



Generation of TMT Infrared Guide Star Catalog (TMT-IRGSC) - Phase III

[Document Number]

May 31, 2023

Sarang Shah (ITCC), Smitha Subramanian (IIA), Avinash C.K (IIA), G. C. Anupama (IIA), Francisco Delgado (TMT PO), Kim Gillies (TMT PO), Maheswar Gopinathan (IIA), A. N. Ramaprabakar (IUCAA), B. E. Reddy (IIA), Warren Skidmore (TMT PO), T. Sivarani (IIA), Annapurni Subramaniam (IIA)

Contents

1	Executive Summary	3
2	Background and Motivation	4
3	Previous Studies	6
3.1	Summary of the IRGSC Phase I	6
3.2	Summary of the IRGSC Phase II	8
4	Requirements of the IRGSC	10
5	IRGSC Phase III	11
5.1	Overview of Phase III	11
5.2	Selection of the test fields and obtaining the data of those fields	12
5.2.1	PAN-STARRS	16
5.2.2	UKIDSS	17
5.2.3	Gaia	18
5.2.4	Hubble Source Catalog	18
5.3	Steps to compute and validate the NIR magnitudes	19
5.3.1	Separate Stars and Galaxies in the optical data-set	19
5.4	Interstellar Extinction	21
5.4.1	Synthetic photometry in the Pan STARRS and UKIDSS filters	22
5.5	Application of the stellar atmospheric models to the probable stellar sources	23
5.5.1	Using S16 methodology and finding benefits of using five optical bands vs. three	25
5.6	Finding an optimal model	27
5.6.1	Is the optimal method applicable for other test fields?	30
6	Results	32
7	Validation of the results by alternate method	35
8	Discussion	36
8.1	Accuracy of the computed NIR magnitudes and the faintness achieved	36
8.2	Source Density criteria	40
8.3	90% Completeness of the catalog	40
8.4	Applying the optimal methodology on sources detected in g, r, and i bands only	41
8.5	Testing the efficiency of star-galaxy separation	41
8.6	Fields close to the Galactic Plane	42
8.7	Fields with declination $< -30^\circ$	43

8.8	Creation of the IRGSC for additional fields	43
9	Nature of the IRGSC	44
10	Nature of the validated IRGSC	45
11	Summary	47
12	Future Development	49
A	Appendix A: How was the optimal methodology devised?	51
A.0.1	Modelling all the sources by calculating d_{dev} and using various Kurucz and Phoenix model templates	51
A.0.2	Combining optimal Kurucz + Phoenix models	52
B	Appendix B: Comparison of the computed and observed NIR magnitudes when the optimal methodology is applied to other test fields	53
C	Appendix C: An SQL query to obtain the PANSTARRS data from CasJobs	53
D	Appendix D: An SQL query to extract the information in the PANSTARRS flags	54
E	Result table after applying the optimal methodology on the twenty test fields but with 1.06 sq. deg. area	54

List of Figures

1	A schematic diagram illustrating the working of an AO system in a telescope. (Image credits: TMT International Observatory)	5
2	<i>Left</i> figure shows the location of NFIRAOS on the TMT while <i>right</i> figure is the illustrative image of the NFIRAOS (Image taken from Crane et al. 2018).	6
3	The image shows the location of the twenty test fields (as red dots) in the sky in "Aitoff" projection and galactic coordinates.	14
4	Location of the twenty test fields (as red dots) in the sky in "Aitoff" projection and equatorial coordinates.	15
5	The <i>left</i> plots show the error vs magnitude plot in the <i>mean</i> photometric data while the <i>right</i> plot shows the same for the <i>stack</i> photometric data. While the errors in both the photometry increase exponentially in the fainter end, the maximum value of the error in the stack photometry is lesser in comparison to the mean photometry. Comparison of the x-axis also shows that stack photometry is deeper than the mean photometry.	17

6	Top: figure shows the flowchart of the steps required to generate an IRGSC while Bottom: figure shows the flowchart of the steps to validate the IRGSC.	20
7	The left-most plot shows the $(i_{psf} - i_{kron})$ vs. i_{psf} relation of all the sources in the optical PS1 data set for TF1 field. The <i>violet</i> color points in the center plot are the most-probable stellar sources that are obtained after applying Equation (1) to the input PS1 photometric data. In contrast, the <i>black</i> points are the extended sources. The right-most plot shows the Color-Color diagram of the objects where the stars follow a locus while the galaxies form a clump.	22
8	Top-left figure shows the response function of the PAN-STARRS g, r, i, z, and y filters while top-right figure shows the response function of the UKIDSS NIR filters. The bottom plot shows the PS1 and UKIDSS response functions convolved with Kurucz models plotted on top of a sample Kurucz spectra for the model parameters (T_{eff} , $\log(g)$, [Fe/H]) = (5500.0, 0.0, 0.0)	24
9	These plots compare the observed and computed NIR magnitudes of the sources that are modeled using the K0 models and S16 formalism. Only g, r, and i bands are used here, and there is a bunch of sources at the fainter end, which are cooler, metal-poor, and either compact or giant. The sources are color-coded according to the model parameters.	28
10	These plots compare the observed and computed NIR magnitudes of the sources that are modeled using the K0 models and S16 formalism. These sources are modeled using the g, r, i, z, and y bands are used here. The clump of the sources in the second sequence at the fainter end in Figure (9) disappears, and the sources are modeled better. The number of sources in this plot is less than in the previous plot because now the S16 formalism is applied to ten colors instead of three. The sources are color-coded according to the model parameters.	29
11	These plots show the comparison of the observed and computed NIR magnitudes of all the sources in the <i>stack</i> photometric data in g, r, and i bands. When we do not apply the condition in Equation (4), all the sources appear in the plot. We also find that the second sequence is prominent.	30
12	These plots show the comparison of the observed and computed NIR magnitudes of all the sources in the <i>stack</i> photometric data in g, r, i, z, and y bands. Although we find the presence of a second sequence, the density of the sources is far less than when only g, r, and i bands are used.	31
13	For a particular star in the second sequence, these plots show the PS1 optical (blue), UKIDSS observed (magenta), and computed NIR (yellow) points plotted on top of the spectral energy distribution (orange polynomial) of the best-fit model parameters used to compute the NIR magnitudes. The left plot shows that only g, r, and i bands are used to compute the NIR magnitudes, while the right plot shows that after adding z and y bands, the accuracy of estimation of NIR magnitudes increases.	32

- 14 These plots show the comparison between the observed and the computed NIR magnitudes (J, H, and K from top to bottom) when *left* A_v is bounded by 3σ limit and *right* is not bounded by any 3σ limit. The colorbar shows the difference in the observed and computed J using the used method. The plots using the former method show a very tight correlation indicating that the computed NIR magnitudes are very similar while the plots using the latter method show a loose correlation indicating that for some sources, the A_V takes unrealistic value to minimize the χ^2 . The mean value of A_v is 0.13. 37

15 The plots in the *left* side of the figures show the comparison of the difference in the computed NIR magnitudes when A_v is fixed and when A_v is kept free but bounded by 3σ limit. The colorbar shows the A_v values that each source takes in order to minimize the χ^2_r . The plots in the *right* side of the figures show the comparison of the difference in the computed NIR magnitudes when A_v is fixed and when A_v is kept free and unbounded by the 3σ limit. The mean value of A_v is 0.13. 38

16 These plots show the distribution histogram of the model parameters using the optimal color method, flux method when A_v is bound by 3σ limit and A_v is free. As expected, the model parameters are not similar. However, the computed magnitudes are similar to the observed magnitudes. The bin size is 250K for T_{eff} and 0.5 dex for log(g) and [Fe/H] respectively. The distribution of the histogram for the optimal color method and the flux method where A_v is bounded by 3σ limit is nearly similar. 39

17 (*Top-left:*) Histogram of I_{psf} when only i-band data is filtered. (*Top-right:*) Histogram of I_{psf} when all the bands are filtered. (*Bottom-left:*) Histogram of I_{psf} after star-galaxy classification (*Bottom-right:*). Histogram of I_{psf} for the sources in the generated catalog. These plots are for stack photometry. 55

18 Histogram showing the 90% completeness of the catalog for the TF1 field. The plot in the *left* shows the i-band completeness of all the probable stellar sources after the star-galaxy classification, while the plot in the *right* shows the 90% completeness of the computed J magnitudes. 56

19 The plots in this figure show the comparison of the observed and computed NIR magnitudes for the TF11 field (one of the fields having a low density of sources) using g, r, and i bands only (multicolored points) which are plotted on top of sources for which the NIR magnitudes are computed using g, r, i, z and y bands (green circles). Some of the sources are poorly modeled when only g, r, and i bands are used and these sources must be used cautiously. 56

20 These figures show the overlap of the HSC data on the PS1 data. The HSC data do not cover the entire fields in all the cases. 57

21	The figure shows the (i_{psf} vs $i_{psf} - i_{kron}$) plot of a number of stars (magenta points) and galaxies (black points) in the TF13 data. The blue triangles represent the sources classified as stars using Equation (1) applied to g, r, i, z, and y bands in the PS1 data but are galaxies in the HSC data. Similarly, green plus symbols represent the sources classified as stars using Equation (1) applied to g, r, and i bands in the PS1 data but are galaxies in the HSC data.	58
22	The plots show the comparison of the observed and computed NIR magnitudes using the <i>stack</i> photometry data and when no $d_{dev,min}$ is computed for each source. The plots in the <i>first-row</i> are modeled using K0 models, while the plots in the <i>second-row</i> are modeled using the K1 models.	59
23	The plots show the comparison of the observed and computed J magnitudes when all the sources are modeled with K1 models and the scatter is color-coded according to the model parameters and the cumulative deviation of the observed color from the model color (see the second row of Figure (22)). Each plot shows the source color-coded according to the best-fitted model parameter and dev	60
24	The plots show the comparison of the observed and computed H magnitudes when all the sources are modeled with K1 models and the scatter is color-coded according to the model parameters and the cumulative deviation of the observed color from the model color (see the second row of Figure (22)). Each plot shows the source color-coded according to the best-fitted model parameter and dev	61
25	The plots show the comparison of the observed and computed K magnitudes when all the sources are modeled with K1 models and by calculating $d_{dev,min}$ for each source (see the second row of Figure (22)). Each plot shows the source color-coded according to the best-fitted model parameter and dev	62
26	The plots show the comparison of the observed and computed J magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using P0 models. .	63
27	The plots show the comparison of the observed and computed H magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using P0 models.	64
28	The plots show the comparison of the observed and computed K magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using P0 models.	65
29	The plots show the comparison of the observed and computed J magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using C1 and C2 models.	66
30	The plots show the comparison of the observed and computed H magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using C1 and C2 models.	67
31	The plots show the comparison of the observed and computed K magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using C1 and C2 models.	68

32	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry. There is a significant improvement in the accuracy of the computed NIR magnitudes (see the titles of each subplot). The outliers are possibly those objects which are faint and cool.	69
33	The plots show the properties of the sources in the plot showing the comparison between observed and computed J magnitudes of the sources.	70
34	The plots show the properties of the sources in the plot showing the comparison between observed and computed H magnitudes of the sources.	71
35	The plots show the properties of the sources in the plot showing the comparison between observed and computed H magnitudes of the sources.	72
36	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF2 field.	73
37	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF3 field.	74
38	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF4 field.	75
39	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF5 field.	76
40	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF6 field.	77
41	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF7 field.	78
42	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF8 field.	79
43	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF9 field.	80
44	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF10 field. There is no J and H band data from UKIDSS for this test field.	81

45	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF11 field.	82
46	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF12 field.	83
47	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF13 field. There is no J and H band data from UKIDSS for this test field.	84
48	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF14 field. There is no J and H band data from UKIDSS for this test field.	84
49	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF15 field. There is no J and H band data from UKIDSS for this test field.	85
50	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF16 field.	86
51	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF17 field.	87
52	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF18 field.	88
53	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF19 field.	89
54	The plots show the comparison of the computed and the observed NIR magnitudes when the <i>optimal</i> method is applied on <i>stack</i> photometry for the probable stellar sources in the TF20 field. There is no J and H band data from UKIDSS for this test field.	90

List of Tables

1	The equatorial and galactic coordinates (epoch J2000) of the twenty test fields used in this study. The last column shows the median value of the foreground reddening towards the line of sight of these fields (Schlafly & Finkbeiner, 2011).	13
2	Basic information about each sub-survey of UKIDSS.	18
3	Grid size of the various Stellar Atmospheric models (SAMs) used in this study.	25
4	The results obtained after applying the optimal combination of K0, C1, and C2 models to the most probable stellar sources in the optical data of the test fields. The entries marked by the '-' symbol are for the test fields that do not have the observed UKIDSS data in the J and H bands while * indicates the fields that are close to the galactic plane. The % values in the brackets are the percentage number of sources within 0.2 mag of the absolute difference in the observed and the computed NIR magnitudes and the density is computed for the NFIRAOS field-of-view.	34
5	Table showing the efficiency of Star-Galaxy Classification (SGC). The first column indicates the field number; the second column indicates the number of probable stellar sources in the optical PS1 data set after applying the SGC (Equation (1)), the third column indicates the number of galaxies in the PS1 data. The next column indicates the number of sources in the Hubble Source Catalog (HSC) data, followed by the column indicating the number of probable stellar sources in IRGSC that are flagged as galaxies in HSC data and the last column indicates the number of stellar sources in IRGSC that find a match in the HSC data. The HSC data is available only for TF4, TF11, TF13, and TF15 fields and the coverage is very small as compared to the PS1 coverage (see Figure (8.5)). The colored numbers in the TF13 row indicate the values when g, r, and i bands are used.	42
6	Table showing the positions of the <i>Medium Deep Survey</i> fields and the mean value of the reddening towards those fields.	44
7	The table shows the name of the columns in the IRGSC. Here <i>pm</i> is the acronym for proper motion, <i>sam</i> stands for the name of the stellar atmospheric model and <i>ps1</i> stands for PANSTARRS. The information of the PANSTARRS flag values can be found in (Flewelling et al., 2020) and a CasJobs query to extract information hidden in each flag value is given in Appendix (D).	46
8	The table shows the name of the additional columns in the validated IRGSC. Here the observed UKIDSS NIR magnitudes are petro magnitudes.	47

- 9 The results obtained after applying the optimal combination of K0, C1, and C2 models to the most probable stellar sources in the optical data of the test fields for 1.06 sq. deg. area of the sky. The entries marked by the '-' symbol are for the test fields that do not have the observed UKIDSS data in the J and H bands while * indicates the fields that are close to the galactic plane. The % values in the brackets are the percentage number of sources within 0.2 mag of the absolute difference in the observed and the computed NIR magnitudes and the density is computed for the NFIRAOS field-of-view.

Acronyms used in this report

1. TMT = Thirty Meter Telescope
2. IRGSC = Infrared Guide Star catalog
3. NFIRAOS = Narrow-Field Near Infrared facility Adaptive Optics
4. AO = Adaptive Optics
5. AOS = Adaptive Optics system
6. WFS = Wavefront Sensor
7. NIR = Near Infrared
8. PS1 = PANSTARRS/PANSTARRS DR2
9. UKIDSS = The UKIRT INFRARED DEEP SKY SURVEY
10. 2MASS = Two-Micron All Sky Survey
11. VISTA = Visible and Infrared Survey Telescope for Astronomy
12. SAM = Stellar Atmospheric Models
13. K0 = All interpolated Kurucz models
14. K1 = Interpolated Kurucz model with $T_{eff} > 4000\text{K}$
15. C0 = Interpolated Phoenix models in the range $2800 < \text{Teff} < 5000$, $\log(g) > 3.0$, $[\text{Fe}/\text{H}] < -1.5$
16. C1 = Interpolated Phoenix models in the range $[2800 < \text{Teff} < 4000, \log(g) < 3.0, [\text{Fe}/\text{H}] > -0.50]$
17. Gaia EDR3 = Gaia Early Data Release 3
18. CMD = Colour-Magnitude diagram
19. UI = User Interface

1 Executive Summary

In this Phase III, we have carried forward the work performed in Phases I and II of DMS-IRGSC and presented a method to generate the IRGSC using the optical data PANSTARRS PS1 survey. We have also validated the computed magnitudes in the generated IRGSC using the observed NIR UKIDSS data. We have identified twenty-test fields across TMT's observable sky that are located at various latitudes and suffer from different amounts of reddening and extinction. These test fields each have a size of 0.785 sq. deg. (30 arcmin radius) and have PANSTARRS DR2 optical data and UKIDSS observed NIR data readily available. The data is filtered for stellar and extended sources based on the *psf* and *kron* photometry available in the optical PANSTARRS data. A methodology to compute the NIR magnitudes of the optical stellar sources using the combination of Kurucz models having $T_{eff} > 4000\text{K}$ (K1 models) and two sets of Phoenix models having $2800\text{K} < T_{eff} < 4000\text{K}$, $\log(g) < 3.0$ and $[\text{Fe}/\text{H}] > -0.50$ (C1 models) and $2800\text{K} < T_{eff} < 5000\text{K}$, $\log(g) > 3.0$ and $[\text{Fe}/\text{H}] < -1.50$ (C2 models) is developed, applied and tested on these test fields. We use the stacked-photometric optical data in g, r, i, z, and y bands from the 3-pi survey of the PANSTARRS DR2 of these twenty test fields and separate the stars and galaxies in the data set. The combination of the stellar atmospheric models is applied to the stellar sources that are detected in all the five optical bands and the NIR magnitudes for these sources are computed. On comparing with Phase II, we find that additional *z* and *y* optical bands from the PANSTARRS help in better modeling of the cooler sources. The computed NIR magnitudes are validated using the observed NIR magnitudes of these sources, obtained from UKIDSS surveys. The developed method is found satisfactory and can be used to generate an IRGSC for the TMT's AO observations. Together these steps mentioned briefly constitute our *optimal methodology*. We have prepared a Python package to generate the IRGSC using PANSTARRS. The name of this package is *irgsctool* and is available on <https://github.com/tmtsoftware/dms-irgsc>. *irgsctool* not only generates the IRGSC for the given equatorial coordinates but also validates the generated IRGSC using the observed NIR UKIDSS data available for those coordinates. Detailed information on the software API and tutorials can be found on <https://tmtsoftware.github.io>. Using the developed methodology we generated a partial catalog for the twenty test fields and checked whether it satisfies the requirements of the IRGSC. We note that some of the requirements of the IRGSC are not fully met and we discuss them below in the report. We also discuss the limitations of PANSTARRS survey data to use as the input catalog for the generation of the IRGSC. The requirement that the catalog needs to have sources as faint as $J_{Vega} = 22\text{mag}$ is not met for several test fields, the reason can be attributed to the limitation of the input optical photometry from the PANSTARRS. While individually, there are fainter sources in the PANSTARRS data, our condition that the source must be detected in all the five bands restricts the depth of the optical data which reflects in the depth of the computed NIR magnitudes. The same reason can be given to the unfulfillment of the requirement of 3 stars per 2 arcmin FoV of NFIRAOS for some test fields. Since the PANSTARRS DR2 data do not provide the proper motion value of the sources, we have cross-matched the position of the optical sources to the sources in the Gaia DR3 data within 1 arcsec and obtained the astrometric information

for them. This information is not available for the fainter sources in the PANSTARRS data. As the information on whether a source is single or a part of a multi-stellar system is not available from the PANSTARRS data, we have added a flag called *ruwe* from Gaia DR3 which gives more information on whether the source is singular photometrically. This information is also not available for the fainter optical sources. We found that the developed method is not very effective in computing the NIR magnitude of sources from their optical magnitude in high extinction regions such as the Galactic plane. Hence we suggest using the directly available NIR observations of the Galactic plane. The PS1 data cover the sky above -30° declination and hence is not sufficient to generate the IRGSC for the entire TMT observable sky (declination $> -45^\circ$). We suggest that the optical data from the upcoming deep optical surveys from Vera Rubin observatory can be used to generate the IRGSC for the regions of the sky between $-45 < \text{declination} < -30$ degrees. Alternatively, the VISTA Hemisphere Survey (VHS) which is a NIR survey and has the coverage of the sky below -30 degrees declination can also be used. Due to the large density of the sources in the fields close to the galactic plane, the size of the input PANSTARRS data for a 30 arcmin radius is very large. Therefore we have downloaded the data for 5 arcmin size for these fields instead of 30 arcmin size and applied the developed optimal methodology to compute the NIR magnitudes of the sources in them. As the requirement was to have test fields of 1.0 sq. deg. size, we have also applied the developed methodology on the test fields of 1.06 sq. deg. size (35 arcmin radius) and the generated IRGSCs for these test fields, as well as the former test fields, are provided in the [TMT's GitHub repository]https://github.com/tmtsoftware/dms-irgsc/tree/main/generated_irgsc. In addition to the test fields, we generate the IRGSC for additional ten PANSTARRS Medium Deep Survey fields using the developed *irgsctool* python package. The generated IRGSCs for these fields are also provided in the GitHub repository.

2 Background and Motivation

The performance of any ground-based optical/near-infrared (NIR) telescope is affected by the turbulence in the atmosphere of Earth. When the light from a distant astronomical source passes through Earth's turbulent atmosphere, it distorts the wavefront of the light. These distortions make the science images appear fuzzy/blurry. To improve the performance of the ground-based optical/NIR telescopes by compensating for the effects of wavefront distortions, astronomers use a technique known as Adaptive Optics (AO). An AO system tries to correct the distortions using a WaveFront Sensor (WFS), which takes some of the astronomical light, a deformable mirror that lies in the optical path, and a computer that receives input from the detector (Refer Figure: 1). A WFS is a high-speed camera that detects the deformity in the incoming light several thousand times per second. After measuring the deformity, the WFS sends feedback to the system-controlling computer, which then changes the shape of the deformable mirror so that the light eventually becomes distortion-free. This distortion-

free light is then fed to the science instruments in the telescopes to obtain high-quality images with a spatial resolution close to the theoretical diffraction limit. A science target to be studied is often too faint or extended to be used as a reference for measuring the shape of the incident wave-fronts. Instead, a nearby brighter guide star can be used for distortion correction. But, sufficiently bright stars are not available in all parts of the sky, which significantly limits the usefulness of natural guide star AO. This limitation can be overcome by creating an artificial guide star by shining a laser into the mesosphere and creating an asterism. However, much fainter natural reference stars are still required for image position or tip/tilt correction.

The Thirty Meter Telescope (TMT) is one of the largest optical and near-infrared (NIR) ground-based telescopes to be built, and the first light is expected in the next decade. The first light NIR instruments of TMT will be assisted by a multi-conjugate AO instrument, known as Narrow Field Infrared Adaptive Optics System (NFIRAOS) (see Figure 2). NFIRAOS will have a field of view (FOV) of 2-arcmin diameter and will use laser guide stars (LGS) for distortion correction (Crane et al., 2018). However, it will require three or more natural guide stars for tip/tilt correction. A catalog of guide stars will thus be a critical resource for TMT operations. It will enable efficient planning and observations, fulfilling a role similar to that of the Guide Star Catalogs I and II, which were created to allow for the acquisition and control of the Hubble Space Telescope. The TMT Infrared Guide Star Catalog (TMT-IRGSC) should be a star catalog consisting of point sources with NIR (J , H , K_s) magnitudes as faint as 22 mag in the J band in the Vega system (refer to Section 3 and the Design Requirements Document for the Observatory Software - Data Management System Image and Objects Catalogs, TMT.SFT.DRD.12.003.REL01.pdf), covering the entire TMT-observable sky. No catalog currently exists with objects as faint as $J_{Vega} = 22$ mag over the entire TMT observable sky to be used as a guide star catalog. Hence it is essential to develop this catalog by computing the expected NIR magnitudes of stellar sources identified in various optical sky surveys using their optical magnitudes.

3 Previous Studies

3.1 Summary of the IRGSC Phase I

The TMT IRGSC Phase I (Subramanian et al., 2013)¹, was divided into two sub-phases, Phase IA and Phase IB during the period May 2012 - August 2014. Here we discuss the overall summary of Phase I, which identified some of the essential requirements of the TMT IRGSC and initiated its development. The requirements identified were:

¹Also TMT IRGSC Phase IA reports:TMT.SFT.TEC.12.017.DRF06.pdf and TMT.SFT.TEC.13.018.REL04.pdf respectively

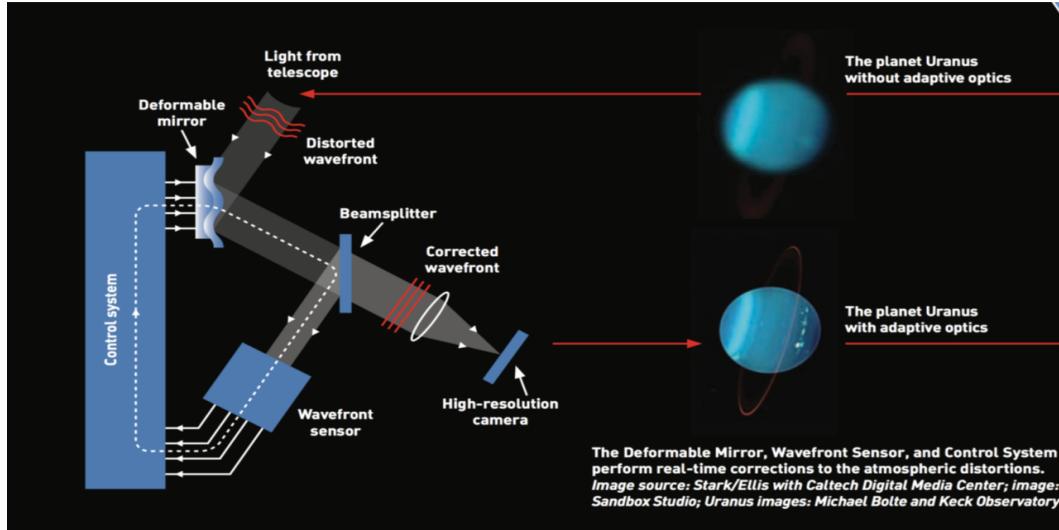


Figure 1: A schematic diagram illustrating the working of an AO system in a telescope. (Image credits: TMT International Observatory)

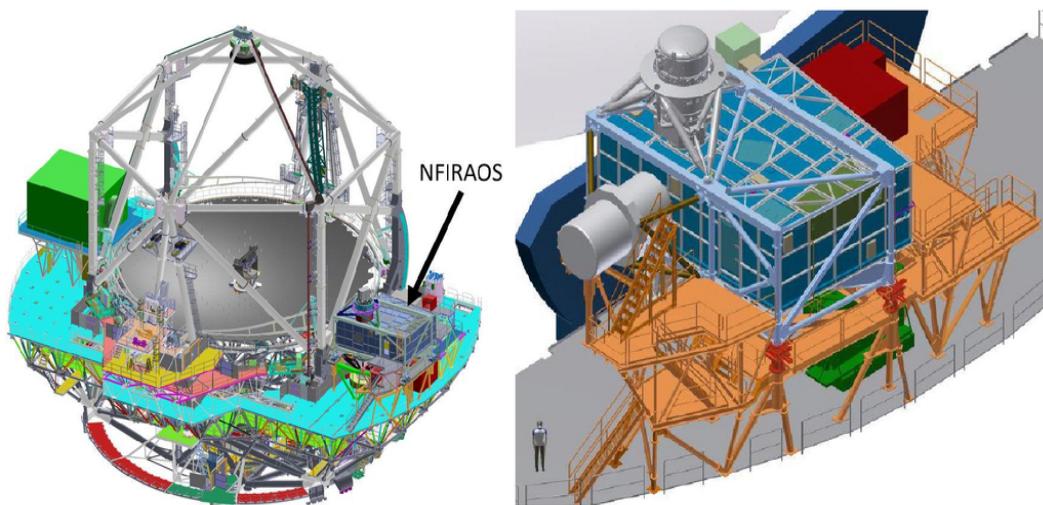


Figure 2: Left figure shows the location of NFIRAOS on the TMT while right figure is the illustrative image of the NFIRAOS (Image taken from Crane et al. 2018).

1. IRGSC shall contain point stellar sources.
2. The position of these sources in the catalog should have an accuracy of the order of milli-arcsec.
3. These sources will have proper motion information, positional information, and brightness in the optical and NIR bands.
4. The required source density is three guide stars per 2 arc-min diameter FOV of the NFIRAOS.
5. The NIR magnitudes should have an accuracy of ≤ 0.2 mag.

In Phase IA, archival optical data (g, r, i bands) from the Canada France Hawaii Telescope (CFHT) Mega-prime camera² were used to compute the NIR magnitudes of the stellar sources. Three test fields (T1, T2, and T3) were selected, with optical and NIR data available. The sources in the optical data were classified as probable stellar sources or galaxies based on their spatial extent in the sky. The temperature of the stellar sources was calculated using the color-temperature relation after correcting the optical magnitudes for extinction. Then assuming a black body model with the computed temperature, the expected J band magnitude for all the probable stellar sources was calculated. All three test fields had NIR observations from UKIRT Infrared Deep Sky Survey (UKIDSS). Hence the computed J band magnitudes were compared with the observed J band magnitudes of probable stellar sources. The comparison was satisfactory for the magnitude range $J_{Vega} = 16\text{--}20$ mag. However, computation of the expected NIR magnitudes of fainter optical sources was difficult due to the difficulty separating the stars and galaxies in the fainter magnitude end. Hence the computed NIR magnitudes of probable stellar sources were mostly brighter than $J_{Vega} = 20$ mag for these fields. It was also found that the source density criteria (3 stars in the two arc-min FOV) were not achieved for these test fields. Another limitation of this method was that the black body color-temperature relation was used to estimate the source's temperature; hence, the expected NIR magnitudes were valid only for a particular range (4200 K to 9600 K) of temperature. Fainter optical sources were included in the analysis by extending the star/galaxy classification, and the expected NIR magnitudes were computed for fainter probable stellar sources. But the comparison of the observed and the computed J band magnitudes of these fainter optical sources was not highly satisfactory. Many sources showed large deviations. Thus it was proposed at the end of Phase IA that a better star/galaxy classification of sources in the fainter magnitude end and usage of Spectral Energy Distribution (SED) template fitting technique to compute the expected NIR magnitudes could be considered in the subsequent phases. In this phase, it was also found that to achieve a depth of $J_{Vega} = 22$ mag, the optical data should go as deep as $i_{AB} \sim 23$ mag. The results of the Phase IA are published in Subramanian et al. (2013) (S13) and a detailed description is available in the TMT IRGSC Phase IA report: TMT.SFT.TEC.12.017.DRF06.pdf. In Phase IB, the developed methodology was applied to compute the H_{Vega} and $K_{S Vega}$ magnitudes of

²This data is available in three optical bands: g, r, i.

probable stellar sources in all three test fields. Though this methodology was satisfactory for the computation of NIR magnitudes, there was a slight difference in the observed and computed magnitudes of the H and K bands. The anomaly was due to the limitation of the black body model since the stars are not perfect black bodies. Therefore, a modified methodology to compute the NIR magnitudes by incorporating the stellar atmospheric models like Kurucz and Next-Gen/Phoenix (Kurucz, 1992a,b, 1993; Castelli & Kurucz, 2003) and Phoenix (Hauschildt et al., 1999a,b) was prepared and applied on the same test fields. This modified methodology was better than the previous methodology and was validated using the UKIDSS data for the test fields till $J_{Vega} = 22$ mag. This method could compute the expected NIR magnitudes for nearly 80 percent of the optical sources. A detailed description of Phase IB is available in the TMT IRGSC Phase IB report: TMT.SFT.TEC.13.018.REL04.pdf.

3.2 Summary of the IRGSC Phase II

Phase II of the IRGSC development started in August 2014. This phase improved the methodology developed in Phase I and put forward the requirements for Phase III. In this section, we summarise the milestones achieved in Phase II.

In Phase II, the methodology developed in Phase IB was modified and fine-tuned to improve the accuracy of the computed NIR magnitudes. In this phase, the Kurucz and Phoenix stellar atmospheric models were interpolated along T_{eff} , log(g), and [Fe/H] axes. A linear interpolation of the Kurucz model was performed by initially reducing the spacing between T_{eff} axis to 125K and, later, 62.5K. Similarly, the spacing was reduced to 0.25 dex along the [Fe/H] and log(g) axes. When these interpolated models were used, the number of stellar sources for which accurate NIR magnitudes could be computed increased significantly. However, when the computed NIR magnitudes of sources in the test fields were compared with the observed NIR magnitudes, it was found that the computed magnitudes of nearly 30% sources were brighter by 0.2 - 0.5 mag. These deviant sources had the model parameters in the range $3500 < T_{eff} < 4000$ K and $\log(g) < 3$. This discrepancy was suggested as the limitation of the stellar atmospheric models in the low-temperature regime. Since removing cooler sources from the data would reduce the source density in the catalog, they tried to balance between the source density criteria and the accuracy of the computed NIR magnitudes by adopting an optimal method to identify an appropriate model for the sources and compute their NIR magnitudes.

In stage I of the optimal method, the interpolated Kurucz models having the parameters $3500 < T_{eff} < 4000$ K and $\log(g) > 3$ and other Kurucz interpolated models with $T_{eff} > 4000$ K were applied to the probable stellar sources. In stage II, the interpolated Next-Gen models with $3000 < T_{eff} < 3500$ K and $3500 < T_{eff} < 4000$ K with $\log(g) < 3$ were applied on the non-retrieved sources in stage I. Finally, in stage III, the two sets of sources were merged, and the method used to generate them was added as a flag in the catalog. The steps to create an IRGSC using the CFHT optical data set that was identified during the IRGSC Phase II are as follows:

1. Identify the stellar sources for which the NIR magnitudes must be computed. The extended half-light radius criteria used in the Phase IB work can be used to separate the stellar sources in the data.
2. De-redden the probable stellar sources before modeling with the stellar atmospheric models.
3. Initially, the probable stellar sources should be modeled with the interpolated Kurucz model ($T_{eff} > 4000\text{K}$ and $3500\text{K} < T_{eff} < 4000\text{K}$ with $\log(g) > 3$) and then the NIR magnitudes must be computed.
4. The sources which are not retrieved from the optical catalog after the previous step are then to be modeled with the interpolated Next-Gen templates ($T_{eff} < 3500 \text{ K}$ and $3500 < T_{eff} < 4000 \text{ K}$ with $\log(g) < 3$). This way, the appropriate model for the remaining sources in the optical catalog can be found, and their NIR magnitudes can be computed.
5. The sources retrieved from the previous two steps should be merged to form the final catalog. The sources retrieved using Kurucz and Next-Gen models must be flagged separately in the catalog.

This fine-tuned methodology was also applied to the publicly available PAN-STARRS reference data. Since this data was not deep enough³, the Phase II study could not check the source density criteria. Nevertheless, the optimal method was applied to get familiarised with the PAN-STARRS data, and the computed NIR magnitudes were validated using the 2MASS NIR data. The study found that the magnitudes were comparable, and the results were promising for the future production of the IRGSC from PAN-STARRS data.

Phase II of the IRGSC also used the HST archival data to analyze the star/galaxy classification accuracy in phase I. It was found that the stellar atmospheric models, along with three bands (g, r, i), could not optimally classify the stars and galaxies in the CFHT data. Prior information on the star-galaxy classification based on the half-light radius or a color-color diagram is required to separate the stars and galaxies optimally. Therefore, this study stated that since the PAN-STARRS data have observations in five bands (g, r, i, z and y), the additional two bands may help to improve the star/galaxy classification during the final production of IRGSC. A future road map was then proposed as follows:

1. To devise and/or fine-tune the existing method for efficient star-galaxy separation and computation of NIR magnitudes for cooler sources from the PAN-STARRS data available in five optical bands - $grizy$.
2. To incorporate the mixture model method developed in this phase and apply it to the PAN-STARRS data set.

³the catalog went up to $i \sim 19$ mag only

3. To validate the results and modify the mixture model method as necessary.
4. To prepare a code for generating IRGSC in Python 3.6 or higher programming language, as the earlier codes were written in FORTRAN/IDL programming languages. This is also one of the requirements of the IRGSC, which is discussed in the next section.

The TMT IRGSC Phase II report results are published in ([Subramanian et al., 2016](#)) (S16), and the details are also available in TMT.SFT.TEC.14.024.REL03.pdf.

4 Requirements of the IRGSC

Based on the analysis and recommendations from the IRGSC Phase I and Phase II, the main requirements of the IRGSC were added in the Design Requirements Document (DRD) for the TMT Observatory Software (OSW)-Data Management System (DMS) - Image and Object Catalogs. The reference document is TMT.SFT.DRD.12.003.REL01.pdf. The main requirements of the IRGSC given in the design requirements document are the following:

1. **Coverage:** The IRGSC should provide coverage for the entire sky that is observable from TMT in Hawaii, in the declination range of -45 to $+90$ degrees.
2. **Point sources:** IRGSC should have point sources as required by the WFS algorithm. Point sources can be stellar and non-stellar (eg, quasars). But as quasars vary in brightness over short periods, it may be difficult to compute their expected NIR magnitudes using appropriate quasar models. In the case of stellar sources, stellar atmospheric models can be used to compute the NIR magnitudes.
3. **Magnitudes:** Based on the NFIRAOS performance simulations and sky coverage studies, the results indicate that the performance requirements can be met using a J band magnitude (in the Vega magnitude system) as faint as 22. The AO simulations that lead to the $J_{Vega} = 22$ limiting magnitude were performed considering median atmospheric conditions for a 30-degree zenith angle. The IRGSC should report the star magnitudes in the J, H, and K_s spectral bands, and the fainter end of the catalog should be $J_{Vega} = 22$ mags, and errors in these magnitudes should be ≤ 0.2 mag.
4. **Source Density:** The source density in the IRGSC should be three stars per 2-arcmin diameter FOV of NFIRAOS, which translates to 3440 stars per square degree.

5. **Source of Magnitudes:** The IRGSC should indicate the source of the NIR magnitudes, whether they are obtained from available surveys or derived using SED.
6. **Positional accuracy:** The IRGSC should contain positional information about the objects, proper motion, and the observations' epoch. It shall limit the objects to those stars with a positional accuracy of 25 mas or better.
7. **Nature of the point source:** It should flag whether the source is a single star or a part of multiple stellar systems. It shall contain the SED and color information.
8. **Input Source Catalog:** It should preserve the information in the input source catalog.
9. **Software Code:** The source code to generate the IRGSC must be written in Python 3.6 or later.

5 IRGSC Phase III

5.1 Overview of Phase III

The third phase of the IRGSC started in August 2019. In this phase, we applied the methodology developed in Phase II on the PAN-STARRS Data Release 2 (DR2) optical data. Different steps involved in this phase are the following:

1. Identify test fields across the TMT's observable sky at different latitudes.
2. Obtain the observed optical data for these test fields from the PANSTARRS to compute the expected NIR magnitudes of stellar sources in these fields.
3. Obtain NIR data for these test fields from the UKIDSS surveys to validate the methodology used to compute the expected NIR magnitudes.
4. Classify stars and galaxies in the PAN-STARRS data.
5. Apply the methodology developed in Phase II for the PAN-STARRS photometric system and validate it.
6. Modify the methodology as necessary to improve the accuracy of the computed magnitudes, maintain the source density, and flag the method applied to the sources to compute the NIR magnitudes.

7. Characterise a new methodology in the low-temperature regime of stars to improve the accuracy of the NIR magnitudes for them.
8. Prepare a software code in Python 3.6 or above.
9. Produce a partial catalog for twenty test fields similar to the final catalog.
10. Flag multiple star systems.
11. Add proper motion values to the stars in the catalog.

The partial catalog generated in this stage of work will be similar to the final IRGSC. It will contain the positional information of the sources, their observed optical, computed NIR magnitudes, SED parameters, and the details of the stellar atmospheric model used to compute the NIR magnitudes. It will also have uncertainties in the computed as well as observed magnitudes. The outline of the report is as follows: We describe the selection of test fields and the data in Section (5.2). Section (5.3) enlists the steps to generate the guide star catalog and describes them in detail. The results table is shown in Section (6), and we discuss the results in Section (8). In Section (7), we validate the results via an alternate method and discuss the pros and cons of the alternate method. In Sections (9 and 10), we discuss the nature of the generated IRGSC and a validated IRGSC and then summarise the work in Section (11). We also propose future development in Section (12). In Appendix (A), we discuss in detail the steps that lead us to the development of the optimal methodology to generate an IRGSC using the PANSTARRS data. In Appendix (B), we show the plots comparing the observed and the computed NIR magnitudes for all the test fields, and in Appendices (C and D), we show an SQL query to obtain the PANSTARRS data from MAST CasJobs and information about the PANSTARRS flags respectively. Lastly, Appendix (E) shows the result table after applying the optimal methodology developed in this phase on the same test fields as used Section (6) but with 1.06 sq. deg. area as per the requirements.

5.2 Selection of the test fields and obtaining the data of those fields

This section discusses the selection of test fields across the sky and the surveys from which we have used the data in this phase. The requirements of the test fields are:

1. They must be located at different galactic latitudes across the TMT's observable sky. This is to understand the effect of varying stellar density/crowding and extinction on the computation of the expected NIR magnitudes using optical data.

2. They must have observed optical (g,r,i,z,y) data for the computation of the expected NIR magnitudes.
3. They must have observed NIR (J, H, K) data to validate the computed NIR magnitudes.

In this study, we selected 20 test fields. The test fields are 30 arc-min in radius or 0.785 deg^2 in area. These fields have five-band optical data from the PANSTARRS and observed J, H, and K NIR data from UKIDSS. In Table (1), we enlist the coordinates and the mean foreground reddening value ([Schlafly & Finkbeiner, 2011](#)) along the line of sight of these fields. In Figure (3), we show the location of the twenty test fields in the sky in galactic frame with the background of the Milky Way while in Figure (4), we show the location of these fields in the sky as red dots in the equatorial frame.

Test Field	α	δ	l	b	$E(B - V)$
TF1	227.26	0.0	359.27	47.24	0.04
TF2	334.27	0.38	63.08	-43.84	0.04
TF3	60.00	1.25	188.72	-36.53	0.26
TF4	30.00	0.50	156.53	-57.82	0.02
TF5	11.16	7.83	120.00	-55.00	0.04
TF6	225.53	2.19	0.0	50.0	0.04
TF7	269.93	-13.48	15.00	5.00	0.98
TF8	334.80	50.96	100.00	-5.00	0.28
TF9	324.09	51.47	95.00	-0.50	2.48
TF10	298.02	34.02	70.00	3.00	1.01
TF11	0.00	0.00	96.33	-60.18	0.02
TF12	34.50	-5.16	169.97	-59.87	0.01
TF13	36.25	-4.50	171.65	-58.22	0.02
TF14	164.25	57.66	148.39	53.43	0.04
TF15	66.75	15.86	180.08	-22.32	0.58
TF16	82.25	-2.60	205.62	-19.48	0.62
TF17	189.83	0.00	296.33	62.71	0.01
TF18	150.25	10.00	227.71	46.40	0.03
TF19	15.00	0.90	127.47	-61.89	0.02
TF20	35.00	-3.50	168.62	-58.28	0.01

Table 1: The equatorial and galactic coordinates (epoch J2000) of the twenty test fields used in this study. The last column shows the median value of the foreground reddening towards the line of sight of these fields ([Schlafly & Finkbeiner, 2011](#)).

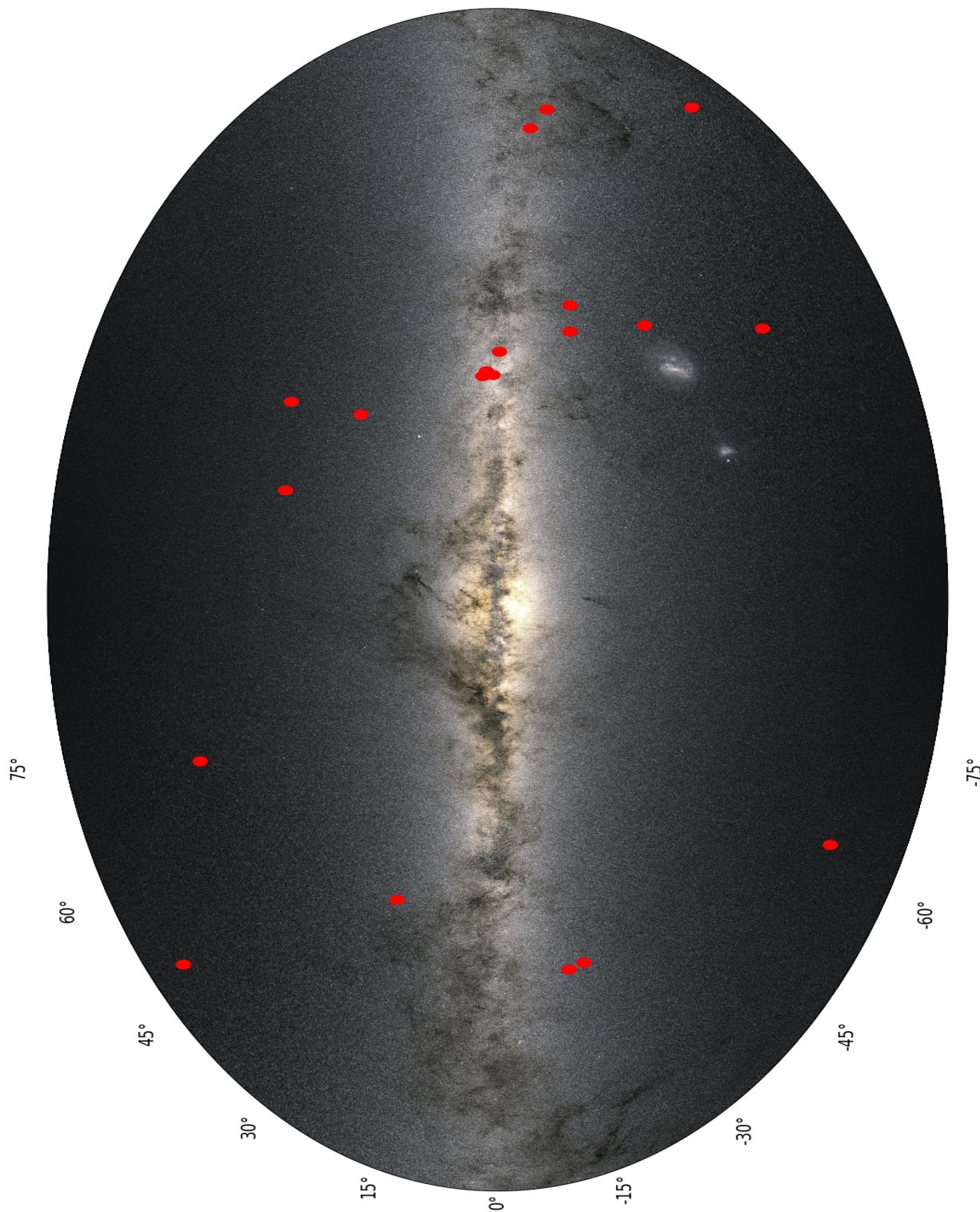


Figure 3: The image shows the location of the twenty test fields (as red dots) in the sky in "Aitoff" projection and galactic coordinates.

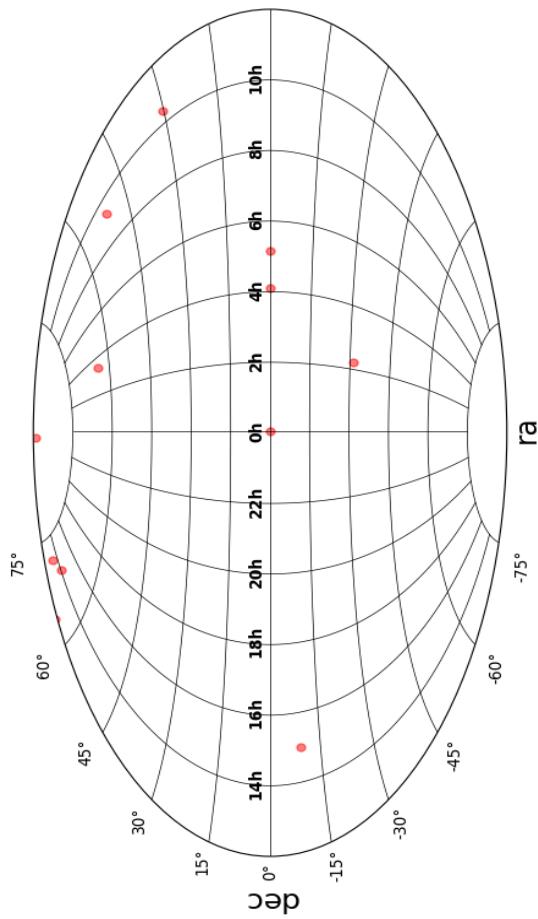


Figure 4: Location of the twenty test fields (as red dots) in the sky in "Aitoff" projection and equatorial coordinates.

5.2.1 PAN-STARRS

The Panoramic Survey Telescope and Rapid Response System (PANSTARRS) (PS1) is an innovative wide-field astronomical imaging and data processing facility developed at the University of Hawaii's Institute for Astronomy (Kaiser et al., 2002, 2010) on the island of Maui where the seeing is 0.83 arcseconds. PS1 is an alt-az telescope whose optical design has a wide field Ritchey-Chretien configuration with a 1.8-m diameter primary mirror and 0.9-m secondary. PAN-STARRS images the sky by using a 1.4 Gigapixel Camera where each pixel has a size of 10 μm that subtends 0.258 arcsecs in the sky.

PS1 survey has several sub-surveys for achieving different science goals. Out of these surveys, we use the data from the "3 π Steradian Survey". PS1 dedicated maximum time to this survey during its operations from 2009 to 2014, covering the sky north of $\delta = -30^\circ$ and conducted the observations in five optical bands - *grizy*. The data for this survey is made public and can be retrieved by various means. The easiest way to retrieve the PS1 data is by using the "web-query" on the PS1 website⁴, which performs a cone search. Here one can input the equatorial coordinates and the radius of the area to be searched and select the columns which contain a variety of observational as well as photometric information⁵. Although this is easy, it gives only a fraction of the complete data set. To obtain the complete data set, we used the "Barbara A. Mikulski Archive for Space Telescopes (MAST) Catalog Archive Server Jobs System (CasJobs)".

The PS1 database provides *mean* and *stack* photometry of the objects detected in its surveys in the AB magnitude system. As the name suggests, the *stack* photometry database is conducted on the objects detected after stacking the images. Therefore, it goes deeper and contains more objects. On the other hand, the *mean* photometry contains the photometry of objects detected in single-epoch observations. The objects in the *stack* photometry also have a higher SNR and lesser uncertainties than the *mean* photometry. In Figure (5), we show the error pattern in the *mean* and the *stack* photometry, which also shows the depth of observations in each filter. The depth and the number of sources in the *stack* photometry are significantly higher, and the magnitudes have lesser uncertainties than the *mean* magnitudes. In addition, the *stack* photometry also provides flags that give more information about the nature of the detected object (refer to Section (9) and Appendix (D)). We, therefore, use *stack* photometric data in our work and develop a methodology to compute the NIR magnitudes of the optical sources in it. We ensure that our data is free from spurious, duplicates, and falsely detected objects. For our work, we have obtained the data from CasJobs by performing a cone search for a 30 arcmin radius centered on the coordinates of the test fields. The query for this search is given in Appendix (C).

⁴<https://catalogs.mast.stsci.edu/PAN-STARRS/>

⁵The PS1 photometry is in the AB system of magnitudes

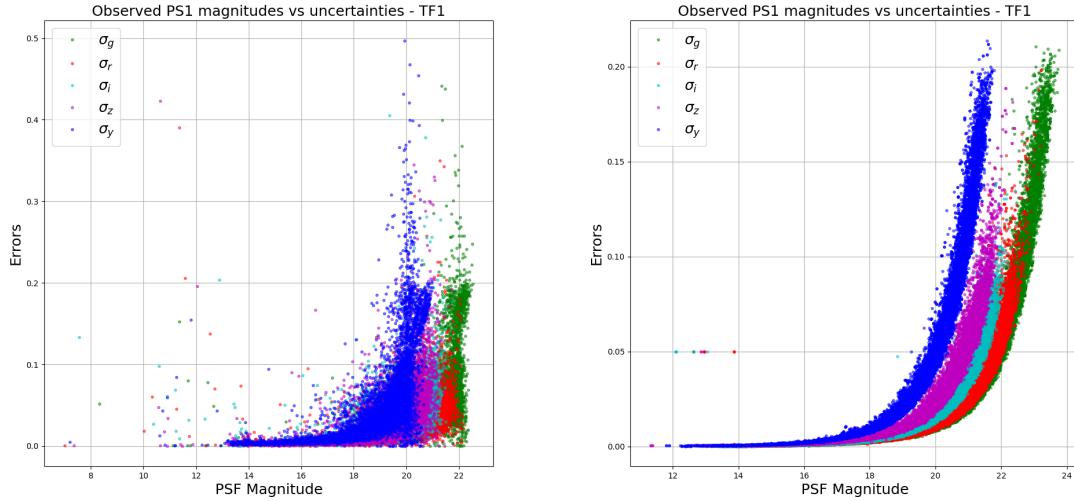


Figure 5: The left plots show the error vs magnitude plot in the mean photometric data while the right plot shows the same for the stack photometric data. While the errors in both the photometry increase exponentially in the fainter end, the maximum value of the error in the stack photometry is lesser in comparison to the mean photometry. Comparison of the x-axis also shows that stack photometry is deeper than the mean photometry.

5.2.2 UKIDSS

The UKIRT Infrared Deep Sky Survey (UKIDSS) was a set of five sub-surveys that had different photometric depths and scanned different parts of the sky depending on the scientific goals that it had to achieve (see Table 2). UKIDSS used the Wide Field Camera (WFCAM) on the 3.8-m United Kingdom Infrared Telescope (UKIRT) that has a field of view of 0.21 deg^2 where each pixel of the camera subtends an angle of 0.4 arcsecs in the sky, and the median seeing is 0.6 arcsec. We use UKIDSS observed NIR data to validate the computed NIR magnitudes. The combination of a large telescope, large field of view, and a camera that provides good image quality make UKIDSS give a larger photometric depth and accuracy than 2MASS, which had a photometric depth of $K = 14.3$ mag (Skrutskie et al., 2006) in the Vega system. Therefore it is a better choice for comparing the computed NIR magnitudes than the 2MASS survey.

The various sub-surveys of UKIDSS conduct observations in either or all of the five filters $ZYJHK$ covering the wavelength range of $0.83\text{-}2.37 \mu\text{m}$. Although UKIDSS, with its sub-surveys, observed a sizeable area and volume of the sky, the data cannot be used for IRGSC because it does not cover the entire observable sky of the TMT. However, as discussed in later sections, the observed NIR UKIDSS, 2MASS, or VISTA data can be readily used for the fields close to the galactic plane as the method to compute the NIR magnitudes from PS1 optical magnitudes fails to provide good results due to

significant amount of extinction.

The UKIDSS data⁶ is given in the Vega system of magnitudes and can be downloaded via a "Region Search" interface⁷ or via a SQL query from the data access page. While the former is a fast and easy way to get the data, the latter helps customize the data columns. We have obtained the UKIDSS data by performing a cone search over 30 arcmin radius regions centered on the coordinates of our test fields. We note that the UKIDSS magnitudes are petro magnitudes and not psf magnitudes.

Survey Name	Passbands	Depth K_s mag	Sky coverage (deg^2)
Large Area Survey (LAS)	$YJHK_s$	18.2	4028
Extra-galactic Deep Survey (DXS)	JK_s	20.8	35
Ultra-Deep Survey (UDS)	JHK_s	22.8	0.77
Galactic-Plane Survey (GPS)	JHK_s	18.8	1868
Galactic-Cluster Survey (GCS)	$ZYJHK_s$	18.6	1067

Table 2: Basic information about each sub-survey of UKIDSS.

5.2.3 Gaia

Gaia is a space mission that has observed billions of stars in the Milky Way with its 1.4-m x 0.5-m space telescope since 2013 and provided their astrometric information. The primary aim of this mission is to generate a 3D model of the Milky Way by accurately measuring parallaxes, positions, and proper motions of all the stars. Gaia DR3 data was released in June 2022 and is based on the data collected between July 2014 and May 2017 (Gaia Collaboration et al., 2022). Amongst a variety of information, DR3 provides proper motions and parallaxes for billions of stars in the galaxy. In this work, we use the Gaia DR3 data to provide the parallax and proper motion information for the sources in IRGSC. The Gaia data can be accessed via ESA's Gaia Data Archive or MAST CasJobs. The Gaia data is also obtained for a region similar to the PANSTARRS and UKIDSS surveys.

5.2.4 Hubble Source Catalog

The Hubble Source Catalog (HSC) is a catalog of sources that was released to help optimize science from the Hubble Space Telescope (HST) by combining several thousands of observations in the Hubble Legacy Archive (HLA) into a single master catalog (Whitmore et al., 2016). It includes WFPC2,

⁶The UKIDSS photometry is in the Vega System of magnitudes.

⁷wsa.roe.ac.uk:8080/wsa/regionform.jsp

ACS/WFC, WFC3/UVIS, and WFC3/IR photometric data generated using SExtractor software to produce the individual source lists. As the space-based observations conducted on HST can achieve a higher spatial resolution than the ground-based telescopes, HSC can distinguish between the point and extended sources or stellar and quasi-stellar sources more effectively than the ground-based telescopes. Therefore, we use HSC to check the efficiency of the star-galaxy classification method that we employ on the PS1 data (see Sections (5.3.1 and 8.5)) for more details). We use a newer released Version (3) of the catalog, which has more sources over a wider area than the previous versions. After several rounds of post-processing of the images, the sources in the catalog have flags 0 for a point source and 1 for an extended source based on the photometry performed. HSC data can also be accessed via MAST CasJobs or by using the "web-query" on the PS1 website⁸. Although we have tried to obtain the HSC data for the 30 arcmin radius region of the test fields, this data is not available for all the test fields. For the test fields that have HSC v3 data, the coverage is very small as compared to the PANSTARRS data.

5.3 Steps to compute and validate the NIR magnitudes

Now we discuss the steps to compute the expected NIR magnitudes of probable stellar sources using their optical magnitudes. We then validate the method by comparing the computed magnitudes of sources with their observed NIR magnitudes. While the initial approach is similar to S16, our analysis changes according to the PS1 data. We show two flow charts in Figure (6), which display the steps we use to compute and validate the NIR magnitudes of sources.

5.3.1 Separate Stars and Galaxies in the optical data-set

We download the PS1 photometry containing the "*psf*" and the "*kron*" photometry of the objects in all five optical bands. The *psf* magnitude is the measure of the total flux of a source fitted by a *Point Spread Function* (*psf*) model which is a Gaussian. Thus this model can estimate the brightness of the point sources accurately. Alternatively, the *kron* magnitudes are the measurement of the total flux of an extended object inside the *Kron* radius⁹ (R_k). R_k is defined as the first moment of the surface brightness light profile (Kron, 1980). An R_k of 2 or 2.5 contains a sufficient amount of the total flux of the galaxy to be a useful measure of total flux. This method can be used to determine the radius of faint extended sources. A *kron* radius for a point source is very similar to the width of the Gaussian

⁸<https://catalogs.mast.stsci.edu/hsc/>

⁹<https://outerspace.stsci.edu/display/PANSTARRS/PS1+Kron+photometry+of+extended+sources>

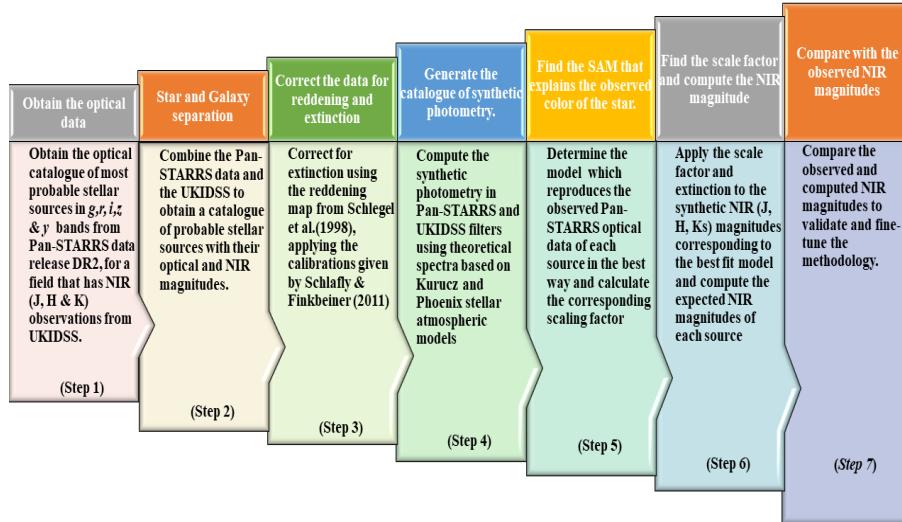
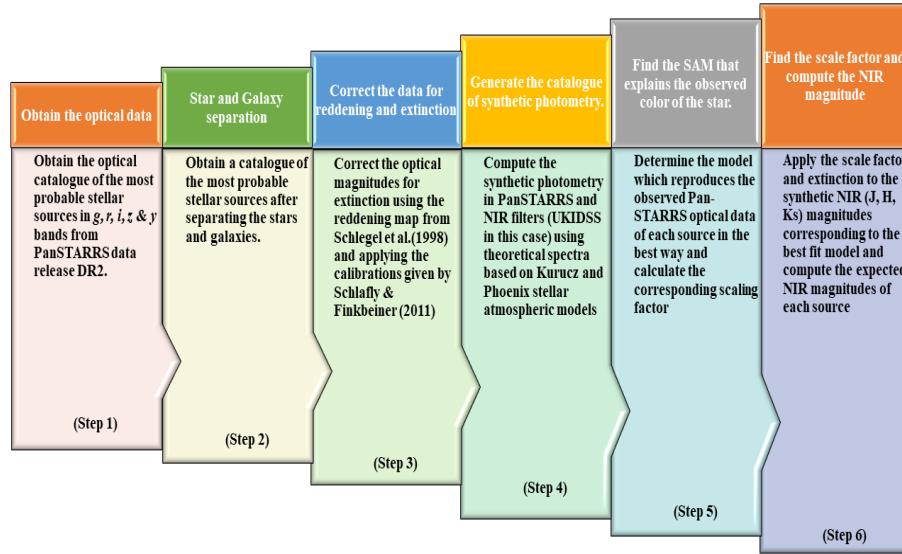


Figure 6: Top: figure shows the flowchart of the steps required to generate an IRGSC while **Bottom:** figure shows the flowchart of the steps to validate the IRGSC.

of the PSF model fitted. Hence, the Kron as well as the PSF magnitudes can be used to distinguish between the stars and the galaxies in the data. As we require only point stellar sources in the catalog, we separate the galaxies from the data set by using the relation:

$$a_{psf} - a_{kron} < 0.05 \quad (1)$$

where "a" ∈ [g,r, i,z, and y] (Chambers et al., 2016)¹⁰. In the *left* plot of Figure (7), we show all the sources in the $(i_{psf} - i_{kron})$ vs i_{psf} diagram in the PS1 optical data for TF1 field. In the consecutive plot, we show how stars and galaxies are separated by using Equation (1) where the *violet* colored points are the most probable point stellar sources and *black* colored points are the galaxies or the extended sources. We also show a CCD of these sources in the *right* most plot where stars can be seen following a particular locus while the galaxies are clumped together. We discuss the efficiency of our method of separating the stars from galaxies in detail in Section (8.5). In principle, many faint sources which may be galaxies will not be separated by this relation. The same is true for brighter sources which might be quasars and not non-extended. Although PANSTARRS provides *objinfoflag*, which can be used to infer whether the source is a quasar (see Appendix (D) for discussion about a CasJobs SQL query to extract the information in this flag), this flag is not available for all the sources at this stage. So there is no other way to remove these sources at this stage. An alternative way is to check the goodness of fit metric of the model fitted to the observed data. While it is expected that non-stellar sources would be fitted poorly by the stellar atmospheric models, we show later that this is not necessarily the case for point sources observed from the ground. So at this stage it is challenging to identify if there is any contamination due to point sources that are non-stellar.

5.4 Interstellar Extinction

For further steps, the observed optical magnitudes of the probable stellar sources have to be corrected for the effects of interstellar extinction. To correct foreground extinction, we use the reddening map provided by (Schlafly & Finkbeiner, 2011), which is the updated map based on Schlegel et al. (1998). The reddening values ($E(B - V)$) are obtained using *dustmaps* python package (Green, 2018). The extinction coefficients in optical bands are calculated by using the relations given by Tonry et al. (2012), which is valid for $-1 < (g-i) < 4$ (see Equations 2). For the validation of the computed NIR magnitudes with the observed NIR magnitudes, we convert them to the apparent magnitudes by adding the extinction coefficients in the respective NIR filter. The NIR extinction coefficients are obtained by multiplying the reddening obtained from the *dustmaps* to the standard reddening-to-

¹⁰also see the PANSTARRS data archive home page <https://outerspace.stsci.edu/display/PANSTARRS/How+to+separate+stars+and+galaxies>

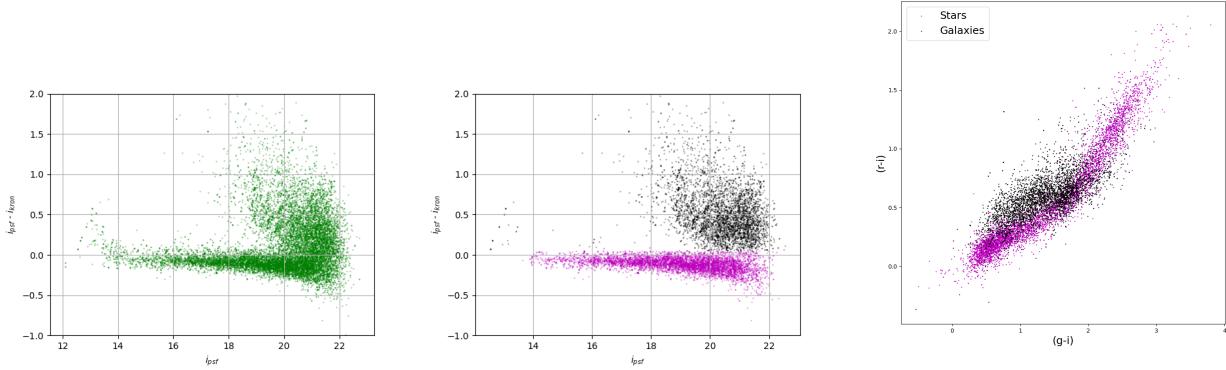


Figure 7: The left-most plot shows the $(i_{psf} - i_{kron})$ vs. i_{psf} relation of all the sources in the optical PS1 data set for TF1 field. The violet color points in the center plot are the most-probable stellar sources that are obtained after applying Equation (1) to the input PS1 photometric data. In contrast, the black points are the extended sources. The right-most plot shows the Color-Color diagram of the objects where the stars follow a locus while the galaxies form a clump.

extinction $A/E(B - V)$ ratio of 0.709, 0.449, and 0.302 in J, H, and K bands respectively¹¹.

$$\begin{aligned}
 a_g/E(B - V) &= 3.613 - 0.0972(g_{psf} - i_{psf}) + 0.01(g_{psf} - i_{psf})^2 \\
 a_r/E(B - V) &= 2.585 - 0.0315(g_{psf} - i_{psf}) \\
 a_i/E(B - V) &= 1.908 - 0.0152(g_{psf} - i_{psf}) \\
 a_z/E(B - V) &= 1.499 - 0.0023(g_{psf} - i_{psf}) \\
 a_y/E(B - V) &= 1.251 - 0.0027(g_{psf} - i_{psf})
 \end{aligned} \tag{2}$$

5.4.1 Synthetic photometry in the Pan STARRS and UKIDSS filters

The light we get from the stars is the radiation that leaves their photo-sphere. This light contains information about the nature and chemical composition of the star. During photometric observations, the light from the source is registered on a telescope detector through a filter that allows only a particular wavelength range to pass through it. The parameters determining the nature of a star can be obtained by comparing the observed fluxes in the optical bands with the synthetic photometric fluxes for different types of stars predicted by the *Stellar Atmospheric Models*. The synthetic photometry of the stars in various filters can be computed from the synthetic spectra taken from different spectral libraries. To compute the synthetic magnitudes in the PANSTARRS and UKIDSS filters, we use syn-

¹¹These values are provided in "NASA IPAC Infrared Science Archive", assuming an extinction law based on Fitzpatrick (1999) and Indebetouw et al. (2005)

thetic high-resolution spectra based on the Kurucz (Kurucz, 1992a,b, 1993), Castelli-Kurucz (Castelli & Kurucz, 2003) and Phoenix (Hauschildt et al., 1999a,b) stellar atmospheric models (SAMs). Stellar atmospheric models provide expected high-resolution spectra from stars with different physical parameter combinations - (T_{eff} , log(g) and [Fe/H]). As we have only observed photometric information, we need to find the model magnitudes in our pass bands (g, r, i, z, y, J, H, and K, in this case). To find the model magnitudes, we convolve the synthetic spectra with the telescope response function, known as "*Effective Stimulus*" (ES). If the model flux emitted by a star is F_λ and the telescope's response function is P_λ , then the ES is given by:

$$ES = \frac{\int F_\lambda P_\lambda \lambda d\lambda}{\int P_\lambda \lambda d\lambda} \quad (3)$$

where, λ is the wavelength. Using Equation (3), we generate the synthetic photometry catalog in PS1 and UKIDSS filters where each set of magnitudes is associated with a [T_{eff} , log(g) and [Fe/H]]. We generate the table of synthetic photometry using a tool known as *pysynphot*¹², which is an object-oriented synthetic photometry package from IRAF in python (STScI Development Team, 2013). In Figure 8, we show the PS1 and UKIDSS response functions for their respective filters plotted on top of a Kurucz spectrum for [T_{eff} , log(g) and [Fe/H]] = [5500.00, 0.0, 0.0] against the wavelength range¹³. Table (3) shows the range of Kurucz, Castelli-Kurucz, and Phoenix model parameters. Although the T_{eff} goes up to 50000K for all the models, we have restricted the T_{eff} to 10000 K to increase the speed of the code to compute the NIR magnitudes. As performed in Phase II of IRGSC development, we refine the grid axes of these model parameters by interpolating linearly between the parameters. These are called "Interpolated Kurucz," "Interpolated Castelli-Kurucz," and "Interpolated Phoenix" models and we use these models in our work. Here, we reduce the spacing between the temperature points to 62.5K and between the log(g) and [Fe/H] points to 0.25 dex and 0.1 dex respectively. While the Castelli-Kurucz models are the newer version of the Kurucz models, the metallicity range covered by them is less than the Kurucz models. We, therefore, use Kurucz model templates in places where the Castelli-Kurucz model templates are not available.

5.5 Application of the stellar atmospheric models to the probable stellar sources

To estimate the parameters that give rise to the flux of a star in the data set, we fit the Spectral Energy Distribution (SED) to the observed star. We do so by fitting the model colors to the de-reddened observed colors and finding the scale factor to compute the NIR magnitudes. We verify the computed NIR magnitudes by comparing them with the observed NIR magnitudes from UKIDSS.

¹²pysynphot.readthedocs.io

¹³The amplitude of these curves represent the percentage of the incident photons that are transmitted by each filter to the detector while the width represents the wavelength range defined for each filter.

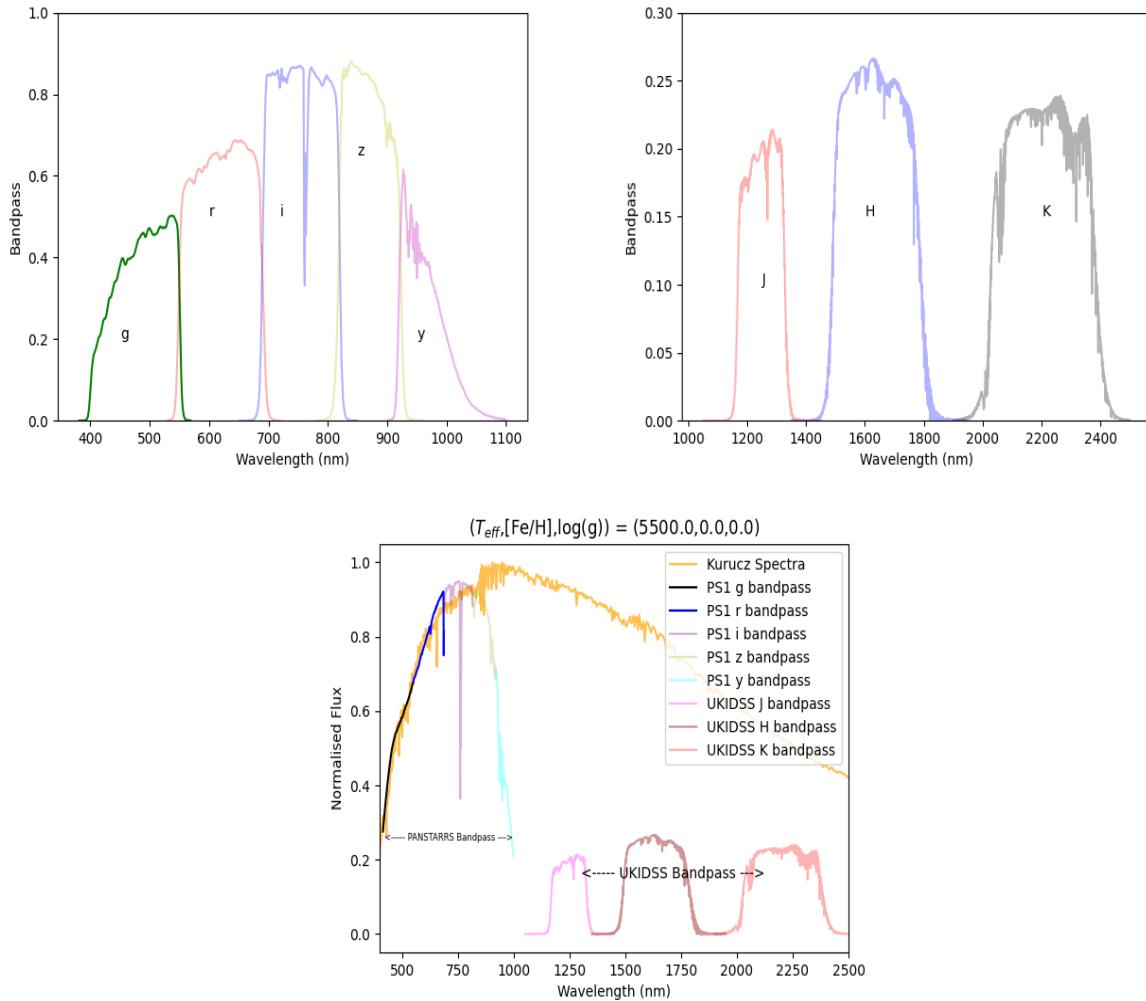


Figure 8: **Top-left** figure shows the response function of the PAN-STARRS *g*, *r*, *i*, *z*, and *y* filters while **top-right** figure shows the response function of the UKIDSS NIR filters. The **bottom** plot shows the PS1 and UKIDSS response functions convolved with Kurucz models plotted on top of a sample Kurucz spectra for the model parameters (T_{eff} , $\log(g)$, $[Fe/H]$) = (5500.0, 0.0, 0.0)

Stellar Atmospheric Model	T_{eff} (K)	$\log(g)$ (dex)	[Fe/H] (dex)
Kurucz	3500 to 10000 in steps of 250K	0.0 to 5.0 in steps of 0.5 dex	+1.0, +0.5, +0.3, +0.2, +0.1, 0.0, -0.1, -0.2, -0.3, -0.5, -1.0, -1.5, -2.0, -2.5, -3.0, -3.5, -4.0, -4.5, -5.0
Interpolated Kurucz	3500 to 10000 in steps of 62.5K	0.0 to 5.0 in steps of 0.25 dex	1.0 to -5.0 in steps of 0.1 dex
Castelli-Kurucz	3000 to 10000 in steps of 250K	0.0 to 5.0 in steps of 0.5 dex	0.0, -0.5, -1.0, -1.5, -2.0, -2.5, +0.5, +0.2
Interpolated Castelli-Kurucz	3000 to 10000 in steps of 62.5K	0.0 to 5.0 in steps of 0.25 dex	0.5 to -2.0 in steps of 0.1 dex
Phoenix	2000 to 7000 in steps of 100K, 7000 to 10000 in steps of 200K	0.0 to 5.5 in steps of 0.5 dex	0.0, -0.5, -1.0, -1.5, -2.0, -3.0, -3.5, -4.0, +0.3, +0.5
Interpolated Phoenix	2000 to 10000 in steps of 62.5K	0.0 to 5.0 in steps of 0.25 dex	0.5 to -4.0 in steps of 0.1 dex

Table 3: Grid size of the various Stellar Atmospheric models (SAMs) used in this study.

In the following sections, we discuss how we generate an optimal method to generate an IRGSC in detail. We have used the data for the TF1 test field for this purpose, and later the best method is extended to all the other test fields.

5.5.1 Using S16 methodology and finding benefits of using five optical bands vs. three

Now that we have our observed data and synthetic photometry ready, we find empirical ways to fit the models to the observed stars. Firstly, we implement a formalism similar to S16 ([Subramanian et al., 2016](#)) on the PS1 optical data, which was decided after checking for errors in the CFHT archival data. Since that method proved to be satisfactory in generating the NIR magnitudes, here we apply it to the PS1 data and fit the interpolated Kurucz models (we call them K0 models) to the observed color of the star. To do so, we first find the difference in the observed and the model colors using the condition:

$$|(\Gamma_{observed,p} - \Gamma_{model,p})| \leq 2 \times \sigma_{observed,p} \quad (4)$$

where Γ is the color for the filter $p \in n$, "n" being all the possible color combinations using the five filters, *grizy*, of PanSTARRS. Once a star satisfies this condition, we calculate the amount of deviation of each observed color from the model color and then calculate a quantity called d_{dev} :

$$d_{quad} = \sqrt{\sum_{p=1}^{p=n} (\Gamma_{obs,p} - \Gamma_{model,p})^2} \quad (5)$$

The model with the minimum value of d_{quad} is the best-fitted model. To compute the expected NIR magnitudes from this best-fit model, we calculate the scaling factor (s.f.), which is defined as the difference between the observed and de-reddened optical magnitudes and the model optical magnitudes;

$$s.f.b = m_b - M_b \quad (6)$$

where "b" denotes different optical bands. It is already shown in S13 and S16s that the *s.f.* is similar for all the bands. Nevertheless, we consider the *average value* of the s.f. in further analysis. We then add the model synthetic NIR magnitudes to the scale factor and account for extinction to get the apparent NIR magnitudes. The positions of the PS1 sources for which the NIR magnitudes are computed are matched to the position of the UKIDSS sources within one arc-second to compare the computed and observed magnitudes.

Discussed in Section (3.2), the comparison plot of "observed NIR magnitudes" vs. "difference in the observed and computed NIR magnitudes" was found to have a dual sequence in the previous study. The sources in the second sequence were found to have computed magnitudes brighter than the observed magnitudes. Based on their best-fit model parameters, they were suggested as cool giants by (Subramanian et al., 2016). To model these sources in a better way, S16 implemented a two-stage modeling strategy. The plots in Figure (9) show the comparison of the observed vs. the difference in the observed and computed magnitudes when S16 formalism is applied to the sources and color-coded according to the model parameters. The sources here are modeled using K0 models, and only g, r, and i bands are considered, and the median and spread of the scatter at 1σ level are 0.006 ± 0.477 . In this plot, we can see the sources in the second sequence are modeled as cool (having $T_{eff} < 4000K$), metal-poor, and either compact or giant. Figure (10) shows that when the sources are modeled using g, r, i, z, and y bands and S16 formalism applied to K0 models, they are modeled better, and the second sequence merges with the primary sequence. The mean and spread of the scatter at 1σ level become 0.017 ± 0.202 . The spread of the scatter is reduced now. Although the median value is slightly higher, it is well within the 1σ level of the former. Thus, adding additional "z" and "y" bands helps us better model the cooler sources when S16 formalism is used. Another thing to note is that there are only 134 sources in the plot when five bands are used vs. 1127 when three bands are used because in the former case, the Equation (4) is applied to ten colors arising out of five optical bands while in case of the later, it is applied to only three colors that can be formed using three optical bands.

Out of the 6939 probable stellar sources for TF1, only 134 satisfy the S16 formalism and find a positional match with the sources in the UKIDSS data. This number of sources very is less because the errors in the *stack* photometry are very small and fail to satisfy the condition in Equation (4) (See Figure 5). Therefore, to increase the number of sources in the plot, we remove the condition given in Equation (4) and model all the sources with K0 models. Analogous to Figures (9 and 10), in Figures (12 and 11), we show the comparison of the observed and computed NIR magnitudes color-coded according to the model parameters when all the sources detected in g, r, and i vs. g, r, i, z, and y bands respectively and the quality cut of S16 formalism is not applied. When only g, r, and i bands are considered, the clump of stars in the second sequence in Figure (9) becomes a prominent second

sequence. These sources are also cooler, metal-poor, and either compact or giant. When we include z and y bands in the modeling, though the displacement of the second sequence from the primary sequence decreases, it is still present. These stars are also cooler, metal-poor, and either compact or giant but many metal-poor and cooler sources from Figure (11) find a better model. This indicates that cooler Kurucz models (having $T_{eff} < 4000K$) are inefficient. In addition to z and y bands, the sources need to be modeled with Phoenix models as they are better in low-temperature regimes.

In Figure (13)), we plot the observed photometric points of a star in the second sequence on the SED using the modeled parameters. The left plot shows that when the star is modeled with only g, r, and i bands, the computed NIR magnitude points (magenta-colored points) do not match the observed UKIDSS NIR points. The right plot shows that when the same star is modeled using g, r, i, z, and y bands, the magenta points better match the observed UKIDSS NIR magnitudes.

5.6 Finding an optimal model

In the previous section, we saw that the condition used by S16 is ineffective in computing the NIR magnitudes of the sources in the PS1 *stack* data because they have very small errors. So we removed this condition and found the best-fit model with the least deviation from the observed colors. Although the source density increased drastically after this, we found that the K0 models are inefficient in modeling the cooler sources (see Figure (12)). Thus, to improve the accuracy of the computed NIR magnitudes of the cooler sources in the second sequence, we now consider using the Phoenix models in addition to the Kurucz models and modifying the S16 methodology. In particular, instead of filtering sources through the Equation (4), we model all the sources to find out the minimum of the sum of the square of the deviation of each observed color with respect to the model color i.e. $d_{dev,min}$ (see Equation 7). After several trials with different combinations of models, we suggest an optimal methodology for the generation of IRGSC, which is to model all the sources by *combining* Kurucz model templates having $T_{eff} > 4000K$ (K1 models), Phoenix model templates in the range [$2800 < T_{eff} < 4000K$, $[Fe/H] > -0.50$, and $\log(g) < 3.0$] (C1 models) and [$2800 < T_{eff} < 5000K$, $[Fe/H] < -1.50$ and $\log(g) > 3.0$] (C2 models) and computing d_{dev} for them. A detailed explanation of how we have devised this methodology can be found in Appendix (A).

$$d_{dev} = \sum_{p=1}^{p=n} |(\Gamma_{obs,p} - \Gamma_{model,p})^2| \quad (7)$$

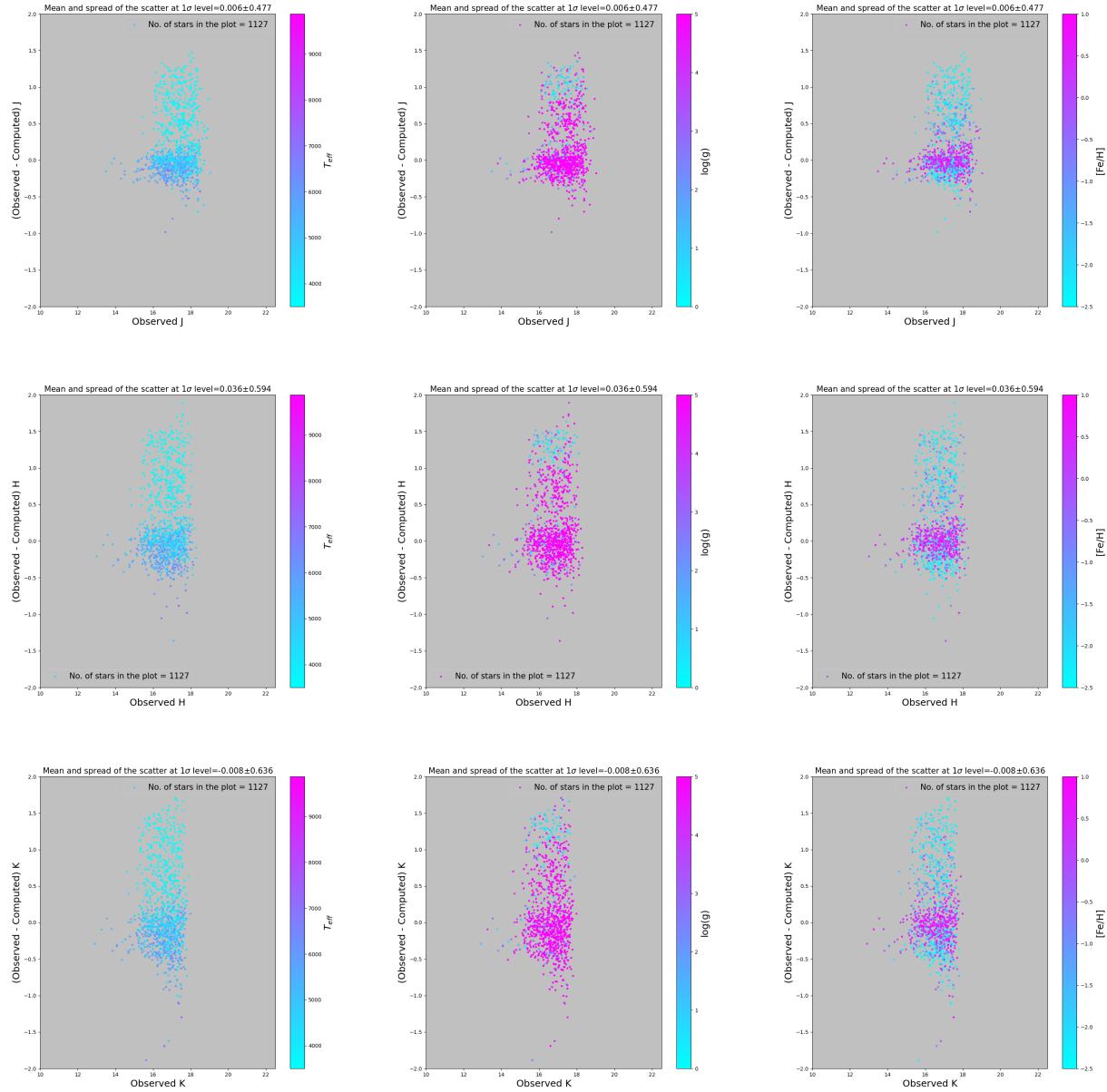


Figure 9: These plots compare the observed and computed NIR magnitudes of the sources that are modeled using the K0 models and S16 formalism. Only g, r, and i bands are used here, and there is a bunch of sources at the fainter end, which are cooler, metal-poor, and either compact or giant. The sources are color-coded according to the model parameters.

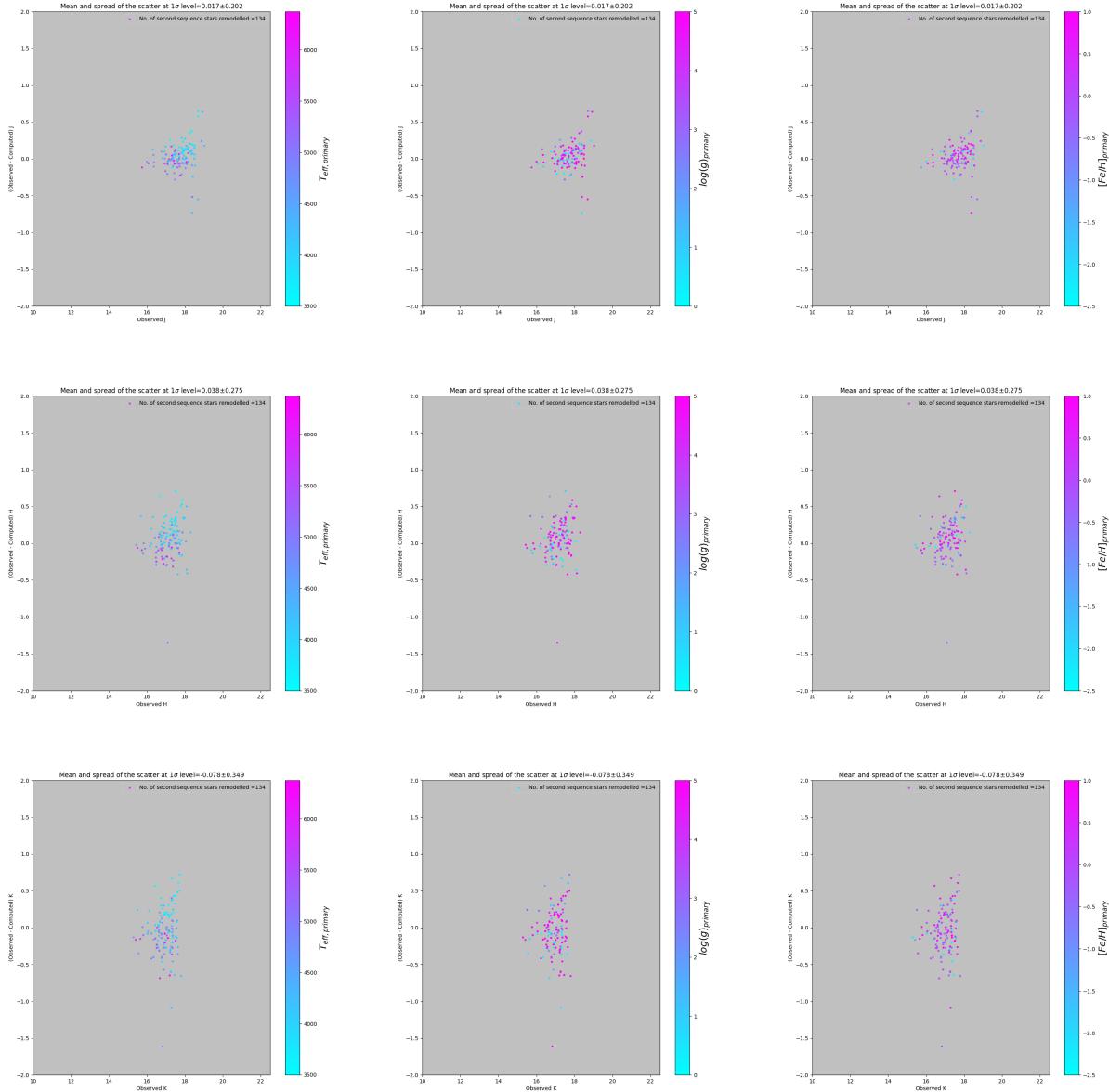


Figure 10: These plots compare the observed and computed NIR magnitudes of the sources that are modeled using the K0 models and S16 formalism. These sources are modeled using the g , r , i , z , and y bands are used here. The clump of the sources in the second sequence at the fainter end in Figure (9) disappears, and the sources are modeled better. The number of sources in this plot is less than in the previous plot because now the S16 formalism is applied to ten colors instead of three. The sources are color-coded according to the model parameters.

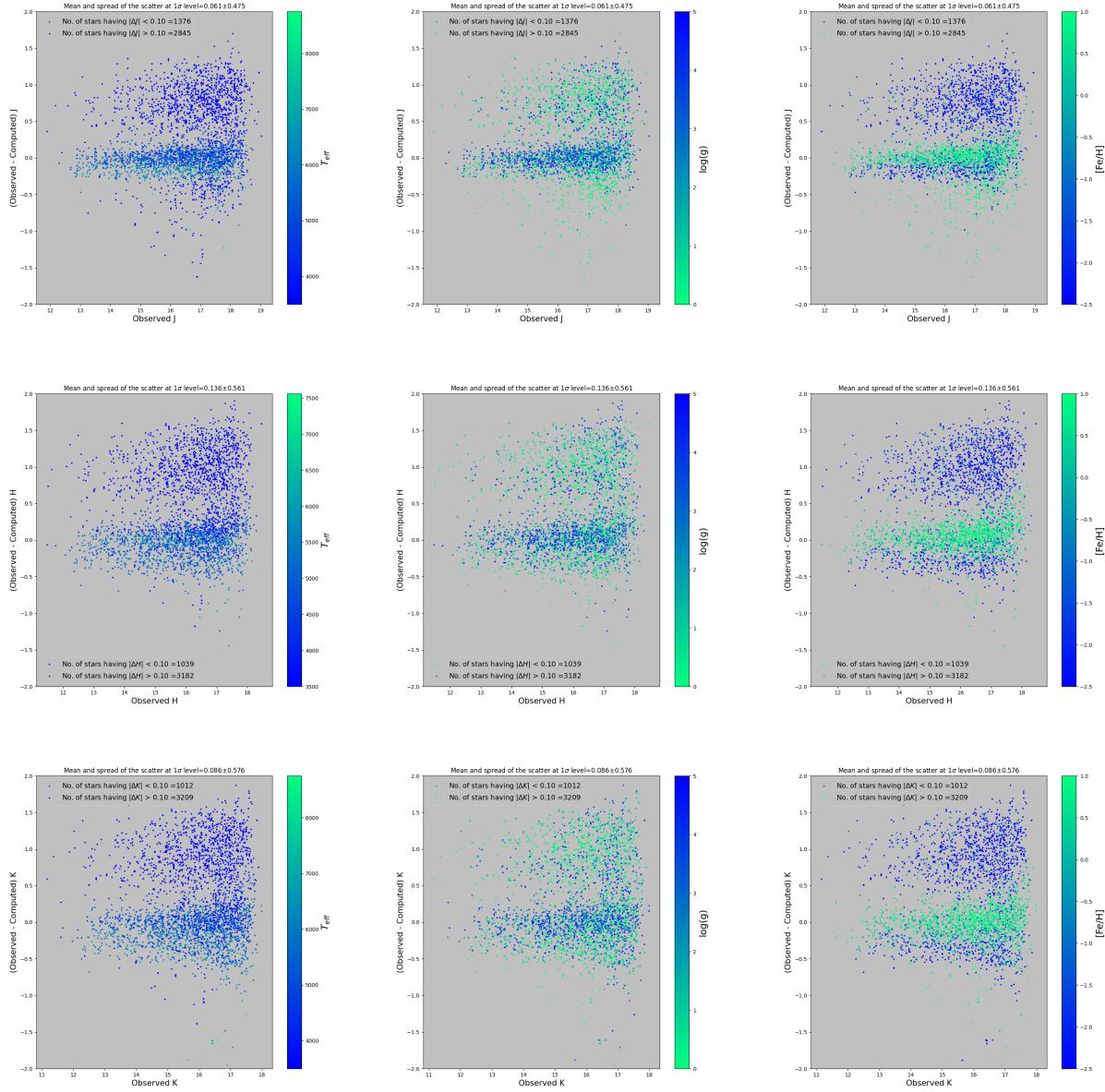


Figure 11: These plots show the comparison of the observed and computed NIR magnitudes of all the sources in the stack photometric data in g , r , and i bands. When we do not apply the condition in Equation (4), all the sources appear in the plot. We also find that the second sequence is prominent.

5.6.1 Is the optimal method applicable for other test fields?

While deciding upon the method to generate an IRGSC, we used the sources from only the TF1 test field. We applied various combinations of the Kurucz and Phoenix models to achieve the required

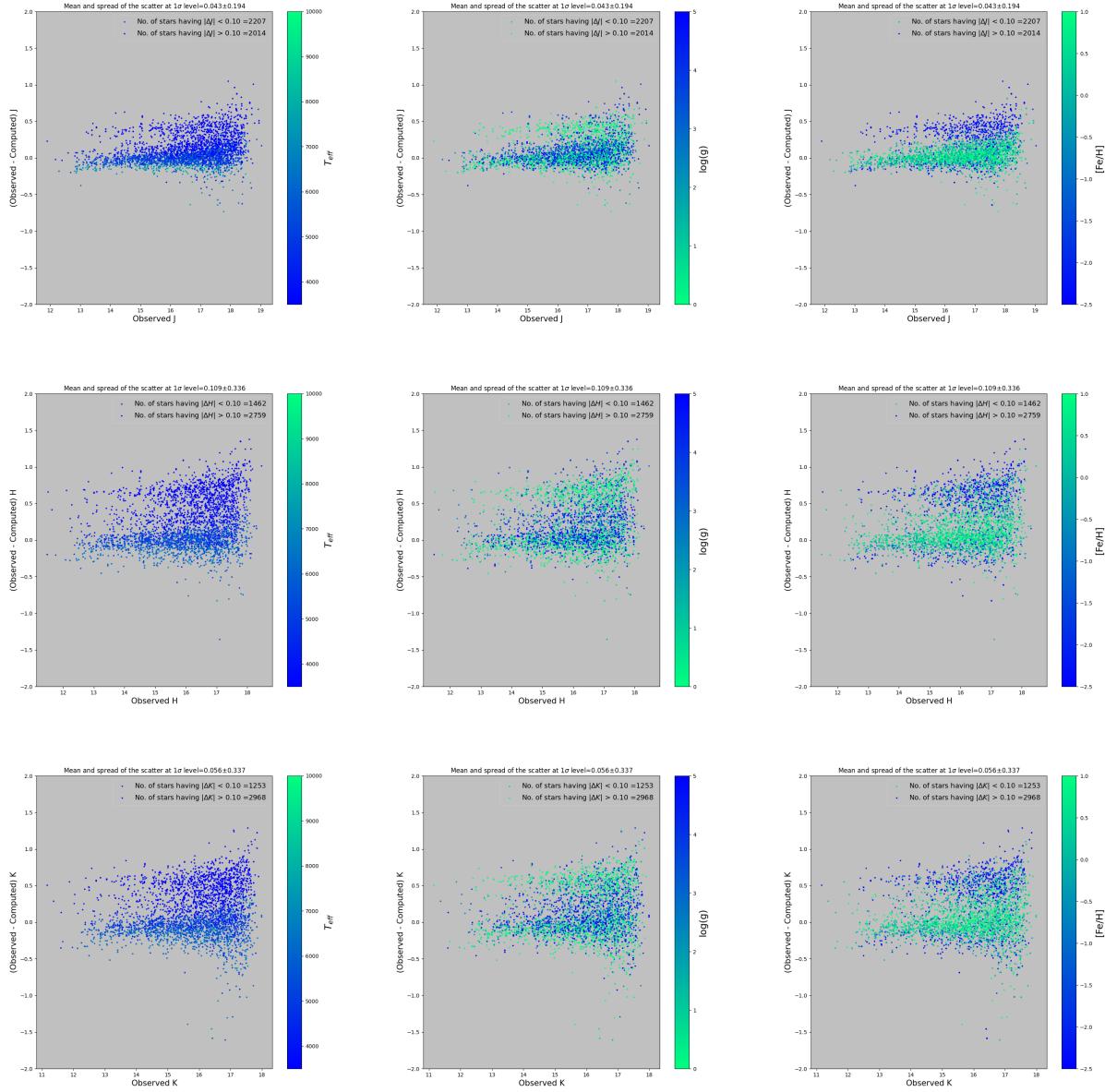


Figure 12: These plots show the comparison of the observed and computed NIR magnitudes of all the sources in the stack photometric data in g , r , i , z , and y bands. Although we find the presence of a second sequence, the density of the sources is far less than when only g , r , and i bands are used.

sources density, accuracy, and depth of the computed NIR magnitudes (discussed in Section (4)). However, does our final method produce satisfactory results for other test fields? In the Figures (36,

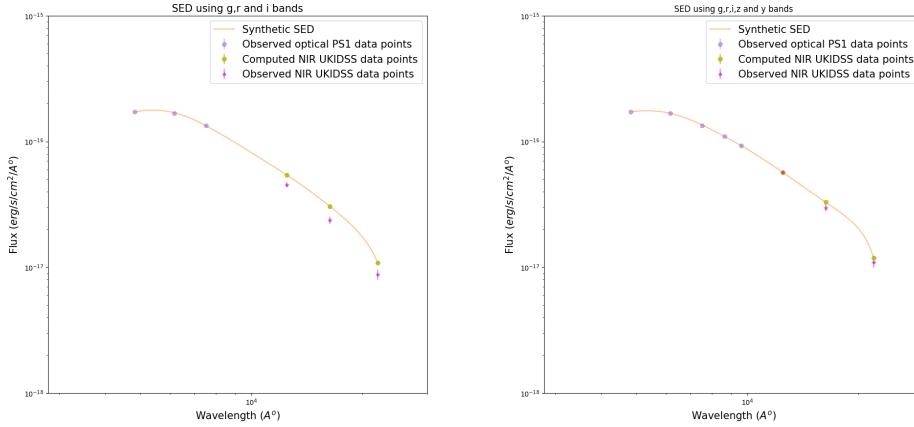


Figure 13: For a particular star in the second sequence, these plots show the PS1 optical (blue), UKIDSS observed (magenta), and computed NIR (yellow) points plotted on top of the spectral energy distribution (orange polynomial) of the best-fit model parameters used to compute the NIR magnitudes. The **left** plot shows that only *g*, *r*, and *i* bands are used to compute the NIR magnitudes, while the **right** plot shows that after adding *z* and *y* bands, the accuracy of estimation of NIR magnitudes increases.

37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53 and 54) we show the comparison between the observed and computed NIR magnitudes that are obtained after applying our optimal method on the remaining 19 test fields. The test fields that do not contain the observed J and H band data in the present UKIDSS data release have only K band comparison plots. The fields TF7, TF8, TF9, and TF10 that lie close to the galactic plane suffer from high levels of interstellar extinction and differential reddening. Also, as the source density in these fields is also very high, we download the PS1 and UKIDSS data for a region of 5 arcmin radius instead of the 30 arcmin radius that we use for other test fields. This not only speeds up the computation of the NIR magnitudes but also improves the accuracy of the predicted NIR magnitudes suffering spatially varying extinction. See Section (8.6 for more discussion).

6 Results

We display the results obtained after applying the optimal set of Kurucz and Phoenix models on the most probable stellar sources of the test fields. There are two result tables: Table (4) for the test fields of size 30 arcmin or 0.785 sq. deg. and Table (9) for 35 arcmin radius or 1.06 sq. deg. region in the sky. In Table (4), $N_{sources}$ is the number of sources in the optical PS1 data for the particular field

having at least two detections, N_{stars} and N_{galax} are the number of stars and galaxies (or extended sources) respectively after the star-galaxy classification, N_{Cat} are the number of sources that are in our catalog, N_{UKIDSS} is the number of sources in the UKIDSS NIR data to validate the computed NIR magnitudes, i_{comp} is the magnitude at which the i band of PS1 data is 90% complete, J_{comp} is the magnitude at which the catalog is 90% complete, $J_{computed}$ denotes the faintest magnitude of the source in the catalog and *Density* denotes the source density in the NFIRAOS FOV (2 arcmin diameter), Δ_J , Δ_H and Δ_K represent the median values of the difference in the observed and the computed J, H, and K respectively at the 1σ level. We note that N_{UKIDSS} contains all the sources in the UKIDSS data that have $SNR>5$. This is a mixture of stars and galaxies and so this number is not equal to the N_{Cat} .

Field	N _{sources}	N _{stars}	N _{galaxy}	N _{cat}	N _{UKIDSS}	i _{convp}	J _{comp}	J _{faint}	Density	Δ _J	Δ _H	Δ _K
TF1	10376	4424	5952	4424	6441	20.94	19.64	22.65	4.91	(90.82%) 0.030 ± 0.131	(61.32%) 0.094 ± 0.254	(65.69%) 0.034 ± 0.277
TF2	13269	3945	4229	3945	4443	20.77	19.42	22.23	4.38	(86.50%) 0.046 ± 0.175	(64.29%) 0.116 ± 0.192	(71.29%) 0.065 ± 0.195
TF3	4628	2108	2520	2108	3298	20.36	18.64	21.00	2.34	(90.65%) 0.051 ± 0.185	(60.23%) 0.128 ± 0.214	(70.37%) 0.037 ± 0.234
TF4	6756	1778	1778	4978	1778	20.75	19.66	20.97	1.97	(85.47%) 0.055 ± 0.162	(53.51%) 0.152 ± 0.264	(60.32%) 0.092 ± 0.303
TF5	6447	1907	4540	1907	2763	20.67	19.31	22.10	2.11	(88.86%) 0.071 ± 0.138	(59.66%) 0.134 ± 0.245	(67.05%) 0.069 ± 0.274
TF6	8089	3327	4762	3327	4121	20.86	19.61	21.98	3.69	(92.45%) 0.031 ± 0.140	(62.51%) 0.085 ± 0.229	(63.46%) 0.003 ± 0.261
TF7*	2400	1700	1700	1700	4775	19.47	17.49	18.59	68.05	(43.38%) -0.208 ± 0.268	(58.87%) -0.068 ± 0.286	(54.90%) 0.106 ± 0.327
TF8*	1508	1281	227	1281	1828	19.99	18.65	20.00	51.25	(83.17%) +0.055 ± 0.178	(66.35%) +0.052 ± 0.253	(58.22%) -0.123 ± 0.276
TF9*	1149	942	207	942	3103	19.82	17.99	19.33	37.69	(13.74%) -0.436 ± 0.700	(11.50%) -0.508 ± 0.877	(17.78%) -0.371 ± 0.862
TF10*	2370	1910	459	1911	5921	19.68	18.01	19.07	76.45	-	-	(42.92%) -0.210 ± 0.275
TF11	4842	1492	3350	1492	4199	20.73	19.37	21.66	1.65	(89.14%) 0.053 ± 0.151	(56.26%) 0.128 ± 0.277	(64.94%) 0.071 ± 0.295
TF12	5388	1752	3636	1752	140204	20.82	19.81	22.57	1.94	(89.01%) 0.060 ± 0.213	(53.01%) 0.123 ± 0.456	(61.05%) 0.0 0.579
TF13	5892	1762	4220	1762	41694	19.83	19.70	20.85	1.95	-	-	(58.24%) 0.061 ± 0.370
TF14	5828	1617	4211	1617	4211	20.73	19.35	21.89	1.80	-	-	(52.95%) 0.089 ± 0.382
TF15	6684	4036	2648	4036	12301	20.42	18.71	21.36	4.48	-	-	(44.09%) 0.054 ± 0.373
TF16	9560	6314	3046	6314	8259	20.39	18.94	21.37	7.23	(70.81%) -0.132 ± 0.168	(60.14%) -0.050 ± 0.259	(59.28%) -0.061 ± 0.262
TF17	8117	2873	5244	2873	6120	21.20	20.05	22.40	3.20	(88.66%) 0.044 ± 0.148	(56.98%) 0.119 ± 0.239	(60.82%) 0.061 ± 0.257
TF18	7243	2230	5013	3884	4057	20.76	19.57	21.85	2.47	(91.92%) 0.040 ± 0.200	(60.38%) 0.106 ± 0.287	(67.06%) 0.063 ± 0.313
TF19	6209	1675	4534	1675	1123	21.01	19.89	22.31	1.86	(80.55%) 0.077 ± 0.237	(50.60%) 0.174 ± 0.253	(62.48%) 0.109 ± 0.220
TF20	5510	1599	43196	1599	43196	20.66	19.58	22.22	1.77	-	-	(61.46%) 0.057 ± 0.429

Table 4: The results obtained after applying the optimal combination of K0, C1, and C2 models to the most probable stellar sources in the optical data of the test fields. The entries marked by the ‘*’ symbol are for the test fields that do not have the observed UKIDSS data in the J and H bands while * indicates the fields that are close to the galactic plane and due to high source density, the data is obtained for 5 arcmin radius only. The % values in the brackets are the percentage number of sources within 0.2 mag of the absolute difference in the observed and the computed NIR magnitudes and the density is computed for the NFIRAOS field-of-view.

7 Validation of the results by alternate method

We validate the computed NIR magnitudes using the *optimal-color method* proposed above by an alternate method called the *flux method*. In this method, we treat A_v and *s.f.* as free parameters and scale the model fluxes to the observed fluxes instead of colors (see Equation 8). As per this method, the model which gives the least value of χ^2_r is the best fitting model for a particular combination of *s.f.* and reddening. The advantage of this method is that it considers discrete extinction values for every star and its distance.

$$\chi^2_r = \frac{1}{N - n_p} \sum_{i=1}^N \sum_{m=1}^M \left(\frac{f_{i,obs} - 10^{-s.f.} \times f_{i,model} \times 10^{-0.4 \left(\frac{A_\lambda}{A_v} \right)_i \times A_v}}{\sigma_{i,obs}} \right)^2 \quad (8)$$

Where N represents the total number of filters, n_p denotes the number of free parameters (*s.f.* and A_v in this case), $f_{i,obs}$ is the observed flux, $f_{i,model}$ is the model flux, $\sigma_{i,obs}$ is the error in the observed flux, *s.f.* is the scale factor, $(A_\lambda/A_v)_i$ is the ratio of the extinction coefficient of a particular filter to A_v . The best-fitted model is found by optimizing the values of the free parameters in the equation for a given f_{obs} and minimizing the χ^2_r hyper-surface. A conventional approach involves using *nested for* loops to iterate through different values of these parameters and the χ^2_r is computed by comparing each f_{model} for each f_{obs} . However, the time complexity for this routine grows exponentially as the number of iterations increases. To address this issue, one common approach is to use *Vectorization* of arrays. This method replaces iterative operations with vector operations, thereby reducing the time required to perform calculations and improving the speed of the fitting routine. Since the dimensions of the stellar atmospheric model array are very large, we implement this algorithm using optical photometry for the TF1 field on the *GoogleColab*. Here we increase the speed of the computation by making use of the parallel computing ability of the GPUs enabled by the *PyTorch* python package.

To vectorize the computation of the χ^2 , we construct a 5D array where, the first axis i.e. $axis_0$ represents the array of models fluxes, $axis_1$ represents *s.f.* array (split into 200 bins in between the range 20.0 to 25.0), $axis_2$ represents A_v values (split into 200 bins in the range from 0.0 to 3.0), $axis_3$ represents the observed fluxes and $axis_4$ represents the filters. Since $axis_3$ has the observed fluxes of different stars, it is possible to find the best-fit parameters for multiple stars in a single run. Since each array has different dimensions, we make use of a Python property called *broadcasting* where the dimensions of each array are made compatible with each other before computing χ^2_r . Once all inputs have become 5D arrays, we use equation (8) for calculating χ^2_r by summing over the filter axis. Then the indices of the minimum χ^2 value across $axis_0$, $axis_1$ and $axis_2$ provide the best-fitted model, *s.f.*, and A_v information.

We apply the flux method to the PANSTARRS optical data of the TF1 field in two ways. Firstly, we keep A_v as a free parameter within the range mentioned above. Since a wider range of A_v may lead to over-fitting and un-realistic prediction of A_v , next, we choose a tight range for A_v where $A_v \in \{A'_v, 3\sigma A'_v\}$ and A'_v is taken from [Schlafly & Finkbeiner \(2011\)](#). We plot the results in Figure (14)

where we show the correlation of the computed NIR magnitudes using these methods and the mean value of A_v is 0.13. The colorbar represents the difference in the observed and the NIR magnitudes computed when A_v is not fixed. In the next Figure (15), we show the comparison of the difference in the computed magnitudes when A_v is fixed in the optimal color method and when A_v is not fixed. The colorbar in these plots represents the A_v value corresponding to χ^2_{min} . Figure (16) shows the distribution of the model parameters using these methods. We find that the NIR magnitudes predicted when A_v is bound by 3σ are similar to the NIR magnitudes computed when A_V is fixed and observed UKIDSS NIR magnitudes. Although placing every star at a different distance by treating *s.f.* and A_v as free parameters is realistic, in the case of NIR bands, the computed magnitudes when A_v is bounded and unbounded by the 3σ limit do not differ significantly from the computed NIR magnitudes using the optimal-color method where A_v is fixed. This is the case, especially for the sources showing larger deviations from the observed NIR magnitudes. Also, keeping A_v unbounded gives a larger scatter with respect to the observed magnitudes, meaning that A_v takes any value to minimize the χ^2_r . Since computing the NIR magnitudes by keeping A_v and *s.f.* free is computationally more extensive and time-consuming, we suggest that using the optimal color method can be used to generate IRGSC.

8 Discussion

Based on the requirements decided after Phase II of the study, we have devised a method to incorporate the PS1 data with sources observed in five optical bands. This method is devised to reach the required faintness of $J_{Vega} = 22$ mag, achieving the source density criteria in the NFIRAOS FOV while maintaining the accuracy of the computed NIR magnitudes. In this section, we discuss the results in detail.

8.1 Accuracy of the computed NIR magnitudes and the faintness achieved

The results shown in Table (4) are obtained after applying the combination of K0, C1, and C2 models discussed in Section (A.0.2) to the most probable stellar sources in the fields. We aimed to compute the NIR magnitudes for the probable stellar sources in each test field by maintaining the required accuracy of 0.2 mag and also reach up to $J_{Vega} = 22$ mag. We see that the faintest J computed for most of the test fields is ~ 22 mag, and the median value of the difference in observed and computed J for all the test fields is 0.01-0.02, which is an acceptable range. However, the spread is larger in

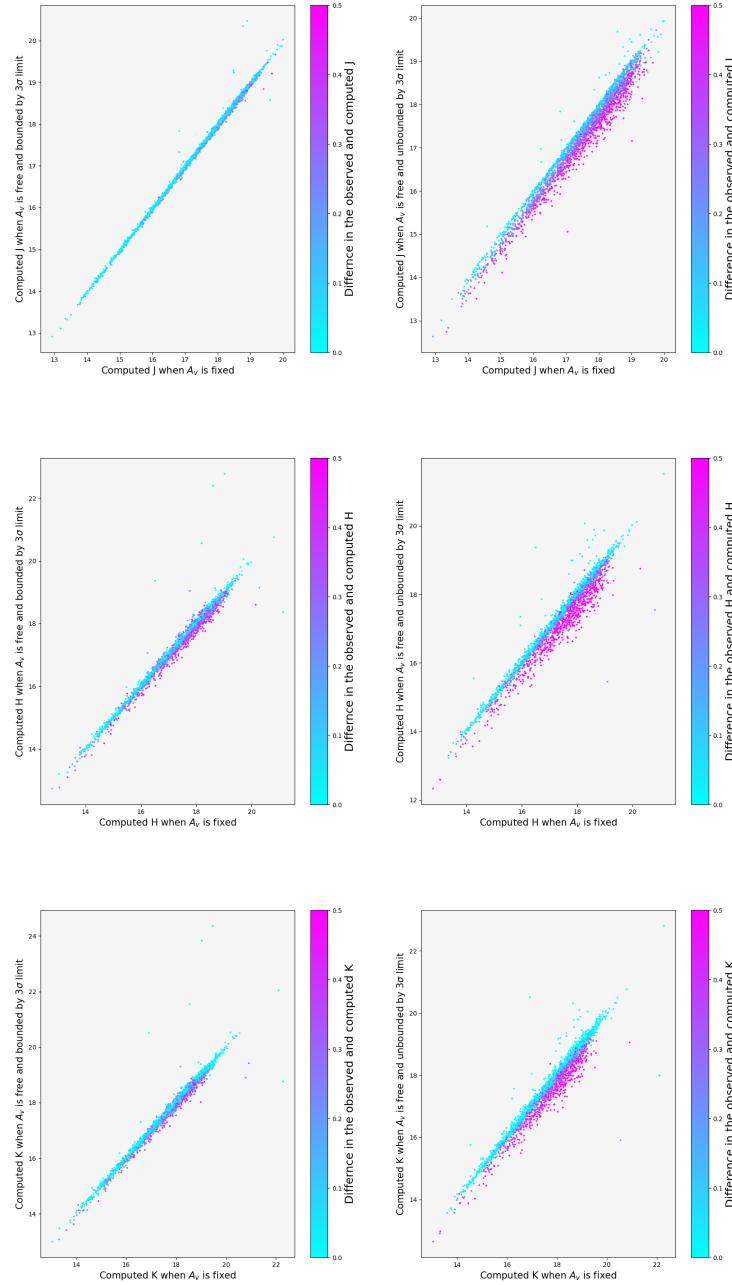


Figure 14: These plots show the comparison between the observed and the computed NIR magnitudes (J, H, and K from top to bottom) when left A_v is bounded by 3σ limit and right is not bounded by any 3σ limit. The colorbar shows the difference in the observed and computed J using the used method. The plots using the former method show a very tight correlation indicating that the computed NIR magnitudes are very similar while the plots using the latter method show a loose correlation indicating that for some sources, the A_v takes unrealistic value to minimize the χ^2 . The mean value of A_v is 0.13.

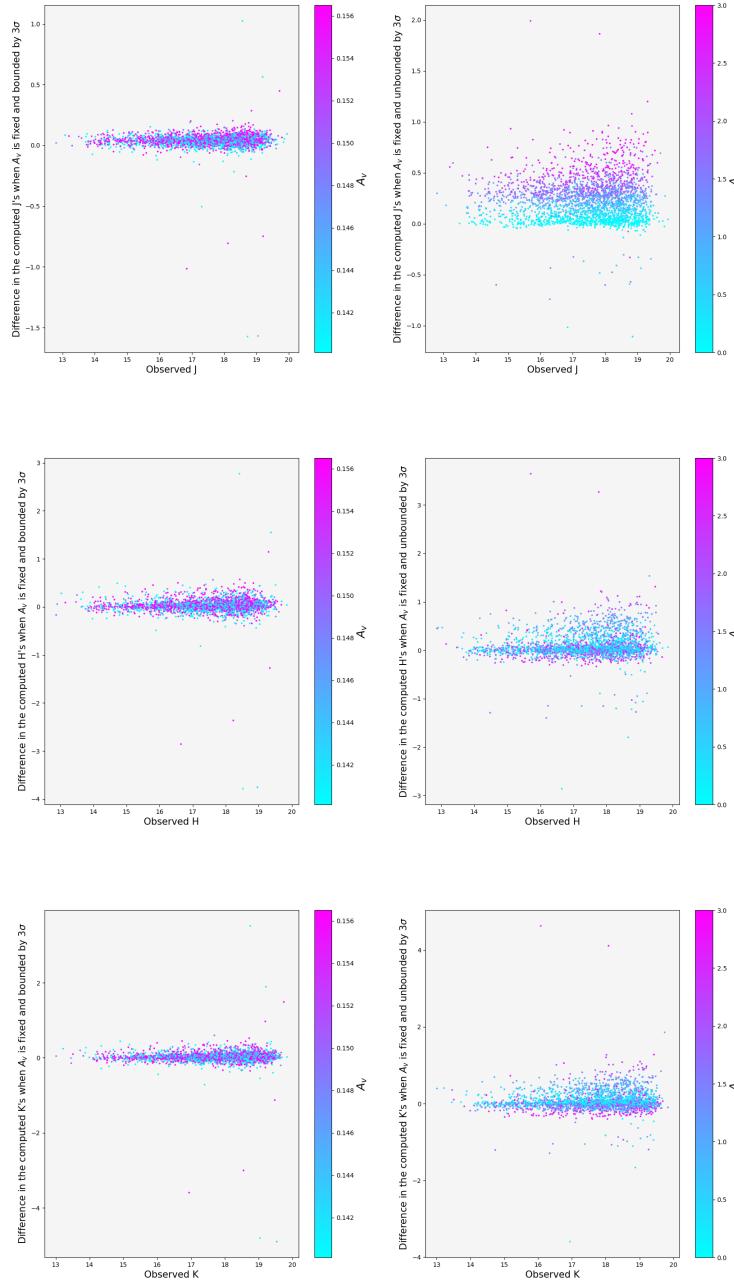


Figure 15: The plots in the left side of the figures show the comparison of the difference in the computed NIR magnitudes when A_v is fixed and when A_v is kept free but bounded by 3σ limit. The colorbar shows the A_v values that each source takes in order to minimize the χ^2_r . The plots in the right side of the figures show the comparison of the difference in the computed NIR magnitudes when A_v is fixed and when A_v is kept free and unbounded by the 3σ limit. The mean value of A_v is 0.13.

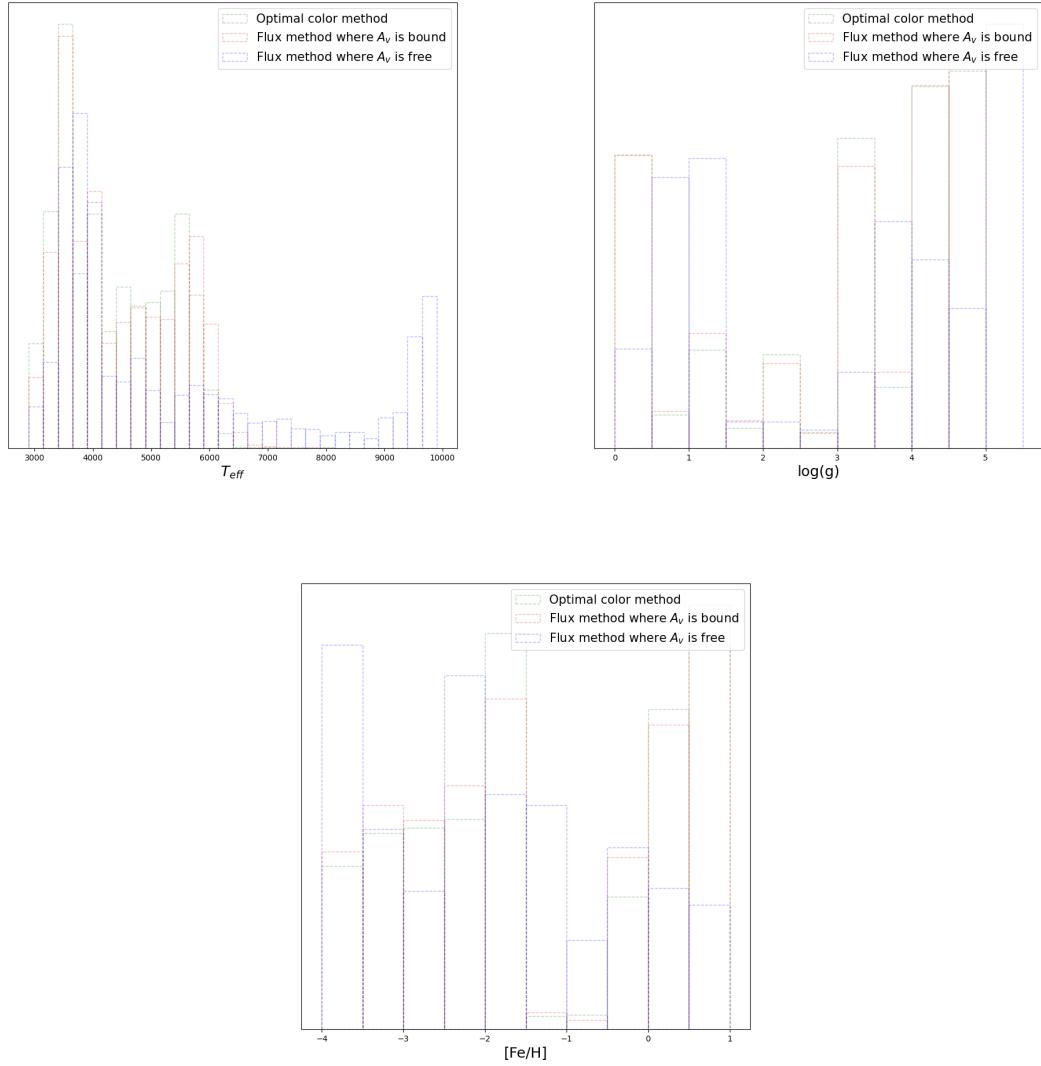


Figure 16: These plots show the distribution histogram of the model parameters using the optimal color method, flux method when A_v is bound by 3σ limit and A_v is free. As expected, the model parameters are not similar. However, the computed magnitudes are similar to the observed magnitudes. The bin size is 250K for T_{eff} and 0.5 dex for $\log(g)$ and $[Fe/H]$ respectively. The distribution of the histogram for the optimal color method and the flux method where A_v is bounded by 3σ limit is nearly similar.

the H and K bands than in the J band. This can be attributed to the poor modeling of the ultra-cool, metal-poor dwarfs by the SAMs in H and K bands.

While the PS1 *stack* data is deep, our condition that the input sources must have at least one detection in each optical band (see Section (5.3.1)) removes many fainter sources from the data which could increase the magnitude depth or the 90% completeness of the IRGSC. This is shown in Figure (17), where we plot the histograms of i_{psf} at various stages for the TF1 field data. We also see in *top-left* plot of Figure (17) that if used individually, the PS1 i-band can go up to 30 mag, which is sufficient to reach $J_{comp} = 22$ mag. However, in the next plot, we can see that as we use the data from all the bands, fainter objects in the i-band do not appear in the data set. The following plot shows the i-band distribution of the probable stellar sources obtained after the star-galaxy classification while the last histogram shows the i-band distribution of the sources in the generated IRGSC for the TF1 field. Thus it is seen that our condition that the source must have at least one detection in each band removes many faint objects, which could lead to $J_{Vega} = 22$ mag.

8.2 Source Density criteria

Along with the source faintness, an important thing to achieve was fulfilling the source density criteria for NFIRAOS FOV. The condition of the source density criteria according to Section (4) is three stars in the one arc-min radius NFIRAOS FOV. This number translated to 3440 stars per sq. deg. FOV of our test field size region of the sky. In the result table, excluding the fields located close to the galactic place, the test fields TF1, TF2, TF6, TF12, TF15, TF16 and TF17 satisfy our requirement of the source density and the others do not. This source density is not very small than the required three stars criteria but it is less because these fields are located at a high galactic latitude and have less source density in the input PS1 data itself (see N_{stars} column in the result Table (4) and the next Section (8.3)).

8.3 90% Completeness of the catalog

The 90% completeness of a catalog is the magnitude at which 90% of the sources are present. This number is important because it gives us an idea of the magnitude depth of a catalog and the number of sources reaching that depth. When we plot a distribution of the sources, 90% completeness also resembles the bin in which a maximum number of sources lie. We compare the distribution of the stars in the input PS1 catalog with the synthetically generated catalog of stars using the latest version of the *Besançon Galactic Model* of stellar population synthesis (Czekaj et al., 2014). This model can be accessed through a web query after creating an account on [their official website](#). Here, to

create a new simulation, one has to set up the model by referring to the version of the model and the photometric system to be used. We used *SDSS + 2MASS JHK* photometric system as PANSTARRS, and UKIDSS photometric systems are unavailable on the web page and resemble the former systems. Since we are only interested in the star counts towards the line of sight of the field, for simplicity, we do not select a model with the kinematics of stars included. The next step is to provide the coordinates and range of magnitudes and colors that the stars should be allowed. We apply no range here so that we get all the sources in the field. The *Besançon* model uses [Marshall et al. \(2006\)](#) extinction law which is valid from $-100^\circ < l < 100^\circ$ and $-10^\circ < b < 10^\circ$. Although some of our test fields lie outside the galactic plane region, we still use this law in the simulation because the dust column density is very less towards these fields. In this way, we generate the simulated stellar counts toward the center of all the test fields without any constraint on the magnitude and color range. We have already listed the completeness value of the generated IRGSC for all of our test fields and show the completeness plot for the TF1 field in Figure (18).

8.4 Applying the optimal methodology on sources detected in g, r, and i bands only

We have seen that the required depth of $J_{Vega} = 22$ mag and the required source density is not satisfied in some of the test fields like TF11, TF114, TF19, and TF20, due to the limitation of the PANSTARRS data. So we apply our optimal methodology for the sources in these test fields detected only in g, r, and i bands and see the results. We note that the star-galaxy separation is based on these three bands only. Naturally, the number of sources in the IRGSC will increase but at the cost of the accuracy of the computed NIR magnitudes. Overall we find that there is an increase of up to 35% in the sources and the completeness of the NIR magnitudes for these fields goes beyond 20 mag in J_{Vega} . Both of these changes are sufficient to satisfy our requirements, but the accuracy of the computed NIR magnitudes decreases. There is a reduction of up to 16% in the number of sources lying within -0.2 to +0.2 of the difference in the computed and observed J mag. The same number is up to 10% for the H band and 11% for the K band respectively (see Figure (19)). In the next section, we also show that the efficiency of star-galaxy separation also reduces when we use only the g, r, and i bands (see Figure (8.5)).

8.5 Testing the efficiency of star-galaxy separation

To test the efficiency of the star-galaxy classification method, we download the HUBBLE Source Catalog v3.1 (HSC v3.1) ([Whitmore et al., 2016](#)) which is available for TF4, TF11, TF13, TF14, and TF15 test fields. The Hubble Source Catalog (HSC) is a list of the sources in the Hubble Legacy

Field	Stars in PS1	Galaxies in PS1	Sources in HSC	IRGSC stars as HSC galaxies	IRGSC stars as HSC stars
TF4	1778	4978	1569	0	4
TF11	1492	3350	478	3	12
TF13	1762 (2282)	4220 (3375)	9414	18 (25)	34 (40)
TF15	4036	2648	502	0	1

Table 5: Table showing the efficiency of Star-Galaxy Classification (SGC). The first column indicates the field number; the second column indicates the number of probable stellar sources in the optical PS1 data set after applying the SGC (Equation (1)), the third column indicates the number of galaxies in the PS1 data. The next column indicates the number of sources in the Hubble Source Catalog (HSC) data, followed by the column indicating the number of probable stellar sources in IRGSC that are flagged as galaxies in HSC data and the last column indicates the number of stellar sources in IRGSC that find a match in the HSC data. The HSC data is available only for TF4, TF11, TF13, and TF15 fields and the coverage is very small as compared to the PS1 coverage (see Figure (8.5)). The colored numbers in the TF13 row indicate the values when g, r, and i bands are used.

Archive (HLA) into a single master catalog. We note that the sky coverage of HSC data is not similar to the sky coverage by PANSTARRS (see Figure (8.5) and covers only a small fraction of the area. So we cannot match all the sources in our PS1 data to the HSC data. However, we carry out our procedure to check the efficiency of the star-galaxy separation on the available data. The sources in the HSC data have a flag '0' for the sources that are stars while '1' for the sources that are galaxies. We match the position of the PS1 sources to the HSC sources. Since the resolution of the Hubble Space Telescope is very high compared to the ground-based PANSTARRS telescope, we might find multiple HSC sources that find a positional match to the HSC sources within 1 second. In that case, we select the HSC source that is the closest to the PANSTARRS source. Using this technique, we select the objects that are classified as stars using our star-galaxy separation criteria (see Equation (1), but are galaxies in the HSC data which have more resolution. Similarly, we list the sources classified as stars using our criteria and also are stars in HSC data. We show the results in Table (5). We also show how the efficiency of the star-galaxy separation is affected when g, r, and i bands are used instead of g, r, i, z, and y bands. In Figure (8.5), we show the comparison of $i_{psf} - i_{kron}$ (on y-axis) and i_{psf} (on x-axis). The magenta-colored points are the most probable stellar sources according to Equation (1), while the black-colored points are galaxies. The points marked in blue and green color are the probable stellar sources according to the PANSTARRS data using the five and three optical bands respectively but are galaxies in the HSC data. We see that the number of green points is larger than the blue points, indicating that the efficiency of the star-galaxy separation reduces with the number of bands.

8.6 Fields close to the Galactic Plane

Fields TF7, TF8, TF9, and TF10 lie close to the galactic plane. These fields also contain high amounts of dust, contributing to high extinction and differential reddening. Therefore none of the methods can

compute the NIR magnitudes up to the required levels of satisfaction eg. See Figures (41, 42, 43 and 44) which show the comparison of the difference in the observed and computed NIR magnitudes vs the observed NIR magnitude for these test fields. Not only the number of sources lying within a difference of 0.2 magnitudes between the observed and the computed NIR magnitudes is significantly less, but also there is a shift in the scatter as a whole towards the brighter side indicating that the computed NIR magnitudes are brighter than the observed. We, therefore, recommend using the stellar sources from the readily available NIR surveys (eg. *Visible and Infrared Survey Telescope* (VISTA) (?)) in the galactic plane regions as guide stars. VISTA has a primary mirror of 4.1m and observes the sky in NIR bands. The latest VISTA survey comprises of six public surveys eg. *UltraVISTA*, *VISTA Kilo-Degree Infrared Galaxy Survey (VIKING)*, *VISTA Magellanic Survey (VMC)*, *VISTA Variables in the Via Lactea (VVV)*, *VISTA Hemisphere Survey (VHS)*, and *VISTA Deep Extragalactic Observations Survey (VIDEO)*. While these surveys have a larger photometric depth, in particular, the data from the VVV survey (Saito et al., 2010) can be particularly helpful in obtaining information about the NIR stellar sources near the galactic plane. A study on the usability of this data can be performed in the future.

8.7 Fields with declination $< -30^\circ$

While the entire observable sky for TMT will be north of -45° in declination, the methodology to generate IRGSC from the PANSTARRS 3-pi survey DR2 is applicable for fields located above -30° declination. This is because of the coverage of the PANSTARRS survey which is north of -30° degrees in declination. So we need to find an alternative way to generate IRGSC for the narrow strip in between -30° and -45° of the sky. The VHS survey (McMahon, 2012) covers the entire sky in the southern hemisphere and the stellar sources from this survey data can be used for IRGSC. However, it is to be noted that the data from this survey is available in J, H, K_s bands instead of the required J, H, K and the depth of the survey is only up to J=20.6 mag. Alternatively, the optical data from *Vera C. Rubin Observatory* (formerly known as Legacy Survey of Space and Time: LSST) which will observe the entire sky that is visible from Chile (Ivezic et al., 2019), could be used to generate the IRGSC. However, the optimal methodology would have to be tested and modified if necessary before computing the NIR magnitudes of the sources in the optical data of LSST.

8.8 Creation of the IRGSC for additional fields

In addition to the test fields used in this work to test and validate our methodology, we have created the IRGSC for additional ten fields using *irgsctool*, a software code created for the generation of IRGSC (see next section for details). These fields belong to the *Medium Deep Survey (MDS)* of PANSTARRS

Field	α	δ	$E(B - V)_{mean}$
F1	10.675	41.267	0.498
F2	35.875	-4.250	0.0223
F3	53.100	-27.800	0.0071
F4	130.592	44.317	0.0229
F5	161.917	58.083	0.0093
F6	185.00	47.117	0.0197
F7	213.704	53.083	0.0084
F8	242.787	54.950	0.4985
F9	334.188	0.283	0.0516
F10	352.312	-0.433	0.0347
F11	270.000	66.561	0.0396

Table 6: Table showing the positions of the Medium Deep Survey fields and the mean value of the reddening towards those fields.

(see Table (6)). Although the completeness of the MDS is higher than the 3pi survey data, it is not publicly available. So we have used the readily available 3pi survey data for these fields and generated the catalog. However, IRGSCs using the deeper data, whenever it is made available, can be created and the results can be compared with the catalog created using the 3pi survey data. This will help us to understand whether deeper into optical data helps in improving the accuracy of the computed NIR magnitudes by satisfying the other requirements of the IRGSC.

In addition to the catalog for our twenty test fields, the catalog for these fields is also available for the public on [TMT GitHub repository](#)¹⁴. It is to be noted that the MD01 field coincides with the Andromeda galaxy. So we exclude it from the analysis.

9 Nature of the IRGSC

The IRGSC generated using the optimal method applied on the *stack* photometric data of the PANSTARRS has various information about the sources shown in Table (7). This table describes the columns in the IRGSC generated for a particular test field. The details of the flags, e.g., infoflags, filterflags, and qualityflags can be found in [Flewelling et al. \(2020\)](#). These flags indicate various values assigned to the source by the PANSTARRS team, which gives further information about the nature of the source and the quality of its detection, which can help understand more about a particular object of interest. A CasJobs SQL query to extract the information in each flag is given in Appendix (D) and users will have to run this query on MAST CasJobs each time to find out the hidden combination of flags that

¹⁴The IRGSC for these fields are 15 arc-min in radius due to the limitation of the python package used by *irgsctool* to retrieve the PANSTARRS data

has generated a particular flag value. It is to be noted that we use the *stack* photometric information in our analysis, but we use the Right Ascension and Declination of the source from *mean* photometric database as they are well calibrated using Gaia DR2 (Magnier et al., 2020). For the generated IRGSC, we include the additional astrometric information from Gaia DR3 like the proper motion and parallax. We also include a flag from Gaia DR3 called *re-normalized unit weight error (ruwe)*. *ruwe* indicates whether the source is single or a part of a multiple-star system. A value less than 1.3 indicates that the star is single. The code to generate IRGSC is in the form of a Python package called *irgsctool* and can be obtained in three ways:

1. Using pip: pip install irgsctool
2. Using .zip file from GitHub: pip install '.' after downloading and unzipping the file.
3. Using the development version from GitHub: pip install git+https://github.com/dms-irgsc/irgsc@main

The *Readme* file can be accessed here <https://github.com/tmtsoftware/dms-irgsc/blob/main/README.md>. *irgsctool* generates the catalog for a given set of R.A., Dec. (both in degrees), and the default search radius is 0.25 degrees due to the limitation of a dependant Python package *pyvo*. *irgsctool* also automatically retrieves the UKIDSS DR11 and Gaia DR3 data, reddening and NIR extinction values along the line-of-site of the field. The generated as well as validated IRGSC is saved in *comma-separated values (.csv)* format with the date and time stamp of when it was generated. In regions where the data cannot be obtained, an appropriate *Warning* is raised. Details of the content of the generated IRGSC can be found in Table (7) and the detailed software/API documentation of the code and tutorials can be found on <https://tmtsoftware.github.io/dms-irgsc/>.

10 Nature of the validated IRGSC

In addition to the generation of IRGSC, *irgsctool* also validates the computed NIR magnitudes with the observed NIR UKIDSS sources obtained for the same region of the sky. The positions of the PS1 and UKIDSS sources in the sky are cross-matched up to 1 arcsecond and the *Validated IRGSC* catalog is generated. Apart from the columns described in Table (7), the validated catalog contains additional columns given in Table (10) which are mainly dependent on the UKIDSS data. This catalog cannot be generated if there is no observed UKIDSS data available for the given field.

Column Name	Description
ps1 objid	Object ID in PS1 data (float).
ps1 ra	R.A. of the source in PS1 single epoch (<i>mean</i> photometry) data (float).
ps1 ra error	Uncertainty in R.A. (float).
ps1 dec	Dec. of the source in the PS1 single epoch (<i>mean</i> photometry) data (float).
ps1 dec error	Uncertainty in Dec. (float).
ps1 gpf	psf magnitude of the source in g-band (float).
ps1 gpf error	Uncertainty in the psf magnitude in g-band (float).
ps1 rpsf	psf magnitude of the source in r-band (float).
ps1 rpsf error	Uncertainty in the psf magnitude in r-band (float).
ps1 ipsf	psf magnitude of the source in i-band (float).
ps1 ipsf error	Uncertainty in the psf magnitude in i-band (float).
ps1 zpsf	psf magnitude of the source in z-band (float).
ps1 zpsf error	Uncertainty in the psf magnitude in z-band (float).
ps1 ypsf	psf magnitude of the source in y-band (float).
ps1 ypsf error	Uncertainty in the psf magnitude in y-band (float).
teff	T_{eff} of the best-fitted model (float).
logg	$\log(g)$ of the best-fitted model (float).
feh	[Fe/H] of the best-fitted model (float).
sam g	Synthetic g magnitude in PANSTARRS filter(float) (in AB system).
sam r	Synthetic r magnitude in PANSTARRS filter(float) (in AB system).
sam i	Synthetic i magnitude in PANSTARRS filter(float) (in AB system).
sam z	Synthetic z magnitude in PANSTARRS filter(float) (in AB system).
sam y	Synthetic y magnitude in PANSTARRS filter(float) (in AB system).
sam j	Synthetic j magnitude in UKIDSS filter(float) (in AB system).
sam h	Synthetic h magnitude in UKIDSS filter(float) (in AB system).
sam k	Synthetic k magnitude in UKIDSS filter(float) (in AB system).
scale factor	The scale factor computed after matching the SAM to the observed data (float).
scale factor error	Uncertainty in the scale factor (float).
d_{dev}	The minimum value of d_{dev} as defined in Section (5.6) (float).
J	Computed J in the Vega system (float).
J error	Uncertainty in the computed J (float).
H	Computed H in the Vega system (float).
H error	Uncertainty in the computed H (float).
K	Computed K in the Vega system (float).
K error	Uncertainty in the computed K (float).
gaia source id	Source ID in Gaia DR3 catalog (float).
gaia ra	R.A. of the source in Gaia DR3 (float).
gaia ra error	Uncertainty in R.A. (float).
gaia dec	Dec. of the source in Gaia DR3 (float).
gaia dec error	Uncertainty in Dec. in Gaia DR3 (float).
gaia parallax	Parallax of the source from Gaia DR3 (float).
gaia parallax error	Uncertainty in the parallax (float).
gaia pm	Proper motion in Gaia DR3 catalog (float).
gaia pm ra	Proper motion along the Right Ascension axis in Gaia DR3 catalog (float).
gaia pm ra error	Uncertainty in the proper motion along the Right Ascension axis (float).
gaia pm dec	Proper motion along the Declination axis in Gaia DR3 catalog (float).
gaia pm dec error	Uncertainty in the proper motion along the Declination axis (float).
gaia ruwe	Re-normalized unit weight Error (float) metric in Gaia DR3. A value less than 1.4 is likely a single star.
objinfoflag	These flag values of the source in PANSTARRS data specify whether the object is a QSO, transient, asteroid, extended, a known solar system object, etc. in nature (float).
objqualityflag	These flag values denote if an object is real or a possible false positive (float).
ndetections	The number of times something is detected from the individual exposures (float).
nstackdetections	The number of stack detections after which the stack photometric measurements are done (float).
ginfoflag	These flags indicate the details of the g filter stack photometry (float).
ginfoflag2	These flags indicate the details of the g filter stack photometry (float).
ginfoflag3	These flags indicate the details of the g filter stack photometry (float).
rinfoflag	These flags indicate the details of the r filter stack photometry (float).
rinfoflag2	These flags indicate the details of the r filter stack photometry (float).
rinfoflag3	These flags indicate the details of the r filter stack photometry (float).
iinfoflag	These flags indicate the details of the i filter stack photometry (float).
iinfoflag2	These flags indicate the details of the i filter stack photometry (float).
iinfoflag3	These flags indicate the details of the i filter stack photometry (float).
zinfoflag	These flags indicate the details of the z filter stack photometry (float).
zinfoflag2	These flags indicate the details of the z filter stack photometry (float).
zinfoflag3	These flags indicate the details of the z filter stack photometry (float).
yinfoflag	These flags indicate the details of the y filter stack photometry (float).
yinfoflag2	These flags indicate the details of the y filter stack photometry (float).
yinfoflag3	These flags indicate the details of the y filter stack photometry (float).
SAM name	The name of the best-fitted Stellar Atmospheric Model (SAM).

Table 7: The table shows the name of the columns in the IRGSC. Here *pm* is the acronym for proper motion, *sam* stands for the name of the stellar atmospheric model and *ps1* stands for PANSTARRS. The information of the PANSTARRS flag values can be found in (Flewelling et al., 2020) and a CasJobs query to extract information hidden in each flag value is given in Appendix (D).

Column Name	Description
diff_J	Difference in the observed and computed J (float).
diff_H	Difference in the observed and computed H.
diff_K	Difference in the observed and computed K (float).
J_UKIDSS	J mag from the UKIDSS observations (petro mag) (float).
err_J_UKIDSS	Uncertainty in J_UKIDSS (float).
H_UKIDSS	H mag from the UKIDSS observations (petro mag) (float).
err_H_UKIDSS	Uncertainty in H_UKIDSS (float).
K_UKIDSS	K mag from the UKIDSS observations (petro mag) (float).
err_K_UKIDSS	Uncertainty in H_UKIDSS (float).

Table 8: The table shows the name of the additional columns in the validated IRGSC. Here the observed UKIDSS NIR magnitudes are petro magnitudes.

11 Summary

In this work, we have demonstrated a way to generate the partial IRGSC on the framework laid by [Subramanian et al. \(2016\)](#). We generate a catalog of synthetic photometry for interpolated Kurucz and Phoenix models by convolving them with PS1 and UKIDSS response functions. We have identified twenty test fields across the sky located at different galactic latitudes with readily available PANSTARRS DR2 optical and UKIDSS NIR observed data and obtained 30 arcmin radius data for them. After applying the several methods discussed here (which are the modification of the previous work), we find an improvement in the computation of the NIR magnitudes when additional two bands - "y" and "z" are used. Adding longer wavelength bands in the future can further help to improve this study.

Due to *stack* photometry, many faint objects are detected, which are cooler giants or dwarfs and metal-poor, e.g., Brown dwarfs, ultra-cool red dwarfs, etc.). Kurucz models cannot compute the NIR magnitudes of these sources effectively. Hence we see the presence of multiple sequences in the comparison plot of the observed and computed NIR magnitudes. Therefore, we model these sources with selective Phoenix models and find that a combination of interpolated Kurucz model with $T_{eff} > 4000\text{K}$ and two different sets of interpolated Phoenix models within the range $[2800 < T_{eff} < 5000, \log(g) > 3.0, \text{[Fe/H]} < -1.50]$ and $[2800 < T_{eff} < 4000, \log(g) < 3.0, \text{[Fe/H]} > -0.50]$ give good results. The interpolated Kurucz models compute the NIR magnitudes of the hotter sources with good accuracy. In contrast, the first set of the interpolated Phoenix models computes the NIR magnitudes of the metal-poor and cooler dwarfs efficiently. Similarly, the other set of interpolated Phoenix models computes the NIR magnitudes of the metal-rich and cool giants efficiently. However, the use of the NIR magnitudes (especially H and K bands) of the sources that are modeled as an ultra metal-poor ($\text{[Fe/H]} < -2.0$) and ultra-cool ($T_{eff} < 3500\text{ K}$) is cautioned.

After applying the combination of the interpolated Kurucz and interpolated Phoenix models to the *stack* photometry data, the condition for the source density in the NFIRAOS FOV (three stars in one arc-min radius) is satisfied for some fields. While for other fields, the number of sources after the star-galaxy classification is less than the required source density criteria. Therefore, we do not meet the required density criteria in these fields. Our requirement that a source must be detected in five bands removes many fainter sources in every test field. While relaxing our criteria that a source

must be detected in three bands not only increases the number of sources but also the depth of the completeness of the catalog, we show that the accuracy of the computed NIR magnitudes and the efficiency of the star galaxy classification is affected. So the sources that qualify to be in IRGSC after the application of the optimal methodology developed here on the sources not detected in five bands but in three or four bands must be used cautiously.

Our method also computes the NIR magnitudes of the sources up to $J_{Vega} = 22$ mag for all the test fields though it goes deeper than 22 mags for some. The 90% completeness of the PS1 optical input catalog is in the range of 20-22 mags in the i-band. Therefore the 90% completeness of the IRGSC is between 19-20 mags in J_{Vega} . A possible reason why the 90% completeness in the i-band is restricted to 22 mags is because of our condition that the source must have detection in all the five PS1 bands. This situation can change in the future with the release of PS1 DR3 or PS2. We also find that the SAMs cannot efficiently compute the magnitudes of the sources that are metal-poor, ultra-cool, and dwarf in the H and K bands, unlike the J band. This situation can also change in the future with further improvements in the SAMs.

While we have assumed the mean value of the reddening towards the line-of-site of our test fields, in principle, every star is located at different distances from us. Due to this, approximating the reddening as a constant entity is valid only for low-extinction fields. Nevertheless, we compute the NIR magnitudes of the sources in the TF1 field by fitting the model fluxes to the observed flux of the given star in each filter and treating scale factor and extinction as free parameters. We also find that keeping A_v free but bounded by the 3σ range gives us better results rather than keeping A_v totally free. We find that the computed NIR magnitudes by this method match the computed NIR magnitudes by assuming reddening as constant. However, computing the NIR magnitudes by keeping the scale factor and reddening as free parameters is computationally extensive and time-consuming. Instead, our assumption that each star within a 30 arcmin radius field undergoes similar amounts of reddening and extinction is a good approximation. It can be used to generate IRGSC for fields lying away from the galactic plane ($-10^\circ < b < 10^\circ$).

The code for generating an IRGSC is made in *Python* 3.11 and is uploaded on the [TMT GitHub repository](#). This code can be installed in various ways and can be run on any system in a fresh environment. The generated IRGSC contains all the required information, as put forward in the work package deliverable requirements document. Because the PS1 has the proper motion values of some of the sources calibrated w.r.t. Gaia DR1, we have provided the astrometry information for the sources that find a match in Gaia DR3. We have also included *Re-normalised Unit Weight Error (ruwe)* quantity from Gaia DR3, indicating whether a source is a single or a multiple star-system. The IRGSC generated using the code is stored in a *comma-separated values* format with the name containing the input *R.A.* and *Dec.* as well as the date stamp indicating when it was generated. The code also provides all the relevant plots discussed in this report e.g., star-galaxy separation, error pattern plots, comparison between computed and observed NIR magnitudes, etc. Due to the requirement of 1.0 sq. deg. size of the test fields, we have also downloaded the data for the 35 arcmin radius region of the test fields (which is equivalent to 1.06 sq. deg.) and generated the IRGSC using this data. The IRGSC generated using this and the former dataset is also provided in the GitHub repository.

12 Future Development

We suggest the following developments that can be potentially included in the future:

1. **Use better input photometry of the optical sources:** More sources in the input catalog must have detection in all the filters to satisfy our initial selection criteria. The $\text{SNR} \geq 5$ will help in better estimation of the NIR magnitudes of the cooler and fainter sources. The data release version 3.0 of PANSTARRS is expected to have more objects and better photometry of the sources. It will also contain the proper motion values of the objects that are detected.
2. **Use improved SAMs:** Large width of the scatter in H and K bands is due to the limitation of the Phoenix models. Since the accuracy of the Kurucz models is better above $T_{\text{eff}} > 4000\text{K}$, we have to rely on the Phoenix models to increase the source density by modeling the cooler and fainter sources. A solution to this problem can be an improvement in the existing SAM templates in the future.
3. **Use the existing NIR data from the VISTA surveys for the fields lying close to the galactic plane:** As mentioned in Section (8.6), the readily available observed NIR data from the existing NIR surveys like the *Visible and Infrared Telescope for Astronomy* (VISTA), which covers the whole southern hemisphere can be used for IRGSC.
4. **IRGSC for fields below -30° declination:** Since the TMT will observe the sky up to a declination of -45° and the PANSTARRS survey does not cover the sky below -30° declination, we can use the data from the LSST and compute the NIR magnitudes for the optical stellar sources in the narrow strip in between the declination range of -30° to -45° . The optimal methodology developed here will have to be tested and modified to meet the requirements.
5. **Use the latest version of Gaia data:** The space-based astrometry of the objects provided by Gaia is more accurate than any other survey. The future releases of Gaia data are expected to have more sources and thus can be used to incorporate the proper motion and parallaxes of the sources that find a match with the sources in IRGSC.
6. **Include sources not detected in all the five optical bands:** In Section (8.4), we have already shown the changes in the results when only g, r, and i bands are used to generate the IRGSC. We suggest using the sources detected in three or four optical bands also and flagging them separately. These sources should be used cautiously.
7. **Release the all-sky IRGSC v1.0:** An all-sky IRGSC can be generated using the available PANSTARRS DR2 data from the 3-pi survey. This catalog will be based on the optimal methodology developed in this phase. This catalog can also be tried and tested on the existing AO systems implemented on telescopes eg. Keck telescope (Wizinowich et al., 2006).

Acknowledgements

This work made use of Astropy:¹⁵ a community-developed core Python package and an ecosystem of tools and resources for astronomy (Astropy Collaboration et al., 2013, 2018, 2022). This work used Hubble Source Catalog (HSC) based on data from the Hubble Legacy Archive (HLA) based on observations made with the NASA/ESA Hubble Space Telescope and obtained from the Hubble Legacy Archive, which is a collaboration between the Space Telescope Science Institute (STScI/NASA), the Space Telescope European Coordinating Facility (ST-ECF/ESAC/ESA) and the Canadian Astronomy Data Centre (CADC/NRC/CSA). This project also acknowledges the use *Astrolib PySynphot* which is an object-oriented replacement for *STSDAS SYNPHOT* synthetic photometry package in *IRAF*. Although this package was developed for HST, it can be utilized with other observatories. *pysynphot* simulates the photometric data and spectra as they are observed with the Hubble Space Telescope (HST). However, users can also incorporate their own filters, spectra, and data. We thank David Andersen, TMT Project office, for constructive comments on an initial version of the report.

¹⁵<http://www.astropy.org>

A Appendix A: How was the optimal methodology devised?

A.0.1 Modelling all the sources by calculating d_{dev} and using various Kurucz and Phoenix model templates

We start our search to find an optimal method by modeling all the sources using K0 models and calculating d_{dev} for each source. The NIR magnitudes are then computed using the model which gives $d_{dev,min}$. The comparison between the computed and observed NIR magnitudes is shown in the first row of Figure (22). In these plots, similar to the previous section, we find that K0 models are inefficient in low-temperature regimes. We also see the presence of dual sequences in the H and K bands. So next we model all the sources using Kurucz models with $T_{eff} > 4000\text{K}$ (K1 models) only. The second row of Figure (22) shows the comparison between the observed and computed NIR magnitudes when K1 models are used. Here, the second sequence similar to the one in Figure (12) does not appear. However, we find 1992 sources have $|J_{observed} - J_{computed}| > 0.1$ and do not belong to the primary sequence. This is because not all the sources in the field are hot enough to be modeled efficiently with K1 models. In the H and K bands, this number is 2304 and 2688, respectively.

In Figures (23, 24 and 25), we show the comparison of the observed and computed NIR magnitudes where the sources are color-coded according to the model parameters (T_{eff} , $\log(g)$, $[Fe/H]$) and the cumulative deviation of the observed color from the best-fitted model color (dev). We see that the majority of the sources having absolute differences between observed and computed NIR magnitudes > 0.1 are cooler, metal-poor, and compact. But we also find that there is also a presence of another population that comprises cooler, metal-rich, and giant sources. The sources in this population also have a smaller dev value as compared to the former population sources. Since Phoenix models can model the cooler sources in a better way than the Kurucz models (shown by previous S16 study), we select the sources having $[T_{eff} < 4500\text{K}, \log(g) > 3.0, [Fe/H] < -1.5 \text{ and } dev > 1.0]$ and $[T_{eff} < 5000\text{K}, [Fe/H] > -0.5 \text{ and } \log(g) < 3.0]$ and re-model them with Phoenix model templates having $T_{eff} < 4000\text{K}$ (P0 models). In Figures (26, 27 and 28), we show the comparison of the observed and computed NIR magnitudes of the remodeled sources where we see the presence of multiple sequences in the plots, particularly in H and K bands where the second and third sequences are largely displaced from the primary sequence. This indicates that Phoenix models with $T_{eff} < 4000\text{K}$ alone cannot compute the NIR magnitudes of certain ultra-cool (having $T_{eff} < 3500\text{K}$) sources effectively. We, therefore, model the compacts and giants using two different sets of Phoenix model templates. From Figure (23), we know that the outliers are two distinct populations of the sources that are cooler, metal-poor, compact, and relatively hotter than the former, metal-rich, and giant. Therefore, we re-model the compacts using the cooler and metal-poor Phoenix models while the giants are re-modeled using metal-rich and slightly larger T_{eff} range Phoenix models. After some trials, the range of the parameters that we select is $[2800 < T_{eff} < 4000\text{K}, [Fe/H] > -0.50, \text{ and } \log(g) < 3.0]$ and $[2800 < T_{eff} < 5000\text{K}, [Fe/H] < -1.50 \text{ and } \log(g) > 3.0]$. We call them C1 and C2 models respectively. The results are shown in Figures (29,

[30](#) and [31](#)).

A.0.2 Combining optimal Kurucz + Phoenix models

In the previous section, we have seen that using K0 models, K1 and then P0 models individually are ineffective in computing the NIR magnitudes of the sources. We find however that re-modeling of the sources that are scattered largely from the J band primary sequence in Figure ([23](#)) using C1 and C2 model templates not only improves the accuracy of the NIR magnitudes but also converts the random scatter into a sequence (see Figures ([29](#), [30](#), [31](#))). This sequence has a slight shift towards the positive side and the spread is larger in H and K bands. In the catalog, we flag these sources and are suggested to be used only if there are not enough sources to be used as guide stars in the field.

To suggest an optimal methodology for the generation of IRGSC, we model all the sources by combining K1, C1, and C2 model templates and computing d_{dev} for them. The results are shown in Figure ([32](#)). On comparing with the previous plots, we see that a proper selection of the range of $\log(g)$ and $[Fe/H]$ values is crucial to reduce the spread of the scatter as well as to reach the depth of $J_{Vega} = 22$ mags. Thus, this method can be called our *optimal* method to generate IRGSC. In Figures ([33](#), [34](#) and [35](#)), we show the properties of these sources where the median and the spread of the scatter in J, H and K bands is 0.031 ± 0.128 , 0.096 ± 0.243 and 0.037 ± 0.265 respectively. The outliers are again metal-poor, dwarf, and cooler sources. Also, the width of the spread increases as the wavelength increases (i.e., increases from J to K). This effect can be seen across all the test fields (see the next section) and can be due to the limitation of the cooler SAMs.

B Appendix B: Comparison of the computed and observed NIR magnitudes when the optimal methodology is applied to other test fields

This section shows the comparison of the computed and observed NIR magnitudes for the remaining test fields after applying the optimal methodology on stack photometry. Due to the varying completeness of the input optical data and the observed NIR data, the number of sources varies in each plot. For the fields that do not have observed J or H band data available, only K band comparison plots are shown.

C Appendix C: An SQL query to obtain the PANSTARRS data from CasJobs

Here we show the SQL query that we have used to fetch the PANSTARRS DR2 data from the 3-pi survey for our test fields. The photometry is from the stack database while the positions are from the mean database. This query performs a cone search over the given set of *ra*, *dec*, and *search.radius*.

```

select o.objid, o.raMean, o.decMean, o.raMeanErr, o.decMeanErr,
m.gPSFMag, m.gPSFMagErr, m.gKronMag, m.gKronMagErr, m.rPSFMag,
m.rPSFMagErr, m.rKronMag, m.rKronMagErr, m.iPSFMag, m.iPSFMagErr,
m.iKronMag, m.iKronMagErr, m.zPSFMag, m.zPSFMagErr, m.zKronMag,
m.zKronMagErr, m.yPSFMag, m.yPSFMagErr, m.yKronMag, m.yKronMagErr,
o.objInfoFlag, o.qualityFlag, o.nDetections, o.nStackDetections,
m.ginfoFlag, m.ginfoFlag2, m.ginfoFlag3, m.rinfoFlag, m.rinfoFlag2, m.rinfoFlag3,
m.iinfoFlag, m.iinfoFlag2, m.iinfoFlag3, m.zinfoFlag, m.zinfoFlag2, m.zinfoFlag3,
m.yinfoFlag, m.yinfoFlag2, m.yinfoFlag3
into mydb.db from fGetNearbyObjEq(ra, dec, search.radius) nb
inner join ObjectThin o on o.objid=nb.objid
inner join StackObjectThin m on o.objid=m.objid
inner join StackObjectView s on s.objid=o.objid
and o.uniquePspsoBid=s.uniquePspsoBid
inner join AstrometryCorrection a on a.objid=o.objid
WHERE s.bestDetection > 0 AND s.primaryDetection > 0
  
```

D Appendix D: An SQL query to extract the information in the PANSTARRS flags

This is a simple SQL query to extract the information about the different flags (flag.value) in the IRGSC obtained from the PANSTARRS database tables (table.name). The *qualityflag* can be obtained from the *ObjectQualityFlags* table, *objInfoFlag* can be obtained from the *ObjectInfoFlags* table while *XinfoFlag*, *XinfoFlag2*, and *XinfoFlag3* can be obtained from *DetectionFlag*, *DetectionFlag2*, *DetectionFlag3* tables respectively where X belongs to the particular filter.

```
select * into mydb.MyTable from table.name
```

```
where (value flag.value) <> 0
```

```
order by value
```

E Result table after applying the optimal methodology on the twenty test fields but with 1.06 sq. deg. area

One of the deliverables of this phase of IRGSC was to create a partial catalog for twenty fields each with an area of ~ 1.0 sq. deg. Although we have used the data for 0.785 sq. deg. (30 arcmin radius) region in the sky, to develop and test the methodology to generate IRGSC using the optical PANSTARRS data, we have also generated the IRGSC for the same twenty test fields, each with an area of 1.06 sq deg (35 arcmin radius). The catalogs generated for these different-sized test fields are available in TMT GitHub repository¹⁶. Results, similar to Table (4), with the parameters obtained for the 1.06 sq. deg. region of the test fields are presented in Table (9).

¹⁶https://github.com/tmtsoftware/dms-irgsc/tree/main/generated_irgsc

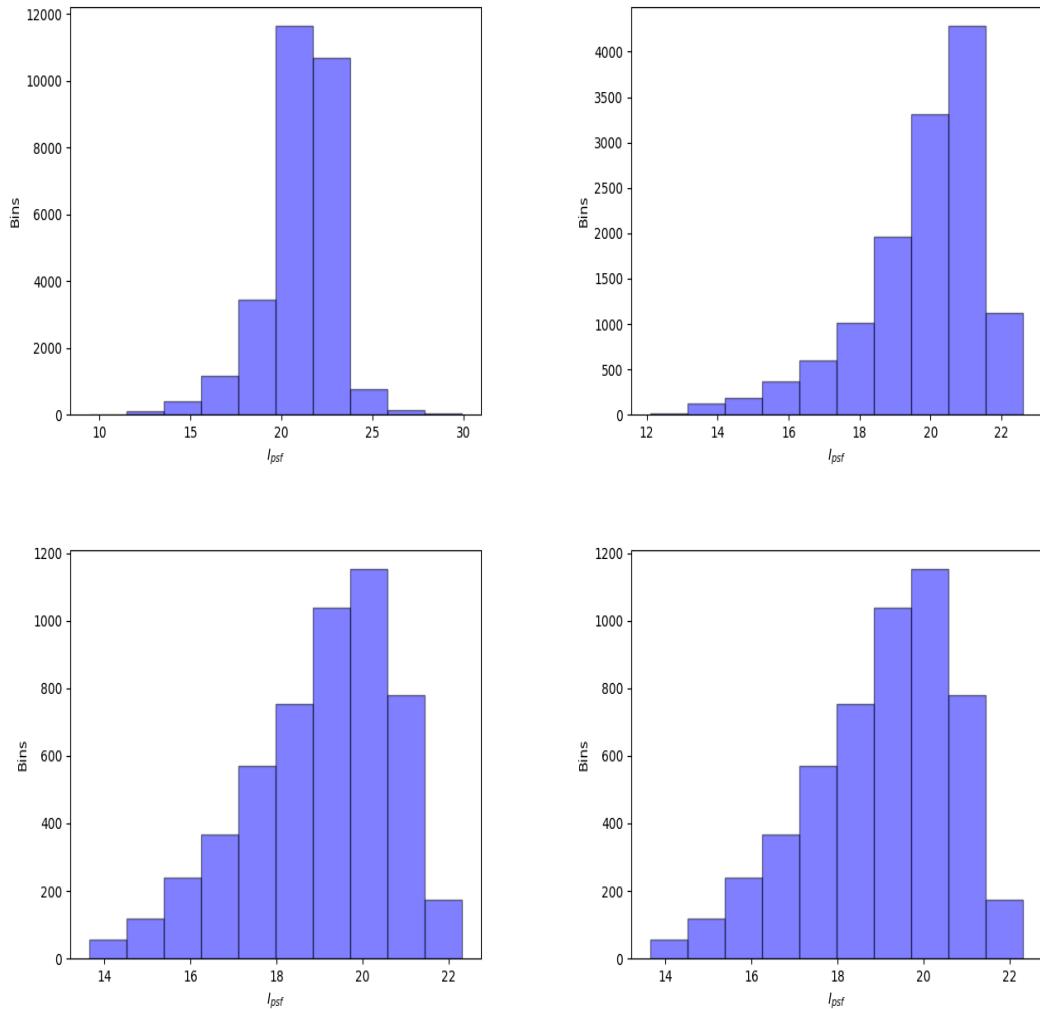


Figure 17: (**Top-left:**) Histogram of I_{psf} when only i-band data is filtered. (**Top-right:**) Histogram of I_{psf} when all the bands are filtered. (**Bottom-left:**) Histogram of I_{psf} after star-galaxy classification (**Bottom-right:**) Histogram of I_{psf} for the sources in the generated catalog. These plots are for stack photometry.

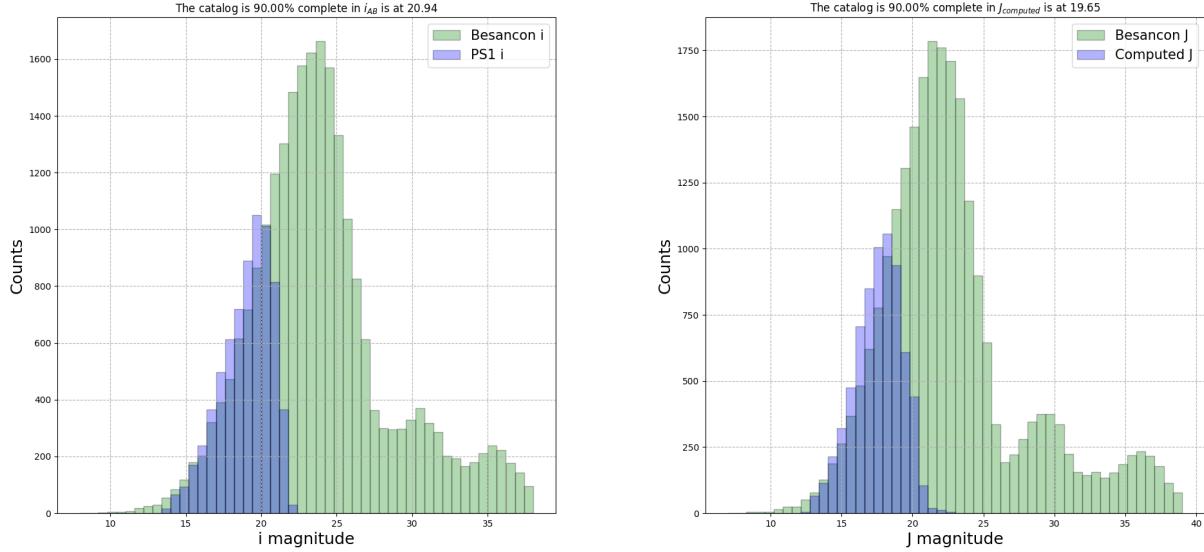


Figure 18: Histogram showing the 90% completeness of the catalog for the TF1 field. The plot in the left shows the i -band completeness of all the probable stellar sources after the star-galaxy classification, while the plot in the right shows the 90% completeness of the computed J magnitudes.

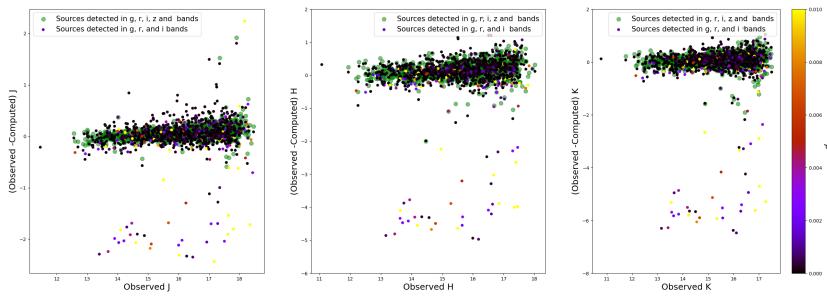


Figure 19: The plots in this figure show the comparison of the observed and computed NIR magnitudes for the TF11 field (one of the fields having a low density of sources) using g , r , and i bands only (multicolored points) which are plotted on top of sources for which the NIR magnitudes are computed using g , r , i , z and y bands (green circles). Some of the sources are poorly modeled when only g , r , and i bands are used and these sources must be used cautiously.

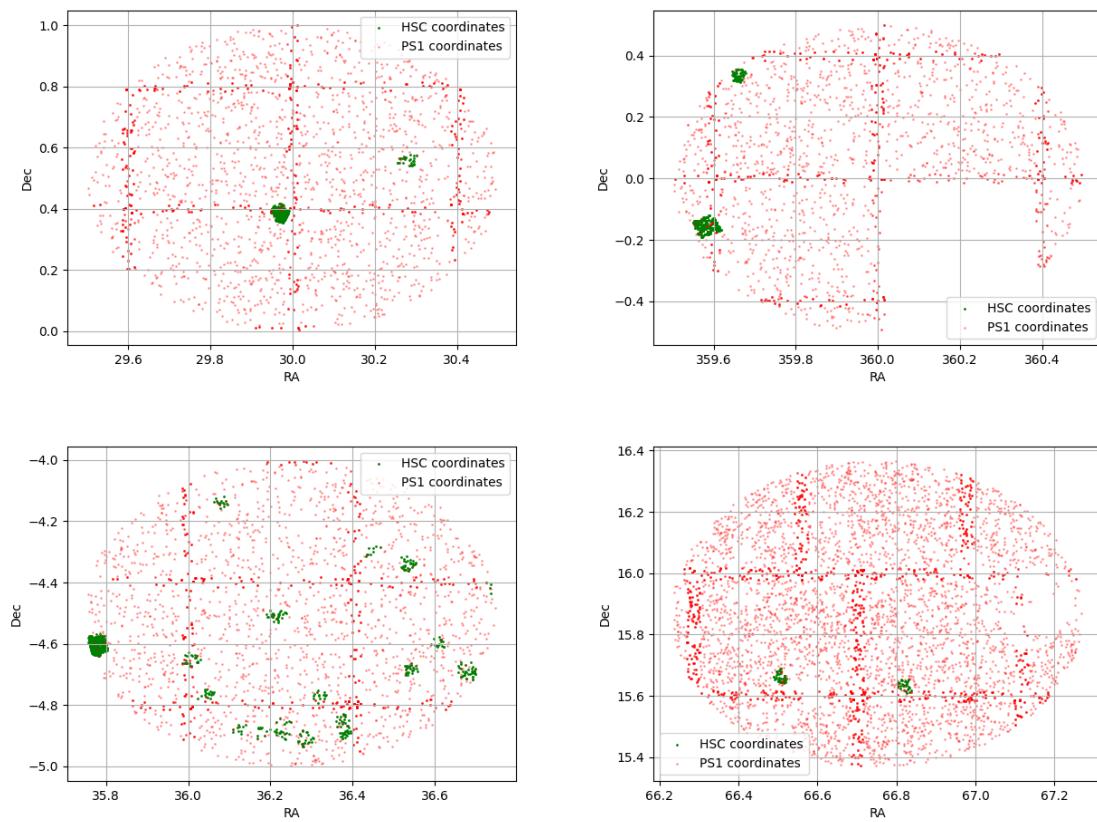


Figure 20: These figures show the overlap of the HSC data on the PS1 data. The HSC data do not cover the entire fields in all the cases.

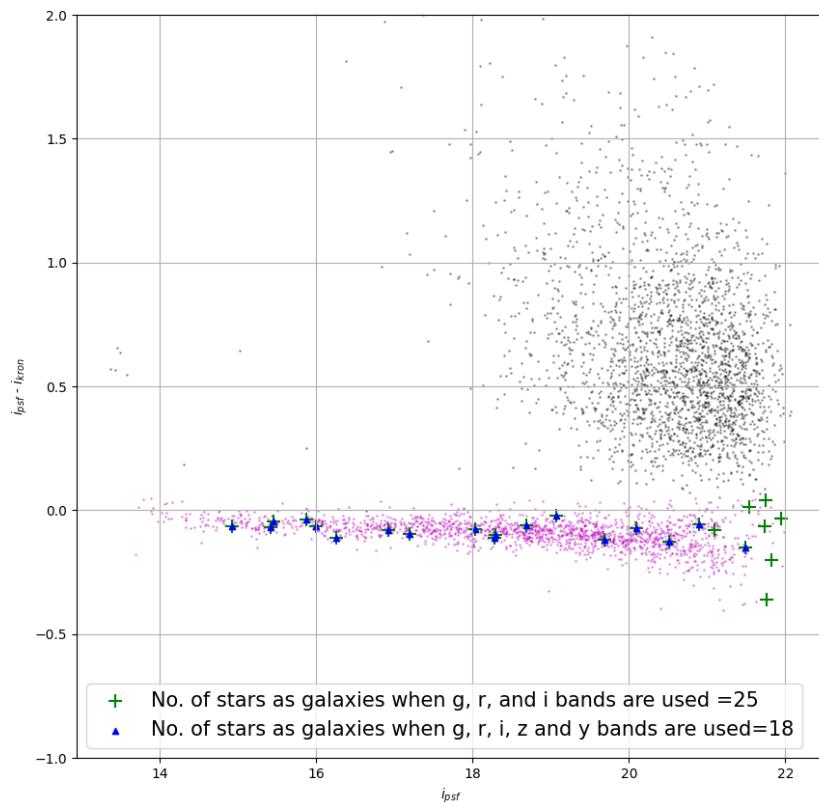


Figure 21: The figure shows the (i_{psf} vs $i_{psf} - i_{kron}$) plot of a number of stars (magenta points) and galaxies (black points) in the TF13 data. The blue triangles represent the sources classified as stars using Equation (1) applied to g , r , i , z , and y bands in the PS1 data but are galaxies in the HSC data. Similarly, green plus symbols represent the sources classified as stars using Equation (1) applied to g , r , and i bands in the PS1 data but are galaxies in the HSC data.

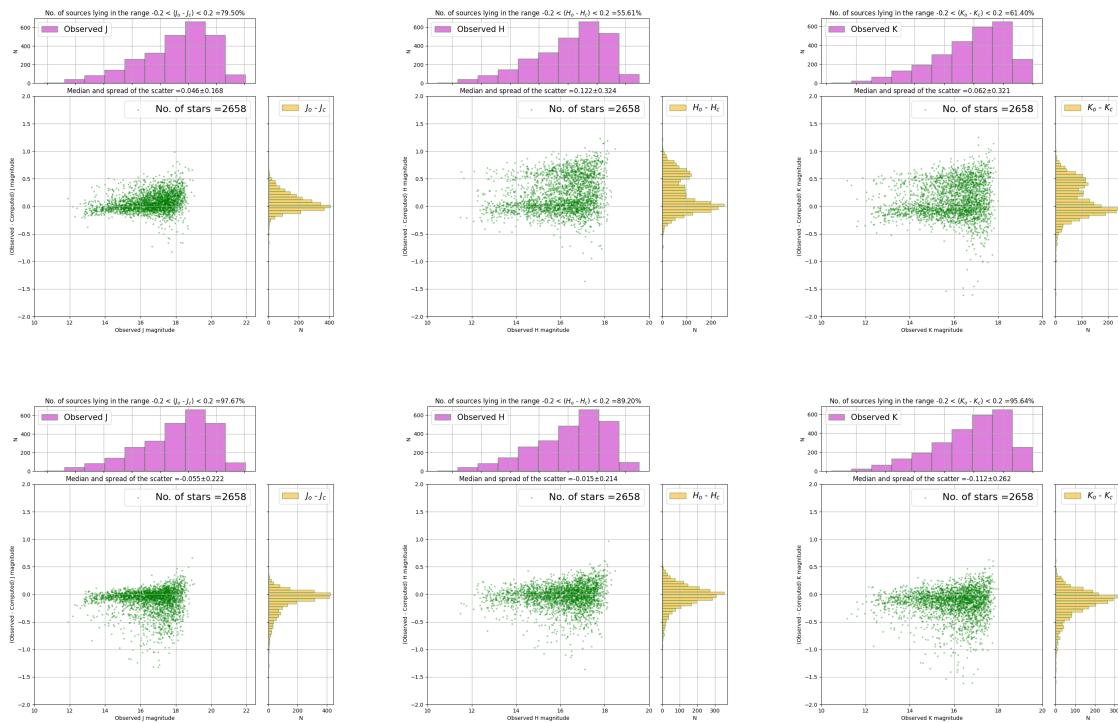


Figure 22: The plots show the comparison of the observed and computed NIR magnitudes using the stack photometry data and when no $d_{dev,min}$ is computed for each source. The plots in the **first-row** are modeled using K0 models, while the plots in the **second-row** are modeled using the K1 models.

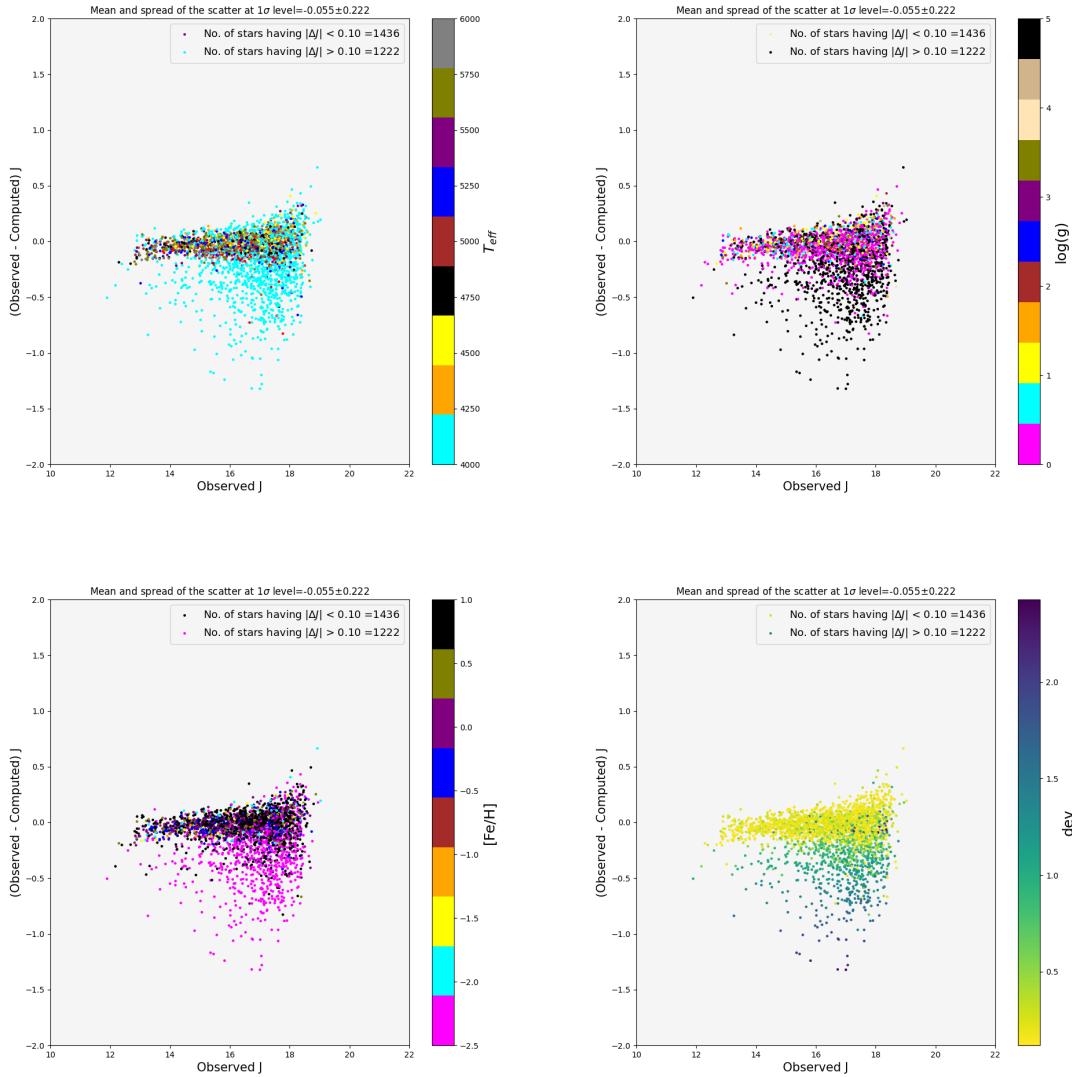


Figure 23: The plots show the comparison of the observed and computed J magnitudes when all the sources are modeled with K1 models and the scatter is color-coded according to the model parameters and the cumulative deviation of the observed color from the model color (see the second row of Figure (22)). Each plot shows the source color-coded according to the best-fitted model parameter and dev.

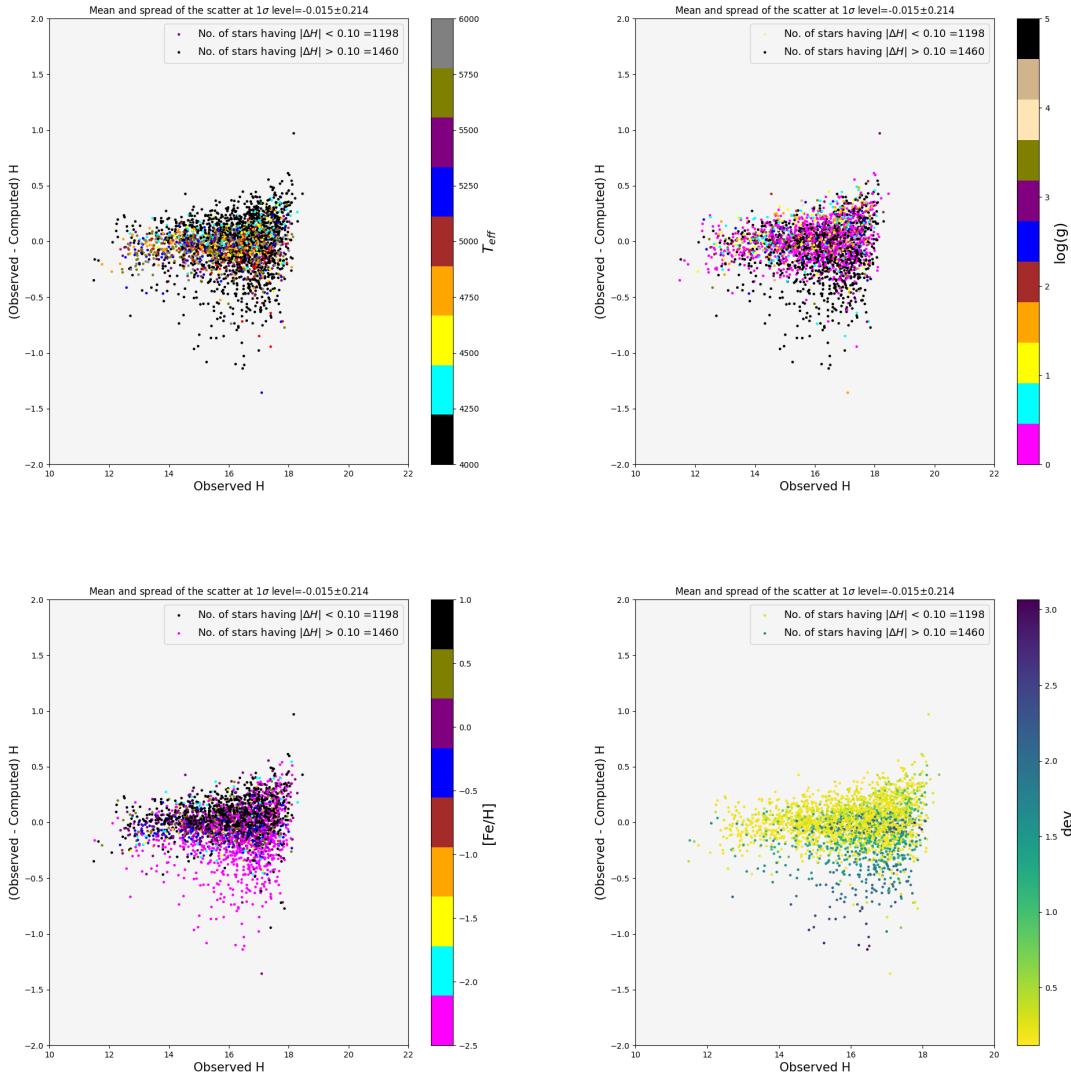


Figure 24: The plots show the comparison of the observed and computed H magnitudes when all the sources are modeled with K1 models and the scatter is color-coded according to the model parameters and the cumulative deviation of the observed color from the model color (see the second row of Figure (22)). Each plot shows the source color-coded according to the best-fitted model parameter and dev.

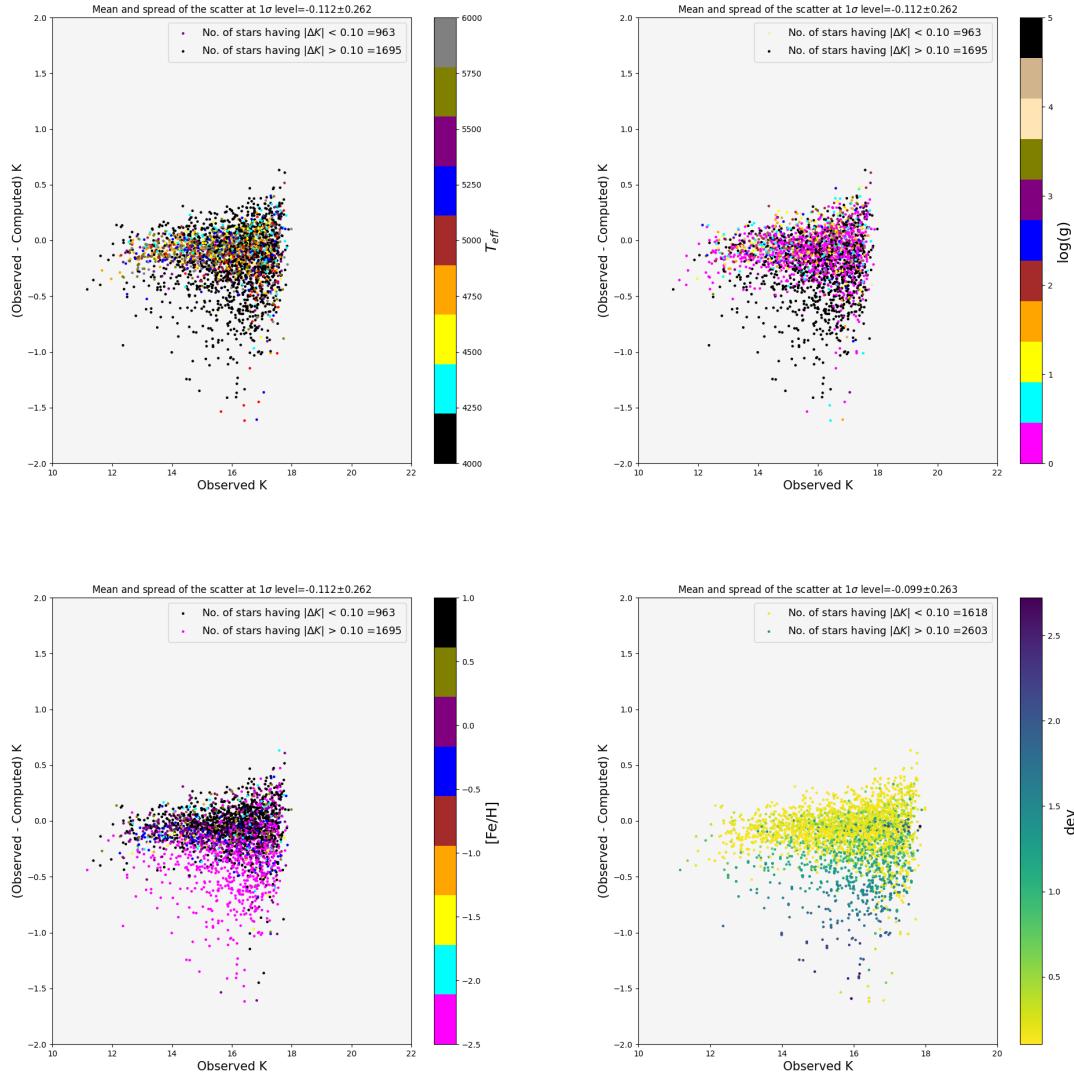


Figure 25: The plots show the comparison of the observed and computed K magnitudes when all the sources are modeled with $K1$ models and by calculating $d_{\text{dev},\text{min}}$ for each source (see the second row of Figure (22)). Each plot shows the source color-coded according to the best-fitted model parameter and d_{dev} .

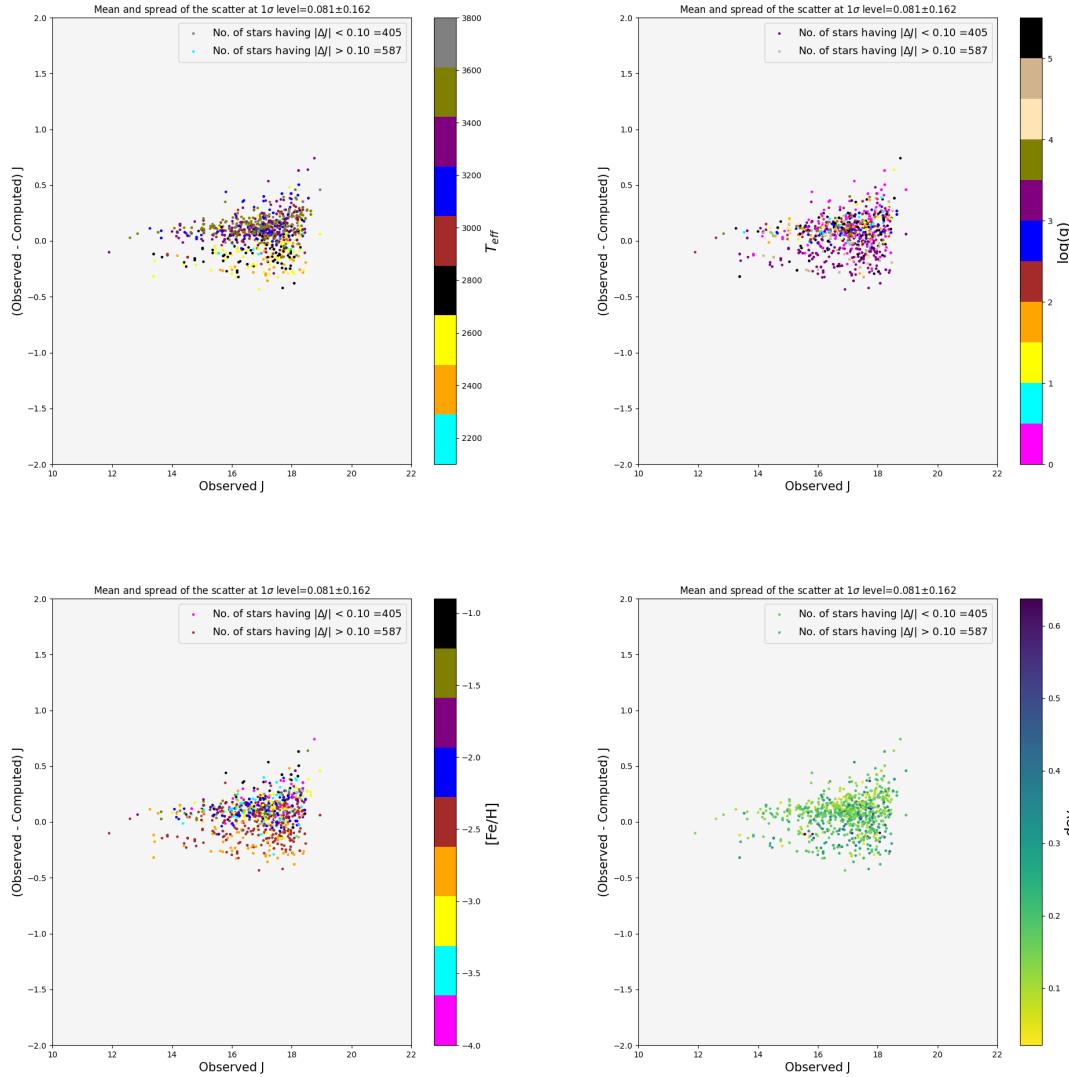


Figure 26: The plots show the comparison of the observed and computed J magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using P0 models.

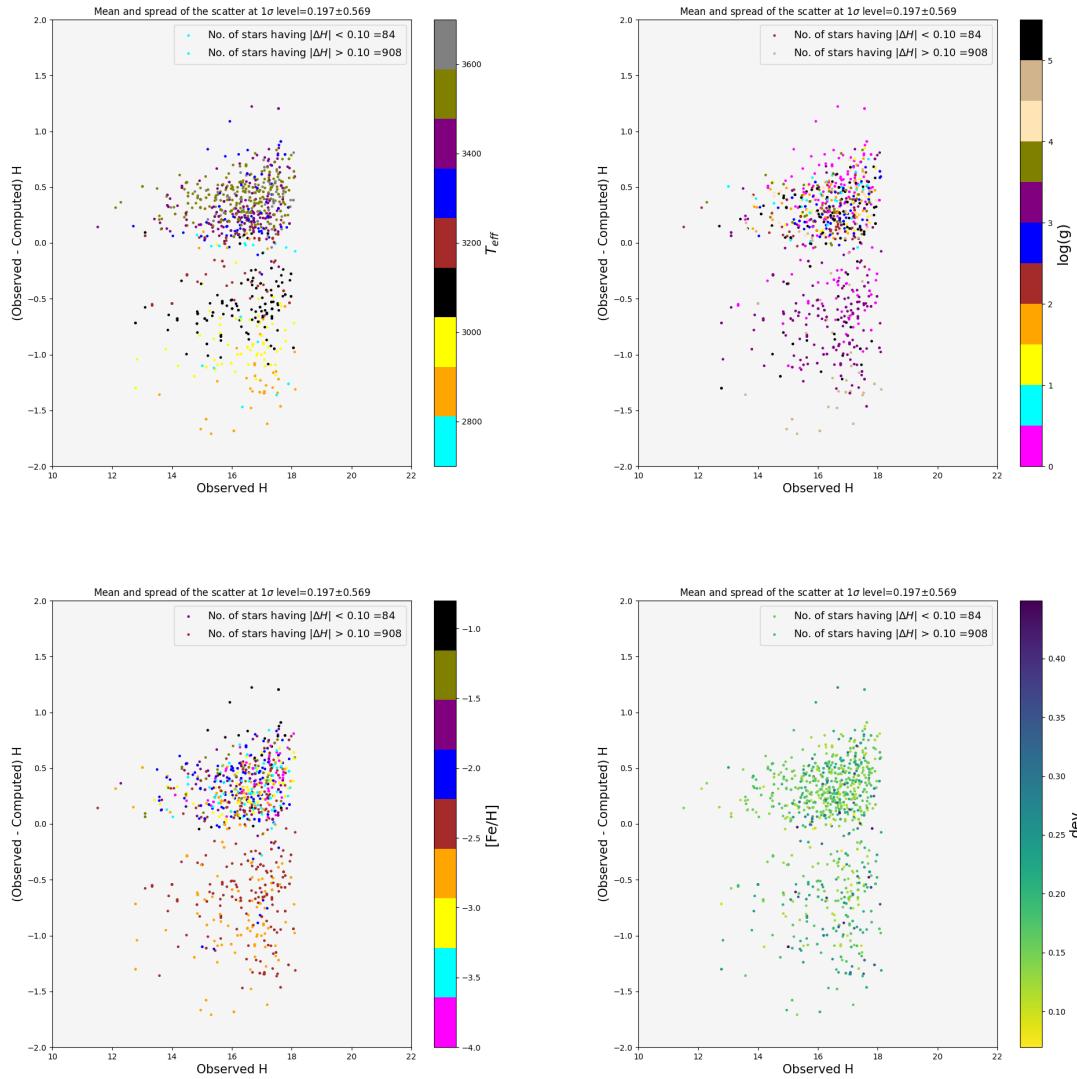


Figure 27: The plots show the comparison of the observed and computed H magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using P0 models.

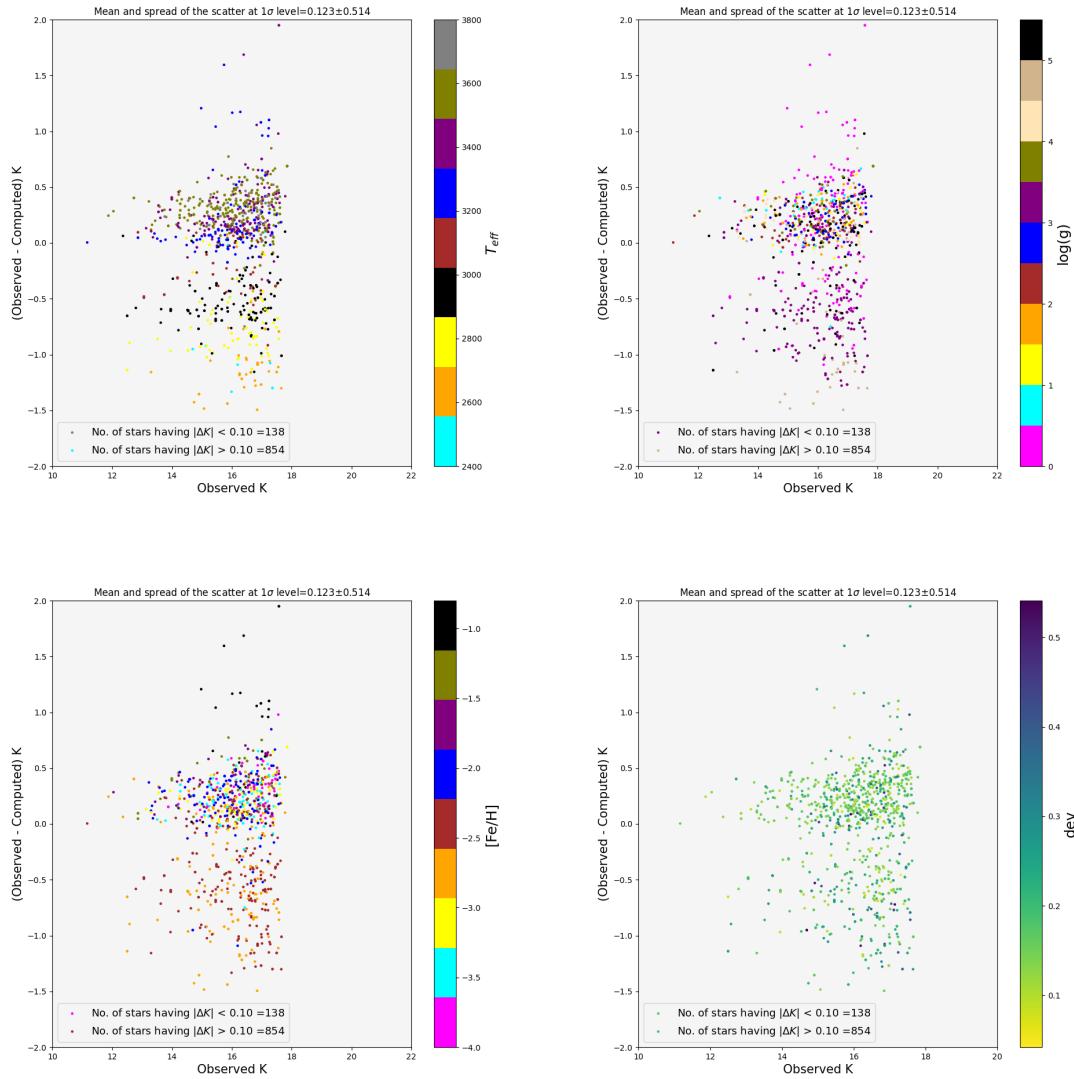


Figure 28: The plots show the comparison of the observed and computed K magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using P0 models.

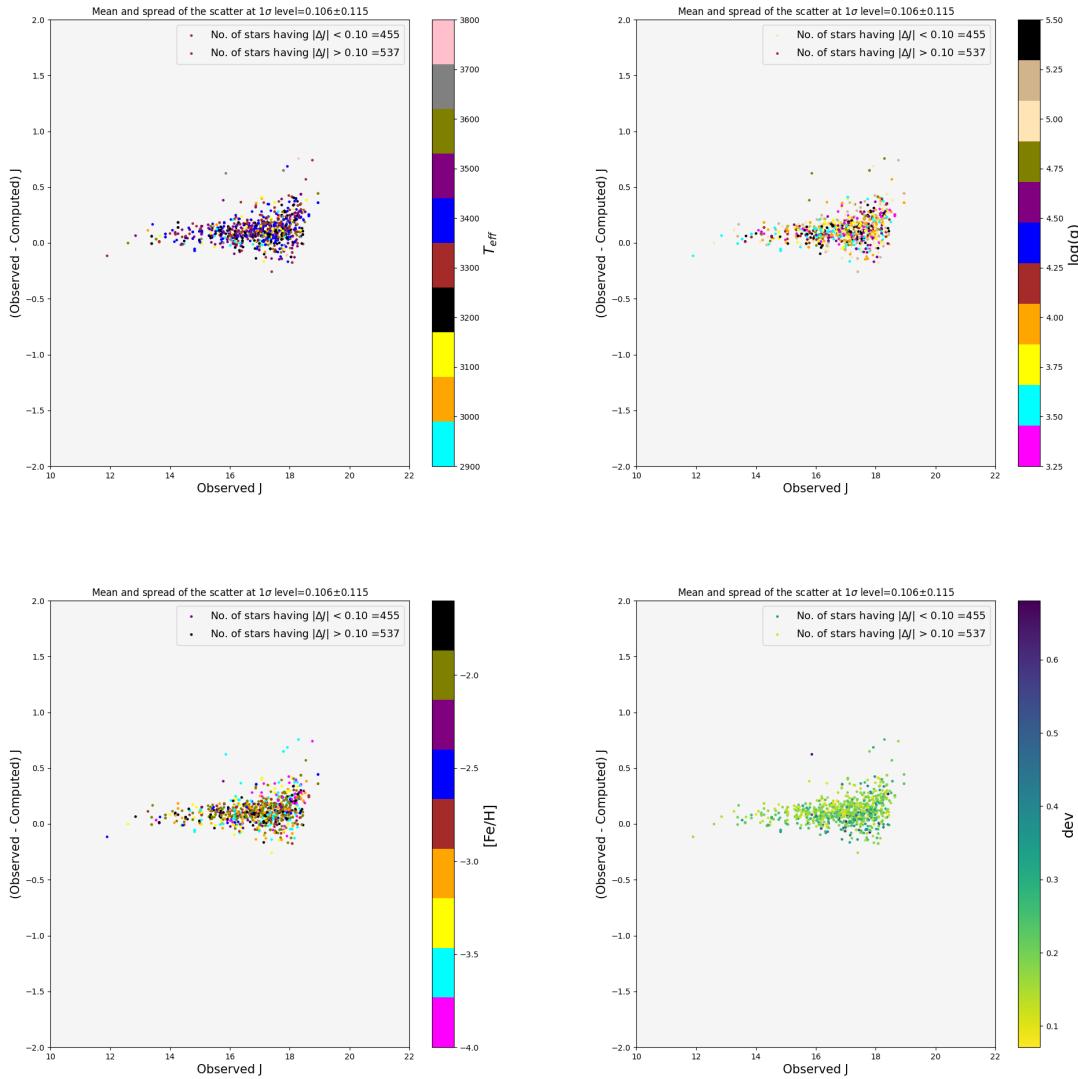


Figure 29: The plots show the comparison of the observed and computed J magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using C1 and C2 models.

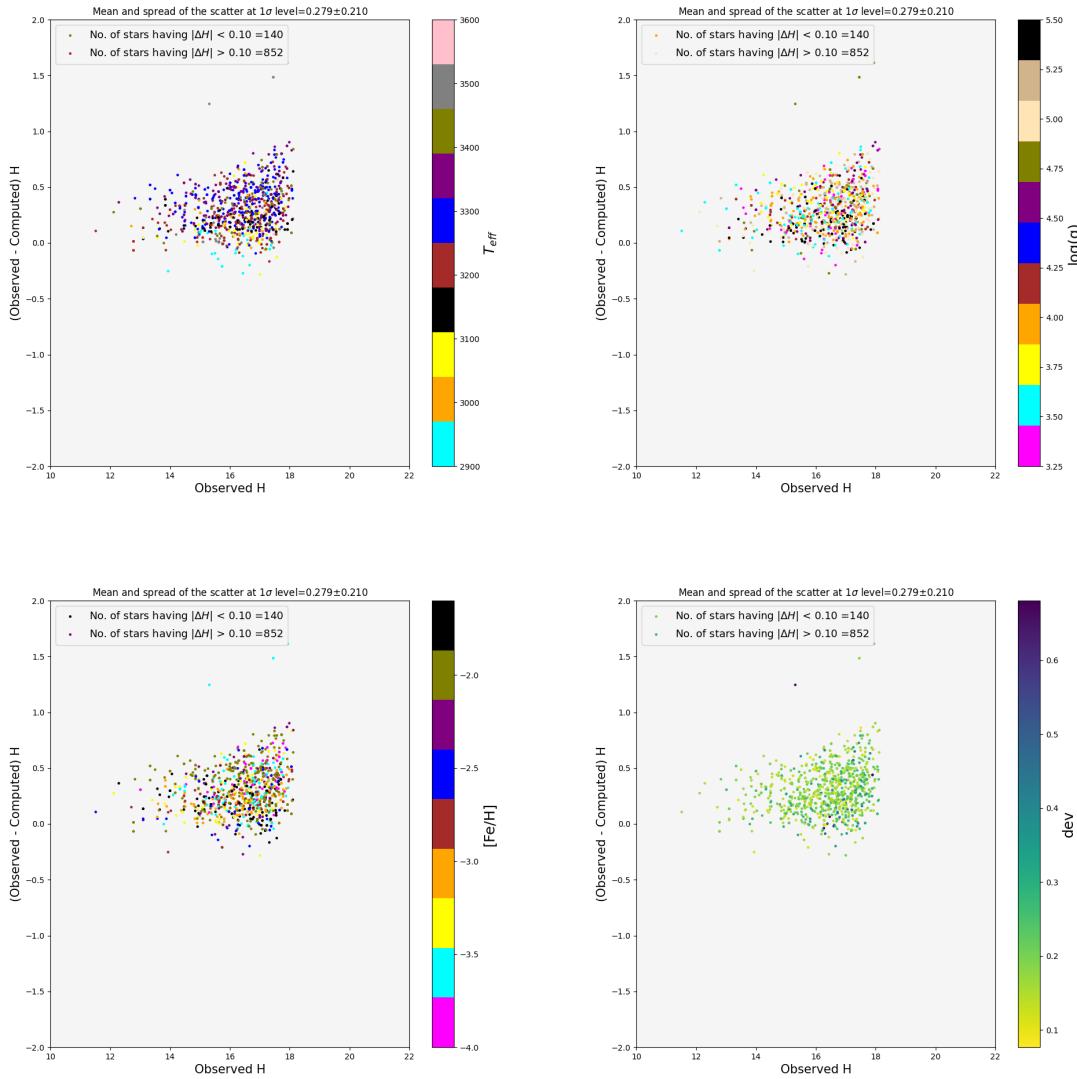


Figure 30: The plots show the comparison of the observed and computed H magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using C1 and C2 models.

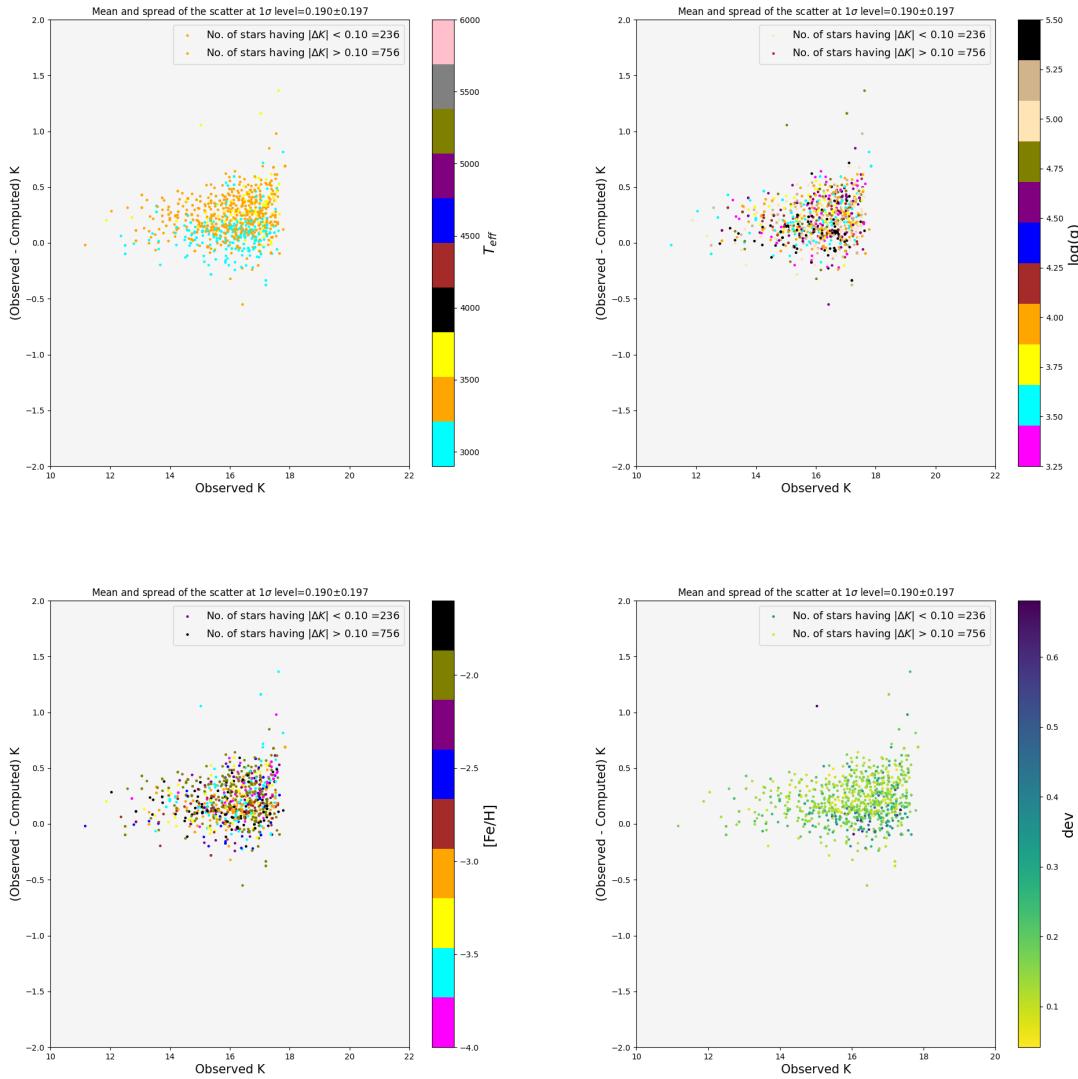


Figure 31: The plots show the comparison of the observed and computed K magnitudes when the scattered sources in the second row of Figure (22) are re-modeled using C1 and C2 models.

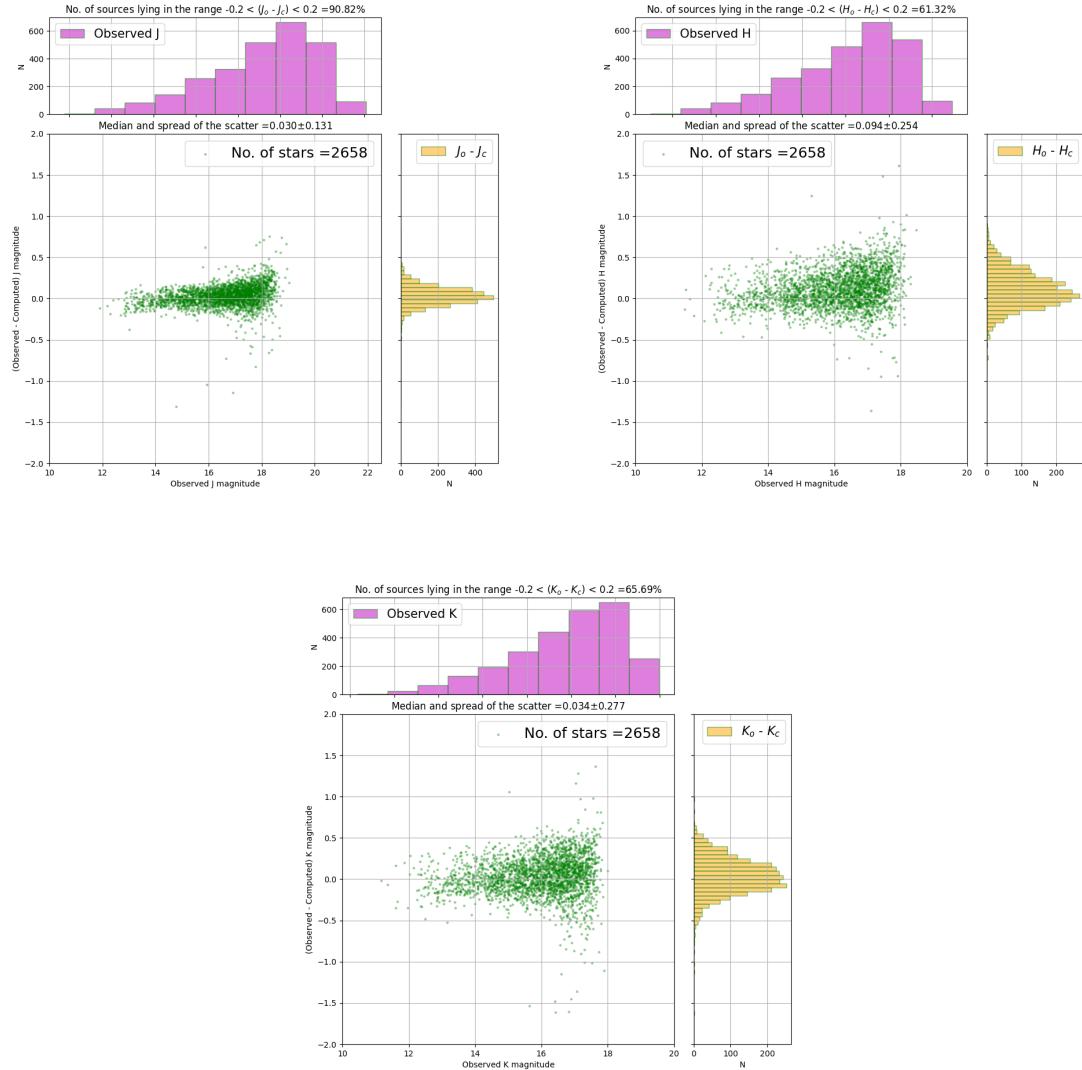


Figure 32: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry. There is a significant improvement in the accuracy of the computed NIR magnitudes (see the titles of each subplot). The outliers are possibly those objects which are faint and cool.

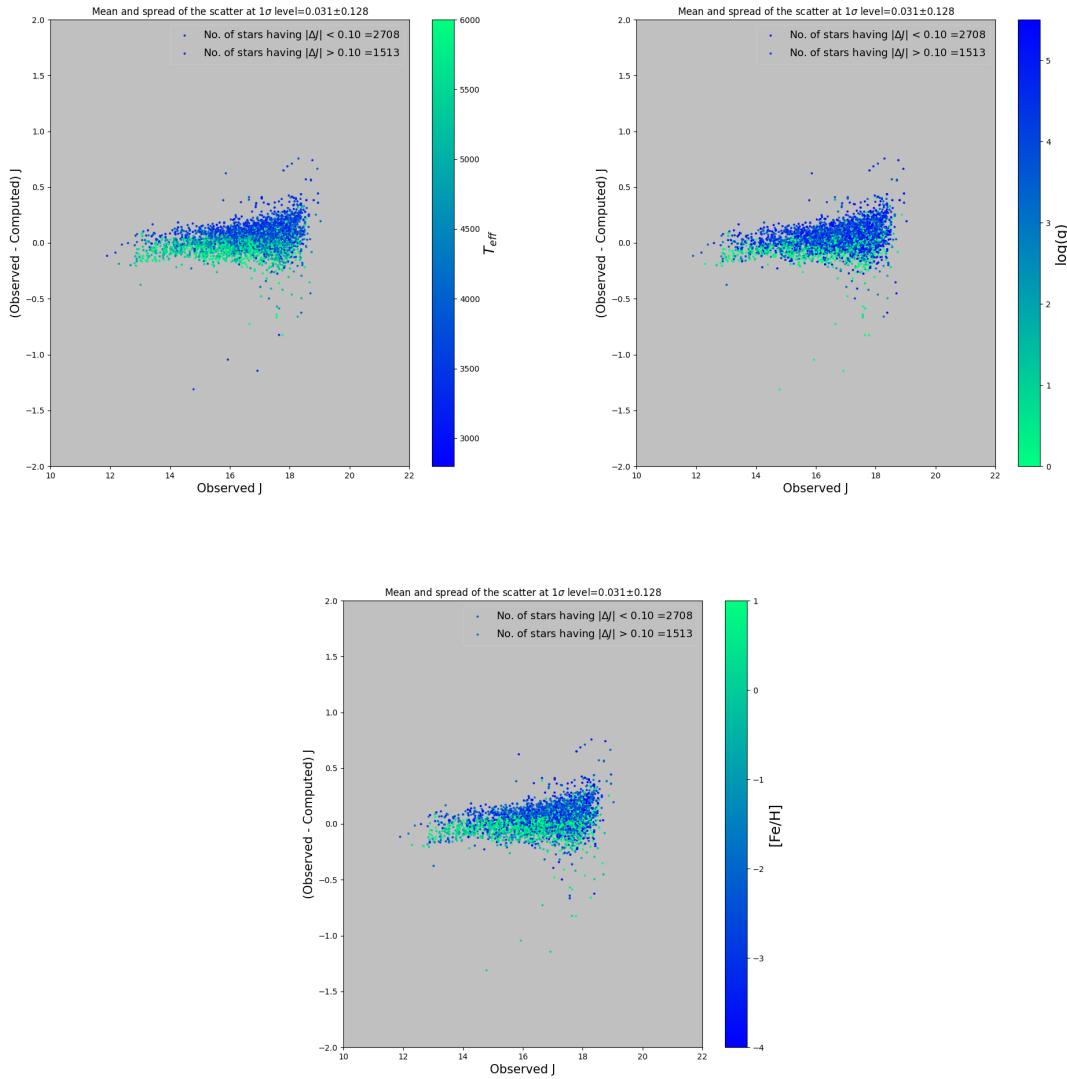


Figure 33: The plots show the properties of the sources in the plot showing the comparison between observed and computed J magnitudes of the sources.

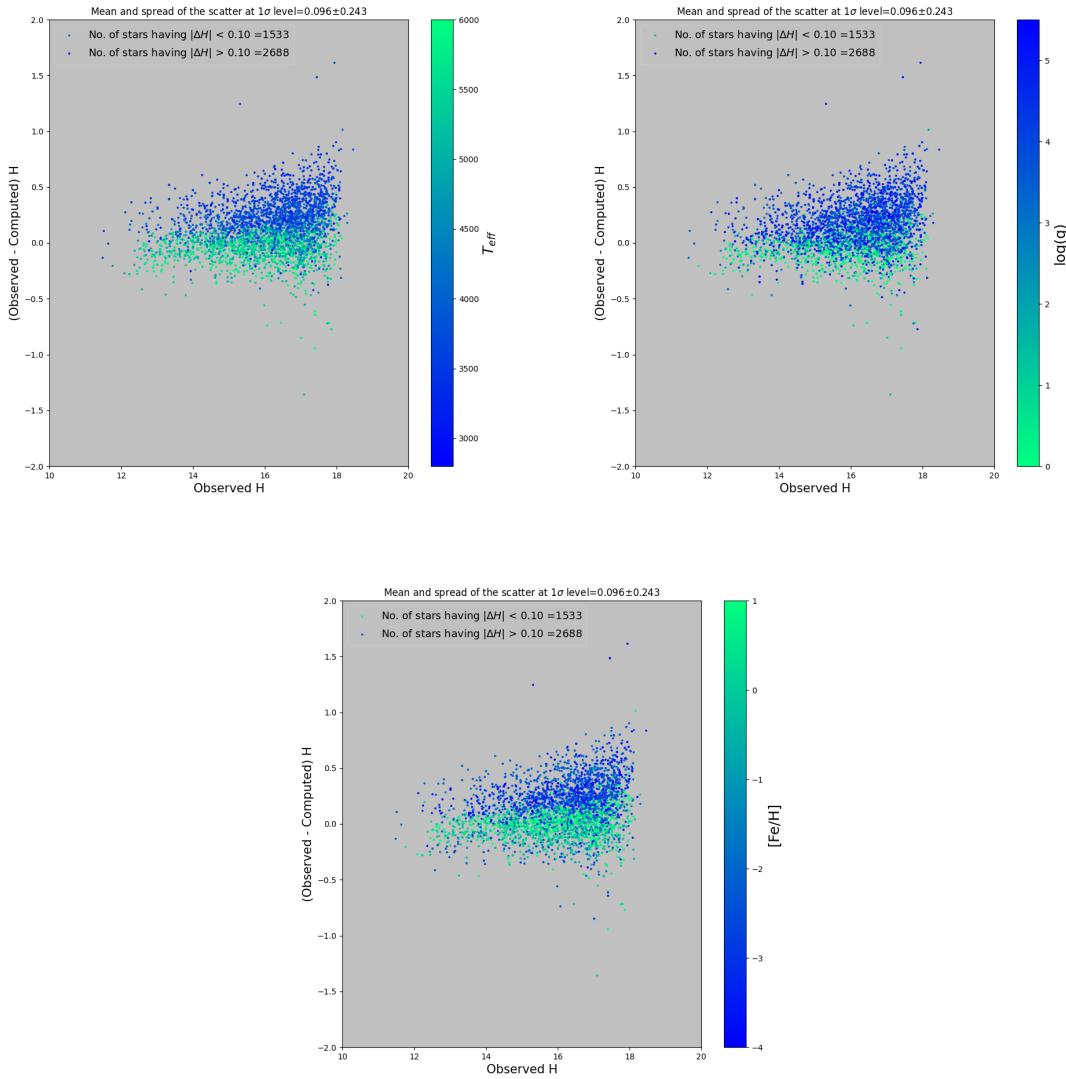


Figure 34: The plots show the properties of the sources in the plot showing the comparison between observed and computed H magnitudes of the sources.

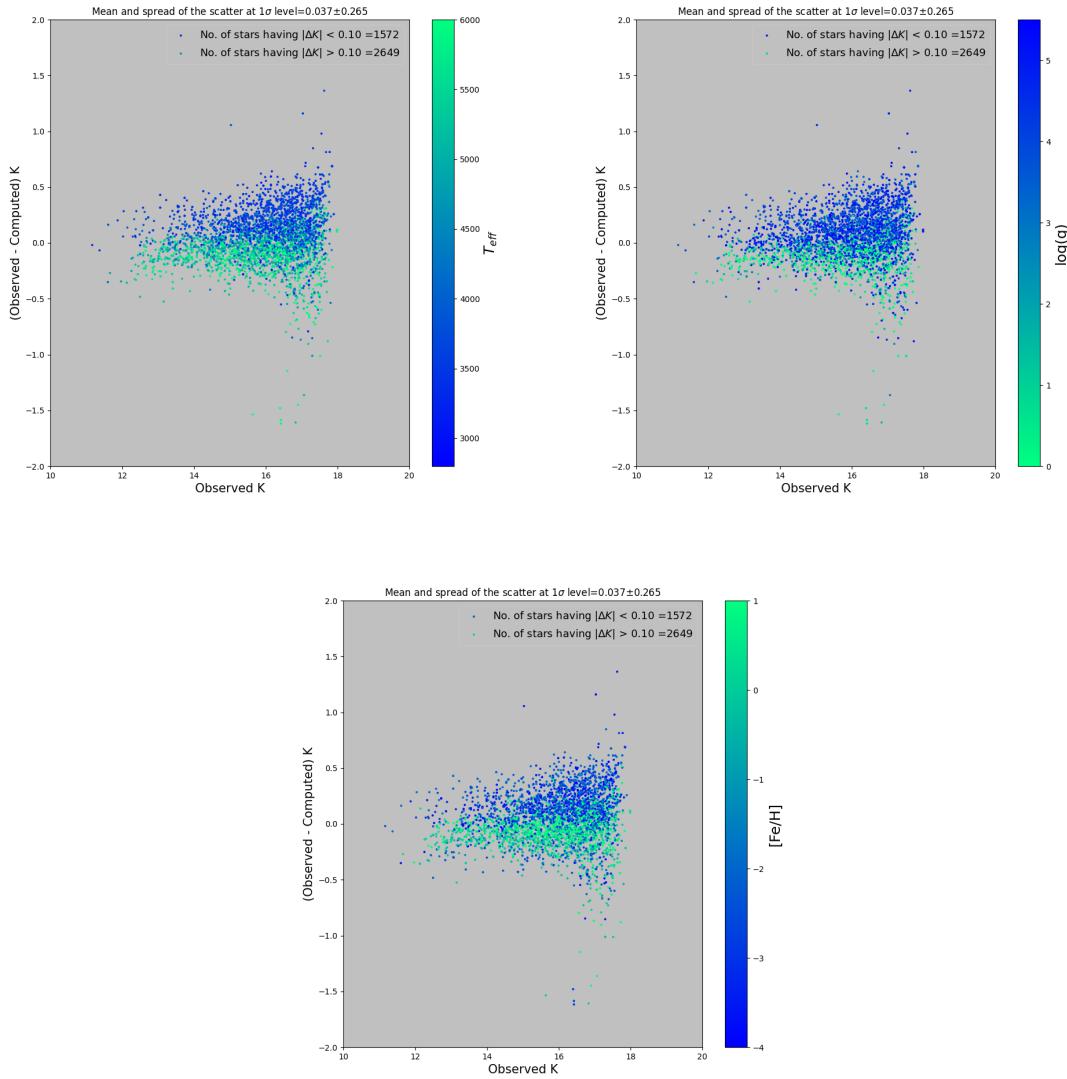


Figure 35: The plots show the properties of the sources in the plot showing the comparison between observed and computed H magnitudes of the sources.

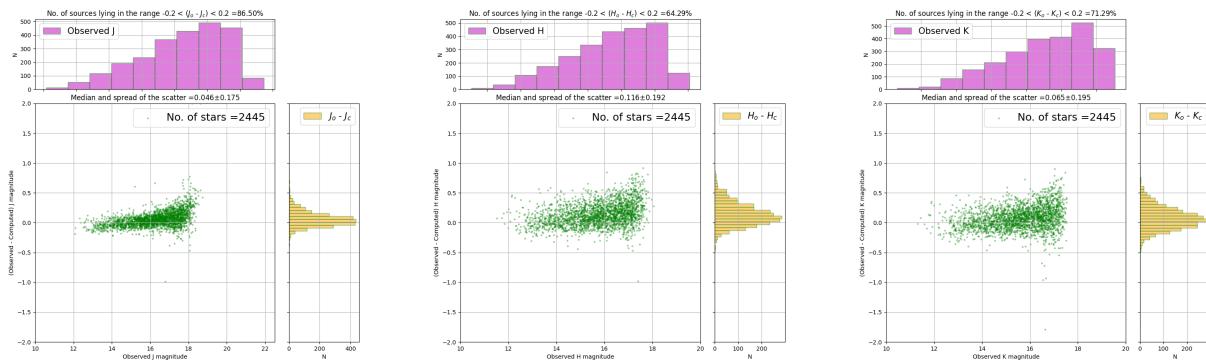


Figure 36: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF2 field.

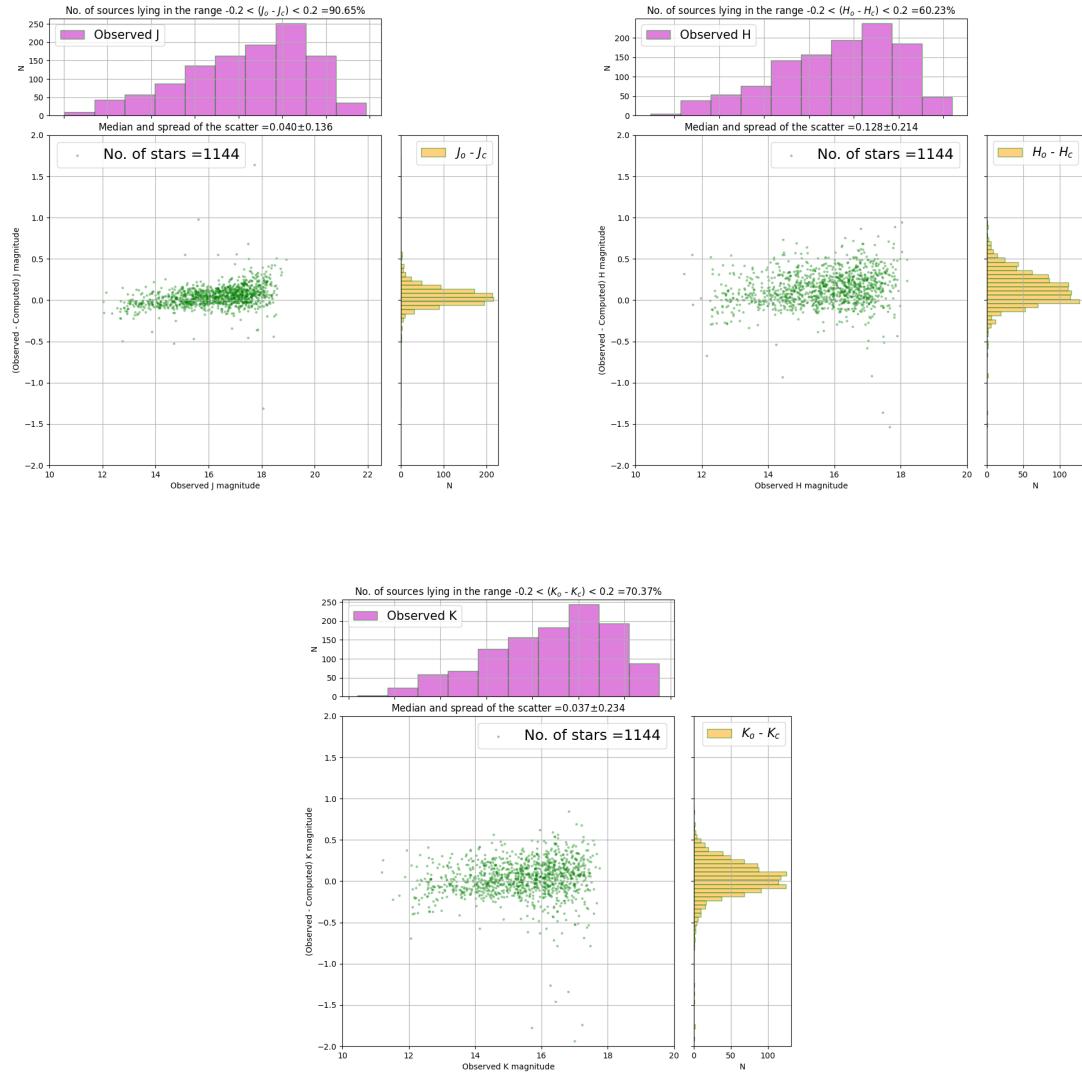


Figure 37: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF3 field.

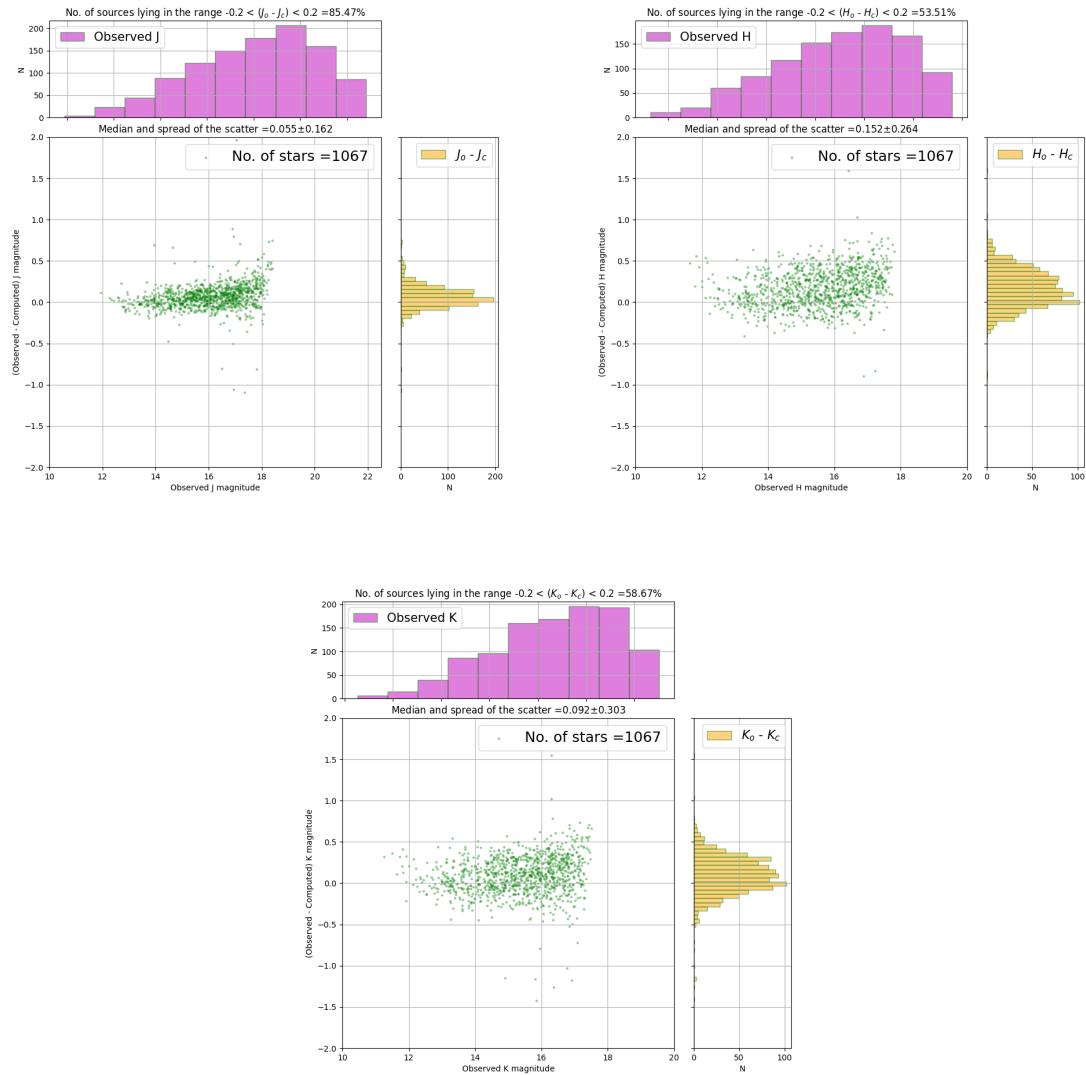


Figure 38: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF4 field.

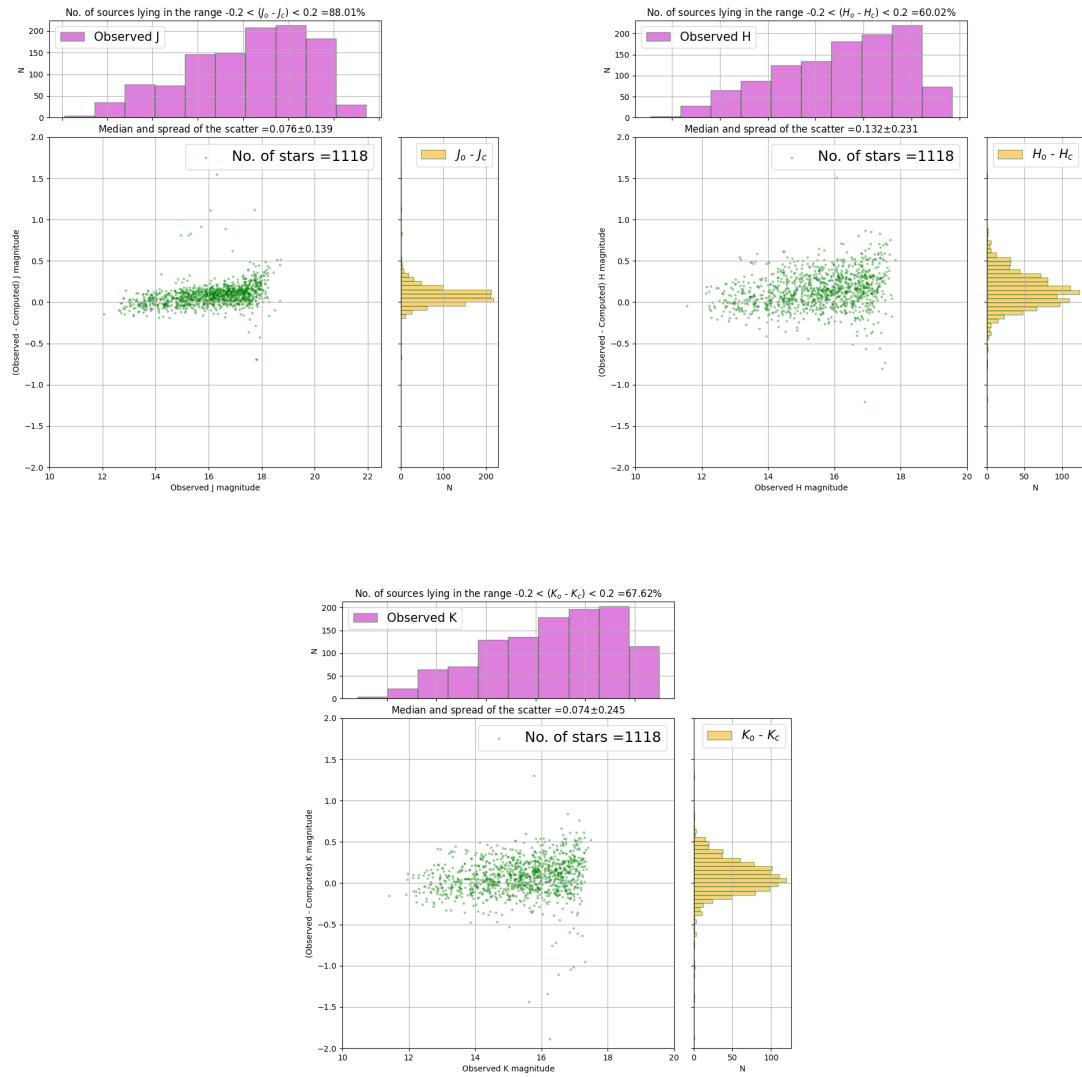


Figure 39: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF5 field.

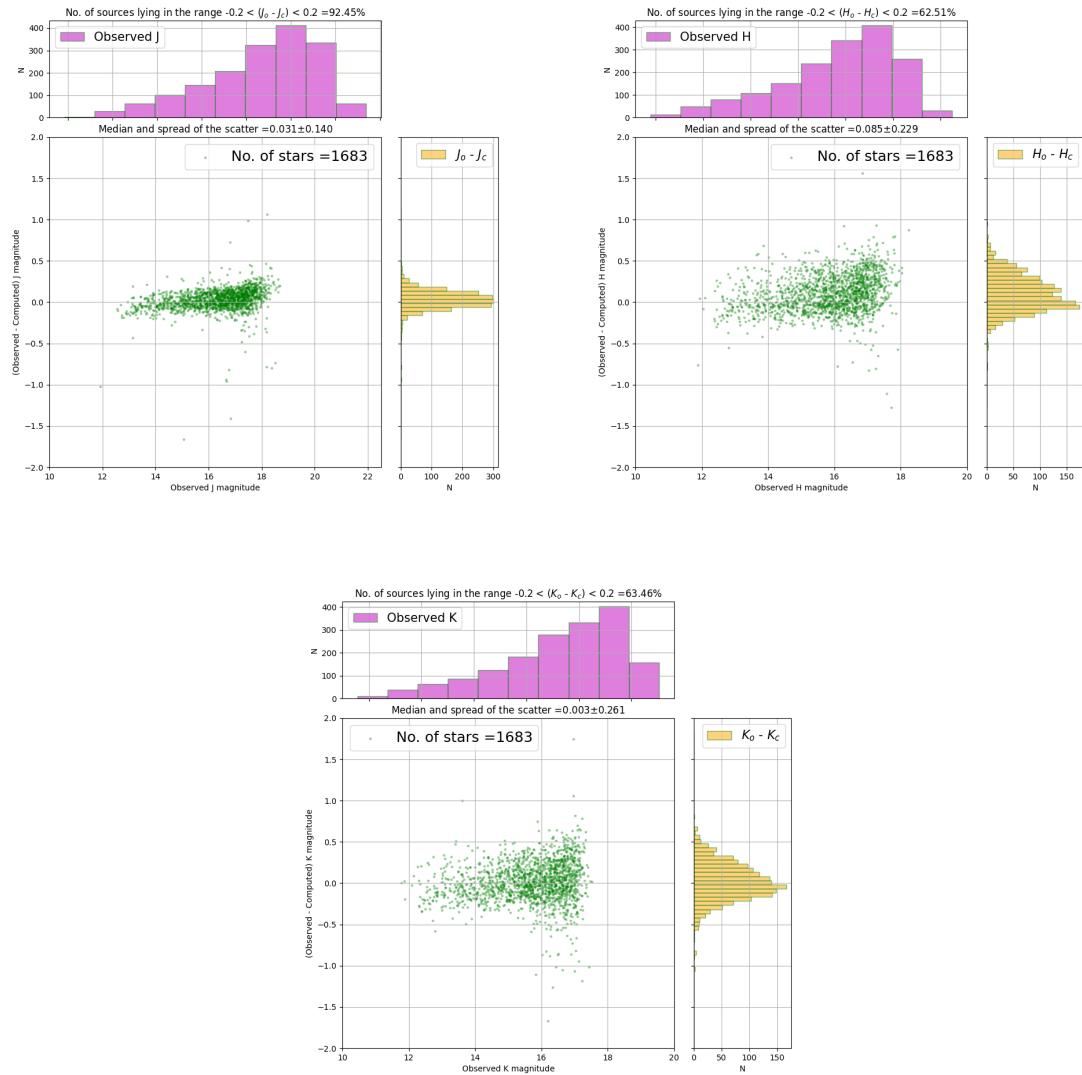


Figure 40: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF6 field.

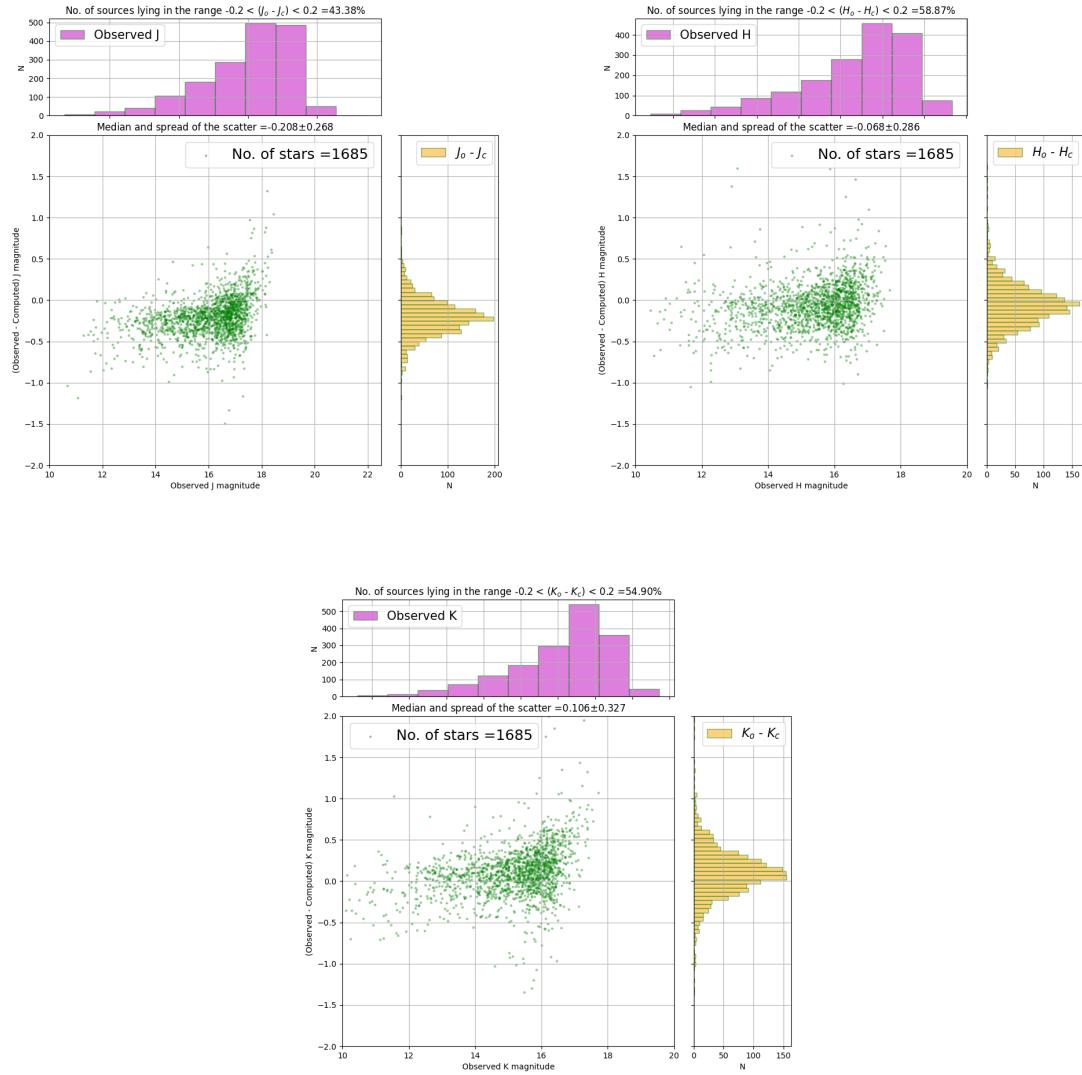


Figure 41: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF7 field.

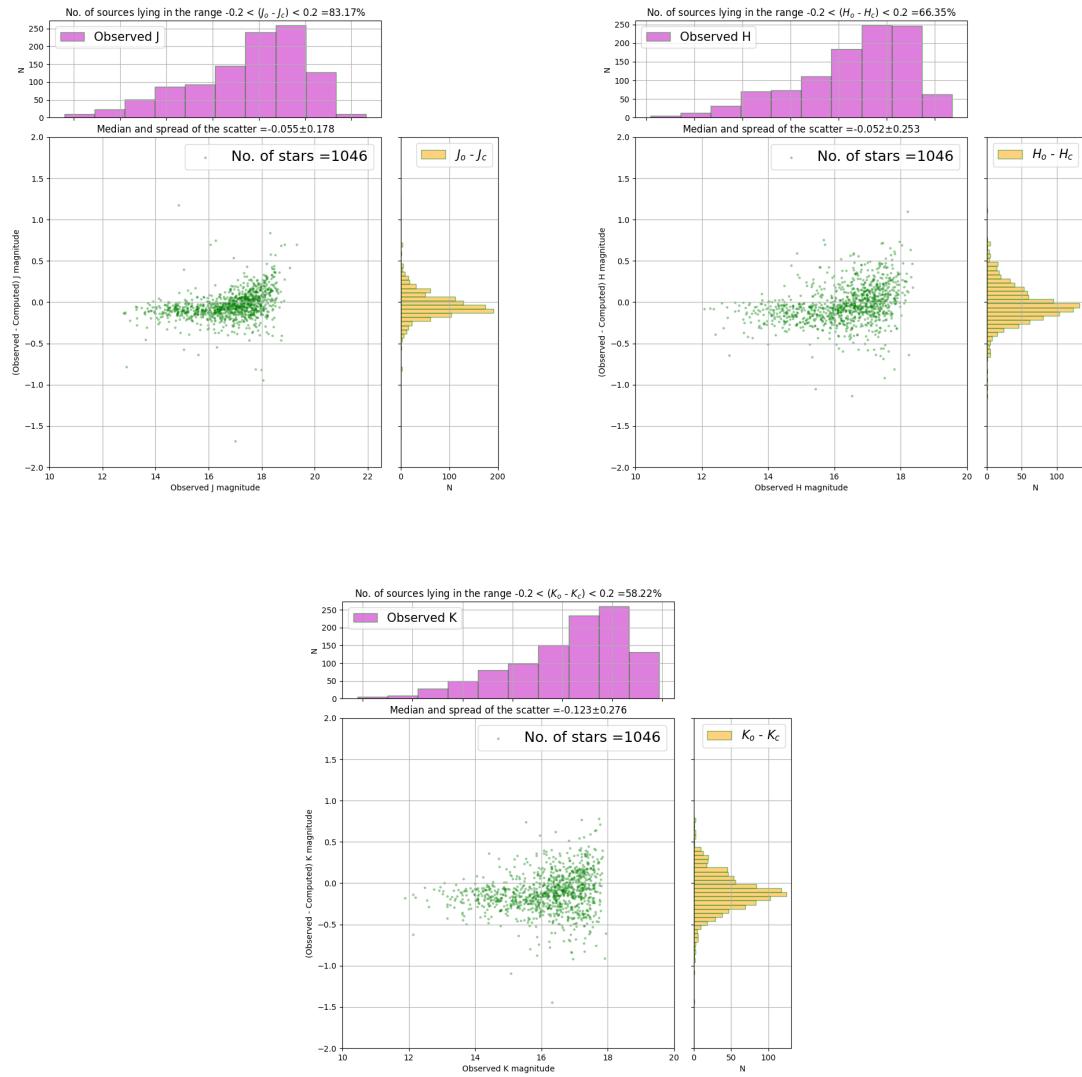


Figure 42: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF8 field.

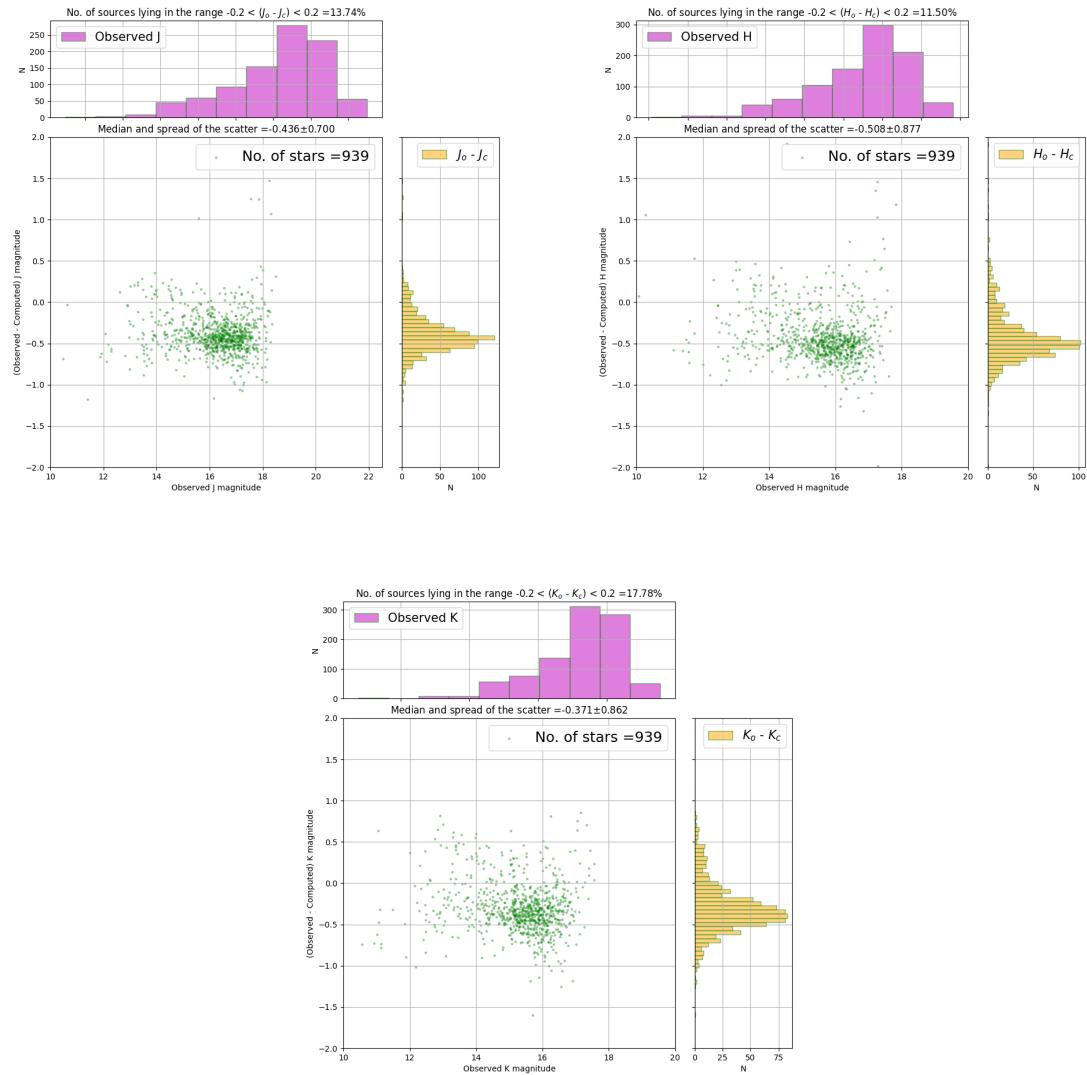


Figure 43: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF9 field.

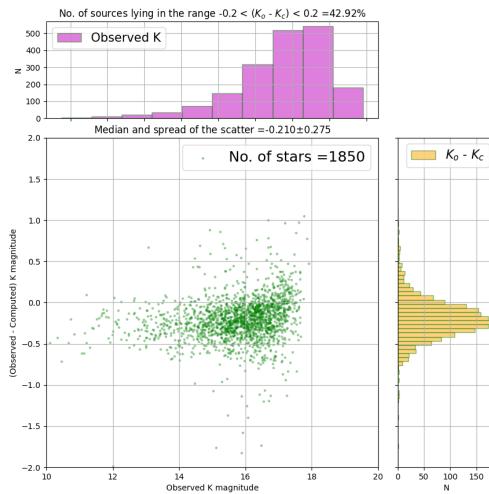


Figure 44: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF10 field. There is no J and H band data from UKIDSS for this test field.

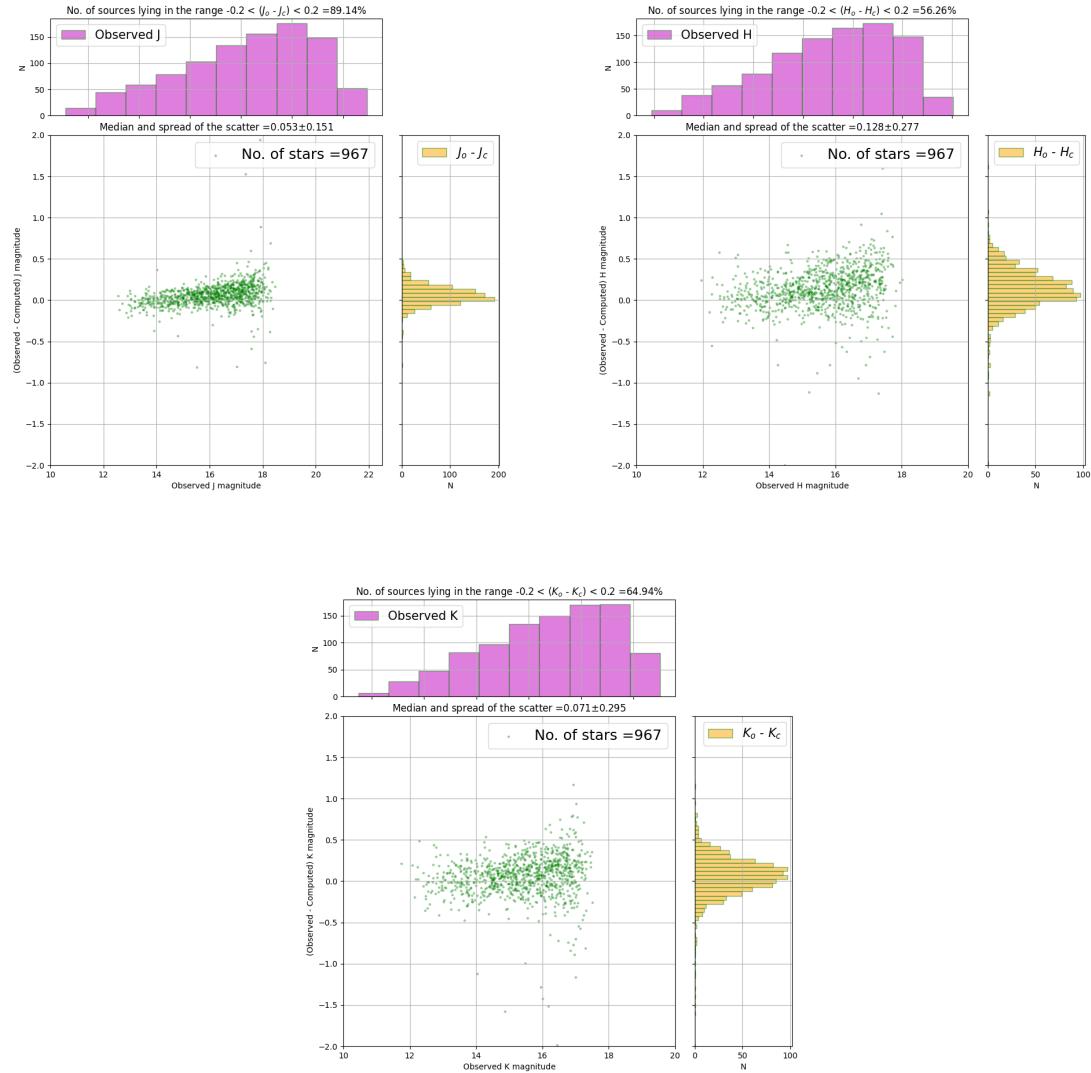


Figure 45: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF11 field.

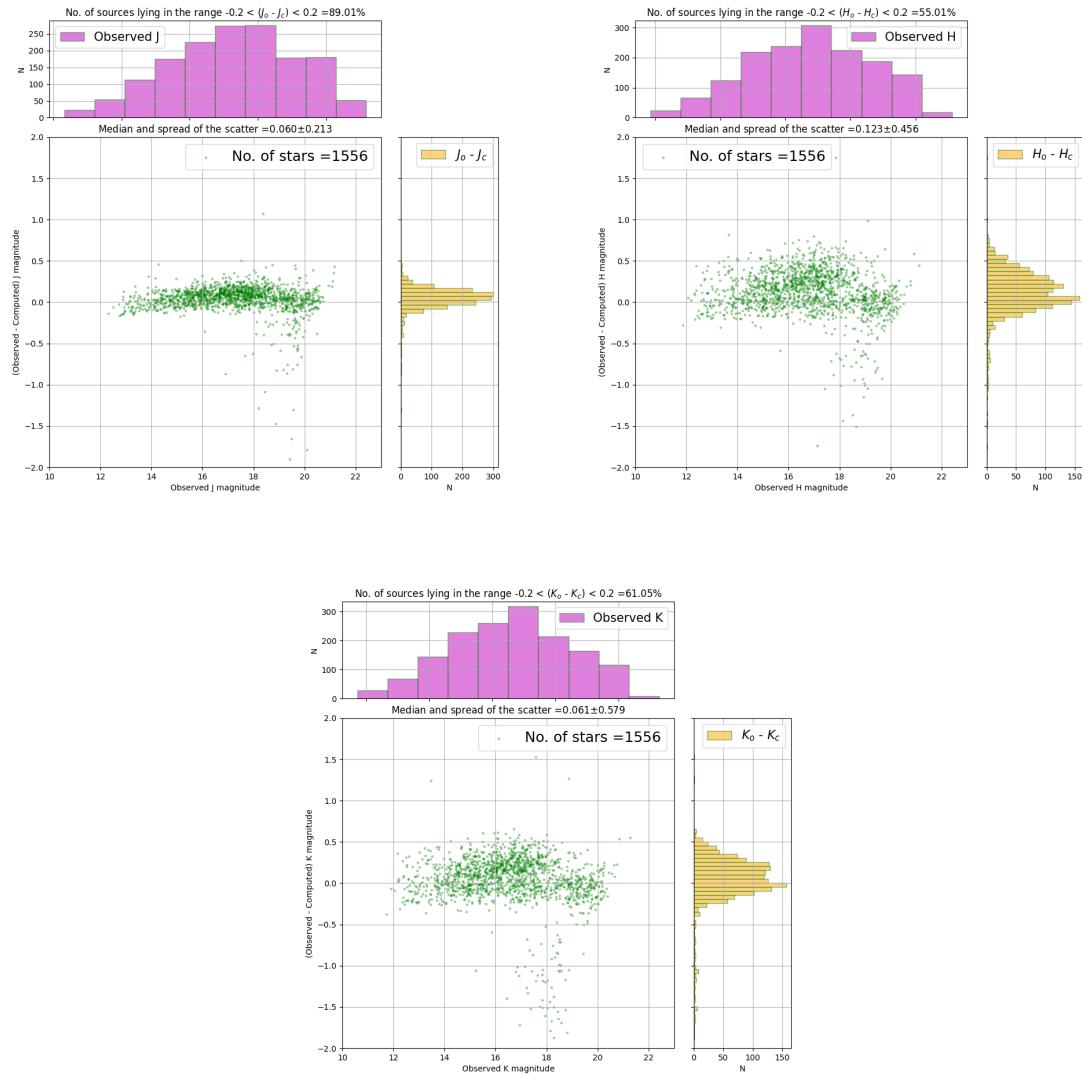


Figure 46: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF12 field.

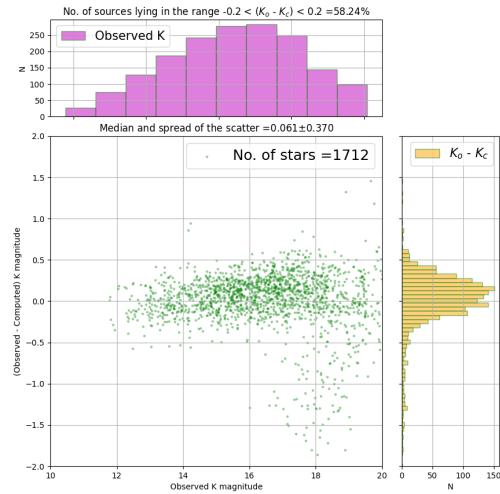


Figure 47: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF13 field. There is no J and H band data from UKIDSS for this test field.

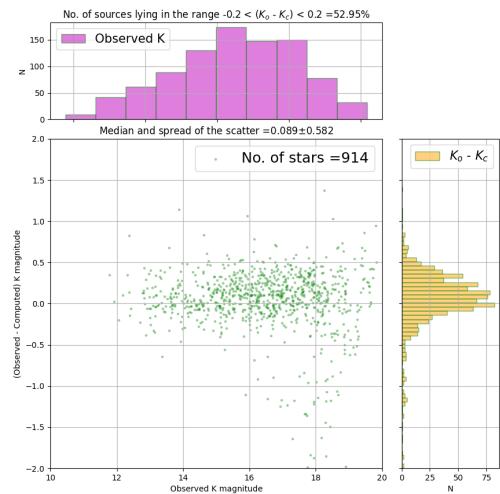


Figure 48: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF14 field. There is no J and H band data from UKIDSS for this test field.

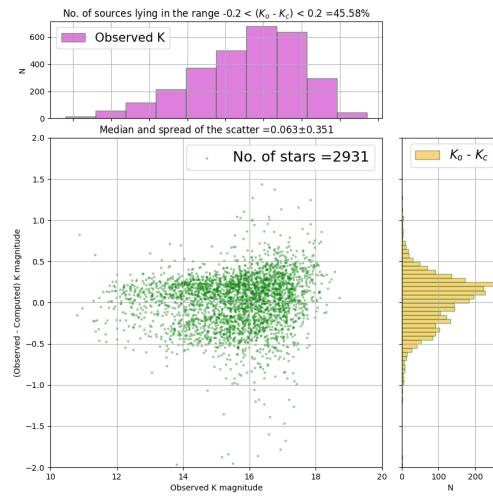


Figure 49: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF15 field. There is no J and H band data from UKIDSS for this test field.

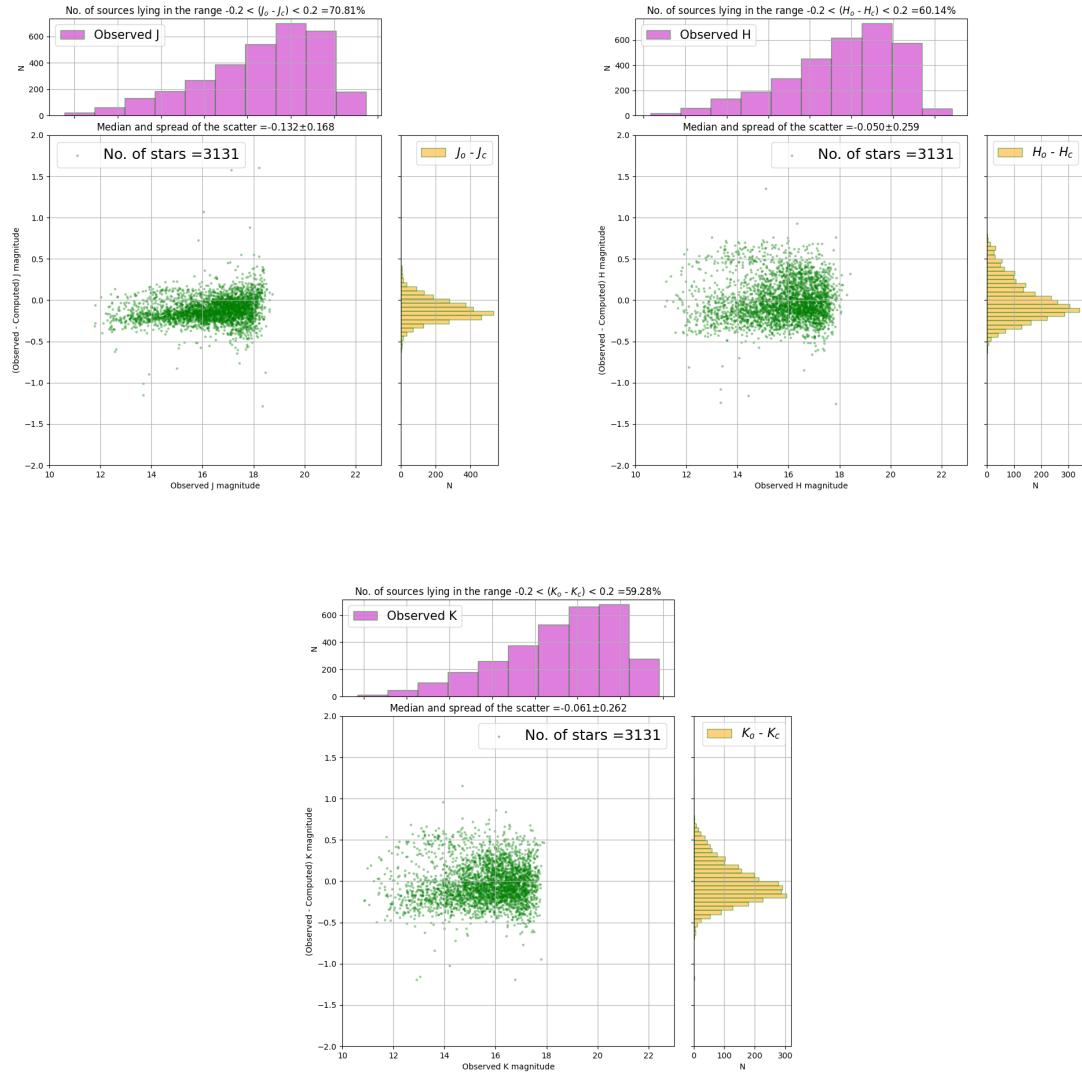


Figure 50: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF16 field.

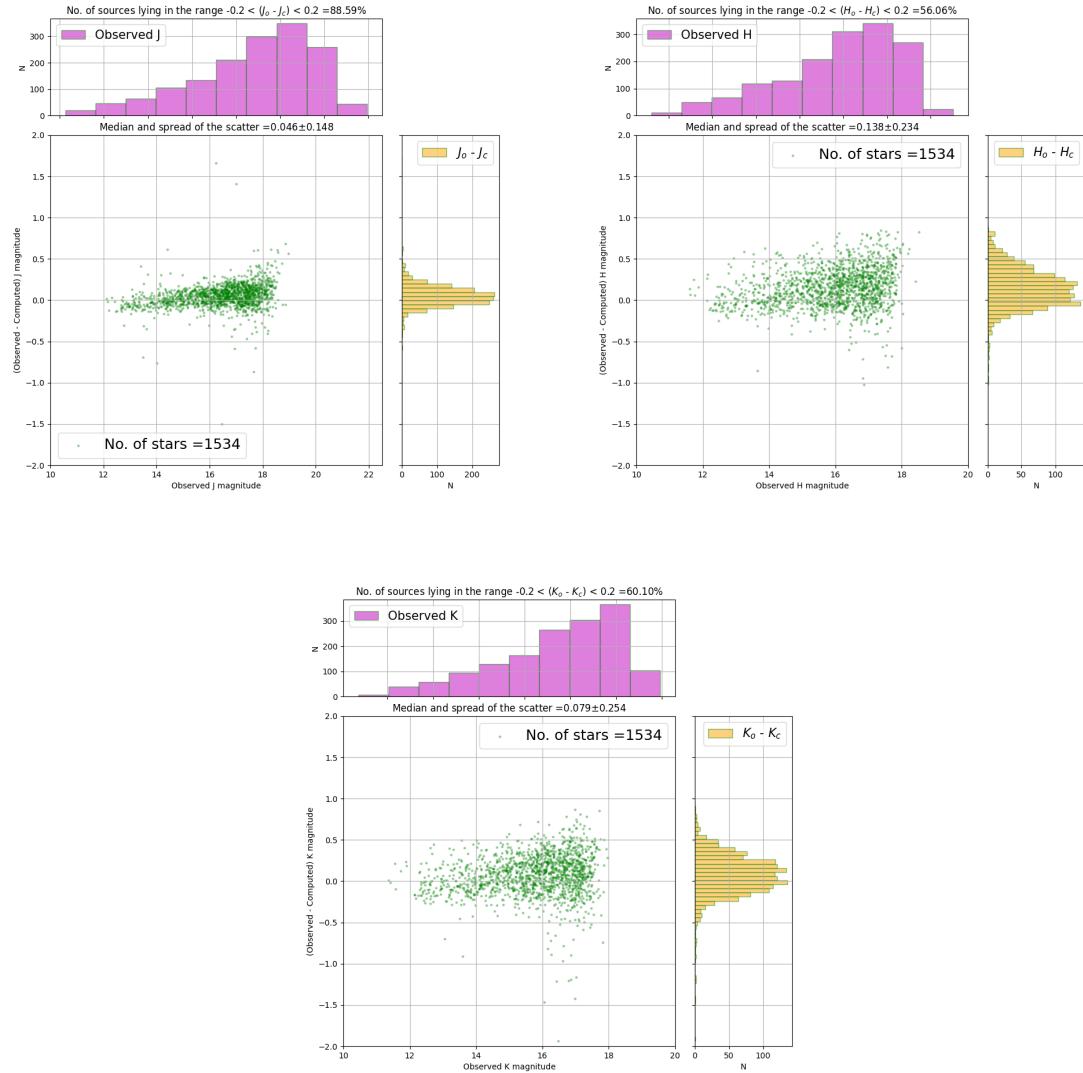


Figure 51: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF17 field.

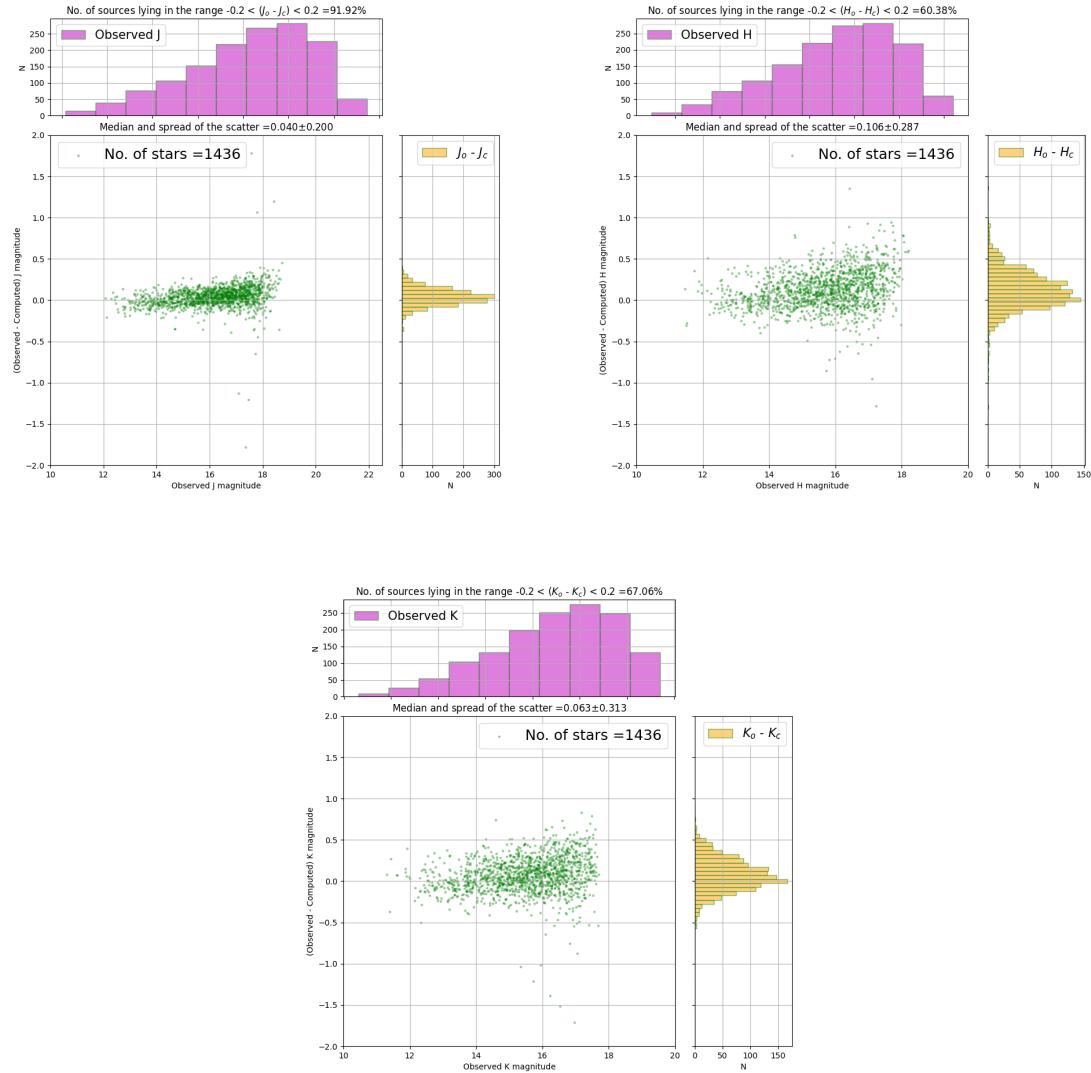


Figure 52: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF18 field.

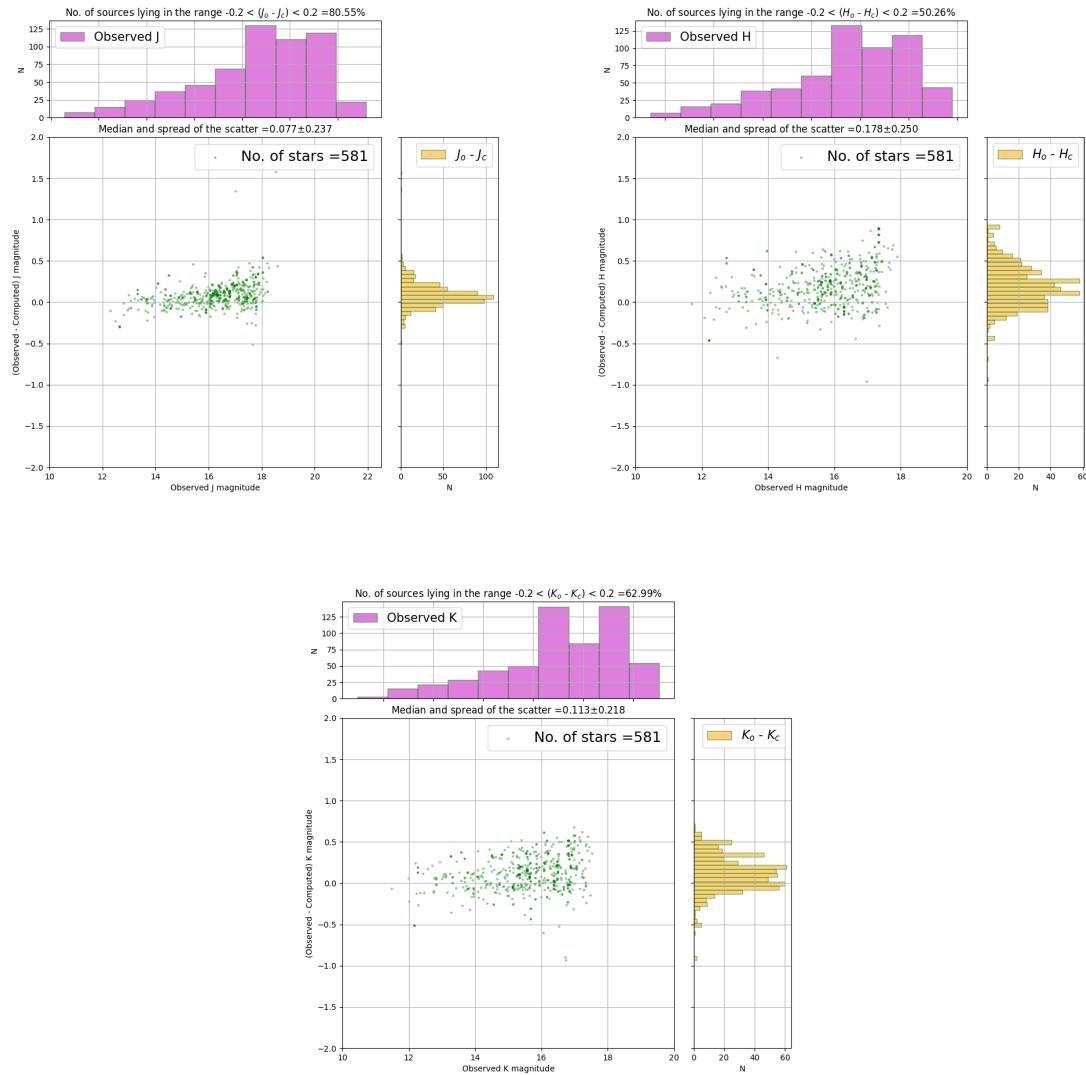


Figure 53: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF19 field.

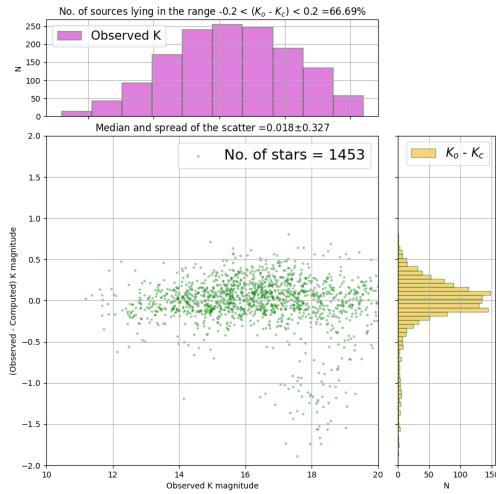


Figure 54: The plots show the comparison of the computed and the observed NIR magnitudes when the optimal method is applied on stack photometry for the probable stellar sources in the TF20 field. There is no J and H band data from UKIDSS for this test field.

Field	$N_{sources}$	N_{stars}	N_{galaxy}	N_{Cat}	N_{UKIDSS}	i_{comp}	J_{comp}	J_{faint}	Density	Δ_J	Δ_H	Δ_K
TF1	14064	6006	8058	6006	6441	20.94	19.62	22.38	4.94	(90.82%) 0.030 ± 0.131	(61.32%) 0.094 ± 0.254	(65.69%) 0.034 ± 0.277
TF2	11104	5310	5794	5310	4443	20.66	19.31	22.23	4.36	(86.50%) 0.046 ± 0.175	(64.29%) 0.116 ± 0.192	(71.29%) 0.065 ± 0.195
TF3	6392	2916	3476	2916	3298	20.36	18.88	21.48	2.40	(90.65%) 0.051 ± 0.185	(60.23%) 0.128 ± 0.214	(70.37%) 0.037 ± 0.234
TF4	8790	2408	6382	2408	2840	20.79	19.74	22.21	1.98	(85.47%) 0.055 ± 0.162	(53.51%) 0.152 ± 0.264	(60.32%) 0.092 ± 0.303
TF5	8729	2588	6141	2588	2763	20.67	19.33	22.10	2.12	(88.86%) 0.071 ± 0.138	(59.66%) 0.134 ± 0.245	(67.05%) 0.069 ± 0.274
TF6	11180	4601	6579	4601	4121	20.87	19.63	21.98	3.78	(92.45%) 0.031 ± 0.140	(62.51%) 0.085 ± 0.229	(63.46%) 0.003 ± 0.261
TF7*	2400	1700	700	1700	4775	19.47	17.49	20.26	68.05	(43.38%) -0.208 ± 0.268	(58.81%) -0.068 ± 0.286	(54.90%) 0.106 ± 0.327
TF8*	1508	1281	227	1281	1828	18.65	19.99	20.58	51.25	(66.25%) -0.052 ± 0.253	(58.22%) -0.123 ± 0.276	(58.22%) -0.123 ± 0.276
TF9*	1149	942	207	942	3103	19.82	17.99	19.21	37.69	(13.74%) -0.036 ± 0.178	(66.35%) -0.050 ± 0.253	(17.78%) -0.371 ± 0.862
TF10*	2370	1910	439	1911	5921	19.68	18.01	19.07	76.45	-	(42.92%) -0.210 ± 0.275	(42.92%) -0.210 ± 0.275
TF11	6674	2125	4549	2125	4199	20.75	20.75	21.47	1.74	(89.14%) 0.053 ± 0.151	(56.26%) 0.128 ± 0.277	(64.94%) 0.071 ± 0.295
TF12	11816	3836	7980	3836	140204	20.85	19.73	22.07	3.15	(89.01%) 0.060 ± 0.213	(53.01%) 0.123 ± 0.456	(61.05%) 0.061 ± 0.579
TF13	8193	2369	5824	1762	41694	20.82	19.71	21.79	1.94	-	(58.24%) 0.061 ± 0.570	(58.24%) 0.061 ± 0.570
TF14	7854	2187	5667	2187	2073	19.35	19.35	21.73	1.80	-	(52.95%) 0.089 ± 0.582	(52.95%) 0.089 ± 0.582
TF15	9193	5556	3637	5556	12301	20.40	18.70	21.36	4.57	-	(44.09%) 0.054 ± 0.373	(44.09%) 0.054 ± 0.373
TF16	12985	8800	4185	8800	8259	20.39	18.96	22.27	7.30	(70.81%) -0.132 ± 0.168	(60.14%) -0.050 ± 0.259	(59.28%) -0.061 ± 0.262
TF17	10995	3870	7125	3870	6120	21.17	20.05	22.68	3.18	(88.66%) 0.044 ± 0.148	(56.38%) 0.119 ± 0.239	(60.82%) 0.061 ± 0.257
TF18	9880	3054	6826	3054	4057	20.80	19.61	22.06	2.51	(91.92%) 0.040 ± 0.200	(60.38%) 0.106 ± 0.287	(67.06%) 0.063 ± 0.313
TF19	8494	2232	6262	2232	1123	20.95	19.85	21.95	1.83	(80.55%) 0.077 ± 0.237	(50.60%) 0.174 ± 0.253	(62.48%) 0.109 ± 0.220
TF20	7491	2161	5330	2161	43196	20.66	19.60	22.05	1.77	-	(61.46%) 0.057 ± 0.429	(61.46%) 0.057 ± 0.429

Table 9: The results obtained after applying the optimal combination of K_0 , $C1$, and $C2$ models to the most probable stellar sources in the optical data of the test fields for 1.06 sq. deg. area of the sky. The entries marked by the '-' symbol are for the test fields that do not have the observed UKIDSS data in the J and H bands while '*' indicates the fields that are close to the galactic plane and due to high source density, the data is obtained for 5 arcmin radius only. The % values in the brackets are the percentage number of sources within 0.2 mag of the absolute difference in the observed and the computed NIR magnitudes and the density is computed for the NFRAOS field-of-view.

References

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, , 558, A33, doi: [10.1051/0004-6361/201322068](https://doi.org/10.1051/0004-6361/201322068)
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *aj*, 156, 123, doi: [10.3847/1538-3881/aabc4f](https://doi.org/10.3847/1538-3881/aabc4f)
- Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, *apj*, 935, 167, doi: [10.3847/1538-4357/ac7c74](https://doi.org/10.3847/1538-4357/ac7c74)
- Castelli, F., & Kurucz, R. L. 2003, in Modelling of Stellar Atmospheres, ed. N. Piskunov, W. W. Weiss, & D. F. Gray, Vol. 210, A20. <https://arxiv.org/abs/astro-ph/0405087>
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv e-prints, arXiv:1612.05560. <https://arxiv.org/abs/1612.05560>
- Crane, J. D., Herriot, G., Andersen, D. R., et al. 2018, in Astronomical Telescopes + Instrumentation
- Czekaj, M. A., Robin, A. C., Figueras, F., Luri, X., & Haywood, M. 2014, , 564, A102, doi: [10.1051/0004-6361/201322139](https://doi.org/10.1051/0004-6361/201322139)
- Fitzpatrick, E. L. 1999, *pasp*, 111, 63, doi: [10.1086/316293](https://doi.org/10.1086/316293)
- Flewelling, H. A., Magnier, E. A., Chambers, K. C., et al. 2020, , 251, 7, doi: [10.3847/1538-4365/abb82d](https://doi.org/10.3847/1538-4365/abb82d)
- Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2022, arXiv e-prints, arXiv:2208.00211. <https://arxiv.org/abs/2208.00211>
- Green, G. 2018, *The Journal of Open Source Software*, 3, 695, doi: [10.21105/joss.00695](https://doi.org/10.21105/joss.00695)
- Hauschildt, P. H., Allard, F., & Baron, E. 1999a, *apj*, 512, 377, doi: [10.1086/306745](https://doi.org/10.1086/306745)
- Hauschildt, P. H., Allard, F., Ferguson, J., Baron, E., & Alexander, D. R. 1999b, *apj*, 525, 871, doi: [10.1086/307954](https://doi.org/10.1086/307954)
- Indebetouw, R., Mathis, J. S., Babler, B. L., et al. 2005, *apj*, 619, 931, doi: [10.1086/426679](https://doi.org/10.1086/426679)
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, *apj*, 873, 111, doi: [10.3847/1538-4357/ab042c](https://doi.org/10.3847/1538-4357/ab042c)
- Kaiser, N., Aussel, H., Burke, B. E., et al. 2002, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, 154–164, doi: [10.1117/12.457365](https://doi.org/10.1117/12.457365)

Kaiser, N., Burgett, W., Chambers, K., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7733, Ground-based and Airborne Telescopes III, ed. L. M. Stepp, R. Gilmozzi, & H. J. Hall, 77330E, doi: [10.1117/12.859188](https://doi.org/10.1117/12.859188)

Kron, R. G. 1980, , 43, 305, doi: [10.1086/190669](https://doi.org/10.1086/190669)

Kurucz, R. L. 1992a, rmxaa, 23, 45

Kurucz, R. L. 1992b, in The Stellar Populations of Galaxies, ed. B. Barbuy & A. Renzini, Vol. 149, 225

Kurucz, R. L. 1993, in Astronomical Society of the Pacific Conference Series, Vol. 44, IAU Colloq. 138: Peculiar versus Normal Phenomena in A-type and Related Stars, ed. M. M. Dworetsky, F. Castelli, & R. Faraggiana, 87

Magnier, E. A., Sweeney, W. E., Chambers, K. C., et al. 2020, , 251, 5, doi: [10.3847/1538-4365/abb82c](https://doi.org/10.3847/1538-4365/abb82c)

Marshall, D. J., Robin, A. C., Reylé, C., Schultheis, M., & Picaud, S. 2006, , 453, 635, doi: [10.1051/0004-6361:20053842](https://doi.org/10.1051/0004-6361:20053842)

McMahon, R. 2012, in Science from the Next Generation Imaging and Spectroscopic Surveys, 37

Saito, R., Hempel, M., Alonso-García, J., et al. 2010, The Messenger, 141, 24

Schlafly, E. F., & Finkbeiner, D. P. 2011, The Astrophysical Journal, 737, 103, doi: [10.1088/0004-637x/737/2/103](https://doi.org/10.1088/0004-637x/737/2/103)

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, apj, 500, 525, doi: [10.1086/305772](https://doi.org/10.1086/305772)

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, aj, 131, 1163, doi: [10.1086/498708](https://doi.org/10.1086/498708)

STScI Development Team. 2013, pysynphot: Synthetic photometry software package. <http://ascl.net/1303.023>

Subramanian, S., Subramaniam, A., Sivarani, T., et al. 2016, Journal of Astrophysics and Astronomy, 37, 24, doi: [10.1007/s12036-016-9401-1](https://doi.org/10.1007/s12036-016-9401-1)

Subramanian, S., Subramaniam, A., Simard, L., et al. 2013, Journal of Astrophysics and Astronomy, 34, 175, doi: [10.1007/s12036-013-9176-6](https://doi.org/10.1007/s12036-013-9176-6)

Tonry, J. L., Stubbs, C. W., Lykke, K. R., et al. 2012, apj, 750, 99, doi: [10.1088/0004-637X/750/2/99](https://doi.org/10.1088/0004-637X/750/2/99)

Whitmore, B. C., Allam, S. S., Budavári, T., et al. 2016, , 151, 134, doi: [10.3847/0004-6256/151/6/134](https://doi.org/10.3847/0004-6256/151/6/134)

Wizinowich, P. L., Le Mignant, D., Bouchez, A. H., et al. 2006, *pasp*, 118, 297, doi: [10.1086/499290](https://doi.org/10.1086/499290)