PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Spitzer/IRAC precision photometry: a machine learning approach

James G. Ingalls, Jessica E. Krick, Sean J. Carey, Patrick J. Lowrance, Carl J. Grillmair, et al.

James G. Ingalls, Jessica E. Krick, Sean J. Carey, Patrick J. Lowrance, Carl J. Grillmair, William J. Glaccum, Seppo Laine, Jonathan D. Fraine, "Spitzer/IRAC precision photometry: a machine learning approach," Proc. SPIE 10698, Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, 106985E (24 July 2018); doi: 10.1117/12.2313640

Event: SPIE Astronomical Telescopes + Instrumentation, 2018, Austin, Texas, United States

Spitzer/IRAC Precision Photometry: A Machine Learning Approach

James G. Ingalls[a], Jessica E. Krick[a], Sean J. Carey[a], Patrick J. Lowrance[a], Carl J. Grillmair[a], William J. Glaccum[a], Seppo Laine[a], and Jonathan D. Fraine[b]

[a]IPAC, Mail Code 314-6, Caltech, 1200 E. California Blvd., Pasadena, CA 91125, USA
[b]Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

ABSTRACT

The largest source of noise in exoplanet and brown dwarf photometric time series made with *Spitzer*/IRAC is the coupling between intra-pixel gain variations and spacecraft pointing fluctuations. Observers typically correct for this systematic in science data by deriving an instrumental noise model simultaneously with the astrophysical light curve and removing the noise model. Such techniques for self-calibrating Spitzer photometric datasets have been extremely successful, and in many cases enabled near-photon-limited precision on exoplanet transit and eclipse depths. Self-calibration, however, can suffer from certain limitations: (1) temporal astrophysical signals can become aliased as part of the instrument model; (2) for some techniques adequate model estimation often requires a high degree of intra-pixel positional redundancy (multiple samples with nearby centroids) over long time spans; (3) many techniques do not account for sporadic high frequency telescope vibrations that smear out the point spread function. We have begun to build independent general-purpose intra-pixel systematics removal algorithms using three machine learning techniques: K-Nearest Neighbors (with kernel regression), Random Decision Forests, and Artificial Neural Networks. These methods remove many of the limitations of self-calibration: (1) they operate on a dedicated calibration database of approximately one million measurements per IRAC waveband (3.6 and 4.5 microns) of non-variable stars, and thus are independent of the time series science data to be corrected; (2) the database covers a large area of the "Sweet Spot, so the methods do not require positional redundancy in the science data; (3) machine learning techniques in general allow for flexibility in training with multiple, sometimes unorthodox, variables, including those that trace PSF smear. We focus in this report on the K-Nearest Neighbors with Kernel Regression technique. (Additional communications are in preparation describing Decision Forests and Neural Networks.)

Keywords: Spitzer, calibration, algorithms, intra-pixel sensitivity, precision photometry, exoplanets, machine learning, k-nearest neighbors

1. INTRODUCTION

The largest source of noise in exoplanet and brown dwarf photometric time series made with the InfraRed Array Camera (IRAC)[1] on board the *Spitzer* Space Telescope[2] is the coupling between intra-pixel gain variations and spacecraft pointing fluctuations. Observers typically correct for this systematic in science data by deriving an instrumental noise model simultaneously with the astrophysical light curve and removing the noise model. Such techniques for self-calibrating Spitzer photometric datasets have been extremely successful, and in many cases enabled near-photon-limited precision on exoplanet transit and eclipse depths.[3] Self-calibration, however, can suffer from certain limitations: (1) temporal astrophysical signals can become aliased as part of the instrument model; (2) for some techniques adequate model estimation often requires a high degree of intra-pixel positional redundancy (multiple samples with nearby centroids) over long time spans; (3) many techniques do not account for sporadic high frequency telescope vibrations that smear out the point spread function.

We have begun to build independent general-purpose intra-pixel systematics removal algorithms using three machine learning techniques: K-Nearest Neighbors with Kernel Regression (KNN-KR), Random Decision Forests,

Further author information: (Send correspondence to J.G.I.)
J.G.I.: E-mail: ingalls@ipac.caltech.edu, Telephone: 1 626 395 8659

Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave, edited by Makenzie Lystrup,
Howard A. MacEwen, Giovanni G. Fazio, Proc. of SPIE Vol. 10698, 106985E
© 2018 SPIE · CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2313640

and Artificial Neural Networks. These methods remove many of the limitations of self-calibration: (1) they operate on a dedicated calibration database of approximately one million measurements per IRAC waveband (3.6 and 4.5 μm) of non-variable stars, and thus are independent of the time series science data to be corrected; (2) the database covers a large area of the "Sweet Spots" on the IRAC cameras, so the methods do not require positional redundancy in the science data; (3) machine learning techniques in general allow for flexibility in training with multiple, sometimes unorthodox, variables, including those that trace PSF smear.

Machine Learning (ML) is a branch of computer science and artificial intelligence that attempts to teach computers to learn like humans do, by generalizing from data. The goal is to derive a map from input to output data, without an explicit model. In this report, we focus on one of the simplest forms of ML, K-Nearest Neighbors with Kernel Regression. (Additional communications are in preparation describing Random Decision Forests and Neural Networks.) This technique has already been used successfully for self-calibration of exoplanet lightcurves.[4,5] Using the technique to "learn" systematics from our independent calibration dataset has also been tried with varying degrees of success. The first version, using a much smaller sample than described here, used KNN-KR to create a table of corrections as a function of subpixel location,[6] which could be interpolated to the position of interest. This method had insufficient resolution to follow all variations in pixel sensitivity. A second version, which interrogated the calibration dataset in real time to estimate the pixel response from the nearest neighbors of a measurement, was better at following variations in sensitivity,[7] but was not as good as self-calibration using the same technique.[3,8]

As stated above, it is desirable to have a systematics removal method that is independent of the measurements being corrected, chiefly because of the problem of knowing how well a method has separated systematics from astrophysics. One approach to measuring the quality of separation is to simulate data with real systematics, as in the 2015 *Spitzer* Exoplanet Data Challenge, whose results were described in Ref. 3. But any such method is limited by the fidelity of the simulations. We still believe that the best approach should be to measure any static, non-evolving systematics independently of time series lightcurves.

Here we introduce a third approach to KNN-KR, which updates the technique so that it attempts to reach a better representation of the intrapixel systematics using a calibration dataset. We first describe the calibration datasets, followed by an overview of the KNN-KR algorithm and a description of how we derive kernel-based estimates from a dataset. We use the technique of Ref. 9 to develop "Steering" Kernels that morph to the data. We visualize the intra-pixel gain map as a function of x and y pixel centroid position, and cross-validate the algorithm on a portion of the dataset removed from consideration, testing the accuracy and precision of KNN-KR in predicting the fluxes of the test subset.

2. PIXEL MAPPING DATASETS

Over approximately 5 years, we have accumulated a set of pixel mapping measurements of two stars: KF09T1 (TYC 4212-1074-1) in IRAC Channel 1 (3.6 μm; 1421280 measurements) and BD+67 1044 (NPM1+67.0536; 874950 measurements) in IRAC Channel 2 (4.5 μm). We define x and y pixel centroid of a point source in terms of the first moment of the brightness image:

$$x_{\rm cen} = \frac{\sum_{i,j}(f-{\rm B}G)_{i,j}\, i}{\sum_{i,j}(f-{\rm B}G)_{i,j}}; \tag{1}$$

$$y_{\rm cen} = \frac{\sum_{i,j}(f-{\rm B}G)_{i,j}\, j}{\sum_{i,j}(f-{\rm B}G)_{i,j}}. \tag{2}$$

Here, i and j are the integer pixel numbers in the x and y directions, $f_{i,j}$ is the brightness of pixel ij, and BG is the average background measured outside the region being summed over. Typically the moment sum is taken over a small number of pixels surrounding the point source.

Figure 1 displays the centroids for our pixel mapping (pmap) datasets, centered on pixel (15,15) of the IRAC Ch 1 and Ch 2 subarray apertures. Most of the measurements were obtained by a target pre-acquisition technique called "Peak-Up" which uses a separate camera on board *Spitzer*, the Pointing Calibration & Reference Sensor (PCRS)[10] to refine the pointing to the target before observing with IRAC. We have defined a "Sweet Spot"

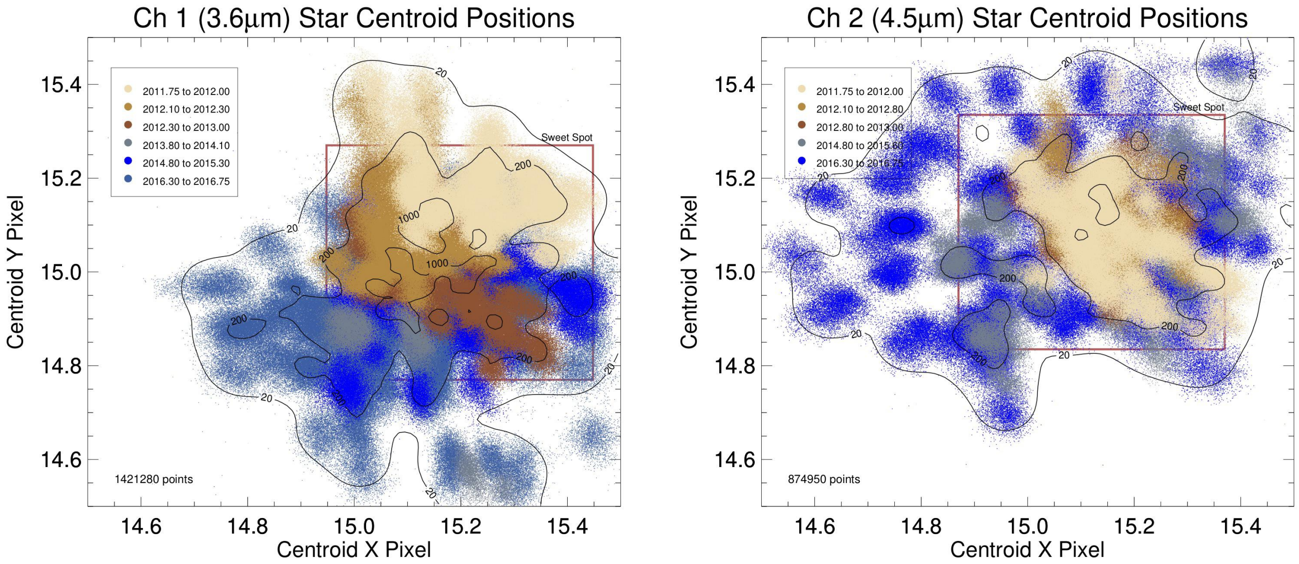

Figure 1. Pixel x and y centroids (subarray coordinates) for the calibration datasets, color coded by observing epoch. Contours of constant occupation number (see text) are indicated on the plots. Red squares label the "Sweet Spot" regions within which PCRS Peak-Up is used to place all time series observations. (*Left*) Channel 1 (3.6 μm) measurements of KF09T1. (*Right*) Channel 2 (4.5 μm) measurements of BD+67 1044.

region on each pixel, towards which the PCRS Peak-up is more than 90% successful at placing a point source (provided self-Peak-up is used).[6,11] The *Spitzer* Science Center recommends that all observers interested in precision time series use the PCRS to acquire their targets and put them on the Sweet Spots.

Each data point on Fig. 1 represents a unique stellar image taken with an IRAC array. In addition to the centroid, as shown in the Figure, we also measure the photometric flux in an aperture, F, the measured FWHM size of the star in the x and y directions, $x_{\rm FWHM}$ and $y_{\rm FWHM}$, and other quantities that we do not discuss here. The flux as a function of time gives the light curve of time series measurements. For these calibration datasets, where the targets are non-variable stars (flat light curve), any variations in F can be taken to be either random noise or instrument systematics. In principle, all features of the dataset can be examined for relationships with F and used to build a predictive model of the systematics given by a score determined from the feature set. This is the province of Machine Learning algorithms.

3. THE KNN-KR ALGORITHM

3.1 Overall method

We follow the kernel regression prescription of Ref. 9. Given a dataset of P measurements of a function $\{f_i\}$ along with D additional data features represented by the coordinate vectors $\{\mathbf{x}_i\}$, construct a model of $F(\mathbf{x})$ such that at point $\mathbf{x}$ *not* in the dataset:

1. The function is Taylor expanded in the local region about $\mathbf{x}$, with the dataset $\{f_i\}$ being used to solve for the first $n_{\rm order}$ derivatives of the function via χ^2 minimization (the zeroth derivative being an estimate of the actual function at $\mathbf{x}$);

2. The fit is localized by weighting each data point by a set of kernel values $\{K(\mathbf{x}_i - \mathbf{x})\}$, which fall off with "distance" from the new point $\mathbf{x}$;

3. The fit is localized further by considering only a subset of $k_{\rm NN}$ "nearest" neighbors to $\mathbf{x}$ in any computations over the dataset.

We refer to "distance" and "nearest" in quotes above because these terms do not always correspond to the Euclidean distance,

$$d_{\mathrm{E}} = \sqrt{(\mathbf{x}_i - \mathbf{x})^T\,(\mathbf{x}_i - \mathbf{x})}, \tag{3}$$

but can often incorporate a warping of the coordinate space via a $(D \times D)$ covariance matrix, $\mathbf{C}$:

$$d_{\mathrm{M}} = \sqrt{(\mathbf{x}_i - \mathbf{x})^T\,\mathbf{C}^{-1/2}\,(\mathbf{x}_i - \mathbf{x})}, \tag{4}$$

also known as the Mahalanobis distance.[12] Takeda et al. describe a method for deriving a set of local "steering" kernels from the dataset,[9] which relies on the pre-computation of covariance matrices at points spanning the coordinate space. The goal here is to maximize the contribution to a given estimate from points that are both local in *coordinates* (Euclidean distance), and have the most similar *fluxes*.

3.2 Steering kernels

For our purposes a kernel is a peaked function that usually falls to zero when its argument is large. There are many possible kernels than can optimize different metrics.[13,14] In our analysis, however, we use Gaussian kernel functions, because they are easy to compute and familiar to most scientists and engineers:

$$K(\mathbf{x}_i - \mathbf{x}) = \frac{\sqrt{\det(\mathbf{C}_i)}}{(2\pi h^2 \mu_i^2)^{D/2}} \exp\left[-\frac{(\mathbf{x}_i - \mathbf{x})^T\,\mathbf{C}^{-1/2}\,(\mathbf{x}_i - \mathbf{x})}{2h^2\mu_i^2}\right]. \tag{5}$$

Here h is an optional global smoothing parameter, often taken to be unity in our analysis, D is the number of dimensions of $\mathbf{x}$, and μ_i is the local "bandwidth" at $\mathbf{x}$, which we take to be the standard deviation in the Euclidean coordinate distances of the k_{NN} neighbors in the calibration dataset from the point of interest, $\mathbf{x}$. This definition of μ_i has the effect of widening the kernel when the k_{NN} neighbors are spread out and narrowing the kernel when they are close together.

The shape and orientation of Gaussian steering kernels are governed by the covariance matrix, $\mathbf{C}$, the computation of which is done iteratively using the following procedure:

1. Define a grid of values for $\mathbf{x}$, covering the range over which we expect the coordinates to vary, with sufficient resolution.

2. Calculate the function $F(\mathbf{x})$ at each grid point using circular Gaussian kernels ($\mathbf{C}$ = Identity Matrix). Repeat N_{smooth} times to obtain a smooth, minimally noisy version of F.

3. Calculate the first derivative of F_{smooth} at each grid point using the Taylor expansion least squares fit, applied to the gridded values of F (not the source dataset).

4. Build a Jacobian matrix at each grid point, consisting of the computed derivatives of the flux as a function of each coordinate, for k_{smooth} nearest grid neighbors. This yields information on the orientation of structures in the data.

5. Following the principles outlined in Ref. 9, estimate $\mathbf{C}$ from a singular value decomposition of the Jacobian matrix.

Selected kernels derived from the pixel mapping datasets are shown in Fig. 2. The matrix $\mathbf{C}$ (together with the local bandwidth μ) was computed for 100×100 grid points across the pixel, but not all kernels are shown for clarity. In machine learning parlance, this stage of the process is where we use the data $(\mathbf{x}_i, F(\mathbf{x}_i))$ to "teach" the computer how to more accurately map inputs $(\mathbf{x})$ to outputs $(F(\mathbf{x}))$. Reference to Fig. 3 shows that the kernels are indeed oriented along structures in the data.

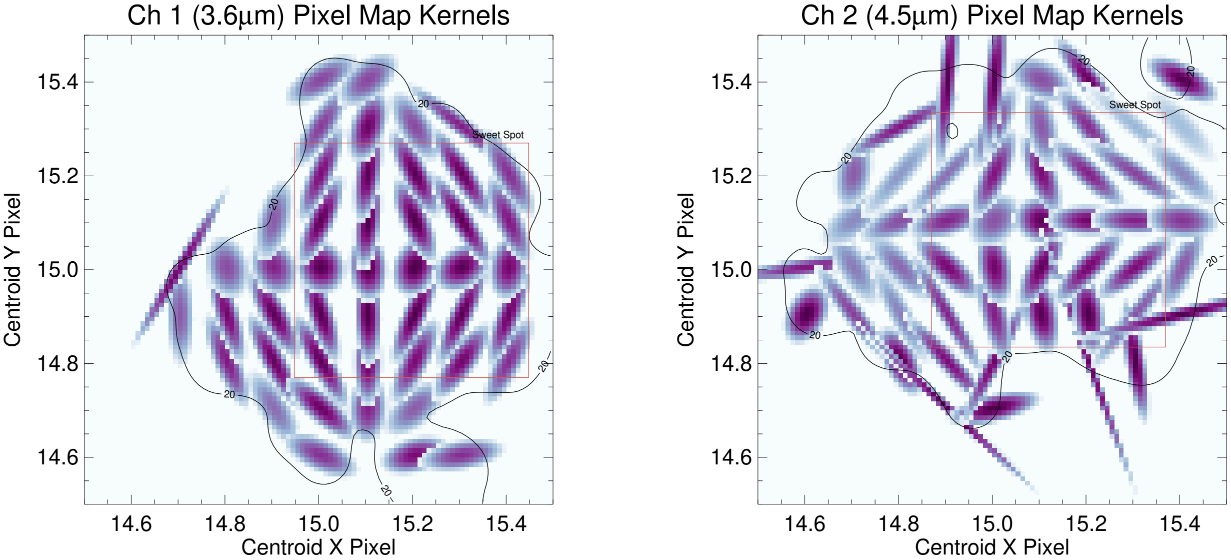

Figure 2. Steering kernels based on analysis of the pixel mapping datasets for Channel 1 (3.6 μm) (*Left*) and Channel 2 (4.5 μm) (*Right*). Contours of occupation number 20 (see text) are indicated on the plots. Kernels are not shown outside this contour. Red squares label the "Sweet Spot" regions within which PCRS Peak-Up is used to place all time series observations.

3.3 Kernel regression estimates

A kernel regression estimate of F is derived at the point $\mathbf{x}$ using the following procedure:

1. Find the $k_{\rm NN}$ nearest neighbors to $\mathbf{x}$ in the calibration dataset, using the Euclidean distance.
2. Find the nearest grid point to $\mathbf{x}$ in the 100×100 table.
3. Use the values of $\mathbf{C}$ and μ at the chosen grid point to estimate the kernel values for the $k_{\rm NN}$ neighbors (Eq. 5).
4. Use the matrix equations given in Ref. 9 to compute the regression value of $F(\mathbf{x})$.

We can estimate the *density* at $\mathbf{x}$ as the sum over the $k_{\rm NN}$ kernel estimates at $\mathbf{x}$. For the *tabulated* kernels that span the range of the dataset, we also compute an *occupation number*, that is the density normalized in such a way that the sum over all tabulated kernels equals the total number of points in the dataset. The occupation number is thus the approximate number of data points contributing to kernel estimates at the grid point. As in Ref. 6 we find that the most reliable kernel estimates result when occupation is greater than 20. In locations where the density falls to lower values, kernels are sampled less isotropically and result in spurious estimates. Contours of occupation number are shown in Figs. 1 and 2.

To visualize the sensitivity, we compute $F(\mathbf{x})$ on our 100×100 grid and show the results in Fig. 3. To normalize these maps we divided kernel-predicted fluxes by the average aperture flux of each star over the entire pixel. As mentioned above, we compute sensitivity only for grid points where occupation > 20. For grid points with lower occupation, the maps show discontinuities in sensitivity, which appear "noisy."

4. CROSS-VALIDATION

The quality of an estimator can be assessed when the value being estimated is already known. The easiest way to obtain a set of known values is by extracting it from the calibration dataset itself. For the pixel mapping datasets

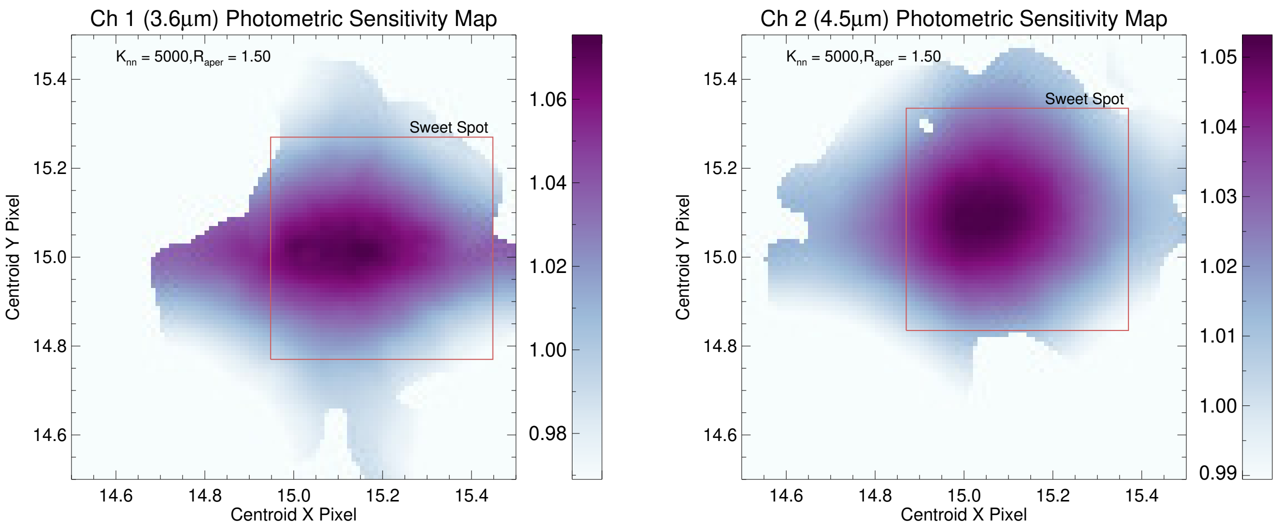

Figure 3. Relative photometric sensitivity as a function of centroid location, based on analysis of the pixel mapping datasets for Channel 1 (3.6 μm) (*Left*) and Channel 2 (4.5 μm) (*Right*). White regions indicate missing data (occupation number < 20). Red squares label the "Sweet Spot" regions within which PCRS Peak-Up is used to place all time series observations.

of each warm IRAC channel, we removed 30,000 measurements randomly but uniformly spaced across each pixel. These "test" samples were not used to compute the kernels. We used the remaining "training" samples to build the kernels and, via KNN-KR estimation, to predict the values of $F(\mathbf{x}_{\rm test})$ (dependent variables) for the set of $\{\mathbf{x}\}_{\rm test}$ coordinates (independent variables). This process is depicted as a flow chart in Fig. 4.

One useful statistic for evaluating the quality of a regression estimator is the coefficient of determination (also known as R^2), which measures the degree to which the estimates account for the variance in the dataset. Given a set of measurements of a function, $\{f_{\rm test}\}$ at coordinates $\{\mathbf{x}_{\rm test}\}$, and a corresponding set of regression

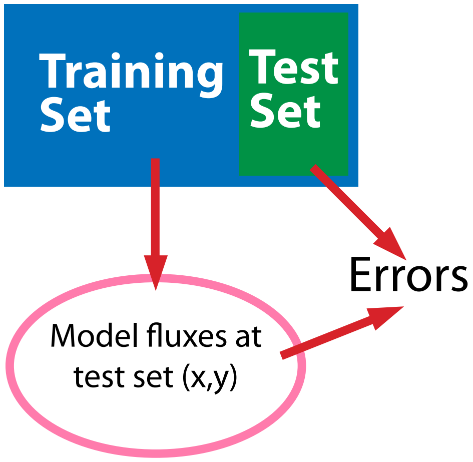

Figure 4. Schematic of cross-validation technique. The calibration data are split into "training" and "test" subsets. The training set is used to estimate the model flux $F(\mathbf{x}_{\rm test})$ at the set of $\{\mathbf{x}\}_{\rm test}$ values. The true fluxes of the test set are then used to compute errors.

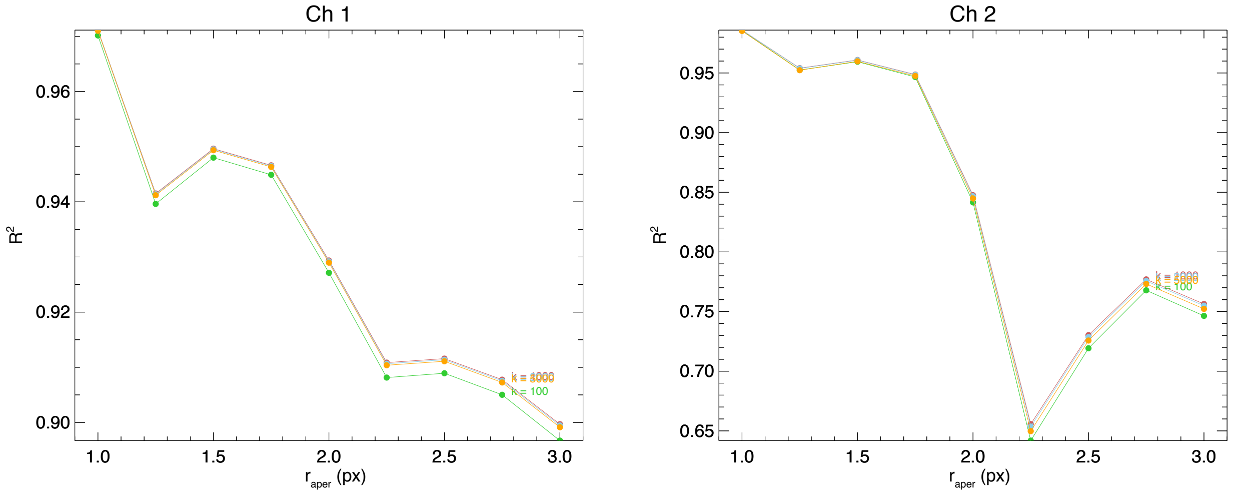

Figure 5. Coefficient of determination (R^2) statistic for cross-validation dataset, for Channel 1 (3.6 μm) (*Left*) and Channel 2 (4.5 μm) (*Right*). This measures the amount of variance in the dataset $\sigma^2(f_{\rm test})$ that is accounted for by the KR-KNN prediction $\sigma^2[f_{\rm test} - F(\mathbf{x}_{\rm test})]$. Colored lines depict different choices for the number of neighbors used, $k_{\rm NN}$. Results are shown for a range of values of the aperture photometry radius, $r_{\rm aper}$.

estimates, $\{F(\mathbf{x}_{\rm test})\}$,

$$R^2 = 1 - \frac{\sum[f_{\rm test} - F(\mathbf{x}_{\rm test})]^2}{\sum(f_{\rm test} - \overline{f_{\rm test}})^2}. \tag{6}$$

We show the values of R^2 for the cross-validation datasets as a function of both aperture photometry radius, $r_{\rm aper}$ and the number of neighbors in the kernel regression, $k_{\rm NN}$. The regression does a better job of capturing the variance in the data (higher R^2) for smaller $r_{\rm aper}$, but $k_{\rm NN}$ does not affect R^2 noticeably.

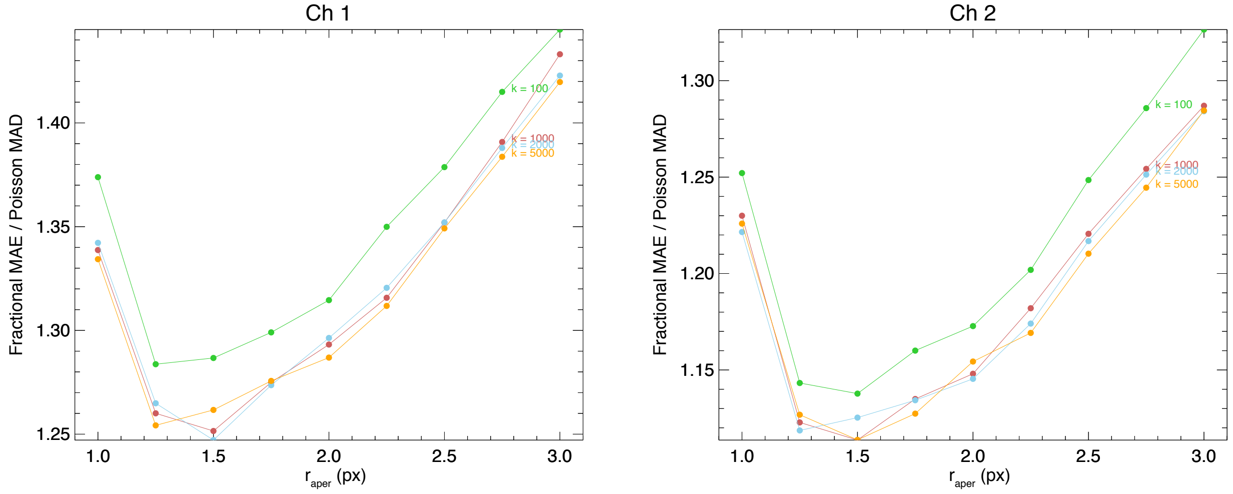

Figure 6. Mean absolute error (MAE) for cross-validation dataset, for Channel 1 (3.6 μm) (*Left*) and Channel 2 (4.5 μm) (*Right*). The MAE has been normalized by the expected mean absolute deviation from 0 (MAD) of a Poisson random variable with the same number of electron counts as the observations. As in Fig. 5, colored lines depict different choices for the number of neighbors used, $k_{\rm NN}$, and the ordinate spans a range of values of the aperture photometry radius, $r_{\rm aper}$.

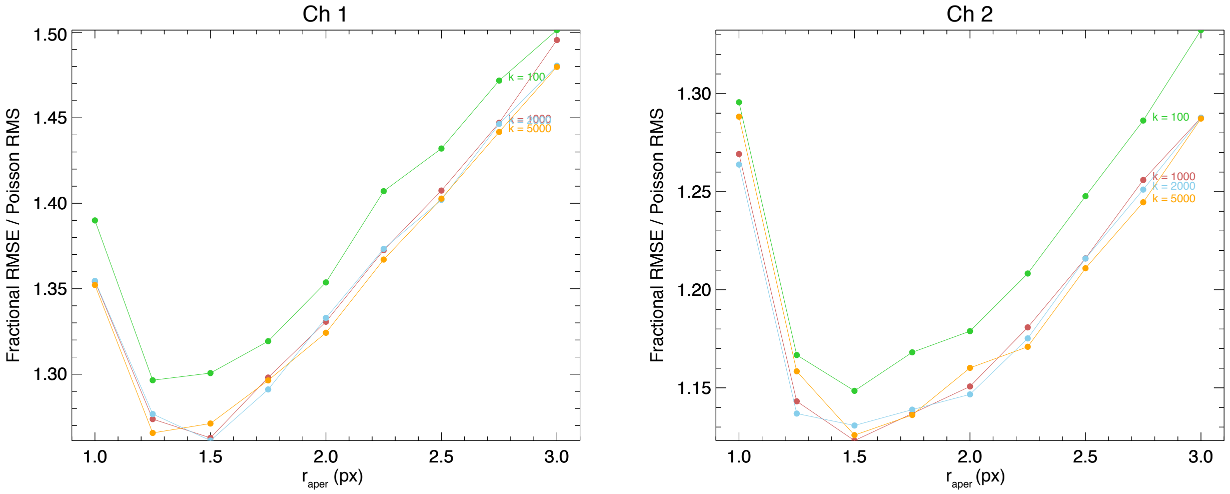

Figure 7. Root mean squared error (RMSE) for cross-validation dataset, for Channel 1 (3.6 μm) (*Left*) and Channel 2 (4.5 μm) (*Right*), which is more susceptible to outliers than MAE. The RMSE has been normalized by the expected root mean squared deviation from 0 (RMSD) of a Poisson random variable with the same number of electron counts as the observations. As in Fig. 5, colored lines depict different choices for the number of neighbors used, $k_{\rm NN}$, and the ordinate spans a range of values of the aperture photometry radius, $r_{\rm aper}$.

Although R^2 can often be a good indicator of regression quality, and especially (as in this case) allows us to constrain control parameters (such as $r_{\rm aper}$), it has some flaws. Firstly, even a perfect fit can have a low value of R^2 if the noise in the data is high. Moreover, R^2 measures the *correlation* between variables, not how well a model captures the *dependence* of one set of variables on another, which is the ultimate goal of any regression.

A better indication of goodness of fit is the mean absolute error (MAE),

$$\mathrm{MAE} = \frac{\sum |f_{\rm test} - F(\mathbf{x}_{\rm test})|}{N_{\rm test}}, \tag{7}$$

which estimates the overall accuracy (including precision and bias) of an estimate. In Fig. 6 we plot MAE for the same range of $k_{\rm NN}$ and $r_{\rm aper}$ as in Fig. 5. Both IRAC Channels show relatively well-behaved MAE curves, with positive concavity and a single minimum near $r_{\rm aper} = 1.25 - 1.5\,$px and $k_{\rm NN} \gtrsim 1000$. The plots shown in Fig. 6 are normalized to the mean absolute deviation of a Poisson random variable with the same number of electron counts as each observation, thus giving a sense of how close to the photon noise floor the residuals are. For Ch 1, the MAE minimum occurs about 25% above the noise floor and the Ch 2 MAE minimum occurs about 11% above the floor. In other words, if we correct the data for systematics (divide the data by $F(\mathbf{x}_{\rm test})$), photon noise accounts for 80% (Ch 1) and 90% (Ch 2) of the scatter in the resulting values.

A third statistic is root mean square error (RMSE),

$$\mathrm{RMSE} = \sqrt{\frac{\sum \left[f_{\rm test} - F(\mathbf{x}_{\rm test})\right]^2}{N_{\rm test}}}. \tag{8}$$

Due to the fact that errors are squared, this metric is weighted more heavily towards data with larger errors than MAE, and so can reveal problems with possible outliers. Indeed, the minimum RMSE of both datasets is only slightly larger than the minimum MAE (26% above the photon limit in Ch 1; 12% above the photon limit in Ch 2), indicating that there are outliers but that they are not a significant component of the dataset. The overall conclusion, that $r_{\rm aper} = 1.25 - 1.5\,$px and $k_{\rm NN} \gtrsim 1000$ minimizes the errors, is the same as derived from the MAE plots.

5. CONCLUSIONS

We have demonstrated a new version of the K-nearest neighbors Kernel Regression (KNN-KR) machine learning technique for systematics removal in post-cryogenic *Spitzer*/IRAC photometry that makes use of an independent calibration dataset that does not require self-calibration and maintains near photon-limited performance on a cross-validation "test" dataset. For this demonstration we confined our systematics estimation to a function solely of the two-dimensional vector $\mathbf{x} = (x, y)$ pixel centroid. The method is extensible to additional dimensions, and in a future work we plan to incorporate the apparent stellar size $(x_{\rm FWHM}, y_{\rm FWHM})$, which has been found to trace the systematics, even at times when the centroid does not.[15]

REFERENCES

[1] Fazio, G. G., Hora, J. L., Allen, L. E., Ashby, M. L. N., Barmby, P., Deutsch, L. K., Huang, J.-S., Kleiner, S., Marengo, M., Megeath, S. T., Melnick, G. J., Pahre, M. A., Patten, B. M., Polizotti, J., Smith, H. A., Taylor, R. S., Wang, Z., Willner, S. P., Hoffmann, W. F., Pipher, J. L., Forrest, W. J., McMurty, C. W., McCreight, C. R., McKelvey, M. E., McMurray, R. E., Koch, D. G., Moseley, S. H., Arendt, R. G., Mentzell, J. E., Marx, C. T., Losch, P., Mayman, P., Eichhorn, W., Krebs, D., Jhabvala, M., Gezari, D. Y., Fixsen, D. J., Flores, J., Shakoorzadeh, K., Jungo, R., Hakun, C., Workman, L., Karpati, G., Kichak, R., Whitley, R., Mann, S., Tollestrup, E. V., Eisenhardt, P., Stern, D., Gorjian, V., Bhattacharya, B., Carey, S., Nelson, B. O., Glaccum, W. J., Lacy, M., Lowrance, P. J., Laine, S., Reach, W. T., Stauffer, J. A., Surace, J. A., Wilson, G., Wright, E. L., Hoffman, A., Domingo, G., and Cohen, M., "The Infrared Array Camera (IRAC) for the Spitzer Space Telescope," **154**, 10–17 (Sept. 2004).

[2] Werner, M. W., Roellig, T. L., Low, F. J., Rieke, G. H., Rieke, M., Hoffmann, W. F., Young, E., Houck, J. R., Brandl, B., Fazio, G. G., Hora, J. L., Gehrz, R. D., Helou, G., Soifer, B. T., Stauffer, J., Keene, J., Eisenhardt, P., Gallagher, D., Gautier, T. N., Irace, W., Lawrence, C. R., Simmons, L., Van Cleve, J. E., Jura, M., Wright, E. L., and Cruikshank, D. P., "The Spitzer Space Telescope Mission," **154**, 1–9 (Sept. 2004).

[3] Ingalls, J. G., Krick, J. E., Carey, S. J., Stauffer, J. R., Lowrance, P. J., Grillmair, C. J., Buzasi, D., Deming, D., Diamond-Lowe, H., Evans, T. M., Morello, G., Stevenson, K. B., Wong, I., Capak, P., Glaccum, W., Laine, S., Surace, J., and Storrie-Lombardi, L., "Repeatability and Accuracy of Exoplanet Eclipse Depths Measured with Post-cryogenic Spitzer," *The Astronomical Journal* **152**, 44 (Aug. 2016).

[4] Ballard, S., Charbonneau, D., Deming, D., Knutson, H. A., Christiansen, J. L., Holman, M. J., Fabrycky, D., Seager, S., and A'Hearn, M. F., "A Search for a Sub-Earth-Sized Companion to GJ 436 and a Novel Method to Calibrate Warm Spitzer IRAC Observations," *Publications of the Astronomical Society of the Pacific* **122**, 1341–1352 (Nov. 2010).

[5] Lewis, N. K., Knutson, H. A., Showman, A. P., Cowan, N. B., Laughlin, G., Burrows, A., Deming, D., Crepp, J. R., Mighell, K. J., Agol, E., Bakos, G. Á., Charbonneau, D., Désert, J.-M., Fischer, D. A., Fortney, J. J., Hartman, J. D., Hinkley, S., Howard, A. W., Johnson, J. A., Kao, M., Langton, J., and Marcy, G. W., "Orbital Phase Variations of the Eccentric Giant Planet HAT-P-2b," *The Astrophysical Journal* **766**, 95 (Apr. 2013).

[6] Ingalls, J. G., Krick, J. E., Carey, S. J., Laine, S., Surace, J. A., Glaccum, W. J., Grillmair, C. C., and Lowrance, P. J., "Intra-pixel gain variations and high-precision photometry with the Infrared Array Camera (IRAC)," in [*Proc. SPIE*], Clampin, M. C., Fazio, G. G., MacEwen, H. A., and Oschmann, J. M., eds., 84421Y, Spitzer Science Ctr., California Institute of Technology (United States), SPIE (Sept. 2012).

[7] Krick, J. E., Ingalls, J., Carey, S., von Braun, K., Kane, S. R., Ciardi, D., Plavchan, P., Wong, I., and Lowrance, P., "Spitzer IRAC Sparsely Sampled Phase Curve of the Exoplanet Wasp-14B," *The Astrophysical Journal* **824**, 27 (June 2016).

[8] Kilpatrick, B. M., Lewis, N. K., Kataria, T., Deming, D., Ingalls, J. G., Krick, J. E., and Tucker, G. S., "Spitzer Secondary Eclipse Depths with Multiple Intrapixel Sensitivity Correction Methods Observations of WASP-13b, WASP-15b, WASP-16b, WASP-62b, and HAT-P-22b," *The Astronomical Journal* **153**, 22 (Jan. 2017).

[9] Takeda, H., Farsiu, S., and Milanfar, P., "Kernel Regression for Image Processing and Reconstruction," *IEEE Trans. on Image Process.* **16**(2), 349–366 (2007).

[10] Mainzer, A. K., Young, E. T., Huff, L. W., and Swanson, D., “Pre-launch performance testing of the pointing calibration and reference sensor for SIRTF,” in [*IR Space Telescopes and Instruments. Edited by John C. Mather . Proceedings of the SPIE*], Mather, J. C., ed., 122–129, Lockheed Martin Advanced Technology Ctr., USA, SPIE (Mar. 2003).

[11] SSC, “Advice for designing high precision photometry observations.” IRAC High Precision Photometry, 28 July 2016 https://irachpp.spitzer.caltech.edu/page/Obs%20Planning. (Accessed: 27 June 2018).

[12] Mahalanobis, P. C., “On the generalized distance in statistics,” *Proceedings of the National Institute of Sciences of India* **2**(1), 49–55 (1936).

[13] Gasser, T., Muller, H. G., and Mammitsch, V., “Kernels for Nonparametric Curve Estimation,” *Journal of the Royal Statistical Society Series B-Methodological* **47**(2), 238–252 (1985).

[14] Messer, K. and Goldstein, L., “A New Class of Kernels for Nonparametric Curve Estimation,” *The Annals of Statistics* **21**, 179–195 (Mar. 1993).

[15] Lanotte, A. A., Gillon, M., Demory, B.-O., Fortney, J. J., Astudillo, N., Bonfils, X., Magain, P., Delfosse, X., Forveille, T., Lovis, C., Mayor, M., Neves, V., Pepe, F., Queloz, D., Santos, N., and Udry, S., “A global analysis of Spitzer and new HARPS data confirms the loneliness and metal-richness of GJ 436 b,” *Astronomy and Astrophysics* **572**, A73 (Dec. 2014).