Supplementary Material for Pipeline Processing and Data Products for Roman Wide Field Spectroscopy

Roman Science Support Center at IPAC

March 2, 2022

Purpose

- The purpose of these slides is to provide additional details on the processing of data from Roman Wide Field Spectroscopy
- The primary description of this pipeline is contained in "SSC Support for Roman ROSES 2022 call for proposals"
- This processing at the Roman Science Support Center (SSC) at IPAC is designed to meet the science goals of the High Latitude Wide Area and High Latitude Time Domain surveys

Roles & Responsibilities

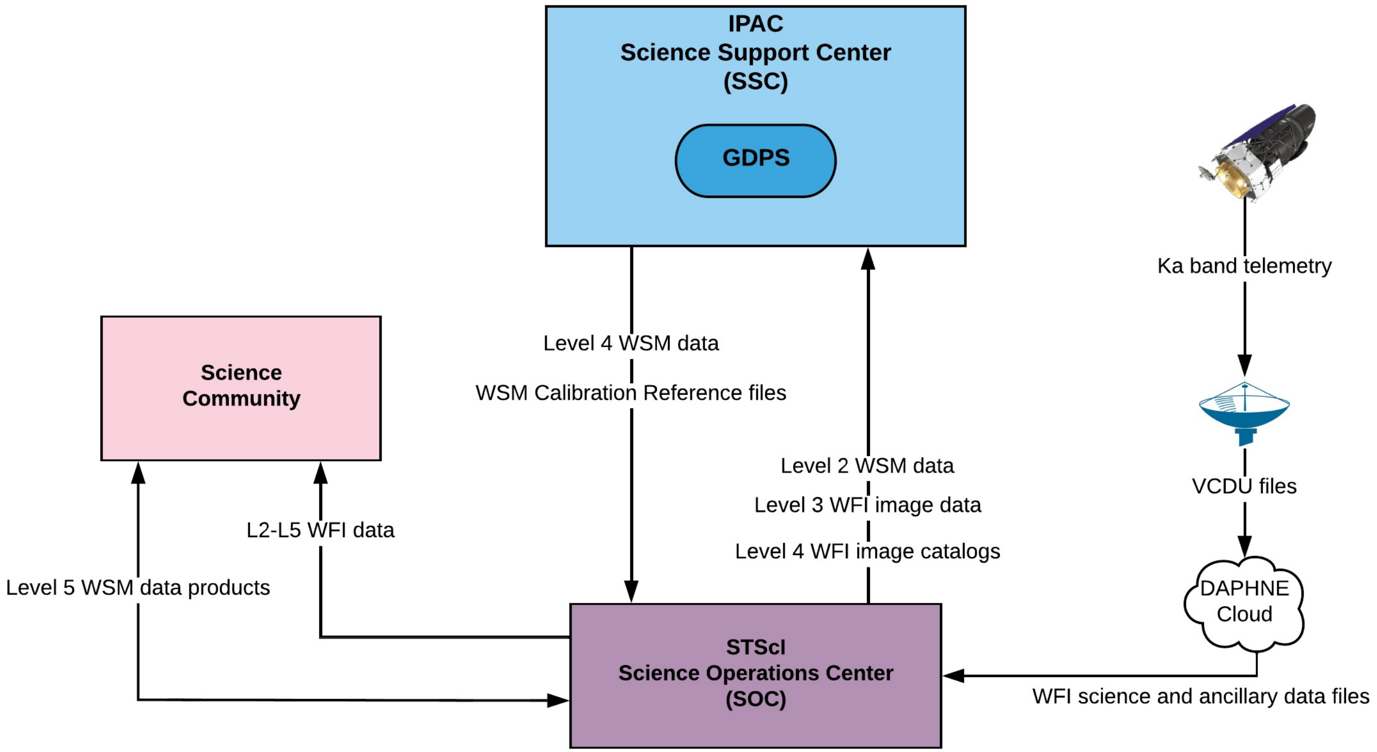
- The SSC and the Laboratoire d'Astrophysique de Marseilles (LAM) are collaborating to create the Roman Grism-Prism Data Processing System (GDPS).

- The GDPS is built on IPAC and LAM expertise, particularly the Euclid NASA Science Center at IPAC (ENSCI), part of the distributed Euclid processing system.

The primary responsibilities of the GDPS are to:

- Create and maintain the Roman pipelines that will operate on all WFI spectroscopic data.
- Reduce all grism and prism science and calibration data starting at Level-2. The spectroscopic pipeline also uses Level-3 and Level-4 associated WFI image data products.
- Deliver Level-4 spectroscopic data products to the Roman archive.
- Release pipeline modules and associated documentation to the scientific community.

GDPS Operational Interfaces



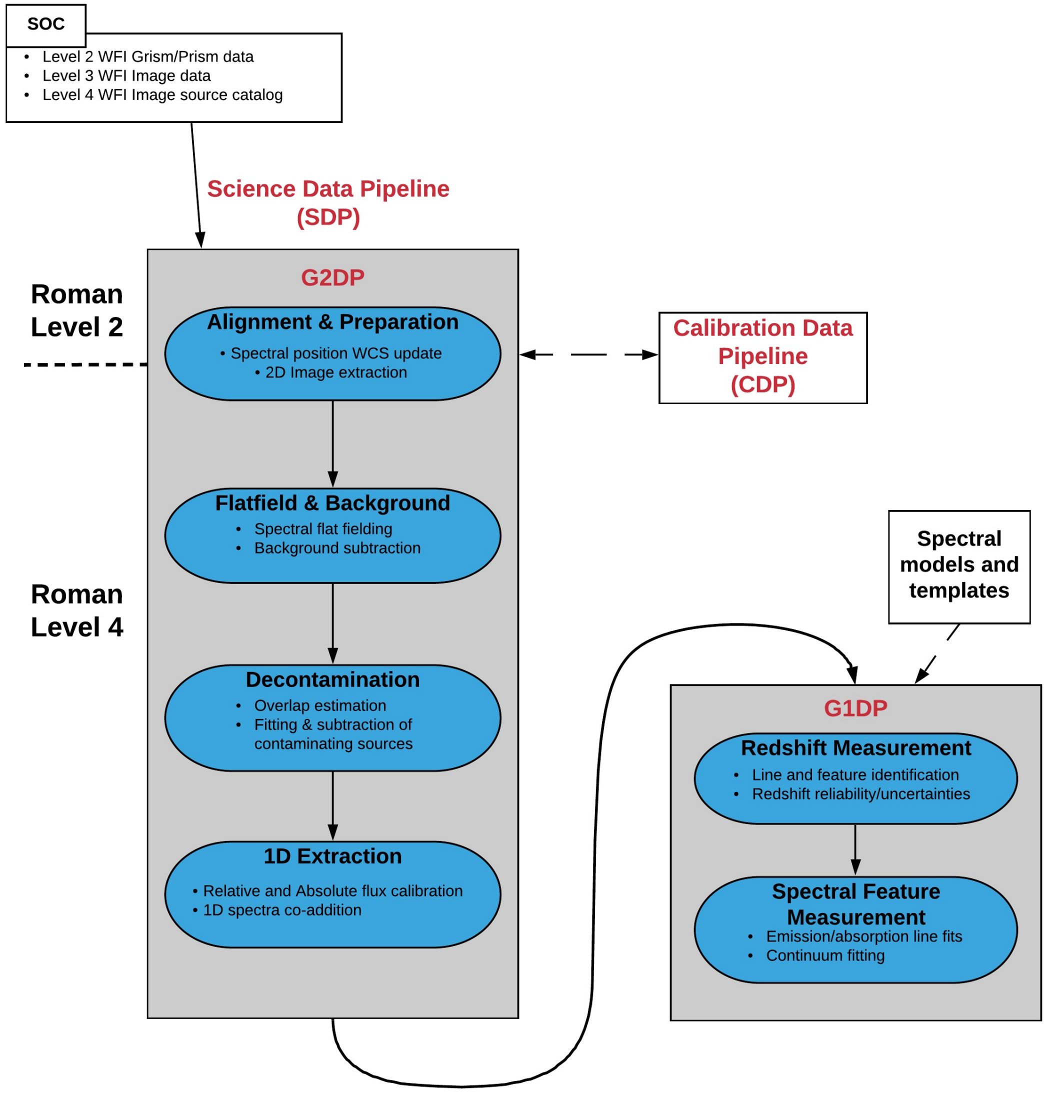
GDPS Design: Pipelines

- The GDPS consists of two main sub-systems:
  - **Science Data Pipeline** (SDP): Processes all WFI grism and prism science data beyond Level-2.
  - **Calibration Data Pipeline** (CDP): Processes grism and prism calibration data beyond Level-2. Produces Calibration Reference Files that become inputs to the SDP, necessary for the processing of the grism and prism science data.

- The SDP consists of two coordinated pipelines to reduce and analyze the 2D and 1D grism and prism spectra:
  - **Grism-Prism 2D Pipeline** (G2DP): reduces the 2D data and produces 1D spectra for each detected source.
  - **Grism-Prism 1D Pipeline** (G1DP): fits the 1D spectra and produces output parameters for each detected source (e.g., redshifts, line fit parameters, etc.)

GDPS Science Data Flow

- *Basic data flow for the SDP.*
- *WFI spectroscopic and image science and calibration data are inputs to the G2DP and CDP*
- *G2DP outputs are inputs for the G1DP.*



GDPS Design: Assumptions

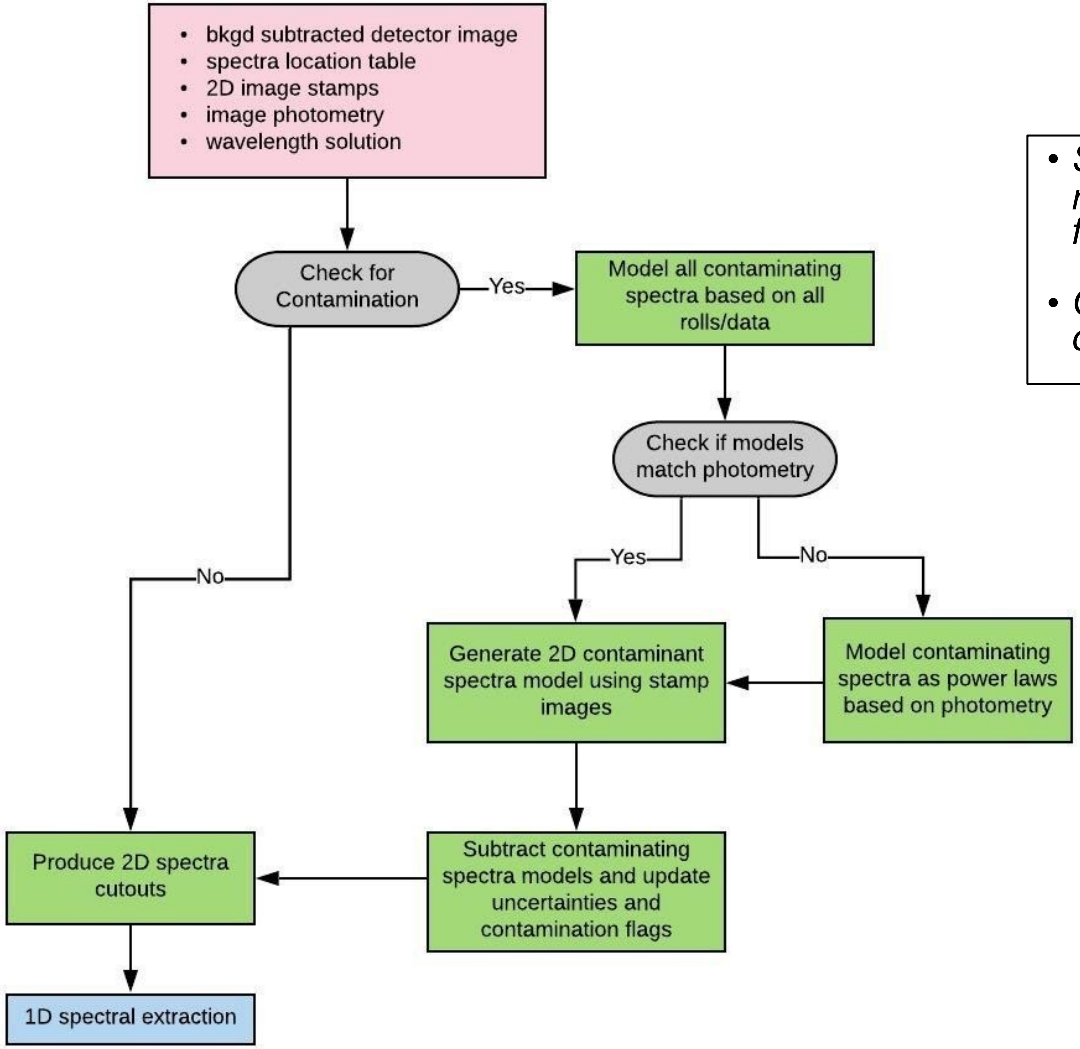
- All grism and prism spectra have corresponding direct imaging data and target source lists available.
- Key calibration observations (e.g. absolute and relative flat field observations, wavelength and flux calibration standards, etc.) are taken during regular facility calibrations.
- Ground test data will be used to derive initial in-flight calibrations and combined with facility calibrations to derive updated calibration reference files used in the science data pipeline.
- Grism and prism spectra are processed similarly, albeit using element specific calibration reference files and pipeline module optimizations. Specialized scientific analysis, such as time dependent analysis of prism spectra (e.g., to detect and separately measure SNe from their host galaxy), will be done by Roman science teams.

GDPS Design: G2DP modules

- Individual pipeline modules in the G2DP include:
  - *CoordMapping:* Convert RA+Dec catalog into bounding boxes for dispersed images
  - *WCSUpdate:* Measure blue edge positions for bright point sources to update positions of spectra
  - *2DStamps:* Cut out small stamp images from direct image for input sources
  - *FlatField:* Apply small-scale flat field (based on background zodiacal light) to input grism/prism
  - *Bkgd_Small:* Mask localized background features using an optical model and input bright sources
  - *Bkgd_Large:* Fit and subtract large-scale diffuse background component
  - *ContamIdentify:* For each 2D spectrum, identify all contaminating sources
  - *Decontamination:* Using all available rolls, model contaminating spectra and subtract them from each source
  - *1DExtract:* Extract 1D spectrum from each 2D spectrum, applying wavelength solution
  - *FluxCal_Relative:* Apply relative flux calibration cube to 1D spectra to correct for large-scale illumination pattern
  - *FluxCal_Absolute:* Apply absolute flux calibration to 1D spectra to produce flux-calibrated 1D spectra.
  - *SpectraCoadd:* Weighted co-addition of spectra from all available dithers/rolls.

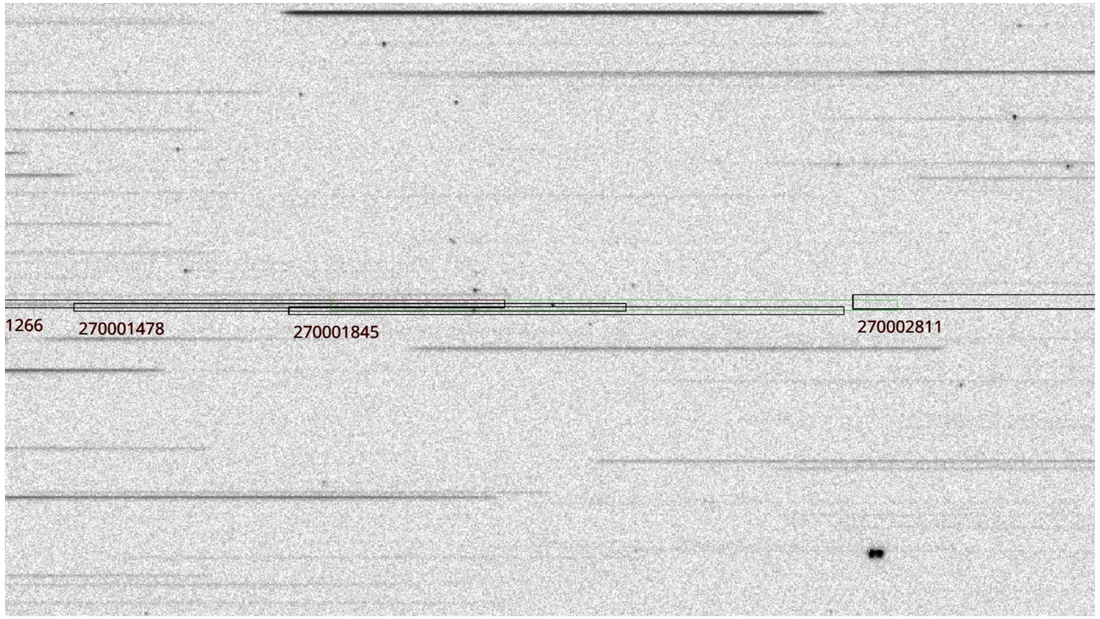
- These modules take as input the WSM images and catalogs, WSM spectra, and the calibration reference files from the CDP

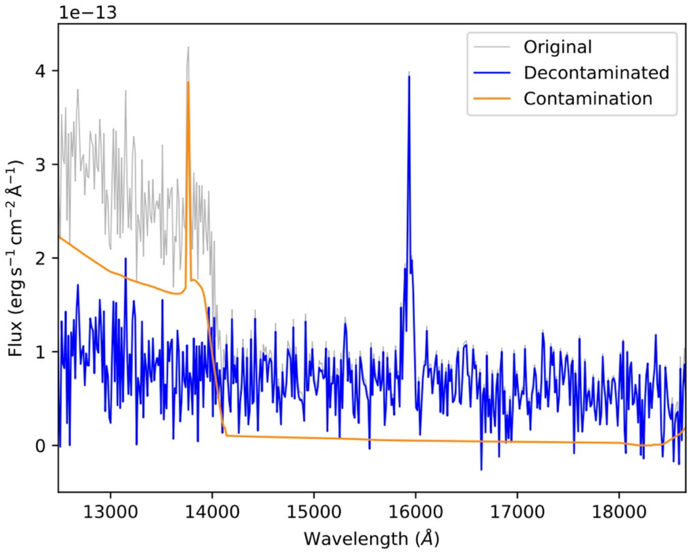
G2DP Details: Spectral Decontamination

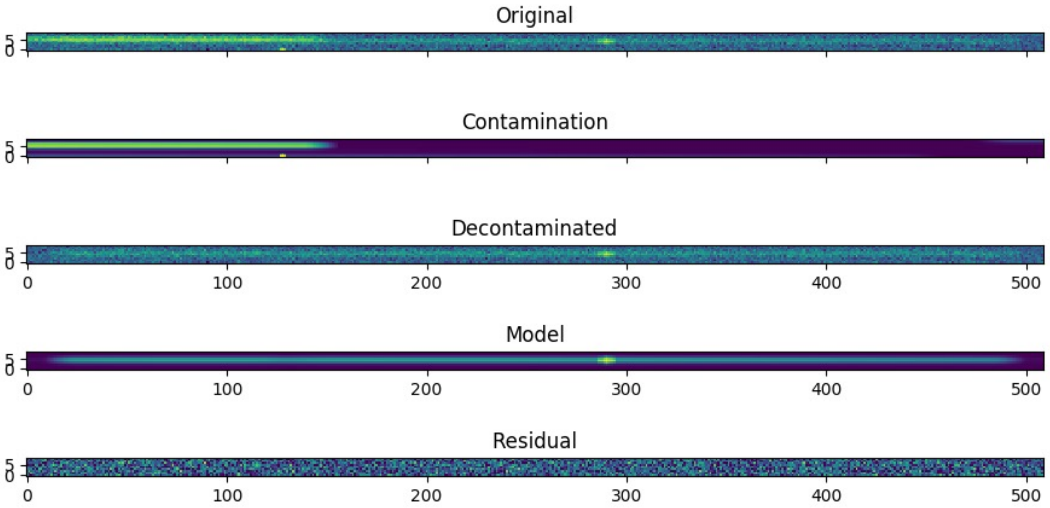


- *Spectral and imaging data are used to remove contaminating, overlapping spectra for each science target.*
- *Output is a clean, 2D spectrum for each detected source.*

Decontamination Example







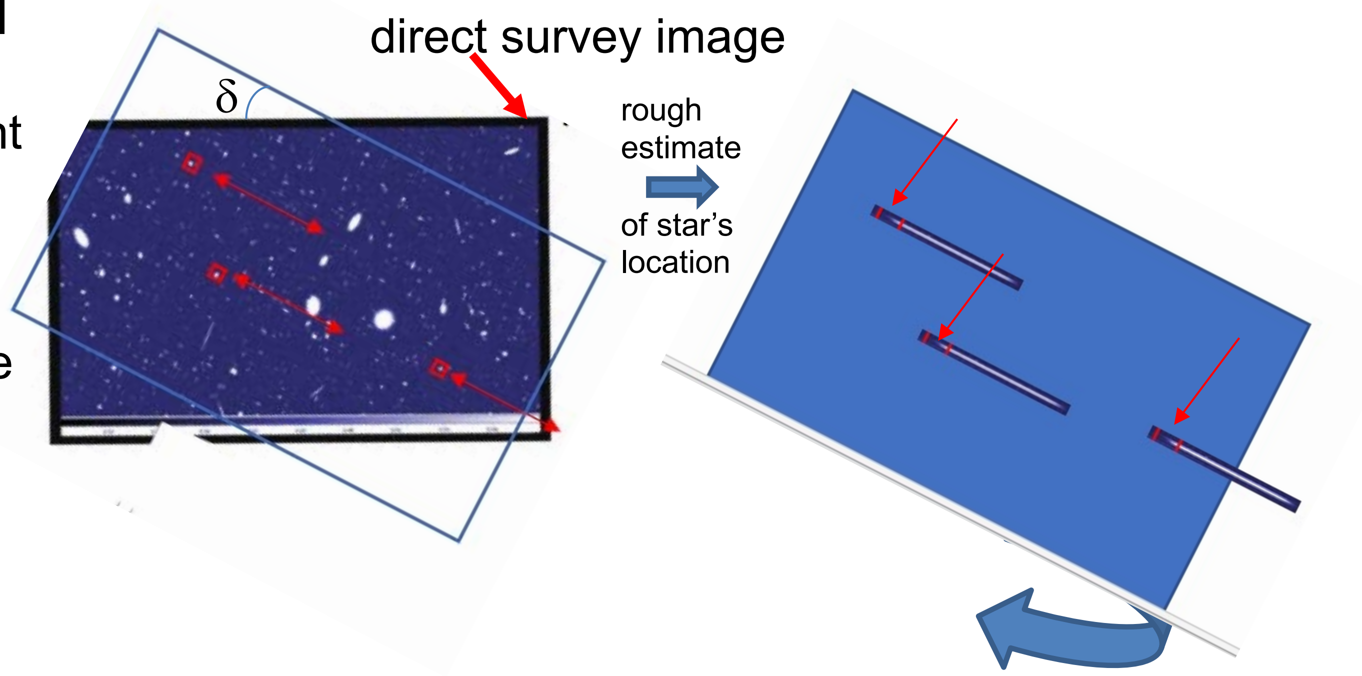
Decontamination example: (above) The faint, green box at the center of the simulated 2D spectral image shows the science target. Black boxes highlight contaminating (overlapping) sources. In the Roman pipeline, photometry from the imaging data and multiple roll angle spectra are used to build a model of each contaminating source (above right) and subtract them from the target spectra (below right). This example is courtesy of the Euclid NASA Science Center at IPAC (ENSCI).

G2DP Details: Image to Spectra mapping

The *CoordMapping* module will convert RA, Dec positions from the direct image catalog into spectra bounding boxes on the dispersed images (using image to spectral mapping derived in CDP).

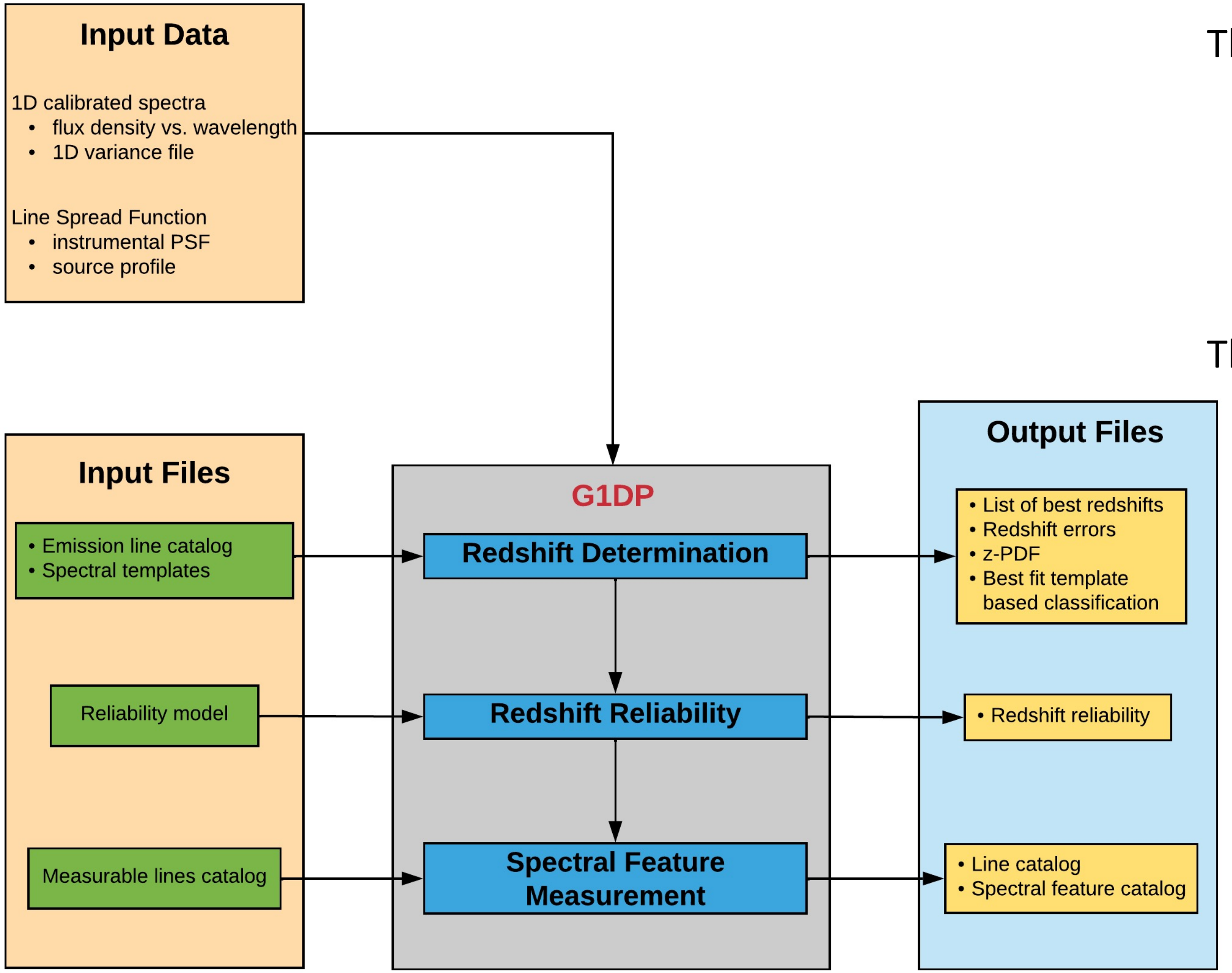
The *WCSUpdate* module will locate the approximate position of the blue/red edges for all bright point sources, accounting for the fact that the spectral images may be offset or at a different roll than the reference direct image data.

The module establishes the precise location of the blue/red edge, fitting the SED with a template stellar spectrum matched to the measured photometry and transmission curve. This produces a final set of (x, y) offsets for every bright, compact source.



From those offsets, we fit a Δx, Δy, ΔPA center offset for the entire field and generate an updated Spectra Location Table to be used throughout the rest of the G2DP pipeline.

G1DP Design and data flow



The individual pipeline modules in the G1DP:

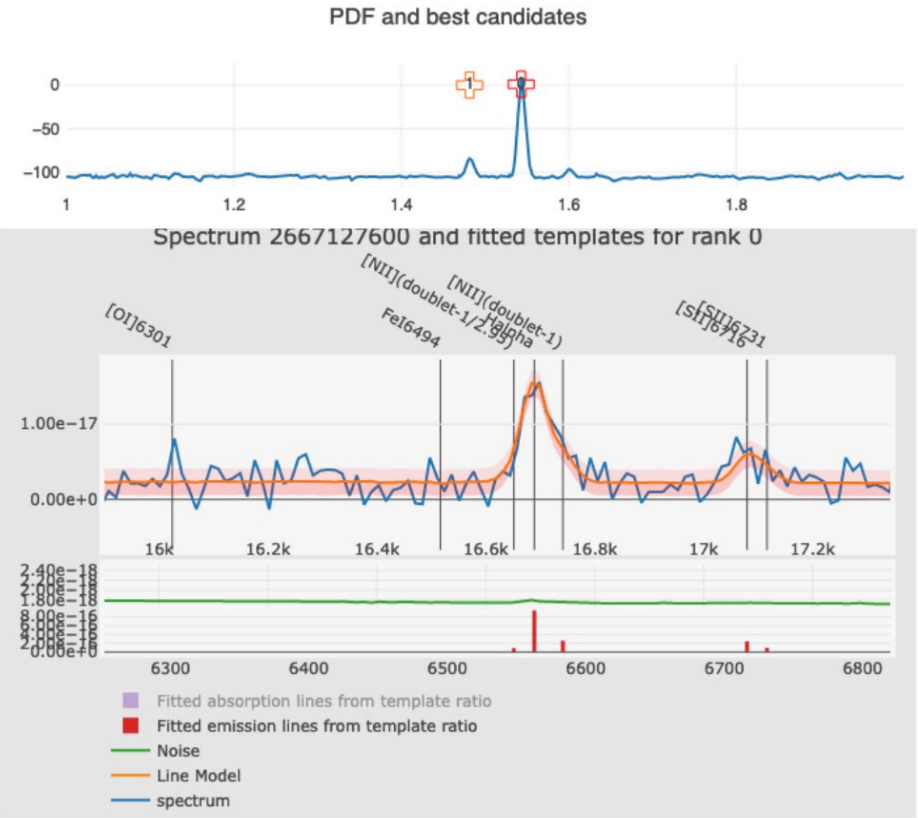
- ***Find Redshift****:* noise weighted, least squares fit of model templates to calibrated 1D spectra
- ***Redshift Reliability****:* determine reliability index (rank) of redshifts from zPDF
- ***Spectral Features****:* fit continuum and detected feature parameters

These modules produce:

- Best fit redshift and redshift probability distribution
- Redshift reliability
- Best fit galaxy template and list of identified features and feature fit parameters

G1DP Details: Redshift and template fitting

- The redshift measurement is based on a noise-weighted, least squares (2-pass) fit of a set of model templates to the calibrated, extracted 1D spectrum. The templates cover a wide range of source types, including:
  - Galaxies, QSOs, stars, type-1a SNe.
- The fitting also includes:
  - Nebular, emission line ratio templates
  - Continuum absorption (ISM, IGM)
  - An independent fit to Lyman alpha
  - Galactic ISM, extragalactic IGM and continuum absorption
- The G1DP is based on the Euclid SPE 1D spectral fitting software written by the LAM team and is being adapted at the SSC for Roman grism and prism data.

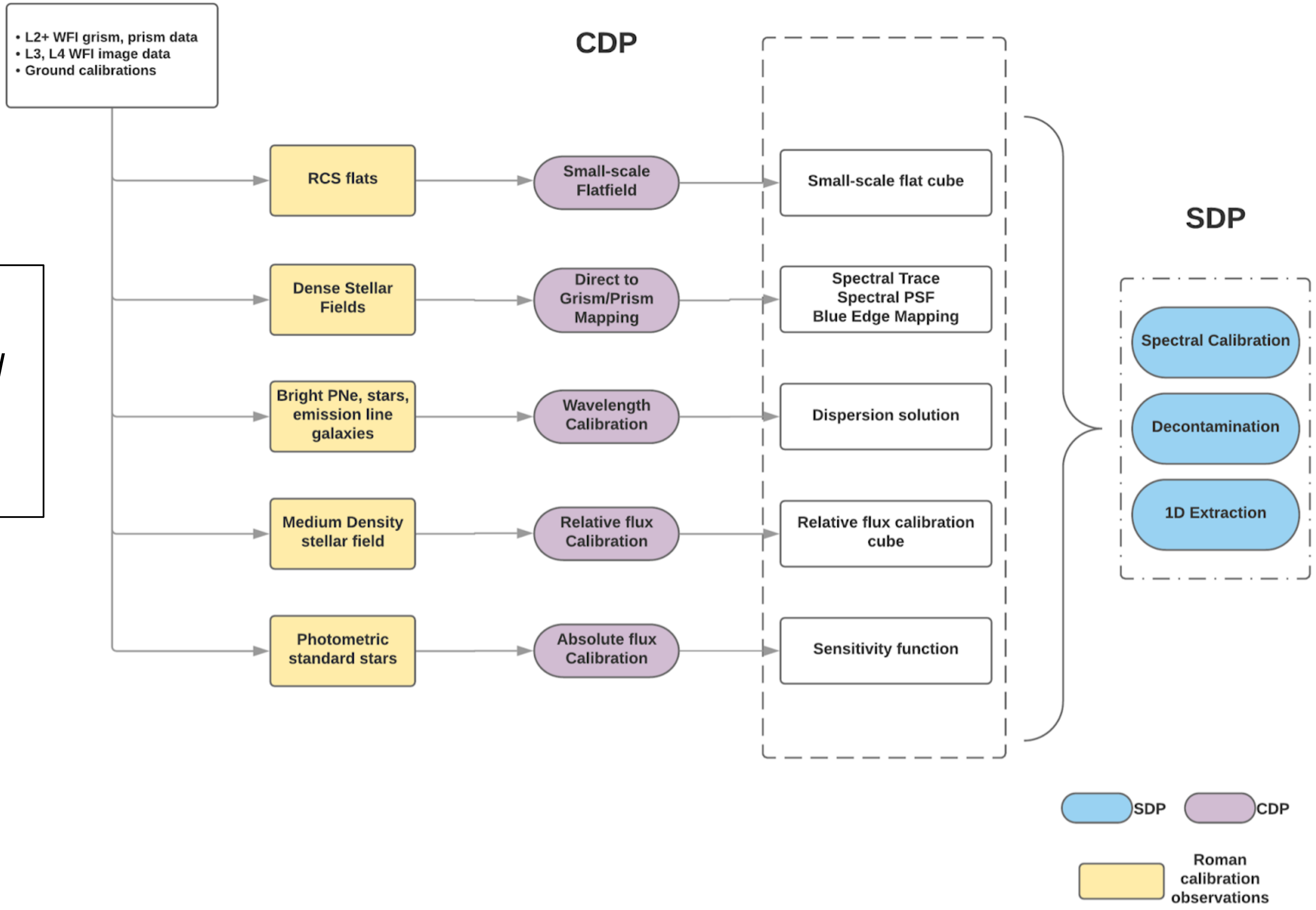


Example of redshift PDF visualization (top) and spectral fitting (bottom) in the 1D pipeline. In the bottom panel, the data are shown in blue and fit to the continuum and the spectral features (along with the uncertainty) are shown in red. The fit to the emission lines are used to generate the zPDF shown in the top plot, here with a strong peak in the solution at a redshift z~1.55. A weaker, secondary maximum is also seen at z~1.48.

Calibration Data Pipeline

Basic data flow for the CDP

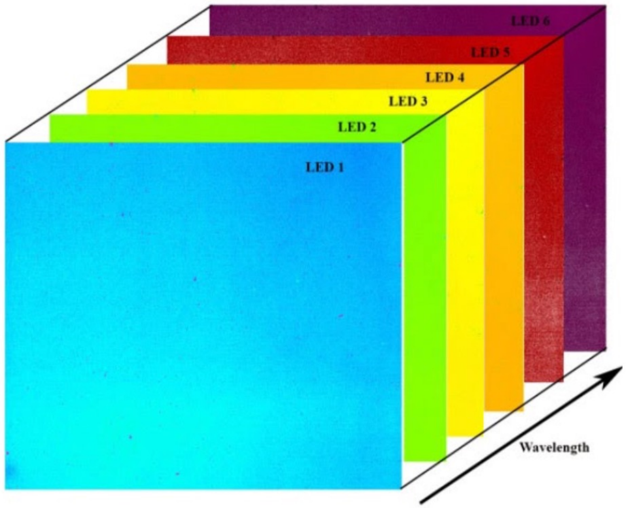
- *Calibration input from ground and in-flight data.*
- *Outputs are calibration reference files used in the SDP.*



Calibration Data Pipeline steps

- **Small Scale Spectral Flat Field**
  - Small scale: spectral cube built via wavelength interpolation of Relative Calibration System (RCS) flat field images of each detector. A pixel-level, detector flat.
- **Direct Image to Grism/Prism Dispersed Image Mapping**
  - Maps blue edge to image locations across each array.
  - 2D spectral traces (and PSF vs. wavelength) are derived from dithered observations of dense stellar fields.
- **Wavelength Calibration**
  - Derived via observations of bright PNe, absorption/emission line stars or emission line galaxies. Modified in-orbit from calibration derived from ground test data.
  - Dispersion solution referenced to blue edge of each spectrum.
- **Relative Flux Calibration** (a form of large-scale flat field)
  - Derived chromatically from repeated observations of one or more stellar "touchstone" fields. Includes effects of telescope optics and filters, sampled over larger scales (~30 arcsec) than the small-scale flat field.
  - Applied to 1D extracted science data based on the source position.
- **Absolute Flux calibration**
  - Derived from repeated observations of a set of ~14-16 mag, spectrophotometric standard stars.
  - Applied to 1D extracted science data.

CDP details: Small Scale Flat Field

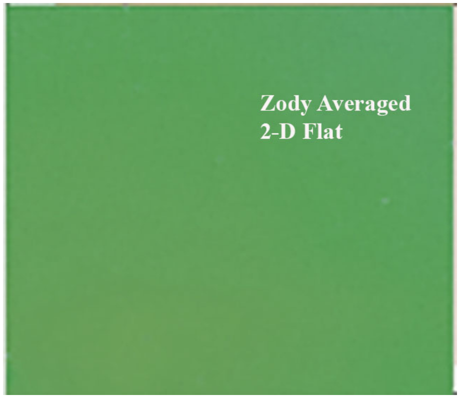


Monochromatic images are used to create a pixel-level, small-scale flat field cube for each detector. Each plane is initially normalized to unity.

Flats initially come from ground test data which is sampled finely in wavelength. After launch, we will use the RCS system to update the ground flats.

Zody Spectrum is used to weight each plane of the normalized cube at each wavelength.

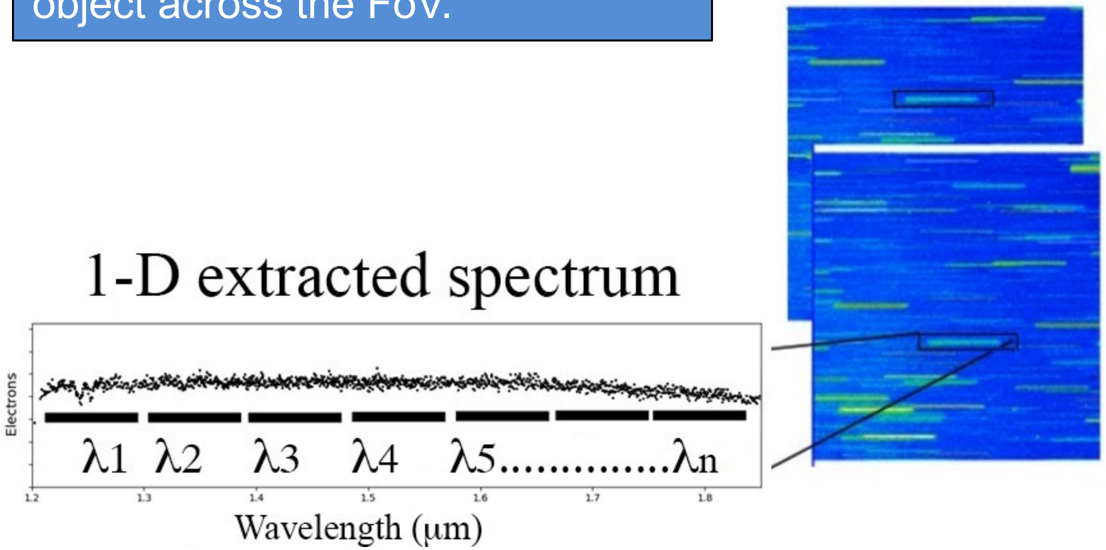
Compress weighted 3D cube to 2D flat field



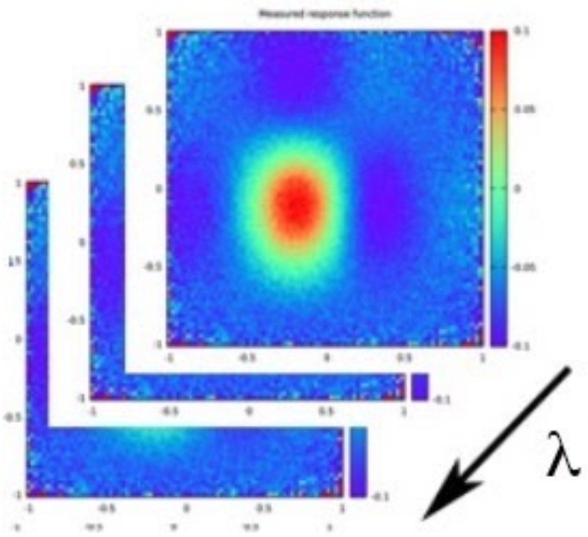
Final 2D flat is representative of response to Zody background light. There will be one image per detector.

CDP details: Relative Flux Calibration

Dithered, well-sampled 1D spectra from one or more “touchstone fields” are used to build up a relative, large-scale flat at each wavelength by comparing selected wavelength regions of a given object across the FoV.



Solve flux calibration matrix for each wavelength separately



and build cube in Field-of-View coordinates

Final relative flux calibration cube is a calibration product. It is used to provide a correction for every extracted 1D science spectrum based on field position.

Roman Data Level Definitions

Level 0	Science telemetry: Packetized data as it arrives from the spacecraft
Level 1	Uncalibrated exposures: include metadata, engineering data
Level 2	Calibrated exposures: remove detector signature, map to scene flux
Level 3	Resampled data: use rectified grid, coadd multiple exposures as appropriate
Level 4	High-level data: mostly source-oriented, include catalogs, extracted spectra, postage stamps, high fidelity flux, shape, morphology measurements
Level 5	Community-contributed products: may include any of the above types

19