Aperture Photometry Uncertainties assuming Priors and Correlated Noise

F. Masci, version 2.0, 10/06/2009

1. Summary

We derive a general formula for the noise variance in the flux of a source estimated from aperture photometry assuming (i) prior pixel-flux uncertainties are available for the image (e.g., computed *a priori* from a noise model), and (ii) noise is correlated amongst pixels. Correlated noise usually occurs in re-sampled and interpolated images (e.g., mosaics), with the degree of correlation depending on the size of the interpolation kernel. The smoothing kernel moves noise-power from high to low spatial frequencies and therefore needs to be recaptured to properly quantify the uncertainty in the flux summed over a region. Ignorance of correlated-noise will lead to an underestimate of the final aperture-flux uncertainty. The 1-σ uncertainty is given by the square root of the variance expression below. A formalism for estimating uncertainties when one does not have access to priors and is content on ignoring correlated-noise is described in:
http://web.ipac.caltech.edu/staff/fmasci/home/wise/ApPhotUncert.pdf

Important Pre-Check:

It is important that the prior pixel uncertainties are statistically compatible with the input image data on which photometry is being performed. One way to do this is to compare the uncertainties with the local RMS pixel-to-pixel fluctuation about the mean or median background level. It is important to ensure that the background region is stationary (spatially uniform) and free of contamination from outliers (including real sources), otherwise, some trimmed version or robust measure of the RMS must be used. Furthermore, it is strongly recommended that the background region used for this comparison has approximately uniform depth-of-coverage. The best place to perform this check is at the raw-image level *prior to mosaicking*. This will ensure that the depth-of-coverage is constant, i.e., unity. If inconsistencies are found between the data-derived (RMS) noise and prior uncertainties, one will need to rescale the pixel-uncertainties according to the ratio RMS/$<\sigma_i>$, where $<\sigma_i>$ is the mean or median pixel-uncertainty over the region of interest.

We first give the final result and define all quantities involved. The derivation is given below.

$$\boxed{\sigma_{src}^2 = F_{corrA}\sum_{i}^{N_A}\sigma_i^2 + F_{corrB}\,k\,\frac{N_A^2}{N_B}\,\sigma_{\overline{B}/pix}^2}$$

where

N_A = number of pixels in source aperture

N_B = number of pixels in background annulus

σ_i = prior pixel flux uncertainty for image, rescaled if necessary

F_{corrA} = correlated noise correction factor for variance in flux in source aperture

F_{corrB} = correlated noise correction factor for variance in flux in background annulus

$\overline{B}$ = estimated background per pixel in annulus (either mean or median):

set $k = 1$ if $\overline{B}$ = mean background/pixel

set $k = \pi/2$ if $\overline{B}$ = median background/pixel

set $k = 0$ if assume $\overline{B} = 0$ or if no background is subtracted

$\sigma_{\overline{B}/pix}^2$ = variance in sky background annulus in [image units]^2/pixel.

Can compute from square of RMS deviation from mean or median, and trimmed versions thereof. Can also approximate using robust estimators of scale:

$\approx \left[0.5(q_{0.84} - q_{0.16})\right]^2 \approx \left[(q_{0.5} - q_{0.16})\right]^2$ where the q are quantiles, or the MAD:

$\approx \left[1.4826\ median\left|p_i - median\{p_i\}\right|\right]^2$, the Median Absolute Deviation from the median

2. Derivation

The above formula is derived as follows. First, the equation for estimating the flux of a source using aperture photometry can be written:

$$F_{src} = F_{tot} - N_A\overline{B}, \tag{1}$$

where F_{tot} is the sum of all pixel fluxes f_i^A in the source aperture:

$$F_{tot} = \sum_{i}^{N_A} f_i^A, \tag{2}$$

and all other quantities were defined above. For the purpose of variance estimation, we assume that the sky-background per pixel is derived using a mean of all pixel fluxes in the annulus:

$$\overline{B} = \frac{1}{N_B}\sum_{i}^{N_B} f_i^B. \qquad (3)$$

Generalization to the median will be described below.

The noise-variance in the estimate from Eq. 1 can be derived using standard error propagation. Ignoring correlations *between pixels in the source aperture and background annulus* (since these are assumed to be well separated), we have:

$$\sigma_{src}^2 = \sigma_{tot}^2 + N_A^2\sigma_{\overline{B}}^2, \qquad (4)$$

The first term on the right is the variance in the total flux in the source aperture. Using Eq. 2, this can be written in terms of the variance in pixel i and the covariance between any two pixels (j, k) in the source aperture:

$$\sigma_{tot}^2 = \sum_{i}^{N_A}\sigma_{iA}^2 + 2\sum_{j}^{N_A}\sum_{k<j}^{N_A}\mathrm{cov}(j,k) \qquad (5)$$

$$\text{with the constraint}:$$

$$(x_j - x_k)^2 + (y_j - y_k)^2 \le D_{PRF}^2$$

Similarly, the variance in the mean background per pixel as estimated from Eq. 3 can be written:

$$\sigma_{\overline{B}}^2 = \frac{1}{N_B^2}\left[\sum_{i}^{N_B}\sigma_{iB}^2 + 2\sum_{j}^{N_B}\sum_{k<j}^{N_B}\mathrm{cov}(j,k)\right] \qquad (6)$$

$$\text{also with the constraint}:$$

$$(x_j - x_k)^2 + (y_j - y_k)^2 \le D_{PRF}^2$$

If the co-add or mosaic was constructed using an interpolation kernel represented by the detector Point Response Function (PRF), then a noise fluctuation in a detector pixel will affect all the co-add pixels in the PRF's domain after interpolation. Therefore, the maximum range over which co-add pixels can be correlated is determined by the maximum linear extent of the PRF, D_{PRF}. This is also called the “correlation length”.

Equations 5 and 6 can re-factored respectively as follows:

$$\sigma_{tot}^2 = F_{corrA} \sum_i^{N_A} \sigma_{iA}^2 \qquad (7)$$

$$\sigma_{\bar{B}}^2 = \frac{F_{corrB}}{N_B^2} \sum_i^{N_B} \sigma_{iB}^2 \qquad (8)$$

where :

$$F_{corrA} = 1 + \frac{2\sum_j^{N_A}\sum_{k<j}^{N_A} \mathrm{cov}(j,k)}{\sum_i^{N_A} \sigma_{iA}^2}; \qquad F_{corrB} = 1 + \frac{2\sum_j^{N_B}\sum_{k<j}^{N_B} \mathrm{cov}(j,k)}{\sum_i^{N_A} \sigma_{iB}^2} \qquad (9)$$

The correlated-noise correction factors in Eq. 9 are further discussed in section 3.

The noise variance in the source flux can then be written be combining Equations 4, 7 and 8:

$$\sigma_{src}^2 = F_{corrA} \sum_i^{N_A} \sigma_{iA}^2 + F_{corrB} \frac{N_A^2}{N_B^2} \sum_i^{N_B} \sigma_{iB}^2, \qquad (10)$$

One would *not* typically use the prior-uncertainties in the sky annulus (σ_{iB}) to estimate the variance contributed by the background. This is because the annulus may be contaminated by sources and other outliers, whose effect is to inflate the σ_{iB} priors. Assuming the background is *stationary* within the annulus, the second summation term in Eq. 10 can be replaced by $N_B\sigma^2_{B/pix}$, where $\sigma^2_{B/pix}$ is the variance in the sky estimate, e.g., as estimated from a histogram of the N_B pixel values. Therefore, if $\bar{B}$ in Eq. 1 were estimated using an arithmetic mean (or a trimmed version thereof), its variance can be written:

$$\sigma_{\bar{B}=\mu}^2 = \frac{1}{N_B^2} \sum_i^{N_B} \sigma_{iB}^2 \approx \frac{\sigma_{B/pix}^2}{N_B}, \qquad (11)$$

If however a median was used for $\bar{B}$ in Eq. 1 and the pixel values are normally distributed, the variance as derived from Eq. 11 will be slightly underestimated by a factor of $\pi/2$. In other words, the median is noisier (*less efficient* in statistical parlance) than the mean for a randomly drawn sample. Nonetheless, given the robustness of the median against outliers, this is a small price to pay. A derivation of this "$\pi/2$ inflation" exists in each of the following references and was used to derive Equation 8 in the following paper:

http://web.ipac.caltech.edu/staff/fmasci/home/statistics_refs/MADstats.pdf

[2] B. L. van der Waerden, Mathematical Statistics, Springer, New York, 1969, section 17.
[3] S. S. Wilks, Mathematical Statistics, Wiley, New York, 1962, section 9.6.

Therefore, under the assumption of normally distributed data (which is usually satisfied in the limit of large N_B with a 'well behaved' astronomical detector), the variance *in the median* sky value per pixel is given by:

$$\sigma^2_{\overline{B}=med} = \frac{\pi}{2}\frac{1}{N_B^2}\sum_i^{N_B}\sigma^2_{iB} \approx \frac{\pi}{2}\frac{\sigma^2_{B/pix}}{N_B}, \tag{12}$$

We can combine Equations 10, 11 and 12 into our final general expression:

$$\sigma^2_{src} = F_{corrA}\sum_i^{N_A}\sigma_i^2 + F_{corrB}\,k\frac{N_A^2}{N_B}\sigma^2_{\overline{B}/pix}, \tag{13}$$

where $k = 1$ corresponds to $\overline{B}$ estimated using an arithmetic mean, and $k = \pi/2$ is for $\overline{B}$ estimated using a median. Incidentally, if the source-flux estimate involved no sky-background subtraction, or the background is known to be negligible *a priori*, one can set $k = 0$.

3. Estimating F_{corrA} and F_{corrB}

F_{corrA} and F_{corrB} in Eq. 9 represent correction factors ≥ 1 to account for an increase in the variance due to correlations between pixels in the source-aperture and sky-annulus respectively. In general, the covariance between any two pixels in either the source-aperture or sky-annulus can be written: $\text{cov}(j, k) = \rho_{jk}\sigma_j\sigma_k$, where ρ_{jk} is the correlation coefficient. Note, $\rho_{jk} = 0$ for pixel separations $d_{jk} > D_{PRF}$ (the correlation length of the smoothing kernel), and $0 < \rho_{jk} \leq 1$ for $d_{jk} \leq D_{PRF}$.

The spatially uniform background-dominated case

For background-limited observations, i.e., where flux in the source-aperture is dominated by background photons, the pixel variance is approximately stationary (spatially uniform) so that $\text{cov}(j, k) \approx \rho_{jk}\sigma^2_{iA}$. For the sky-annulus, we can also safely assume $\text{cov}(j, k) \approx \rho_{jk}\sigma^2_{iB}$. Furthermore, there is good reason to believe that pixel-to-pixel covariances are stationary over the aperture and annulus regions so that $\rho_{jk} \approx$ constant for *any* pixel pair j,k with fixed separation $d_{jk} \leq D_{PRF}$. Therefore, for the *background-photon dominated case*, either F_{corrA} or F_{corrB} in Eq. 9 can be reduced to:

$$F_{corr} \approx 1 + \frac{2}{N}\sum_j^N\sum_{k<j}^N \rho_{jk} \tag{14}$$

where $N = N_A$ or N_B, the number of source-aperture or sky-annulus pixels respectively. We expect (and simulations confirm it) that:

$$F_{corr} = \text{ constant} \approx N_{PRFeff} \text{ for } N \geq N_{PRFeff}, \tag{15}$$

where N_{PRFeff} is the *effective* number of pixels in the smoothing kernel, i.e., within some effective correlation length. This is not necessarily the total number of pixels spanned by the kernel, unless however the kernel is a top-hat (see below). We also expect:

$$F_{corr} \approx \alpha N \text{ for } N < N_{PRFeff}, \tag{16}$$

where α is a (non-trivial) constant of proportionality. Therefore, the correction factor is a linear function of the number of pixels in the aperture up to some effective $N \approx N_{PRFeff}$, and then levels off to some constant value N_{PRFeff} for $N \geq N_{PRFeff}$.

If one has no knowledge of the smoothing kernel, the correlation coefficient ρ_{jk} in Eq. 14 can be calibrated as a function of pixel separation d_{jk} from the image using some robust estimator of the autocorrelation function (ACF), preferably within stationary regions. An analytic function can then be fit to $\rho(d_{jk})$ for use in evaluating Eq. 14. If the smoothing kernel is known, $\rho(d_{jk})$ can be derived from the kernel directly, either numerically or analytically. It can be shown that for a pixelized *PRF* kernel where a value r_{ij} therein is defined as the response at pixel j when the *PRF* is centered on pixel i,

$$\rho_{jk} = \frac{\sum_{i=1}^{N_{PRF}} r_{ij} r_{ik}}{\sum_{i}^{N_{PRF}} r_i^2} = N_p \sum_{i=1}^{N_{PRF}} r_{ij} r_{ik}, \qquad (17)$$

where N_p is the 'infamous' *effective number of noise pixels* for the *PRF* kernel in question. The derivation shall be added in future. For an explanation of N_p, the reader is referred to: *http://web.ipac.caltech.edu/staff/fmasci/home/wise/noisepix_specs.pdf*. For a top-hat *PRF* volume-normalized to unity, the (constant) values are $r_{ij} = 1/N_{PRF}$, and it's not difficult to show from Eq. 17 that $N_p \equiv N_{PRF}$ and $\rho_{jk} = 1 \; \forall \; j,k$. Unless one wants to satisfy their mathematical curiosity, explicit calculation of F_{corr} using ρ_{jk} for any *PRF* in general is unnecessary as we shall show. The full derivation is deferred to a future paper.

Consider the simple case of a top-hat *PRF* as used when interpolating detector pixels onto a finer co-add grid using overlap-area weighting. Here we have $\rho_{jk} \approx 1$ over the span of the N_{PRF} co-add pixels overlapping with an input pixel. The double summation in Eq. 14 is just the number of distinct co-add pixel pairs in this span and evaluates to $N(N\text{-}1)/2$. Eq. 14 then simplifies to $F_{corr} \approx N = N_{PRF} = N_p$ (the number of 'noise pixels' as shown above)! This is also the resampling factor, i.e., the number of output (co-add) pixels per input pixel. For the general *PRF* case, it can also be shown (the full derivation is deferred) that the *effective* number of pixels in the smoothing kernel, i.e., N_{PRFeff} in Eq. 15, is $\approx N_p$. Hence in general, $F_{corr} = N_p$ as the aperture size increases beyond the effective correlation length of the kernel, i.e., contains $N > N_p$ pixels. It's important to note that N_p for the kernel should be measured in terms of the number of target image (co-add) pixels, not native detector pixels.

<u>*The source-photon dominated case*</u>

When flux in the source-aperture is dominated by actual photons from the source, the pixel-variance and covariances $\rho_{jk}\sigma_j\sigma_k$ (at fixed pixel-to-pixel separation) therein can no longer be assumed to be stationary since the Poisson variance generally follows the profile of the source. Eq. 14 will no longer hold, however, some workable approximation based on averaging correlations within the aperture is still possible. Nonetheless, it's comforting that simulations also show that $F_{corr} \approx N_p$ is a good approximation in the large aperture limit.

Consequently, for the source-photon dominated case, F_{corrA} is expected to depend on the relative contribution of source-to-background flux in the source aperture. We have parameterized this using the parameter R_{sb}. In the notation of Equations 1-3, this is defined as:

$$R_{sb} = \frac{F_{tot}}{N_A\overline{B}} = \frac{N_B}{N_A}\frac{\sum_i^{N_A} f_i^A}{\sum_i^{N_B} f_i^B}. \qquad (18)$$

To account for the complications noted above: (i) estimating correction factors for the source-photon dominated (spatially non-uniform) case, and (ii) the dependence of correction factors for small apertures with < N_{PRFeff} pixels (i.e., α in Eq. 16), we have resorted to a Monte-Carlo simulation. Here are the simulation steps used to estimate F_{corrA} and F_{corrB}:

1. Create a test 'truth' detector image containing some background $\overline{B}$ per pixel and a point source spike with flux F_{src} in the middle;
2. Convolve this truth image with the detector PRF, volume-normalized to unity;
3. Add Poisson noise to the detector image. We assume we're in the Gaussian limit and sample our pixel errors ε_i from a normal distribution: $\varepsilon_i \sim N(0, \sigma^2 = p_i)$ for pixel value p_i;
4. Interpolate the detector image to a new grid (i.e., the "co-add" image grid) using your favorite interpolation kernel. For WISE this will be the detector PRF as implemented in AWAIC [see: *http://web.ipac.caltech.edu/staff/fmasci/home/wise/awaic_adass08.pdf*];
5. Compute and store values of F_A ($\equiv F_{src}$), F_B ($\equiv N_B\overline{B}$), V_A, and V_B, defined as respectively:

   $$F_A = \sum_i^{N_A} f_i^A; \quad F_B = \sum_i^{N_B} f_i^B; \quad V_A = \sum_i^{N_A} \sigma_{iA}^2; \quad V_B = \sum_i^{N_B} \sigma_{iB}^2;$$

   V_A and V_B represent the sum of squares of the input prior 1-sigma uncertainties.
6. Go back to step 3 and re-simulate a new realization of Poisson distributed noise, saving the values from step 5 at each trial;
7. After 500-1000 noise realizations, compute the variance in F_A and F_B over all N_t trials via:

   $$\sigma^2(F_A) = \frac{1}{N_t - 1}\sum_{m=1}^{N_t}\left(F_A^m - \left\langle F_A^m \right\rangle\right)^2$$
   $$\sigma^2(F_B) = \frac{1}{N_t - 1}\sum_{m=1}^{N_t}\left(F_B^m - \left\langle F_B^m \right\rangle\right)^2$$

8. The source-aperture and sky-annulus correction factors are computed from the ratios:

   $$F_{corrA} = \frac{\sigma^2(F_A)}{V_A}$$
   $$F_{corrB} = \frac{\sigma^2(F_B)}{V_B}$$

So how do we know that the values for F_{corrA} and F_{corrB} computed in this manner are correct? The trick is that we also compute the variance in the source flux over all trials, $\sigma^2(F_{src})$, where F_{src} is estimated at each trial (noise realization) using Eq. 1. When this is compared to the predicted variance using Eq. 13, values of F_{corrA} and F_{corrB} as precisely computed above are needed for consistency. F_{corrA} and F_{corrB} will be computed for each band-dependent PRF and a range of N_A, N_B, and R_{sb} values. R_{sb} is defined by Eq. 18 and is only applicable to F_{corrA}. A user performing aperture photometry can then select (or approximate) the appropriate F_{corrA} and F_{corrB} to use in the boxed equation in section 1.

We also note that F_{corrA} (as well as F_{corrB}) can be computed (or verified) directly from the post-smoothed image by throwing many apertures at random, retaining those apertures which fall within regions with a spatially-uniform, source-free (and confusion-free) background, computing their summed-flux variance as in step 7 above, and then comparing to the pixel variance, V_A. The difficultly here is having enough random samples that are not significantly skewed by contaminating sources, confusion, and/or a varying background.

4. Example using a WISE test PRF

The example below is for a band-1 WISE *PRF* (FWHM ~ 6″) used in creating an interpolated (co-add) image with resampling factor of 4 (= number of output co-add pixels per input native pixel with linear scale 2.75″ arcsec). Note, the *PRF* used here dates back to July 2008. Results are shown for two test source-apertures and one sky annulus.

Ap. Radius (coadd pix)	Rsb (Eq. 18)	FcorrA
6	0.99993001330583	27.0115684322831
6	1.23171885262548	28.3402138075576
6	1.46350319270776	29.177824695401
6	1.69528559355028	29.7575425004733
6	1.92706696650331	30.182409774697
6	2.85418713529996	31.1364509404322
6	3.78130296465519	31.5880144532236
6	7.48974884015606	32.2096004158001
6	14.9066046154863	32.4674530971021
6	29.7402667820869	32.5596714958217
6	59.4074963888088	32.5798367932639
6	118.741845827015	32.5714647056869
12	0.999822875462531	29.7736948915878
12	1.25630020660288	30.937000934271
12	1.51277377621095	31.7747186767907
12	1.76924612740915	32.3876481487989
12	2.02571799730685	32.8537565183179
12	3.05160093677735	33.9616488259818
12	4.0774798439602	34.5275668375048
12	8.18097952895852	35.3987282525947
12	16.3879429638275	35.8529450916661
12	32.8018249605264	36.0914961279744
12	65.6295199442466	36.2182614825564
12	131.284829080113	36.2867436648321

```
annulus correlated-noise correction factor for 36 -> 51 co-add pixels:
FcorrB = 31.8995915784158
```

For comparison, the effective number of 'noise pixels' for this *PRF* (in number of co-add pixels) is $N_{PRFeff} \approx 34.261$, consistent with Eq. 15.