Flash-Type Discrimination

GOES-AWG/R3 GLM Science Meeting
September 29-30, 2009
Huntsville, AL

Dr. William Koshak; NASA-MSFC

National Aeronautics and Space Administration

Very Brief Summary of Results From …

1. Koshak, W. J., Optical Characteristics of OTD Flashes and the Implications for Flash-Type Discrimination, to be submitted to JTECH, 2009.

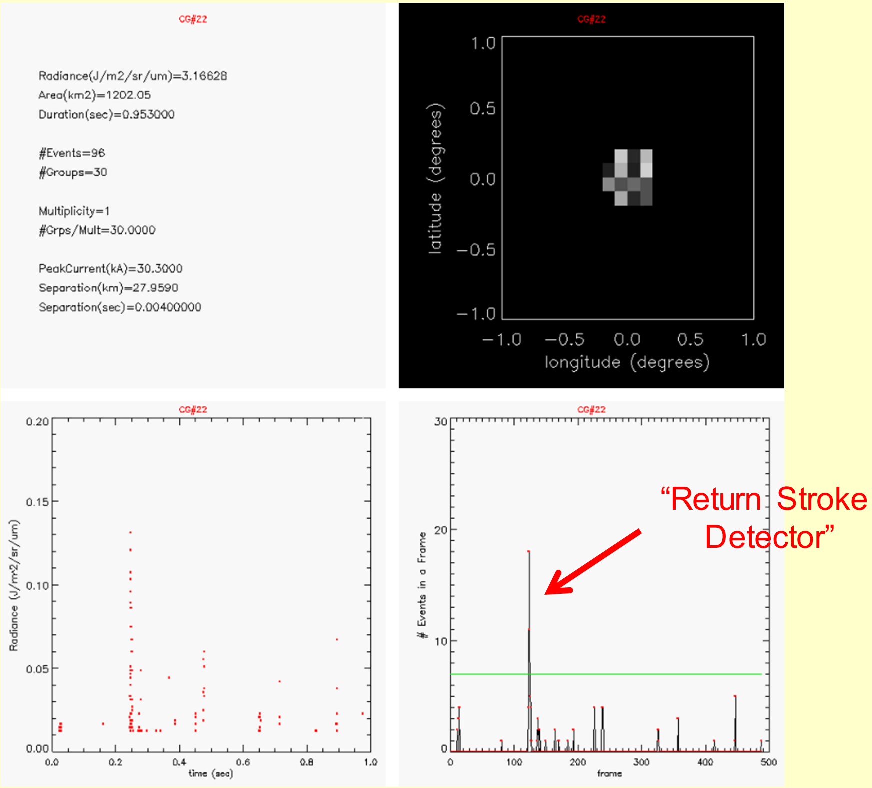
2. Koshak, W. J., R. J. Solakiewicz, Retrieving the Fraction of Ground Flashes from Satellite Lightning Imager Data Using CONUS-Based Optical Statistics, to be submitted to JTECH, 2009.

3. Koshak, W. J., A Mixed Exponential Distribution Model for Retrieving Ground Flash Fraction from Satellite Lightning Imager Data, to be submitted to JTECH, 2009.

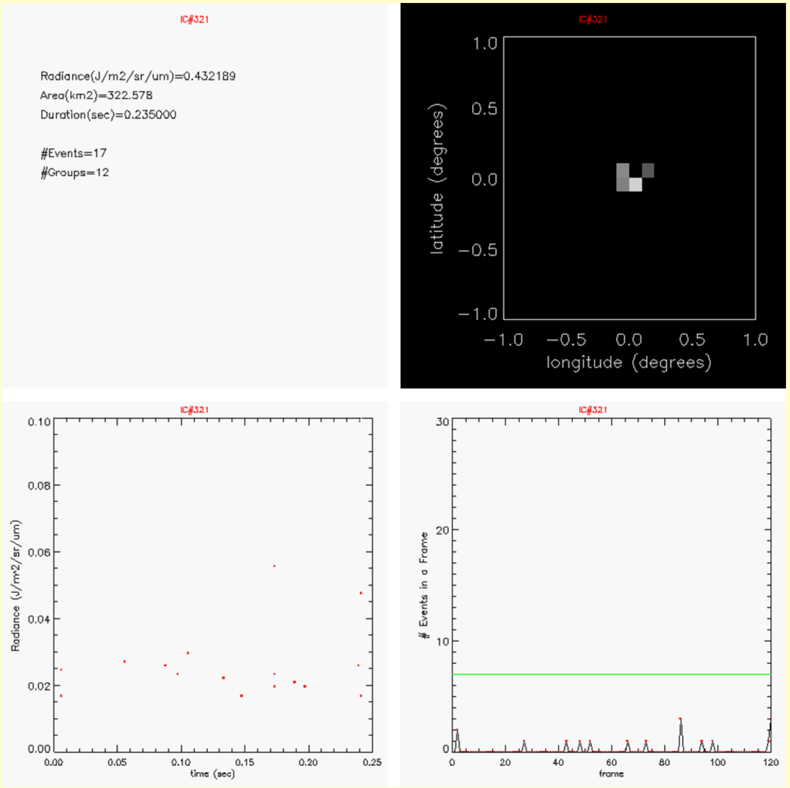
Paper #1 Highlights

IDL FlashMovie.pro Analysis

Typical Ground Flash:



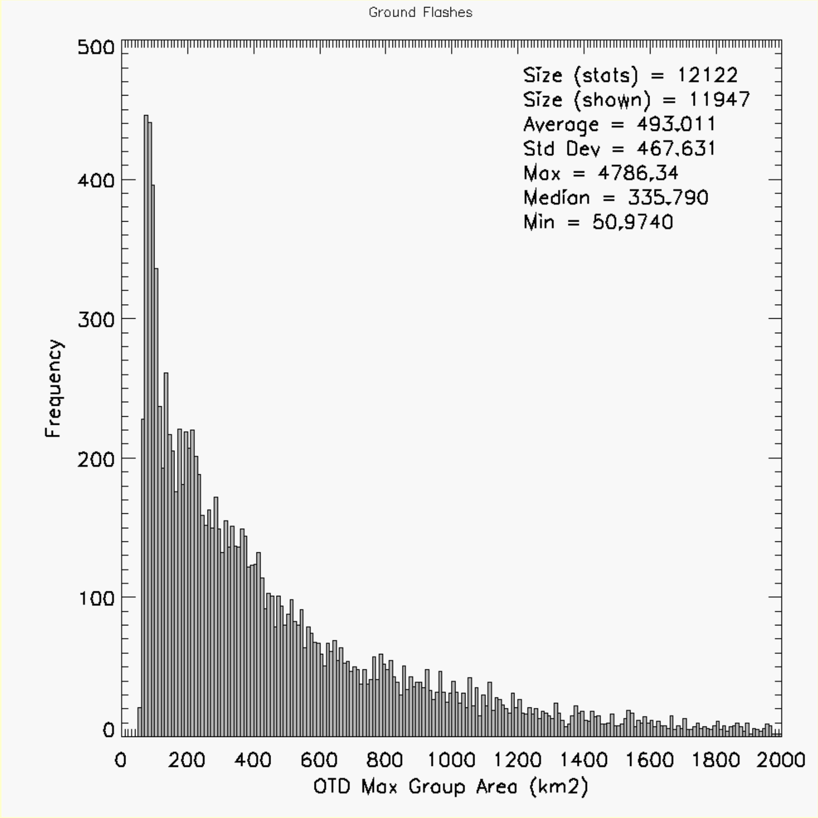
Typical Cloud Flash:



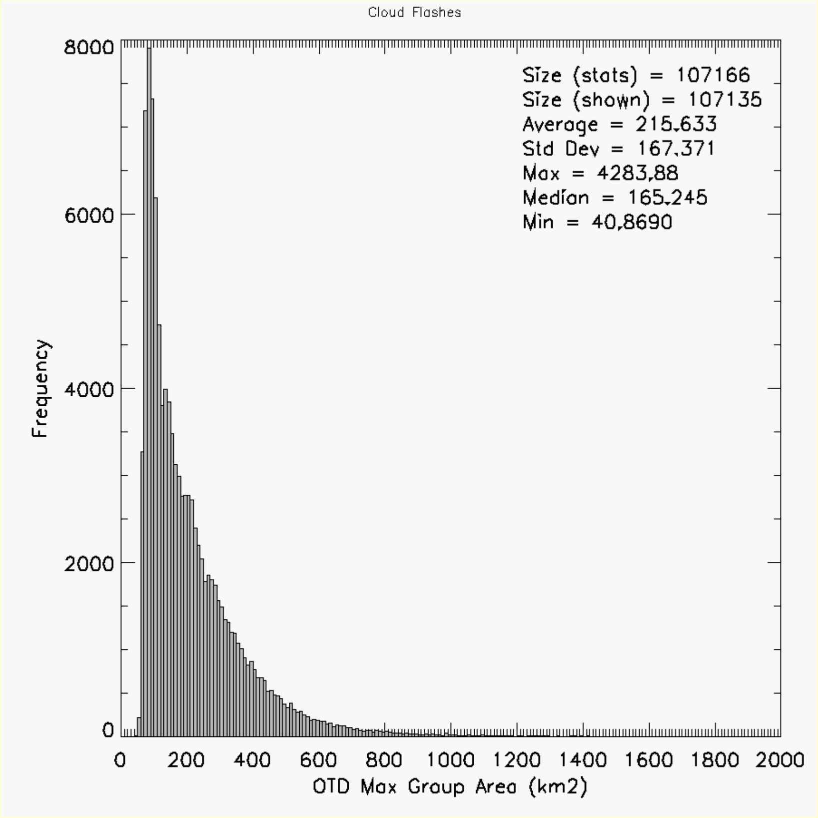
… large <u>group</u> areas seem to indicate presence of a return stroke

Distributions of MGA

Ground Flashes:

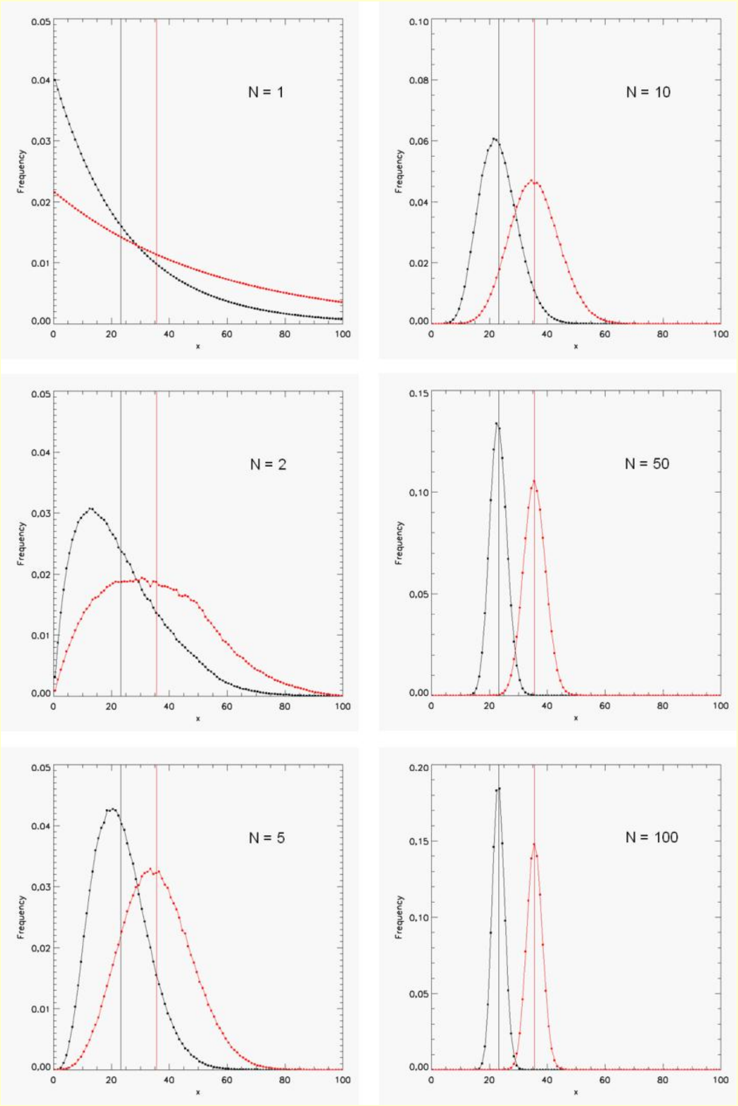


Cloud Flashes:



Use of Mean Data

Example of CLT:



Mean optical data could be used to discriminate flash type since Central Limit Theorem removes distributional overlap.

- Using mean data implies examining several (say *N*) flashes.
- Examining *N* flashes implies you are looking for the fraction of the *N* flashes that are ground flashes.

Paper #2 Highlights

Paper #2 Highlights

Mean of k^th optical charactertistic:

$$\bar{x}_k = \frac{1}{N}\sum_{i=1}^{N} x_{ik} = \frac{1}{N}\left[\sum_{j=1}^{N_g} x_{gjk} + \sum_{l=1}^{N_c} x_{clk}\right] = \frac{1}{N}\left[N_g \bar{x}_{gk} + N_c \bar{x}_{ck}\right],$$

$$\Rightarrow \left\{ \begin{aligned} \bar{x}_k &= \left[\alpha \bar{x}_{gk} + (1-\alpha)\bar{x}_{ck}\right], \\ \alpha &\equiv \frac{N_g}{N}. \end{aligned} \right.$$ (Ground Flash Fraction)

Multiple optical characteristics:

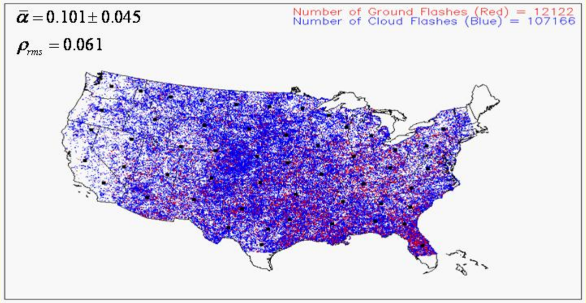
$$\alpha = \frac{\sum_{k=1}^{n} (\bar{x}_{gk} - \bar{x}_{ck})(\bar{x}_k - \bar{x}_{ck})}{\sum_{k=1}^{n} (\bar{x}_{gk} - \bar{x}_{ck})^2} \cong \frac{\sum_{k=1}^{n} (\mu_{gk} - \mu_{ck})(\bar{x}_k - \mu_{ck})}{\sum_{k=1}^{n} (\mu_{gk} - \mu_{ck})^2} \quad (for \;\; n > 1),$$

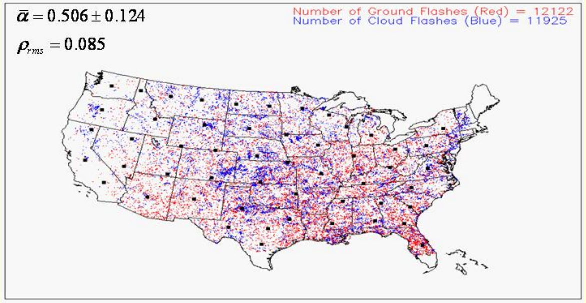
One optical characteristic:

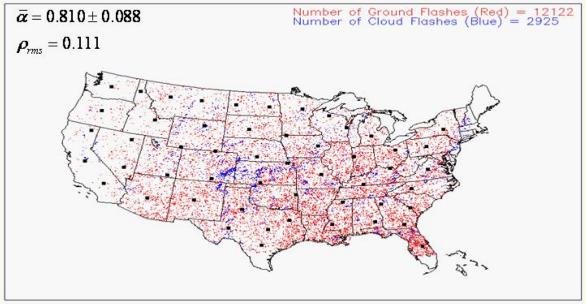
$$\alpha = \frac{(\bar{x} - \bar{x}_c)}{(\bar{x}_g - \bar{x}_c)} \cong \frac{(\bar{x} - \mu_c)}{(\mu_g - \mu_c)} \qquad (for \;\; n = 1).$$

"Poor Man" Retrievals

$$\rho = \alpha_{retrieved} - \alpha_{true}$$



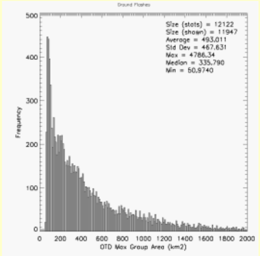




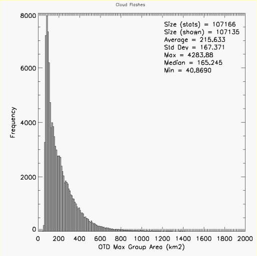
Paper #3 Highlights

Mixed Exponential Distribution Model

Ground MGAs



Cloud MGAs



Results from Paper #1

Distribution of MGA modeled as a <u>Mixed</u> Exponential Distribution:

$$p(y) = \alpha\, p_g(y) + (1-\alpha)\, p_c(y) = \frac{\alpha}{\mu_g} e^{-y/\mu_g} + \frac{(1-\alpha)}{\mu_c} e^{-y/\mu_c}, \quad y \geq 0 \quad ,$$

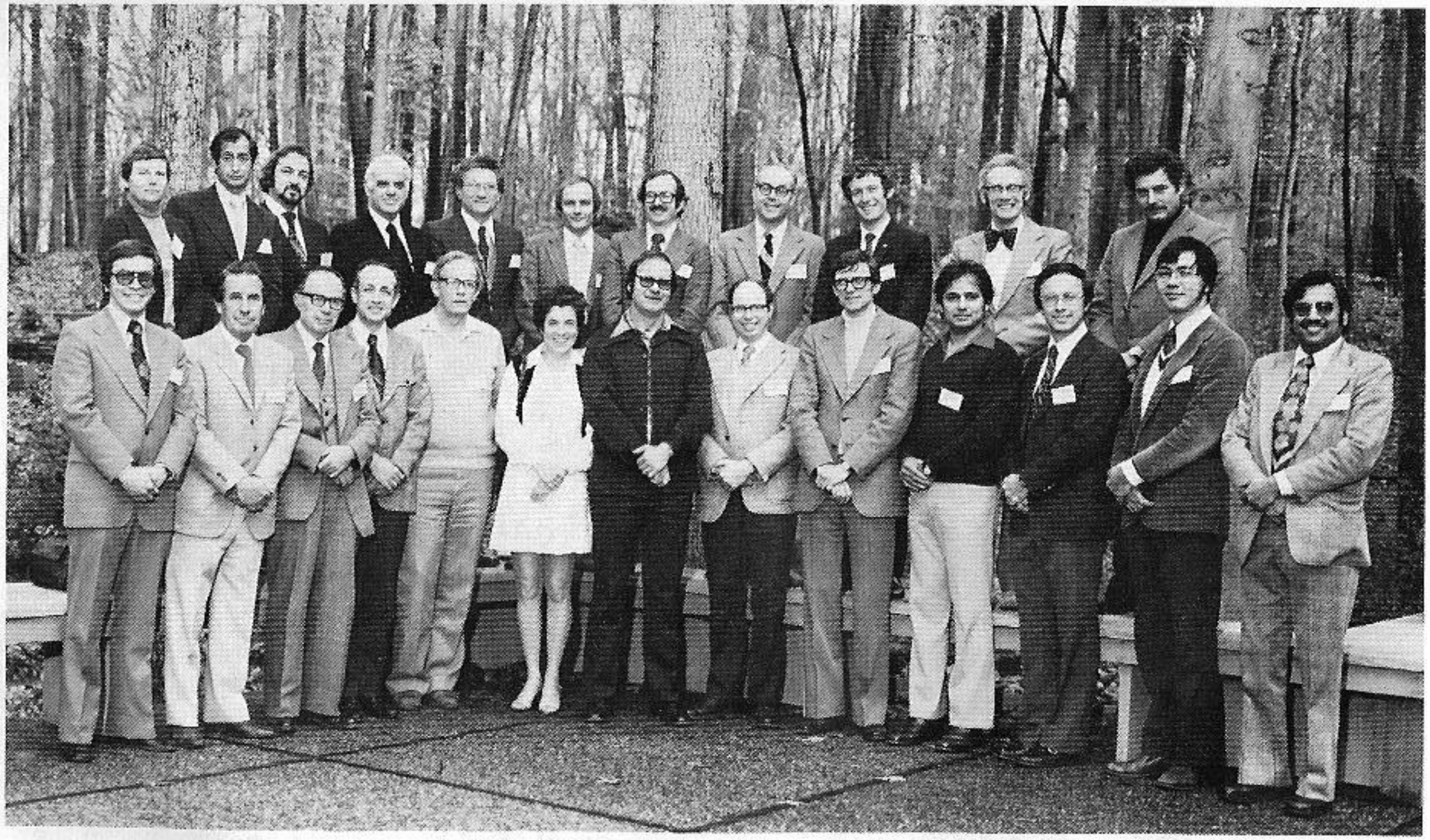
Population Means of MGA:

$$\mu_g \equiv \int_0^\infty y p_g(y) dy, \quad \mu_c \equiv \int_0^\infty y p_c(y) dy \ .$$

Require that:

$$\mu_g > \mu_c$$

1st International Interactive Workshop on Inverse Methods 1976



Workshop Speakers and Chairmen: (Left to right) Front row: M. P. McCormick (Associate Chairman), Langley Research Center; S. Twomey, U. Arizona; L. Kaplan, U. Chicago; M. Chahine, JPL/Cal. Tech; H. van de Hulst, U. Leiden, Netherlands; C. Whitney, C. S. Draper Lab; E. Westwater, NOAA/WPL; D. Staelin, MIT; B. Conrath, Goddard SFC; J. Kuriyan, UCLA; J. Gille, NCAR; W. Chu, Old Dominion U.; A. Deepak (Chairman), Old Dominion U. Second row: J. Lenoble, U. de Lille, France; B. Herman, U. Arizona; A. Fymat, JPL/Cal Tech; J. King, AFGL; A. Green, U. Florida; H. Malchow, C. S. Draper Lab; W. Irvine, U. Massachusetts; H. Fleming, NOAA/NESS; C. Rodgers, U. Oxford, UK; C. Mateer, Atmos. Environ. Serv., Canada; T. Pepin, U. Wyoming.

Bayesian Inversion

Bayes' Law:

$$P(\alpha,\mu_g,\mu_c\,|\,\mathbf{y}) = \frac{P(\mathbf{y}\,|\,\alpha,\mu_g,\mu_c)P(\alpha,\mu_g,\mu_c)}{P(\mathbf{y})} \quad ,$$

Find parameters $\mathbf{v} = (\alpha,\mu_g,\mu_c)$ that maximize the probability on LHS.

Invoking ignorance prior, means you just maximize the following :

$$S(\mathbf{v}) \equiv \ln P(\mathbf{y}\,|\,\mathbf{v}) = \ln \prod_{i=1}^{m} p(y_i\,|\,\mathbf{v}) = \sum_{i=1}^{m} \ln\left[p(y_i\,|\,\mathbf{v})\right] = \sum_{i=1}^{m} \ln\left[\frac{\alpha}{\mu_g} e^{-y_i/\mu_g} + \frac{(1-\alpha)}{\mu_c} e^{-y_i/\mu_c}\right] \quad ,$$

Formally :

$$\frac{\partial S(\mathbf{v})}{\partial \mathbf{v}} = \mathbf{0} \quad \Rightarrow \quad \mathbf{v} = \text{"Maximum A Posteriori (MAP) Solution"}$$

Practically :

Use Broyden-Fletcher-Goldfarb-Shannon variant of Davidon-Fletcher-Powell numerical method to minimize $-S(\mathbf{v})$.

Initialization for Numerical Search

Population mean and variance of the mixture:

$$\mu \equiv \int_{-\infty}^{\infty} yp(y)dy = \alpha\mu_g + (1-\alpha)\mu_c \quad (line),$$

$$\sigma^2 \equiv \int_{-\infty}^{\infty} (y-\mu)^2 p(y)dy = \alpha(2-\alpha)\mu_g^2 - 2\alpha(1-\alpha)\mu_g\mu_c + (1-\alpha^2)\mu_c^2 \quad (rotated\ ellipse).$$

Using 1st equation to solve for μ_c gives an equation quadratic in μ_g, hence:

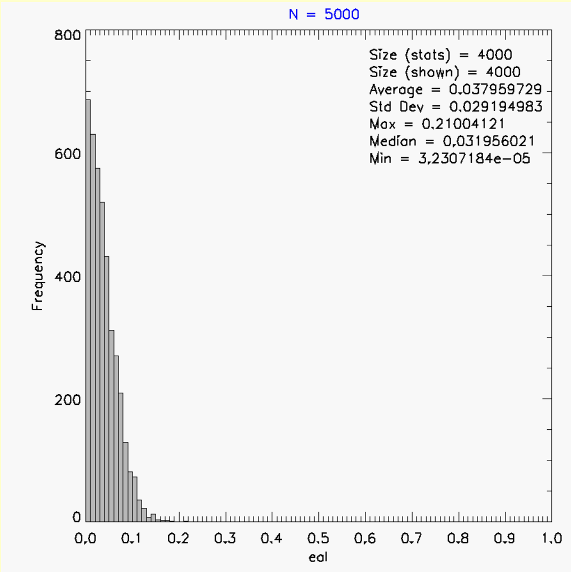
$$\mu_g = \mu + \sqrt{\frac{1}{2}\left(\frac{1-\alpha}{\alpha}\right)\left(\sigma^2 - \mu^2\right)}$$

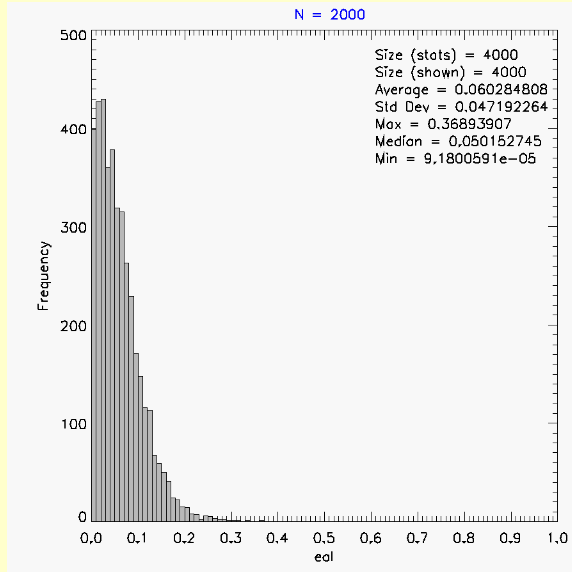
$$\mu_c = \mu - \sqrt{\frac{1}{2}\left(\frac{\alpha}{1-\alpha}\right)\left(\sigma^2 - \mu^2\right)}$$

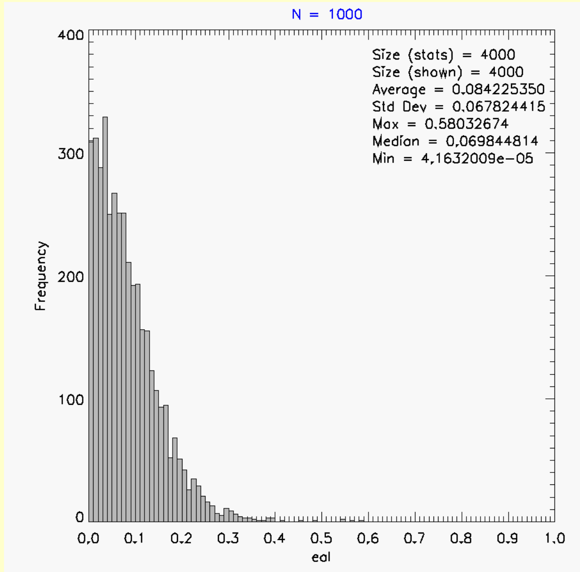
where initialization as follows results in an initialization of (μ_g, μ_c) :

$$\alpha = 0.5, \quad \mu = \bar{y}, \quad \sigma = s_y$$

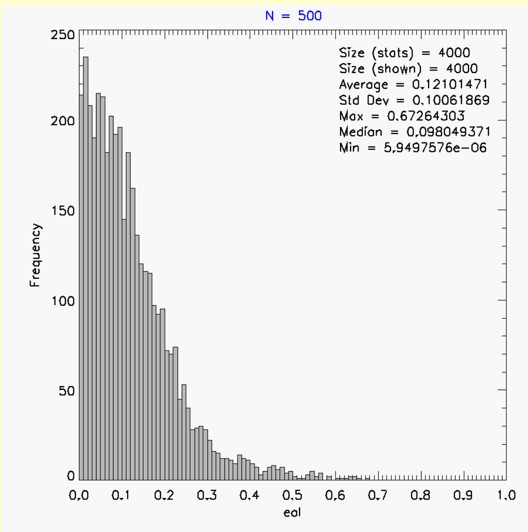
Retrieval Errors: Effect of Finite Sampling



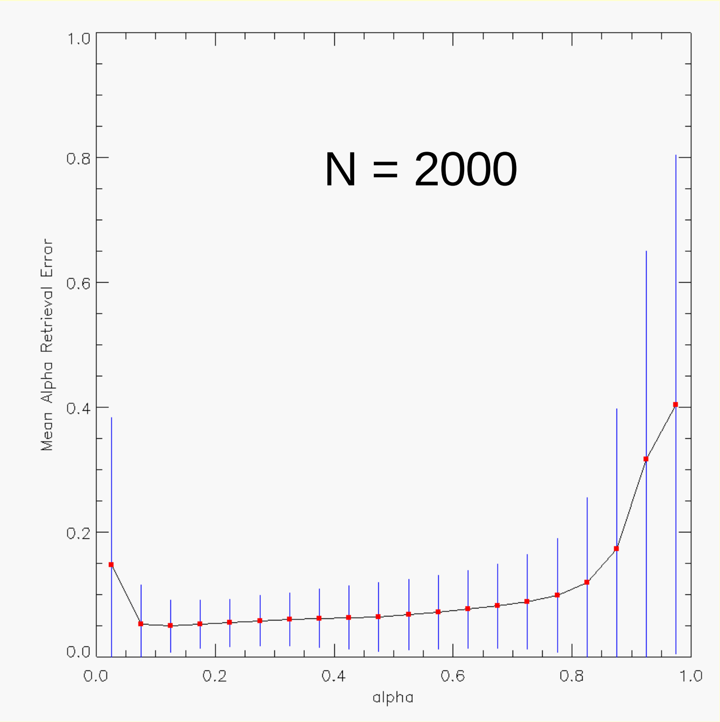


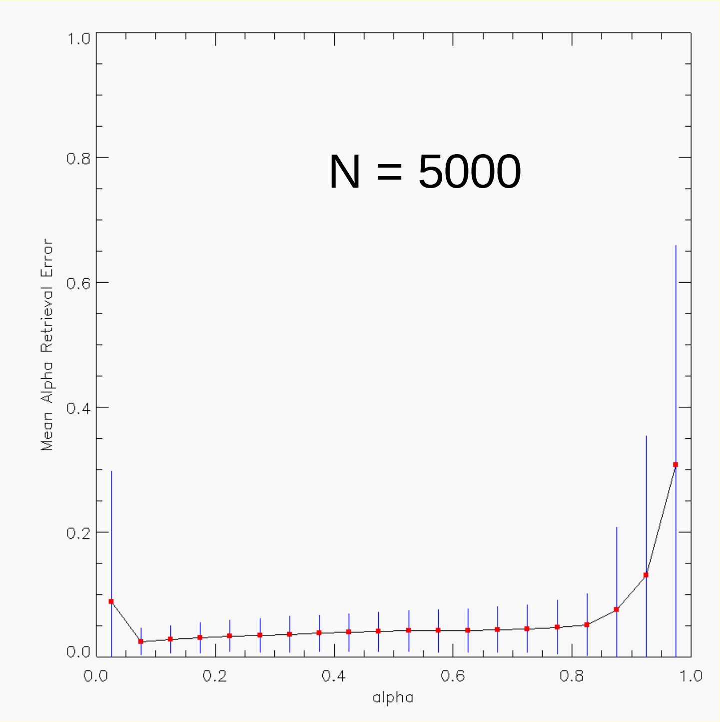


- ❑ Simulated tests
- ❑ 4000 known “mixtures” of CGs/ICs.
- ❑ Retrieval errors of <u>ground flash fraction</u> shown for different values of *N* (# flashes analyzed)



Retrieval Errors: Sensitivity to Alpha





- Simulated tests
- 100 known “mixtures” of CGs/ICs for each ground flash fraction (alpha) bin
- Mean ground flash fraction retrieval errors shown for 2 values of *N*
- Errors largest for alpha near 0 or 1 due to well-known “label switching” ambiguity

“Label Switching” Problem

$p(y)$ is invariant under the following transformation:

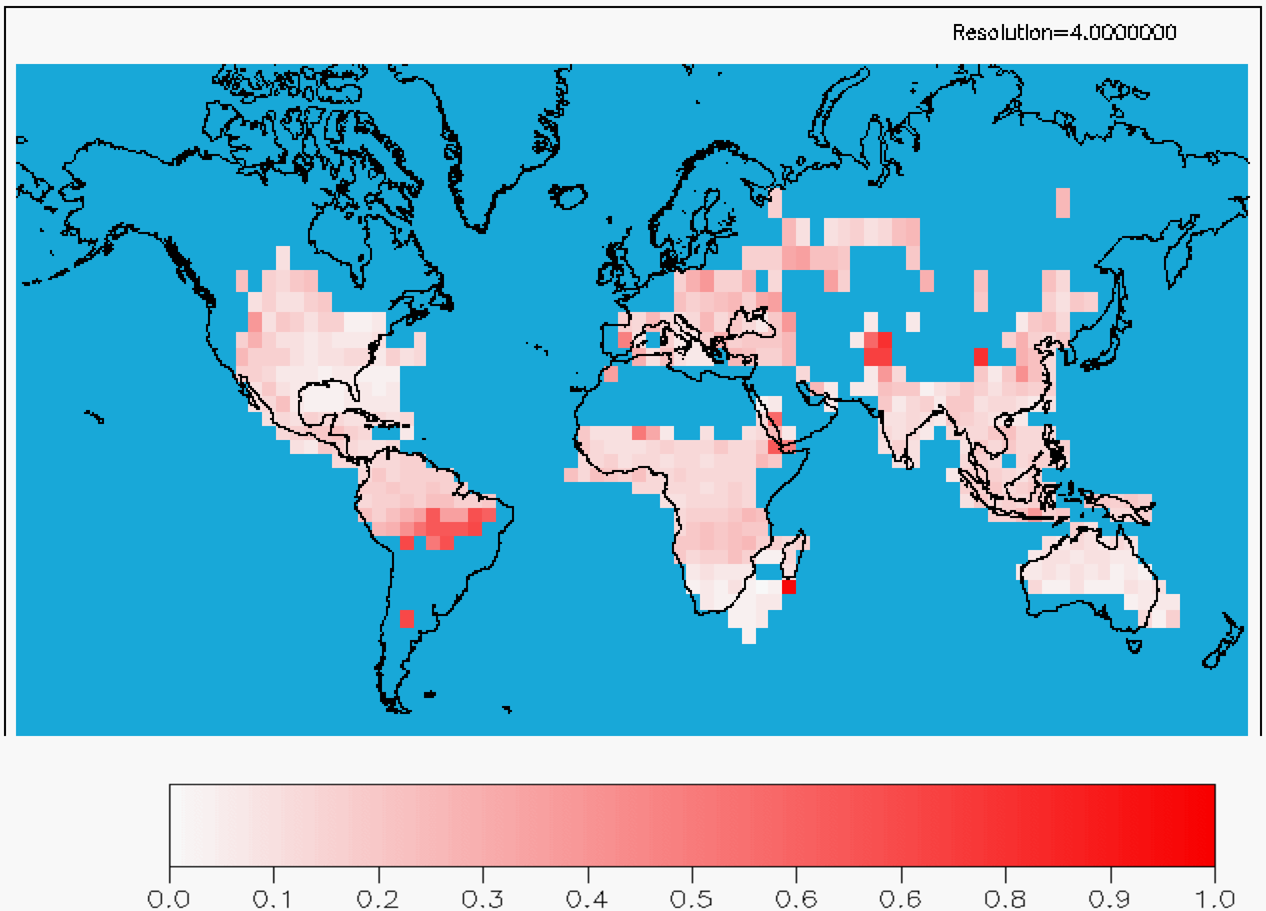
$$\tilde{\alpha} = 1-\alpha\,, \quad \tilde{\mu}_g = \mu_c\,, \quad \tilde{\mu}_c = \mu_g \quad .$$

Specifically:

$$\tilde{p}(y) = \frac{\tilde{\alpha}}{\tilde{\mu}_g} e^{-y/\tilde{\mu}_g} + \frac{(1-\tilde{\alpha})}{\tilde{\mu}_c} e^{-y/\tilde{\mu}_c} = \frac{(1-\alpha)}{\mu_c} e^{-y/\mu_c} + \frac{(1-(1-\alpha))}{\mu_g} e^{-y/\mu_g}$$

$$= \frac{\alpha}{\mu_g} e^{-y/\mu_g} + \frac{(1-\alpha)}{\mu_c} e^{-y/\mu_c} \equiv p(y)$$

Global Retrieval of Ground Flash Fraction (Preliminary)



Ground flash fraction (alpha)

- 5 yrs OTD data
- 4,365,395 flashes
- 4^{o} resolution bins
- 563 bins
- >= 2000 flashes/bin
- Mean Alpha = 0.158
- Min Alpha = 0.009
- Max Alpha = 0.95