DC3b Data Quality Requirements

Tim Axelrod
DRAFT Feb 8, 2010

This document summarizes the data quality requirements for the outputs of DC3b. The requirements are divided into those that are most naturally assessed on the outputs of particular processing stages, and those that are assessed on the integrated science database produced by DC3b. Numerical thresholds are part of most of these requirements. At this stage of requirements development, the actual values should be viewed as provisional and subject to change. The structure of the requirements is what matters most, and the intention is that these be fixed early. Note that computing speed requirements are **not** covered here.

Stage Output Requirements

1. Instrument Signature Removal (ISR)
    1. The task of ISR is to produce an image that approximates the one that would result from an ideal detector system exposed to the same sky flux at the telescope entrance pupil. The properties of the ideal detector system (which includes telescope optics as well as CCD detectors) are:
        1. The telescope and camera optics maps incoming photons within each differential solid angle to the corresponding detector pixels with an efficiency that is independent of the angular coordinates.
        2. The detector quantum efficiency is spatially uniform at all wavelengths
        3. A zero-exposure image from the detector is spatially uniform except for gaussian read noise which may vary in level from segment to segment. Note that (2) and (3) together imply that there are no detector "defects"
        4. Detector readout channels are independent - there is no crosstalk
        5. There are no electrons generated by cosmic rays
        6. Fringing characteristics - TBD
        7. Electron diffusion in the detectors is present at the expected level.
    2. The ISR performance will be verified using a subset of simulated images that cover a representative range of fields and observing conditions (to include all filters). This subset of simulated images will be generated in pairs. Both members of a pair will have the same sky flux at the entrance pupil. Both members of the pair will have the same PSF and distortion map (function that maps incoming photon angles to detector coordinates). The first member of the pair ("realistic") will be generated in the normal way, including full detector physics and vignetting of the focal plane by the telescope and camera optics. The second member of the pair ("ideal") will be generated using the ideal detector system specified in (1) above.
    3. The realistic image includes in its FITS header a list of input cosmic rays each with
        1. x, y detector coordinates where added
        2. "flipped" and "rotation" quantities that specify the orientation of the cosmic ray track
        3. an identifier of the cosmic ray in the Imsim library of cosmic rays
    4. A set of simulated images will be supplied which are generated in conditions that duplicate, as much as practical, the taking of monochromatic dome flats and bias (zero) exposures. These flats and biases will be processed to form the calibration inputs to the ISR.

5. With this information in hand, the performance of the ISR will be determined by comparing, for the identified subset of simulated inputs, the ISR output exposure ("output") generated from the "realistic" simulated input to the "ideal" simulated image. The output image will be divided by the ideal image to form the "ratio" image. The following metrics will be utilized, each on a per-amplifier basis:
    1. Flatness. A second order spatial polynomial will be fit to the ratio image. The polynomial coefficients shall fall within these bounds (needs work)
        1. a00 = 1.0 +/- Pflat0
        2. ax, ay = 0.0 +/- Pflat1
        3. axx, axy, ayy = 0.0 +/- Pflat2
    2. Cosmic ray identification: Fraction of pixels affected in Imsim realistic image unmasked as cosmic in output image < PCR %
    3. Bad pixel identification: Fraction of bad pixels in Imsim XXX file unmasked by ISR < Pbad %
    4. RMS deviation over pixels unmasked by ISR < Pdev %

2. Image Characterization Pipeline (ICP)
    1. WCS

    Definition: The error of a WCS is the RMS distance over a set of stars between the position predicted from the WCS and the catalog value of the position. The set of stars shall be all stars in a catalog of bright, isolated stars whose positions are within the image.
        1. Failure rate: A WCS shall be termed a failure when either the stage that produces the WCS raises an exception, or when the error > Pwcs1 arcsec. The failure rate, averaged over all images in the DC3b input set, shall be less than Pwcs2 %.
        2. Median error: The median WCS error, including all images in the DC3b input set that do not result in WCS failures, shall be less than Pwcs3 mas.
    3. Photometric Zeropoint
        1. For simulated images, the photometric zeropoint determined by the ICP shall agree with the zeropoint applied by the simulator, spatially averaged over the image, to within Pzp1 mag. (NOTE: this requires some coordination work with Imsim team to extract average cloud extinction)
    3. PSF Determination
        1. Failure rate: PSF determination failure, as signaled by an exception being raised, shall occur in less than Ppsf1 % of the input images
        2. Average error: The PSF error shall be determined from a specialized set of simulated images. These images shall be generated with sky catalogs that contain only isolated stars with a wide range of magnitudes, and a zero background. The "ideal" detector shall be used, as specified in Section 1 of this document. For these ideal images, the actual PSF will be extracted as a subimage centered on each catalog star with a size deemed sufficient to encompass all PSF features, including the scattering halo. The input to the pipeline shall be for the same sky catalog and atmosphere, but with realistic sky background and detector physics. The PSF model determined by the ICP from this input shall be directly compared with the actual PSF at each star location by forming the ratio of the PSF model to the actual PSF. For all pixels in the ratio images unmasked by ISR the pixel value shall be bounded by 1.0 +/- Ppsf2*noise, where the

        noise is the RSS of the variance in the ideal image and the real image at that pixel location.

    4. Background Determination
        1. Average error: For simulated images, (TBD - we need to work out with the Imsim team how to determine the background to which we should compare)

3. Produce Coadd for Detection
    1. Registration of images in stack: TBD
4. Produce Coadd for Subtraction Template
    1. Registration of images in stack: TBD
    2. PSF characteristics:
        1. PSF width, as defined by FWHM, shall be narrower than that of at least Ptempl1 % of the input images in the stack
        2. PSF eccentricity: TBD
5. Image Subtraction

   The residual image is defined as the subtracted image divided by the square root of the predicted variance image.
    1. Failure rate: A subtraction shall be termed a failure when a stage that produces the subtraction raises an exception, or when the absolute value of the average residual > Psub1. The failure rate, averaged over all images in the DC3b input set that generate a valid WCS, shall be less than Psub2 %.
    2. Residuals in object footprints: For each object in the template image with magnitude < Psub3 (probably need to make this filter dependent), define the footprint pixels to be those where the amplitude of the PSF, centered on the object is > Psub4 % of its maximum value. For the union of all object footprint pixels in an image, a subtraction is deemed of bad quality if the fraction of those pixels with abs(resid) > Psub5 is > Psub6 percent. The fraction of bad quality subtractions shall be < Psub7 percent.
3. Detect Sources in Subtracted Image
    1. Completeness: For simulated input images, a minimum of Pdet1 % of sources that differ from their template flux by at least Pdet2 times the square root of the variance of the sky flux in a PSF in both exposures in a visit shall result in valid DIASources. Note: this is much weaker than expected from Gaussian statistics.
    2. False detection rate: For the same values of detection parameters as chosen to meet the completeness requirement, above, and for simulated images, no more than Pdet3 % of the DIASources generated for a visit shall be false, ie not resulting from an actual flux change in a simulated source. Again, much weaker than Gaussian.
4. Deep Detection and Measurement
    1. Completeness: For simulated images, the 5 sigma detection limit shall be determined by the LSST ETC for each image stack using the simulated values of the photometric zero point, sky background, and PSF widths as determined by the ICP, yielding M5. The completeness of the generated object catalog shall be determined as a function of magnitude by ratioing, in each magnitude bin, the number of objects detected by the pipeline which are matched to objects in the simulated catalog, to the number of objects in the simulated catalog in that magnitude bin (how to handle systematic magnitude errors here?). The completeness for M < M5 shall be at least Pdeep1 %
    2. Reliability: For detected objects with magnitudes < M5, the fraction of objects that do not match simulated catalog objects within Pdeep2 arcsec

    (and are therefore deemed to be false detections) shall be less than Pdeep3 %
    3. Accuracy of shape measurements: for simulated images, TBD [It's unclear the best thing to do here. The image simulator models galaxies as the sum of a bulge and disk model, each with its own Sersic model and inclination. The Multifit model of a galaxy at this point is considerably simpler]
5. Day MOPS
    1. Adopt metrics from PS-MOPS
    2. For simulated input images, calculate completeness and reliability of the MOPS generated orbit catalog to the simulated solar system catalog [Need details from Imsim on their solar system catalog]
6. Association (AP)
    1. Accuracy of association of DIASources to Objects: For simulated images, DIASources resulting from actual flux changes shall be correctly associated with their template object at least Pap1 % of the time on average.
    2. Completeness of MOPS matching: For DIASources that fall within the MOPS-predicted error ellipse of a solar system object, the AP shall match the DIASource to that solar system object at least Pap2 % of the time on average.
    3. Classification correctness: Cosmic rays that result in a DIASource in one exposure of a visit shall be correctly classified by the AP at least Pap3% of the time on average.
7. Photometric Calibration
    1. Accuracy of photometry (basically, can do SRD tests): TBD
    2. For simulated images, comparison of derived to input atmospheric model: TBD
    3. Limitations on systematics: TBD
8. Astrometric Calibration
    1. Accuracy of astrometric models: TBD
    2. Limits on systematics: TBD

<u>Integrated Database Requirements</u>

1. Exposure table shall contain an entry for every exposure processed, and each shall include a minimal set of valid metadata (TBD)
2. Database shall pass consistency checks: eg, no orphaned Sources, DIASources

<u>Table of Thresholds</u>

Threshold name	Value	Units
Pwcs1	2.0	arcsec
Pwcs2	2	percent
Pwcs3	40	milli-arcsec
Pzp1	0.05	mag
Ppsf1	2	percent
Ppsf2		
Psub1	0.5	

Psub2	2	percent
Psub3	22	mag
Psub4	5	percent
Psub5	3	
Psub6	10	percent
Psub7	3	percent
Pdet1	95	percent
Pdet2	5	
Pdet3	2	percent
Ptempl1	95	percent
Pap1	99	percent
Pap2	99	percent
Pap3	90	percent