



WP3.7: Prototype LSB-Optimised Sky-Subtraction Package for LSST Software Stack

Deliverable 3.7.4

Project Acronym LUSC-B
Project Title UK Involvement in the Legacy Survey of Space and Time
Document Number LUSC-B-41

Submission date	24/May/2023
Version	1.0
Status	Final
Author(s) inc. institutional affiliation	Aaron Watkins (Hertfordshire), Chris Collins (LJMU), Sugata Kaviraj (Hertfordshire)
Reviewer(s)	John Stott (Lancaster), Lee Kelvin (Princeton)

Dissemination level	
Public	<i>Public</i>

Version History

Version	Date	Comments, Changes, Status	Authors, Contributors, Reviewers
0.1	30/MAR/23	First Draft	Watkins, Collins, Kaviraj, Stott, Kelvin
0.2	28/APR/23	Second Draft	Watkins, Collins, Kaviraj, Stott, Kelvin
1.0	10/MAY/23	Final Draft	Watkins, Collins, Kaviraj, Stott, Kelvin
1.0	24/MAY/23	Approved by LSST:UK Exec	Watkins, Collins, Kaviraj, Stott, Kelvin

Table of Contents

Version History	2
1. Executive Summary	5
2. Introduction	6
3. Synthetic sources	7
3.1. Photometry	10
4. Experimental procedure	11
4.1. Summary of the HSC pipeline sky-subtraction	11
4.2. Experimental setups	12
5. Deliverable 3.7.4	13
5.1. Metrics	13
5.2. Slopes, offsets, and flux change per unit area of injected models	15
5.3. The magnitude of the problem	17
5.4. Conclusions and recommendations	19
Annex A. Acknowledgements	24

List of Figures

1. 3D scatter plots of synthetic source catalogues. The left panel shows our standard catalogue, with parameters magnitude (m_λ), half-light-radius (R_{eff}), and Sérsic index (n) distributed into a $7 \times 7 \times 7$ grid (m_λ and n with linear spacing, R_{eff} in logarithmic spacing). We also injected models with larger stamp sizes, with parameters shown in the right panel. Because the models were larger, we injected fewer models.	7
2. Distribution of injected models in celestial coordinates. Our default catalogue is shown in the top panel, while the large model catalogue is shown in the bottom panel. Colored boxes show full stamp sizes of injected models. Red x's show the center of visit 26060, into which we injected the models, and the red circles show roughly the HSC footprint ($45'$ radius).	9
3. Example images of a single CCD (49 in visit 26060). From left to right: <code>calexp</code> , an image without injected sources; <code>injected_calexp</code> , the same image with sources injected; <code>injected_calexp-calexp</code> , isolated models pre-sky-subtraction; <code>(injected_calexp - injected_skyCorr)-(calexp - skyCorr)</code> , isolated models post-sky-subtraction. The color scales in the right two images are set to showcase structure in the background (-0.5 to 0.5 counts).	10
4. Showcasing the effect of a global linear offset in the background on the surface brightness profile of a single model galaxy. The top-left panel shows the ground-truth (pre-sky-subtraction) surface brightness profile as a solid black line, and the post-sky-subtraction profile as a dotted red line. In the top-right panel, we again show these two profiles, but we have added to the post-sky-subtraction profile the mean value of the linear surface brightness difference between the two curves (0.039 counts). This linear difference is nearly constant across the model profile, hence adding the value back in joins the two curves, resulting in negligible differences. The small panels below show the difference between the black and red curves (ground-truth-post-sky-subtraction).	13

5.	Linear surface brightness difference profiles for a single model under four different experimental cases: <code>bgModel2</code> bin size of 128px, 256px, 512px, and 1024px. For larger bin sizes, the linear surface brightness differences between the ground-truth and post-sky-subtraction models are nearly constant, suggesting that the models are influencing the sky solution only on large scales, whereas for small bin sizes the profile has a definite slope, suggesting the sky solution is influencing as well the shape of the model profile.	14
6.	Linear surface brightness difference profile slopes values for our default model catalogue. Each panel shows this value for one of our eight experiment types, as a function of stamp size. Stamps were drawn to the $\mu_\lambda = 30$ mag arcsec ⁻² isophote, hence stamp size is equivalent to the 30 mag arcsec ⁻² diameter, or D_{30} , which is a function of the models' Sérsic parameters. We label each experiment by its defining characteristic. Units are surface brightness per unit radius, or counts per pixel per pixel.	15
7.	As Fig. 6, but showing values of surface brightness offset vs. stamp size.	16
8.	As Fig. 6, but showing the total flux change per unit area within the largest photometry aperture vs. stamp size.	17
9.	As Fig. 6, but for our larger model catalogue. Most points lie underneath the red line, with values very near 0.	18
10.	As Fig. 7, but for our larger model catalogue.	19
11.	As Fig. 8, but for our larger model catalogue.	20
12.	Difference images of the FFP final <code>skyCorr</code> output for all experimental runs with <code>bgModel2</code> active, with and without models present. Each image is scaled in colorbar between -0.4 and 0.4 counts per pixel, to showcase the small-scale changes incurred by the models' presences.	21
13.	As Fig. 12, but showing the remaining four experiment types.	22
14.	As Fig. 12, but showing <code>skyCorr</code> difference images when injecting the larger model catalogue.	23
15.	Maximum relative change in model surface brightness profiles as a function of stamp size, corrected for global flux offsets induced by the entire model catalogue. In most models, the maximum offset occurs at the lowest surface brightness (roughly 30 mag arcsec ⁻²). The large deviations in the mean and standard deviation curves are due to a lack of models with sizes ~ 300 px.	24
16.	As Fig. 15, but for the large model catalogue. Trends are less clear in this catalogue due to the smaller number of models.	25
17.	Total change in flux (nJy) within the largest photometry apertures around each model, as a function of stamp size, corrected for global offset.	25
18.	As Fig. 17, but for the large model catalogue injection.	26

1. Executive Summary

Low-surface-brightness (LSB) science composes a large fraction of LSST’s potential discovery space, yet it is also uniquely sensitive to the LSST data reduction pipeline. Specifically, sky-subtraction can be biased by LSB flux, resulting in over-fitting and removal of said flux from images. As part of an ongoing project working with Data Management, we tested the existing sky-subtraction used for processing Subaru Hyper Suprime-Cam images, which is likely to serve as a model for the algorithm used for LSST, to determine whether or not it can be modified at run-time to ensure that most or all LSB flux is preserved. Modifying the existing algorithm in this way would serve as a prototype LSB-optimised sky-subtraction package for the LSST software stack.

We conducted a variety of experiments on the pipeline sky-subtraction software (Section 4) by running the task `skyCorr` using a variety of pre-set parameters on images with synthetic sources. Doing this, we have isolated the third step in the `skyCorr` task, `bgModel12`, as the most impactful on LSB (or other extended) sources (Section 5.1). Objects with sizes greater than or equal to the bin size used to produce `bgModel12` tend to systematically lose flux to the sky subtraction, with the amount lost showing increasing scatter about the mean flux loss with increasing model stamp size (Figures 6 through 18). While using larger bin sizes in this step does mitigate these effects, uncertainty in the photometry of the injected models increases for all models types over that for models processed using only `skyCorr`’s first two steps (`bgModel` and `doSky`). Therefore, if some form of the HSC algorithm is to be implemented for LSST data, we conclude the following (Section 5.4):

- To deliver on the full potential of LSST galaxies science, `bgModel12` should not be implemented. Implementing this step both limits the sizes of objects for which reliable photometry will be available in the final coadds, and it decreases the accuracy of photometry for all extended objects, regardless of size.
- If `bgModel12` is to be used, our recommendation to the DM team is that the bin size should be tuned to ensure that LSB flux is preserved on the largest scales possible without sacrificing the quality of the final generation coadds.

Regardless of which procedure is implemented, the impact on LSST’s key science goals will be minimal, as we recommend only altering how the `deepCoadd` is produced. The `deepCoadd_calexp`, used for point source and high-redshift object photometry, will not be affected, and difference images produced for time-series photometry should change little assuming appropriate template images are adopted.

2. Introduction

In our past work ([1], [2], [3]), we demonstrated and quantified how the injection of synthetic model galaxies into Subaru Hyper Suprime-Cam (HSC) images impacts the performance of the HSC data reduction pipeline (DRP; [5]) sky-subtraction algorithm, which is expected to serve as the basis for the sky-subtraction algorithm implemented in LSST. We found that, at the time of that writing, the presence of extended objects resulted in an over-estimate of sky brightness local to said objects, often resulting in a removal of diffuse light from the objects' outskirts. Such performance could potentially have devastating effects on low-surface-brightness (LSB) science. While high-redshift dwarf galaxies (objects with small sizes and low fluxes, (e.g., [6, 9]) are likely preserved by the current pipeline, local Universe dwarfs, intracluster light ((e.g., [7, 8]), and tidal debris with large spatial distribution (e.g., [10]) could potentially be negatively impacted, leading to unreliable photometry of such objects.

Since we concluded our initial assessments, development has continued on the HSC pipeline, and we have performed a variety of tests and experiments on known LSB-flux-preserving sky-subtraction algorithms (Watkins et al., in prep.) in order to determine the best way forward for LSB science with LSST. We found both through our experimentation, and through a close examination of the current (version 6.0.0 of the Rubin Environment, with the w.2023_12 of the LSST Science Pipelines) HSC DRP architecture, that the current HSC pipeline potentially has the machinery to preserve LSB flux down to the survey's expected depths via a change to some parameter values input at runtime. This implies that no new flux-preserving algorithm is required, and that facilitating LSB science with LSST would require only some incremental changes to the current procedure. To this end, we have again tested the current version of the pipeline, examining now in-detail the impact of each of the algorithm's three basic tasks, which allows us to recommend an ideal set of parameters under which LSB flux at different scales could be preserved without the need to write entirely new software. We present these results as follows: in Section 3, we describe our new synthetic source catalogues used for this testing; in Section 4, we describe our procedure for testing the different stages of the sky-subtraction algorithm; and in Section 5, we describe our modified metrics for quantifying the pipeline performance and discuss the results of our various experiments, ultimately offering a recommendation for a set of input parameters which best preserves LSB flux.

3. Synthetic sources

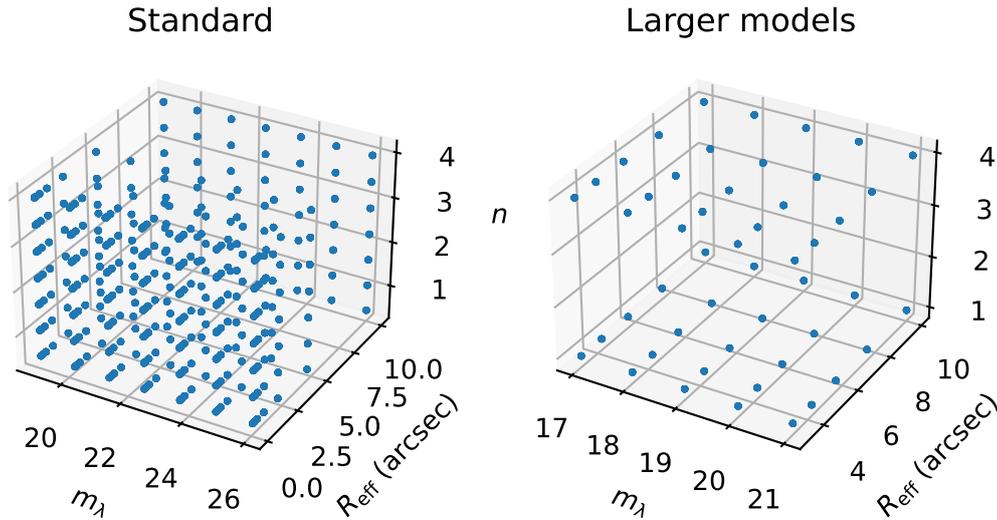


Figure 1: 3D scatter plots of synthetic source catalogues. The left panel shows our standard catalogue, with parameters magnitude (m_λ), half-light-radius (R_{eff}), and Sérsic index (n) distributed into a $7 \times 7 \times 7$ grid (m_λ and n with linear spacing, R_{eff} in logarithmic spacing). We also injected models with larger stamp sizes, with parameters shown in the right panel. Because the models were larger, we injected fewer models.

To test the HSC pipeline sky-subtraction, we developed a new set of synthetic model galaxies, injected using the `source_injection` package (for this report, run using the LSST pipeline Git branch `tickets/DM-34253`, up-to-date as of 24/Mar/2023), specifically the following two tasks:

- `ingest_injection_catalogue` (to ingest the model catalogue)
- `make_injection_pipeline` (to generate the script to run the injection task in the pipeline)

Due to the extensive time required to run the full pipeline (up to the `deepCoadd`), we limited ourselves to running the sky-subtraction software on single exposures, using the `inject_visit` task. This limits the space in which we can inject models. To make the fullest use of this limit, our default catalogue includes only face-on models, with parameters skewed toward much smaller sizes than our previous model catalogue ([1, 2]). To validate the trends we uncovered with this default catalogue, we created a second model catalogue skewed toward larger sizes, albeit with fewer models. We show the distribution of model parameters for both catalogues in Figure 1.

These catalogues take the form of FITS tables, with models spanning $7 \times 7 \times 7$ values ($5 \times 5 \times 2$) in magnitude, half-light-radius, and Sérsic index space for our default (larger model) catalogue, with 20 duplicates of each model type. We also truncated each model to its $\mu_\lambda = 30 \text{ mag arcsec}^{-2}$ isophote, roughly the expected surface brightness limit of the 10-year survey; this determines the stamp size of each model. This was a compromise, allowing us to inject a large number of models with a broad parameter space without significant model-to-model overlap. Regardless, we found that the sky-subtraction algorithm can have noticeable impacts on the models at and above this depth, from which we can draw broader conclusions.

After generating our catalogues, we rearranged all models within the tables by index, to ensure that similar models did not always lie close together on the sky. We then assigned each model

celestial coordinates throughout HSC Tract 9615 (part of the GAMA-15 survey footprint) following a pseudo-random distribution (Fig. 2). The set of models we analyze is thus a subset of the full catalogue we inject. We chose visit 26060 (a g -band image) for our injection as it lies close to the center of this tract, and therefore ensures that all detectors in the image contain some injected models. Specifically, we used this visit from weekly run `w_2022_44/DM-36763`. Of the 6860 (1400) models we created, we ultimately analyzed 3755 (541) within visit 26060 for the default (large model) catalogue. Once injected, we ran the sky correction task on images with models injected and without, and performed comparative photometry on these images to measure our chosen metrics testing the sky-subtraction's performance.

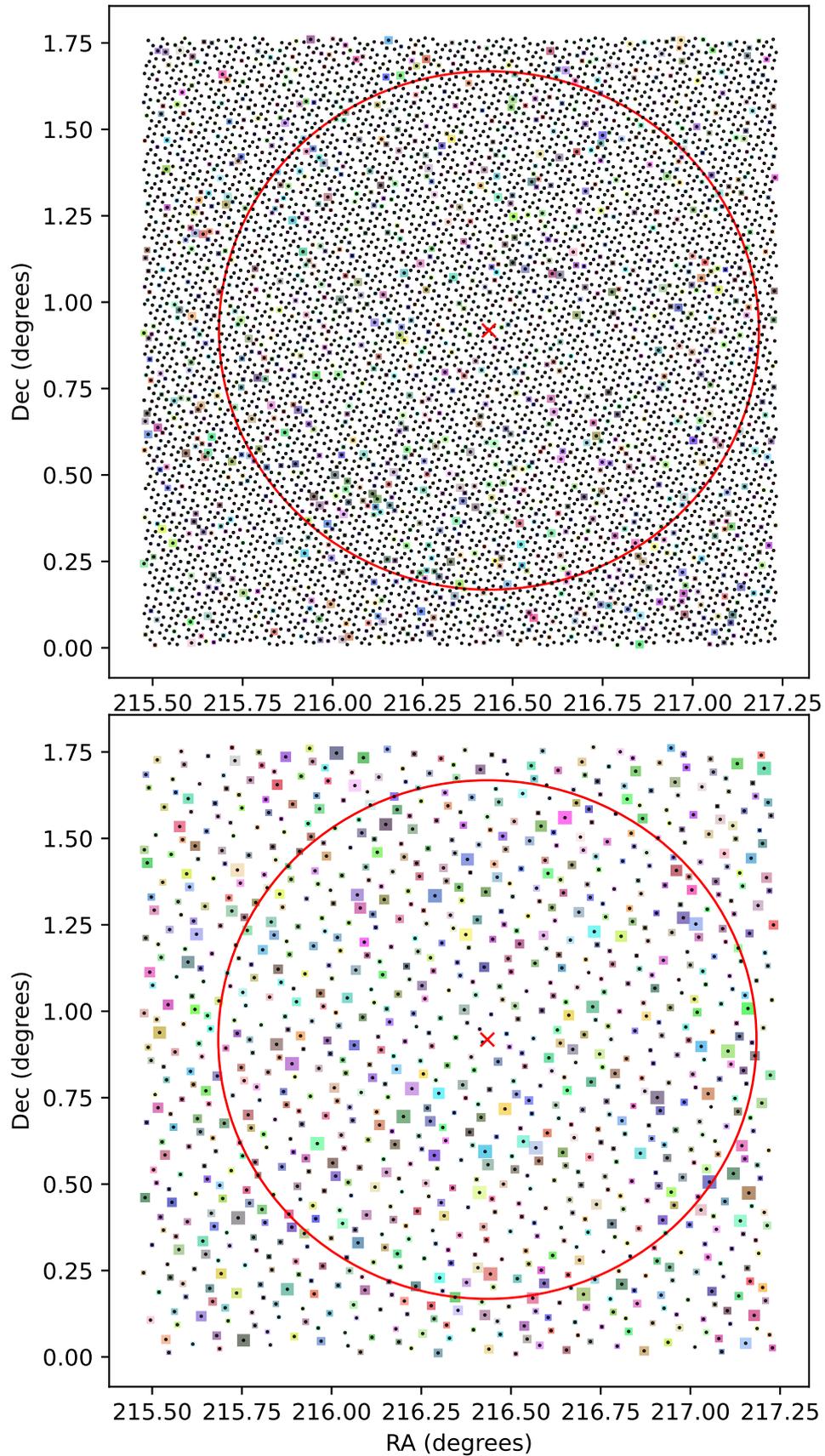


Figure 2: Distribution of injected models in celestial coordinates. Our default catalogue is shown in the top panel, while the large model catalogue is shown in the bottom panel. Colored boxes show full stamp sizes of injected models. Red x's show the center of visit 26060, into which we injected the models, and the red circles show roughly the HSC footprint ($45'$ radius).

3.1. Photometry

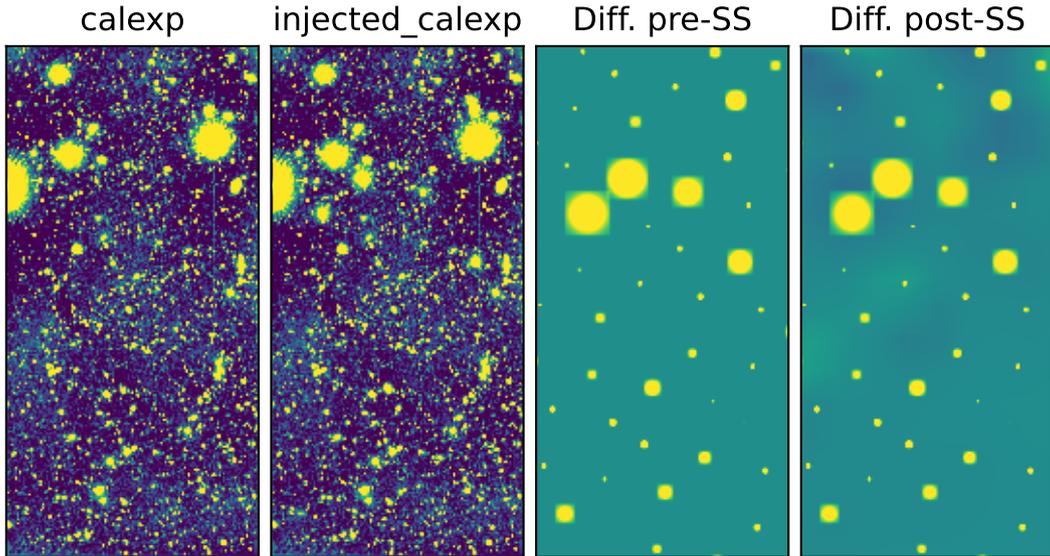


Figure 3: Example images of a single CCD (49 in visit 26060). From left to right: `calexp`, an image without injected sources; `injected_calexp`, the same image with sources injected; `injected_calexp - calexp`, isolated models pre-sky-subtraction; $(\text{injected_calexp} - \text{injected_skyCorr}) - (\text{calexp} - \text{skyCorr})$, isolated models post-sky-subtraction. The color scales in the right two images are set to showcase structure in the background (-0.5 to 0.5 counts).

We derived our metrics via comparative surface photometry, similar to the process described in [3]. To generate ground-truth photometry, we isolated our injected models by subtracting the `calexp` images from the `injected_calexp` images for the same detectors, then performed our surface photometry on these images using circular annulus apertures (all models being face-on) with widths of one pixel, storing the results in Python dictionaries. To generate photometry of the models post-sky-subtraction, we created difference images isolating the models and the sky subtraction through the following process: $(\text{injected_calexp} - \text{injected_skyCorr}) - (\text{calexp} - \text{skyCorr})$. This results in images containing just the injected models, and $\Delta\text{skyCorr}$, the difference between `skyCorr` with the models present and that without. Comparative photometry therefore showcases the impact that the models' presence has on the sky models generated by `skyCorr`, not necessarily the direct amount of flux lost or gained by the sky subtraction procedure. We stored all photometric profiles for all detected sources as pickle files for later retrieval.

We show example images and difference images (on which we performed the photometry) in Figure 3.

4. Experimental procedure

Here we describe our procedure for investigating the existing HSC sky-subtraction algorithm, to determine how well it can be tuned to preserve LSB flux in images. We begin with a brief summary of the existing algorithm, then describe the variety of experiments we conducted on this pipeline by tuning different parameters.

4.1. Summary of the HSC pipeline sky-subtraction

Pipeline task	Description
skyCorr	Full focal plane sky estimation task
bgModel	8192 × 8192px binned estimate of full focal plane sky (large-scale gradient)
doSky	Scaling and subtraction of pre-generated sky frames
bgModel2	$N \times N$ px binned correction of artifacts left from previous two steps

Table 1: Summary of HSC `skyCorr` task

Post instrumental signature removal, each visit-level image in the HSC pipeline is processed through a task called `skyCorr`, used to estimate and remove night-sky emission from HSC images. This process begins at the `calExp` data type, from which one sky estimate has already been subtracted, retrievable as a separate data product via the Butler called `calExpBackground`. This `calExpBackground` was estimated by binning each detector into 128×128 px bins, generating a spline interpolation function from this binned image, then generating a detector-level background model at full resolution using the interpolation function—it is used only for detection and visit-level photometry.

The first step of `skyCorr` itself is to subtract the inverted image of this local sky estimate (hence adding it back into the image), generate object and artifact masks for each detector using the `calExp`-level detections it served to produce, then re-estimate these local sky models from masked images to improve the detection depths. This is iterated three times, generating final generation masks, which are used to ignore flux from astrophysical objects in each subsequent step of the algorithm.

The next step is to estimate the overall gradient in the sky across the entire focal plane, generating a model called `bgModel`. First, the masked detector images generated from the previous step are binned into 8192×8192 superpixels, whose values reflect the clipped mean intensities of unmasked pixels in these bins. These bins are then transformed to focal plane coordinates, to ensure that the model is not influenced by gaps between detectors, and a spline interpolation model is generated from this binned focal plane image. This spline model is then used to reproduce the FFP sky model in detector coordinates, which is then finally subtracted from each detector.

Following this, pre-generated sky frames—maps of transmission and reflectivity variations in the filters and CCD substrates across the focal plane ([4])—are scaled to the sky-subtracted background and subtracted off as a flux-dependent flat-field correction ([4]). This task is called `doSky`.

Finally, a third sky estimation is performed to clean up defects incurred by the previous two ([5]). This step, called `bgModel2`, performs the same task as `bgModel` but uses a much smaller bin size (by default, 256×256 px). A new FFP model is derived from the images already processed by `bgModel` and `doSky`, which is then subtracted off the images in detector coordinates. The full data product this `skyCorr` task produces is also called `skyCorr`, and contains four aspects: an

inverted `caIexpBackground`, and three sky models generated by `bgModel`, `doSky`, and `bgModel2`. Therefore, subtracting a `skyCorr` image from a `caIexp` image produces an image of that detector with the original background restored and the FFP model local to that detector removed. We provide a summary table for this process in Table 1.

4.2. Experimental setups

Name	skyCorr tasks	Bin size	Interpolation
<code>bgModel2-1</code>	<code>bgModel</code> , <code>doSky</code> , <code>bgModel2</code>	1.92 (128px)	Akima
<code>default</code>	<code>bgModel</code> , <code>doSky</code> , <code>bgModel2</code>	3.84 (256px)	Akima
<code>bgModel2-2</code>	<code>bgModel</code> , <code>doSky</code> , <code>bgModel2</code>	7.68 (512px)	Akima
<code>bgModel2-3</code>	<code>bgModel</code> , <code>doSky</code> , <code>bgModel2</code>	15.36 (1024px)	Akima
<code>linear</code>	<code>bgModel</code> , <code>doSky</code> , <code>bgModel2</code>	3.84 (256px)	Linear
<code>noBgModel2</code>	<code>bgModel</code> , <code>doSky</code>	–	–
<code>noDoSky</code>	<code>bgModel</code>	–	–
<code>onlyBgModel2</code>	<code>bgModel</code> , <code>bgModel2</code>	3.84 (256px)	Akima

Table 2: Summary of experiment parameters testing `skyCorr`.

When running the `skyCorr` task, steps can be turned on or off, and different parameters such as bin size and interpolation type can be adjusted at runtime by setting flag values. For example, to adjust the bin size used during `bgModel2`, one can add the following statement to the `pipetask` command: `-c skyCorr:bgModel2.xSize=NN.NN`, where `NN.NN` is a floating point value in units of millimeters. (The scale factor between mm and pixels is 0.015, hence, for example, 1.92 mm=128 px.) In order to test the impact of each sky-subtraction step on our synthetic source catalogue, we injected our synthetic sources into visit 26060, a *g*-band image (chosen because it lies near the center of Tract 9615; see Sec. 3), then ran the `skyCorr` task on both this model-injected image and on an unpopulated image of visit 26060 drawn from one of the weekly pipeline runs. We ran the task eight times in total, changing different parameters each time and recording the resulting `skyCorr` images in separate collections, then performed comparative photometry on the injected models as described in Section 3.1.

We summarize these experimental setups in Table 2. The third and fourth columns of this table refer to `bgModel2`; we chose not to adjust these parameters for `bgModel`. In short, we tested four different bin sizes in the final step `bgModel2`, as well as the impact of turning off `bgModel2` entirely, turning off both `bgModel2` and `doSky`, and using a linear spline interpolation rather than the default Akima. Finally, we tested how much work is being done by `bgModel2` by running only this step but excluding the sky frame subtraction task `doSky`. We now summarize the results of these experiments, ultimately aiming to recommend a course of action for best mitigating the impact of `skyCorr` on LSB science without significantly hampering the data quality of the `deepCoadds`.

5. Deliverable 3.7.4

5.1. Metrics

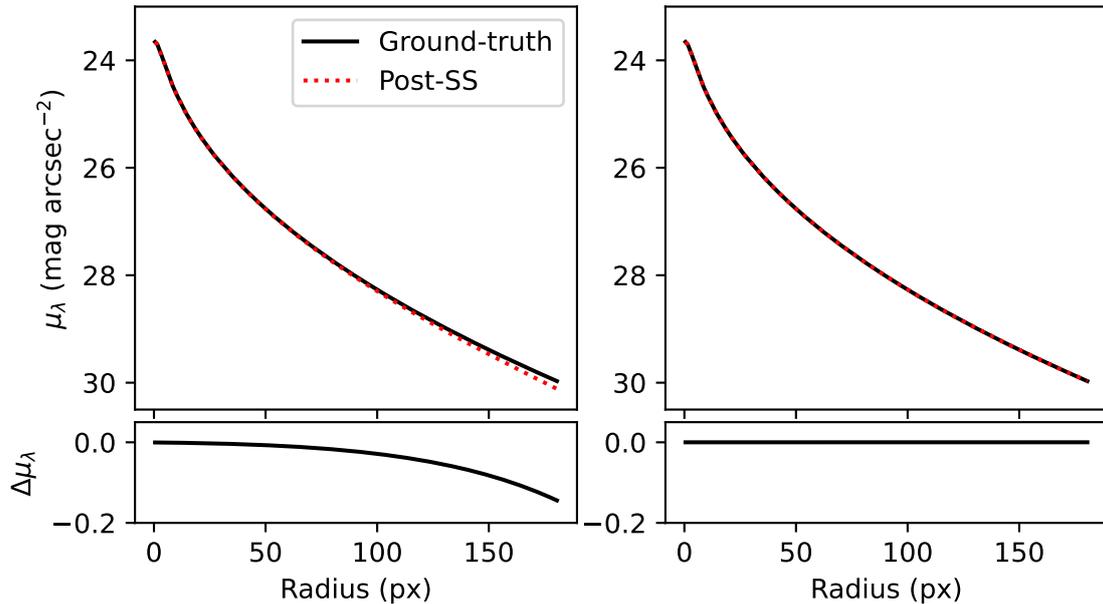


Figure 4: Showcasing the effect of a global linear offset in the background on the surface brightness profile of a single model galaxy. The top-left panel shows the ground-truth (pre-sky-subtraction) surface brightness profile as a solid black line, and the post-sky-subtraction profile as a dotted red line. In the top-right panel, we again show these two profiles, but we have added to the post-sky-subtraction profile the mean value of the linear surface brightness difference between the two curves (0.039 counts). This linear difference is nearly constant across the model profile, hence adding the value back in joins the two curves, resulting in negligible differences. The small panels below show the difference between the black and red curves (ground-truth–post-sky-subtraction).

In our previous experiments ([2]), we measured the performance of the sky subtraction by comparing pre- and post-sky-subtraction surface brightness profiles, identifying the average isophote among all injected models at which the surface brightness profile changed by $0.1 \text{ mag arcsec}^{-2}$ ($\mu_{\Delta\mu=0.1}$) relative to the ground-truth profile. However, in measuring this same metric across models for this new test, we discovered a subtlety which alters the interpretation of the metric. We demonstrate this in Fig. 4. The left panel shows the difference between the surface brightness profiles for an arbitrary large model from our default catalogue ($m_\lambda = 19$, $R_{\text{eff}} = 10$, $n = 2.25$), from experiment `noBgMode12` (turning off the third sky-subtraction step). We show the ground-truth surface brightness profile and the post-sky-subtraction profile as a solid black and dotted red line, respectively. At low surface brightness, the red line falls below the black, suggesting an over-subtraction of flux local to the model. However, we found by examining the linear surface brightness difference (in counts per pixel) between the two profiles that this offset was not due to a systematic over-subtraction as a function of radius, but due to a constant offset difference present across the entire stamp. The right panel in Fig. 4 shows what occurs when we add the mean value of the linear difference profile back to the surface brightness values of the red curve: the two profiles match. As such, a difference in surface brightness is not necessarily indicative of a local change in the sky solution—it could also include the effects of a global change across the image.

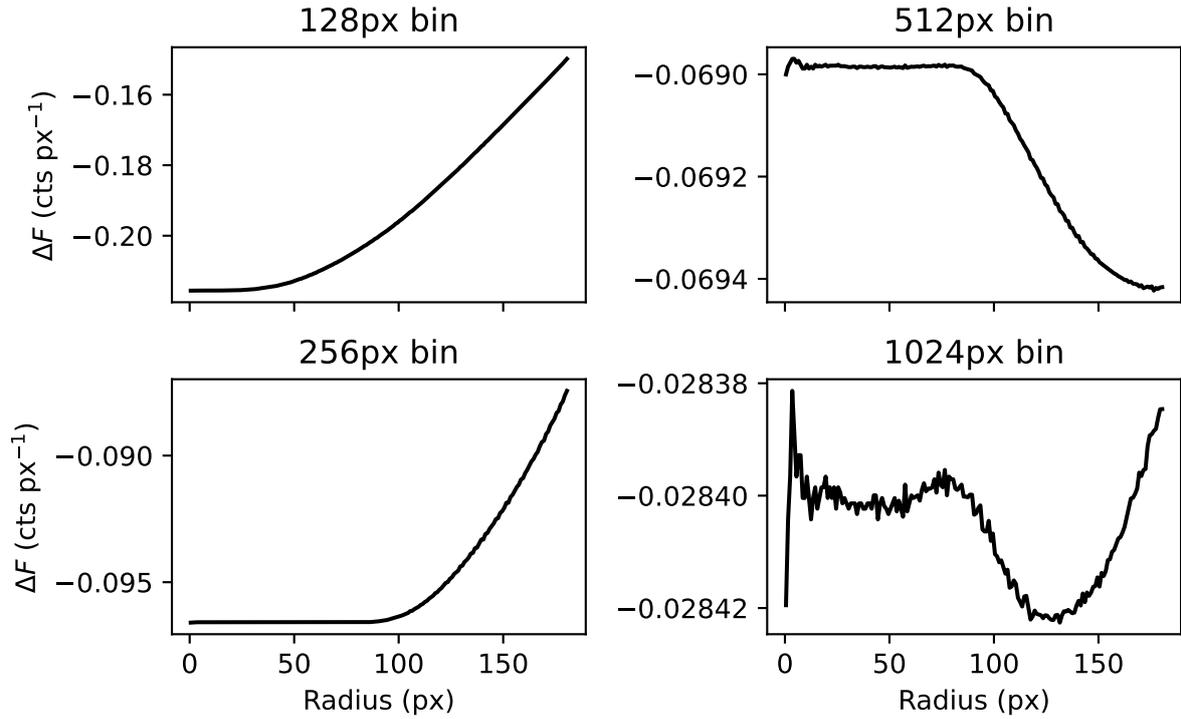


Figure 5: Linear surface brightness difference profiles for a single model under four different experimental cases: `bgModel2` bin size of 128px, 256px, 512px, and 1024px. For larger bin sizes, the linear surface brightness differences between the ground-truth and post-sky-subtraction models are nearly constant, suggesting that the models are influencing the sky solution only on large scales, whereas for small bin sizes the profile has a definite slope, suggesting the sky solution is influencing as well the shape of the model profile.

Our previous experiments did show that local changes can occur, however, mostly around extended objects of order or larger than the bin size used to estimate the sky. Also, a single offset in flux within the model stamp can be a local effect; see, for example, the rightmost panel of Fig. 3. We thus tested whether the sky subtraction was still influencing model profile shapes under the current version of the pipeline by examining the linear difference profiles of models under the four experimental cases in which we altered the bin size in `bgModel2`. We show this test in Fig. 5. Each panel of this figure shows the linear flux difference between the pre- and post-sky-subtraction surface brightness profile of a single model galaxy (the same shown in Fig. 4) under four different experiment cases: `bgModel2-1`, `default`, `bgModel2-2`, and `bgModel2-3` (see Table 2), increasing in bin size from the top-left to the bottom-right. As the bin size increases, the slope of the linear surface brightness difference profile decreases, such that for the largest bin size the profile becomes nearly flat (save for some interpolation artifacts at the level of ~ 0.0001 counts per pixel). The mean value of each profile also decreases with increasing bin size, implying that as larger and larger bins are used, the sky solution local to that model approaches the sky solution when the model was not present at all. Therefore, to preserve the nuances of this behavior, we have opted to examine the slopes and offset values of each model’s linear surface brightness difference profile (hereafter, slope and offset). We also recorded the total flux change per unit area within the largest photometry aperture surrounding each model, $\Delta F/A$, in units of counts pixel⁻², as a better estimate of how much flux is being added or subtracted when the model is present to bias the sky-subtraction. We measure this merely as the total flux within the largest measurement aperture divided by the area of that aperture. We demonstrate the behavior of these metrics for our model catalogues in the following section.

5.2. Slopes, offsets, and flux change per unit area of injected models

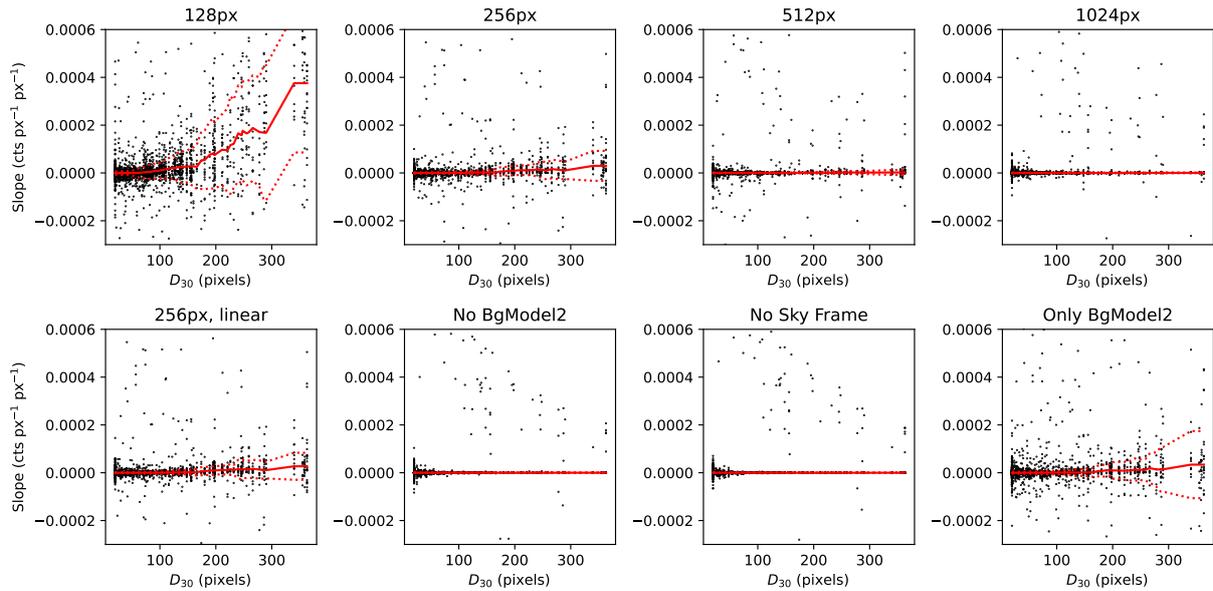


Figure 6: Linear surface brightness difference profile slopes values for our default model catalogue. Each panel shows this value for one of our eight experiment types, as a function of stamp size. Stamps were drawn to the $\mu_\lambda = 30$ mag arcsec $^{-2}$ isophote, hence stamp size is equivalent to the 30 mag arcsec $^{-2}$ diameter, or D_{30} , which is a function of the models’ Sérsic parameters. We label each experiment by its defining characteristic. Units are surface brightness per unit radius, or counts per pixel per pixel.

Figures 6 through 11 show the values of our three metrics described in the previous section plotted against model stamp size. Black points show the value of the chosen metric—slope, offset, and $\Delta F/A$ —while solid red lines and dotted red lines show the running mean and equivalent standard deviation of the points.

We chose stamp size due to how we defined it: each model is truncated to its $\mu_\lambda = 30$ mag arcsec $^{-2}$ isophotal radius, meaning stamp size is equal to the $\mu_\lambda = 30$ mag arcsec $^{-2}$ diameter D_{30} . This value is a function of m_λ , R_{eff} , and n . We found through exploring plots of these metrics against all other parameters that trends with D_{30} incorporate trends with all three other parameters more succinctly. For example, from the top-left panel of Fig. 6, it is clear that larger values of D_{30} correlate with larger values of slope (as well as higher scatter around the mean); we also found that the mean slope and the scatter increase with decreasing m_λ (brighter models) and with increasing R_{eff} (larger models), with no obvious correlation with n . Mechanically, this means that the two factors impacting the sky estimation local to the models are the fraction of unmasked pixels in each sky bin which contain contaminating model flux, and the average intensity of said model flux within these pixels. The same story is thus told using D_{30} alone: the most extended models are the most affected, locally, by the sky subtraction algorithm.

With this in mind, we can examine each of these figures in more detail to obtain a fuller picture of how `skyCorr` is influencing, and being influenced by, the injected models. Evidently, for small bin sizes in `bgModel2`, we see large local influence from the models on the sky solution using all three metrics. For the default catalogue skewed toward smaller models, with bin size equal to 128px, models with $D_{30} \gtrsim 128$ px show increasing slope with increasing size, more negative offset values with increasing size, and more flux lost per pixel with increasing size. Scatter in these three parameters also generally increases with increasing model size. For bin size equal to 256px, we see similar trends above $D_{30} \sim 256$ px, albeit less pronounced. By bin size of

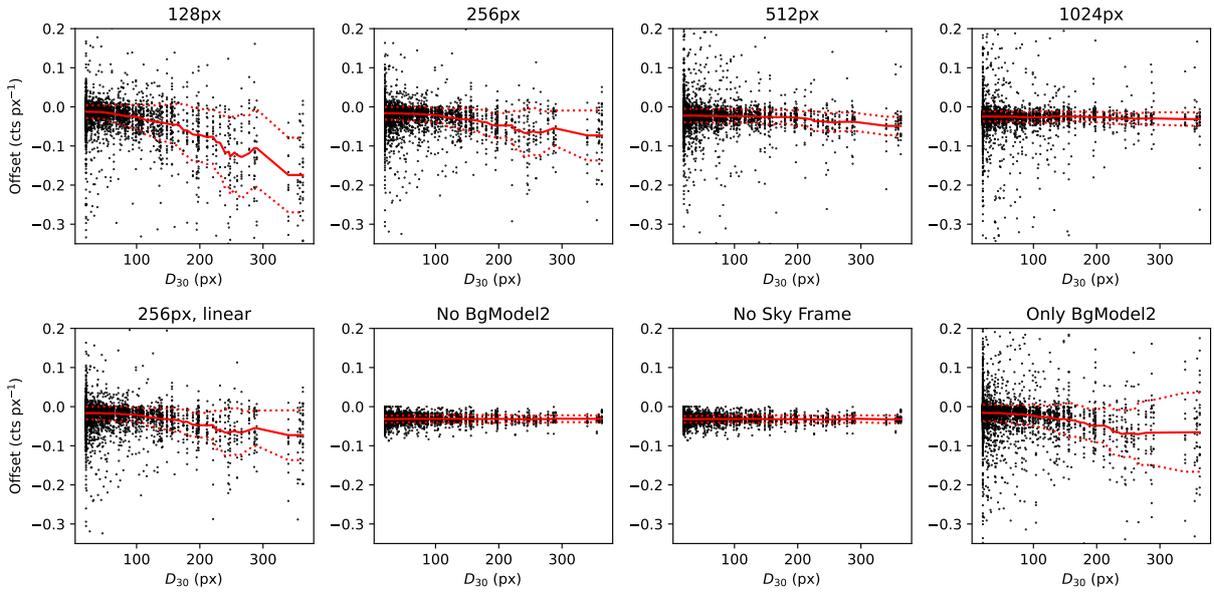


Figure 7: As Fig. 6, but showing values of surface brightness offset vs. stamp size.

512px and higher, the trends have mostly vanished. For the larger model catalogue, we see similar trends beginning at bin size of 256px for models with sizes $D_{30} \gtrsim 300\text{--}400\text{px}$, albeit with more noise given the smaller model sample at any given size. We also see a mild increase in slope, decrease in offset, and decrease in $\Delta F/A$ above $D_{30} \sim 500\text{px}$ when using a bin size of 512px, and no clear trends with either metric for bin size 1024px or when excluding `bgModel2`. Linear interpolation seems to make no difference, nor does using `bgModel2` in lieu of including a sky frame subtraction (although the latter does tend to increase the scatter for all model sizes, hence seems ill-advised). We see the smallest influence on the models when turning off `bgModel2` entirely—without `bgModel2`, or `doSky`, there is no trend with any parameter and stamp size, and the scatter among models is consistent across models and the smallest out of all experiments. The most reasonable conclusion, then, is that `bgModel2` is the most influential step regarding the photometry of extended sources in HSC data.

That said, the models do appear to have some influence on the first two steps as well (`bgModel1` and `doSky`). While the mean value of slope among all models is zero for the `noBgModel2` and `noDoSky` experiments, the mean offset for the `noBgModel2` model injection is ~ -0.023 counts. For the large model catalogue, this increases to ~ -0.031 . Similarly, the mean flux loss per unit area is ~ -0.032 and ~ -0.039 counts for the default and large model catalogues, respectively. This suggests that, while the sky local to the models is not being influenced by the models’ presence when `bgModel2` is excluded, the global mean sky value is being influenced by the presence of extra sources within the image. Potentially this is a result of flux leakage through masks, or possibly undetected sources adding a small pedestal value to the measured sky background, or a combination of both. Because the mean offset is larger for the larger model injection, mask leakage may be the primary cause.

We can see similar behavior when examining the `skyCorr` images themselves, resulting from each experiment run. We show this in Figures 12 through 14, in which, in each panel, we show the difference between `skyCorr` with and without injected models for visit 26060 (retrieved through the Butler as `calexpBackground_skyCorr_visit_mosaic` for images without models, and the same but pre-pended with `injected_` for images with models). Each bright and dark spot indicates a relative under- or over-subtraction of flux imposed by the presence of the models. Such spots tend to trace the locations of injected models, with smaller spots appearing for smaller

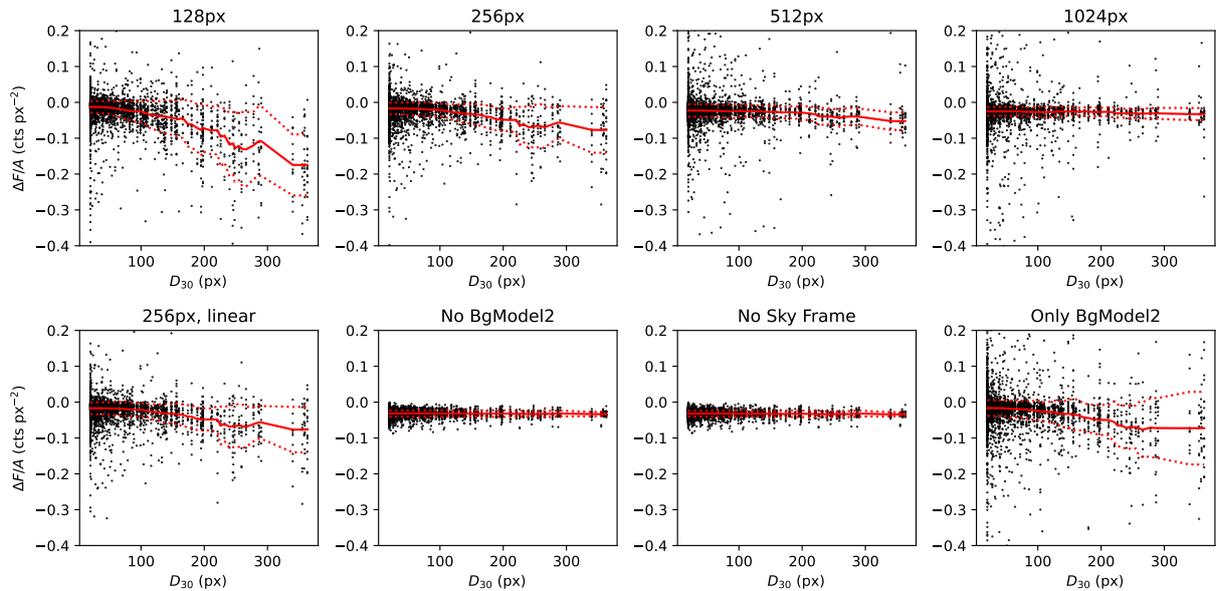


Figure 8: As Fig. 6, but showing the total flux change per unit area within the largest photometry aperture vs. stamp size.

bin size experiments. Interestingly, dark and light spots often appear side-by-side, suggesting that if a local over-subtraction is induced by a model, this local divot carries itself into adjacent regions as an over-correction via the spline interpolation. When `bgModel2` is excluded, the differences between sky solutions are minimal.

5.3. The magnitude of the problem

In the previous section, we demonstrated that `bgModel2` is the step most influenced by, and most influential on, the presence of the synthetic galaxies we injected into our chosen image. Larger bin sizes in this step result in smaller local impacts from the models, although the scatter in the photometry among all model types does increase above runs without `bgModel2` entirely. Evidently, any algorithm involving masking, binning, and local sky estimation will be influenced to some extent by large, bright, and densely populated objects. Whether this influence is considered a problem, however, depends on the magnitude of the effect, or how much flux is lost or gained due to the sky subtraction compared to the null case.

We have been working in counts up until now, as we have been comparing photometry across a visit-level image. However, while the photometric zeropoint differs slightly across the field of view, this difference is fairly small, so we can assume a mean zeropoint for the purposes of converting counts into a more useful diagnostic unit. Among all CCDs in visit 26060, the mean photometric zeropoint is 32.59 mag. These zeropoints are meant to convert instrumental flux into AB magnitudes, which are themselves directly convertible into physical flux units (nJy), which allow us to compare the impact of the sky-subtraction for different broad LSB science cases.

We can begin with global offsets. Even when `bgModel2` is excluded, the injection of additional sources into visit 26060 results in an estimated sky brightness $\sim 0.04\text{--}0.05$ cts px⁻¹ brighter than the estimate when the models were not present. Using the above conversion, this is equivalent to a change of $\sim 0.014\text{--}0.018$ nJy px⁻¹, or a surface brightness change of ~ 32 mag arcsec⁻². Evidently, this is a small change, far below the level of concern for the LSST 10-year survey

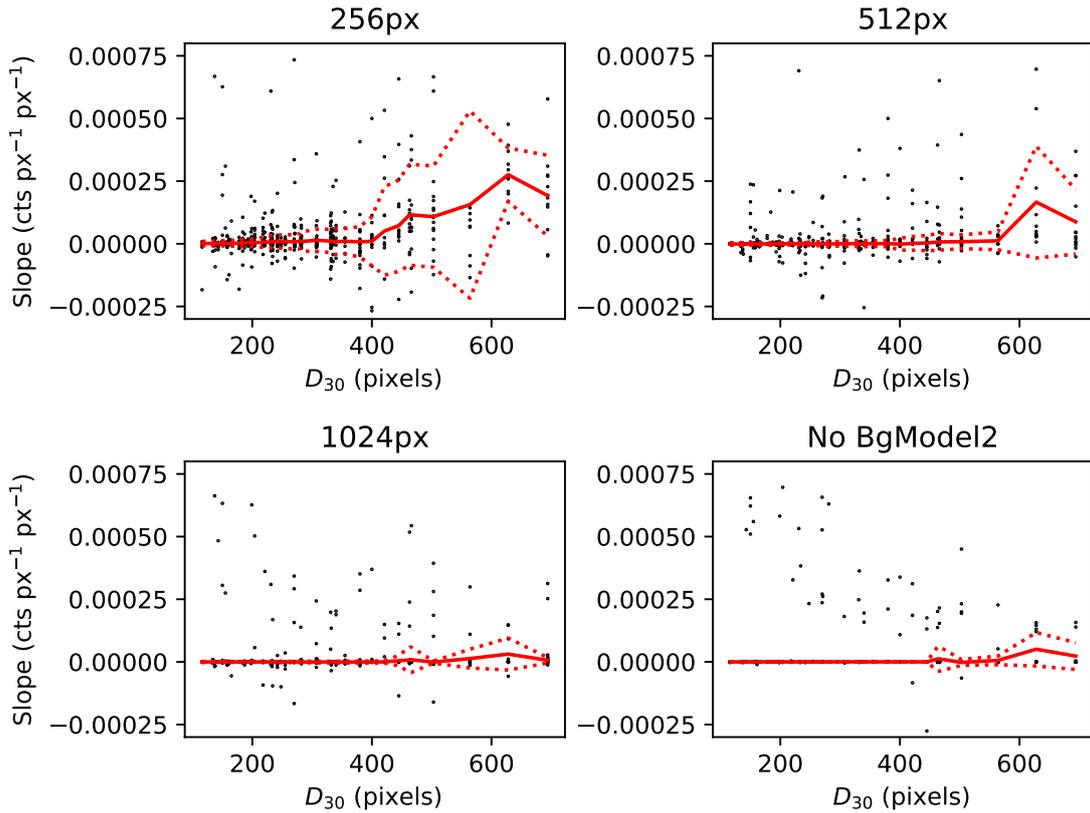


Figure 9: As Fig. 6, but for our larger model catalogue. Most points lie underneath the red line, with values very near 0.

(expected to reach depths of ~ 30 mag arcsec $^{-2}$). Assuming this value continues to scale with the density and brightness of sources populating the images (as hinted at by our two model injection experiments), this offset may become problematic for very dense regions, but for the GAMA field this effect seems negligible.

Regarding local offsets, though our original metric of $\mu_{\Delta\mu=0.1}$ is complicated by the behavior discussed in Section 5.1, we can exploit the global offsets measured from the `noBgModel2` experiment to correct this metric for any global offsets imposed by `bgModel1`, thereby isolating both local changes to the surface brightness profiles and broader local changes to the sky solution in the vicinity of specific models. We show this in Figures 15 and 16. Here we have measured the maximum of the difference between the ground-truth and post-sky-subtraction surface brightness profiles (in mag arcsec $^{-2}$), with all fluxes offset by the mean offset value of the `noBgModel2` run. Because the maximum deviation in surface brightness typically occurs at the largest isophote ($\mu_\lambda = 30$ mag arcsec $^{-2}$), and because we truncated the models at roughly the LSST 10-year depth, these figures show the level of accuracy of photometry at the survey’s noise limit for different kinds of models.

Aside from a sharp jump in between $D_{30} \sim 275$ – 325 caused by a relative lack of models, our default catalogue (Fig. 15) demonstrates the same behavior as discussed in the previous section. For models with sizes close to or larger than the bin size, post-sky-subtraction surface brightness profiles begin to systematically deviate from their ground-truth counterparts, where we see no systematic deviation when excluding `bgModel2`. Profiles begin to deviate by ~ 0.25 mag arcsec $^{-2}$ at their $\mu_\lambda = 30$ mag arcsec $^{-2}$ isophote on average when that isophote falls between 1.5 – $2\times$ the bin size, while the scatter increases sharply beyond the bin size itself. Without including `bgModel2`, the scatter in these values is only ~ 0.02 mag arcsec $^{-2}$, a factor of 3–5 smaller than the mean scatter among other experiment runs. The bigger model catalogue shows somewhat

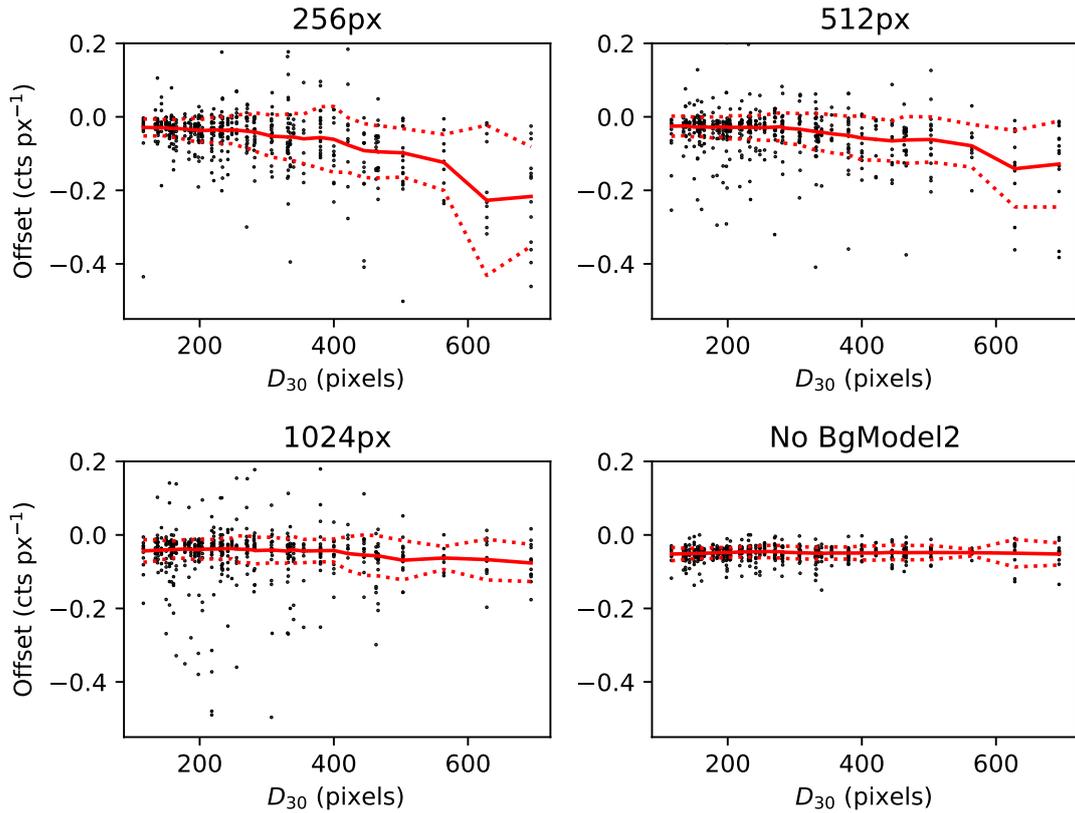


Figure 10: As Fig. 7, but for our larger model catalogue.

similar trends, although the relative lack of injected models introduces significant noise into these measurements (as with our other metrics). Regardless, the `noBgModel2` run shows the same scatter as the default model catalogue experiment.

Finally, we measured the total change in flux within each model’s largest photometry aperture, corrected for the global offset as described above. We show these trends in Figures 17 and 18. We see trends consistent with all previous trends here as well, such that models as large as or larger than the bin size systematically lose flux due to the sky subtraction, with the amount of flux lost increasing with increasing model size. The large model tests show similar behavior, again with more scatter due to the smaller number of injected models. At worst, the models are losing thousands of nJy in total flux (equivalent to a $m_\lambda \sim 25\text{--}23$ mag object).

5.4. Conclusions and recommendations

We can use these results to explore how the current pipeline might affect different LSB sources. The relevant scale here is 256px, or $43''$ at the HSC pixel scale of $0.168 \text{ arcsec px}^{-1}$ —given the above results, anything with sizes between $1'\text{--}1.5'$ would likely suffer significant changes to radial profile shape, as well as significant losses in total flux, due to the sky subtraction routine.

For scale, at $z = 0.05$, $1 \text{ kpc} \sim 1''$. Dwarf galaxies are likely safe from systematic flux changes under the default pipeline; a typical dwarf in the Fornax cluster, for example, has a half-light-radius of $\sim 0.85 \text{ kpc}$ ([11]), hence $\sim 0.85''$ at $z = 0.05$, evidently decreasing in size with redshift. The outer disk of a Milky Way-sized galaxy, at which radius the stellar halo begins to dominate, is typically of order 20–30 kpc (hence $\sim 20''\text{--}30''$ at $z = 0.05$); some low-redshift stellar halos therefore may begin to suffer flux loss, as would tidal features with lengths at the scale of these

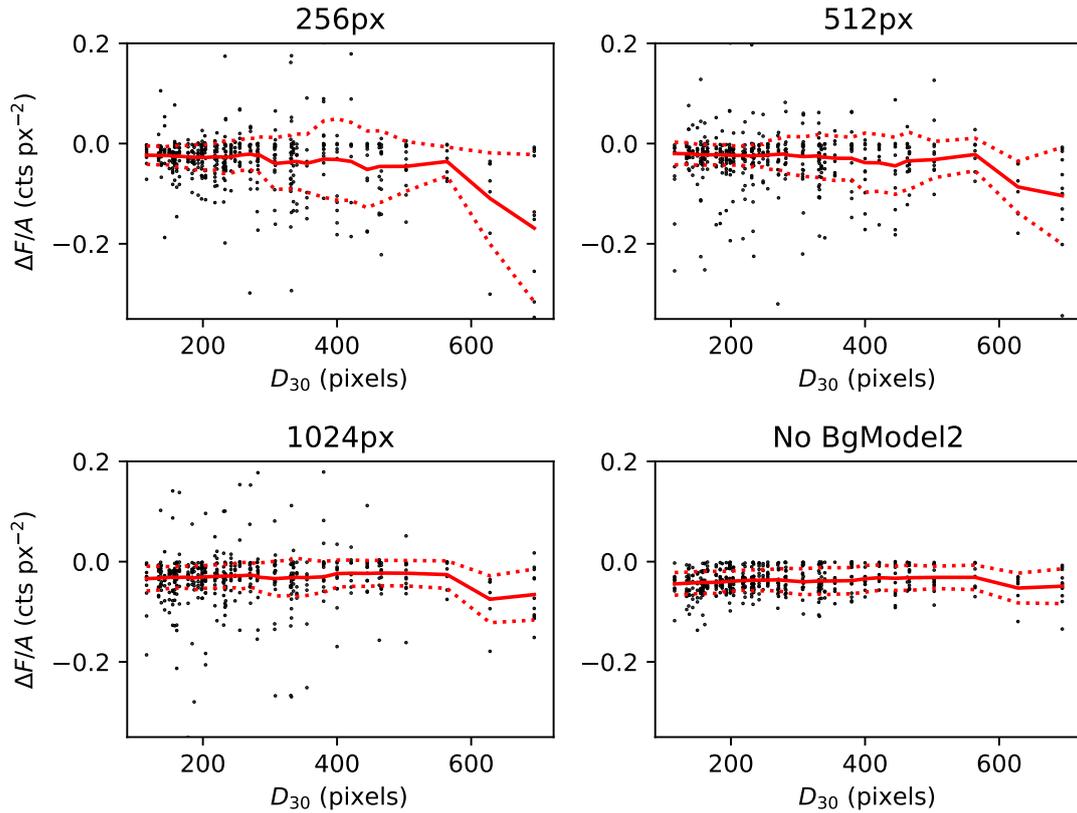


Figure 11: As Fig. 8, but for our larger model catalogue.

stellar halos. The algorithm’s impact would be largest for intracluster and intragroup light: even at $z = 0.3$, ICL tends to have a half-light-radius of $\sim 100''$ (C. Collins, private communication), hence ICL even at this high a redshift will likely lose a significant fraction of its total light, with stronger losses at lower redshift, where it will be more easily detectable.

The overall conclusion of this work, then, is that while `bgModel2` serves to clean the backgrounds of images, it does so at the expense of LSB flux for objects with large angular size. Should the current HSC sky-subtraction algorithm be implemented without alteration as part of the LSST pipeline, LSB science will be possible, but it will be limited in scope, excluding ICL, IGL, or other large diffuse objects such as extremely large LSB disk galaxies or large stellar halos at low redshift. Such objects will suffer systematic losses of flux and deviations in the slopes of their surface brightness profiles, as well as decreased reliability of photometry given the increased uncertainty induced by variations in the sky solutions from image to image. Additionally, smaller objects found near such large LSB objects (e.g., background galaxies embedded in projection within ICL) will suffer similar effects, with the increase in scatter being perhaps the most concerning.

Ultimately, the ideal solution for LSB science would be to do away with `bgModel2` entirely—for injected sources, `bgModel` and `doSky` showed minimal systematic impact on photometry for all model types we injected, and all of our chosen metrics showed the lowest scatter for these experiments among all experiments attempted.

Whether it is feasible to turn off this step in the final pipeline depends on why it was necessary in the first place (i.e., the reason such a cleanup step was required: [5]). For example, should some of the background variation vanish with the implementation of extended point spread function subtraction and reflection removal (currently the subject of an in-kind contribution: PI Sarah Brough), `bgModel2` may become redundant, and so can safely be excluded. This is the

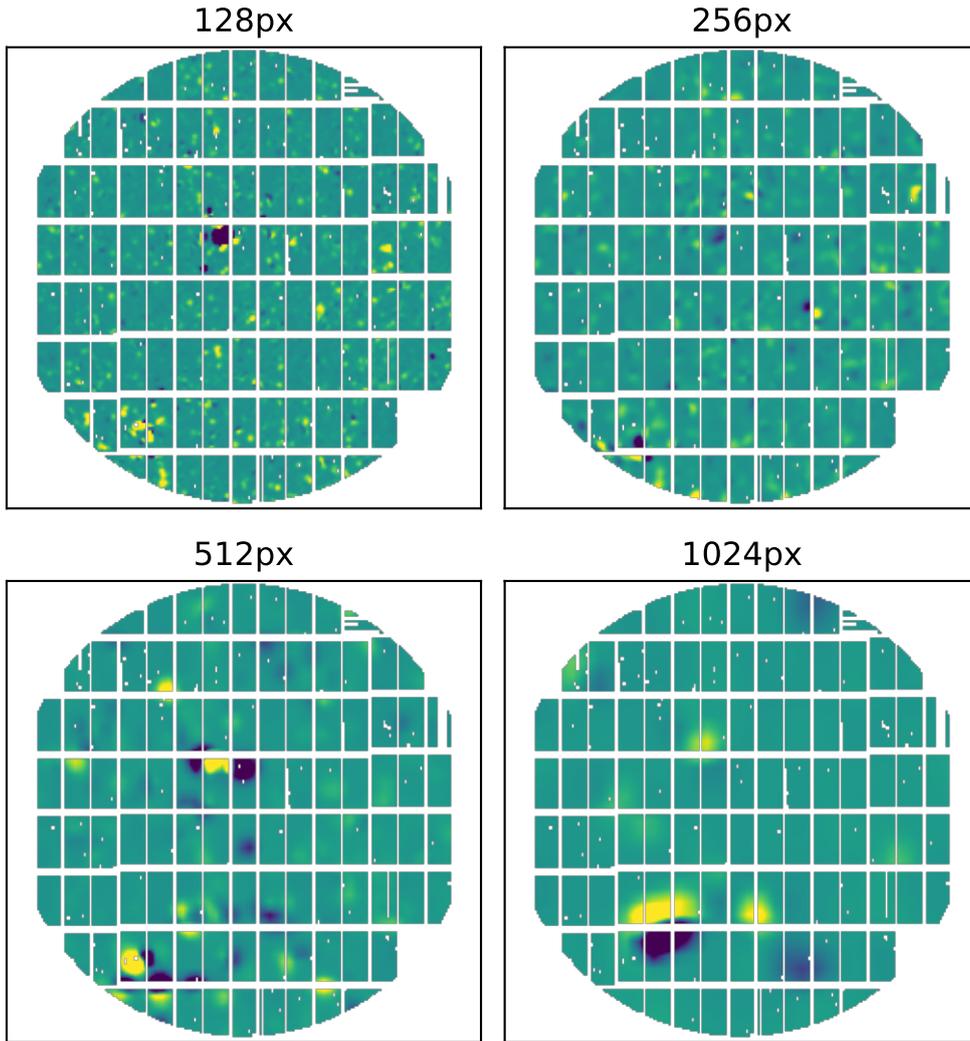


Figure 12: Difference images of the FFP final `skyCorr` output for all experimental runs with `bgModel12` active, with and without models present. Each image is scaled in colorbar between -0.4 and 0.4 counts per pixel, to showcase the small-scale changes incurred by the models' presences.

ideal case for LSB science, as the current pipeline with the addition of PSF-subtraction and the exclusion of `bgModel12` mimics closely sky-subtraction procedures already used in LSB-oriented observations currently achieving depths similar to what is expected for LSST (e.g., [12, 13]). We note that DM has discussed this issue as well and have come to similar conclusions [14], suggesting that our recommendations could be accommodated without impacting LSST's key science goals. In the absence of such a solution, however, we recommend at least using the largest bin size possible to balance the flatness of the visit-level backgrounds with the loss of flux and reliability in the photometry of large objects.

The synthetic source catalogue FITS tables used in this work are available upon request. Images generated using these catalogues are stored at the US Data Facility for Rubin Operations at the SLAC National Accelerator Laboratory, hence are accessible only by those with the appropriate permissions.

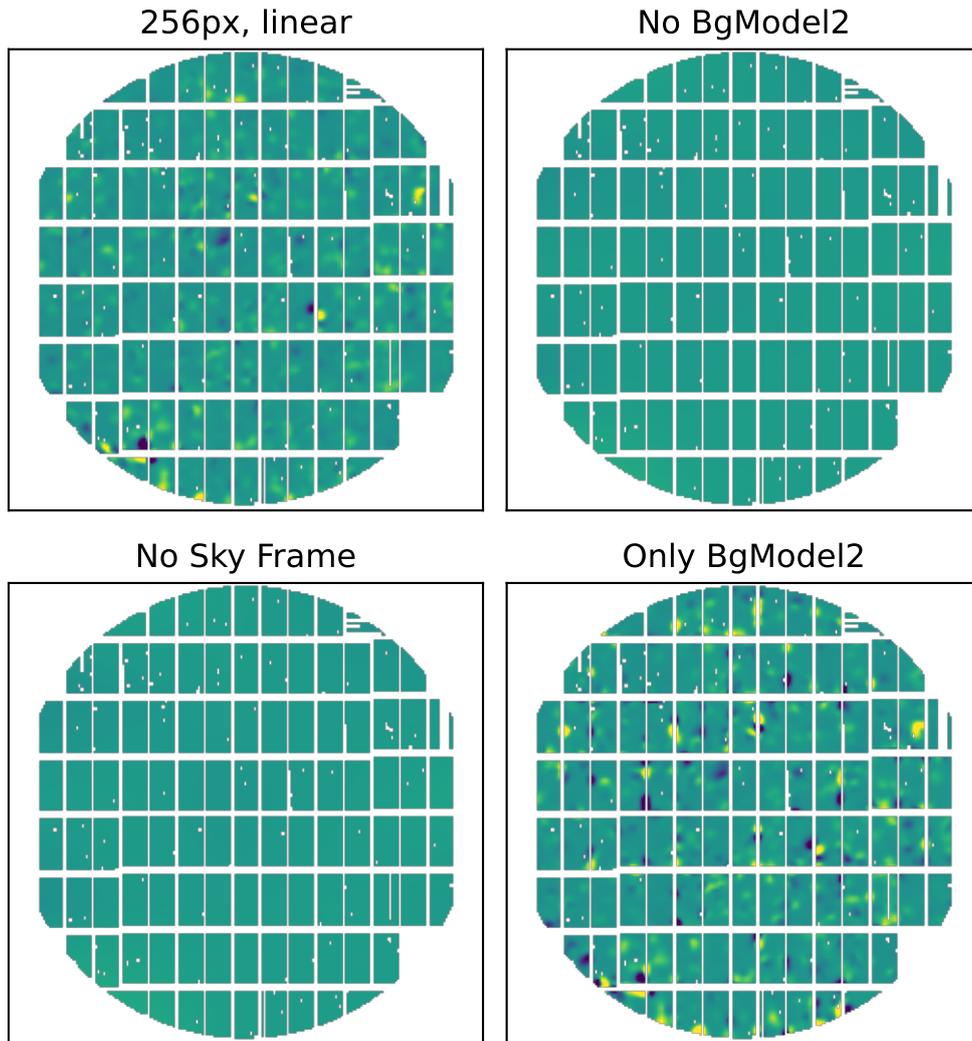


Figure 13: As Fig. 12, but showing the remaining four experiment types.

References

- [1] Report on optimal metrics for measuring the impact of the LSST pipeline sky subtraction on low-surface-brightness flux at different spatial scales, Project Deliverable D3.7.1
- [2] Report on Mock Testing Results, Project Deliverable D3.7.2
- [3] Software to Output Metrics That Keep Track of Improvements to the Pipeline Sky Subtraction, Project Deliverable D3.7.3
- [4] Aihara, H. et al., PASJ, 71, 114
- [5] Aihara, H. et al., PASJ, 74, 247
- [6] Dalcanton J. J., Spergel D. N., Gunn J. E., Schmidt M., Schneider D. P., 1997, AJ, 114, 635
- [7] Burke C., Collins C. A., Stott J. P., Hilton M., 2012, MNRAS, 425, 2058
- [8] Montest, M., Nature Astronomy, 6, 308

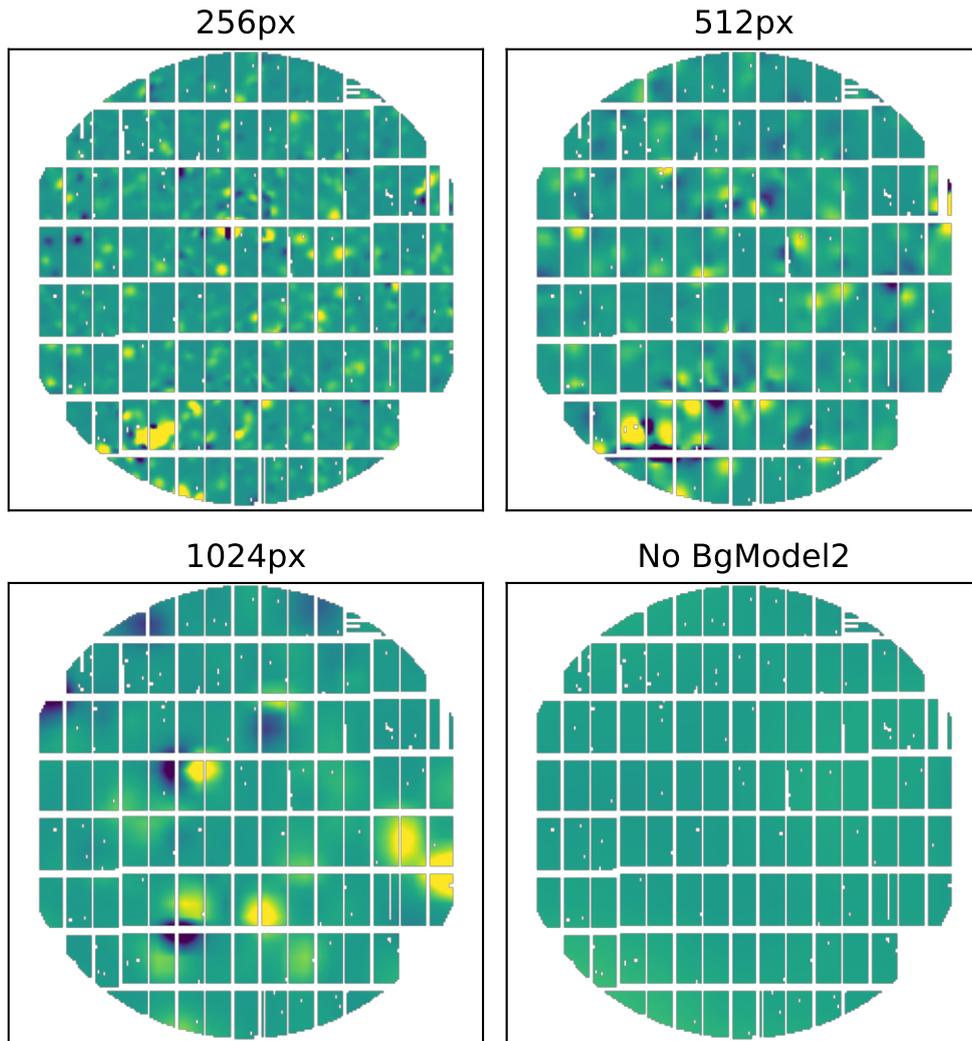


Figure 14: As Fig. 12, but showing skyCorr difference images when injecting the larger model catalogue.

- [9] Martin G., et al., 2019, MNRAS, 485, 796
- [10] Kaviraj, S., 2014, MNRAS, 440, 2944
- [11] Watkins, A., Salo, H., Kaviraj, S., Collins, C. A., Knapen, J. H., Venhola, A., Román, J., MNRAS, 521, 2012
- [12] Mihos, J. C., Harding, P., Feldmeier, J. J., Rudick, C., Janowiecki, S., Morrison, H., Slater, C., Watkins, A., ApJ, 834, 16
- [13] Trujillo, I., et al., AA, 654, 40
- [14] Kelvin, L, et al., Jira Ticket, <https://jira.lsstcorp.org/browse/DM-37653>

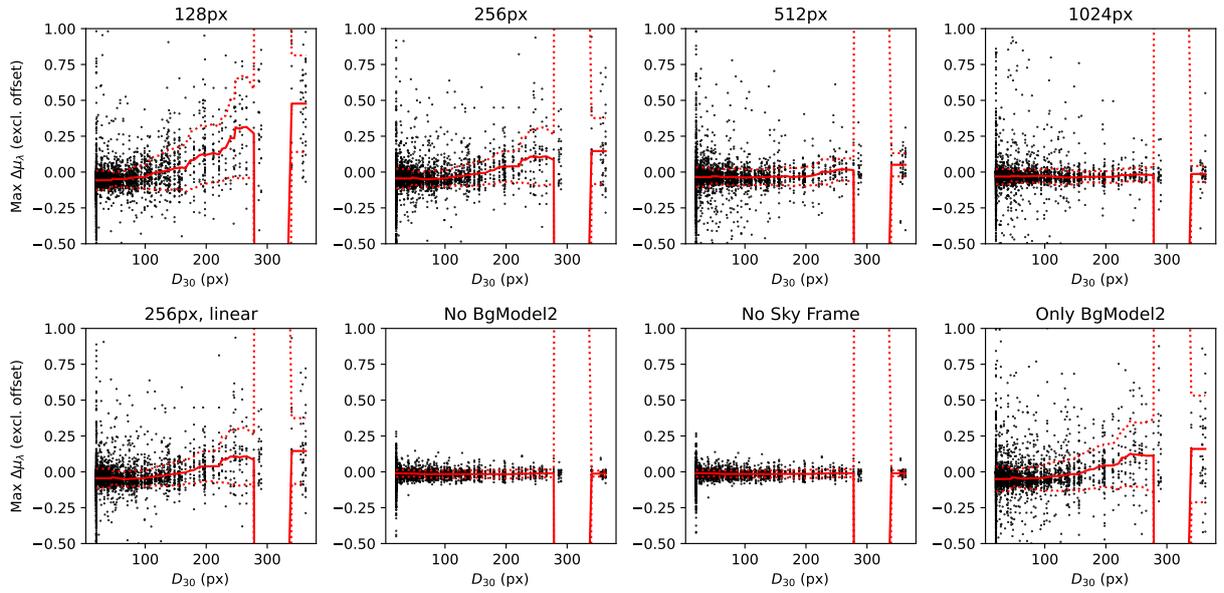


Figure 15: Maximum relative change in model surface brightness profiles as a function of stamp size, corrected for global flux offsets induced by the entire model catalogue. In most models, the maximum offset occurs at the lowest surface brightness (roughly $30 \text{ mag arcsec}^{-2}$). The large deviations in the mean and standard deviation curves are due to a lack of models with sizes $\sim 300\text{px}$.

Annex A Acknowledgements

We are grateful to the members of the DM pipeline team who have worked with us throughout this project: Lee Kelvin, Yusra AlSayyad, Robert Lupton, Sophie Reed, and Dan Taranu. Without them, this project will not be possible.

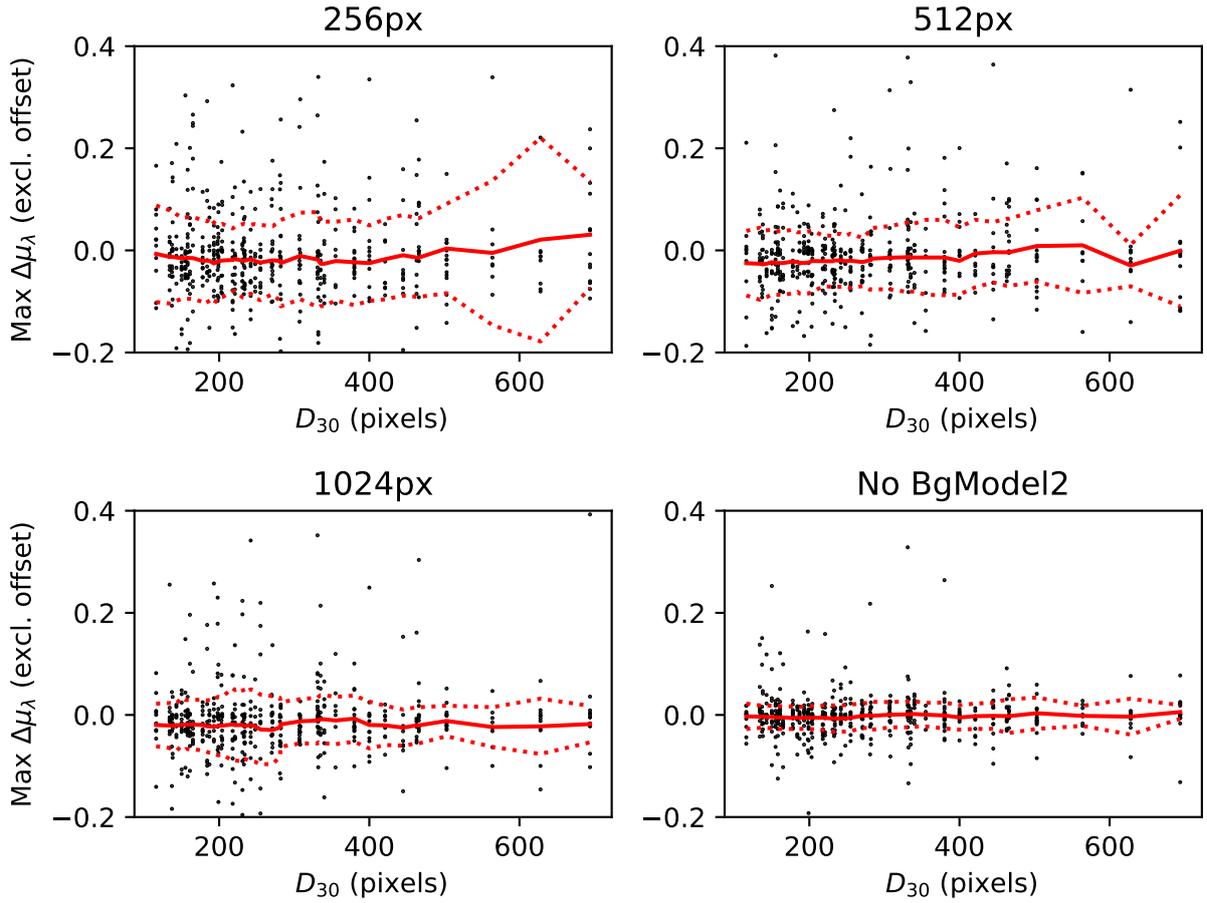


Figure 16: As Fig. 15, but for the large model catalogue. Trends are less clear in this catalogue due to the smaller number of models.

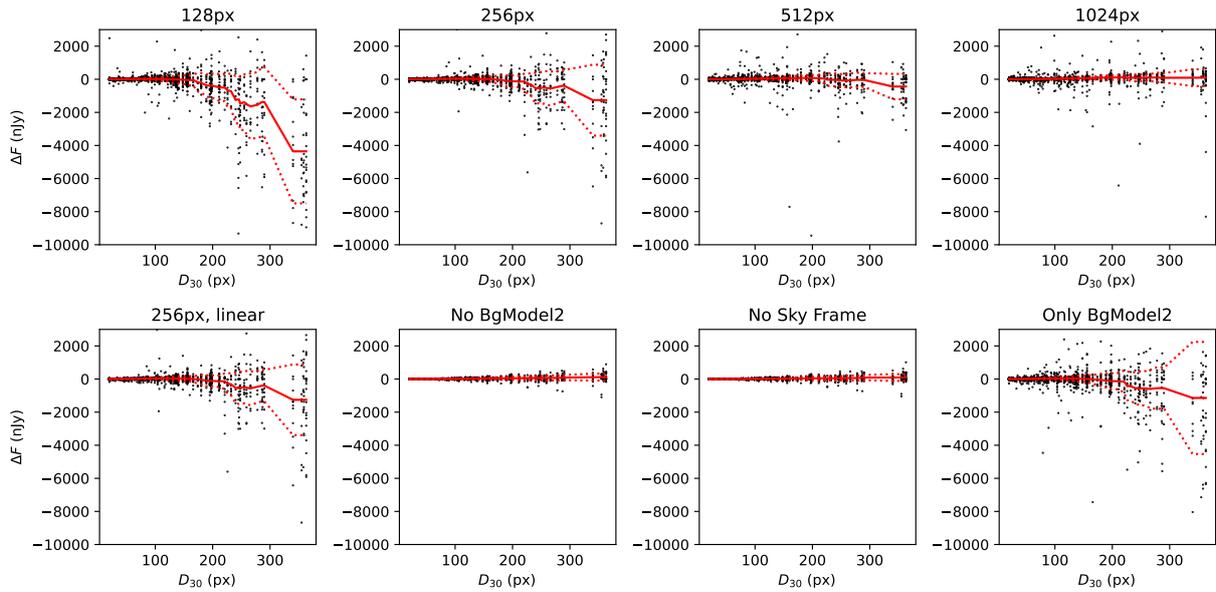


Figure 17: Total change in flux (nJy) within the largest photometry apertures around each model, as a function of stamp size, corrected for global offset.

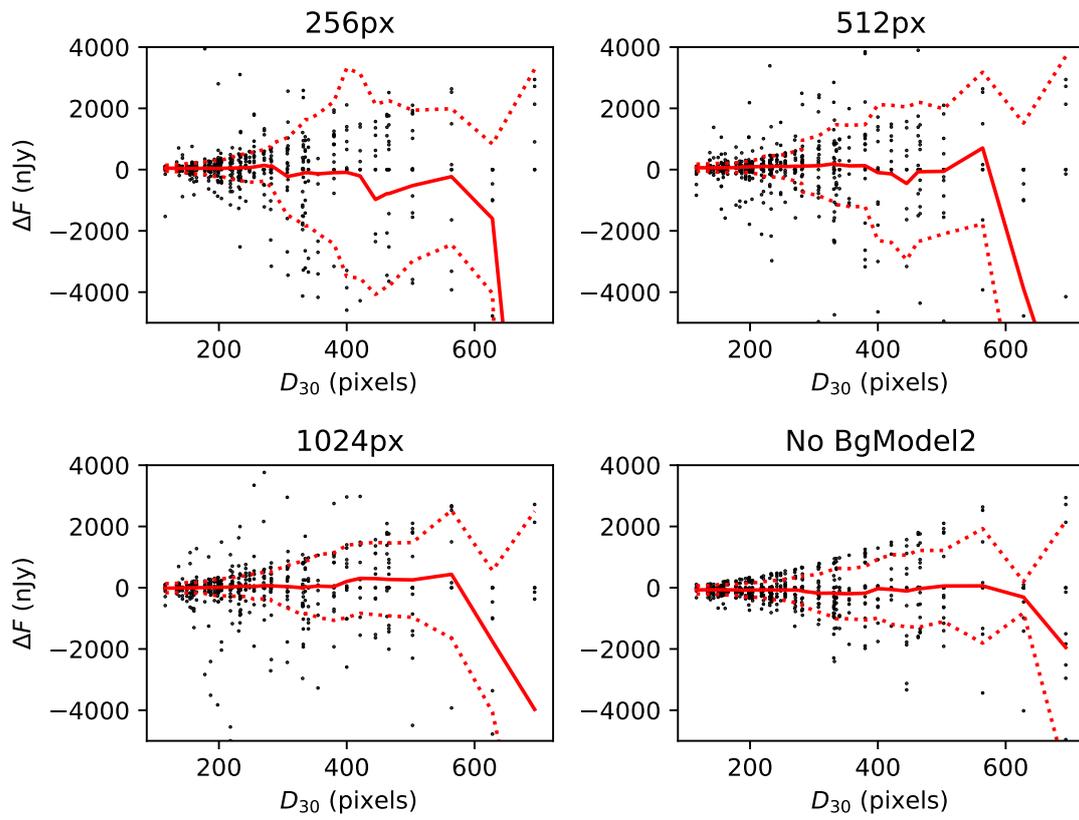


Figure 18: As Fig. 17, but for the large model catalogue injection.