Status of Ground Flash Fraction Retrieval Algorithm

Dr. William Koshak, NASA-MSFC

GLM Science Meeting
Huntsville, AL
September 19, 2012

Space Shuttle Video (STS-48)

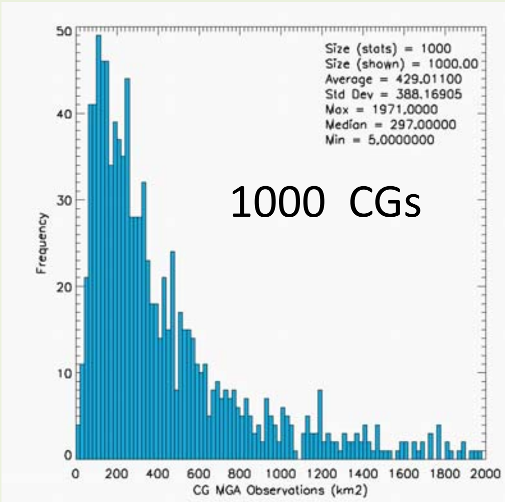
Algorithms Use

Maximum Group Area (MGA)

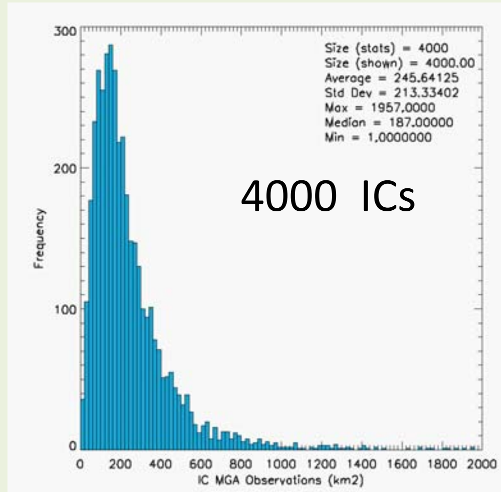
to retrieve

Ground Flash Fraction (α)

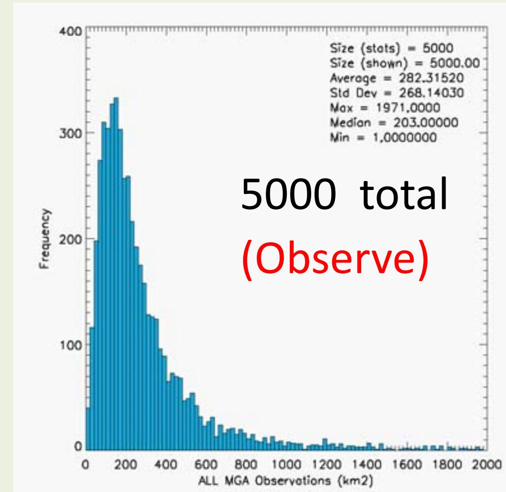
GLM Observes a Mixture of CG and IC MGAs



+



=



INVERT

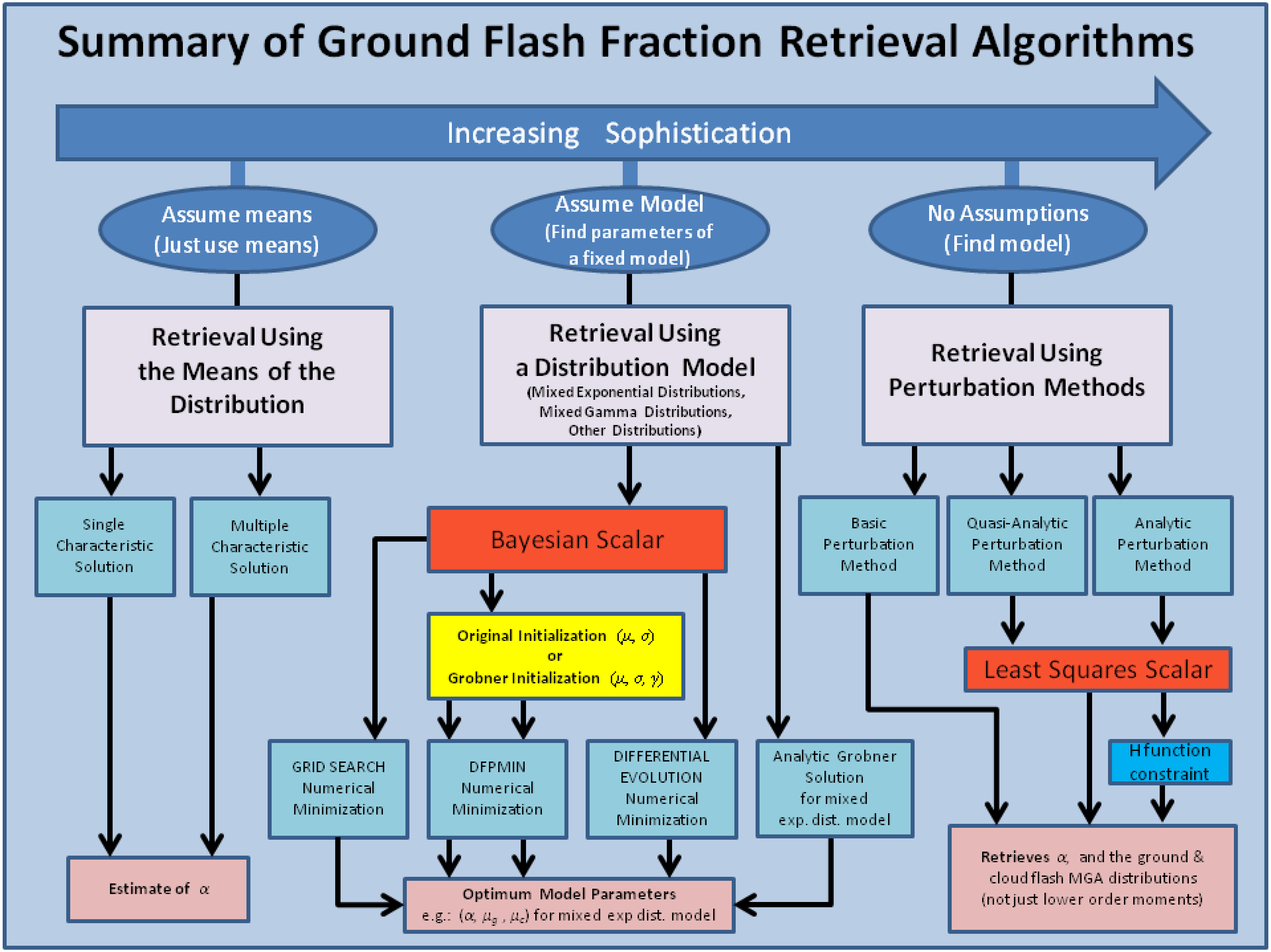
Retrieve characteristics about these.

Three Basic Algorithms:

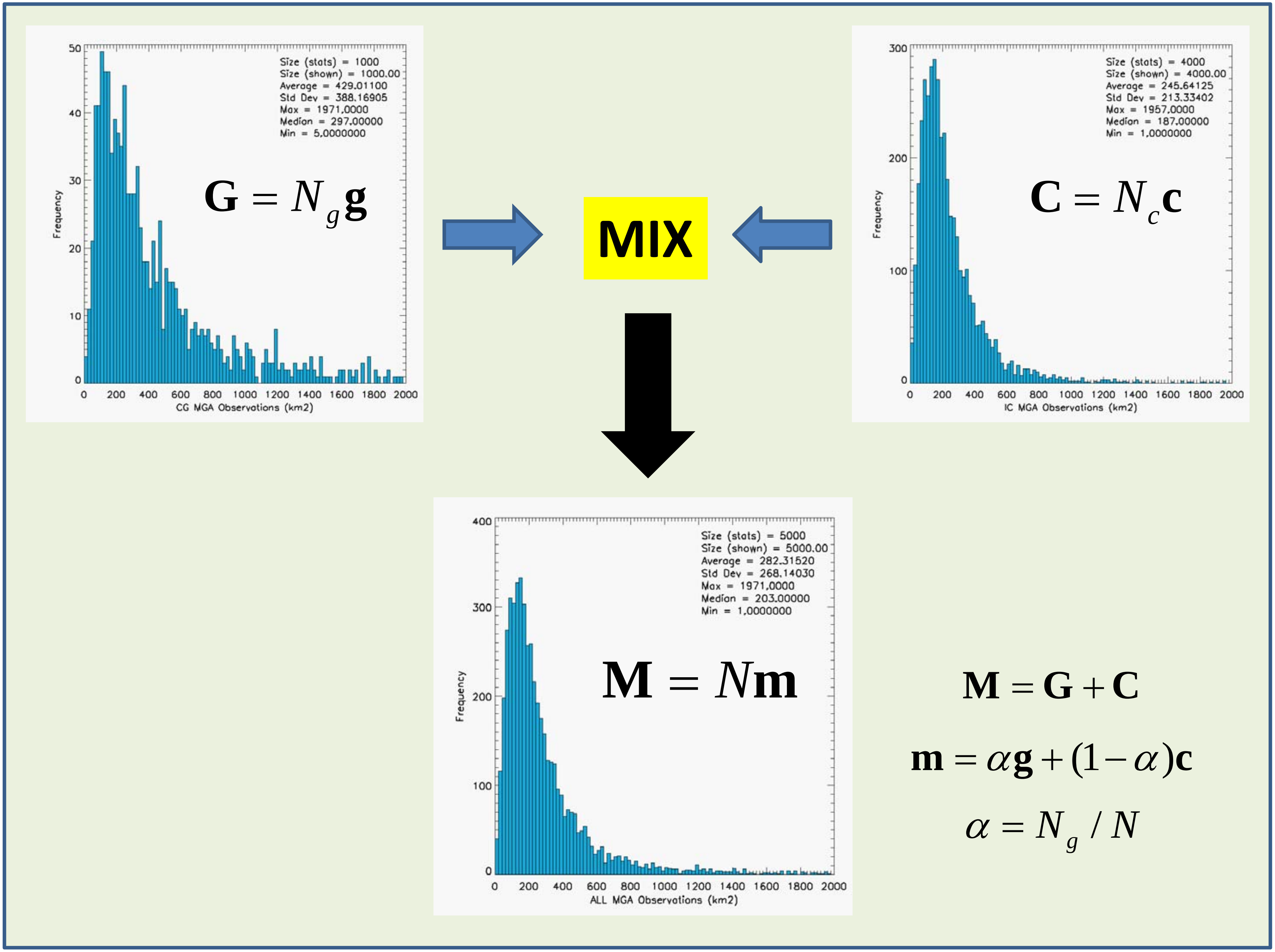
Simple: $\bar{x} = \alpha \bar{x}_g + (1-\alpha)\bar{x}_c \quad \Rightarrow \quad \text{get } \alpha \quad (\text{assume } \bar{x}_g \ \& \ \bar{x}_c)$

Bayesian: $p(x;\alpha, \mu_g, \mu_c, \ldots) = \alpha\, p_g(x;\mu_g,\ldots) + (1-\alpha)\, p_c(x;\mu_c,\ldots) \ \Rightarrow \text{get } (\alpha, \mu_g, \mu_c, \ldots)$

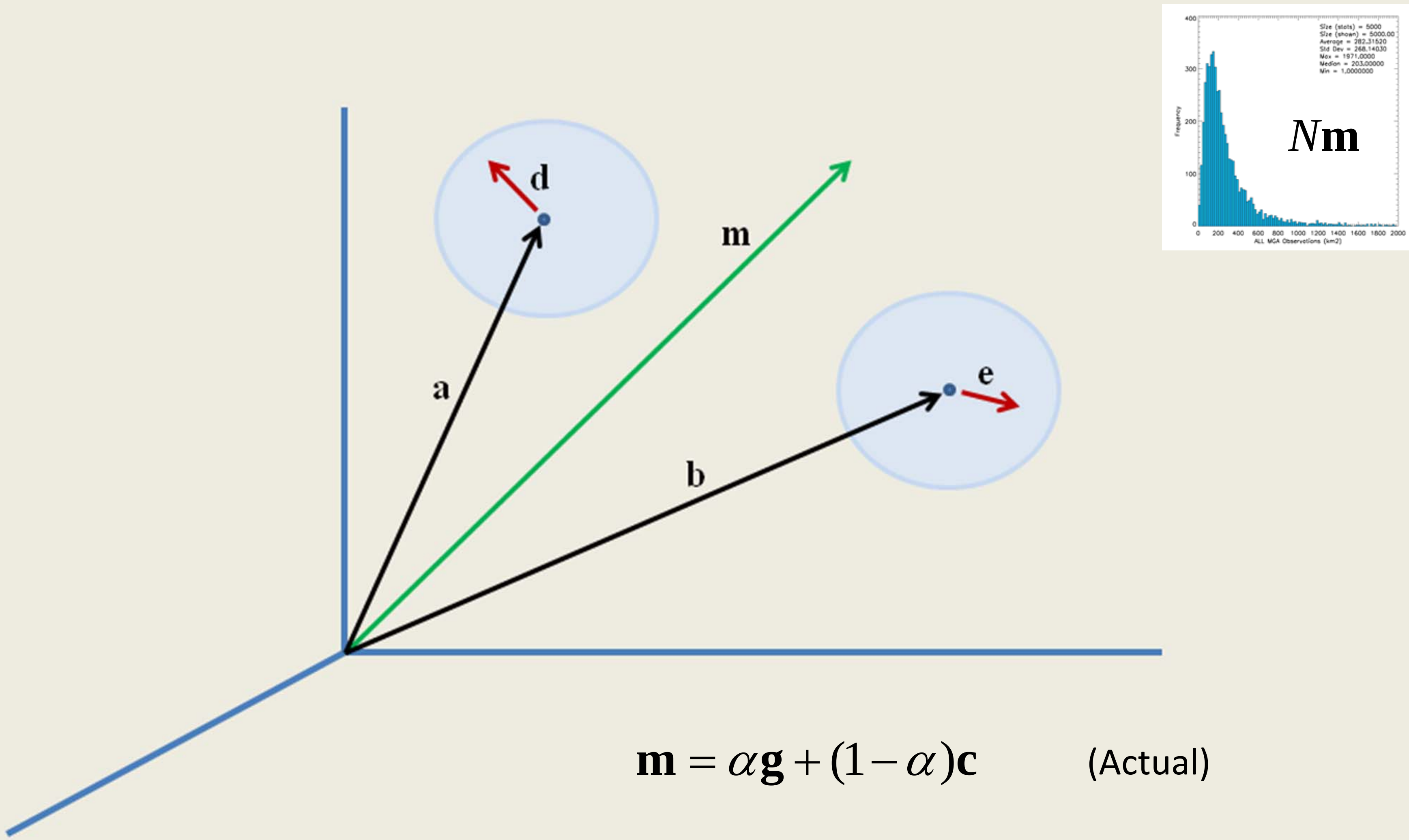
Perturbation Methods: $p(x;\alpha) = \alpha\, p_g(x) + (1-\alpha)\, p_c(x) \ \Rightarrow \text{get } \{\alpha,\ p_g(x), p_c(x)\} \ !$


Summary of Ground Flash Fraction Retrieval Algorithms
Increasing Sophistication
Assume means (Just use means)
Assume Model (Find parameters of a fixed model)
No Assumptions (Find model)
Retrieval Using the Means of the Distribution
Retrieval Using a Distribution Model (Mixed Exponential Distributions, Mixed Gamma Distributions, Other Distributions)
Retrieval Using Perturbation Methods
Single Characteristic Solution
Multiple Characteristic Solution
Bayesian Scalar
Basic Perturbation Method
Quasi-Analytic Perturbation Method
Analytic Perturbation Method
Original Initialization (μ, σ) or Grobner Initialization (μ, σ, γ)
Least Squares Scalar
H function constraint
GRID SEARCH Numerical Minimization
DFPMIN Numerical Minimization
DIFFERENTIAL EVOLUTION Numerical Minimization
Analytic Grobner Solution for mixed exp. dist. model
Estimate of α
Optimum Model Parameters e.g.: (α, μg, μc) for mixed exp dist. model
Retrieves α, and the ground & cloud flash MGA distributions (not just lower order moments)

The pathway to the
Analytic Perturbation Method



Size (stats) = 1000
Size (shown) = 1000.00
Average = 429.01100
Std Dev = 388.16905
Max = 1971.0000
Median = 297.00000
Min = 5.0000000
$\mathbf{G} = N_g \mathbf{g}$
Frequency
CG MGA Observations (km2)
MIX
Size (stats) = 4000
Size (shown) = 4000.00
Average = 245.64125
Std Dev = 213.33402
Max = 1957.0000
Median = 187.00000
Min = 1.0000000
$\mathbf{C} = N_c \mathbf{c}$
Frequency
IC MGA Observations (km2)
Size (stats) = 5000
Size (shown) = 5000.00
Average = 282.31520
Std Dev = 268.14030
Max = 1971.0000
Median = 203.00000
Min = 1.0000000
$\mathbf{M} = N\mathbf{m}$
Frequency
ALL MGA Observations (km2)
$\mathbf{M} = \mathbf{G} + \mathbf{C}$
$\mathbf{m} = \alpha \mathbf{g} + (1-\alpha)\mathbf{c}$
$\alpha = N_g / N$



$$\mathbf{m} = \alpha \mathbf{g} + (1-\alpha)\mathbf{c} \qquad \text{(Actual)}$$

$$\mathbf{m}' = \alpha_r \mathbf{u} + (1-\alpha_r)\mathbf{v} \qquad \text{(Dynamic Model)}$$

Figure 1. Basic geometry of the Perturbation Method depicted in 3-space.

BASIC PERTURBATION METHOD

(Vary both **d** and **e**)

Find $(\alpha_r, \mathbf{d}, \mathbf{e})$ that minimizes:

$$\boxed{S(\alpha_r, \mathbf{d}, \mathbf{e}) = (\mathbf{m} - \mathbf{m}')^2}$$

$$\mathbf{m} = \alpha \mathbf{g} + (1-\alpha)\mathbf{c}$$

$$\mathbf{m}' = \alpha_r[\mathbf{a}+\mathbf{d}] + (1-\alpha_r)[\mathbf{b}+\mathbf{e}] = \alpha_r \mathbf{u} + (1-\alpha_r)\mathbf{v}$$

$(\mathbf{a}, \mathbf{b}) =$ fixed climate

$(\mathbf{d}, \mathbf{e}) =$ perturbants (perturbation vectors)

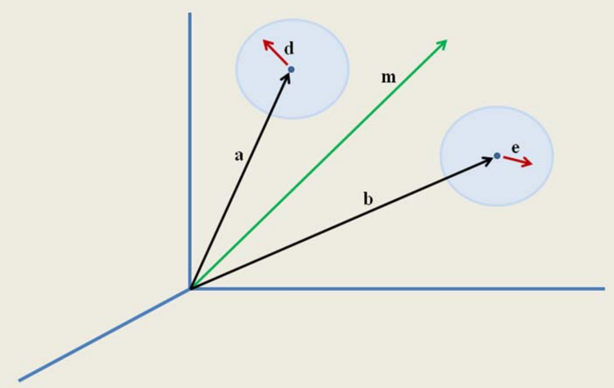
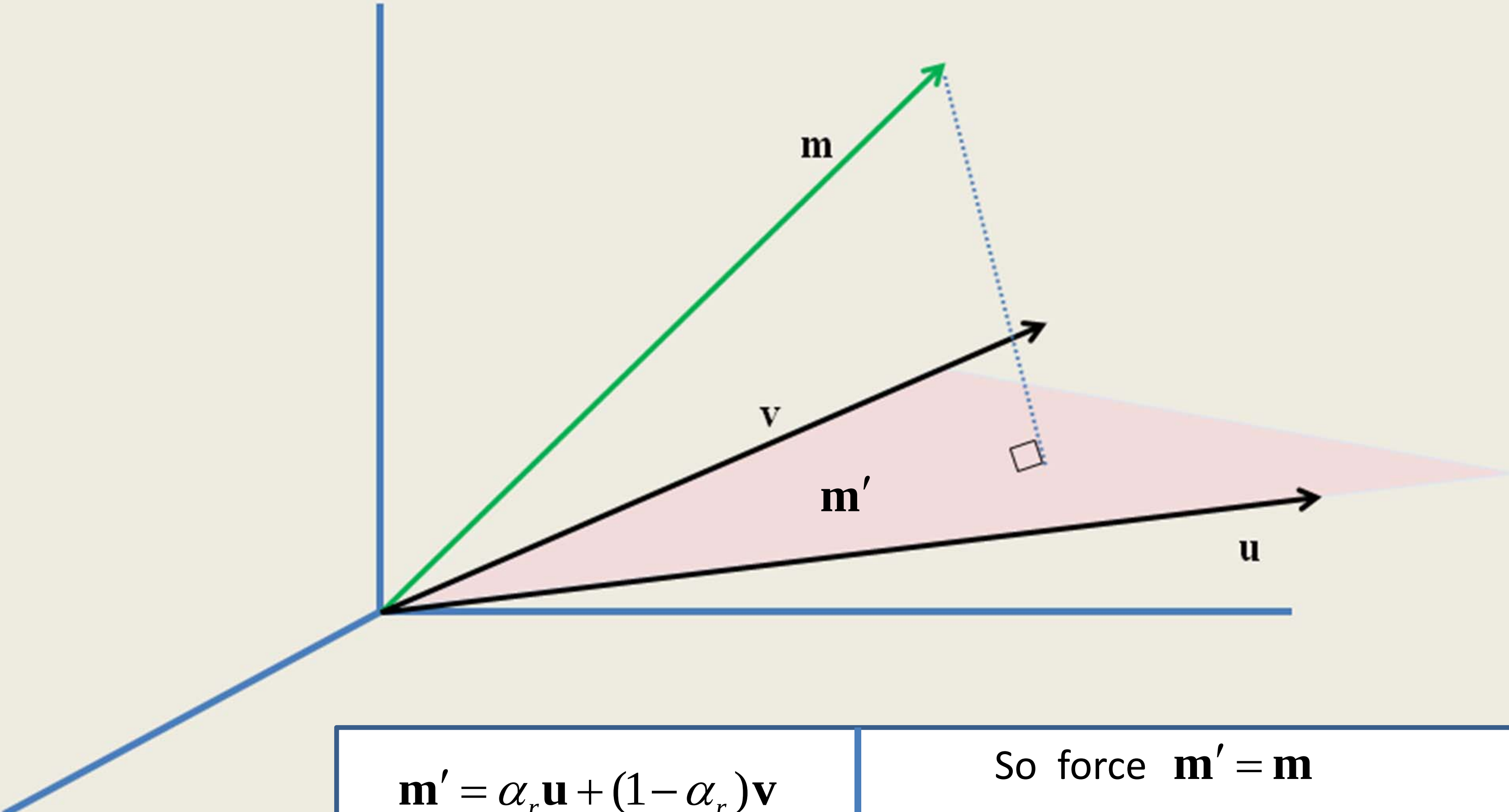


Figure 1. Basic geometry of the Perturbation Method depicted in 3-space.

QUASI-ANALYTIC PERTURBATION METHOD

(Vary just **e**; **d** obtained analytically)



$$\mathbf{m}' = \alpha_r \mathbf{u} + (1-\alpha_r)\mathbf{v}$$ lives in ($\mathbf{u}$, $\mathbf{v}$) plane	So force $\mathbf{m}' = \mathbf{m}$ $$m_i = \alpha_r(a_i + d_i) + (1-\alpha_r)(b_i + e_i) \; .$$ $$d_i = \frac{m_i - (1-\alpha_r)(b_i + e_i)}{\alpha_r} - a_i \quad \Rightarrow \quad S = 0.$$

Figure 2. The solution plane depicted in 3-space. The plane must be moved so that it contains the mixture vector $\mathbf{m}$.

ANALYTIC PERTURBATION METHOD

(**d** and **e** obtained analytically)

Define: $H = H(\alpha_r;\mathbf{m},\mathbf{a},\mathbf{b}) = p(\alpha_r)\,|\,\mathbf{d}\,|^2 + q(\alpha_r)\,|\,\mathbf{e}\,|^2 = p(\alpha_r)\sum_{i=1}^{n} d_i^2 + q(\alpha_r)\sum_{i=1}^{n} e_i^2$.

But: $d_i^2 = Ae_i^2 + B_i e_i + C_i$.

Hence: $H = \sum_{i=1}^{n}\left[(pA+q)e_i^2 + pB_i e_i + pC_i\right]$.

Where:

$$A = A(\alpha_r) = \frac{\beta_r^2}{\alpha_r^2}\,, \quad B_i = B_i(\alpha_r,\mathbf{m},\mathbf{a},\mathbf{b}) = \frac{2\beta_r(\alpha_r a_i - \phi_i)}{\alpha_r^2}\,, \quad C_i = C_i(\alpha_r,\mathbf{m},\mathbf{a},\mathbf{b}) = \frac{\phi_i^2 - 2\alpha_r a_i \phi_i + \alpha_r^2 a_i^2}{\alpha_r^2}$$

$$\phi_i = m_i - \beta_r b_i\ , \quad \beta_r = 1 - \alpha_r$$

$$p = p(\alpha_r)\ , \quad q = q(\alpha_r)\ .$$

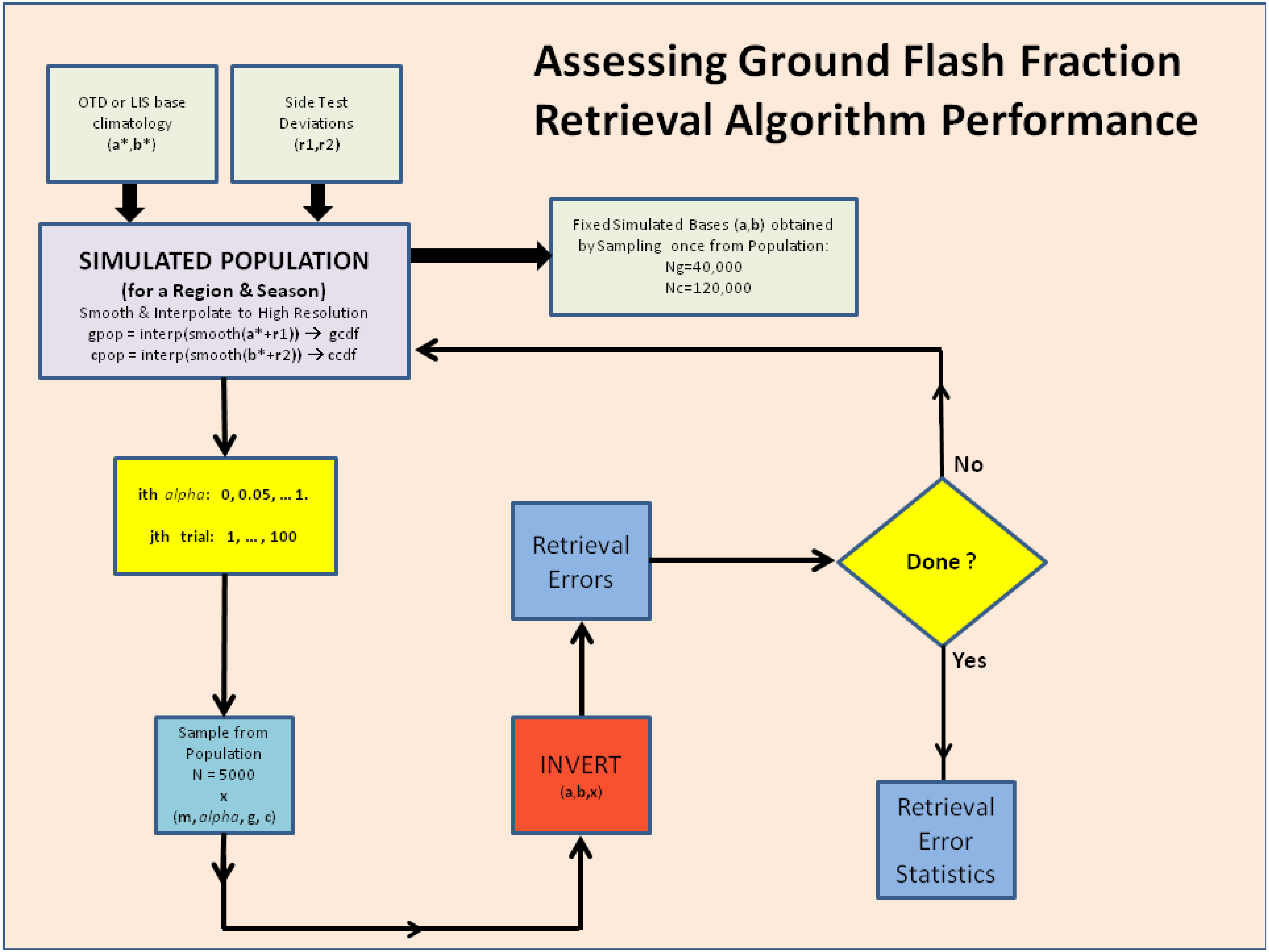
Find **e** that minimizes H:

$$\frac{dH}{de_k} = 2(pA+q)e_k + pB_k = 0\ . \qquad \Rightarrow \qquad \boxed{e_k = \frac{-pB_k}{2(pA+q)}\ , \qquad k = 1,\ldots,n\ .}$$

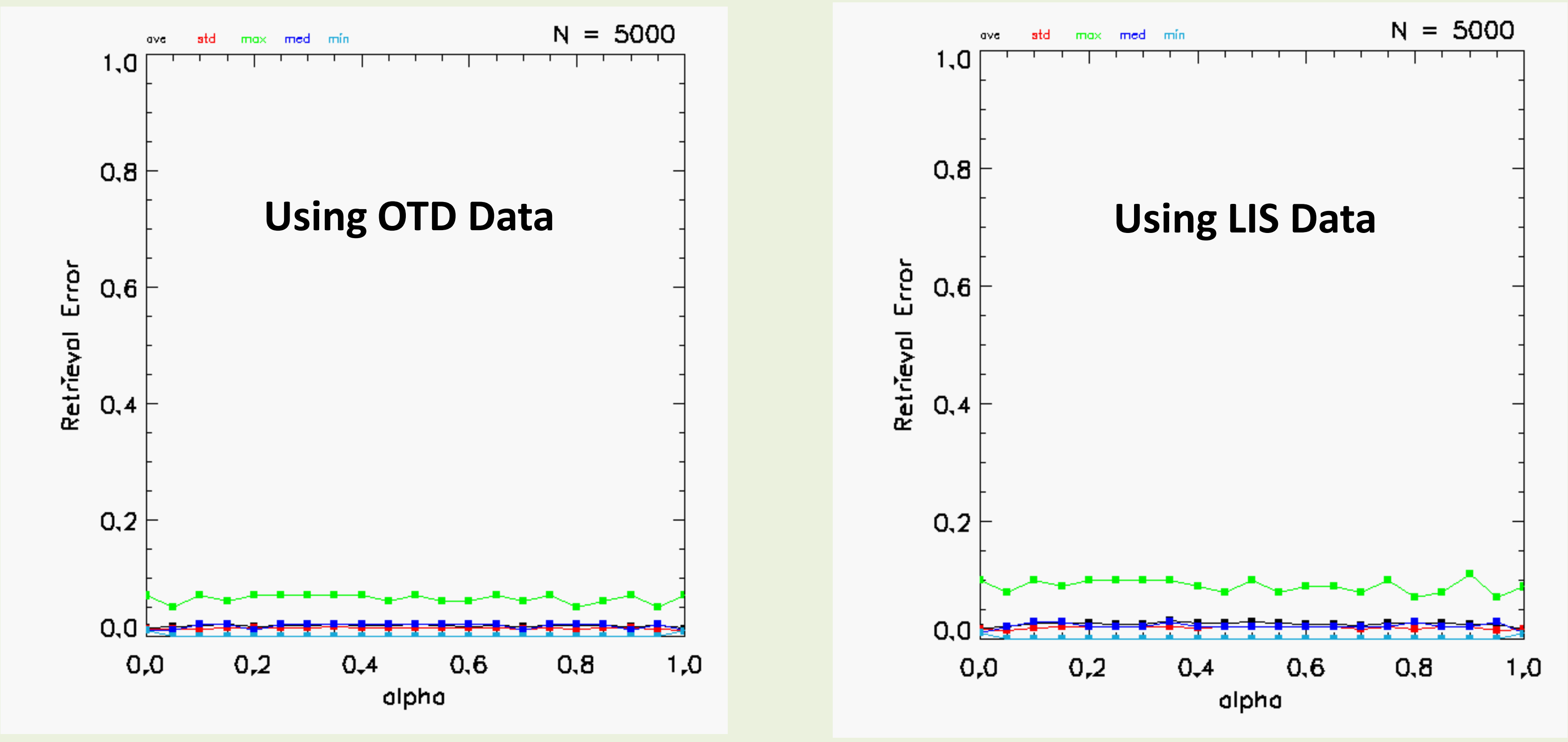
Substitute optimal **e** back into expression for H:

$$\boxed{H(\alpha_r;\mathbf{m},\mathbf{a},\mathbf{b}) = \sum_{i=1}^{n}\left[pC_i - \frac{p^2B_i^2}{4(pA+q)}\right]\ .}$$

… Quick plot of H vs. α_r provides the solution (i.e. value of α_r than minimizes H)


Assessing Ground Flash Fraction Retrieval Algorithm Performance
OTD or LIS base climatology (a*,b*)
Side Test Deviations (r1,r2)
SIMULATED POPULATION
(for a Region & Season)
Smooth & Interpolate to High Resolution
gpop = interp(smooth(a*+r1)) → gcdf
cpop = interp(smooth(b*+r2)) → ccdf
Fixed Simulated Bases (a,b) obtained by Sampling once from Population:
Ng=40,000
Nc=120,000
ith alpha: 0, 0.05, ... 1.
jth trial: 1, ... , 100
Sample from Population
N = 5000
x
(m, alpha, g, c)
INVERT
(a,b,x)
Retrieval Errors
Done ?
No
Yes
Retrieval Error Statistics

RETRIEVAL ERRORS



Sampling from population [to create climate (**a**, **b**)]:

- Ng = 40,000 flashes
- Nc = 120,000 flashes
- Z = Nc/Ng = 3

Thank You

(Questions?)