Roberto Sartori
roberto.sartori@gmail.com

Weighted Batch Preprocessing

an improved version of PixInsight Batch Preprocessing

This project is the result of a collaboration with

Tommaso Rubechi

(www.tommasorubechi.it).

This document reports the motivations behind this extension and the technical changes made to the standard BatchPreprocessing script.

SUMMARY

Why this update?

Many astrophotographers don't use BatchPreprocessing (**BPP**), they prefer to run the whole stacking process manually in order to have full control of each step. Indeed, shrewd astrophotographers tune parameters at each step trying to get the best from each of it. Despite the effectiveness of this manual approach, Batch Preprocessing is capable of generating a master light image which overall quality is, in general, quite close if not undistinguishable from the manual process. Nevertheless, manual execution allows to deviate from a standard integration workflow interposing intermediate steps functional to obtaining a specific result. One additional step, commonly performed before the registration, consists in running the Subframe Selector script in order to measure light frame's FWHM, eccentricity and SNR to discard poor frames that do not match some acceptance criteria (by requiring for example that FWHM must be lower than a given value) and to generate a meaningful weight for the remaining to be used during the Image Integration process. Generally speaking, the objective is to improve one or a combination of the mentioned measurements on the final master light.

The purpose of this script is embedding this extra step into the current BPP in order to cover the most common missing gap between manual and batched stacking process.

This project is a collaboration with **Tommaso Rubechi** which is deeply involved as a supervisor responsible of defining the weighting presets and being the principal tester and main results analyst.

We named this script **Weighted Batch Preprocessing** - **WBPP** in order to highlight its capability of automating the mentioned light frame weighting.

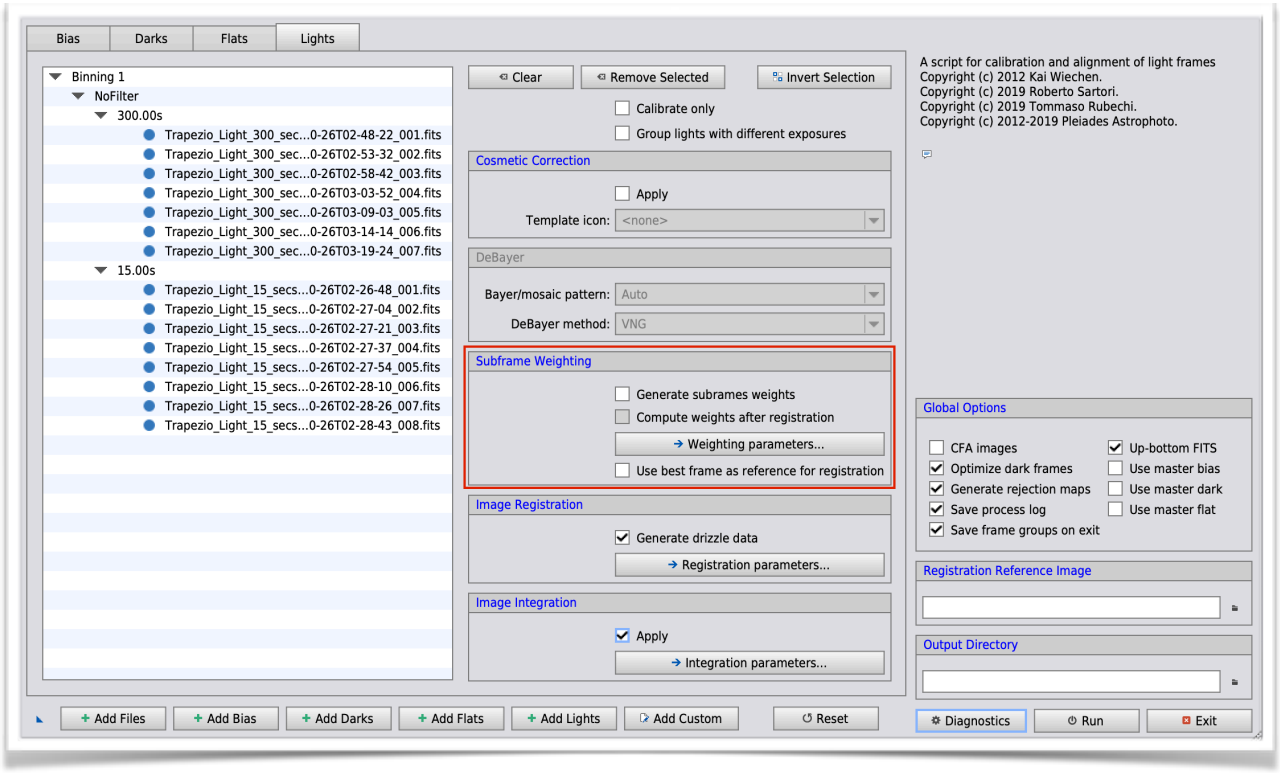
New Features

On top of the main weighting capability, several new features has been implemented welcoming some meaningful feedbacks and nice-to-have features pulled by the testers. The full features list is the following:

- automatic light frames weight computation based on a combined contribution of FWHM, eccentricity and SNR measurements
- automatic selection of the best reference frame for registration
- extended light frame grouping
- smart calibration of light frames
- new "auto" rejection algorithm option
- smart file path naming
- enhanced diagnostic
- smart reporting
- new "Save frame groups on exit" option

Each feature is described and detailed in the following chapter.

Light Frames Weighting and Best Reference Frame Selection



These two features work in synergy. Both are based on the measurement of the so-called **image descriptors**: FWHM, eccentricity and SNR. The JS code that performs these measurements has been extracted directly from the SubframeSelector script in order to precisely provide the same measurements.

The weighting formula

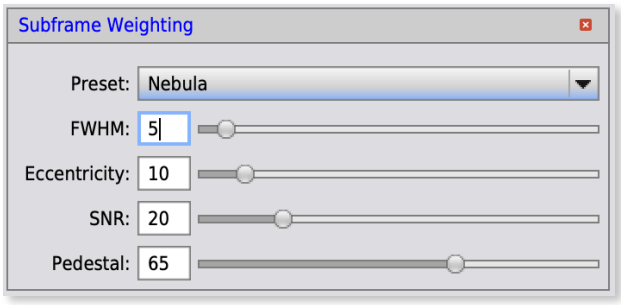
Light Frame Weighting is a step executed for each group of light frames. The principle is to firstly compute light frames descriptors and extract FWHM, eccentricity and SNR for each frame along with min/max values within the whole group. Once the analysis is completed each weight is computed using the well known formula:

$$w = a\cdot\left(1 - \frac{FWHM - FWHM_{min}}{FWHM_{max} - FWHM_{min}}\right) + b\cdot\left(1 - \frac{ecc - ecc_{min}}{ecc_{max} - ecc_{min}}\right) + c\cdot\left(\frac{SNR - SNR_{min}}{SNR_{max} - SNR_{min}}\right) + P$$

where **w** is the final light frame weight, **(a, b, c)** are tunable coefficients that balance the FWHM, eccentricity and SNR contributions and and **P** is the tunable pedestal of the weight.

Usually, descriptors are computed on light frames **before** their registration; nevertheless, we decided to provide the option to postpone the measurement **after** the registration by flagging the new option "*Compute weights after registration*", in case the user would take into account the unavoidable descriptors value changes of a registered frame.

The following image shows the interface to be used to assign descriptor's coefficients and pedestal value. The more a contribution is high, the higher the correlation between this single descriptor and the total image weight should be expected. To simplify the job, we provided three templates that works well in few general cases, each of which is addressing a specific type of target: **Nebula** (which weights more SNR because of an expected presence of fainter details), **Globular Cluster** (which obviously weights more the star properties since stars are the kind of objects we're willing to highlight) and **galaxy** (which sets a compromise in between SNR and stars quality).



Best frame selection

The selection of the best reference frame works on the basis of finding the light frame with the "most suitable stars" to be a reference for the others i.e. the light frame with the smallest and roundest stars. The reference frame is always selected among the whole set of light frames provided despite their grouping. Such that, the strategy is to measure the

descriptors for the whole set of light frames and pick the frame with the highest weight that equally combines FWHM and eccentricity. The weighting function used in this step, thus, adopts the following coefficients:

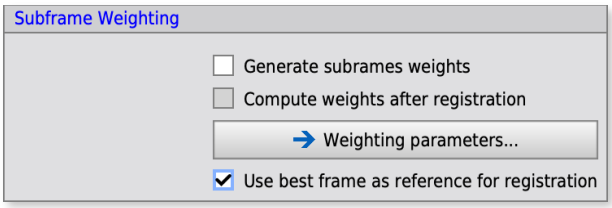
(a: 50, b: 50, c: 0, P: 1)

The best reference frame selection must be performed before the registration in order to provide the reference frame to that next step. Because of that, all the light frame groups need to be calibrated, cosmetic corrected and debayered first and then, once the whole set of debayered images are measured and best frame is chosen, the registration and integration steps can continue. This change required to restructure the sequence of operations performed by the standard BPP as explained in the following section.

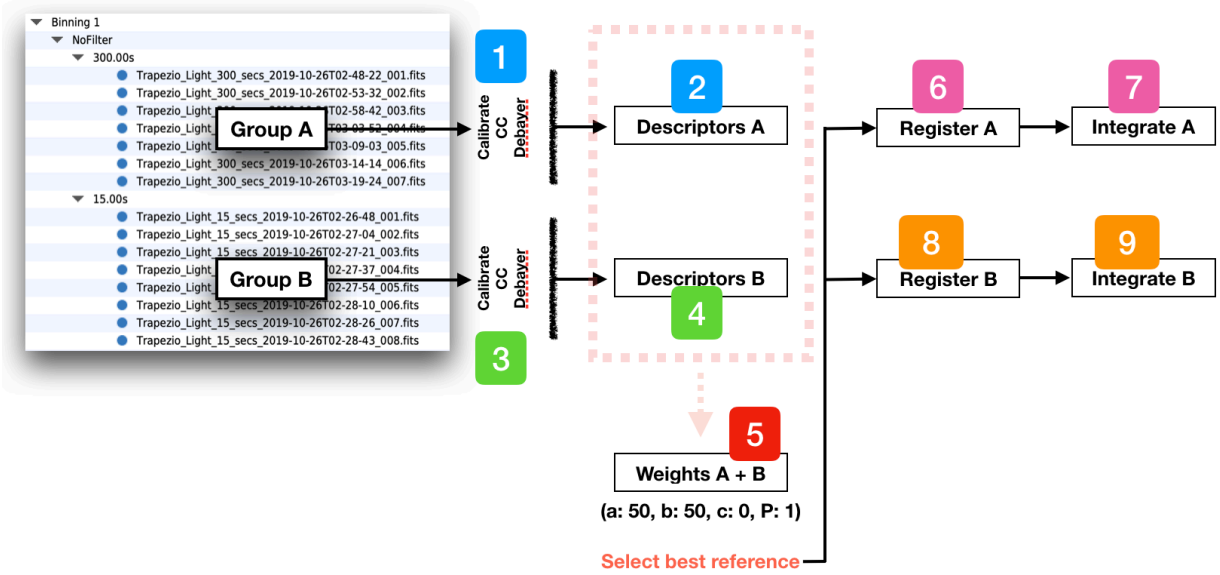
A new light frames processing flow

While the standard BPP processes each group one by one executing the entire processing flow from the initial calibration to the final integration (with the precedence of the group that contains the reference frame) the WBPP needs to work differently, pre processing the whole light frames with calibration, cosmetic correction and debayer to subsequently select the best frame among the whole set of pre processed images. Moreover, depending on the settings, light frame weights needs to be computed before or after the registration step. As a consequence, the processing flow of WBPP needs to be different and follows a sequence that may change depending on the configuration selected by the user.

All the possible cases are detailed below taking as an example the integration of two groups of light frames with different exposures.



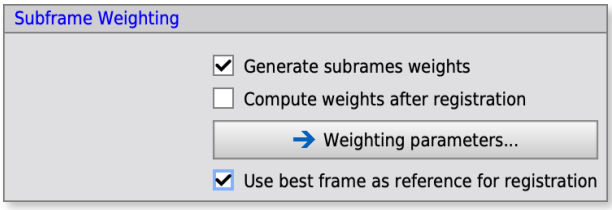
Auto-selection of reference frame, no weights computed



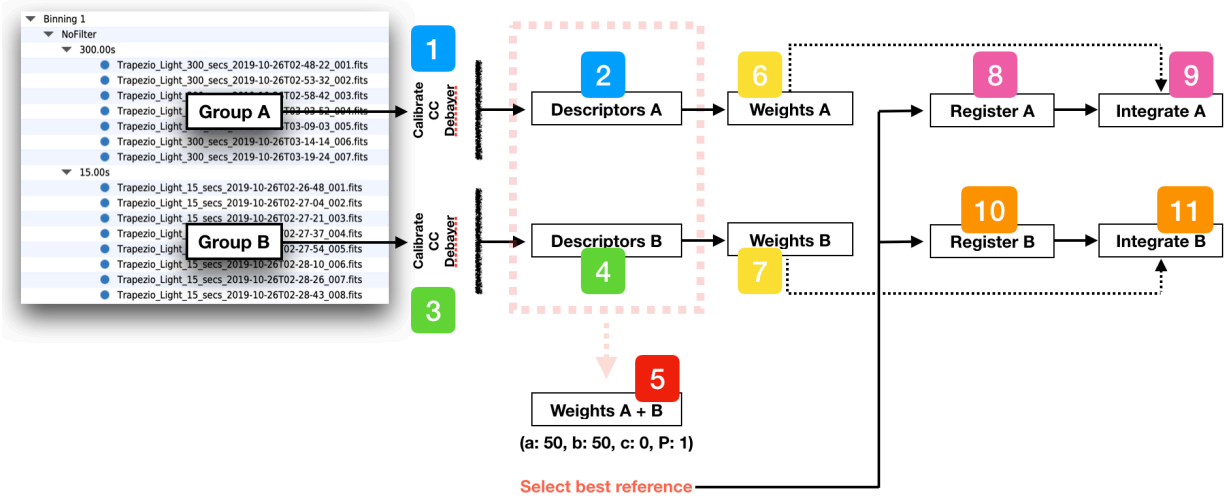
Every light frame group is calibrated, cosmetic corrected and debayered and frames descriptors are measured (step **1-2**, **3-4**).

Subsequently, the descriptors of all groups are grouped together and weights of all light frames are computed taking into account only stars properties. The best reference frame is the one with the highest weight (step **5**).

The process continues registering and integrating each light frame group (step **6-7**, **8-9**) using as a reference the light frame previously selected.



Auto-selection of reference frame, weights computed before registration

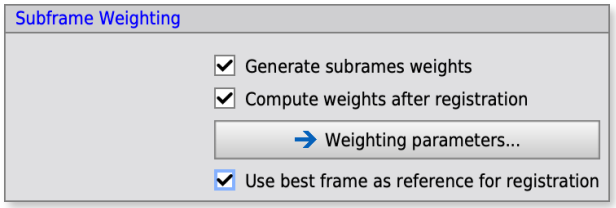


Every light frame group is calibrated, cosmetic corrected and debayered and frames descriptors are measured (step **1-2**, **3-4**).

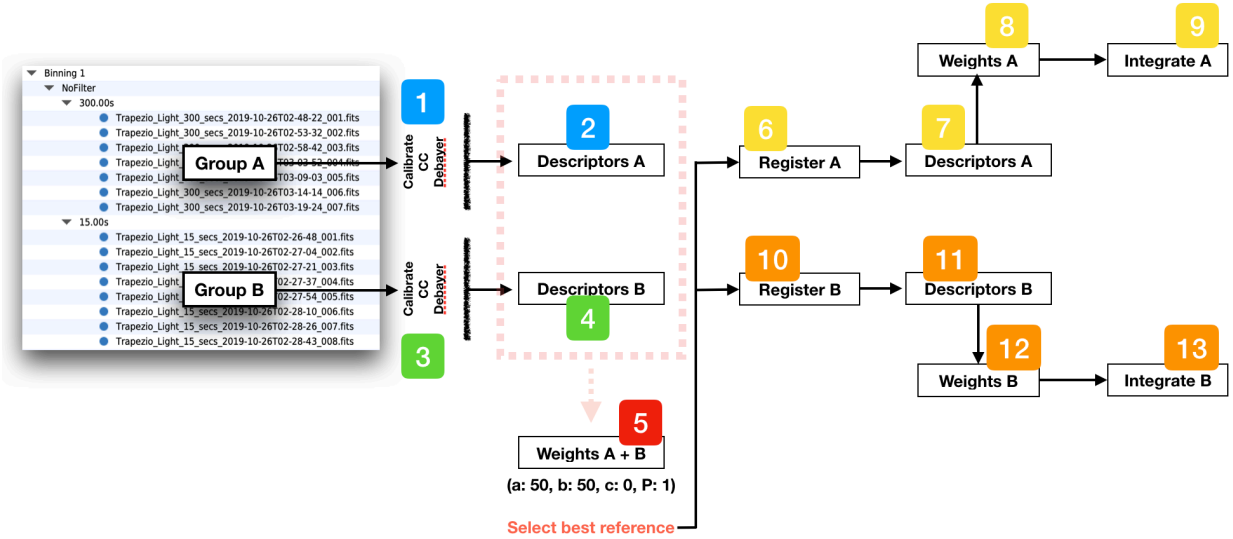
Subsequently, the descriptors of all groups are grouped together and weights of all light frames are computed taking into account only stars properties. The best reference frame is the one with the highest weight (step **5**).

Then, weights are computed for each group of light frames (step **6-7**).

The process continues registering and integrating each light frame group (step **8-9**, **10-11**) using as a reference the light frame previously selected and taking into account the previously computed weights.



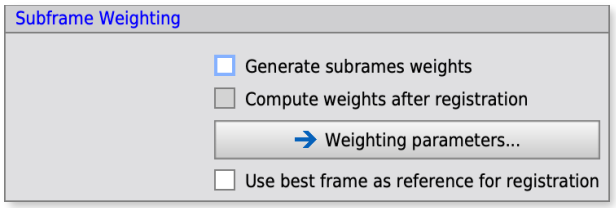
Auto-selection of reference frame, weights computed after registration



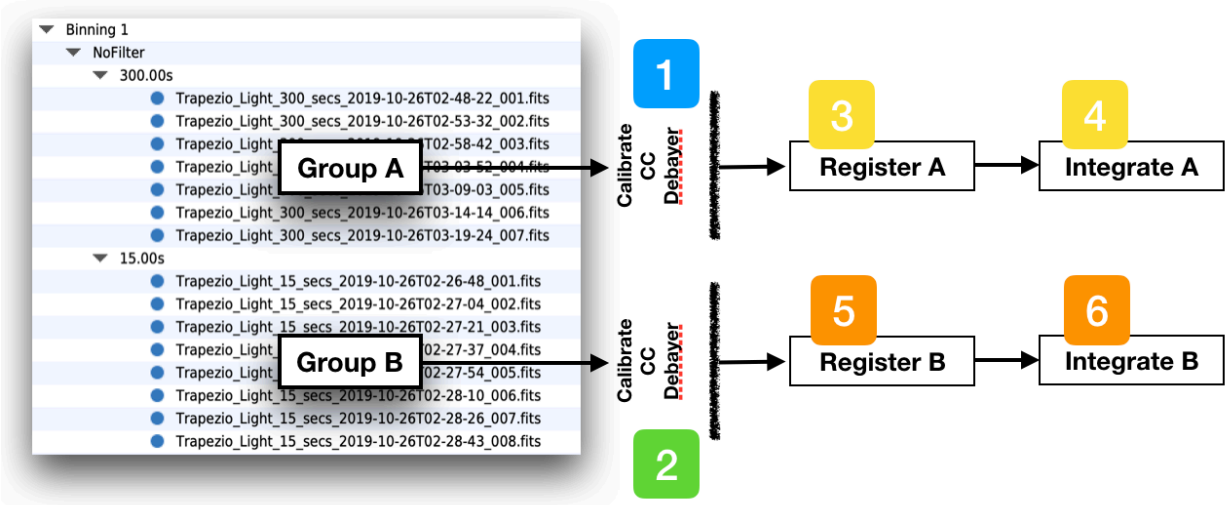
Every light frame group is calibrated, cosmetic corrected and debayered and frames descriptors are measured (step **1-2**, **3-4**).

Subsequently, the descriptors of all groups are grouped together and weights of all light frames are computed taking into account only stars properties. The best reference frame is the one with the highest weight (step **5**).

The process continues by registering, computing descriptors and weights of registered images and integrating each light frame group (step **6-7-8-9**, **10-11-12-13**) using as a reference the light frame previously selected.

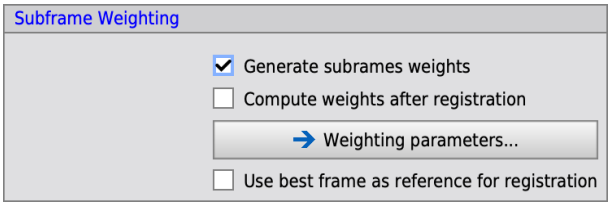


NO Auto-selection of reference frame, no weights computed

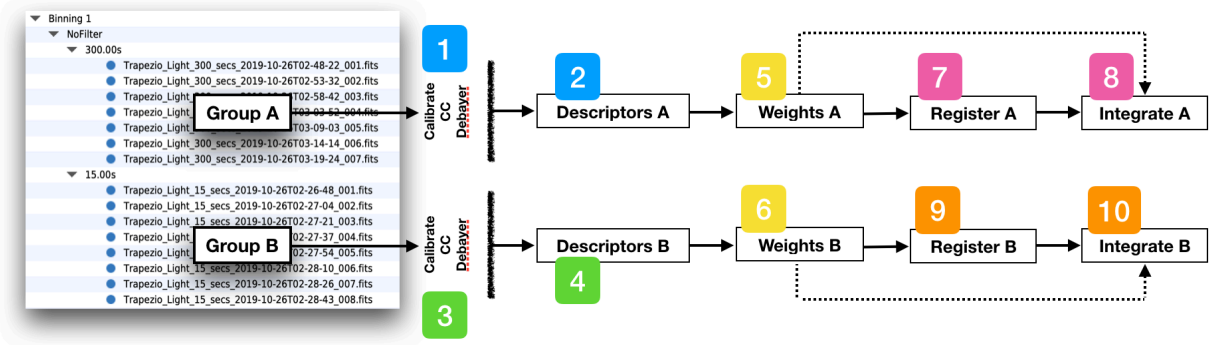


Every light frame group is calibrated, cosmetic corrected and debayered (step **1**, **2**).

The process continues registering and integrating each light frame group (step **3-4**, **5-6**).



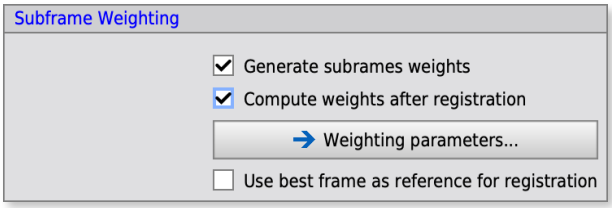
NO Auto-selection of reference frame, weights computed before registration



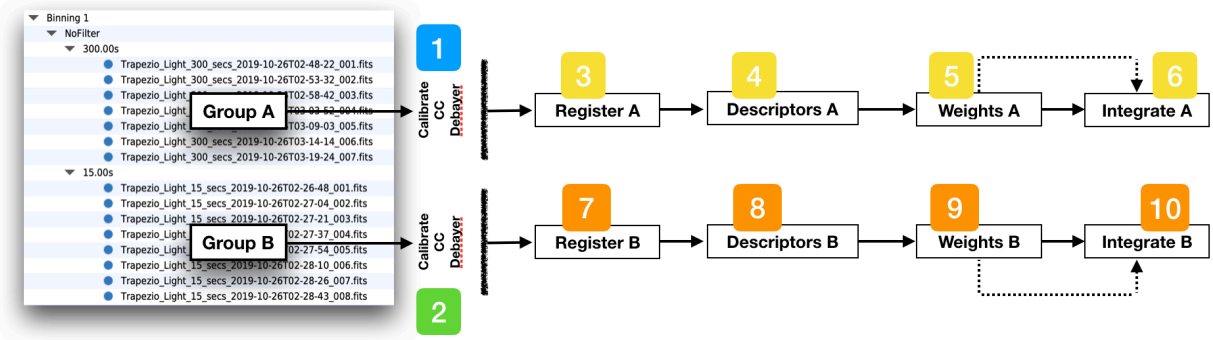
Every light frame group is calibrated, cosmetic corrected and debayered and frames descriptors are measured (step **1-2**, **3-4**).

Then, weights are computed for each group of light frames (step **5-6**).

The process continues registering and integrating each light frame group (step **7-8**, **9-10**) taking into account the previously computed weights.



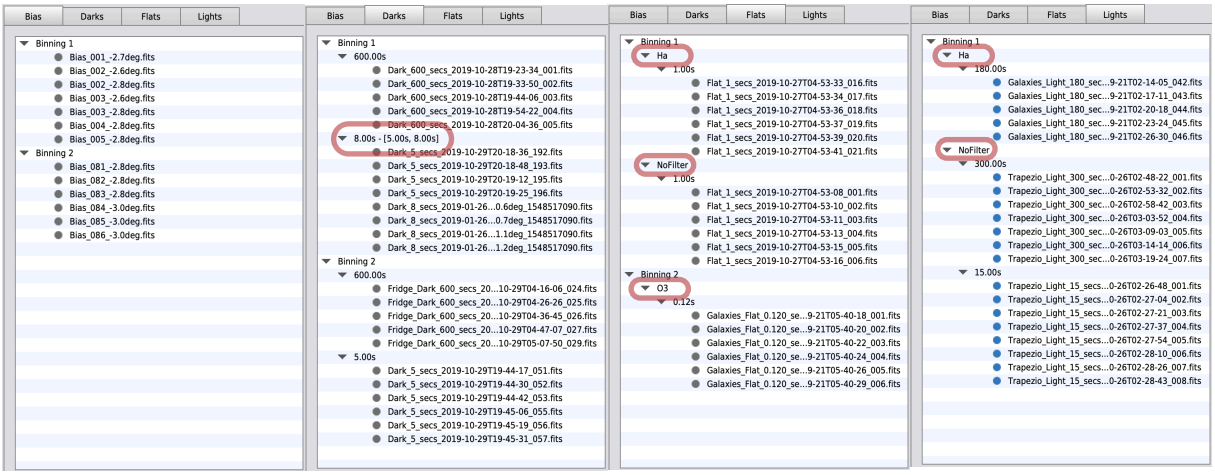
NO Auto-selection of reference frame, weights computed after registration



Every light frame group is calibrated, cosmetic corrected and debayered (step **1**, **2**).

The process continues registering, computing descriptors and weights of registered images and integrating each light frame group (step **3-4-5-6**, **7-8-9-10**).

New frames grouping



Light frame grouping has been improved in order to take into account not only **BINNING** and **filter name** but also the **exposure**. This approach has the advantage of making WBPP capable of integrating an HDR session of images with different exposures in a single run. Frame grouping is hierarchical and it works with the following rules:

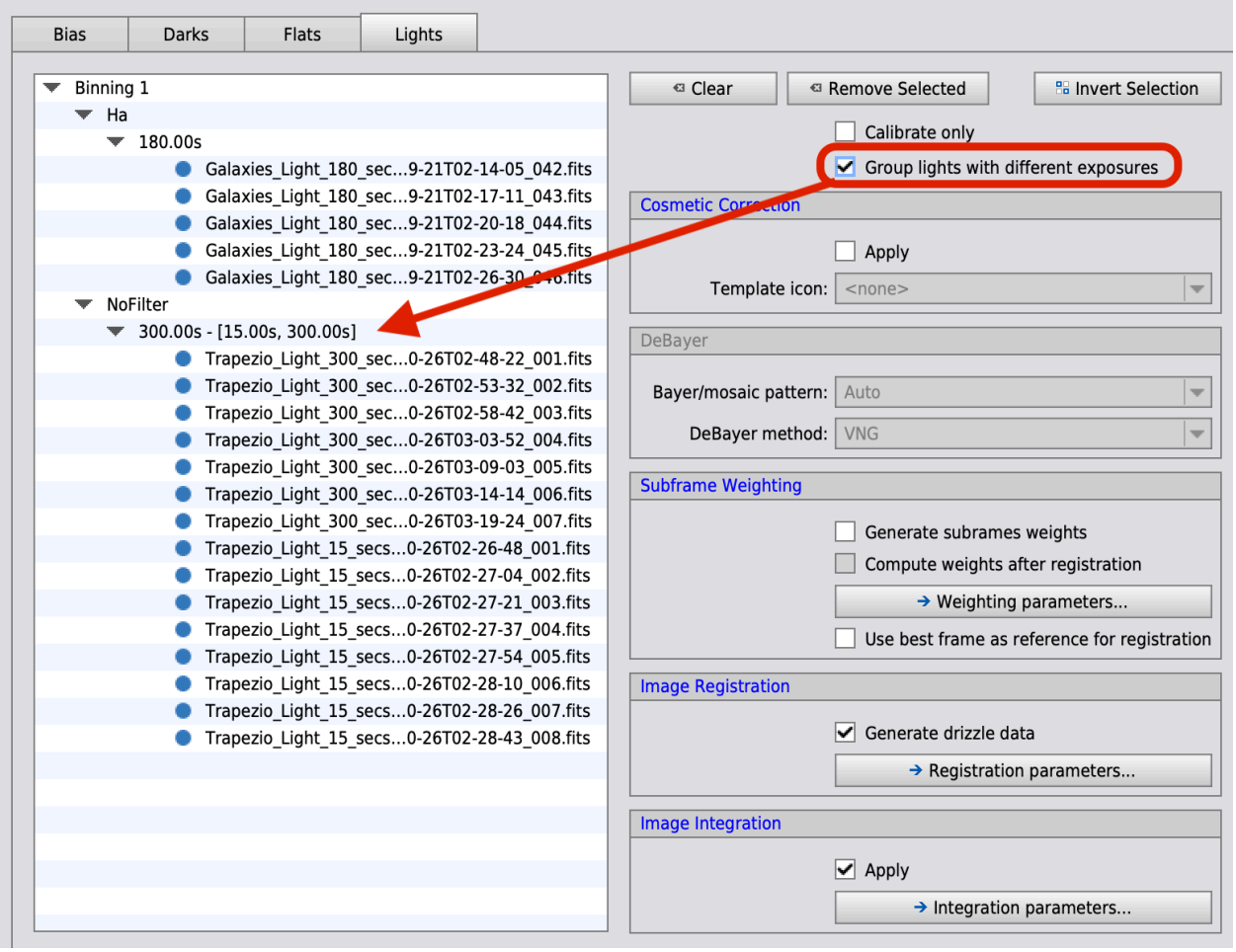
BIAS: grouped by BINNING (ascendant).

DARK: grouped by BINNING (ascendant)-> EXPOSURE (descendant). By setting the exposure tolerance, dark frames with different exposures are grouped together and the exposure of each group is assigned, by convention, to the highest exposure of its frames. The new WBPP shows the exposure assigned to the group and the set of the different exposures of its frames using the following notation:

MAIN EXPOSURE - [exposure 1, exposure 2, exposure 3]

FLAT: grouped by BINNING (ascendant) -> FILTER (alphabetical) -> EXPOSURE (descendant). Frames will be grouped if they exactly match all three parameters (given a tolerance on the exposure of 10 milliseconds)

LIGHTS: grouped by BINNING (ascendant) -> FILTER (alphabetical) -> EXPOSURE (descendant). Light frames are be grouped if they exactly match all three parameters (given a tolerance on the exposure of 10 milliseconds). If the new option "*Group lights with different exposure*" is selected then light frames will be grouped only by BNINNNG (ascendant) -> FILTER (alphabetical): light frames with different exposures will be grouped together and group exposure will be shown adopting the same notation as for dark frames. The example is shown below:



Moreover, as mentioned, a **tolerance of 10 milliseconds** is always taken into account when grouping by exposure since some softwares like APT do not save the exposure precisely and a small difference between light frames of about one / two milliseconds has been found during out tests.

Smart Calibration

Calibration in BPP works by finding the master dark with the closest exposure to the light frame group exposure. Since WBPP allows to group light frames with different durations together, we need to make this matching smarter. In WBPP the calibration is called “smart” since it does not rely on frame group exposure anymore but, differently, on single light frame duration only.

To select the calibrating master files, WBPP implements the following strategies:

	Master Calibration Files		
Groups	**master BIAS**	**master DARK**	**master FLAT**
FLAT FRAMES	use the master bias with the same binning	use the master with the closest exposure. If “calibrate with flat dark only” is flagged then exposure cannot differ more than = 0.5s otherwise no master dark are be used.	
LIGHT FRAMES	use the master bias with the same binning	use master dark with same binning and closest exposure	use master flat with same binning and same filter name

When calibration is performed on a light frame group, light frames are always **sub-grouped by equal exposure** and the most suitable master dark frame is chosen accordingly. This is mandatory to properly calibrate each light frame with the master dark that best matches each light frame exposure.

New “auto” rejection option

Pixel rejection strategy is, in general, selected depending on the number of frames to integrate. Standard BPP allows to set only one pixel rejection algorithm for each image type (bias, dark, flat and light). Considering that light frames can be grouped by filter and exposure, several groups of different size can be integrated each of which should require a proper rejection algorithm depending on its size. This new option solves this problem by letting WBPP to select the most suitable rejection algorithm for each group separately.

The auto selection adopts the following criteria:

N < **8**: Percentile Clip

N < **11**: Average Sigma Clip

N < **20**: Winsorized Sigma Clipping

N < **25**: Linear Fit Clipping

N >= **25**: Generalized Extreme Studentized Deviate

Smart Naming

In order to group light frames by BINNING, EXPOSURE and FILTER, WBPP needs to identify these three properties for each file.

The correct approach is to get their values from the FITS file header by reading the associated keywords. Anyway, some (if not all) of these information could be missing from the header file.

Smart naming provides the possibility for the user to properly rename folders and/or file names in order to embed these information into the file path letting WBPP to retrive them in case they would be missing from the file’s metadata.

It’s worth to highlight that the smart naming **does not override the image FITS data** so any value available from the header file will not change despite any file/folder naming convention.

Naming conventions are detailed below.

Binning

Binning is extracted by searching one keyword among **XBINNING**, **CCDBINX** or **BINNING** (case insensitive) followed by a space, an underline or a dash character and next a sequence of numerical digits.

Examples are :

FILE PATH	DETECTED BINNING
/Users/ME/2019-09-21_Galaxies/Dark **xbinning 1**/dark_01.fits	**1**
/Users/ME/2019-09-21_Galaxies/Dark **BINNING 04**/dark_02.fits	**4**
/Users/ME/2019-09-21_Galaxies/DarkS/dark_**ccdbinx_1**_078.fits	**1**
/Users/ME/2019-09-21_Galaxies/DarkS/master_dark_**XBINNING-2**.fits	**2**

Filter

Filter is extracted by searching one keyword among **FILTER** or **INSFLNAM** (case insensitive) followed by a space, an underline or a dash character and next a sequence of alphanumeric digits.

Examples are:

FILE PATH	FILTER
/Users/ME/NGC2903/FLAT **FILTER Halpha**/flat_01.fits	**Halpha**
/Users/ME/NGC2903/LIGHT **INSFLNAM O3**/light_02.fits	**O3**
/Users/ME/NGC2903/LIGHT/light_**FILTER-SII**_078.fits	**SII**
/Users/ME/NGC2903/FLAT/master_flat_**ISFLNAM-UHCS**.fits	**UHCS**

Exposure

Exposure is extracted in several formats:

1. by searching one keyword among **EXPTIME** or **EXPOSURE** (case insensitive) followed by a space, an underline or a dash character and next a valid integer or float number representation.
2. by searching a valid integer or float number representation not preceded by any alphabetic character and followed by **s**, **sec** or **_secs** if the previous step failed.

Examples are:

FILE PATH	EXPOSURE
/Users/ME/NGC2903/LIGHT **180s**/flat_01.fits	**180 sec**
/Users/ME/NGC2903/LIGHT **2.4_secs**/light_02.fits	**2.4 sec**
/Users/ME/NGC2903/LIGHT/light_**EXPTIME-60**_078.fits	**60 sec**
/Users/ME/NGC2903/FLAT/master_flat_**EXPOSURE-20.0**.fits	**20 sec**

Extended diagnostic

There are many scenarios where an improper set of images could generate a master light with issues. For this reason, the diagnostic has been extended in order to warn the used from the very beginning about potential issues and configuration errors. In particular, the new issues detected and reported by WBPP are:

Generic configurations

- **a group has less than 3 frames**: this is a blocking error since ImageIntegration needs at least 3 frames to proceed. WBPP will not continue.
- **a group contains CFA images but "CFA Images" flag is not set**: this is an ignorable warning. CFA images, in general, needs always to be debayered so the flag bust be selected.

Flat frames checks

• **neither matching master bias nor master dark flat found**: this is an ignorable warning reported when neither suitable master bias nor master dark have been found to calibrate flat frames. The master flat frame will remain entirely uncalibrated. This could be the way the user decided to proceed otherwise he should investigate why there is a binning and/or exposure mismatching between flat groups and bias/dark groups.

• **matching master bias but no matching master dark found**: this is an ignorable warning. In this case flat frames are calibrated with master bias only and no master dark will be used.

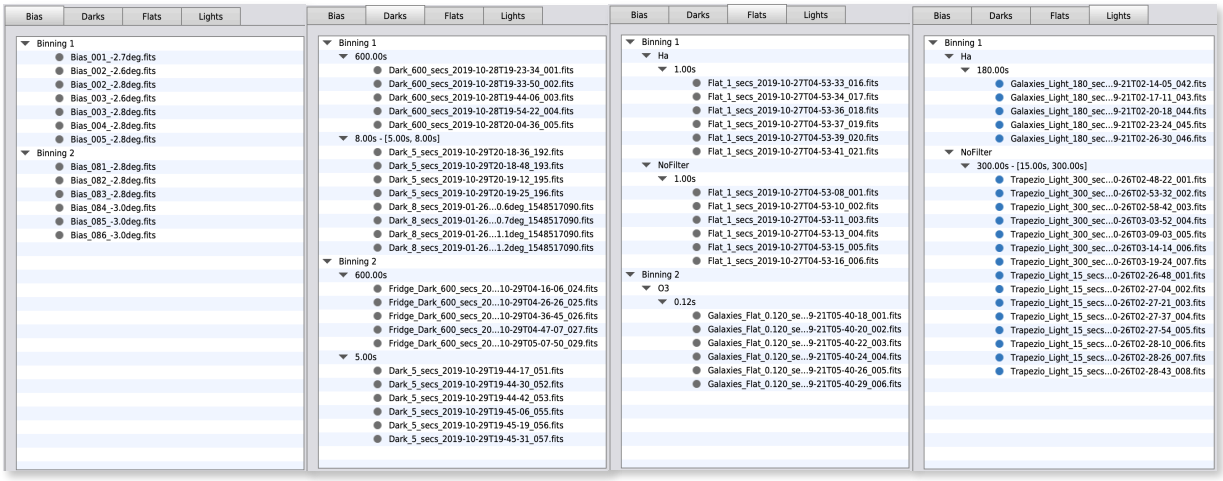
• **selected master dark frame has different exposure**: this is an ignorable warning reporting that the master dark used to calibrate flat frames has a different exposure time. As a double check, the user will be informed if “Optimize dark frames” flag is activated or not. Depending on the case, the user is free to proceed, to stop and activate the “Optimize dark frames” flag or to stop and investigate on the reason of this binning and/or exposure mismatching.

Light frames checks

• **selected master dark frame has different exposure**: this is an ignorable warning reporting that the master dark used to calibrate a subgroup of light frames has a different exposure time. As a double check, the user will be informed if “Optimize dark frames” flag is activated or not. Depending on the case, the user is free to proceed, to stop and activate the “Optimize dark frames” flag or to stop and investigate on the reason of this binning and/or exposure mismatching.

• **no master flat frame has been found**: this is an ignorable warning reporting that WBPP did not find any suitable master flat frame to calibrate light frames. The user is free to proceed (for example i f no flat frames have been provided) or stop and investigate on the reason why the provided master flat mismatches the mentioned light frame group.

As an example, given the following set of frames:



We assume that “Calibrate with flat dark only” is selected for Flats so only a perfect matching master dark can be selected to calibrate flat frames. WBPP will try to match the following files:

1. master Bias
   1.1. BINNING 1
   1.2. BINNING 2
2. master Dark
   2.1. BINNING 1, 600 sec
   2.2. BINNING 1, 8 sec
   2.3. BINNING 2, 600 sec
   2.4. BINNING 2, 5 sec
3. master Flat
   3.1. BINNING 1, Ha, 1 sec: **no master dark matching 1 sec exposure is provided**
   3.2. BINNING 1, NoFilter, 1 sec: **no master dark matching 1 sec exposure is provided**
   3.3.BINNING 2, O3, 0.12 sec: **no master dark matching 0.12 sec exposure is provided**
4. master Light

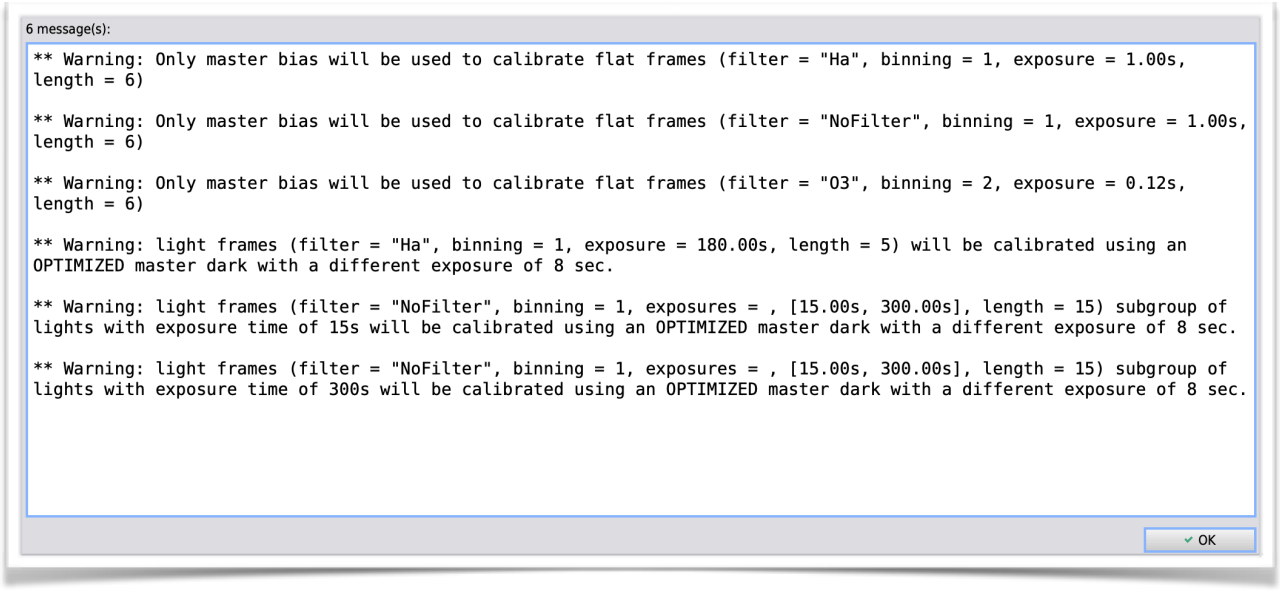
4.1.BINNING 1, Ha, 180 sec: master flat of (BINNING 1, Ha) is found but **master dark with BINNING 1 and 8 sec is selected but exposure is different**

4.2. BINNING 1, Ha, 300 sec: this group is divided into two subgroups

4.2.1. BINNING 1, Ha, 15 sec: master flat of (BINNING 1, Ha) is found but **master dark with BINNING 1 and 8 sec is selected but exposure is different**

4.2.2. BINNING 1, Ha, 300 sec: master flat of (BINNING 1, Ha) is found but **master dark with BINNING 1 and 8 sec is selected but exposure is different**

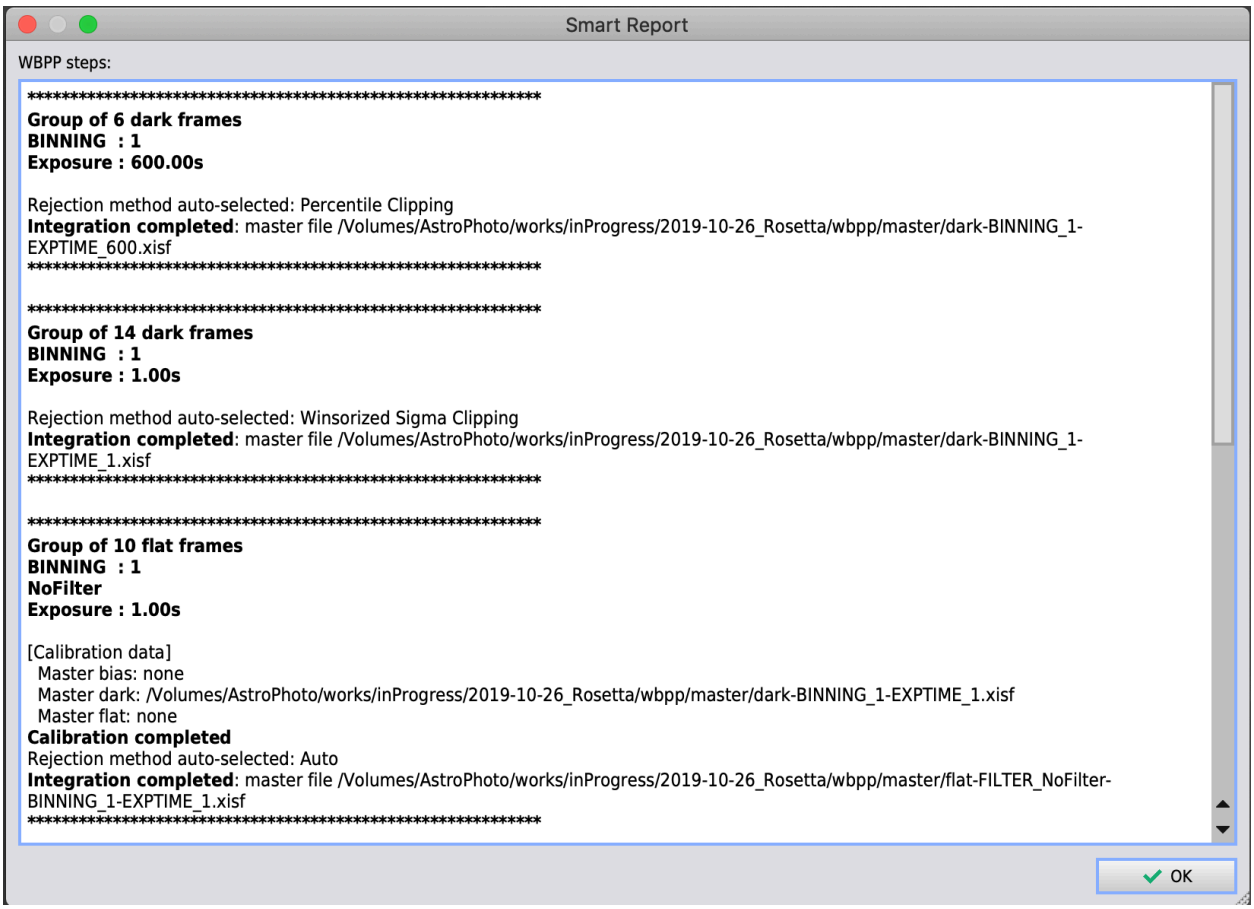
All these warnings are reported in the extended diagnostics.

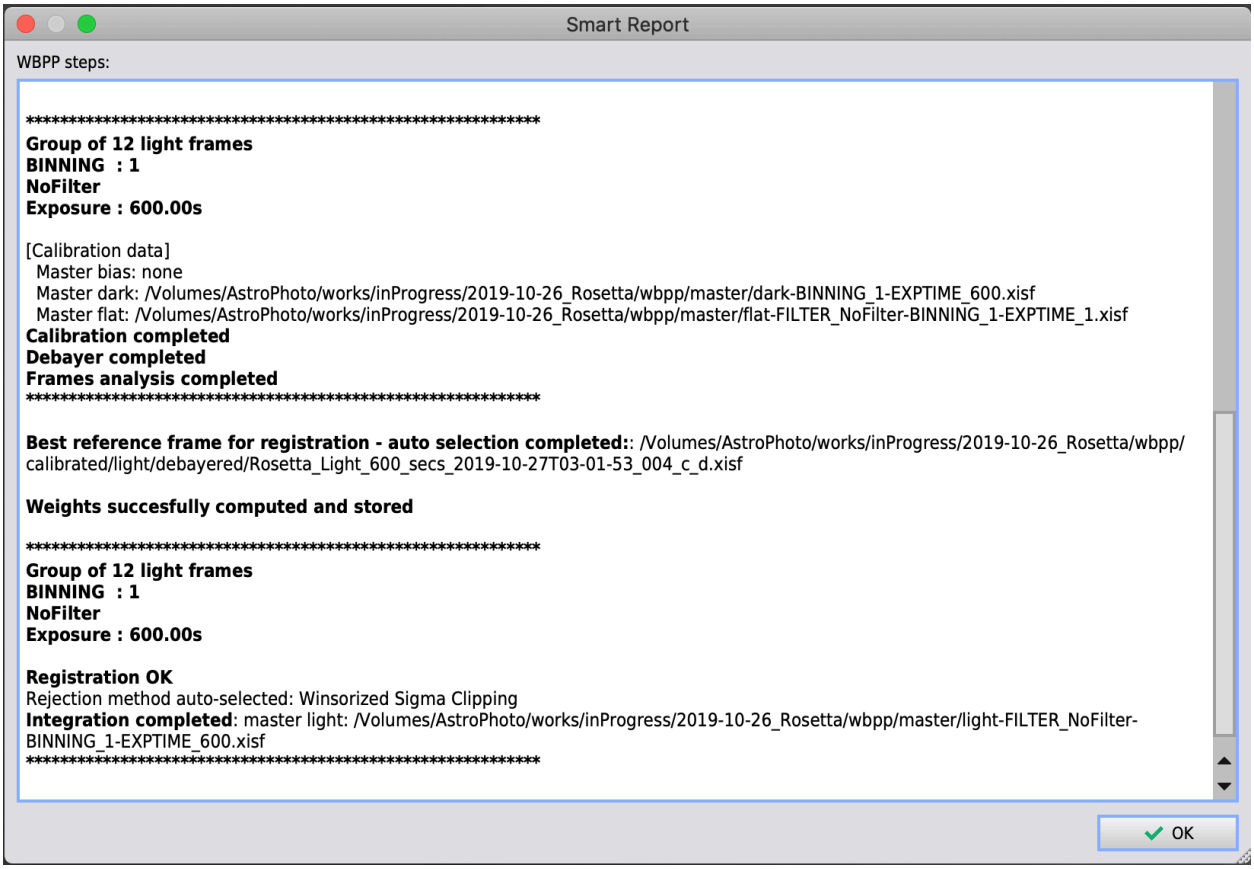


Smart Reporting

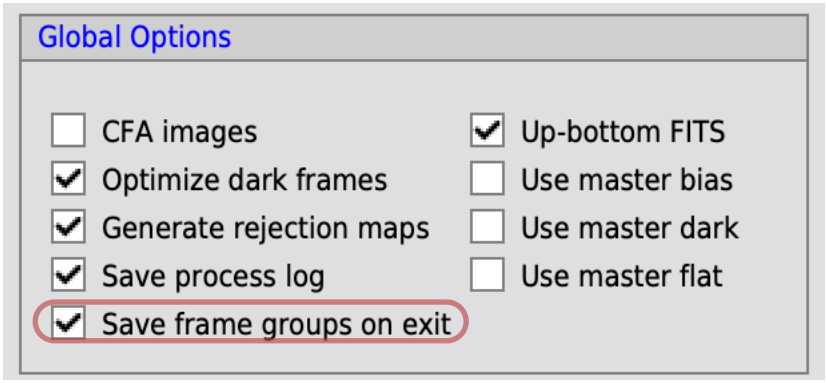
Standard BPP prints a verbose logging on the console during its execution and inspecting the log file to follow the execution means flowing through thousands of text lines. This is awkward and we decided to provide a high-level overview reporting called **Smart Reporting** presented at the end of the WBPP execution. This screen resumes in a human-readable fashion the steps executed on each frame group and few significant details about calibration, registration and integration such as the calibration master files choosen, which reference frame have been selected for the registration, which rejection algorithm has been selected and the location of master files.

Below few screenshots as an example:





Save frame groups on exit



It is foreseeable that given a set if images to stack, the attempt of using WBPP is not generating the best possible master light that can be achieved with this tool. By performing an iterative fine tuning of the weighting parameters the result can be refined. Presets help to get to a good result quickly but an advanced user is probably willing to do more than one attempt changing the weighting parameters.

To inspect the master lights generated, the WBPP dialog needs to be closed in order to load the masters in PixInsight and inspect them before executing a new WBPP integration. Since standard BPP does not remember the files loaded in the previous session, the user needs to reload every time the entire set of bias, dark, flat and light frames (eventually using the masters from the first run).

This option addresses this use case and, if active, lets the WBPP to reload the whole set of images (including the masters if they was generated previously) on the next launch such that there is no need to manually reinsert all files during this iterative tuning process.