users (NHSC)**
- **Autumn 2013** **HIPE Forum 2013: Expanding the Herschel community (ESAC)**
- **2013** **Lessons Learned “Full scope”**

Data Processing Overview – Events

Herschel PACS and SPIRE map-making workshop

<http://herschel.esac.esa.int/2013Mapmaking.shtml>

- **SPIRE will release a report (likely final) in ~June.**
- **PACS will release a preliminary report on a similar timescale (~May) and a final one in ~1 year. Currently PACS plans to implement an error map within MADMap and JScanam for HIPE 11. In HIPE 12 a new baseline drift corrector (MADMap pre-processing) will be implemented. I expect that the merits and avenues for future improvements of PhotProject, MADMap, Unimap and Scanamorphous will be discussed in detail during a PACS photometer meetings preceding the next PACS ICC meeting (23rd of April 2013)**



Data Processing Overview – Events

Bulk reprocessing of Herschel observations with HCSS 9.1

- **Level 2 generation started November 22nd 2012 and was completed 27th of January 2013**
- **All 50319 observations reached the expected processing level**
- **Best sustained system throughput around 40 Operational Days/day**
- **Level 2.5 generation started February 6th and finished February 27th**
 - **99.99% of 10516 PACS maps were processed successfully in 12 days**
 - **90.3% of 670 PACS spectra were processed successfully in 1 hour**
 - **All 1229 SPIRE maps were processed successfully in 12 hours**

An overall throughput of 13 Operational Days/day was reached



Data Processing Overview – Events

- **Herschel Calibration workshop “Only the best data products for the Legacy Archive”**

<http://herschel.esac.esa.int/CalibrationWorkshop5.shtml>

**Herschel reached a very impressive calibration accuracy already!
See Tony’s presentation for more information.**

- **HIPE Forum 2013 “Expanding the Herschel community”**
 - **quality of Herschel pipeline products, including quality of quality control information**
 - **additional high level products: Point source lists, spectra, SEDs – multi observation, multi instrument**
 - **accessibility of Herschel pipeline products outside of HIPE**
 - **documentation (quality, organisation, accessibility)**
 - **increasing the user-friendliness with reduced manpower**
 - **expanding the depth of testing and code quality with reduced manpower**

Data Processing Overview – personnel news

- **12/11/2012** **Jose Maria Castro Ceron was recruited as DP TA until December 2013**
- **1/1/2013** **The DP TAs Asier Abreu and Eugenio Salguero left Herschel to work full time for GAIA**
- **15/02/2013** **Davide Rizzo, the DP documentalist left Herschel. David Shupe (NHSC) and Mark Kidger (HSC) will bridge the gap until a replacement is recruited**
- **01/04/2013 – 31/05/2013** **Paul Balm (developer) will work only 50% for Herschel, and afterwards 100% for Gaia**
- **01/05/2013 –** **Javier Diaz (developer) will work only 50% for Herschel**

This leads to a 50% reduction of current HSC/DP development resources

Other main activities in reporting period

- **Wrap-up telecons with Key Programmes.** One goal of this meeting is to identify possible bottlenecks in terms of data processing, calibration, documentation and science analysis
 - SPIRE photometry gives very little problems
 - SPIRE spectroscopy, HIFI and PACS photometry is more tricky
 - PACS spectroscopy gives the most issues
- **Topical Webinars** (What's new in HIPE 9.0/9.1, Faster and simpler data access with myHSA, SPIRE mapping and photometry (*2), HIFI data reduction (*2), SPIRE spectroscopy (*2), PACS mapping and photometry, PACS spectroscopy (*2), What's new in HIPE 10, PACS map-making follow-up)
- **Major new issues identified during those telecons / webinars are raised and tracked as HCSS tickets**
- **Negotiations with NAOC to secure 12 man-years of effort to ESAC/HSC (7 man years) and SPIRE (5 man years) during Herschel Post-ops**

Future HCSS Releases

HCSS version	Branch-off Point	Installation	Rationale
11.0	11 April 2013	June 2013	Full staff complement still available
12.0	26 Sept. 2013	December 2013	Last bi-annual release
13.0	18 Sept. 2014	December 2014	Annual release
14.0	17 Sept. 2015	December 2015	Annual release
14.1	February 2016	March 2016	End of HIFI POPs
14.2	June 2016	July 2016	End of SPIRE POPs
15.0	15 Sept. 2016	December 2016	End of PACS POPs; Legacy version for Herschel Science Archive

**Bulk reprocessing of all Herschel observations
will be performed with each of these versions**

Selection of Highlights of HCSS 10

➤ **HIPE**

- Optimised the registration and initialisation of tasks. This can reduce HIPE start-up time by up to 20%
- Displaying images uses less memory and is much faster
- Data fitting is now more robust

➤ **HIFI**

- Introduction of a pipeline task to allow comparison of emission in DBS chop positions. It differences the chop positions in all DBS observations and after applying the band-pass correction stores them in the ReferenceSpectra product
- Major updates to flagTool

Selection of Highlights of HCSS 10

➤ **PACS**

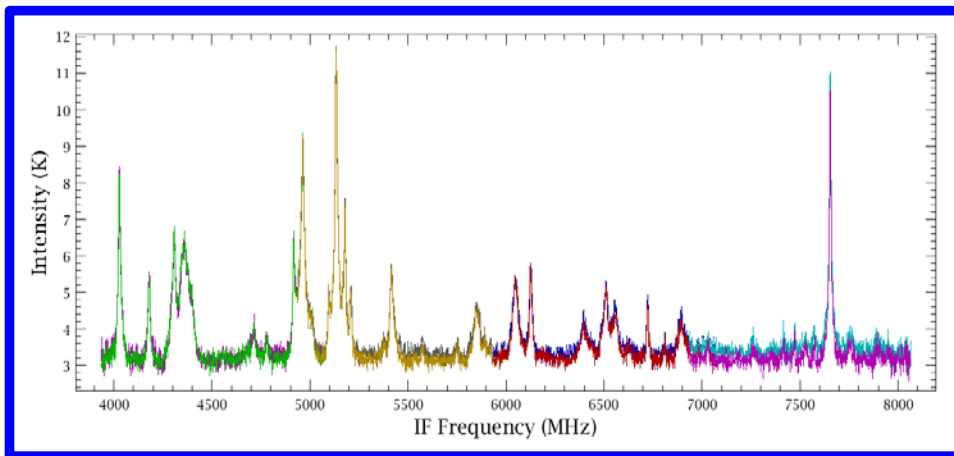
- New task available to derive error maps for maps produced with photProject task on highpass filtered timelines
- Level 2 products tunes for point source observations
- Optimising the processing the memory needed to process PACS scan maps could be halved
- Spectral drizzling is now offered for the line-scan observations

➤ **SPIRE**

- New zeroPointCorrection task for cross-calibration with Planck HFI maps producing extended emission maps
- Automatic correction to SSO frame for moving objects
- Improved photometry consistency across all Source Extraction and Photometry tasks
- Spectrometer semi-extended sources correction tool (SECT)



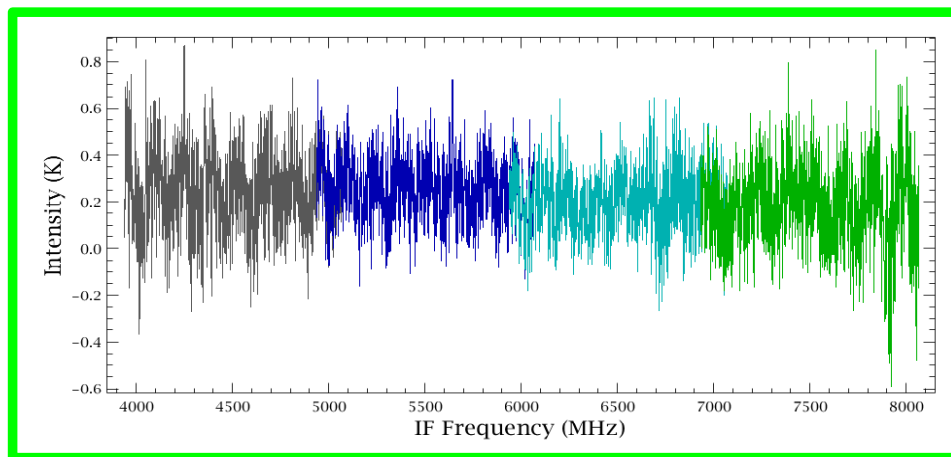
Improvements in HCSS 10 for HIFI products



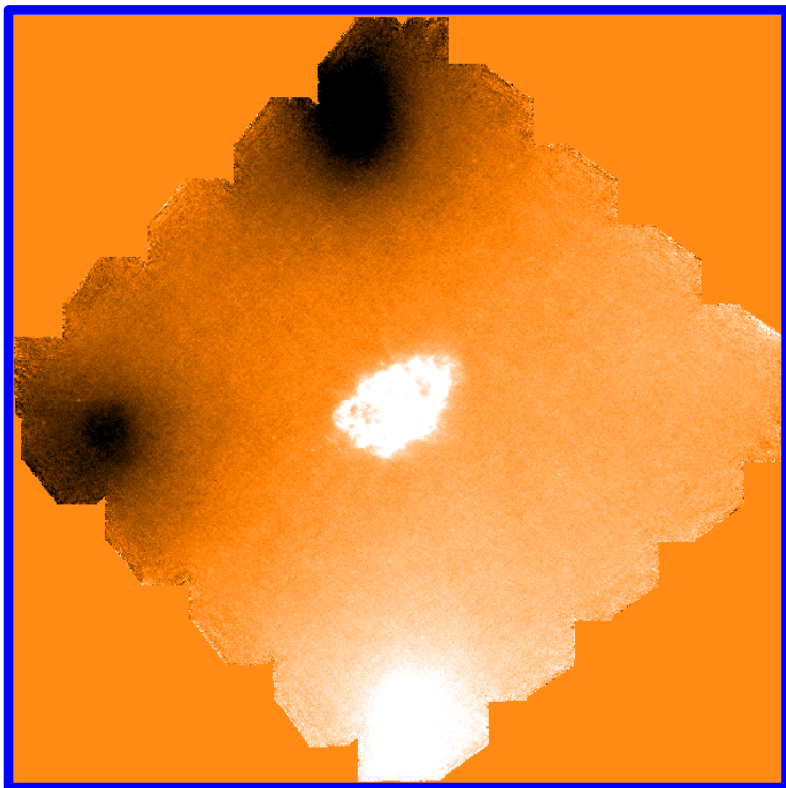
HIFI Level 1 and Level 2 SPG products did *not* change from HCSS 9 to HCSS 10

The spectrum OFF-target (always subtracted to the ON-target) is now stored at the calibration context

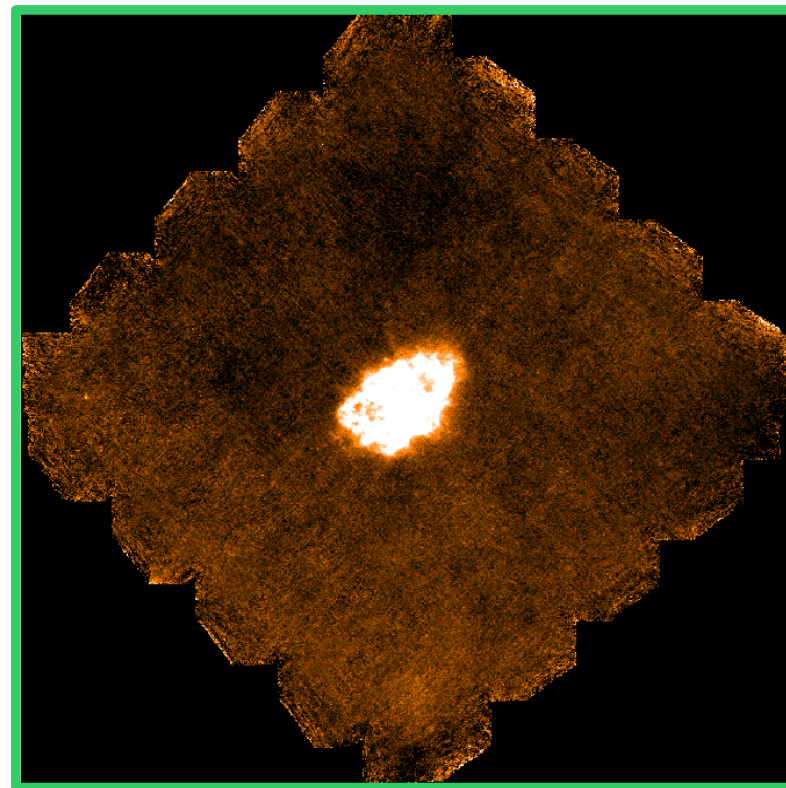
HIFI SPG Level 1 product
(ON-OFF)



Improvements in HCSS 10 for PACS products



HCSS 9 Level2.5 MadMap
SPG product

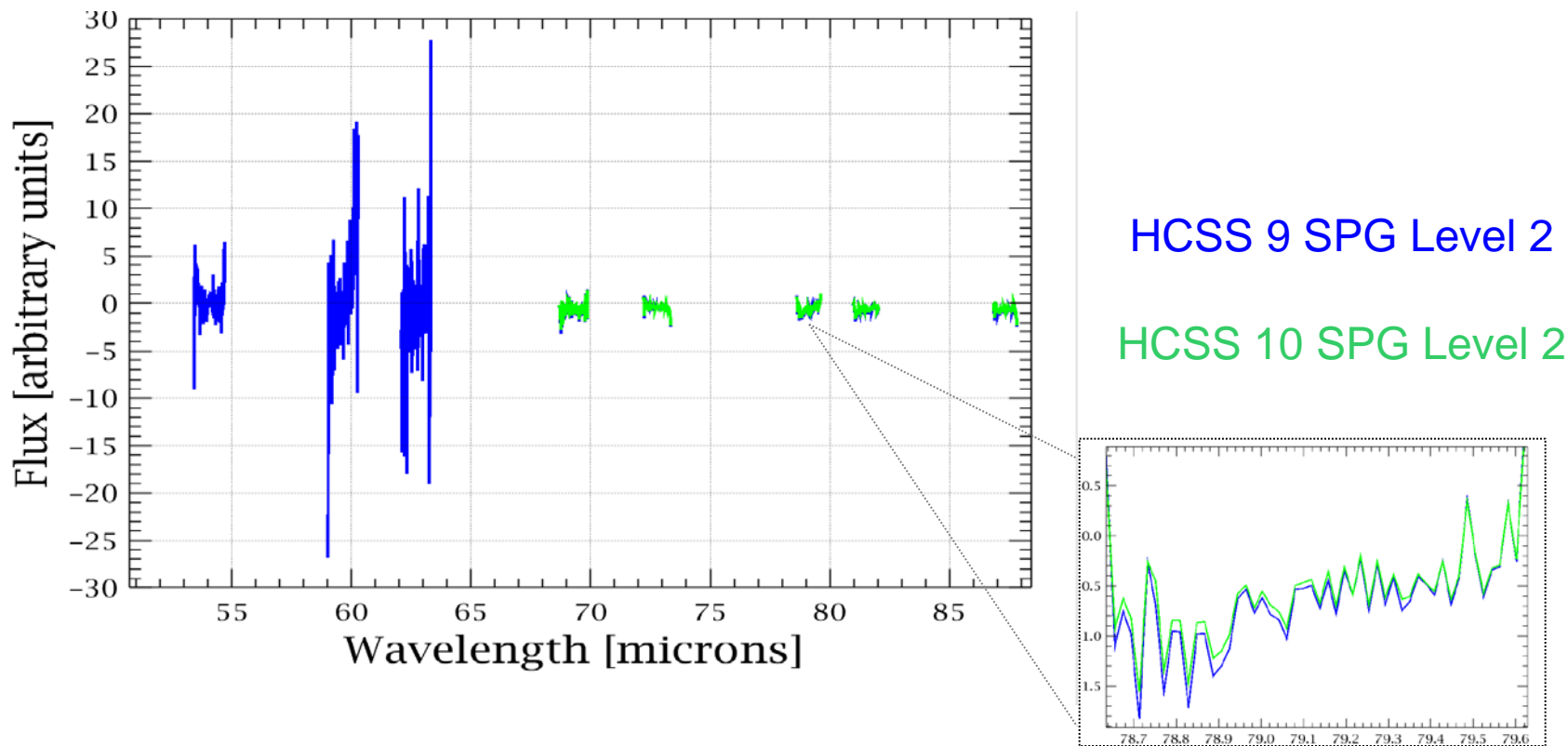


HCSS10 Level2.5 MadMap
SPG product

Drift Correction improved from HCSS 9.0 to HCSS 10. Main work done in HCSS 9.1

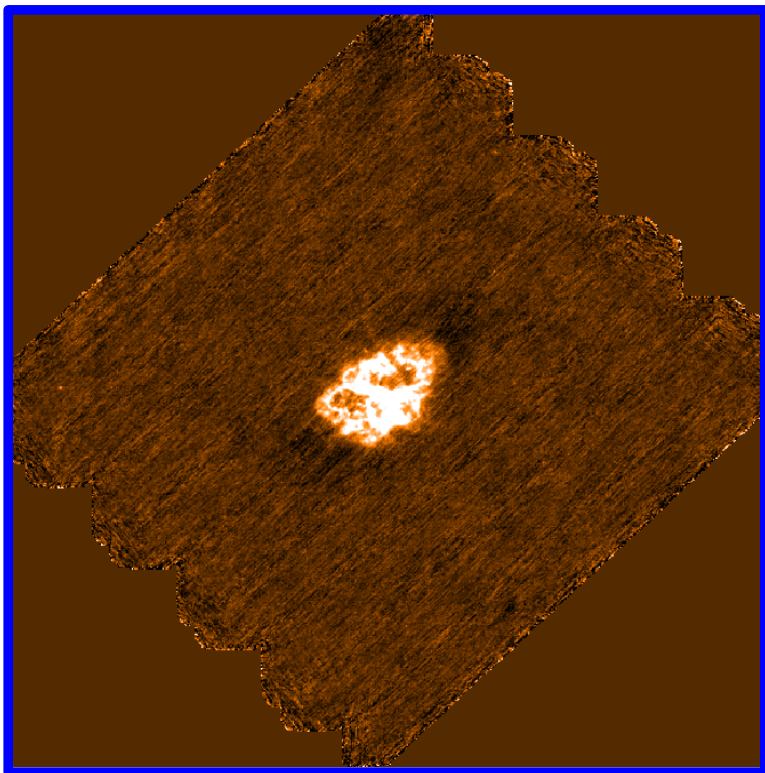


Improvements in HCSS 10 for PACS products

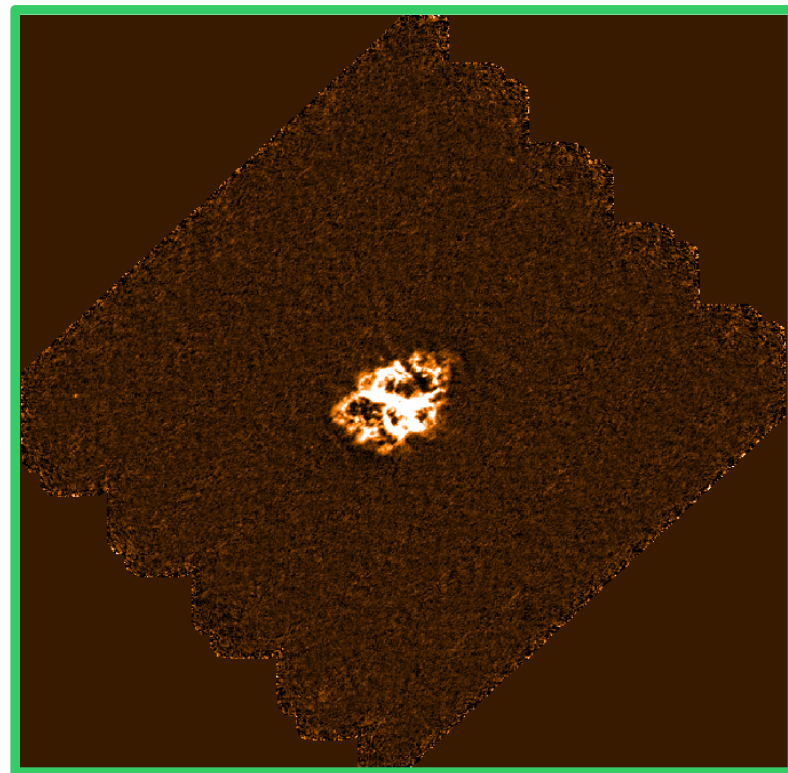


PACS spectrometer data flagged as “Out of band” in the blue range is not provided anymore at Level 1 or Level 2 PACS SPG products

Improvements in HCSS 10 for PACS products



HCSS 9 Level2 SPG product



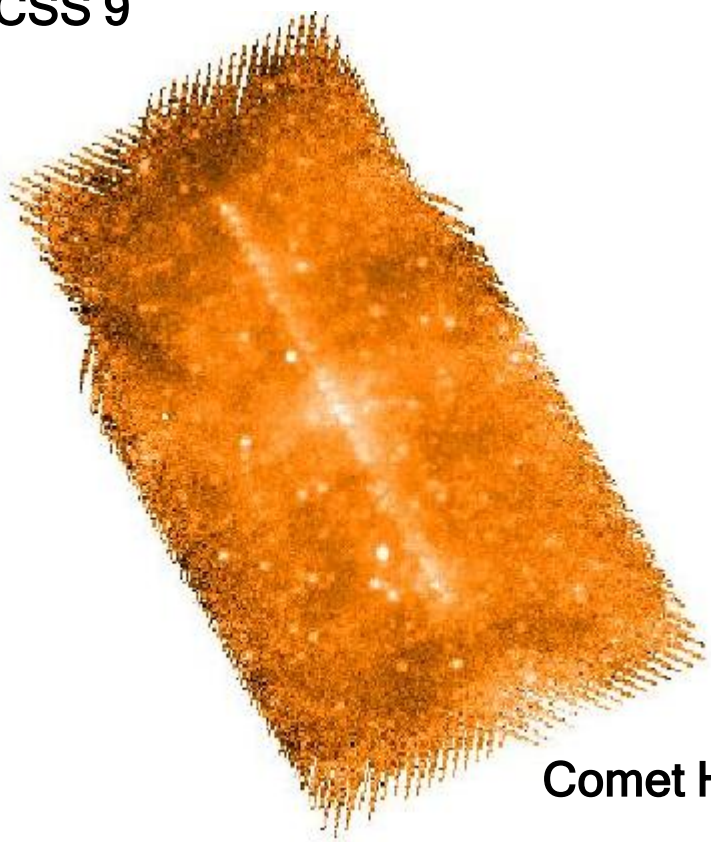
HCSS10 Level2 SPG product

Impact on L2 SPG products for extended sources (smaller highpassfilter)
It is recommended to use L2.5 products for the analysis of extended objects

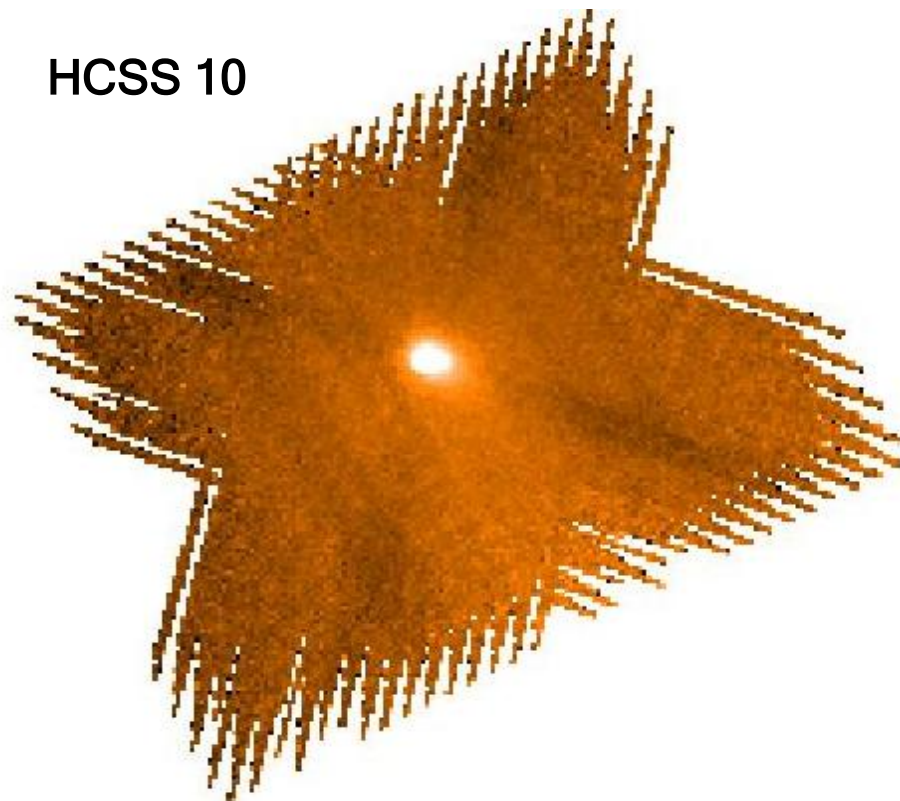
Improvements in HCSS 10 for SPIRE products

Photometer: Automatic correction of the timeline to the reference frame of moving objects

HCSS 9



HCSS 10



Comet Honda-Mrkos-Pajdušáková KPGT_pharto01_1

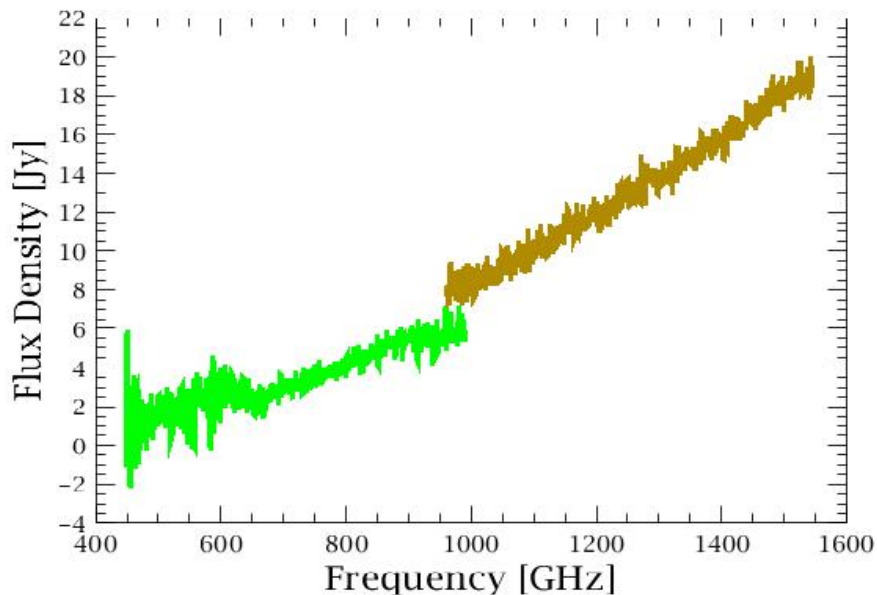
Improvements in HCSS 10 for SPIRE products

Spectrometer: Continuous improvement in calibration

HCSS 9

OD 189, 4 Vesta, HR, 50 reps

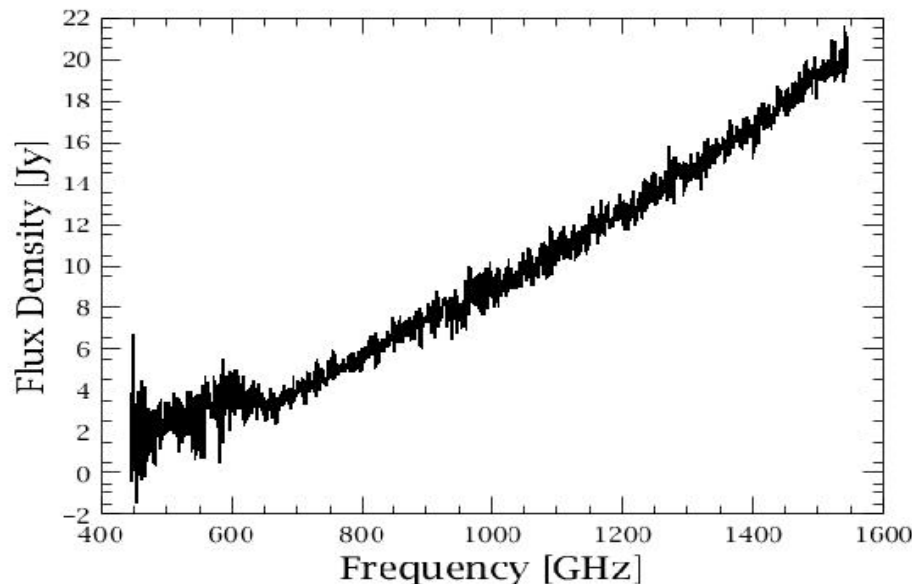
0x5000264C = 1342187084



HCSS 10

OD 189, 4 Vesta, HR, 50 reps

0x5000264C = 1342187084



Inputs to Post-Ops review: Post-ops plans for Data Processing @ HSC (Development priorities)

- **Most important are the archive products that come from standard pipeline processing of the science and calibration observations that were made throughout the mission lifetime. Therefore the support to the pipeline processing environment and the generation and quality control of products has the highest priority**
- **The second highest priority is the maintenance of the interactive data processing system so that the astronomical community has the best means to exploit Herschel data in detail, and the experts from the HSC, ICCs and NHSC can continue to improve algorithms and calibration**

Inputs to Post-Ops review: Post-ops plans for Data Processing @ HSC (Development priorities)

- **Documentation is the third highest priority as the analysis software and data products need to be well described so that the best science results can be extracted**
- **Maintenance of data visualisation and reduction tools, and any changes of basic infrastructure not needed for the SPG or HIPE framework are done on a best efforts basis**
- **All very desirable and important work packages like support to ICCs on s/w matters (profiling, memory leaks, technical advise), remote access to HIPE for European Astronomers, re-use of HIPE for other missions, inclusion of community provided software and ideas into HIPE can not be addressed**

Inputs to Post-Ops review: Post-ops plans for Data Processing @ HSC (Further steps taken)

- **Frontloading of manpower profile to minimise task swapping and excessive reprioritisation discussions as well as possible**
- **Herschel data processing activities end in 2016 (in line with ICCs, one year earlier than other HSC activities)**

Inputs to Post-Ops review: Post-ops plans for Data Processing @ HSC (Further steps foreseen)

- **ICCs were actioned to confirm that their postops WPs will not imply major work for HSC/changes to the DP infrastructure (SPG and HIPE) and tools**
- **HSC will not take over IA core packages from ICCs/NHSC if the developer continues to work for Herschel**
- **When an ICC/NHSC developer leaves the project HSC will become owner of his IA core package. Such packages will be minimally maintained – basically fixes of crash/critical SPRs only**
- **Instrument packages can't be, and are not foreseen to be maintained by the HSC DP team**



Post-Operations readiness review: Board recommendations

- **Review board emphasised importance of products as lasting legacy of Herschel. All ICC managers confirmed this as in line with ICC priorities**
- **Review board emphasised importance of prioritisation as we are resource limited and key personnel might leave the project earlier than envisaged**
- **Review board urged to develop the Herschel Legacy in conjunction with the emergence of other major facilities (ALMA, SOFIA)**
- **Review board asked to raise the awareness of the community to the resource limited character what might be achieved in post-operations**



Herschel post-operations readiness review

Manpower (staff years) available in post-operations

	DP developers	Calibration	Documentation	end
HIFI	11	29	3	31/03/2016
DP @ HSC	7.0		3.9	28/02/2017 *
CS @ HSC			26	31/12/2017
NSHC	14.2	5	7.8	30/09/2017
PACS	17.5	17.5	15	31/12/2016
SPIRE	29.5	1	3	31/07/2016

Handwritten annotations on the table:

- Red circle around 7.0 with arrow pointing to "Think small"
- Red circle around 17.5 with arrow pointing to "Think big"
- Red circle around 15 with arrow pointing to "Think big"
- Red double-headed arrow between 14.2 and 5 with text "Best information"
- Red double-headed arrow between 17.5 and 1 with text "Best information"

* Completion of bulk reprocessing, last development activities end mid 2016

Data Processing Overview: Priorities for HCSS 11 (Framework & system)

- 1. Improvements of calibration, pointing reconstruction and data reduction pipelines**
 - **Startracker pointing derivation**
 - **Restructuring of ACMS to support improved pointing reconstruction**
 - **Adding the Solar Aspect Angle to the pointing product**
- 2. Migration to Java 7**
- 3. Support for MAC OS X 10.8 (Mountain Lion)**
- 4. Framework to generate level 2.5 and level 3 products**
- 5. Reduction of start-up time for HIPE**
- 6. Improvement of documentation search**

Data Processing Overview: Priorities for HCSS 11 (Framework & system)

- 7. Improvements to ASCII tables reading and writing**
- 8. Batch mode for HIPE (instead of jylaunch)**
- 9. Removal of ia_toolbox_spectrum_gui (old spectrum explorer)**
- 10. Replacement of HSA perspective by data access perspective**
- 11. Improvements in user-friendliness (GUI layout, tooltips, harmonisation of Python syntax in HIPE)**
- 12. Pagination of HSA query results to improve response time for users (avoiding a long response time for HIPE)**
- 13. Improvements to task registration and management**



Data Processing Overview: Priorities for HCSS 11 (Framework & system)

- 14. Editing functionality of table and array datasets,
including insertion of changed product into observation
context and history recording of the changes**
- 15. Completion of CalSDB interface into HIPE**
- 16. Irregular grid 2-d interpolation**
- 17. Display**
 - **aperture photometry improvements**
 - **WCS matching of images**
 - **source fitting errors**
 - **import of RGB images**

Data Processing Overview: Priorities for HCSS 11 (Framework & system)

18. Cube Spectrum Analysis Toolbox

- improvement of the moment algorithm for computeVelocityMap
- support for the new ParameterCube product (viewer)
- rectangular selection is added

19. Improvements to PAL performance

20. Support to stabilise system for future reprocessing exercises

21. Improvement of SPG interface for DP TAs to ease future reprocessing exercises

Data Processing Overview: Priorities for HCSS 11 (HIFI)

- 1. Investigation of performance issues for HIFI pipeline**
- 2. HIFI uplink product**
- 3. Update of spur-finder to improve data quality**
- 4. Update of calibration table**
- 5. Improvements to HEB standing wave mitigation prototype tool**
- 6. Improvements to sideband ratio**

Data Processing Overview: Priorities for HCSS 11 (PACS)

System

- 1. Search for further memory leaks / reduction/optimisation of memory use**
- 2. Package / code clean-up**
- 3. Test harness coverage improvements**

Photometer

- 1. Improvements in baseline drift correction**
- 2. Scanamorphos HIPE implementation**
- 3. Level 3 products**
- 4. Improved error calculation for level 2.5 products**
- 5. Improved calblock filtering**

Data Processing Overview: Priorities for HCSS 11 (PACS)

Spectrometer

- 1. Flatfielding improvements**
- 2. Calibration improvements**
- 3. Background subtraction improvements**
- 4. specProject/drizzle improvements**
- 5. Product improvements**
- 6. lpipe script for high flux pointing loss correction**
- 7. User scripts: mapping of line fit parameters and more improvements**

Data Processing Overview: Priorities for HCSS 11 (SPIRE)

Photometer

- 1. New definition of Level 2.5 and Level 3 products**
 - **Level 2.5 contains single pairs of nominal and orthogonal SPIRE destriped parallel maps**
 - **Level 3 will be any overlapping region from a given program using the absolute calibrated maps from the Planck zero point**
- 2. New extended calibration framework**
 - **improved calibration scheme for extended sources**
 - **Correction coefficients in the form of a Calibration table**
- 3. Improved cross-calibration with Planck (TBC)**

Data Processing Overview: Priorities for HCSS 11 (SPIRE)

Photometer

- 4. Map Projections for Naïve Mapper: Implementation of different WCS and pixel minimization into map maker**
- 5. Source subtractor within Source Extraction task**
- 6. Interactive Level-2 deglitcher tool**
- 7. Documentation how to perform source extraction for extended emission (TBC)**

Data Processing Overview: Priorities for HCSS 11 (SPIRE)

Spectrometer

- 1. Bright mode processing to be done in voltage pipeline**
- 2. Updates to the telescope model calibration product**
- 3. Improvement of pipeline for H+L - LR processing**
- 4. Updated instrument and telescopeRsrf calibration products**
- 5. Level-1 spectral products to be averaged**
- 6. Tuning of spectrometer parameters for wavelet deglitcher**
- 7. Update MAD deglitching (2nd level deglitching) parameters for observations with >30 repetitions**
- 8. Improvement of memory usage when threading**

Data Processing Overview: Priorities for HCSS 11 (SPIRE)

Spectrometer

- 9. Improvements for semi-extended tool**
- 10. Sinc-Gauss function for the fitter**
- 11. Convolution Mapper**
- 12. Updates for WCS handling**
- 13. Updated error propagation for naive mapper algorithm**

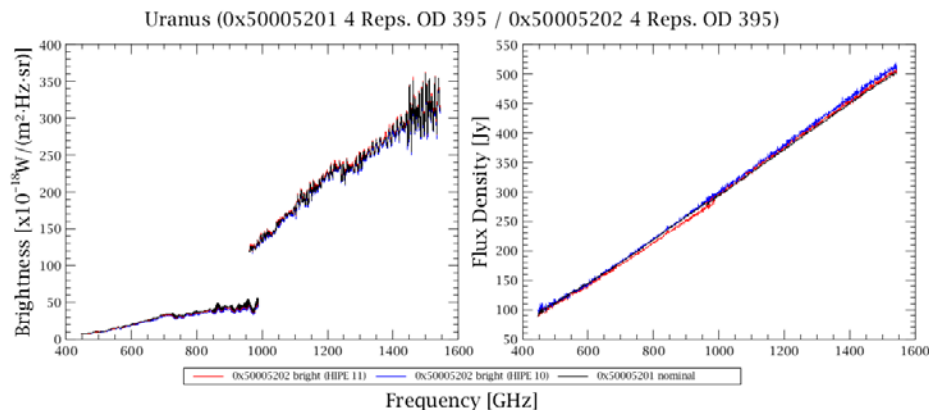


herschel

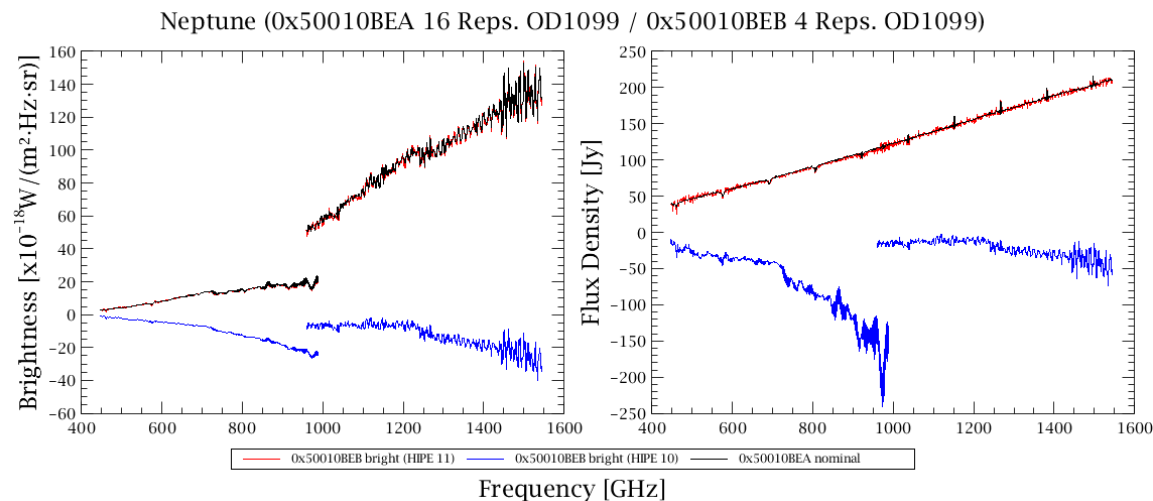
DP Development



Sneak preview of improvements in **HCSS 11** for SPIRE products



**SPIRE bright
mode FTS spectra**



Data Processing Overview: Priorities for HCSS 12 (Framework & system)

- 1. Gyro-propagation for pointing products**
- 2. List-return syntax for tasks (reduction of memory leaks)**
- 3. Improvements in user-friendliness (GUI layout, tooltips, harmonisation of Python syntax in HIPE)**
- 4. Improvement of SPG interface for DP TAs to ease future reprocessing exercises**
- 5. Creation of a multi-observation context to reduce download time for users**

Data Processing Overview: Priorities for HCSS 12 (HIFI and PACS)

HIFI

- 1. Performance issues for HIFI pipeline**
- 2. HEB standing wave mitigation**

PACS

- 1. Tailored level 2.5 products (extended vs. pointsource)**

Data Processing Overview: Priorities for HCSS 12 (SPIRE)

- 1. Alternative error maps for naïve mapper: Provide the user with a set of alternative error maps including standard (no weighting), bolometer noise weighted and hybrid cases for low coverage maps**
- 2. Destriper improvements**
 - improved Level 2 deglitching algorithms
 - destriper working on entire timelines
 - implementation of weighting into restripping
 - detection of bolometer jumps
- 3. Automatic correction of data affected by cooler burps**

Data Processing Overview: Priorities for HCSS 12 (SPIRE)

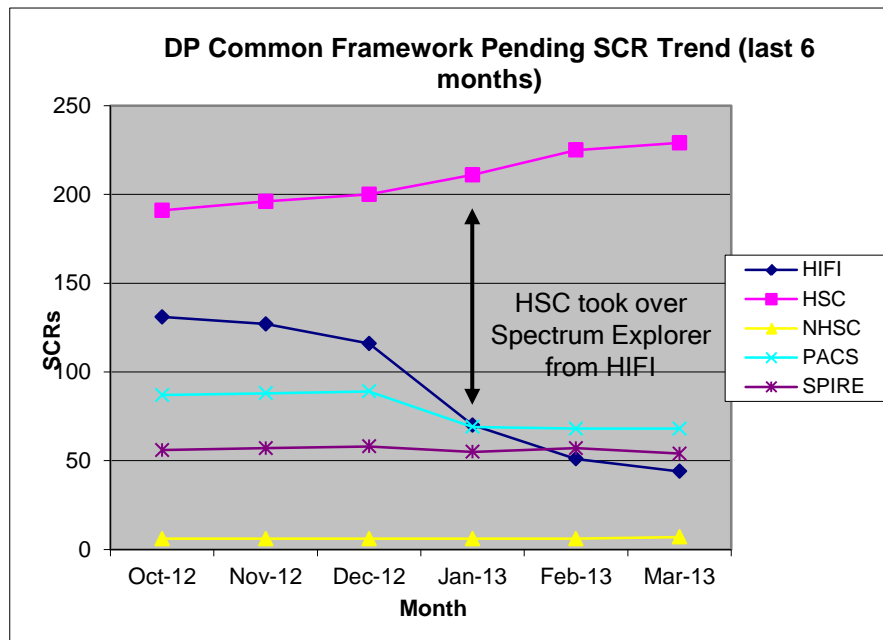
- 4. Continued improvements on Zero Point Calibration with Planck**
- 5. Flux Calibration: New flux calibration using new Neptune fluxes from (ESA_xx) models**
 - new flux calibration files
 - new Temp Drift Correction calibration files
 - new Relative gain correction
- 6. Super Resolution Maps: Port the HiRES code into HIPE as a task for Super resolution mapping**
- 7. Iterative pipeline models to improve deglitching and jump detection**

Data Processing Overview: Priorities for HCSS 12 (SPIRE)

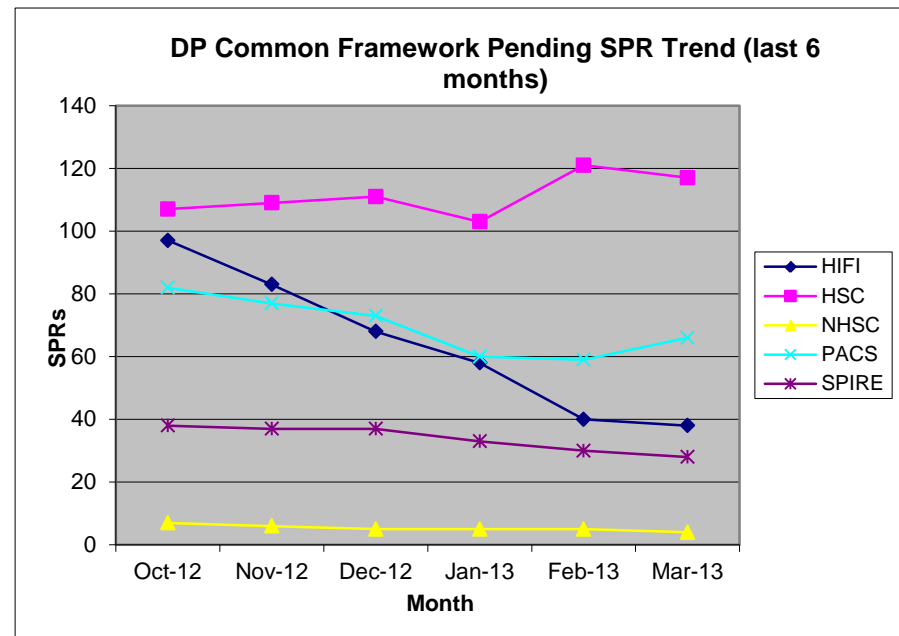
- 8. Quick pipeline without FFT modules up to the Level 2 deglitcher in the destriper that flags glitches in the Level 0.5 timelines, using these results when the full pipeline is run again**

Project development statistics

Trend of s/w tickets for pipeline system, HIPE, documentation and user tools



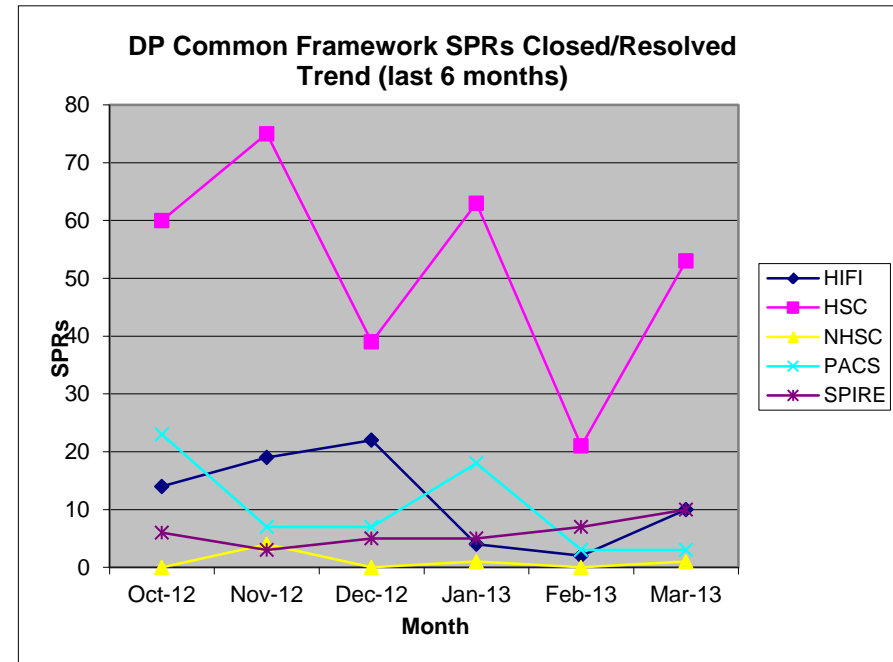
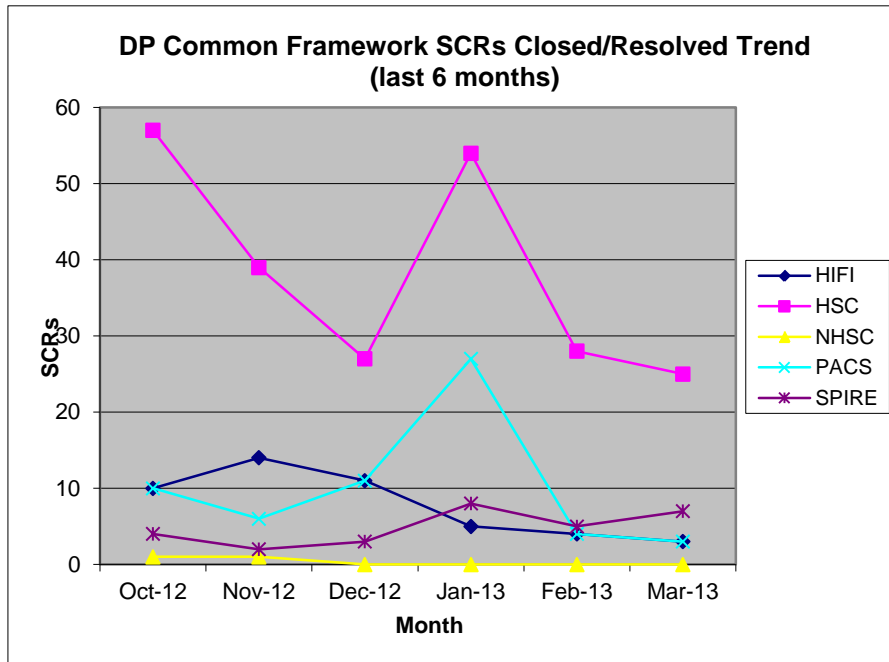
number of change requests
reduced by 17% during last
6 months



number of problem reports
reduced by 30% during last
6 months

Project development statistics:

Closure rate of s/w tickets for Trend of s/w tickets for pipeline system, HIPE, documentation and user tools



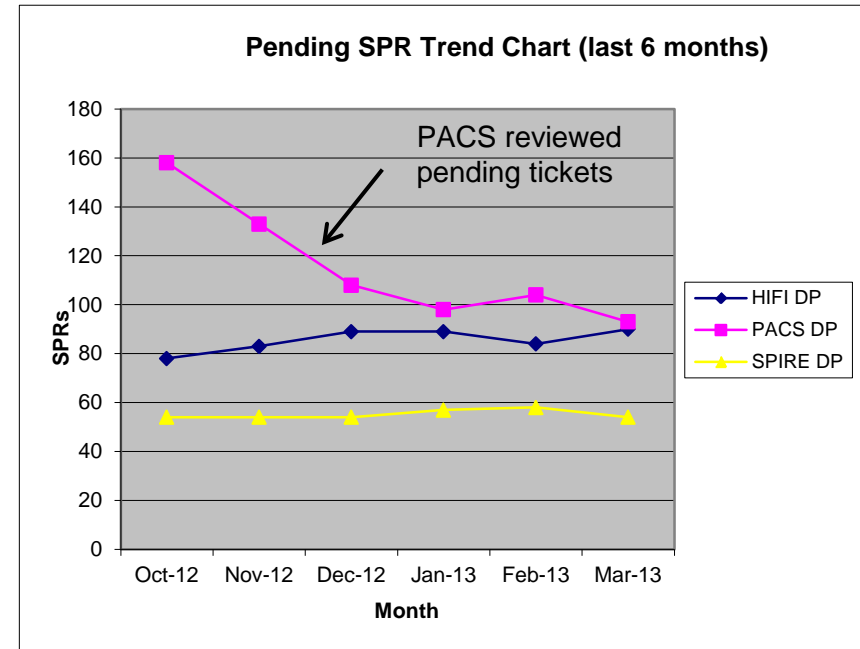
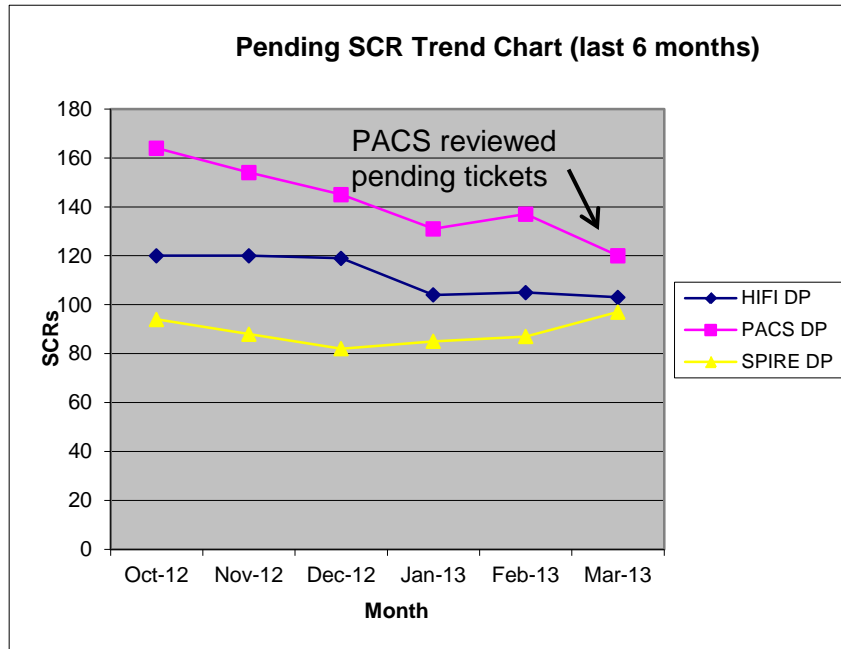
Change Requests

150 s/w tickets were implemented each month

Problem Reports



Project development statistics: Trend of instrument DP s/w tickets

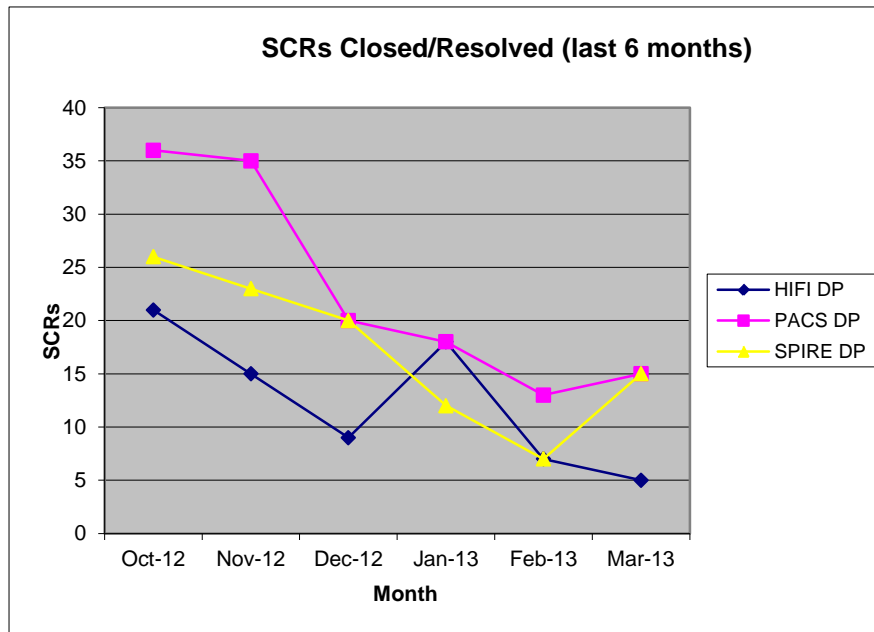


number of change requests
reduced by 18% during last
6 months

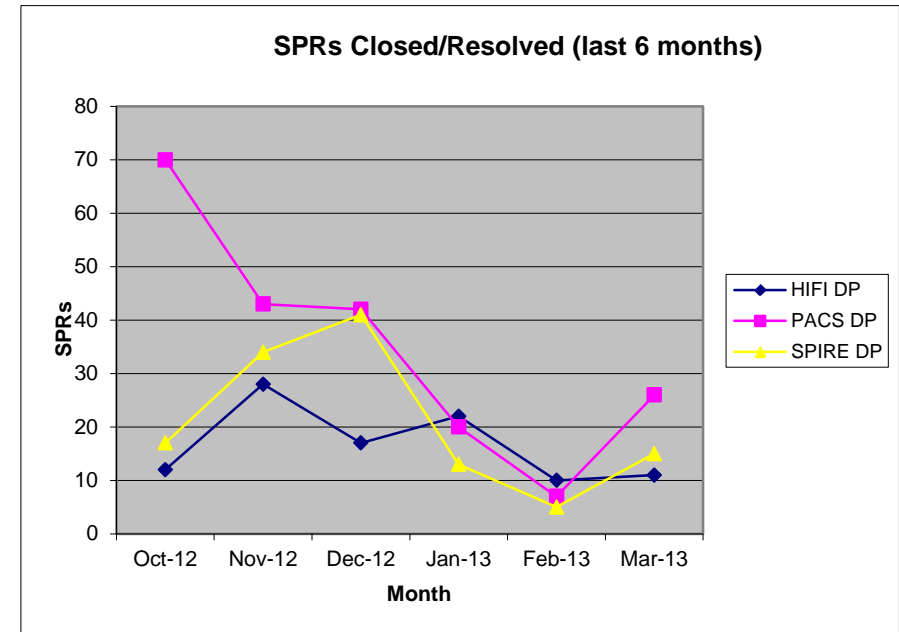
number of change requests
reduced by 22% during last
6 months



Project development statistics: Closure rate of instrument DP s/w tickets



Change Requests

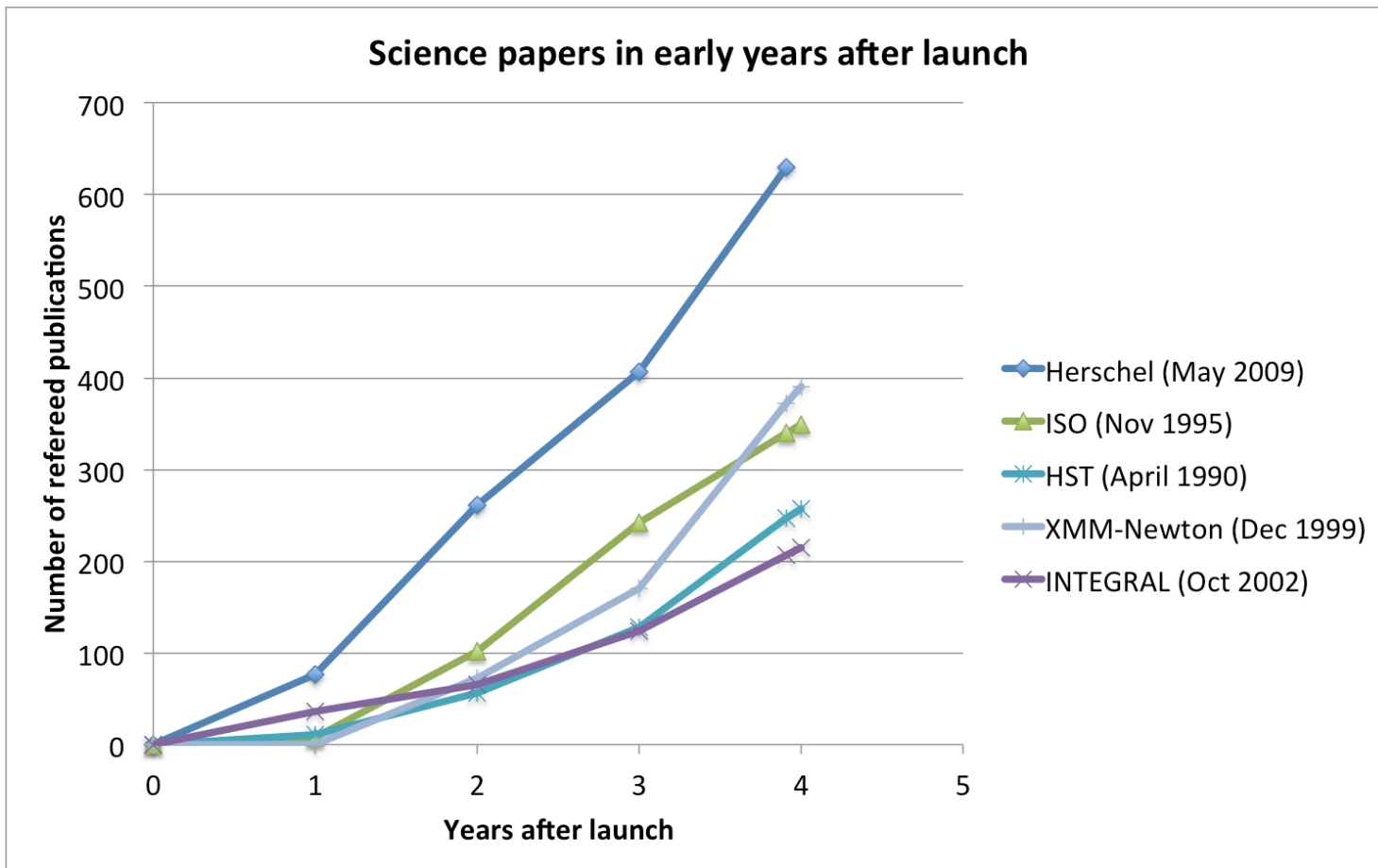


Problem Reports

120 s/w tickets were implemented each month



Payback time for the efforts of the Herschel Science Ground Segment continues!



On-demand processing profiles

- **Rationale: Expand the Herschel community**
 - **Casual users and non-experts do not want to get acquainted with Herschel instruments, calibration and data processing**
 - **Casual users and non-experts do not want to install, learn or run HIPE**
 - **Yet casual users and non-experts want to get products which are tuned to their scientific objectives, improved vs. the current standard pipeline products?**



On-demand processing profiles

- **Assuming this need exists, might this then require something other than even the legacy standard pipeline products?**
- **Opinions offered:**
 1. **No changes, as one does not expect that any significant science can be performed using pipeline products**
 2. **No changes, as one hopes that the final legacy standard pipeline product has adequate quality to address all science goals**
 3. **By provision of all-in products (bulk processed version stored in HSA, on-demand version delivered on the fly)**
 4. **By provision as a combination of static and several, user-selected flavours of on-the-fly generated product**



On-demand processing profiles

- **Assuming the need exists, can be fulfilled, and opinions (1) and (2) are discarded, what is the best way**
 - **to produce static level 2+ products tuned to scientific objectives for the legacy archive**
 - **and, potentially, support the generation of additional high level products: Point source lists, spectra, SEDs – multi-observation, multi instrument**



On-demand processing profiles

- **Option (3): All-in products:**
 - **No adaptations for HCSS and HSA**
 - **Series of different level 2+ products produced with different algorithms (i.e. PhotProject for point-sources, MADMap for extended sources) delivered as one data set that might be confusing for casual users**
 - **Generation of multiple level 0 and level 1 products will multiply data volume and download time**
 - **Executing all options to generate those products will multiply processing time**



On-demand processing profiles

- **Option (4): standard and several on-the-fly products:**
 - **Adaptations for HCSS and HSA needed**
 - **Self-contained dataset tuned to scientific objective**
 - **No increase in data volume and download time**
 - **No increase in processing time**
 - **Increased modularity permits to distribute the development, possibly leading to an increased sense of ownership and a faster improvement cycle**
 - **First investigations indicate that only a series of fixed profiles are affordable (no simple GUI to set switches; no complex GUI to control each pipeline step to change defaults to a limited number of options)**



On-demand processing profiles



Discussions in SGS showed divergence of opinions as to the preferred approach – could the HUG give recommendations?

Even more questions / wishes to the HUG

- **Could you look at the suggested ToR/membership of and volunteer for the Herschel Archive Group (HAG)**
<http://herschel.esac.esa.int/twiki/bin/view/HSC/HAG>
to maximise Herschel's impact for the whole scientific community, e.g. also ALMA, SOFIA, archival research?
- **Could you encourage and support the extension of DPUG with members outside of the current Herschel community (ALMA, SOFIA, archival research)?**
- **Could you encourage and support HIPE acceptance test campaigns with scientists outside of ICCs/HSC/NHSC?**



Your questions, please?