Computing ‘Number of Noise Pixels’ from Data

F. Masci, version 1.0 (2/26/2009)

1. Introduction

We derive a general expression for the “effective number of noise pixels” N_p that can be used on any Point Response Function (PRF) image. It is assumed the PRF image is made *a priori* by combining and re-sampling a number of *background-subtracted* point source images. The better the PRF sampling, the more accurate the estimate of N_p.

In a nutshell, think of N_p as the effective number of pixels contributing to the flux-variance of a point source. Sometimes this is referred to as the number of pixels *per beam*. The bigger this number, the broader the PRF (hence the ‘poorer’ the image quality - resolution-wise), and the lower the sensitivity overall, i.e., sources will exhibit bigger variance on repeated observation. This should become clearer from the derivation below (I hope). The final result is given by Equation 10 (boxed). I’ve never been a fan of this concept since it glosses over some details of the Poisson-noise dependence across the profile of a point source. Life can still proceed without it. In the end, it’s just a number that’s more-or-less proportional to the *square* of the PRF Full-Width at Half-Max (FWHM), e.g., see Equation 13 for the simple case of a Gaussian PRF.

2. Derivation

The signal from a source with true flux F as measured in detector pixel i is given by:

$$D_i = r_i F, \qquad (1)$$

where r_i is the PRF volume normalized to unity: $\Sigma_i r_i = 1$. The r_i are usually estimated by combining point source profiles on an upsampled grid. For the derivation below, we assume the r_i were estimated on a grid with the same pixel size as the detector pixels. We will generalize to different pixel sizes for the r_i and D_i below. Furthermore, if R_i is the *un-normalized* point source response,

$$r_i = \frac{R_i}{\sum_i^N R_i}. \qquad (2)$$

The true source flux F can be estimated from an un-weighted linear least-squares by minimizing the cost function:

$$L = \sum_i^N \left(D_i - r_i F\right)^2. \qquad (3)$$

Differentiating Eq. 3 with respect to F and setting to zero, the least-squares solution is given by:

$$F = \frac{\sum_i^N r_i D_i}{\sum_i^N r_i^2}. \tag{4}$$

The noise-variance in F can be derived by adding errors to the true values in Eq. 4:

$$F + \varepsilon_F = \frac{\sum_i^N r_i \left(D_i + \varepsilon_i\right)}{\sum_i^N r_i^2}. \tag{5}$$

Subtracting true values, squaring, taking expectation values, we have:

$$\left\langle \varepsilon_F^2 \right\rangle = \frac{\sum_i^N r_i^2 \left\langle \varepsilon_i^2 \right\rangle}{\left[\sum_i^N r_i^2\right]^2}, \tag{6}$$

where we assumed the detector pixels are *uncorrelated*: $<\varepsilon_i \varepsilon_j> = 0$ for $i \neq j$. The noise-variance in the flux estimate F can then be written:

$$\sigma_F^2 = \frac{\sum_i^N r_i^2 \sigma_i^2}{\left[\sum_i^N r_i^2\right]^2}, \tag{7}$$

Now for the approximation that makes me uncomfortable. We assume that the pixel variance σ^2_i across a source is approximately constant and there exists some effective variance σ^2_{eff} such that:

$$\sigma_F^2 \approx \frac{\sigma_{eff}^2}{\sum_i^N r_i^2}. \tag{8}$$

A constant variance is only true when the counts are dominated by a spatially uniform *background* and not when the counts are source-photon dominated. For the latter case, the σ^2_{eff} must correspond to some average or intermediate value of the pixel variance across the source. From Eq. 8, the number of noise pixels N_p is identified with

$$N_p = \frac{1}{\sum_i^N r_i^2}, \tag{9}$$

so that $\sigma^2_F \approx N_p\, \sigma^2_{eff}$.

We include two generalizations to Eq. 9: first, the PRF values r_i may be available in un-normalized form, e.g., as the R_i in Eq. 2; second, they may have been estimated on a much finer grid than the detector pixels. We would like to know the value of N_p for the detector pixels, not the PRF pixels. Using Eq. 2 and the fact that the PRF and detector pixel scales can be different, the number of noise pixels can be written:

$$N_p = \frac{\left[\sum_i^N R_i\right]^2}{s \sum_i^N R_i^2}, \qquad (10)$$

where s accounts for differences between the PRF (R_i) pixel size and the native detector pixel size. In general, this is the ratio of the native detector pixel area to PRF pixel area:

$$s = \frac{\mathrm{CDELT1_D\ CDELT2_D}}{\mathrm{CDELT1_{PRF}\ CDELT2_{PRF}}}, \qquad (11)$$

where the CDELT1, CDELT2 refer to standard FITS header keywords for pixel scales along the *X* and *Y* axes respectively.

As a simple illustration, let's assume $s = 1$ and we have an un-normalized *top-hat* PRF spread over *N* pixels with constant value *c*. Equation 10 then gives:

$$N_p = \frac{\left[\sum_i^N c\right]^2}{s \sum_i^N c^2} = \frac{N^2 c^2}{N c^2} = N. \qquad (12)$$

Therefore, the number of noise pixels is exactly equal to the number of pixels in a top-hat PRF. For a PRF with tails that decay fast enough, Eq. 10 will converge to some effective value that's characteristic of the PRF. For a Gaussian PRF, it is not difficult to show that

$$N_p \approx 2.226(\mathrm{FWHM/pixels})^2. \qquad (13)$$

3. Testing

A program was written to test Eq. 10 on WISE PRFs modeled by Ned Wright using lab data (May 2008). These are currently used for profile fit photometry in the WSDS. The program is available from the WISE processing environment and the synopsis is as follows:

```
caustic% noisepix -help2

parameters:

-in                          : string
# Required input PRF FITS file
```

```
-n1                      : double          = 1
# Scale factor for NAXIS1 dimension: detector pix X-scale/PRF X-scale

-n2                      : double          = 1
# Scale factor for NAXIS2 dimension: detector pix Y-scale/PRF Y-scale
```

Example runs on band 1, 2, 3 & 4 PRFs:

```
noisepix -in $inpdir/simcal-w1-psf-wpro-01x01-01x01.fits -n1 16 -n2 16
noisepix -in $inpdir/simcal-w2-psf-wpro-01x01-01x01.fits -n1 16 -n2 16
noisepix -in $inpdir/simcal-w3-psf-wpro-01x01-01x01.fits -n1 16 -n2 16
noisepix -in $inpdir/simcal-w4-psf-wpro-01x01-01x01.fits -n1 16 -n2 16
```

A comparison of the N_p predictions from Eq. 10 using the WSDS PRFs and independent optical modeling by Mark Larsen (SDL) in June 2008 is shown below. The results are in very good agreement.

Source	W1	W2	W3	W4
Equation 10	11.649	14.564	42.124	28.778
Mark Larsen (Jun'08)	11.4 → 14.7	14.6 → 18.3	39.0 → 45.1	27.3 → 28.9