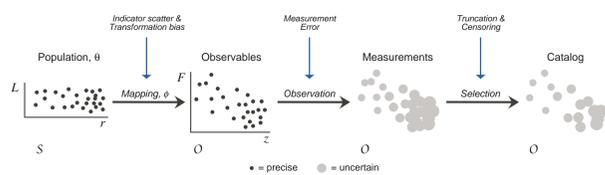


# Modeling GRB (and other) populations: Lessons from multilevel modeling

Tom Loredo and Ira Wasserman (Dept. of Astronomy, Cornell University)

## Survey Scenario



- **Source properties:** Each GRB has properties  $\mathcal{S}$  (e.g., peak luminosity, distance, direction, spectral parameters). The *population distribution* is  $f(\mathcal{S}; \theta)$  (e.g., luminosity function), with parameters  $\theta$ .
- **Observables:** Source properties are not directly observable. The observables  $\mathcal{O}$  (e.g., peak flux, redshift, direction, spectrum) are related to  $\mathcal{S}$  via some mapping  $\mathcal{S} \rightarrow \mathcal{O}$  (inverse-square law, cosmology, extinction), which may have unknown parameters  $\phi$ . This implies an *observable distribution*  $\Sigma(\mathcal{O}; \theta)$ .
- **Measurement error:** Observables must be estimated from analyses of “raw” survey data; estimates are *uncertain*. Uncertainties may be summarized, e.g., by  $\chi^2$  or likelihood contours.
- **Selection:** Measurements enter the final source catalog only if they meet detection and/or classification criteria (burst trigger criteria; solar flare rejection).

Primary goal of population modeling:

Learn about  $f(\mathcal{S}; \theta)$  or  $\Sigma(\mathcal{O}; \theta)$  from the survey catalog

(We may alternatively seek to infer the  $\mathcal{S} \rightarrow \mathcal{O}$  mapping, e.g., infer the cosmology.)

## Multilevel Modeling

*Inverse methods* try to undo the survey process by *correcting* catalog estimates. Models are assessed by comparison with corrected data.

*Forward methods* apply the survey process to candidate models to *predict* the catalog. Models with superior predictions are favored.

**Multilevel modeling:** A forward approach that handles measurement uncertainties by dividing the prediction process into two pieces:

- **Upper level:** Consider the sources’ true  $\mathcal{O}$  values to be drawn independently from  $\Sigma(\mathcal{O}; \theta)$ , adjusted by a *detection efficiency*  $\eta(\mathcal{O}) \rightarrow \{\mathcal{O}_i\}$  for  $i = 1$  to  $N$  (catalog size).
- **Lower level:** With  $\mathcal{O}_i$  given, model the data for each source,  $D_i$ , via independent sampling distributions,  $p(D_i|\mathcal{O}_i)$ . Considered as a function of  $\mathcal{O}_i$ , this defines a *source likelihood*,  $\ell_i(\mathcal{O}_i)$ , for each source (e.g.,  $\propto \exp[-\chi^2(\mathcal{O}_i)/2]$  for  $\chi^2$  fitting of observables).

**Parameter count:** The number of parameters in this model is

$$\nu = \nu_{\text{pop}} + \nu_{\text{obs}} \times N$$

# of population parameters in  $\theta$   
# of uncertain observables per source

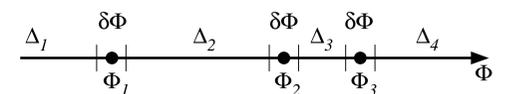
The number of parameters *grows with catalog size*; this makes accurate inference challenging when there is significant measurement error.

*Multilevel modeling is now seen as a framework uniting many previously distinct parts of statistics: measurement error models, shrinkage estimation, ridge regression, latent variable methods, and empirical and hierarchical Bayes methods.*

## Bayesian Multilevel Modeling

We can use a multilevel model to calculate the joint likelihood for all the population and source parameters, i.e., the probability for predicting all the data,  $D$ , if these parameters are known:  $\mathcal{L}(\theta, \{\mathcal{O}_i\}) \equiv p(D|\theta, \{\mathcal{O}_i\})$ .

For astronomical surveys, we adopt an *inhomogeneous Poisson point process* model. We divide the  $\mathcal{O}$  space into empty regions  $\Delta_j$ , and small regions  $\delta_j$  populated by a single source. The figure shows the construction for a single observable, peak flux  $\Phi$ .



The likelihood is found by multiplying independent Poisson probabilities for empty intervals, and intervals with a single event:

$$\mathcal{L}(\theta, \{\mathcal{O}_i\}) \propto \exp \left[ - \int d\mathcal{O} \eta(\mathcal{O}) \Sigma(\mathcal{O}; \theta) \right] \prod_{i=1}^N \ell_i(\mathcal{O}_i) \Sigma(\mathcal{O}_i; \theta)$$

In a Bayesian calculation, we summarize the information about the population parameters by *marginalizing* (integrating over) the uncertain source parameters:

$$\mathcal{L}(\theta) \propto \exp \left[ - \int d\mathcal{O} \eta(\mathcal{O}) \Sigma(\mathcal{O}; \theta) \right] \prod_{i=1}^N \int d\mathcal{O}_i \ell_i(\mathcal{O}_i) \Sigma(\mathcal{O}_i; \theta)$$

This is the *marginal likelihood for the population parameters*.

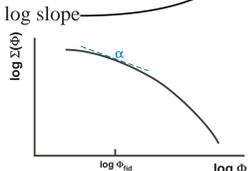
## Simulations

Simulate a (peak) flux survey catalog (producing a “number counts” or “log  $N$ -log  $S$ ” distribution):

- Draw source fluxes from a 3-parameter slowly rolling power-law distribution ( $\Phi_0$  is a fiducial flux):

$$\log \Sigma(\Phi) = \log \Sigma_0 - C_1 \alpha \log \frac{\Phi}{\Phi_0} - C_2 \alpha' \left( \log \frac{\Phi}{\Phi_0} \right)^2$$

Amplitude at  $\Phi_0$   
Log slope at  $\Phi_0$   
Rate of change of log slope



- For each source, simulate a measurement from a photon counting detector with Poisson distributed counts.
- Consider the measurement a detection only if the counts are above a fixed threshold.

Instrument parameters were chosen so the dimmest detected sources had  $\approx 15\%$  flux uncertainties.

Though clearly a “caricature” of real surveys, this setup allows us to generate *and* analyze *hundreds* of simulated survey catalogs relatively quickly.

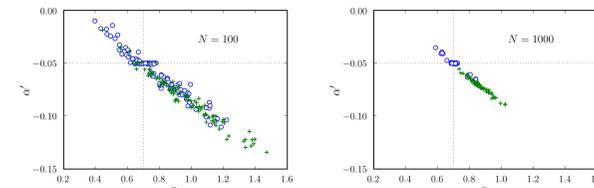
## Accounting For Source Uncertainties

### Ignoring Uncertainties Corrupts Inferences

An important virtue of *Bayesian* multilevel modeling is its ability to account for source uncertainties via marginalization.

It is tempting to argue that, as catalog size grows, the measurement errors should “average out,” so they may be ignored. One would then just plug in the best-fit estimates for the observables in  $\mathcal{L}(\theta, \{\mathcal{O}_i\})$ . The majority of published GRB population studies do this (or something similar).

We simulated 100 data sets as described above, and found maximum likelihood estimates for  $(\alpha, \alpha')$ , using both the marginal likelihood, and just plugging in the best-fit fluxes. The figure shows resulting estimates (blue circles = Bayesian, green crosses = plug-in), for catalogs of  $N = 100$  and 1000 bursts. (Crosshair shows true values.)



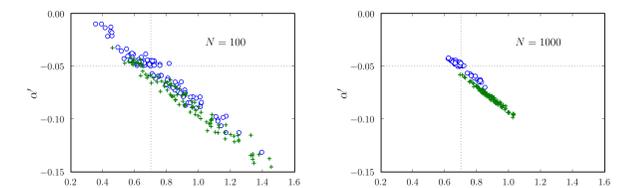
For  $N = 100$ , the Bayesian estimates are distributed roughly symmetrically about the truth; plug-in estimates are biased toward large  $\alpha$  and small  $\alpha'$ , though the uncertainties are large enough that plug-in is sometimes accurate.

For  $N = 1000$ , the Bayesian estimates have converged closer to the truth. The plug-in estimates have converged *away* from the truth.

### Inflating Uncertainties Corrupts Inferences

We have adapted this same framework to the analysis of other populations, e.g., trans-Neptunian objects (TNOs, including Kuiper Belt Objects, KBOs). There we have encountered the practice of surveyors sometimes reporting “inflated” uncertainties, in an effort to be “conservative.” Despite the laudable motivation of this practice, it corrupts inferences.

We repeated the previous exercise with a new set of simulations. But here the green crosses are estimates from a Bayesian calculation, only with uncertainties inflated by 15%.



Clearly, the uncertainties must be accurately calculated (and reported) to guarantee sound inferences.

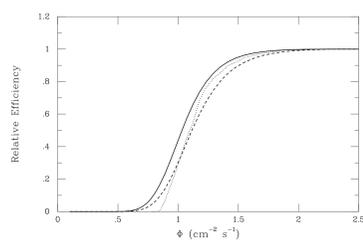
## A BATSE Issue: Detection Efficiency Accuracy

Accurate population modeling requires accurate efficiency functions for analyzed catalogs.

Calculating the survey efficiency is one of the most challenging tasks for surveyors, typically requiring extensive and detailed Monte Carlo simulation of the instrument and analysis pipeline.

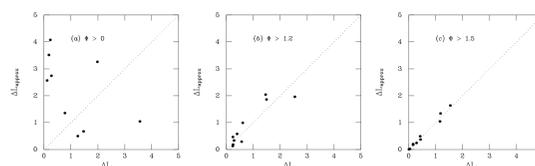
**Limitations of the BATSE efficiency:** For BATSE, two effects that would have significantly complicated the calculation were omitted: *atmospheric scattering*, and *counting uncertainties*.

We have simulated a significantly simplified BATSE mission, from photon counting and triggering through a multilevel analysis, drawing GRB peak fluxes from a power law distribution. This allows exploration of the importance of these two effects. The figure shows three different efficiency calculations: dotted—ignores scattering and counting uncertainties; dashed—incorporates counting uncertainties; solid—adds atmospheric scattering (true efficiency).



We find the faint end of the efficiency function is significantly altered by omitting either effect. Using an efficiency that ignores atmospheric scattering and counting uncertainties seriously corrupts inferences.

As an attempted remedy, we explored *thresholding* the catalog: only analyzing GRBs with fluxes  $> \Phi_{\text{th}}$ , and finding a threshold so the approximate and accurate likelihoods were close. We simulated several surveys, calculating the true and approximate likelihoods for bursts above a threshold, both for the true population parameters, and the resulting maximum-likelihood fit. Accurate inference requires getting the change in log-likelihood,  $\Delta L$ , correct to  $\ll 1$ . The figure shows scatterplots of the true and approximate  $\Delta L$ , for various threshold choices.



The threshold must be disappointingly high, corresponding to  $\eta \approx 0.95$ . Thus many faint bursts cannot be accurately modeled with tabulated efficiencies. This cuts the useable BATSE catalog size by nearly 1/2.

*The 4B catalog team recognized this problem and an improved algorithm was developed. Unfortunately work on a more accurate efficiency was not completed.*

## Nonparametric Methods?

Astronomers often use nonparametric (inverse) methods for analyzing survey data: Lynden-Bell’s  $C^-$  method, Efron-Petrosian estimators, step-wise maximum likelihood (SWML). Unfortunately, *measurement error greatly compromises the performance of nonparametric methods*.

The best-studied such approach—mixture deconvolution—has a depressingly slow asymptotic convergence rate: it is *logarithmic* in  $N$  for Gaussian errors (compare to  $\sqrt{N}$  for conventional parametric modeling).

Recent research finds promise in nonparametric Bayesian multilevel modeling, but the field is still young.

## Further Information

Loredo, T. J., and Wasserman, I. M. (1998) “Inferring the spatial and energy distribution of gamma-ray burst sources. II. Isotropic models,” *ApJ*, 502, 75

Loredo, T. J. (2007) “Analyzing data from astronomical Surveys: Issues and Directions,” in *Statistical Challenges in Modern Astronomy IV* (ed. J. Babu and E. Feigelson), ASP Conference Series, vol. 371, 121

Loredo, T. J., and Hendry, M. A. (2008) “Bayesian Multilevel Modelling of Cosmological Populations,” a chapter in *Bayesian Methods in Cosmology* (ed. Andrew Liddle et al.), Cambridge University Press, in press