Uncertainty Estimation in MCM Images

John Fowler

May 19 2008

Notation:

	r_{ij}	response of the i^{th} measurement at the j^{th} pixel
	D_i	data value ("flux") of the i^{th} measurement
	f_j	"flux" value of the j^{th} pixel
	σ_i	one-sigma uncertainty in D_i
	σ_j	one-sigma uncertainty in f_j
	σ_{ij}	one-sigma uncertainty of error in f_j due to D_i
	w_{ij}	weighting factor for D_i at pixel j

Two types of uncertainty will be considered: uncertainty based on priors (input uncertainties σ_i) and uncertainty derived from the data. The former allows hypothesis testing based on chi-squares, while the latter does not, because chi-squares based on noise derived from the data are tautological.

The weighting factors may be defined with or without inverse-variance factors. In the former case,

$$w_{ij} = \frac{\dfrac{r_{ij}}{\sigma_i^2}}{\sum_i \dfrac{r_{ij}}{\sigma_i^2}} \tag{1}$$

In the latter case,

$$w_{ij} = \frac{r_{ij}}{\sum_i r_{ij}} \tag{2}$$

The AWAIC SDS contains a description of how uncertainty is computed in a single-iteration image using priors, based on the averaging process:

$$f_j = \sum_i w_{ij} D_i \tag{3}$$

In the usual "true-plus-error" notation,

$$\hat{f}_j + \varepsilon_j = \sum_i w_{ij} \left(\hat{D}_i + \varepsilon_i \right) \tag{4}$$

Subtracting true values, squaring, and taking expectation values yields the following.

$$\begin{aligned} \varepsilon_j &= \sum_i w_{ij} \varepsilon_i \\ \sigma_j^2 = \left\langle \varepsilon_j^2 \right\rangle &= \left(\sum_i w_{ij} \varepsilon_i \right)^2 \end{aligned} \tag{5}$$

If the input errors are uncorrelated, this just yields

$$\sigma_j^2 = \sum_i w_{ij}^2 \sigma_i^2 \tag{6}$$

This is equivalent to Equation 11 in the AWAIC SDS. The PRF tends to correlate errors in neighboring measurements, however, so this may not be an acceptable approximation, and the cross-terms in Equation 5 may have to be kept and used with a covariance matrix or some local-correlation approximation.

If more than one iteration is performed, the accumulation of multiplicative correction factors that move flux around in the image makes detailed bookkeeping effectively intractable. The following approximation was used in the IRAS MCM program (Aumann, Fowler, and Melnyk, 1990, Astronomical Journal Vol. 99 No.5). The contribution of the i^{th} measurement to the error in pixel j is taken to be the total error in the i^{th} measurement scaled by the ratio of the flux in pixel j to the total predicted flux for the i^{th} measurement:

$$\begin{aligned} \sigma_{ij} &= \sigma_i \frac{f_j}{F_i} \\ F_i &\equiv \sum_j r_{ij} f_j \end{aligned} \tag{7}$$

Here F_i is the flux predicted for D_i from the image. The summation of the error whose uncertainty is given by the first line in Equation 7 over all pixels in the domain of the PRF of the i^{th} measurement must yield the total error in the i^{th} measurement, because all these errors came from that one measurement and hence are 100% correlated. This summation is defined identically to that for the flux itself on the second line of Equation 7, and so

$$\sum_j r_{ij} \sigma_{ij} = \sum_j r_{ij} \left(\sigma_i \frac{f_j}{F_i} \right) = \frac{\sigma_i}{F_i} \sum_j r_{ij} f_j = \sigma_i \tag{8}$$

The approximation involves an assumption that may be pictured as follows: if pixel j has a flux about equal to the total predicted for measurement i, then it probably got its flux from measurement i, and therefore the error contributed by measurement i is approximately the full error in measurement i. This should be a good approximation, because if a point source is involved and pixel j can account for the flux in measurement i all by itself, it must be surrounded by pixels of much lower flux, and it is not plausible that it received its flux from some measurement other than i. On the other hand, if extended emission is involved, it does not matter very much which measurements contributed to which pixels, since the errors in Equation 7 are all about equal anyway. In both cases, Equation 8 shows that flux error is conserved.

MCM enforces a requirement that a measurement's effect on a pixel must be in proportion to the corresponding response function, and vice versa. The expression for the reduced variance (assuming the Gaussian approximation) is

$$\frac{1}{\sigma_j^2} = \sum_i \frac{r_{ij}}{\sigma_{ij}^2} \tag{9}$$

where measurement error correlations are again assumed negligible but may have to taken into account. Equation 9 is patterned after the ordinary inverse-variance weighted Gaussian uncertainty for a scalar average, but with measurements that are less coupled through the response function contributing less reduction to the uncertainty of the averaged scalar.

An approximate expression for the uncertainty derived from the data without use of priors may be constructed as follows. The correction factor variance (CFV) in MCM is:

$$V_j = \sum_i w_{ij} \left(\frac{D_i}{F_i} \right)^2 - \left(\sum_i w_{ij} \frac{D_i}{F_i} \right)^2 \tag{10}$$

If no flux bias is in effect for the imaging, then V_j is effectively the inverse of the square of the signal-to-noise ratio *S/N* in pixel *j*. The error variance in pixel *j* is therefore

$$(\sigma_{RMS_j})^2 = f_j^2 V_j \tag{11}$$

Taking the uncertainty variance *in* the mean pixel flux (as opposed to *about* the mean pixel flux) to be the error variance above divided by the "number of measurements", where the latter is the "coverage", i.e., the sum of the response at pixel *j* over all measurements *i*, we obtain the alternative (data-derived) estimate for the one-sigma uncertainty in f_j :

$$\sigma'_j = f_j \sqrt{\frac{V_j}{\sum_i r_{ij}}} \tag{12}$$

Comparison of the results from Equations 12 and 9 should yield some insight into the accuracy of the model (prior uncertainties, error correlation, PRFs, etc.) similar to a chi-square analysis.

If a flux bias *was* used in the imaging, then depending on whether a positive or negative bias was used, the CFV is an optimistic or conservative indicator of *S/N*. For example, if a flux bias was subtracted from the data, then there is really more signal involved than is apparent in the image, and the S/N is really a bit better than $1/\sqrt{V_j}$. If a flux bias was added to the data, then there really is not as much signal as there seems to be. An approximate adjustment based on known bias relative to mean flux should be possible without destroying the approximation represented by Equation 12.