PRECISION RADIAL VELOCITY SPECTROMETER

Document Title	**Data Reduction Pipeline Sub-system Design**
Document Number	PRVS-TRE-00005-0001
Issue	1.0
Date	15th September 2006

Document Prepared By:	Andrew J. A. Vick	**Signature and Date**	A.J.A.Vick 15th Sept. 2006
Document Approved By:	D.W.Lunney	**Signature and Date**	D.W.Lunney 15th September 2006
Document Released By:	D.W.Lunney	**Signature and Date**	D.W.Lunney 15th September 2006

CHANGE RECORD

Issue	Date	Section affected	Change Description
0.9	3/7/2006		Initial document
0.91	7/7/2006	All	General corrections
0.92	7/8/2006	All	General corrections and Red Team review changes
0.93	9/8/2006	All	Re-write of various sections
0.94	13/8/2006	5.3.6, 6.1.1	Updates from Head of Group comments
1.0	15/9/2006	All	Final review before release to Gemini by DWL

PRECISION RADIAL VELOCITY SPECTROMETER

Document Number:	PRVS-TRE-00005-0001
Issue:	1.0
Category:	Technical Report
Status:	Issued
Author:	AJ Vick
Date:	15th September 2006

Table of Contents

List of Abbreviations

A+G	Gemini acquisition and guidance system
FPRD	Functional and Performance Requirements Document
FOV	Field of view
FV	Fibre viewer
GCAL	Gemini facility calibration unit
HR	High (spectral) resolution
OAP	Off-axis parabola
OCDD	Operation Concepts Definition Document
PRVS	Precision Radial Velocity Spectrometer
PSF	Point spread function
R	Spectrograph resolving power
RV	Radial velocity
SRF	Spectral Response Function
XD	Cross-dispersed
HET	Hobby-Eberly Telescope
DR	Data reduction
OLDP	Gemini On line data processing environment
IRAF	NOAO Image Reduction Facility

Table of Definitions

TBD	To Be Defined : a requirement to be developed during the preliminary design stage of the instrument.
TBC	To Be Confirmed : a requirement that is correct with the current design information but requires confirmation during the preliminary design stage of the instrument.
TBR	To Be Reviewed : a requirement specified to meet the PRVS top-level requirements, but which might over-constrain the design.

1 PURPOSE

This document describes the initial design of the Data Reduction (DR) pipeline for the Gemini Precision Radial Velocity Spectrometer (PRVS) as required to carry out the observations described in the science case. It includes an analysis of the derived requirements on the DR pipeline and suggested methods for implementing the pipeline within the Gemini OLDP system, with reference to existing systems that can be re-used as part of the design. It also provides a development strategy for the pipeline.

2 APPLICABLE AND REFERENCE DOCUMENTS

Reference	Document Title	Document Number	Issue & Date
AD01	A treatment for background correction on the Hamilton Echelle Spectrograph, Churchill, Christopher W.; Allen, S. L.	Astronomical Society of the Pacific, Publications (ISSN 0004-6280), vol. 107	No. 708, p. 193-198 (PASP Homepage) Publication Date: 02/1995
AD02	Reliable elimination of telluric lines from stellar spectra, Hrudková, M.; Harmanec, P.	Astronomy and Astrophysics, Volume 437	Issue 2, July II 2005, pp.765-768 (A&A Homepage) Publication Date: 07/2005
AD03	Disentangling telluric lines in stellar spectra, Hadrava, P.	Astronomy and Astrophysics, Volume 448	Issue 3, March IV 2006, pp.1149-1152 (A&A Homepage) Publication Date: 03/2006
AD04	Wavelength Calibration in Physical Model Based Pipelines, Fiorentino, M.; Bristow, P.; Kerber, F.; Rosa, M.	Astronomical Data Analysis Software and Systems XIV ASP Conference Series, Vol. 347	ASP, 2005., p.619
AD05	Optimal Techniques in Two-dimensional Spectroscopy: Background Subtraction for the 21st Century, Kelson, Daniel D.	The Publications of the Astronomical Society of the Pacific, Volume 115	Issue 808, pp. 688-699.
AD06	Precision radial velocities with an iodine absorption cell, Marcy, G.W; Butler, R.P.	Astronomical Society of the Pacific, Publications (ISSN 0004-6280), vol. 104	No. 674, April 1992, p. 270-277.
AD07	Guidelines for Designing Gemini Aspen Instrument Software, Gillies, K.	Gemini Observatory Engineering Software Report	AspenSoft-03072004-6, May 13th 2004.
AD08	The Gemini on-line data processing system, Walker, Shane; Gillies, Kim; Brighton, Allan	Ground-based Telescopes. Edited by Oschmann, Jacobus M., Jr. Proceedings of the SPIE, Volume 5493, pp. 432-443	2004

3 OVERVIEW

The data reduction pipeline for any Gemini Aspen instrument is required to provide data that is "science ready". In the case of PRVS this means data that is ready for use in radial velocity estimation codes. This in turn means that any artefacts from telescope, instrument etc. that can be removed are removed, and that the quality and calibration of the data are known.

Since the final output is for Radial Velocity (RV) estimation, the most crucial factor is the centroid of lines (which is dependent on factors such as calibration and line shape) and this has to be the focus of the DR recipes. Alongside this is the issue of calibration and quality; PRVS is designed to be a very stable instrument but that stability has to be constantly measured and checked.

3.1 STRATEGY

Developing a data reduction pipeline in parallel with an instrument is by no means an impossible task, but it does require careful planning and forethought. Whilst many of the hurdles to be crossed can be identified from prior-knowledge and risk analysis, the "unknown unknowns" of the final instrument tend to affect the data reduction package above all others, since it is within the data reduction that the final solution has to be provided.

In order to provide a data reduction package that is ready for commissioning with the instrument and takes account of any features that are identified during the development, we have chosen a strategy that combines a managed development with flexibility, through three distinct approaches;

1. the development of a test bed system, based on the PRVS Pathfinder, that will allow us to test DR recipes with data similar to PRVS.
2. a work plan that focuses on iterative cycles throughout the project (see Section 4) allowing us to refine our pipeline and provide sub-system and system tests as the project progresses.
3. a close integration with the Gemini DR software teams, to ensure that final integration is as smooth as possible. Gillies outlines several support options in "Guidelines for Designing Gemini Aspen Instrument Software", of which we would prefer Enhanced Support but using electronic communication rather than travelling to the site.

It is likely that last minute changes to the DR pipeline will have to be made during commissioning in response to emerging details. We believe that our strategy will minimise these and allow the instrument to be scientifically productive as soon as possible.

3.2 REQUIREMENTS

The Gemini document "Guidelines for Designing Gemini Aspen Instrument Software" calls for delivery of a data reduction package that will reduce data from all major instrument modes and provide testing, quality analysis and, ultimately, results free from instrument or atmospheric artefacts, i.e., science ready data.

The Aspen document (AD07) outlines the following requirements for the DR sub-system;

Basic Gemini Requirements

Source	Requirement
AspenSoft-03072004-6, Section 12	Data processing code must be delivered with the instrument to maximize data quality during observing and to deliver science quality results in normal operating modes.
AspenSoft-03072004-6, Section 12.1	Builders must provide data processing code that provides science quality results in all the major observing modes during commissioning and science operations.
AspenSoft-03072004-6, Section 12.1.1	The data processing code should operate upon and generate FITS files natively. No translation between different formats is acceptable
AspenSoft-03072004-6, Section 13.1	The software development schedule should include several milestones where components of the instrument software are ready for collaborative testing.

3.2.1 Extended Gemini requirements

These requirements also come from text within the "Aspen Guidelines" (AD07), but they are not quoted as formal requirements within that document.

1. data processing code must be integrated with the Gemini OLDP system.
2. normal/major observing modes within the OCDD should be supported with processing code for
3. quality checking at the telescope
4. production of science quality results
5. OLDP should not be used to provide very low-level instrument processing.
6. the builder must provide any new algorithms required for processing instrument data in the major observing modes.
7. an automated test suite demonstrating proper operation of the data processing code in all major observing modes must be provided.
8. the test suite should be standalone and automated
9. all data processing code must operate upon and generate FITS standard files.
10. all data processing code must be delivered in source code format and must be licensable such that Gemini can make changes and distribute the code to interested parties.
11. results should be viewable with any image display program that can view FITS images of the type written by the instrument.
12. tasks should be written to operate independently, and care should be taken in the granularity of tasks to maximise efficiency in the OLDP.

3.2.2 IRAF specific guidelines

These requirements also come from text within the "Aspen Guidelines" and relate specifically to IRAF data reduction tasks, though they could also be applied to any other language or system used for data reduction processing.

- all IRAF scripts must be capable of running in non-interactive mode. This means that;

  - the only required/positional parameter is the primary input.
  - All other parameters are given default values.
- all Gemini IRAF tasks used with the OLDP contain two parameters in common:
  - logfile - specifies the name of the log file the job will produce
  - status - specifies the exit status of the script

3.2.3 Task descriptor objects

The following text is taken from the "Aspen Guidelines" and relates to the task descriptions that the workflow system in OLDP uses. This is a useful specification since it (partially) defines the interface between the OLDP and the tasks independently of the underlying system used (e.g. IRAF);

An OLDP task descriptor object includes the following information

- the inputs that the task expects to act upon (a primary input parameter and any number of calibration input parameters)
- a simple algorithm for mapping the primary inputs to outputs (some tasks combine their inputs, others produce an output per input, etc.)
- a collection of parameters, their types, default values, whether they are required or optional, common or advanced, etc.
- the name of a Java job adapter class which knows how to execute the task

3.2.4 PRVS Requirements

The following are requirements taken directly from the PRVS FPRD. The FPRD also contains a derived list of requirements originating from the Gemini Aspen software requirements, these are not re-iterated here. The requirements listed here are all derived from science or operational requirements, and may have to be taken into consideration when developing DR recipes.

Requirement

The instrument shall be capable of measuring the centroid of a spectral line to better than 0.001 pixels with a goal of 0.0005 pixels. In the course of a 1 hour observation the spectrum shall remain stable to 0.1 pixels without calibration.

Source

FPRD 3.2.2.10.2

Effect

The proposed design must deliver data which allows centroiding to this accuracy; it should make maximal use of the information available and the recipes should in no way degrade the line centroid in favour of other factors.

Requirement
Observational Efficiency shall be better than 70% for a one hour observation. This does not include the time taken to perform flat fielding.
Source
FPRD 3.2.2.2
Effect
If the DR system is required to reduce acquisition data for viewing or automatic acquisition, it must be capable of doing so fast enough to allow this.

Requirement
The flat field calibration of the instrument shall be performed to better than 0.2 % deviation from the mean value.
Source
FPRD 3.2.2.13

Requirement
The spectral resolving power shall be greater than R=50000 (essential) with a pixel sampling of 2.5 pixels per resolution element or greater. The goal for spectral resolving power is R= 70000.
Source
FPRD 3.2.2.10

3.3 DESIGN

In order to progress the design of PRVS, derived requirements that should fulfil the primary science and operational requirements have been developed. It should be noted that some (or possibly many) of these requirements may change as the instrument design is developed, but this set offers a good starting point from which to progress the testing of the design.

3.3.1 The Gemini On-line data processing (OLDP) system

The Gemini OLDP system is based on a Java workflow toolkit. Recipes (data reduction tasks) can be written in a number of systems and run under the control of the OLDP. This allows the OLDP to run recipes based on the data available, which maximises throughput and minimises latency. It also provides a platform for parallelisation at the macro level; recipes can be run in parallel when possible.

In designing recipes for the OLDP several factors need to be taken into consideration;

- The system that the recipes are written in; the trade-off between re-use and efficiency (especially of parallel tasks) needs to be taken into account
- The size and scope of tasks, which can again affect efficiency
- Support and re-use of the recipes; the instrument will run for several years and the pipeline will need to be supported and possibly enhanced during this period.

All of these factors affect the development strategy for the software, but at this stage in the project the important aspect is the definition of recipe algorithms. The development itself, the languages, toolkits etc., will be decided on later in the project (at PDR level) once the recipes themselves have been analysed further.

3.3.2 Primary science requirements

The primary science case for PRVS is the detection of planetary radial velocities. In order to achieve this, the data should be processed to remove any artefacts caused by

- the atmosphere
- the telescope
- the instrument

The data should also be calibrated and reduced to a form that is suitable for a scientist to use as input to further data reduction which will calculate radial velocities. Note that the radial velocity code is not included as part of the OLDP based pipeline for PRVS; it is not a deliverable of the PRVS project except as a test of the functionality of the instrument.

In order to achieve these requirements, the following data reduction procedures must be applied (see Task Descriptions for a more detailed study of the tasks involved);

3.3.2.1 Flat fielding and Bias correction

Initial processing of the data will be similar to image data processing - bias subtraction and (possibly) flat field division will be applied to the whole data frame as a single image; these tasks can be accomplished using existing DR routines from, e.g., IRAF. These tasks will also be applied to most of the calibration frames.

3.3.2.2 Order tracing and slit determination

Tracing of the orders, along with background estimation and order extraction, is very much dependant on SRF determination, and this is especially important in RV estimation (see Marcy&Butler). Arc line and gas cell calibration measurements will enable a precise evaluation of the instrumental SRF across all the orders of PRVS, but initial extraction will still require an accurate estimation of the slit orientation etc. Several methods of tracing orders from a fibre fed spectrograph are available (Gaussian fits, centre of gravity etc.). Initial results from the PRVS Pathfinder suggest that fitting a single Gaussian is more consistent with multiple fibres, since fitting of multiple Gaussians (or similar profiles) is inherently non-linear and prone to significant errors.

Final order tracing can be accomplished using normal methods – fitting a spline (or similar) to the estimated centres along the dispersion direction for each order. However, the work of Fiorentino et al suggests that a faster and more accurate solution can be obtained if a ray optics model is used as the first step in building a physical model of the spectrograph, and considering the stability we are required to achieve in the PRVS this method will be investigated up to PDR level design.

3.3.2.3 Background subtraction and Order extraction

Background subtraction is usually performed as part of the order extraction step. This method has advantages, especially since several versions of both normal and optimal extraction exist in source code and can be applied to the PRVS data. However, Kelson has reported on an alternative method for fitting a two dimensional night sky background, which avoids the problem of sharp residuals at the edges of features with sharp gradients – an effect that could have a big impact on the RV estimation code. This method will be examined using the PRVS Pathfinder data before PDR.

Order extraction for PRVS is likely to be done using an optimal extraction algorithm. Several of these exist in source code, so it is unlikely that a new code need be developed.

3.3.2.4 Wavelength calibration

Exact wavelength calibration is not required by the primary science case. The use of the calibration spectrum alongside the object spectrum, coupled with the calibration measurements between the object and calibration fibres, means that a highly accurate wavelength calibration is available for all observations. Since the radial velocity codes will make use of this information directly, it is not necessary to provide a wavelength calibration routine as would normally be done in spectrograph data reduction. However, wavelength calibration will be a vital component of the test and verification procedures, and may be required by the secondary science case.

3.3.2.5 Telluric spectrum estimation

A special feature of the PRVS spectrograph will be the reduction of fainter telluric lines in the data. Because of the number of these lines in the wavelengths observed by PRVS, and the change in velocity between observations, large amounts of the spectrum taken by PRVS have to be masked out during radial velocity estimation. Although little can be done about the stronger Telluric lines (greater than 10% of the continuum level) they can be used to fit a telluric model spectrum (see Hrudkova & Harmanec, Hadrava). Code is available from Hadrava that uses the stronger lines, and a model atmosphere, to significantly reduce the fainter lines. This method increases the amount of spectrum that can be used to determine radial velocities, and thereby increase the accuracy of the instrument. Data from the PRVS Pathfinder experiment will be used to test this method before PDR.

3.3.3 Secondary science case(s)

The secondary science cases for PRVS can (mainly) be handled using some of the data reduction routines developed for the primary case. Considering the secondary cases as more traditional spectrographic data sets, there are few routines that are not in the primary cases list, such as wavelength calibration, although all of the RV calibration data will be available if required.

The main differences in reducing the secondary science data is the possible need to provide the data calibrated for wavelength and combining of the orders to form a continuous data set.

Wavelength calibration: because of the simultaneous calibration spectrum a very accurate wavelength calibration could be provided. Whether any existing routines could be adapted to provide this is uncertain, so the inclusion of this task will need to be re-evaluated at PDR or FDR depending on the amount of work involved.

Combining orders: due to the gap between the two arrays, and the areas that overflow at longer wavelengths, it will not be possible to construct a single continuous data set, so the advantages of providing combined orders need to be evaluated. It may be better to include the determined blaze function (etc) from the DR pipeline in the output data in order to facilitate re-combination of those orders that are of interest at a later stage.

3.3.4 Test and verification requirements

Throughout the development of the system we plan to provide test and verification procedures to aid both in the development of the instrument and of the pipeline. The basic list of identified test cases is outlined below, but it should be noted that the actual test plans will be developed and expanded during the development stages.

- Array uniformity, flat field and bias responses.
- Read noise estimates
- Array alignment and position estimates
- Order traces and optical distortions
- Blaze function
- Instrumental SRF

- Slit orientation and position
- Wavelength calibration

Of all of these tests the instrumental SRF and the wavelength calibration are considered to be the most important. PRVS is designed to achieve a stability commensurate with achieving a prescribed accuracy in determining radial velocities. As such it is imperative that this accuracy (or at least its dependent variables) be measured throughout the instrument's lifetime. To do this we intend to build up a set of quality control recipes that can be used to monitor the instruments stability on both short and long scales.

These tests will build up until the data processing suite is complete. The final deliverable will include two distinct sets of tests;

1. Tests of the functionality and correctness of the pipeline. These tests will (initially) be based on existing laboratory or simulated data and will act as a test of the correct functioning of the pipeline, independent of the instrument.
2. Tests of the functionality and correctness of the instrument. These will be developed from the test plans above, and will seek to determine the status of the instrument at any time after its verification.

It will be necessary to identify a complete set of recipes to facilitate this by PDR.

3.3.5 Operational use requirements

3.3.5.1 Acquisition

Acquisition for the PRVS spectrograph will be accomplished using the fibre acquisition camera. Simple DR processing of the images will be required for display (flat fielding, bias subtraction). In addition, the images will need accurate WCS data added, so that the atmospheric dispersion between the visible CCD image and the IR science image can be taken into account during acquisition. It may also be necessary to identify guide stars on the image automatically.

3.3.5.2 Calibration Line selection

The camera system requires that the locations of bright calibration lines be identified and downloaded to the camera, so that the camera can provide fast read out of windows around these bright lines. This method stops bright lines from saturating and also reduces persistence problems on the arrays. Bright lines will need to be identified, and a window calculated based on the halo of the line on the array.

3.3.5.3 (Slow) Guiding

The camera system will deliver slow guiding commands directly to the Gemini A&G unit (via a socket interface). It will not be possible to use the DR system to process these frames, all processing will have to be part of the instrument control system. But the DR system could be responsible for providing suitable flats and biases to the camera, and also in identifying guide star locations on the array (i.e. setting the origin for the centroiding function).

3.3.6 Quality Control

The test cases developed during the project, from FDR stage through to integration, will be used to specify an initial set of quality control (QC) measures. Recipes for these QC measures will be developed alongside the operational use concepts and will be provided for commissioning.

QC Parameters will be needed to monitor the stability of the instrument, which again resolves down to monitoring of the wavelength references; the line (SRF) profiles and centroid positions on both the calibration and object fibres. The calibration unit itself, the strengths of the various lamps, and lamp to lamp wavelength calibration will need to be monitored. Along with these, parameters will be required to monitor throughput, absolute and relative temperatures and IR array characteristics.

4 WORK BREAKDOWN

4.1 WORK PLAN

During the phases of development of PRVS we intend to iterate the DR pipeline development several times, providing greater robustness and usability at each iteration. The output of this during the stages of the project is expected to be;

- Test data and primitives, together with some test recipes during PDR and FDR design
- Primitives and test recipes during sub assembly integration
- Science recipes and test scripts during Integration and Test

At each iteration the simulated data will be used to provide unit tests of the work produced, and the data itself will be updated to reflect the design and what is known of the performance of the instrument.

4.1.1 PDR Design

In order to commence work on the DR pipeline design and development as early as possible we intend to follow an iterative design process. We will start by taking the design study and producing a series of example simulated science data and calibration frames, utilising as much information as we can from the various disciplines on the project. Using this as test data, we will investigate available data reduction packages and primitives and produce a PDR design for the data reduction, incorporating

1. Existing packages and primitives that can be re-used
2. Existing packages and primitives that can be re-worked to provide compatibility with PRVS requirements
3. New code that significantly enhances the data reduction for PRVS

In all three cases we will interact regularly with Gemini staff, instrument and project science teams to ensure that the reduction products and the recipes themselves meet requirements for the science, the instrument, and the observatory infrastructure.

At the end of this stage we will have identified the main reduction package(s) and primitives, along with specifications for new or modified code that together will meet the DR pipeline requirements.

4.1.2 Detailed design

Once the PDR phase of the project is complete we will re-work the test data to incorporate new information from the current design, and produce a second level design of the DR, expanding on the depth covered and working towards those recipes required during the sub system build and integration phase for testing of array etc. This work will become the FDR level design, and will include code and test datasets to be used by the instrument team during the integration phase. A particular focus of the design at this stage will be an estimate of run time efficiency, which will be used to determine the best granularity of tasks within the OLDP.

Once the FDR report has been produced, we would hope to spend a short period re-working the test data and recipes from feedback received, then issue these both for the instrument team and the science team.

4.1.3 Sub-assembly build and test

During this phase of the project the initial work would be support for the sub assembly testing. At the end of such tests we would compare the test datasets with any data produced by the sub assemblies, and use this together with any other information gained to re-work test datasets and recipes.

The final piece of work during this stage would be development of an integration plan and test recipes to aid the integration phase.

4.1.4 Integration

Integration would be a similar work flow to sub assembly, with support for testing followed by feedback. At the end of the integration period there will again be an opportunity to review and re-work test data and recipes, this time with realistic ideas of detector and optical instrument effects on the data quality. We would spend some time analysing these results, and at the end of the integration period a full suite of commissioning recipes and test results would be produced in line with the commissioning plan.

4.1.5 Commissioning and beyond

The expectation is that this project will complete at the stage of science verification. Though we have outlined a plan to maximise the amount of work that can be accomplished before this point, it is likely that some re-working of the DR pipeline will be required once verification is complete.

Whilst we will endeavour to apply good software engineering principles to the design and coding of the pipeline in order to provide a maintainable and robust infrastructure, we note here that the knowledge and experience gained by the development team could form major resource for Gemini during the initial period of the instruments observational use.

5 APPENDIX: DATA REDUCTION TASKS

5.1 INPUT SOURCES

- SPECFLAT: Spectral flat exposure
- DETCFLAT: Detector Flat exposure
- BIAS: Dark/Bias exposure
- CALSRC: Calibration source exposure
- OHSKY: Sky exposure for OH identification
- SKY: Sky exposure with OH reduction in IR camera
- STDSTAR: Standard star exposure
- SCIENCE: Science object exposure

5.2 POSSIBLE OUTPUTS

- OH line table
- Trace table
- Bad/Hot pixel map
- Background estimate
- Image data
- Extracted object spectrum
- Extracted calibration orders (from calibration fibre)
- Extracted calibration spectrum (from cal source illumination of the object fibre)
- Various QC parameters;
    - Instrumental SRF
    - Wavelength scale (per order or global) to be compared to model
    - Background; flat and bias variation statistics
    - Throughputs
    - Lamp intensities
    - Temperature and pressure data from control software

5.3 INDIVIDUAL TASKS

5.3.1 Simple tasks

- Bias frame preparation
    - Median several frames to remove anomalies
- Flat field frame preparation
    - Subtract Bias frame from each
    - Median to remove anomalies
    - There will be two types of flat fields available for PRVS;
        1. Flat fields produced by a wide band source in the calibration unit. This flat field is produced via the full spectrograph optics, and will provide high SNR in the orders and very low SNR outside. It will also be affected by the spectrum of the wide band source and the blaze function.
        2. Flat fields produced by a narrow band source directly illuminating the IR arrays. This flat field will not go via the dispersive optics, and will therefore provide an accurate pixel to pixel variation across the entire array.

    The dispersed flat field (1) will be used in tracing the orders as well as it's normal use; flat fielding the order data. The direct illumination flat field will be provide accurate measurements of the blaze function and improve the fitting of background outside of the orders (sky or scattered light).
- Order location
    - Simple location of orders by taking a cut through them
    - Possibility of using physical model to determine the location
- SRF estimation code
    - SRF estimation via re-sampling (“Drizzle” method)

5.3.2 Higher level

- Calibration Line location
    - required inputs: CALASRC
    - other inputs: BIAS, DETCFLAT, SPECTFLAT
    - output: Calibration line table
- Bias subtraction, pixel scaling (division by flat)
    - Required inputs: BIAS, FLAT,
    - output: Bias subtracted image
- Order tracing

    - determine the path of each order across the array. Possibly tracing each fibre output separately, or by a global fit to all fibres.
    - Required inputs: Bias subtracted image
    - other inputs: Fit parameters
    - output: List of order locations

- Flat field normalisation
    - Required inputs: FLAT, BIAS, CALSRC
    - Before the flat field can be utilised to reduce inter-pixel variations it has to be normalised to remove blaze and other variations. This is normally done by fitting a polynomial along the dispersion direction and (possibly) the spatial direction and then setting pixels that lie in inter order gaps to 1.0. There is also the possibility of using a flat produced from an even illumination of the array in estimation of these polynomials - this might improve the fit at the ends of the orders, where the blaze function reduces the signal.
- Background fitting/removal
    - Required inputs: Bias subtracted image, List of order locations
    - other inputs: Physical model of spectrograph
        - See also Churchill&Allen
- Order extraction and background removal
    - Required inputs: Bias subtracted image, List of order locations
    - other inputs: Physical model of spectrograph
    - output: Extracted (1D) orders
    - Extraction methods
        - linear extraction: the simple integration of all pixels in the orders profile with the corresponding background channel removed
        - optimal extraction: where the detectors characteristics (gain, & readout noise), photon statistics and profile are used to weight contributions to the signal. Since we intend to have access to readout noise estimated from NDR sampling, along with accurate gain and profile information, the benefits of this method will be investigated.
- Calibration order extraction
    - The same code as science object order extraction