

A Maximum Likelihood Method to Improve Faint-Source Flux and Color Estimates

DAVID W. HOGG^{1,2} AND EDWIN L. TURNER^{2,3}

Received 1997 November 12; accepted 1998 February 2

ABSTRACT. Flux estimates for faint sources or transients are systematically biased high because there are far more truly faint sources than bright. Corrections that account for this effect are presented as a function of signal-to-noise ratio and the (true) slope of the faint-source number-flux relation. The corrections depend on the source being originally identified in the image in which it is being photometered. If a source has been identified in other data, the corrections are different; a prescription for calculating the corrections is presented. Implications of these corrections for analyses of surveys are discussed; the most important is that sources identified at signal-to-noise ratios of 4 or less are practically useless.

1. INTRODUCTION

Given a noisy photometric measurement of a very faint source, what is the best estimate of its true flux?

The best answer to this question is “I don’t know—integrate longer to reduce the noise!” However, in some cases, this is not possible. For example, the ultradeep images of the Hubble Deep Field (HDF) (Williams et al. 1996) represent so much *Hubble Space Telescope* (HST) observing time that in practice they cannot be much improved. From the ground they cannot even in principle be improved because any ground-based images significantly deeper than existing ones (see, e.g., Djorgovski et al. 1995; Metcalfe et al. 1995; Smail et al. 1995; Hogg et al. 1997a, 1997b) would be totally confusion-limited (see, e.g., Condon 1974). As another example, observations of transients, such as gamma-ray bursts or supernovae, cannot be improved because they cannot be repeated, even in principle.

Given that in some cases deeper imaging is not an option, the reason the question does not have a trivial answer is that the number counts of faint sources tend to rise with decreasing flux, so more sources are available for “upscattering” to a given measurement than are available for “downscattering.” A familiar analogy is with trigonometric parallaxes, where low signal-to-noise ratio measurements are biased large, since given any observed parallax π_o and associated error, there is a finite range of true parallaxes π consistent with it, but there are far more sources in the sky with small parallaxes $\pi < \pi_o$ than large $\pi > \pi_o$. The “Lutz-Kelker” corrections that account for this are easy to compute and apply (Lutz & Kelker 1973; Hansen 1979);

they have been essential in providing unbiased distances in astronomy.

There is a conceptually similar set of corrections for low signal-to-noise ratio measurements of faint-source fluxes. In this article, these corrections are computed and discussed. As in the case of parallaxes, the corrections depend on how the sources are selected and on the intrinsic distribution of the measured quantity, in this case the true number-flux relation. Unfortunately, the number-flux relation is not exactly known in most cases of interest, since the faint-source photometry is usually being performed in order to determine this very relation! Furthermore, in many cases of interest, the correction presented here does not represent the largest source of systematic error. However, unlike the other sources of error, this correction applies to all flux-selected sources, independent of instrumentation or analysis technique.

One note of terminology: the systematic bias discussed here is often improperly referred to as the “Malmquist bias.” The Malmquist bias is the effect that in a flux-limited sample, there is a larger than representative fraction of high-luminosity sources because they can be seen to greater distances and hence over a larger volume (Malmquist 1924; Mihalas & Binney 1981). It is because of the *intrinsic* scatter in source luminosities. Malmquist bias is removed, e.g., when one computes a luminosity function from star counts. It does not involve any kind of measurement error; it does not go away if one obtains more precise photometry! The bias corrected for here results from the *observational* scatter in fluxes; the measurement errors. It does indeed go away when the fluxes are remeasured at much higher precision; it only needs to be considered when low signal-to-noise data are being used. What is discussed in this paper is closely related to Eddington bias, the effect of low signal-to-noise flux measurements on faint-source number-magnitude relations. Statistical corrections to observed number-magnitude relations are computed by Eddington (1913); flux corrections for individual survey sources are computed here.

¹ Theoretical Astrophysics, California Institute of Technology.

² Institute for Advanced Study, Olden Lane, Princeton, NJ 08540; hogg@ias.edu.

³ Princeton University Observatory, Princeton University, Peyton Hall, Princeton, NJ 08544.

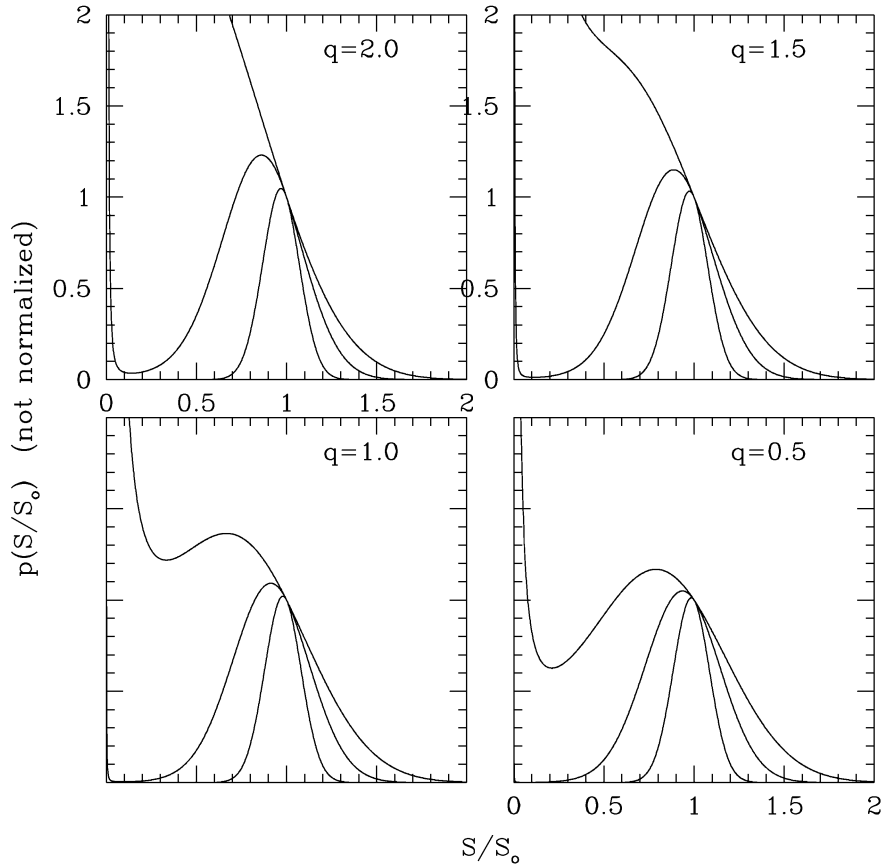


FIG. 1.—Likelihood curves for several number-magnitude exponents q and signal-to-noise ratios $r = 3$ (worst), 5 (middle), and 10 (best)

2. COMPUTATION OF CORRECTIONS

Consider the simplest case, in which a source is being photometered in the image in which it was first detected. That is, it is being measured in the data in which it was selected. The likelihood $p(S|S_o)$ (probability per unit flux) that a source has true flux S given that it is observed to have flux S_o is related to the likelihood $p(S_o|S)$ that it is observed to have S_o when it has S by Bayes's theorem:

$$p(S|S_o) \propto p(S_o|S)p(S), \quad (1)$$

where a proportionality is used because the normalization is being ignored (for now) and $p(S)$ (probability per unit flux) is the true underlying distribution of fluxes, given by the (true, not observed) number-flux relation. If the number of sources $N (< m)$ brighter than magnitude m as a function of m is a power

law,

$$2.5 \frac{d \log N}{dm} = - \frac{d \log N}{d \log S} = q, \quad (2)$$

then the conditional probability becomes

$$p(S|S_o) \propto \left(\frac{S}{S_o}\right)^{-(q+1)} \exp\left[-\frac{(S-S_o)^2}{2\sigma^2}\right], \quad (3)$$

where it is assumed that the observational error is Gaussian-distributed and σ is the uncertainty in the observed flux S_o , or S_o/σ is the signal-to-noise ratio r . Figure 1 shows these likelihood curves for number-flux exponent $q = 2.0, 1.5, 1.0,$ and 0.5 and signal-to-noise ratios $r = 3, 5,$ and 10 . This figure demonstrates that measurements at a signal-to-noise ratio of 3 do not strongly constrain the true flux, whatever the slope of the number counts, but particularly if the counts have the Eu-

clidean⁴ slope of $q = 1.5$ (or greater). It is worth emphasizing that the above equation and the curves plotted in Figure 1 are essentially identical to those computed for parallax corrections (Lutz & Kelker 1973; Hansen 1979), except that the parallax corrections are computed for only one particular exponent value.

If the flux measurement was unbiased, the peak in the likelihood function $p(S|S_o)$ would be at $S/S_o = 1$. However, taking the derivative dp/dS , it is found that the maximum likelihood true flux S_{ML} is in fact

$$\frac{S_{ML}}{S_o} = \frac{1}{2} + \frac{1}{2} \sqrt{1 - \frac{4q + 4}{r^2}}, \quad (4)$$

where q is the number-magnitude exponent defined above and r is the signal-to-noise ratio. There is no finite maximum likelihood value at all if $r^2 < 4q + 4$; an example is the $q = 1.5$, $r = 3$ curve in Figure 1. *The above equation specifies a correction that should in principle be applied to all flux measurements in a flux-limited sample.* When the signal-to-noise is good enough ($r^2 \gg 4q + 4$) the correction can be approximated as

$$\frac{S_{ML}}{S_o} \approx 1 - \frac{q + 1}{r^2} \quad \text{when } r^2 \gg 4q + 4, \quad (5)$$

or in terms of the magnitude correction $\Delta m \equiv m_{ML} - m_o$,

$$\Delta m \approx \frac{1.086q + 1.086}{r^2} \quad \text{when } r^2 \gg 4q + 4. \quad (6)$$

Things change slightly if the likelihood is computed in the magnitude (i.e., log flux rather than flux) domain; after all, maximum likelihood techniques are sensitive to the “metric” of the space in which the likelihood is computed. When computed purely in the magnitude or log flux domain, the correction is

$$\Delta m \approx \frac{1.086q + 2.171}{r^2} \quad \text{when } r^2 \gg 4q + 8. \quad (7)$$

The fact that the solution depends on the type of data space (log or linear) demonstrates that the specific value of the correction is not completely specified, because it only provides a “best guess” (a subjective estimate) for the true flux. It is worthy of note that the Lutz-Kelker parallax corrections are similarly subjective, as are essentially all statistical estimators.

⁴ Another note on terminology: what is called the “Euclidean” slope really ought to be called the “no-evolution, nonexpanding” slope, because even in a Euclidean space, the number counts have $q \neq 1.5$ at large distance if either the universe is expanding or the sources are evolving.

More robust than maximum likelihood estimates are confidence intervals, because these do not depend on the choice of “metric.” Confidence intervals are found by integrating the likelihood curves. Unfortunately, the areas under the curves shown in Figure 1 do not converge; the likelihood distributions are not normalizable! This nonnormalizability comes from the divergence of $p(S|S_o)$ as $S \rightarrow 0$ (not visible in some of the curves in Fig. 1 simply because at high r the divergence happens at very small S/S_o). There are two respects in which this divergence or nonnormalizability is unphysical: first, there cannot be an infinite number of sources in the visible universe; there are not even an infinite number of *atoms* in the universe! Second, most ultradeep images of the sky, including the HDF, are close to their confusion limits, beyond which the observed number counts have to “cut off” no matter how much integration time is employed. Neither of these effects can be simply taken into account in general; they depend on the data quality and the sources under study.

The equations in this section have assumed that observational errors are Gaussian-distributed, which is not true for all photometric measurements. The equations are easily generalized (although they do not necessarily remain analytic) with the Gaussian in equation (3) replaced by whatever error distribution is appropriate for the measurement in question.

3. AN EMPIRICAL TEST

The correction can be tested with any imaging data in which the number-flux relation is known. Here, the *HST* HDF data in the F606W (0.6 μm) bandpass are used. Noise was added to the 1024×1024 “Version 2” mosaics of the *HST* images of the HDF (Williams et al. 1996) to make the pixel-to-pixel sky noise 10 times as bad as in the original mosaics. The higher noise mosaics will be referred to as the “bad” images and the originals as the “good” images. A catalog of sources was chosen in the bad images down to very faint levels using the “SExtractor” source detection package (Bertin & Arnouts 1996) in essentially its default mode: smooth with a 2 pixel FWHM triangular filter and select sources whose central pixel in the smoothed image is above a given threshold. These sources were then photometered with the NOAO “IRAF” software in matched $0''.16$ (2 pixel) diameter apertures in both the bad and good images. The bad/good flux ratios are plotted against signal-to-noise ratio in Figure 2 along with the expected correction computed with equation (4) and the (known) count slope $q = 0.5$ (Williams et al. 1996). The correction does very well down to signal-to-noise ratios $r \approx 3$. At $r < 4$, a significant number of spurious (zero flux in the good image) sources start to appear. Figure 2 shows, as with the Lutz-Kelker corrections, that the corrections are on the same order as the intrinsic scatter due to measurement error, so some sources have underestimated rather than overestimated fluxes. However, the correction is

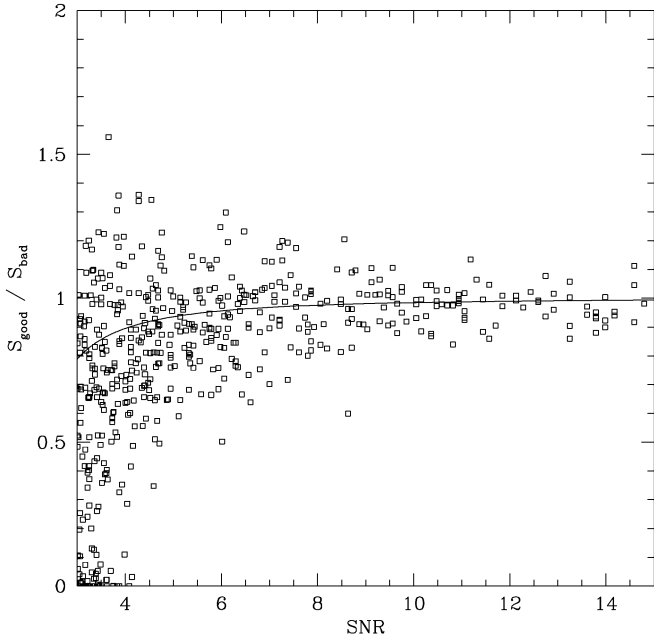


FIG. 2.—Ratios of flux measurements in the “bad” image, in which sources in the HDF were detected, to the flux measurements in the “good” image. The bad image is simply the good image plus additional noise (see text). The solid line is the expected ratio S_{ML}/S_o of the maximum likelihood flux to the observed flux given by eq. (4) for the number-magnitude exponent $q = 0.5$ (see text).

still necessary if an unbiased estimator of the true flux is desired.

4. CHANGING THE SELECTION TECHNIQUE

The next case to consider is photometry of a source in one image (say the I band) after it is detected (and its position is known) in another image (say the V band). In this case, Bayes’s theorem is still used, but for $p(S)$ the true distribution of $V - I$ colors is used rather than the I -band number counts. Actually, rather than the color distribution, it is better to think of the conditional I -band flux distribution $p(S^{(I)}|S^{(V)})$ (probability per unit flux), given that the V -band flux $S^{(V)}$ is known (in what follows, it is assumed that the V -band detection is at very high signal-to-noise ratio so the V -band flux is well known). Because, unlike the number counts, these conditional distributions are not generally power laws, the flux correction depends not only on the shape of the distribution but also on where in the distribution the observed flux S_o lies.

Fortunately, when the signal-to-noise ratio r is large enough, it is possible to linearize Bayes’s formula around the observed flux $S_o^{(I)}$ so only the local power-law slope

$$Q \equiv - \left. \frac{d \log p(S^{(I)}|S^{(V)})}{d \log S^{(I)}} \right|_{S_o^{(I)}} \quad (8)$$

of the distribution of source fluxes is important. The likelihood function (probability per unit flux) for the true I -band flux $S^{(I)}$ given the observed flux $S_o^{(I)}$ and the known V -band flux $S^{(V)}$ is then

$$p(S^{(I)}|S_o^{(I)}, S^{(V)}) \propto \left(\frac{S^{(I)}}{S_o^{(I)}} \right)^{-(Q+1)} \exp \left[- \frac{(S^{(I)} - S_o^{(I)})^2}{2\sigma^2} \right], \quad (9)$$

which leads to the maximum likelihood correction

$$\frac{S_{\text{ML}}^{(I)}}{S_o^{(I)}} \approx 1 - \frac{Q+1}{r^2} \quad \text{when } r^2 \gg 4Q+4, \quad (10)$$

or in terms of the magnitude correction $\Delta m \equiv m_{\text{ML}} - m_o$,

$$\Delta m \approx \frac{1.086Q + 1.086}{r^2} \quad \text{when } r^2 \gg 4Q+4. \quad (11)$$

Note that the correction can be positive or negative, depending on the sign of the local slope Q .

Corrections applicable when more complicated selection procedures have been used can be computed in analogous ways.

5. SUMMARY AND DISCUSSION

Maximum likelihood corrections for faint-source flux measurements have been computed for the case in which the sources are measured at low signal-to-noise in the data in which they were originally selected. It is found that since the number-flux relation tends to be rising at the faint end, the low signal-to-noise flux measurements are usually overestimates of the true flux. At signal-to-noise ratios $r < 4$, flux measurements (of this type—i.e., in the data in which the sources were selected) are almost meaningless because they are consistent with almost any true flux between zero and the measured value.

The bias considered here tends to “steepen” measured number-flux relations at the faint end; i.e., the measured $d \log N/d \log S$ is more negative than the true value because the very numerous faint sources are scattered up to brighter levels. This effect is only significant at very faint levels, where it is usually mitigated or in fact canceled out by incompleteness. The best way to correct measured number-flux relations for both the flux bias and incompleteness is to perform full completeness simulations, which, if done correctly, will account for both effects simultaneously (see, e.g., Smail et al. 1995; Hogg et al. 1997b), and of course a full accounting for all systematic errors requires detailed modeling of every stage in the observing and analysis procedures. Although the corrections presented here do not comprehensively account for most of these systematic biases, they are very general, improving flux estimates for individual sources independent of observational technique.

These corrections ought to be applied to the source fluxes at the faint end of the catalogs from all huge (and therefore difficult to improve upon) surveys, such as the Palomar Observatory Sky Surveys, the *Infrared Astronomical Satellite* (*IRAS*) survey, and from future huge surveys such as the 2 Micron All Sky Survey and the Sloan Digital Sky Survey. In fact, the *IRAS* catalogs were corrected at the faint end for some related biases but not this bias per se (*IRAS Explanatory Supplement* 1988). Also, all transients discovered at low signal-to-noise ratios in transient searches, such as faint gamma-ray

bursts, of which no additional measurements can be made after the fact, should have these corrections applied.

We thank Roger Blandford, John Gizis, Gerry Neugebauer, Neill Reid, and Yun Wang for useful discussions and the HDF team for planning, taking, reducing, and calibrating the HDF data. Some financial support was provided by NSF grant AST-9529170.

REFERENCES

- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
 Condon, J. J. 1974, *ApJ*, 188, 279
 Djorgovski, S., et al. 1995, *ApJ*, 438, L13
 Eddington, A. S. 1913, *MNRAS*, 73, 359
 Hansen, R. B. 1979, *MNRAS*, 186, 875
 Hogg, D. W., Neugebauer, G., Armus, L., Matthews, K., Pahre, M. A., Soifer, B. T., & Weinberger, A. J. 1997a, *AJ*, 113, 474
 Hogg, D. W., Pahre, M. A., McCarthy, J. K., Cohen, J. G., Blandford, R. D., Smail, I., & Soifer, B. T. 1997b, *MNRAS*, 288, 404
IRAS Catalogs and Atlases: Explanatory Supplement. 1988, ed. C. A. Beichman, G. Neugebauer, H. J. Habing, P. E. Clegg, & T. J. Chester (Washington: GPO)
- Lutz, T. E., & Kelker, D. H. 1973, *PASP*, 85, 573
 Malmquist, K. G. 1924, *Medd. Lund. Astron. Obs. Ser. II*, 32, 64
 Metcalfe, N., Shanks, T., Fong, R., & Roche, N. 1995, *MNRAS*, 273, 257
 Mihalas, D., & Binney, J. 1981, *Galactic Astronomy* (2d ed.; New York: Freeman)
 Smail, I., Hogg, D. W., Yan, L., & Cohen, J. G. 1995, *ApJ*, 449, L105
 Williams, R. E., et al. 1996, *AJ*, 112, 1335