

Calibrating and correcting Charge Transfer Inefficiency in CCDs using Pyxel

Bradley Kelman^a, Thibaut Prod'homme^b, Jesper Skottfelt^a, Frédéric Lemmel^b, Matej Arko^b, Patricia Liebing^c, Peter Verhoeve^b, Ben Dryer^a, David Hall^a, and Michael Hubbard^a

^aCentre for Electronic Imaging, Open University, Walton Hall, Milton Keynes, MK7 6AA, UK

^bEuropean Space Agency, ESTEC, Keplerlaan 1, 2201 AZ, Noordwijk, The Netherlands

^cMullard Space Science Laboratory, Holmbury Hill Rd, Dorking RH5 6NT, UK

ABSTRACT

To tackle the ever-more demanding requirements of upcoming astronomical instruments, emphasis is being put on accurate, reliable, and reusable models to simulate detector effects on images. The open-source python package Pyxel aims at solving these issues by providing a simulation framework where detector effects models can be easily implemented, pipelined and calibrated or validated against test data. In this contribution, we detail how by using the Pyxel framework, it is possible to calibrate ArCTIC – a model for simulating and correcting Charge Transfer Inefficiency in CCDs – and check its correction efficiency for realistic galaxy images acquired using an irradiated Teledyne e2v CCD273.

Keywords: Radiation damage, simulation, calibration, Euclid, CCD, trap, Charge Transfer Inefficiency

1. INTRODUCTION

To tackle the ever-more demanding requirements of upcoming astronomical instruments in particular towards detectors, more emphasis is being put on accurate, reliable, and reusable models to simulate different detector effects on images collected by those instruments. The open-source Python package Pyxel^{1,2} aims at solving these issues by providing a simulation framework where models of detector effects can be easily implemented, pipelined and calibrated or validated against experimental test data.

This contribution focuses on a detector effect known as Charge Transfer Inefficiency (CTI), a phenomenon whereby defects in CCDs trap charge and cause spurious trailing events in the resulting images. This is particularly a problem in space, where the radiative environment can cause a significant amount of damage to the CCD over time. Due to the non-linear nature of CTI, numerous computer aided models have been designed to correct for it. One such model is the Algorithm for CTI correction (ArCTIC).^{3,4} This model is to be used by future space missions to correct for CTI in their images and has been proven to be highly accurate for synthetic data. Presented here is the first steps in a general attempt to check the correction efficiency of ArCTIC for realistic test data.

In particular, it is detailed how, by using the Pyxel framework, it is possible to calibrate ArCTIC's parameters against flat fields acquired using an irradiated Teledyne e2v CCD273. Then, by using these parameters found via calibration, the ArCTIC can be leveraged to correct for CTI in realistic galaxy images acquired using the same irradiated CCD and a smartphone projector experiment.⁵

2. EXPERIMENTAL SETUP AND DATA

This section briefly describes how the test data used in this study has been obtained. The test data is composed of very low signal level flat field illuminations corresponding to Euclid's faintest signal range of interest of [0-1000] e⁻. The flat fields have been projected using ESA's Science Payload Validation section (SCI-FIV)

For contacting the authors send correspondence to: bradley.kelman@open.ac.uk

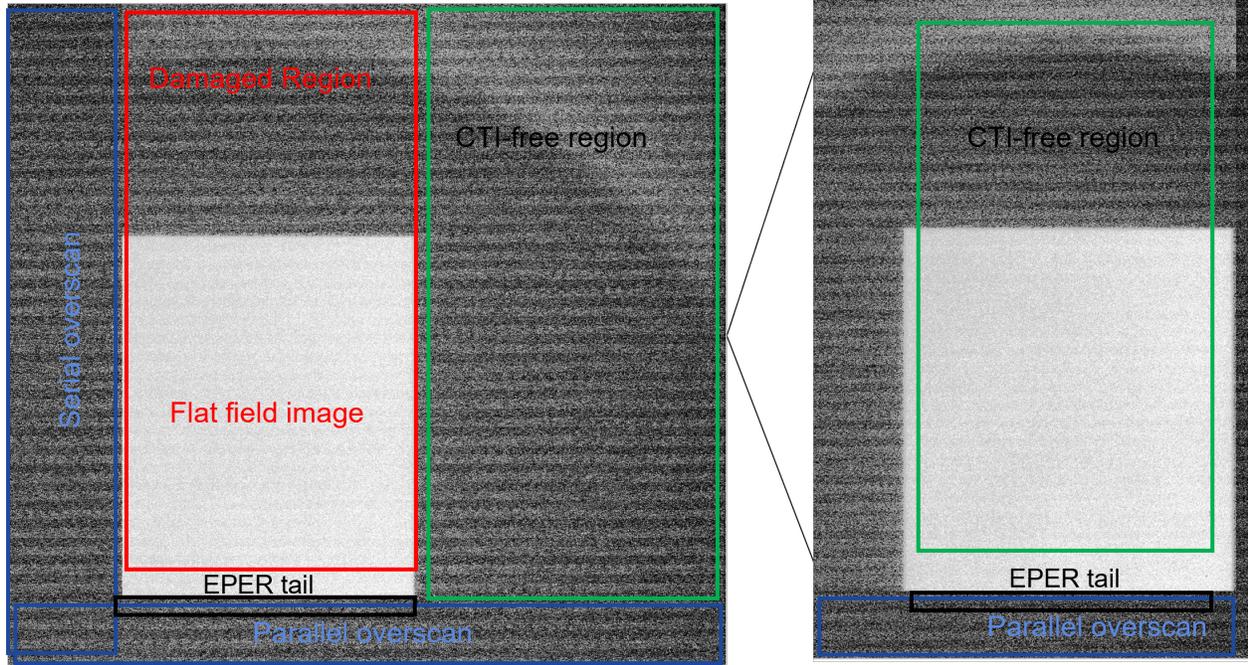


Figure 1. Example flatfield images as using the smartphone setup showing only the G quadrant: the damaged (left) and CTI-free (right) regions of the EM3 CCD. The areas of relevance for this work are highlighted: the damaged region (red), the CTI-free region (green), serial and parallel overscans (blue) and the EPER tail regions (black). The fixed-pattern noise visible in the raw images is removed by our offset removal procedure based on subtracting the average serial overscan profile.

smartphone projection system as described in⁵ and acquired using a Teledyne e2v CCD273 14-62-17-01,⁶ manufactured for the Euclid VIS instrument. This CCD has previously been proton irradiated at room temperature to end-of-life fluence ($4.8 \cdot 10^9 \text{p}^+ \text{cm}^{-2}$ 10 MeV equivalent), see also⁵ for more information on the irradiation.

Each signal level corresponds to a different phone screen brightness setting; Table 1 shows the relationship between brightness setting and the measured average signal in electrons (note that although the increase in brightness setting is linear, the increase in electrons is logarithmic).

The CCD is a $4\text{k} \times 4\text{k}$ pixels ($12 \mu\text{m}$ pitch), back-illuminated full-frame device with a read noise $< 3.6 \text{e}^-$ and 150ke^- full well capacity. It comprises 4 output nodes, dividing the image area into 4 quadrants named E, F, G and H. The data discussed here were obtained over the G quadrant solely. Only half of the G quadrant has been irradiated, the damaged region is referred to as G1 and control/CTI-free region as G0. The optical setup was chosen such that the flat illumination is projected over either G0 or G1 with the distance between two LEDs from smartphone screen corresponding roughly to the CCD pixel pitch. As can be seen in figure 1, only a bit more than half the G0 and G1 regions are illuminated. This means that both First Pixel edge Response (FPR) and Extended Pixel Edge Response (EPER) from respectively the flat leading and trailing edges (parallel overscan).

CCD273 was operated with the same timings than for the Euclid VIS instrument with a total line transfer duration of 38.94 ms. The CCD was overclocked to give sizable overscan regions to enable both electronic offset correction and EPER analysis ; one image quadrant size corresponds to a total of 2500 serial pixels (rows) and 2300 parallel pixels (columns).

Phone signal levels	20	30	50	60	70	80	90	100	110	120	130
Electron values	21	26	61	105	143	200	273	391	512	670	846

Table 1. Screen brightness settings processed within this work and their corresponding measured signal level in electrons. Note that the screen brightness of 40 has been discarded due to a very high level of fixed pattern noise.

3. THE CALIBRATION PROCEDURE

The objective of this study is to obtain CTI model parameters calibrated against test data and to proceed with the CTI-correction of Euclid-like scenes obtained using the same CCD and in the same conditions as the calibration dataset. In the following an overview is given describing the devised calibration procedure by giving a short description of its three main key ingredients:

- ArCTIC^{3,4} (Algorithm for Charge Transfer Inefficiency Correction), the CTI model of reference for the Euclid scientific consortium and the one used in this study.
- Pyxel^{1,2} - an open-source end-to-end detector simulation framework capable of replicating numerous detector effects through the use of plug-and-play models. Pyxel includes a calibration mode that is the main tool behind the work reported.
- The different data processing steps used to prepare the simulation inputs and calibration target test data from the smartphone flat field measurements.

3.1 ArCTIC

ArCTIC is a semi-analytical CTI model with the ability to simulate the transfer of charge to the readout node(s) with a tuneable time resolution through its 'express' factor. It is possible to set this number up to the number of parallel pixels within the simulated image (time resolution corresponding to the line shift duration). The express factor dictates the total number of times the electron and trap densities are recalculated throughout the simulation. The simplest mode is setting the express factor to 1 which only calculates volumes and densities at the start of the simulation, similar to that of the Charge Distortion Model (CDM).⁷ Express=0 gives the highest fidelity, with the recalculation occurring after every pixel transfer. One of the objectives of this study is to investigate the trade off between the increased accuracy of a high express factor and the additional computation resources taken to calibrate and correct the lab data.

Although ArCTIC has been proven to be highly accurate at correcting CTI in synthetic data produced computationally⁸ it has not yet been fully utilised against images taken by detectors in a laboratory environment representative of the future CCD operation and observations. This study is a first step in testing the level of accuracy reachable by ArCTIC for the CTI correction of the Euclid VIS data in realistic conditions.

The ArCTIC free model parameters to be calibrated against the test data are: the β parameter (that dictates how the electron density and volume evolve with the number of electrons), the total number of trap species, and for each trap species, the trap density, ρ , and the trap release time constant, τ .

3.2 Model calibration using Pyxel

Pyxel has three main running modes: exposure, observation and calibration. Exposure mode is the simplest, running the models specified just once, a simple pipeline. If a user wants to run multiple pipelines, each with slightly different argument values, observation mode has that capability. For more information on Pyxel and its running modes, see⁹⁻¹¹.

The calibration mode is designed to find the set of model free parameters such that simulations will best reproduce a set of test data (herein called target). It runs multiple pipelines both sequentially and in parallel to obtain these values, leveraging the genetic algorithm approach in the Pygmo python package to optimise the parameters^{12,13}.

In calibration mode, a pipeline is run using random model parameter values within set boundary conditions on an initial input image, which is then compared to the desired target image. An example of a calibration mode pipeline is presented in Figure 2. Each pipeline is known as an individual. The arguments analysed and used in this work are the trap parameters, i.e. ρ , τ and β . The ArCTIC model is applied to the CTI-free input image to create a new image with CTI. The goodness of fit (fitness value) between simulated and target images is then evaluated by computing the sum of the squared residuals.

It is possible to set the number of individuals in a calibration (*population*), the number of times individuals in a population get compared ((generations)) and the number of times the best individuals in a generation get

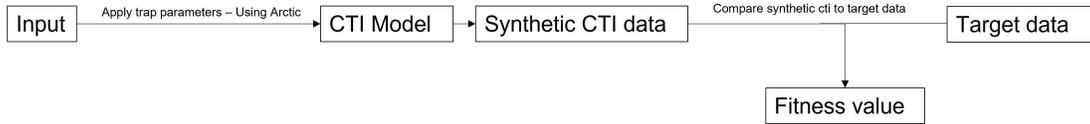


Figure 2. An example of a single simulation (pipeline) within Pyxel’s calibration mode. These pipelines are run numerous times over the course of a single calibration.

compared (*evolution*). When the evolution stage is complete, a set of champion individuals are selected and mutated for use in the subsequent evolution according to pre-set parameters within the genetic algorithm used (SaDE).¹⁴ Details of how an evolution works can be seen in figure 3. It is also possible to run multiple evolutions in parallel with the use of the *islands* parameter, which can then also be compared once each evolution has been simulated, allowing for greater investigation into the parameter space.

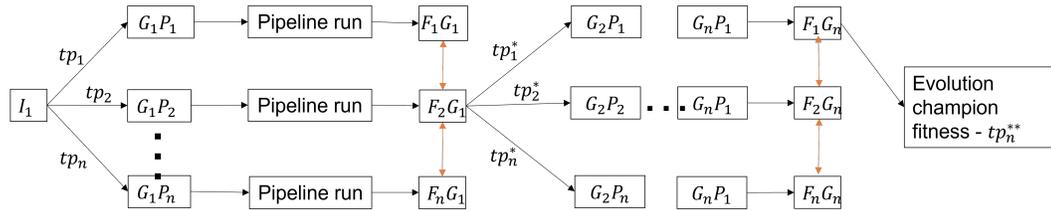


Figure 3. Evolution run of Pyxel calibration mode. An island I_1 contains n number of individuals to form a population in the first generation. These individuals are each given a random set of trap properties within pre-determined boundary conditions, a pipeline is run and then a fitness values denoting each individuals accuracy are compared to their peers. The best fitness values have their trap properties taken and mutated to be used in the next generation, this is then repeated until the number of generations that have been defined is reached and the champion individual is outputted.

3.3 Data processing

This section details the data processing steps to convert flat field images in digital units into one-dimensional parallel charge profiles in electron units used in the ArCTIC calibration as targets, namely: (i) conversion from digital units to electrons including the removal of the average and fixed pattern electronic offset, (ii) determination of the region of interests, (iii) generation of the average parallel charge profiles.

An average image is produced from the 100 frames of raw data for each screen brightness setting used. For each CCD quadrant, a 100x100 pixel region of interest in the serial-parallel overscan corner was averaged to obtain the mean electronic offset value to be subtracted from the average image. To remove any residual fixed pattern offset (easily visible in figure 1) from the mean image, an average serial overscan profile was calculated and subtracted from the mean image on a per column basis. Using previously measured system conversion gain values (obtained using the Fe55 X-ray technique), the image was then converted into electrons.

To produce representative parallel charge profiles, it is first required to define regions of interest. This is done based on the uniformity in both the illuminated flat field and dark current data. The latter to ensure the determined region of interest as a uniform level of radiation-induced damage.

Dark images with an integration time of 3000 s (see an example in figure 4), were collected at the same time as the flat field exposures. Figure 4 (right) shows the mean serial profile for a dark image from which the level of uniformity is derived. A similar process was followed for the flat fields using a screen brightness of 100. Table 2 shows the resulting regions’ serial coordinates.

To produce the final parallel profiles, mean values are produced for each row by averaging over the convolved regions of interest. Parallel charge profiles are produced for each phone brightness setting, producing a CTI-free (simulation input) and a damaged parallel charge profile (calibration target) per quadrant. An example of such parallel charge profiles can be seen in figure 5.

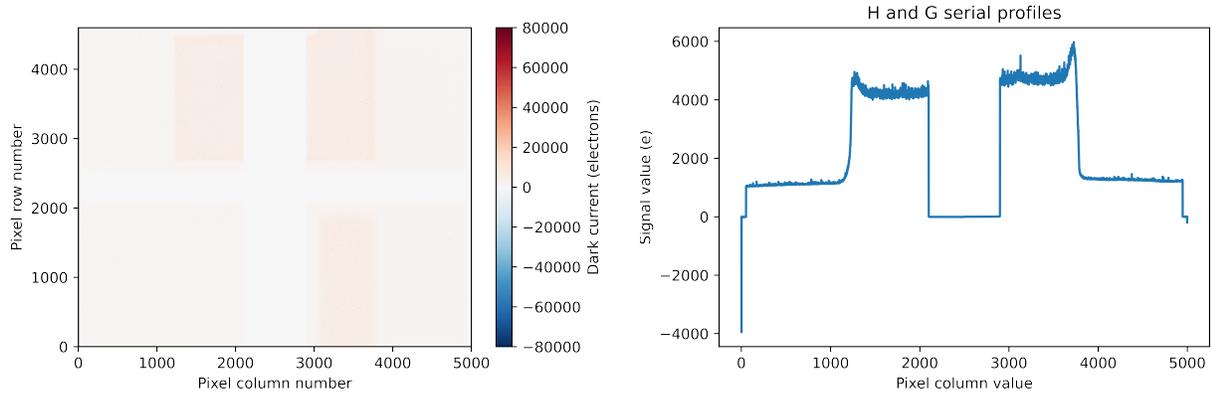


Figure 4. Left: Example of a dark image obtained for 3000 s integration using the EM3 CCD273: three regions with elevated dark current are visible corresponding to the irradiated part of the CCD. Note that the G quadrant analysed in this study is located at the top right corner. Right: The average serial profile for this dark image for the top part of the CCD only i.e. H (left) and G (right) quadrants.

Quadrant	Dark current	Flat field	Convolved ROI
G0	3900-4700	4000-4900	4000-4700
G1	3000-3600	2950-3850	3000-3600

Table 2. ROI serial coordinates. Obtained from looking at restrictions due to the uniform dark current and flat field illumination values, together they give the serial pixel column numbers that can be used to get the final parallel charge profile.

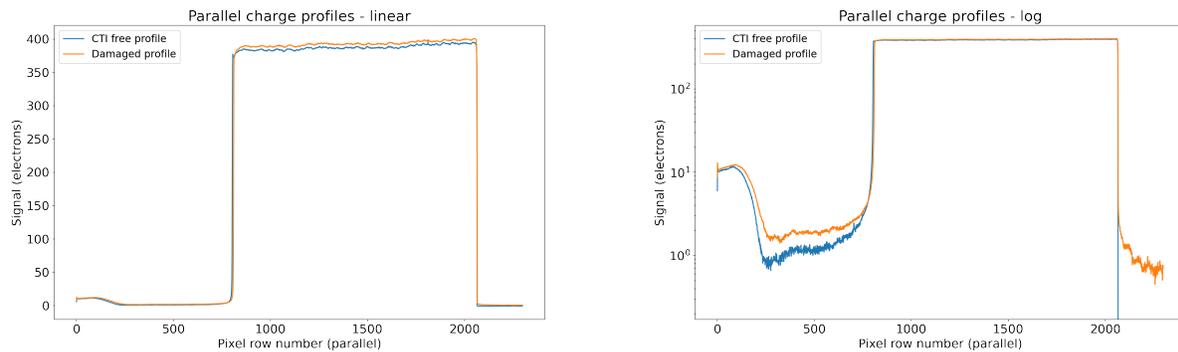


Figure 5. Parallel charge profiles in both linear (left) and logarithmic (right) scales for the CTI-free and damaged regions of the G quadrant which were used within this work. The profiles are for a phone screen brightness setting of 100, corresponding to roughly a 400 e^- signal level. Both FPR and EPER are visible in the images for the damaged profile (orange): bending of the leading edge and charge trailing in the overscan. Note that there is a slight but noticeable mismatch (1 to 3 % depending on the signal level) in the average signal level between damaged and CTI-free profiles that have to be corrected before one can use the CTI-free profile as a simulation input.

4. CALIBRATION SCHEME VALIDATION USING SYNTHETIC DATA

Before applying the explained calibration scheme to target data obtained in the laboratory, tests using synthetic data were conducted, for which CTI is simulated by ArCTIC with known trap parameters. In this way, an investigation can be conducted on how well can the known trap parameters can be retrieved and if there is any bias present, and ultimately validate the procedure.

The input file for the synthetic calibration are the original CTI-free parallel charge profiles already processed (see 3.3), while the target profiles are the CTI-free profiles that have gone through a Pyxel exposure pipeline to

add a known amount of CTI to the profile using ArCTIC.

Two sets of damaged profiles are produced for a given signal level: one, for which CTI is simulated using three trap species and the other using four. In this way one can understand how the number of trap species present within a profile can affect the calibration results, and quantify the known degeneracies, for instance, how well one can reproduce an EPER trail with only three trap species although it was generated with four. For the three trap species profile the trap release timescales are 1, 10, and 100 in units of line transfer duration. For the four trap species profiles, the fourth trap species is a shorter trap with release timescale of 0.1. All trap densities are set to unity. It is important to note that readout noise is added to the profiles to mimic the level of noise present in the experimental target data. A value of $3e/\sqrt{100}$ was chosen, with 3 corresponding roughly to the Euclid CCD readout noise as measured for the SCI-FIV Euclid test bench to the profiles. $\sqrt{100}$ corresponds to the reduction in the statistical uncertainty when averaging over the 100 images, as per the calibration experimental target data.

In the calibration pipeline, the CTI-free parallel profile data is used as the input and one set of the synthetically-damaged profiles (3 or 4 traps) is used as the target. Multiple calibration pipelines are run for each set of target files to investigate the effect of the number of traps within the calibration and the difference between different weighting schemes (anticipating a bias in the goodness of fit towards larger signal levels). The calibrations that investigated weighting showed minimal differences in effectiveness, the following discussion of the results hereby contain no application of weighting.

As already mentioned in section 3.1, the express factor of ArCTIC can be used to speed up the CTI simulation time at the expense of realism. To perform a calibration, it is necessary to run several thousands of simulations, and the computational load becomes a key factor. Numerous simple Pyxel exposure pipelines were benchmarked using different express factors for ArCTIC to measure not only speed but also realism: `express=1` was the quickest (7 ms) but gave non-physical profile results (strange features in both the flat fields and EPER profile yet unexplained). `Express=0` (highest level of realism) simulated realistic profiles but takes too long to be used in the calibration (850 ms). 50 was found to be the smallest express factor showing no erratic behaviour hence simulating realistic EPER profiles, with a run time of 22 ms. This is the express setting used for all calibrations and correction pipelines.

Reduced chi squared ($\chi^2_\nu < 1$), a statistical figure of merit that illustrates normalised goodness of fit is also calculated for the parallel profiles of each phone brightness setting used in the calibration to aid analysis. Figure 6 shows the best ($\chi^2_\nu < 1$) values obtained as a function of signal level in case of 3 (left) and 4 (right) trap species have been used to generate the target synthetic data. Overall the level of fitness is satisfactory ($\chi^2_\nu < 1$) and similar in both case. As already mentioned no significant trend with signal level is observed. Independent of the number of trap species in the target dataset, four trap species always yields a marginally better value. However, simulations using three trap species only can reproduce 3 trap targets well, illustrating the degeneracy problem of fitting CTI trails.

Figure 7 shows fitness maps (fitness as a function of trap densities and release time constants) for a target containing 3 trap species and calibrated using 3 and 4 trap species respectively. The red crosses indicate the location of the true parameters i.e. the trap parameters used to simulate the synthetic target. This provides a visual assessment of the calibration method or the degeneracy of the parameter space. The maps show only the best 10% of all individuals simulated during the calibration. One can clearly identify three regions that converge around the correct trap parameters for the 3 to 3 trap species case (fig. 7 left). Whereas the 4 to 3 trap species case (fig. 7 right) shows a denser area for the trap with the lowest trap release timescale but the remaining data does not show any clear convergence. This suggests that a calibration that overestimates the number of trap species in a target shows a continuum of trap parameter values with no convergence. This, along with the low χ^2_ν values and converging fitness maps, suggests the best parameters found by the calibration scheme should agree with the number of trap species within the target itself.

The next step obtains the values and uncertainties of the release timescales and densities for each trap parameter found and compare this to the real trap parameters used. To do this, the fitness map data is converted into a histogram for each converging area, the mean and standard deviation of which corresponding to the trap species value and error respectively. The table 3 show the best parameters found by each calibration

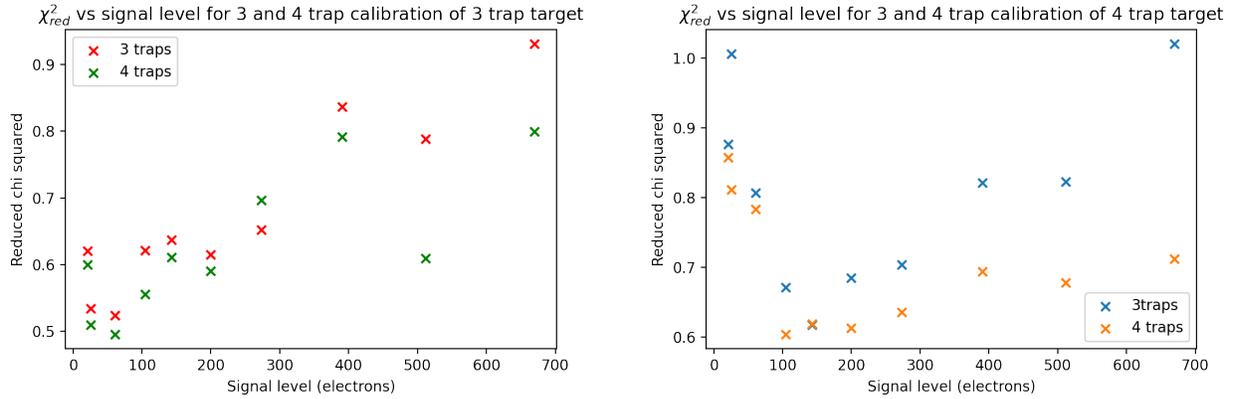


Figure 6. χ_{red}^2 as a function of signal level for the calibration of 3 and 4 traps species against target synthetic data simulated using 3 trap species (left) and 4 trap species (right). A very good level of fitness is achieved for all signal levels (with no clear bias towards lower or higher signal levels). Using 4 traps species shows a marginal improvement in the case of a 3 trap species target and a more significant for the 4 trap species target (right).

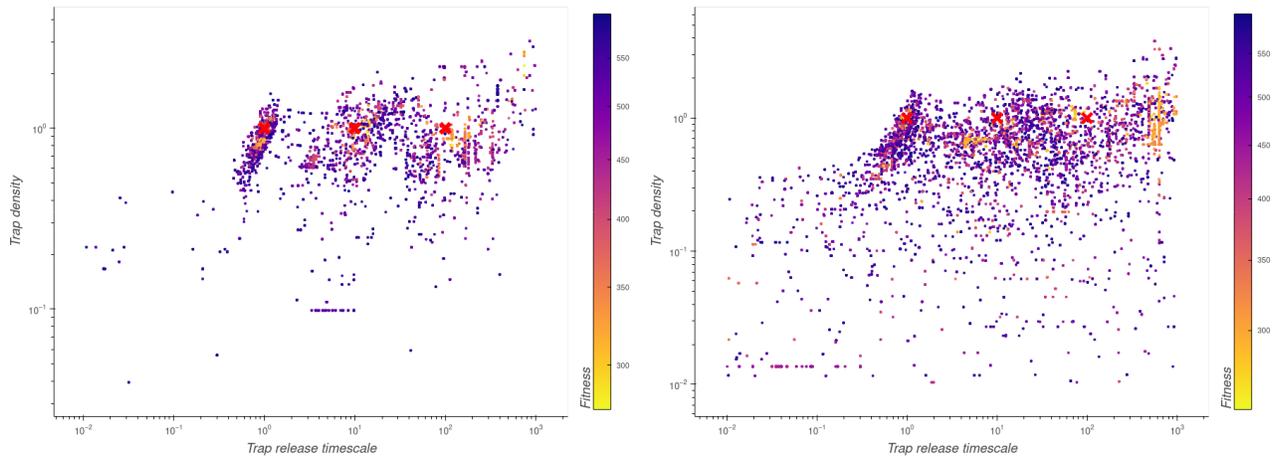


Figure 7. Fitness maps showing the trap parameters (release timescales vs trap density) for the best 10% of all individuals simulates within the 3 (left) and 4 (right) trap species calibration against a target simulated with 3 trap species. For the 3 to 3 trap species case, there are three clear regions of convergence that line up rather well with the true trap parameters i.e. the parameters used to simulate the target profile data (red crosses). Whereas no clear convergence but a larger spread of individuals throughout the parameters space is seen when overestimating the number of trap species (right).

mode. Note that the calibration trying to fit 4 trap species parameters to a 3 trap species target showed no area of convergence, which means it was not possible to obtain values for this calibration run.

The volume parametrization factor (β - the value that relates charge volume to charge density within the simulation) for all calibrations show good agreement with the given value of 0.3 indicating ArCTIC's ability to successfully retrieve this parameter. The 3 trap target, 3 trap calibration simulation indicates good agreement with the known trap parameters added to the target data. The release timescales added to the target all fit within the uncertainties of the release timescales calculated from the calibration. The first two density values also agree and fall within the uncertainties of the calculates densities, however the final trap density lies just outside the upper bound of the calculated trap density. This may indicate an underestimation of the uncertainties calculated at higher release timescale. The 4 trap target 4 trap calibration simulation doesn't show as good an agreement as the 3 trap target 3 trap simulation values, despite the correct number of trap species being

	3 trap target, 3 trap calibration	4 trap target, 3 trap calibration	4 trap target, 4 trap calibration
τ_1	0.9±0.1	0.4±0.1	0.2±0.2
τ_2	9 ± 1	5±2	0.5±0.1
τ_3	134±47	107±183	7±2
τ_4	-	-	92±55
ρ_1	0.86±0.16	1.80±0.2	0.18±0.02
ρ_2	1.01±0.16	1.19±0.1	1.68±0.02
ρ_3	0.82±0.16	1.47±0.5	0.96±0.02
ρ_4	-	-	0.47±0.3
β	0.28±0.04	0.32±0.02	0.3±0.02

Table 3. Trap parameters retrieved from the Pyxel calibration mode for different numbers of traps in the target and in the calibration. Note that the true density is unity for all trap species, and that τ equals 1, 10, 100 in the 3 trap case and 0.1, 1, 10, 100 in the 4 trap one. A β value of 0.3 was used to simulate the targets.

calibrated. The release timescales are underestimated for the first three values, with the final release timescale values calculated showing an agreement with the given release timescale of 100 pixel^{-1} . The trap densities also show similar inaccuracies, the calculated trap densities on traps 1,2 and 3 all indicate a large underestimate, with trap 2 giving a large overestimate.

The way ArCTIC simulates capture and release of charge from traps may hold the reason as to why these values for the 4 trap target 4 trap calibration are rather inaccurate. ArCTIC has the ability to capture released charge from within the same pixel. To simulate this the code first calculates the release of charge from traps and then simulates the capture, then moving to the next pixel and doing the same calculation again. However, this means ArCTIC has a limitation in being able to reproduce release timescales smaller than one line transfer duration. This limitation is indicated most clearly in the 4 trap target, 4 trap calibration, with 2 trap species having release timescales close to or less than 1. This is backed up by the fact that in table 3 the first two release timescale values calculated are the least accurate. It is clear there is greater uncertainty in the release timescale values if they are less than a single line transfer, and future calculations should reflect that. It is unclear if this poses a major problem in the lab data. Traps with a release timescale value smaller than a single line transfer period are more likely to be captured within the same pixel they were produced in. This could mean CTI resulting from this specific trap would not be detected in an image. These smaller release timescale traps does not contribute to CTI as much as higher release timescales traps may do, meaning the ability to discover smaller release timescales may be of lower importance than the ability to accurately simulate the capture of released charge. A trade-off ArCTIC makes.

The 3 trap target 3 trap calibration does show however that the Pyxel pipeline with ArCTIC as a model is capable at reproducing the correct trap parameters to a good degree of precision. Figure 8 (left) show the EPER tails of the simulated and target data for a selection of signal values. The simulated profiles overall show a good agreement and match closely with the target profiles. Figure 8 (right) shows the residual data from all signal levels used in the calibration. Here there is further evidence to suggest a good agreement with the target data and confidence in the ability of ArCTIC to reproduce target files to an accurate level. The residuals for every signal level centred around 0 and are small, indicating only minimal differences, mostly due to the added readout noise. The only region of concern is the first few pixels for higher signal levels, with residuals significantly higher here than at any other point. This artefact could again be caused by the way ArCTIC first releases and then captures charge.

5. CALIBRATING ARCTIC AGAINST REPRESENTATIVE LAB DATA

With the calibration procedure validated and a greater understanding of the level of accuracy that can be expected to be achieved, the next stage is to leverage the calibration mode whilst using lab data. The same calibration, i.e. same simulation inputs and the same number of signal levels as in the synthetic study, procedure is repeated. However, this time with a higher number of individuals (see section 3.2) and this time using target profiles. The parallel charge profiles were extracted from the flat field data of the damaged region of the CCD G quadrant. These profiles, therefore, contain CTI induced from the irradiation of the CCD. The set of signal

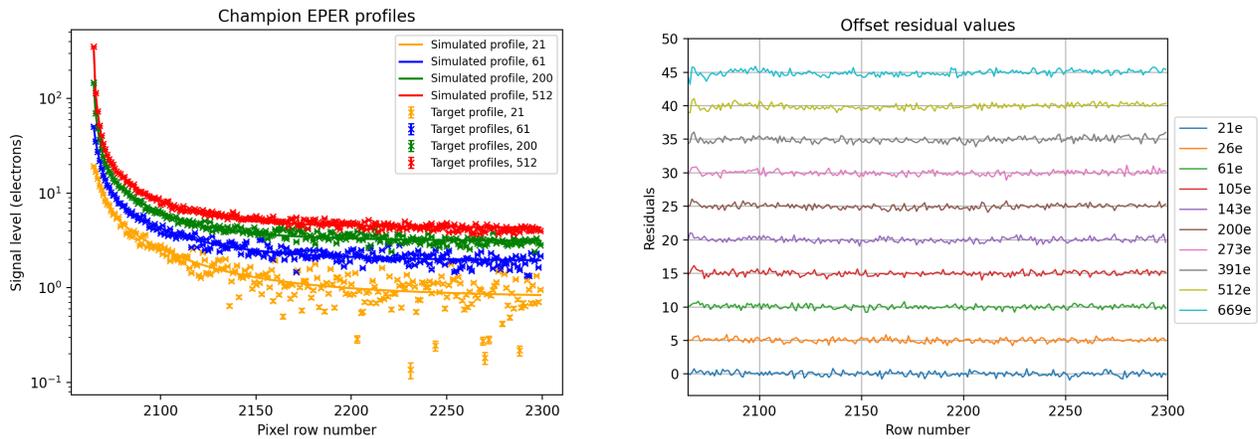


Figure 8. Left: A selection of the EPER tail region parallel charge profiles used as the target in the calibration compared to the best simulated charge tails found by the calibration. The profiles shown are in electrons and are the calibrations best estimate at what the profile should look like using the best trap parameters found. Each set of simulated and target values are offset by 0.5 electrons for clarity Right: Residuals for all signal levels fitted within the calibration (shown in electrons). Each brightness setting is offset by 5 electrons for clarity.

levels considered for the inputs and targets profiles are the same as those used in the synthetic model study (11 in total).

Figure 9 (left) shows a comparison between a subset of the lab target profiles and the best fitting simulations obtained using 3 traps species: the simulated data matches very closely at each signal level. This is compounded by Figure 9 (right) that shows the residuals for every signal level used within the calibration. The residuals are still centered around 0 but are far less noisy than the residuals of the synthetic data, leading to the possibility that the readout noise added to the synthetic target data was too large (indeed the parallel profiles obtained were not only averaged over 100 images but also over 700 columns).

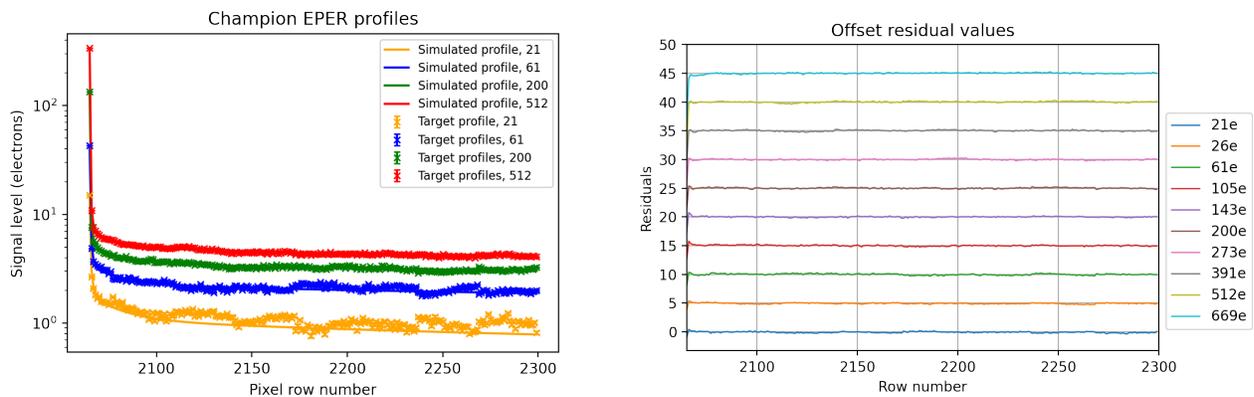


Figure 9. Left: The EPER charge profiles for the lab target data along with the best simulated charge profile that the calibration of three trap species can find. Note that each signal level is offset by 0.5 electrons for clarity. Right: Residual data for all signal levels for the best trap parameters found by the calibration using 3 traps. Each signal level is offset by 5 electrons for clarity.

With confidence that the resulting profiles of the calibrations result in accurate reconstructions of the lab target profiles, the next stage is to investigate the number of trap species present and to extract a set of trap

parameters to proceed with the CTI correction.

Figure 10 shows the χ^2_{red} values for the calibration of 3 and 4 traps species as a function of signal level; the same level of fitness is achieved. From the trends seen previously in the synthetic study, this suggests that the real number of trap species within the CCD with an effect on the EPER data is 3 and not 4.

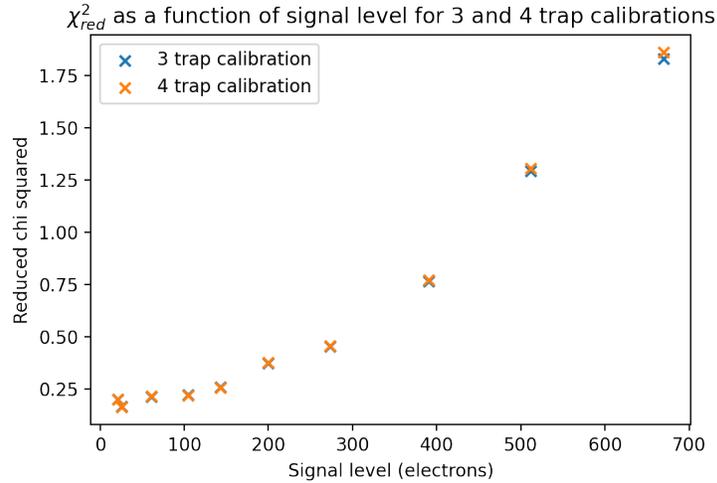


Figure 10. Reduced chi squared values as a function of signal level resulting from the calibration procedure for the laboratory target data when using 3 and 4 trap species. No significant difference in the level of fitness is visible.

To identify the trap species present, fitness maps were once again used to explore the parameter space for the three and four trap species, respectively figure 11 left and right. For 3 trap species case one can clearly distinguish 3 areas of convergence, whilst the 4 trap species case shows a continuum of individuals. This is comparable to what was observed in the synthetic study, strengthening the case for the irradiated CCD containing only 3 trap species. What is also interesting to note is that these fitness maps show less spread than their synthetic counterparts. This could be explained by shortcomings in the generation of the synthetic test data such that the too high level of readout noise, or the higher trap densities.

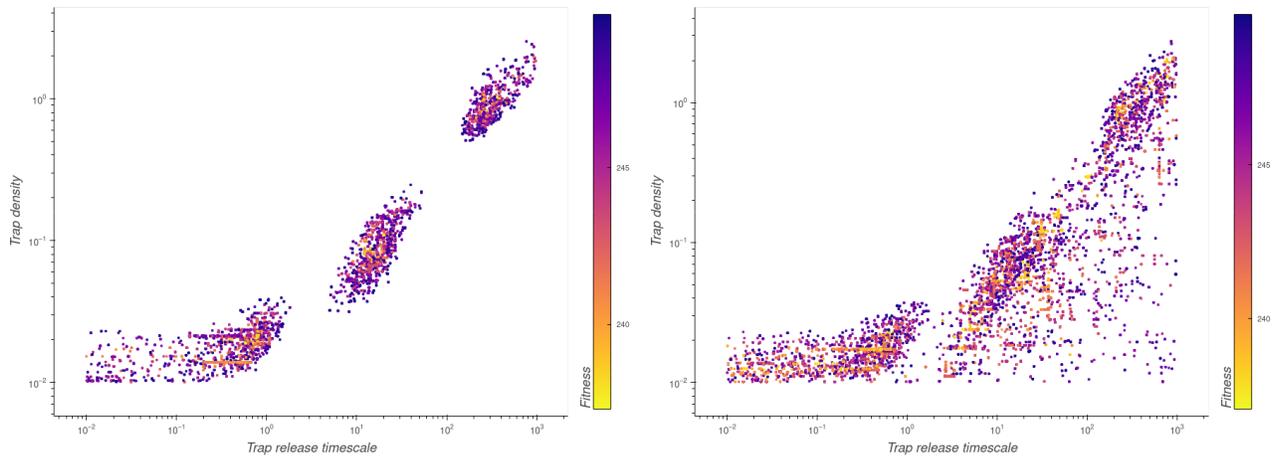


Figure 11. Fitness maps for the calibration of the lab data using 3 (left) and 4 (right) trap species. Only for 3 trap species one can identify clear areas of convergence.

Figure 12 shows the trap density and release time histograms for each trap species generated from the 3 trap species fitness map. The largest spread shown for both density and release time constant is for the fastest trap,

while the longer the release time constant is the more well-behaved the histograms are. As discussed previously, this is expected due to the nature of ArCTIC which simulates less accurately faster trap species close or below the line transfer duration (ArCTIC ultimate resolution).

From the histograms the mean trap parameters and associated uncertainties are extracted. The trap species as well as the β parameter found are listed in table 4. In figure 13, there is a comparison between the results obtained via calibration to the trap species most commonly found in the literature.¹⁵ Unfortunately no trap parameters calibrated in this work match up well with known trap species.

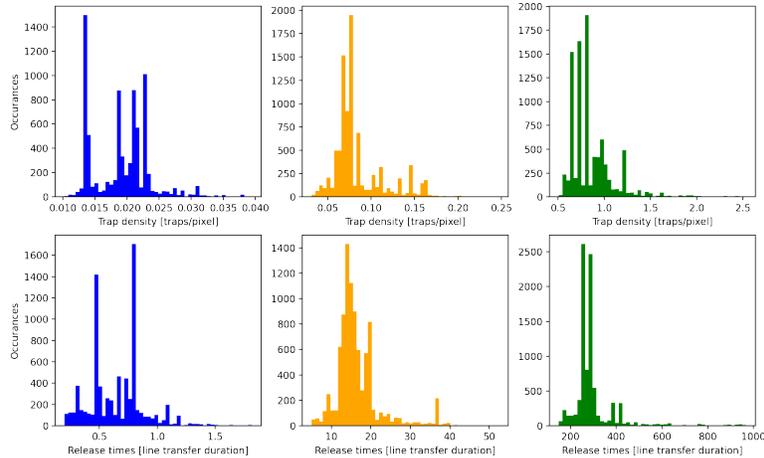


Figure 12. Histograms representing the best individuals, each graph is either the trap density and release timescale of a single trap species.

Release time constants		Densities	Beta
in transfer duration unit	in s	in traps/pixel	No unit
0.66 ± 0.24	0.026 ± 0.009	0.02 ± 0.05	0.21 ± 0.04
16 ± 6	0.623 ± 0.234	0.08 ± 0.03	-
295 ± 97	11.489 ± 3.777	0.9 ± 0.2	-

Table 4. Best trap parameters found from calibration of the lab flat field data.

6. CORRECTING GALAXY IMAGES

The ultimate goal is to perform the correction of a large sample of Euclid-like scenes (field of realistic galaxies) that were projected using the same smartphone projection setup as for the calibration test data and ultimately verify that the weak lensing measurement can be performed on CTI-corrected images to the required level of accuracy. This step is still on-going at the time of writing. The correction uses the calibrated trap parameters. The following represents the first attempt at correction of lab data. First by applying correction the EPER trails in the damaged CTI parallel charge profiles (the same as those first obtained and then calibrated in sections 3.3 and 5 respectively) and then an attempt at correcting faint galaxy images.

The correction principles are explained in,³ on top of the ArCTIC model parameters, there is one other free parameter, the number of iterations. This is set to 3, in accordance to previous studies.⁸ In Pyxel, profiles and images go through a simple Pyxel exposure pipeline, using the ArCTIC remove_cti function as the charge transfer model (instead of the previous add_cti function), with the model parameters used being the trap parameters as calibrated in the previous section. In the first attempt to evaluate the correction efficiency, only a simple comparison between the corrected and the CTI-free EPER trails and images are conducted.

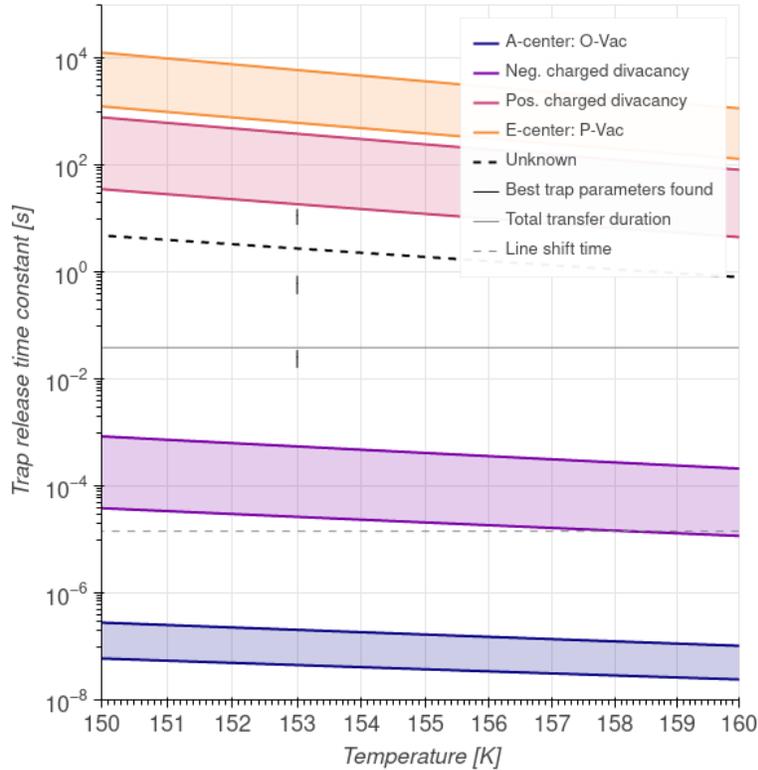


Figure 13. Comparison between the calibrated set of trap parameters and trap species commonly found in the literature (extracted from¹⁵).

Figure 14 (left) shows the EPER profile before (coloured symbols) and after correction (solid line) for two signal levels: 20 (red) and 200 e^- (green). For each signal level, an extra profile is shown (grey symbols), the CTI-free EPER profile. Although a correction clearly took place, the corrected profiles show a small ($<0.5 e^-$) yet constant offset compared to the CTI-free profile limiting the correction efficiency. This requires further investigation.

Figure 14 (right) shows for all signal levels the correction residuals i.e. corrected minus CTI-free parallel charge profiles (in the EPER region): a direct measure of the correction efficiency. At first glance, the overall correction efficiency achieved seems reasonable for the entire profile. But two main limitations require further investigation: (i) the already mentioned offset which is confirmed at all signal levels (despite not being visible on the plot), (ii) the second pixel is over-corrected, meaning the CTI-free profile has higher values than the corrected one (clearly visible in the residuals after correction plot).

Figure 15 show the result of the CTI correction once applied to galaxy images. Once again, trails are removed but it also shows some level of over-correction. A closer look at the galaxy profiles also showed that the correction is not accurate enough and need further work.

7. CONCLUSIONS AND FUTURE WORK

To investigate the best achievable CTI correction efficiency of the ArCTIC model on experimental data in Euclid-like conditions, ArCTIC has been implemented in Pyxel and a calibration procedure using Pyxel's calibration mode has been devised. The devised calibration procedure has been validated using synthetic data generated by the same model to be calibrated. Once validated, the procedure was applied to experimental test data: EPER profiles were extracted from flat fields obtained at different signal levels for a Euclid irradiated CCD. The calibrated trap parameters were then used to correct CTI from the target EPER profiles and Euclid-like galaxy images were obtained using the same experimental setup.

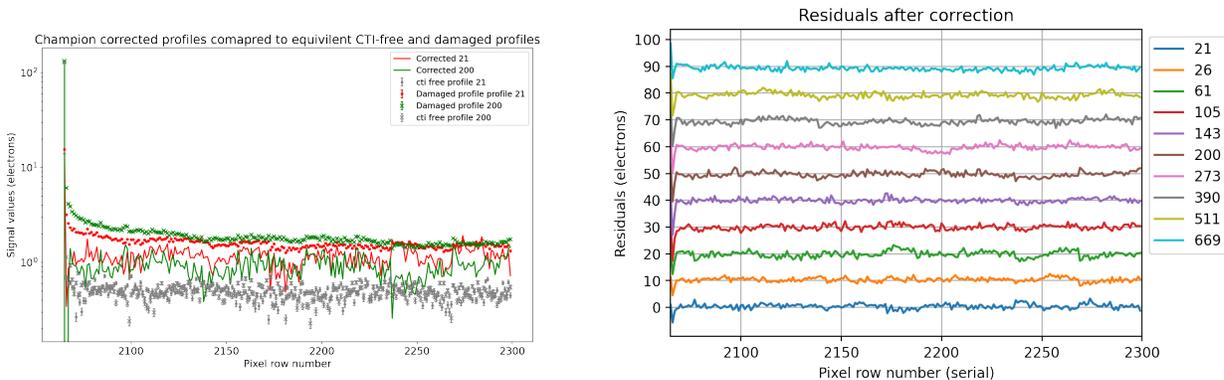


Figure 14. Left: EPER profiles before and after correction for two signal levels. Note that one electron has been added to avoid zero values and enable logarithmic plotting). Right: Residuals after correction i.e. corrected minus CTI-free profiles for all signal levels. Note that each each signal level has been offset by 10 electrons for clarity.

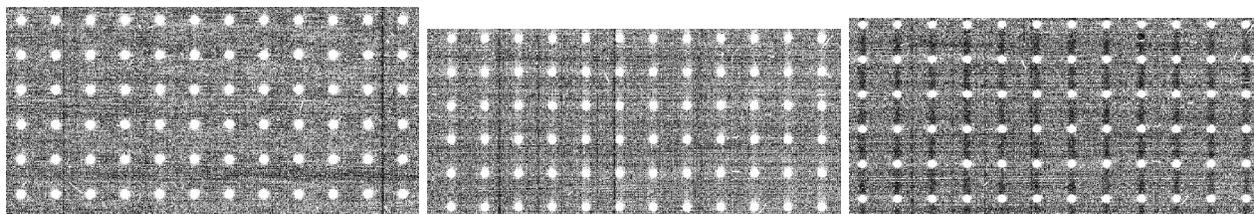


Figure 15. Snippets of Euclid-like galaxy images (average for 100 exposures) obtained using ESA’s SCI-FIV smartphone experiment setup for the CTI-free (left) and damaged (center) sections of the CCD quadrant. In the latter, CTI-induced trailing can be seen despite the very low signal level (less than $20 e^-$ peak). The image on the right show the resulting image after correction.

The calibration procedure validation exercise showed that in ideal but representative conditions (similar data and noise level), the calibration procedure achieve a very high level of fitness for all signal levels and returns trap parameters close to the true parameters (the one used to simulate CTI in the synthetic target data). It also showed that when overestimating the number of trap species (e.g. using four trap species in the model when the data only contains 3), a good level of fitness is achieved but no area of convergence is found in the parameter space, giving us a tool to understand the relevant number of trap species in experimental data.

The calibration procedure applied to the experimental data, enabled us to extract a calibrated set of trap parameters capable of reproducing accurately the target test data. From the fitness maps produced using three or four trap species in the calibration procedure, it can be inferred that only three trap species seem relevant to reproduce the test data. There is also a comparison of the set of trap parameters discovered against well-known trap species commonly found in the literature but there was no clear match.

A first attempt at correcting CTI was then performed on the EPER profiles used to calibrate ArCTIC and Euclid-like scenes. Although some level of correction is visible, the correction efficiency is not yet high enough. A plan to further investigate this issue is underway, which includes stepping back and verifying the correction procedure on synthetic data, validating the correction procedure in a similar fashion than the calibration procedure.

8. ACKNOWLEDGMENTS

The authors would like to acknowledge the whole Pyxel team for their support in assisting in the production of valid Pyxel pipelines for CTI simulations and CTI model calibration.

REFERENCES

- [1] Prod'homme, T., Lemmel, F., Arko, M., Serra, B., George, E., Biancalani, E., Smit, H., and Lucsanyi, D., "Pyxel: the collaborative detection simulation framework," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **11454**, 1145408 (Dec. 2020).
- [2] Arko, M., Prod'homme, T., Lemmel, F., Serra, B., George, E., Bradley, K., Pichon, T., Biancalani, E., and Gilbert, J., "Pyxel 1.0: an open source Python framework for detector and end-to-end instrument simulation," (July 2022).
- [3] Massey, R., Schrabback, T., Cordes, O., Marggraf, O., Israel, H., Miller, L., Hall, D., Cropper, M., Prod'homme, T., and Matias Niemi, S., "An improved model of charge transfer inefficiency and correction algorithm for the Hubble Space Telescope," *Monthly Notices of the Royal Astronomical Society* **439**, 887–907 (Mar. 2014).
- [4] Kegerreis, J., "ArCTIC," (Oct. 2021). original-date: 2020-12-13T09:07:08Z.
- [5] Prod'homme, T., Liebing, P., Verhoeve, P., Menyaylov, I., Lemmel, F., Smit, H., Blommaert, S., Breeveld, D., and Shortt, B., "A smartphone-based arbitrary scene projector for detector testing and instrument performance evaluation," *arXiv:2012.05025 [astro-ph]* (Dec. 2020). arXiv: 2012.05025.
- [6] Short, A. D., Barry, D., Berthe, M., Boudin, N., Boulade, O., Cole, R., Cropper, M., Duvet, L., Endicott, J., Gaspar Venancio, L., Gow, J., Guttridge, P., Hall, D., Holland, A., Israel, H., Kohley, R., Laureijs, R., Lorenzo Alvarez, J., Martignac, J., Maskell, J., Massey, R., Murray, N., Niemi, S. M., Pool, P., Pottinger, S., Prod'homme, T., Racca, G., Salvignol, J. C., Suske, W., Szafraniec, M., Verhoeve, P., Walton, D., and Wheeler, R., "The Euclid VIS CCD detector design, development, and programme status," in [*High Energy, Optical, and Infrared Detectors for Astronomy VI*], Holland, A. D. and Beletic, J., eds., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **9154**, 91540R (July 2014).
- [7] Short, A., Crowley, C., de Bruijne, J. H. J., and Prod'homme, T., "An analytical model of radiation-induced Charge Transfer Inefficiency for CCD detectors," *Monthly Notices of the Royal Astronomical Society* **430**(4), 3078–3085 (2013).
- [8] Israel, H., Massey, R., Prod'homme, T., Cropper, M., Cordes, O., Gow, J., Kohley, R., Marggraf, O., Niemi, S., Rhodes, J., Short, A., and Verhoeve, P., "How well can charge transfer inefficiency be corrected? A parameter sensitivity study for iterative correction," *Monthly Notices of the Royal Astronomical Society* **453**, 561–580 (Oct. 2015).
- [9] Prod'homme, T., Lemmel, F., Arko, M., Serra, B., George, E., Biancalani, E., Smit, H., and Lucsanyi, D., "Pyxel: the collaborative detection simulation framework," in [*X-Ray, Optical, and Infrared Detectors for Astronomy IX*], **11454**, 1145408, International Society for Optics and Photonics (Dec. 2020).
- [10] "Pyxel wesbite." <https://esa.gitlab.io/pyxel>.
- [11] "Pyxel gitlab repository." <https://gitlab.com/esa/pyxel>.
- [12] "PyGMO, the Python Parallel Global Multiobjective Optimizer." <https://esa.github.io/pygmo2/>.
- [13] Biscani, F. and Izzo, D., "A parallel global multiobjective framework for optimization: pagmo," *Journal of Open Source Software* **5**(53), 2338 (2020).
- [14] Yang, X.-S., [*Nature-Inspired Optimization Algorithms*], Elsevier, New York, UNITED STATES (2014).
- [15] Prod'homme, T., Verhoeve, P., Lemmel, F., Smit, H., Blommaert, S., van der Luijt, C., Visser, I., Beaufort, T., Levillain, Y., and Shortt, B., "Comparative Study of Cryogenic Versus Room-Temperature Proton Irradiation of N-Channel CCDs and Subsequent Annealing," *IEEE Transactions on Nuclear Science* **66**, 134–139 (Jan. 2019).