7th IAA Planetary Defense Conference – PDC 2021
26-30 April 2021, Vienna, Austria

IAA-PDC-21-05-13
Image simulation for the NEO Surveyor

J. A. Surace[(1)], **X. Liu**[(2]
[(1)]*IPAC, MS 100-22, Caltech, 1200 E. California Blvd. Pasadena, CA 91125, 626-395-1919, jason@ipac.calteh.edu*
[(2)]*IPAC, MS 100-22, Caltech, 1200 E. California Blvd. Pasadena, CA 91125, 626-395-1898, xliu@caltech.edu*

Keywords: *infrared, image, simulator, NEO Surveyor*

Extended Abstract—

We describe the survey image simulator for the Near-Earth Object Surveyor. The purpose of the simulator is to produce realistic raw image data, with properties similar enough to real instrument data as to exercise all components of the survey data processing system, and sufficient to characterize the likely performance of the processing algorithms relative to the mission requirements. The simulator is intended to support testing of data processing algorithms and system throughput. It is not intended for forward physical modeling and comparison to real data, or any similar procedure.

The NEO Surveyor Mission

The Near-Earth Object Surveyor is a mission to locate and characterize potentially hazardous objects near the Earth. Conceptually based on the NEOWISE pathfinder, NEO Surveyor will consist of a 50cm unobstructed mid-infrared space telescope orbiting the Sun-Earth L1 point. Eight 2048x2048 H2RG detectors with 3 arcsecond pixels observe simultaneously a 1.7 x 7.1 degree swath of sky at 4.6 and 8μm. NEO Surveyor repeatedly surveys the entire sky within 40 degrees of the ecliptic plane, with cadences of minutes, hours, and days. The telescope and detectors are passively cooled and use no consumables. The fundamental observation unit is called a “visit”, which consists of six 30-second exposures in a small hexagonal dither pattern. Visits are executed in a circulating “loop” that revisits a given location on the sky roughly every two hours. Four successive loops constitute a “quad” that forms the basis for detecting moving objects.

NEO Surveyor data will be processed by IPAC at Caltech. In addition to basic image processing and calibration, the survey images will be coadded into reference images of the static sky. These will be subtracted from the individual epochs in order to facilitate detection of moving objects. It is critical to understand the completeness and reliability of these algorithms. The image simulator is thus needed to create data where the “truth” of the images is exactly known in order to produce detailed comparisons to observed results.

Image Simulator Design

The simulator is written entirely in Python. The development procedure involved scientific investigation by a lead scientist (e.g. translation of colors for flux prediction, etc.) as well as the writing of design specification documents and example code prototypes. The lead developer then translated these ideas into production code. In many cases this resulted in direct code validation since the prototype IDL code was completely independent of the production Python code. Comparison of the output verified proper execution.

By design the image simulator is embarrassingly parallel. The simulator exists in two different environments. The primary development version resides in a CentOS linux cluster located at IPAC. In this environment, SLURM is used as the executor, scheduling parallel jobs for execution. The simulator has also been containerized via Docker for deployment within the Amazon Web Services (AWS) compute cloud. The simulator eliminates external dependencies to the extent possible.

Image Survey Pointing Expansion

The basic survey plan is supplied by the NEO Surveyor Survey Simulation Team as a series of visit-level sky pointings of the telescope boresight (optical center axis) and associated times of observation. The image simulator begins by attaching the observation dither pattern, and then expanding this to each of the eight detectors by attaching the detailed information about the detector geometry. The end result is a file containing a single line for each detector exposure. We also generate and attach additional metadata to each line describing the current observation in the context of the overall survey structure, for example, which of the dither positions a current image is in relation to a visit. This file is the underlying basis for the job parallelization that makes simulating large (weeks) of data possible. Subsets of these images are split out and spun off to individual independent drone computers in a cluster.

zodiacal light (reflected sunlight and thermal emission from dust in the inner Solar System), and emission from the ISM (gas and dust in the Milky Way). The zodiacal emission is the brightest of these. It varies depending on the relative position of the spacecraft and the sun. The model is based on that of Wright (1998) and Gorjian, Wright, and Chary (2000). It tends to be very

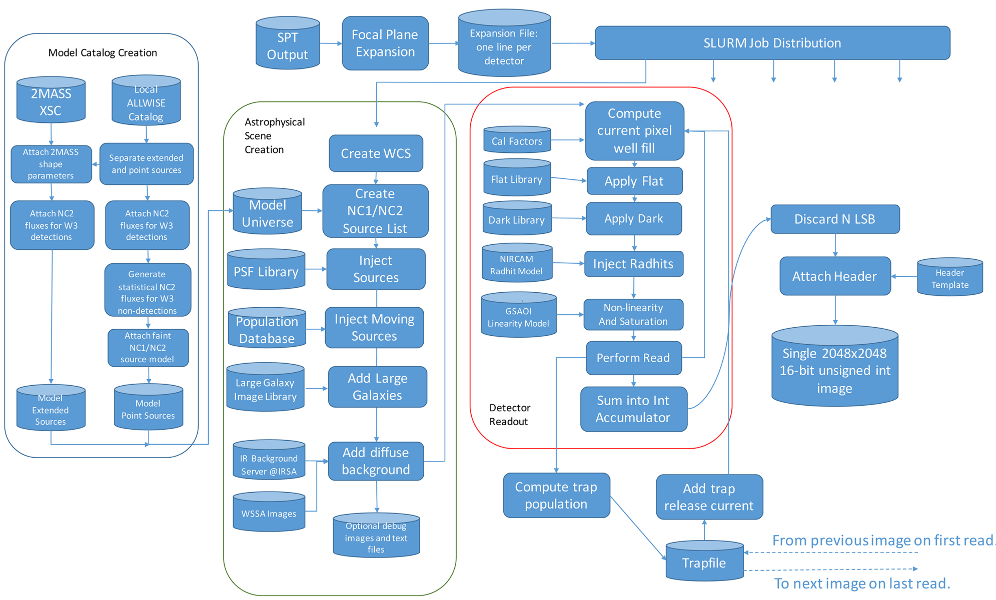

Figure 1 - architecture of the NEO Surveyor image simulator

The Astrophysical Scene

Each image creation job starts with construction of the *astrophysical scene*. This is a noiseless representation of the infrared sky constructed on exactly the same pixel grid as the detectors. The scene is calibrated in units of Jy/pixel. It is generated on-the-fly for every image. It has many components including diffuse and structured backgrounds, "static" background sources such as stars and galaxies, and moving solar system objects. It is primarily based upon results from the WISE mission, and wherever possible reflects the actual sky, primarily limited by the depth of the WISE observations. The objects contained in the scene are mostly modeled, allowing them to be noiseless and perfectly characterized. The scene itself is stored for reference, as are detailed catalogs of all of the scene contents. This establishes a ground truth against which the data processing can be compared.

Objects

The simulator begins by adding the extended diffuse infrared background. Three such backgrounds are explicitly simulated: the cosmic infrared background (CIB: combined light of distant unresolved galaxies), the

smooth, and is modeled as a tilted plane whose exact surface brightness is tied to the specific circumstances of the observation. The CIB is also based on the same model, but is included as a constant across the field. The structured ISM component is derived from the WISE-based "WSSA" images of Meisner & Finkbeiner (2014). These have been heavily reprocessed to scrub them of artifacts and small extended objects, and smoothed with an adaptive wavelet algorithm to suppress correlated noise left over from the WISE coaddition.

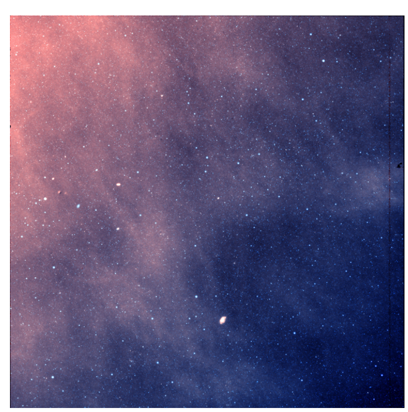

Figure 2 - simulated single NEO Surveyor detector image illustrating the large-scale structure of the ISM

The majority of the discrete objects in the scene are unresolved stars and galaxies. Objects bright enough to be detected in the ALLWISE survey are inserted at their ALLWISE positions and with ALLWISE-derived fluxes. The very brightest stars, which saturate the WISE detectors, are derived from 2MASS. Because NEO Surveyor will detect objects much fainter than ALLWISE, there is an additional synthetic population extending to 20th magnitude (Vega) at 4.6μm, which adds 14x more objects than seen by ALLWISE. This synthetic population is based on the 4.5μm number counts seen in deep Spitzer surveys. A probabilistic method is used to create the synthetic population such that it exactly reproduces the color distribution and number counts of known real sources. This model is stored in a flat-file database for rapid access during scene construction.

A shared database known as the Reference Small Body Population Model (RSBPM) of Main Belt asteroids and Near-Earth Objects (Amor, Aten, and Apollo families) is supplied by the NEO Surveyor Survey Simulation team. These objects are purely synthetic populations. The image simulator accesses the RSBPM via a software API that returns all of the synthetic objects within the field of view at the time of observation, their projected positions on the sky, and their fluxes in the two NEO Surveyor filters. The objects are inserted into the astrophysical scene as point sources. A future upgrade will include comets, which will be embedded in the scene as extended objects.

The planets, as well as the Galilean satellites and Titan, are added to the simulation as appropriately sized disks, taking care to handle scattered light and the occultation of the background stars and galaxies properly. Their positions and sizes are derived from ephemerides computed by JPL-HORIZONS. With few exceptions, these objects are so bright that they saturate the NEO Surveyor detectors.

Figure 3 - simulated NEO Surveyor image of the large galaxy NGC 300.

Finally, extended galaxies are added to the scene. A quarter million small extended galaxies are inserted into the scene using elliptical gaussian profiles derived from 2MASS. The largest one thousand galaxies are embedded using detailed 2D image models based on deep coadded WISE images (Jarrett et al. 2019).

Point Spread Function and Distortion

The point spread function (PSF) used to insert objects into the scene is based on optical modeling of the telescope. The simulator allows the PSF to vary across each detector. In addition, the effects of pixel phase (the centering of an object relative to the actual detector pixel borders) is explicitly handled. For speed, a pre-computed and binned PSF library allows insertion of objects with an accuracy of 0.01 pixels.

The effects of image distortion (the displacement of an object from its expected location based on a pure tangent plane projection) are also explicitly included. Distortion is again based on optical modeling and is expressed using the TAN-SIP methodology, which is a polynomial correction to the tangent plane projection in x-y detector pixel space. A single description is used for the entire focal plane; the individual detectors are located within it based on the planned mechanical design.

Star positions are also corrected for proper motion, derived from the GAIA Data Release 2. A cross-match between ALLWISE and GAIA was used to attach proper motion rates to 14 million sources which will have moved by more than 0.1 detector pixels over the 15 year time gap between ALLWISE and NEO Surveyor. The size of the motion applied is based on the mean time of the WISE observations and the specific time of the NEO Surveyor observations.

Finally, additional convolutions are applied to account for the spacecraft jitter and also the detector transfer function (effects that cause photons to be detected in pixels other than where they actually landed). The latter is a significant contributor to the expected image quality.

Detector Readout

The astrophysical scene is then passed to the detector readout segment. The detector readout reproduces as closely as possible the algorithms by which the detectors are read and how these reads are combined into a single image to be returned to Earth.

The calibrated astrophysical scene is used to determine the rate at which photons fall on each detector pixel, taking into account system throughput, gain, etc. This includes the dark current as a function of position, and the flatfield (pixel-to-pixel gain variation), both of which are based on testing of existing H2RG detectors. The software tracks the linear number of incident photons, and applies a non-linearity function to derive the observed number of electrons (also based on existing H2RG-equipped instruments).

A latent image model is included in the detector readout based on exponential trap decay. During the readout the image simulator tracks the total number of populated traps in each pixel. The timing of the start and end of each readout is used to determine the number of traps released between reads. The trap population model is passed serially from one image creation job to the next. Because this limits the parallelization of the simulator, for practical reasons latents are arbitrarily tracked only within a survey loop (192 images over a period of 2 hours).

The software implements the Sample-Up-The Ramp (SUTR) mode for the 4.6μm detectors as used by WISE. After a detector reset, a series of 18 non-destructive reads are simulated, separated by 1.5-second intervals. The simulated reads include Poisson noise and read noise. Radiation hits are included at each read based on a simple model developed for JWST (and are recorded in a diagnostic file for future analysis). The readout algorithm returns a linear fit to these samples. Saturation is detected during the readout and encoded in the output image. All operations are carefully implemented using integer math to reflect the quantized nature of the reads. A future upgrade is expected to implement a readout that can correct for saturation (Zemcov et al. 2016).

The 8μm detectors use a different readout scheme: correlated double sampling (CDS). Here the detector is reset, followed by a non-destructive read, and then a destructive read 3 seconds later. The difference is stored, and repeated eight times. The final output is effectively an average of the eight pedestal/signal read pairs. Poisson and read noise, as well as saturation and radhits, are handled in a fashion similar to the 4.6μm detectors.

Simulator Output

The data produced by the detector readout is processed and packaged to produce a 16-bit FITS image, including reproduction of the planned compression algorithms to be used to conserve downlink bandwidth. At this time we also generate additional metadata such as simulated housekeeping telemetry, pointing information, and information about the simulation itself (e.g. background values, number of sources simulated, etc), which are included in the image header.

Many additional files are also generated. As mentioned, the astrophysical scene is stored. The actual time-ordered simulated reads (ramps and CDS pairs) may be output as FITS datacubes. The trap population model is supplied as an image. Numerous table files contain the exact star and galaxies in the simulation, along with their positions and fluxes. Similar tables are produced for the moving objects such as asteroids, planets, and moons. Table files also detail the exact position and intensity of the radiation hits.

The data are stored using the naming convention planned for the actual mission, and in the planned processing directory tree. Together with the metadata in their headers, they are capable of being ingested by the data system exactly as the real survey data will be.

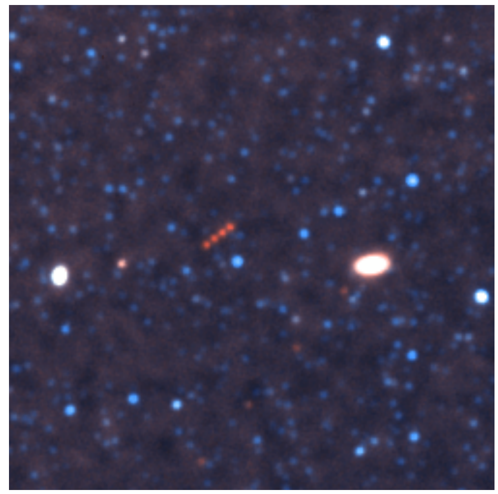

Figure 4 - simulated coadded NEO Surveyor data showing the track of a moving object. The four "visits" were coadded separately, then summed to make this image. The moving objects are very red at [4.6μm-8.0μm] colors.

References

Gorjian, V., Wright, E., & Chary, R. 2000, ApJ 536, 550.
Jarrett, T., Cluver, M., et al., 2019, ApJS, 245, 25
Meisner, A., & Finkbeiner, D., 2014, ApJ, 781, 5
Wright, E., 1998, ApJ, 496, 1
Zemcov, M., Crill, B., et al. 2016, Journal of Astronomical Instrumentation, 5, 3