



**The Abdus Salam
International Centre for Theoretical Physics**



2141-34

**Joint ICTP-IAEA Workshop on Nuclear Reaction Data for Advanced
Reactor Technologies**

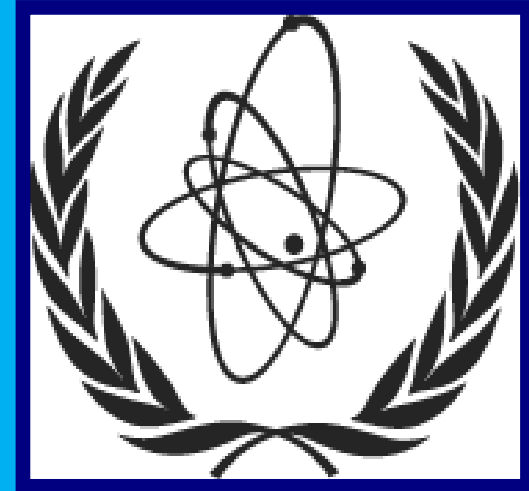
3 - 14 May 2010

Monte Carlo Approaches for Model Evaluation of Cross Sections and Uncertainties

CAPOTE R.

*IAEA
Vienna
AUSTRIA*

Monte Carlo approaches for Model evaluation of cross sections and uncertainties



Roberto Capote and Donald Smith (*)

International Atomic Energy Agency, NAPC - Nuclear Data Section

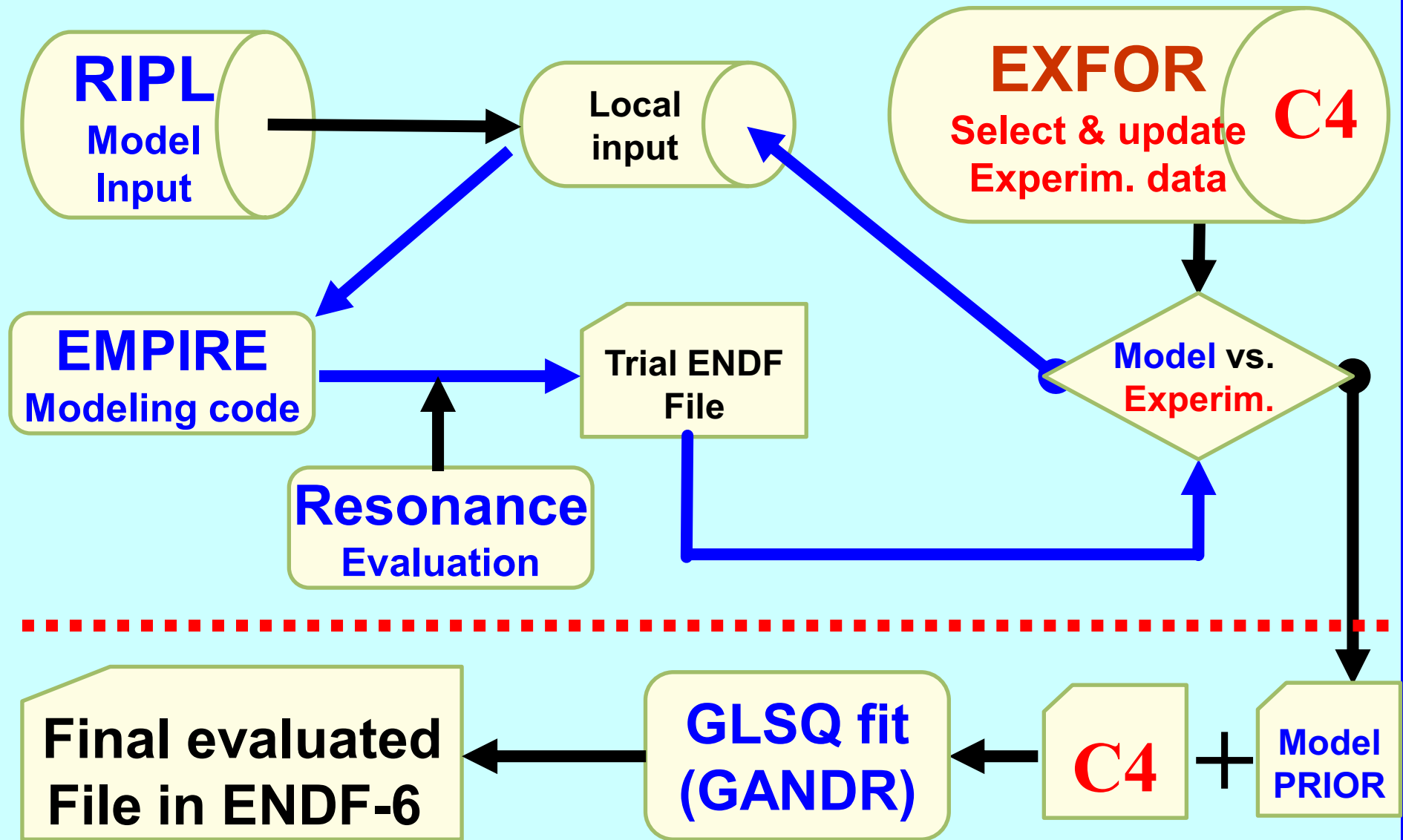
(*) Argonne National Laboratory

OUTLINE

- ❑ Introduction: Monte Carlo method
- ❑ UMC formulation (normal distributions)
- ❑ UMC sampling: Brute Force, Metropolis
- ❑ Toy models: Linear
- ❑ UMC convergence
- ❑ Toy models: Ratio data
- ❑ Log transformation
- ❑ Non-normal PDFs (& examples):
 - 1) lognormal distribution;
 - 2) uniform interval distribution;
 - 3) exponential distribution.
- ❑ Results and discussions for each PDF



Nuclear Data Evaluation process



MONTE CARLO METHOD

D.L. Smith, “Covariance Matrices for Nuclear Cross-Sections Derived from Nuclear Model Calculations”.

Report **ANL/NDM-159**, Argonne National Laboratory, 2005

$$\bar{\sigma}_i = \frac{1}{K} \sum_{k=1}^K \sigma_{ik} \quad V_{ij} = \overline{\sigma_i \sigma_j} - \bar{\sigma}_i \times \bar{\sigma}_j \quad i,j$$

- energy indexes

Monte Carlo calculation of covariance first tested by A. Koning

Monte Carlo prior

+

GANDR (GLS)

D.W. Muir, **GANDR** project (IAEA),
Online at www-nds.iaea.org/gandr/.

A. Trkov and R. Capote, “Cross-Section Covariance Data”, Th-232 evaluation for ENDF/B-VII.0 (MAT=9040 MF=1 MT=451); Pa-231 and Pa-233 evaluations for ENDF/B-VII.0 (MAT=9133 and 9137 MF=1 MT=451), National Nuclear Data Center, BNL (<http://www.nndc.bnl.gov>), 15 December 2006.



Merging of Model Calculated and Experimental Results ... More

(a.k.a. "Auto Repair Shop" Solution)

\bar{f} = collection of functions that related $\bar{\sigma}$ to the data,
i.e., given $\bar{\sigma}$ we can calculate the equivalent to \bar{y}



$\bar{\sigma}_E$ = model calculated cross sections \bar{V}_E = corresponding cov. matrix

$\therefore p(\bar{\sigma} | E, C)$ = probability density function for $\bar{\sigma}$ given experimental data "E" and calculated model-calculated prior results "C"

$$p(\bar{\sigma} | E, C) = C \exp \left\{ (-\frac{1}{2}) [\bar{y}_E - \bar{f}(\bar{\sigma})]^T \bar{V}_E^{-1} [\bar{y}_E - \bar{f}(\bar{\sigma})] + (-\frac{1}{2}) (\bar{\sigma} - \bar{\sigma}_C)^T \bar{V}_C^{-1} (\bar{\sigma} - \bar{\sigma}_C) \right\}$$

$$\bar{\sigma} = \sigma_1, \sigma_2, \dots, \sigma_i, \dots, \sigma_N$$

Donald L. Smith
Argonne National Laboratory
WPEC 2007 - SG24

UNIFIED MONTE CARLO (UMC)

D.L. Smith, “A Unified Monte Carlo Approach to Fast Neutron Cross Section Data Evaluation,” *Proceedings of the 8th International Topical Meeting on Nuclear Applications and Utilization of Accelerators*, Pocatello, July 29 – August 2, 2007, p. 736.

BAYES THEOREM & PRINCIPLE OF MAXIMUM ENTROPY

$$p(\boldsymbol{\sigma}) = C \times \mathcal{L}(\mathbf{y}_E, \mathbf{V}_E \mid \boldsymbol{\sigma}) \times p_0(\boldsymbol{\sigma} \mid \boldsymbol{\sigma}_C, \mathbf{V}_C)$$

$$p_0(\boldsymbol{\sigma} \mid \boldsymbol{\sigma}_C, \mathbf{V}_C) \sim \exp\{-(1/2)[(\boldsymbol{\sigma}-\boldsymbol{\sigma}_C)^T \cdot (\mathbf{V}_C)^{-1} \cdot (\boldsymbol{\sigma}-\boldsymbol{\sigma}_C)]\}$$

$$\mathcal{L}(\mathbf{y}_E, \mathbf{V}_E \mid \boldsymbol{\sigma}) \sim \exp\{-(1/2)[(\mathbf{y}-\mathbf{y}_E)^T \cdot (\mathbf{V}_E)^{-1} \cdot (\mathbf{y}-\mathbf{y}_E)]\}, \mathbf{y}=f(\boldsymbol{\sigma})$$

$\mathbf{y}_E, \mathbf{V}_E$: measured quantities with “n” elements

$\mathbf{y}_C, \mathbf{V}_C$: calculated using nuclear models with “m” elements

UMC based on $p(\boldsymbol{\sigma})$, GLS on the peak of the distribution



UMC sampling schemes

BF approach: A set of independent $\{\sigma\}$

$$\bar{\sigma}_{Ck} - \psi [(\mathbf{V}_C)_{ii}]^{1/2} \leq \sigma_{ik} \leq \bar{\sigma}_{Ck} + \psi [(\mathbf{V}_C)_{ii}]^{1/2}$$

$$\sigma_{ik} = \bar{\sigma}_{Ck} + (2\gamma - 1) \psi [(\mathbf{V}_C)_{ii}]^{1/2}$$

METROPOLIS approach: An stochastic Markov chain $\{\sigma\}$ distributed following $p(\sigma)$

$$\sigma' = \sigma(t) + (2\gamma - 1) \delta [(\mathbf{V}_C)_{ii}]^{1/2}, \text{ being } \sigma(t=0) = \bar{\sigma}_C$$

If $p(\sigma') > \gamma p(\sigma(t))$ then $\sigma(t+1) = \sigma'$; else $\sigma(t+1) = \sigma(t)$

$$p_0(\sigma | \sigma_C, \mathbf{V}_C) \sim \exp\{-\frac{1}{2}[(\sigma - \sigma_C)^T \cdot (\mathbf{V}_C)^{-1} \cdot (\sigma - \sigma_C)]\}$$



Central Limit Theorem

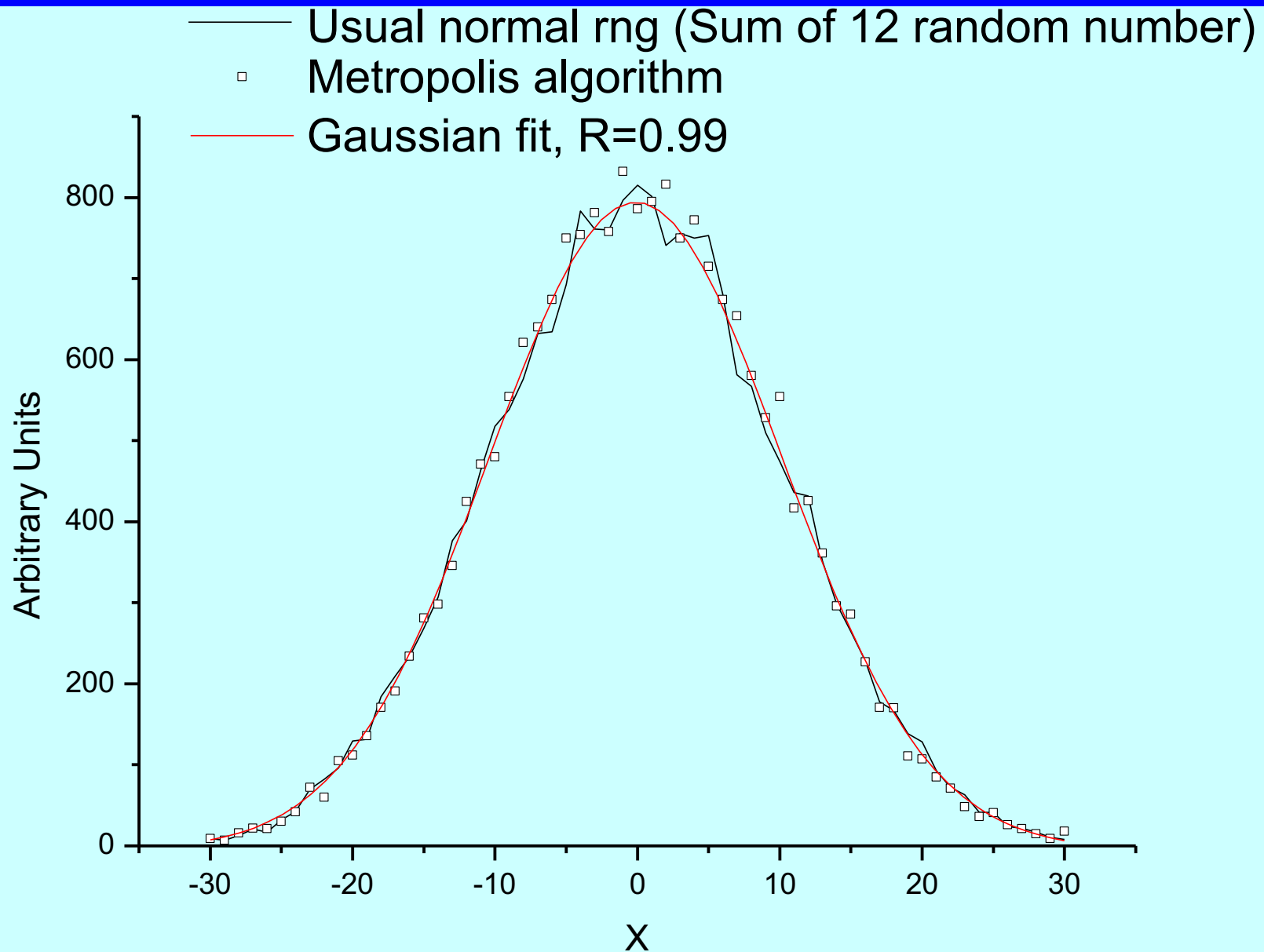
```
do MCsteps=1,naccept,K
  rnorm=-6
  do i=1,12
    CALL RANDOM(rnd)
    rnorm=rnorm+rnd
  enddo
  rnorm are normally distributed
  RANDOMS1(MCSTEPS)=rnorm
enddo
```



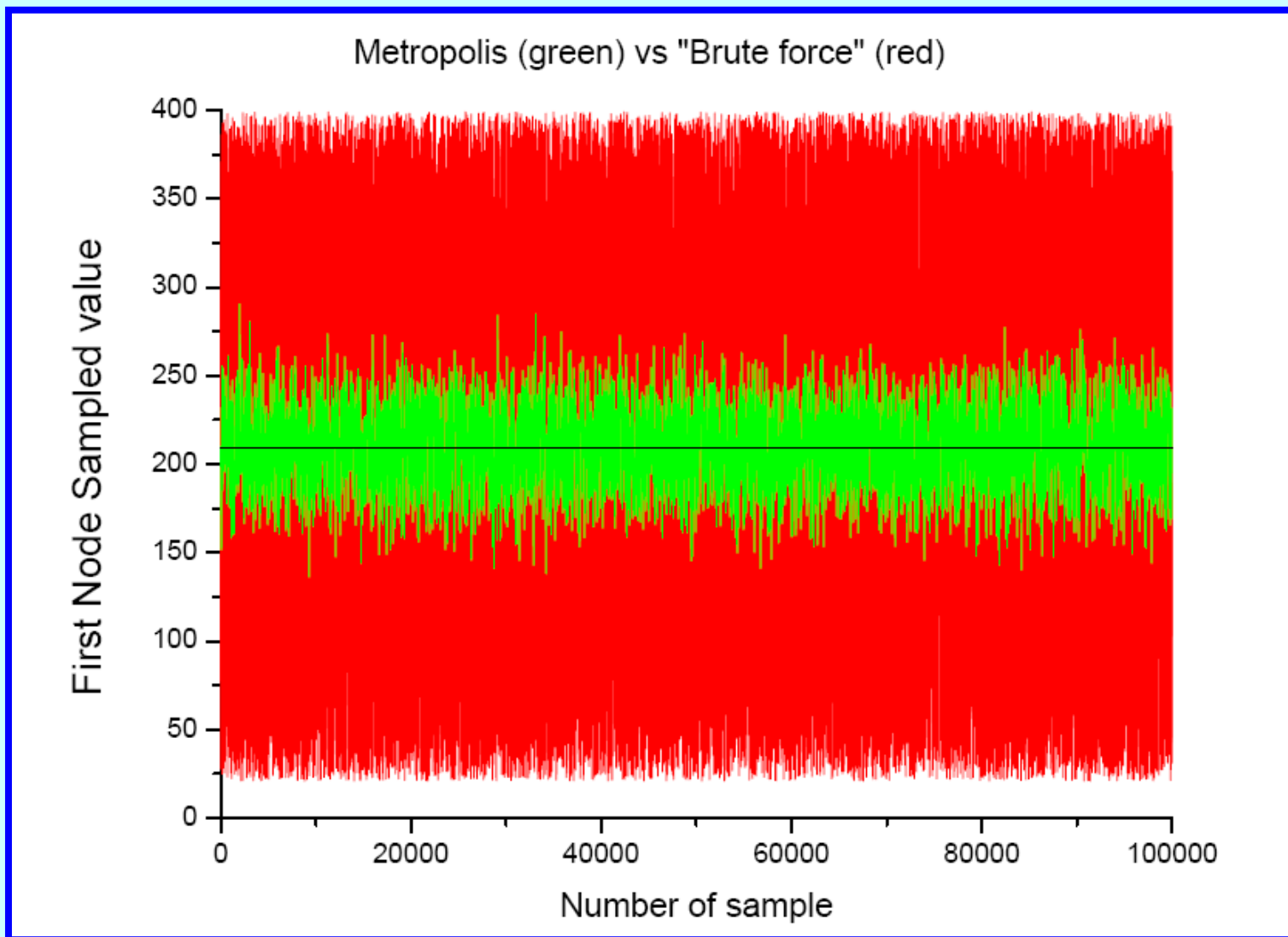
Metropolis algorithm

```
naccept=0
do MCsteps=1,NEVENTS
  CALL RANDOM(rnd)
  xt=xn+(2*rnd-1)*3
  wt=f(xt)
  CALL RANDOM(rnd)
  if (wt.GT.(rnd*wn)) then
    xn=xt
    wn=wt
    naccept=naccept+1
  endif
  xn are normally distributed
  RANDOMS(MCSTEPS)=xn
enddo
```





SPREAD OF SAMPLED VALUES



LINEAR MODEL: $y = \sigma$

<u>Node</u>	<u>Model</u>	<u>Expt</u>	$\sigma_E(\%)$	<u>Expt / Model</u>	<u>Comments</u>
1	210	205.6	30.0%	0.979	within error
2	40	39.3	2.0%	0.983	within error
3	20	26	30.0%	1.300	marginal
4	10	14	5.0%	1.400	discrepant
5	7	6.7	3.0%	0.957	marginal
6	6	8.5	50.0%	1.417	within BIG error
7	6				No exp.data

$\text{covexp}(3,1) = \text{covexp}(1,3) = 0.2 \sigma_E(1) \sigma_E(3)$ - weak correlation

$\text{covexp}(5,2) = \text{covexp}(2,5) = 0.8 \sigma_E(2) \sigma_E(5)$ - strong correlation



MODEL DATA & CORRELATION

***** MODEL DATA

Pmod(1) = 210.0000 +/- 63.0000 (30.0%)	Ymod(1) = 210.0000
Pmod(2) = 40.0000 +/- 12.0000 (30.0%)	Ymod(2) = 40.0000
Pmod(3) = 20.0000 +/- 6.0000 (30.0%)	Ymod(3) = 20.0000
Pmod(4) = 10.0000 +/- 3.0000 (30.0%)	Ymod(4) = 10.0000
Pmod(5) = 7.0000 +/- 2.1000 (30.0%)	Ymod(5) = 7.0000
Pmod(6) = 6.0000 +/- 1.8000 (30.0%)	Ymod(6) = 6.0000
Pmod(7) = 6.0000 +/- 1.8000 (30.0%)	Ymod(7) = 6.0000

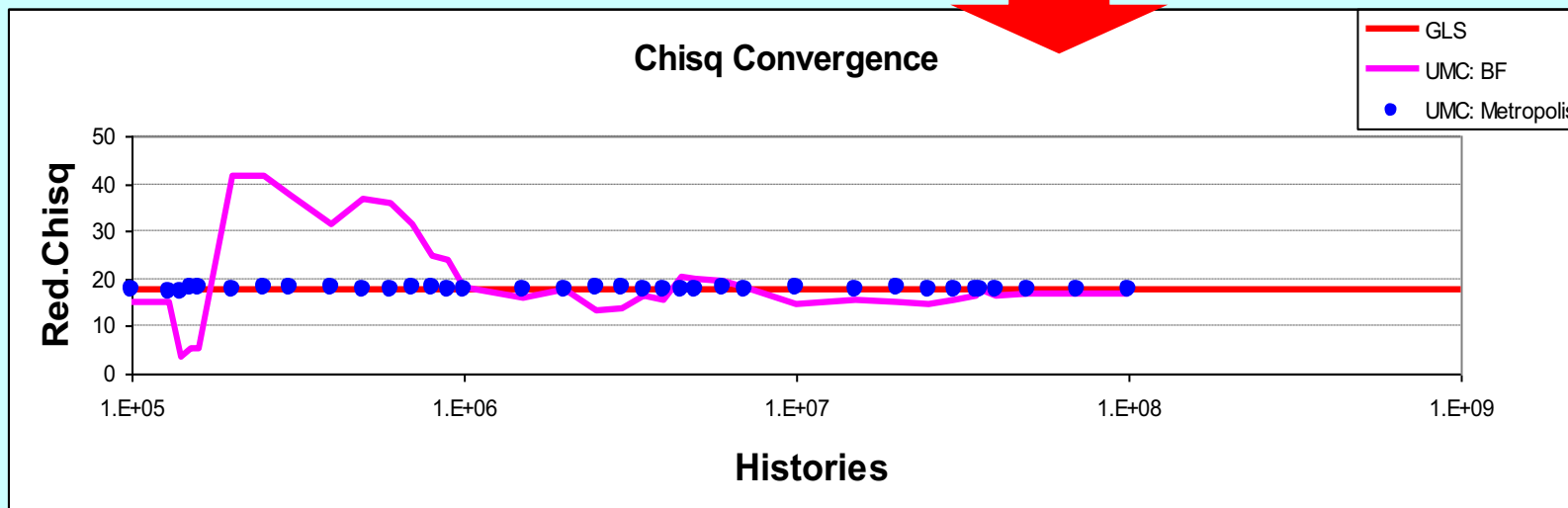
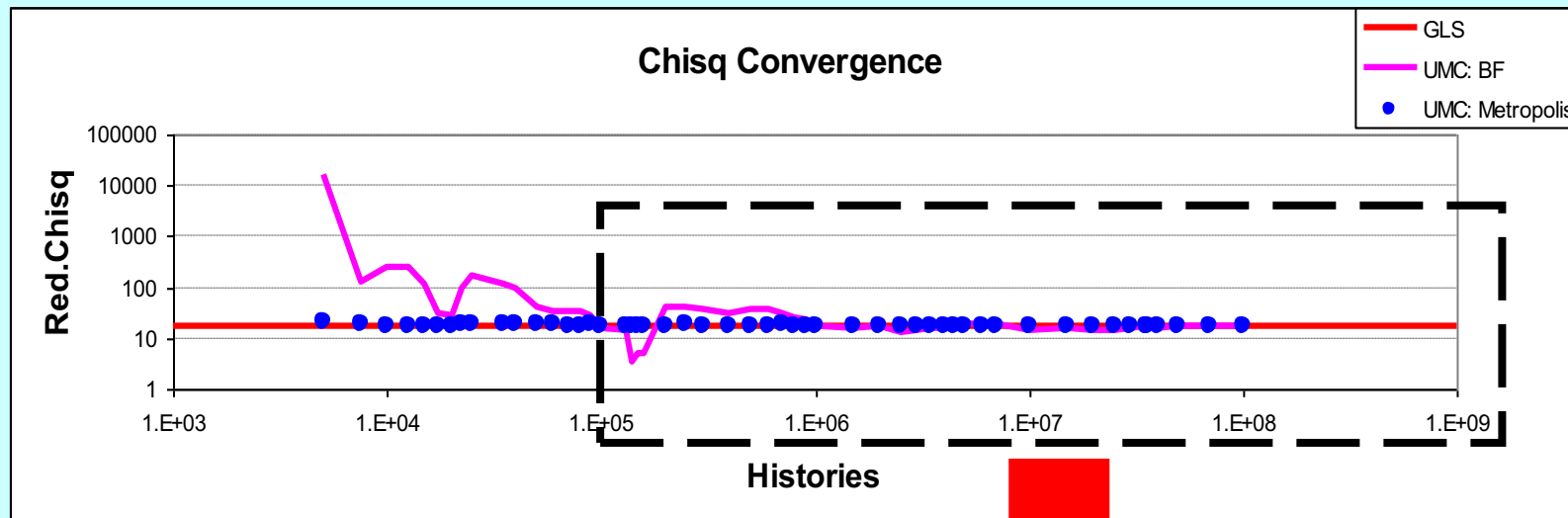
MODEL CORRELATION MATRIX (PRIOR):

```
0.1000000E+01
0.9500000E+00 0.1000000E+01
0.9000000E+00 0.9500000E+00 0.1000000E+01
0.8500000E+00 0.9000000E+00 0.9500000E+00 0.1000000E+01
0.8000000E+00 0.8500000E+00 0.9000000E+00 0.9500000E+00 0.1000000E+01
0.7500000E+00 0.8000000E+00 0.8500000E+00 0.9000000E+00 0.9500000E+00 0.1000000E+01
0.7000000E+00 0.7500000E+00 0.8000000E+00 0.8500000E+00 0.9000000E+00 0.9500000E+00 0.1000000E+01
```

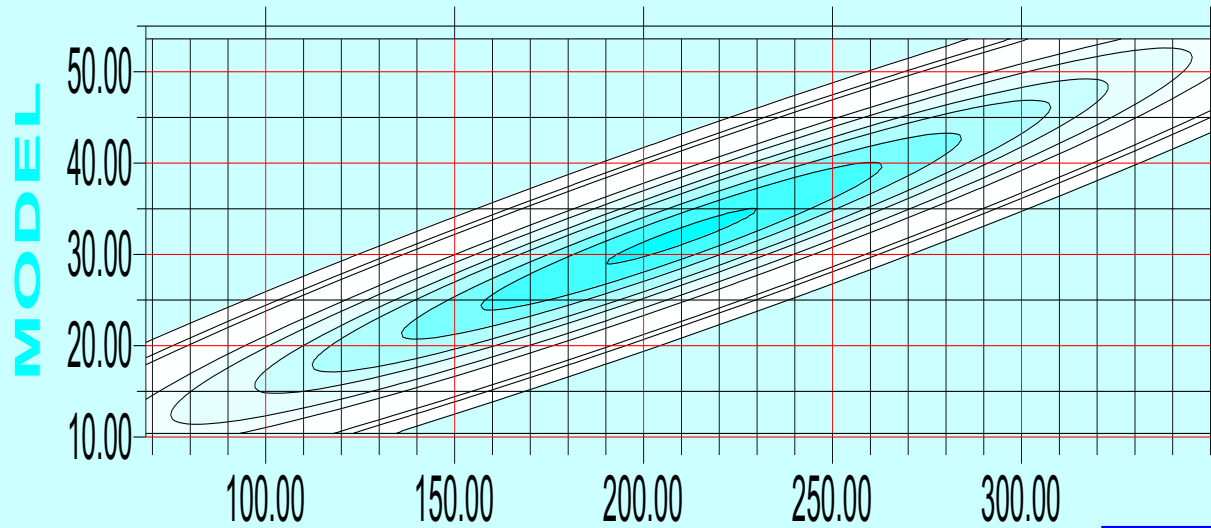
~ 95% correlation



UMC convergence

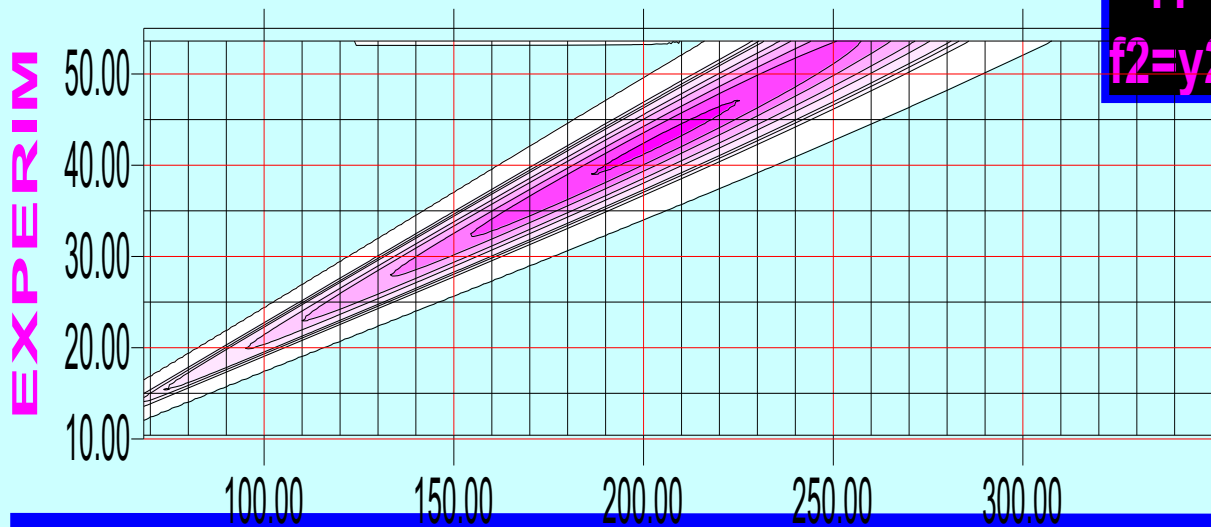


RATIO CASE



MODEL
 $y1=210 \pm 63$ (30%)
 $y2=32 \pm 9.6$ (30%)

$Cov(1,2) = 0.95$

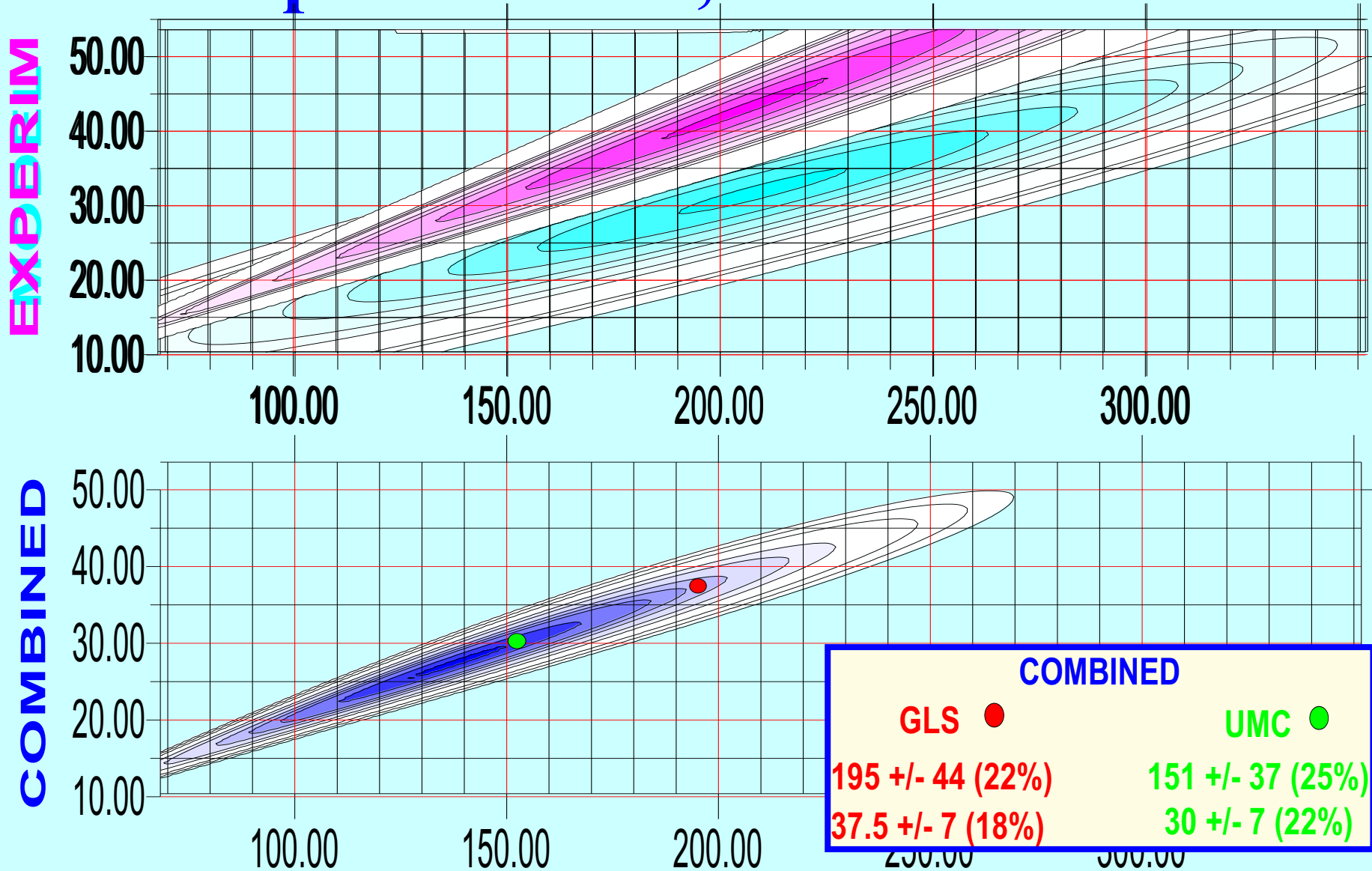


EXPERIM
 $f1=y1=205.6 \pm 61.7$ (30%)
 $f2=y2/y1=0.209 \pm 0.010$ (5%) ~ 43

$Cov(1,2) = 0.$



5% exp. ratio unc., 95% model correl.



GLS FAILURE: ANALYSIS

	<u>Quantity</u>	<u>Node 1</u>	<u>Node 2</u>	<u>Ratio</u>
5% exp. ratio unc.	BF/GLS	0.7767	0.7929	1.0209
95% model correlation	METR/GLS	0.7728	0.7891	1.0210
	METR/BF	0.9950	0.9951	1.0001

	<u>Quantity</u>	<u>Node 1</u>	<u>Node 2</u>	<u>Ratio</u>
5% exp. ratio unc.	BF/GLS	1.0180	0.9795	0.9622
no model correlation	METR/GLS	1.0232	0.9850	0.9626
	METR/BF	1.0051	1.0056	1.0004

	<u>Quantity</u>	<u>Node 1</u>	<u>Node 2</u>	<u>Ratio</u>
30% exp. ratio unc.	BF/GLS	1.0002	1.0007	1.0005
95% model correlation	METR/GLS	0.9995	0.9998	1.0004
	METR/BF	0.9992	0.9991	0.9999



LOG TRANSFORMATION

```
***** LOG(MODEL DATA)
Pmod( 1)= 5.3471 +/- .6000 ( 11.2% )
Pmod( 2)= 3.4657 +/- .6000 ( 17.3% )
```

```
***** LOG(MODEL FUNCTION)
Ymod( 1)= 5.3471
Ymod( 2)= -1.8814
```

```
MODEL CORRELATION MATRIX (PRIOR):
.1000000E+01
.9500000E+00 .1000000E+01
```

```
***** LOG(EXPERIMENTAL DATA)
Yexp( 1)= 5.3259 +/- .3000 ( 5.6% )
Yexp( 2)= -1.5654 +/- .0100 ( -.6% )
```

```
***** ORIGINAL MODEL DATA
Pmod( 1)= 210.0000 +/- 126.0000 ( 60.0% )
Pmod( 2)= 32.0000 +/- 19.2000 ( 60.0% )
```

```
***** ORIGINAL MODEL FUNCTION
Ymod( 1)= 210.0000
Ymod( 2)= .1524
```

```
EXPERIMENTAL CORRELATION MATRIX:
.1000000E+01
.0000000E+00 .1000000E+01
```

```
***** ORIGINAL EXPERIMENTAL DATA
Yexp( 1)= 205.6000 +/- 61.6800 ( 30.0% )
Yexp( 2)= .2090 +/- .0021 ( 1.0% )
```

RESULTS FOR GLS METHOD (LOG TRANSFORMATION):

Mean	Sigma[%]	Red.Chisq
1.999575E+02	2.676447E+01	-1.261845E-02
4.175392E+01	2.677929E+01	1.020535E+02

RESULTS FOR UMC Metropolis (model + exp)

Mean	Sigma[%]	Red.Chisq
1.999971E+02	2.665413E+01	-1.266861E-02
4.176194E+01	2.666949E+01	1.029021E+02

RESULTS FOR GLS METHOD (DIRECT):

Mean	Sigma[%]	Red.Chisq
1.988118E+02	2.779634E+01	2.383451E+03
4.212220E+01	2.001335E+01	-1.662790E+01

RESULTS FOR UMC Metropolis (DIRECT):

Mean	Sigma[%]	Red.Chisq
1.859502E+02	2.391304E+01	5.271764E+01
3.881475E+01	2.387615E+01	-3.378792E-03



SUMMARY

- ❑ UMC is a viable tool for cross-section data evaluation
- ❑ Metropolis sampling scheme is recommended for UMC calculations
- ❑ When data values are cross sections or cross sections and integral (spectrum-averaged) cross sections, GLS and UMC are equivalent so GLS is recommended
- ❑ If ratio or other explicitly non-linear data are introduced UMC may be preferable to GLS



PDF 1. Lognormal distribution

There is a problem associated with utilizing normal distributions if the uncertainties in the variables are large (e.g., typically $> 30\%$).

$$q(z) = (C/z) \exp[-(\ln z - v)^2/(2\sigma^2)] \quad \text{where } 0 < z < +\infty ,$$

$$\text{mean value} = \langle z \rangle = \exp[v + (\sigma^2/2)] ,$$

$$\text{variance} = V_z = \exp(2v + 2\sigma^2) - \exp(2v + \sigma^2) .$$

Best approach: $x = \ln z$

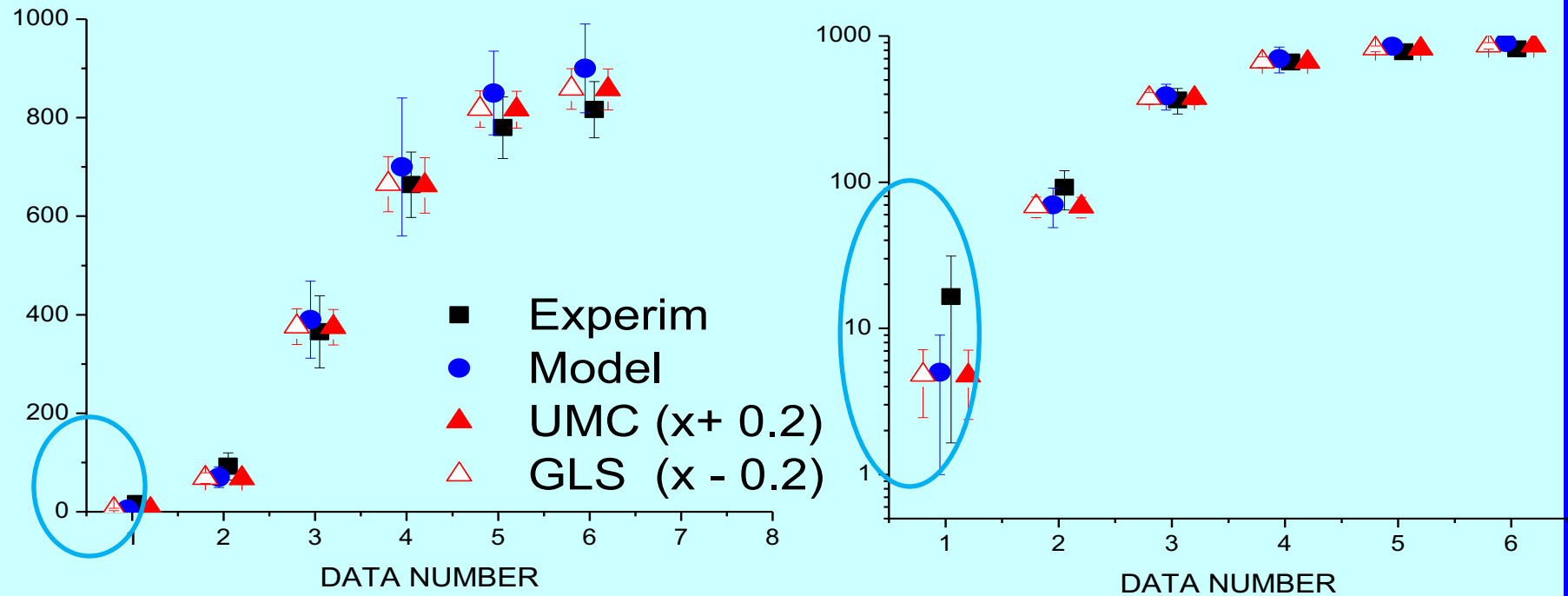
Lognormal PDF in $z \rightarrow$ a normal PDF in x .

If the variable z has an uncertainty E_z , then the uncertainty in $E_x = E_z/z =$ relative error in z .

See paper by Capote R and Smith D in NDST2010 (Korea)



Normal – Lognormal CASE



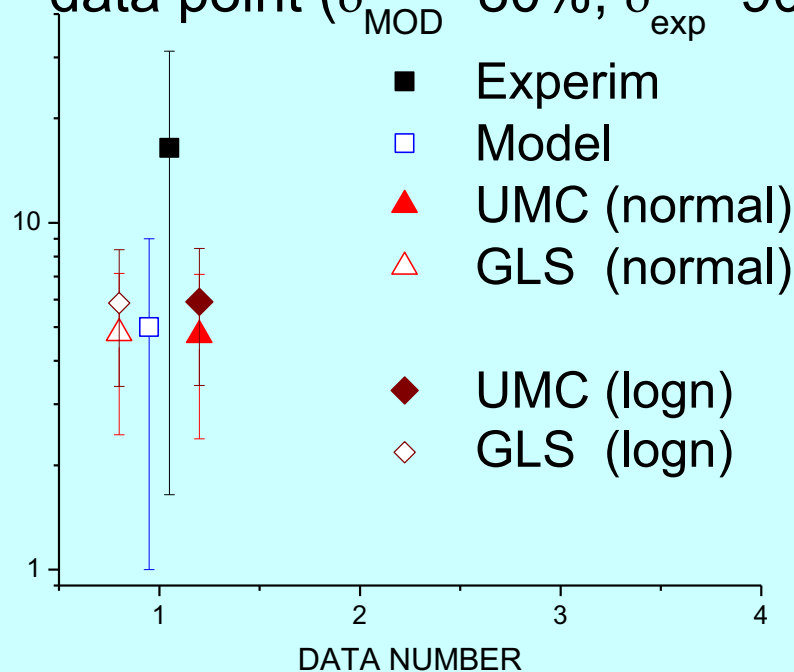
```

***** MODEL DATA
MODEL CORRELATION MATRIX (PRIOR):
Pmod( 1)=  5 +/-  4 ( 80.0% ) .1000000E+01
Pmod( 2)=  70 +/- 21 ( 30.0% ) .9500000E+00 .1000000E+01
Pmod( 3)= 390 +/- 78 ( 20.0% ) .9000000E+00 .9500000E+00 .1000000E+01
Pmod( 4)= 700 +/-140 ( 20.0% ) .8500000E+00 .9000000E+00 .9500000E+00 .1000000E+01
Pmod( 5)= 850 +/- 85 ( 10.0% ) .8000000E+00 .8500000E+00 .9000000E+00 .9500000E+00 .1000000E+01
Pmod( 6)= 900 +/- 90 ( 10.0% ) .7500000E+00 .8000000E+00 .8500000E+00 .9000000E+00 .9500000E+00 .1000000E+01
    
```



RESULTS

1st data point ($\delta_{MOD} = 80\%$, $\delta_{exp} = 90\%$)



GLS (normal)	GLS(logn)	Diff(%)
4.80017	5.87065	22.30102
68.22307	72.07464	5.64556
375.88501	385.8047	2.63902
664.97919	675.7487	1.61953
817.72153	822.9619	0.64085
858.30519	862.9309	0.53894

UMC method

Mean	Sigma [%]
5.9143153E+00	4.2513903E+01
7.2285629E+01	1.4169891E+01
3.8663715E+02	8.9229268E+00
6.7760062E+02	8.1024980E+00
8.2389039E+02	4.4594776E+00
8.6352086E+02	4.7502835E+00

GLS method

Mean	Sigma [%]
5.870653E+00	4.248953E+01
7.207464E+01	1.419907E+01
3.858047E+02	8.931793E+00
6.757487E+02	8.114550E+00
8.229619E+02	4.460754E+00
8.629309E+02	4.751476E+00



PDF 2. Uniform distribution

$a, b \geq 0$ (cross sections)

Uniform distribution

$[1 / (b-a)]$ if $a < z < b$,

$q(z) =$

0 otherwise .

Mean value $\langle z \rangle = (a+b)/2$;

Variance $E_z^2 = [(b-a)^2/12]$

Rel. Error = $(b-a)/(a+b)/\sqrt{3}$

If $a=0 \Rightarrow$ Rel. Error = $1/\sqrt{3} = 57.7\% !!!$



Uniform – normal (no correlations)

$$\langle z \rangle = (a+b)/2$$

1st point – uniform
2nd point – normal

Case 1 (identical intervals):

$$a_{C1} = 10, b_{C1} = 40, \sigma_{C2} = 100 \text{ (20\%)}$$

$$a_{E1} = 10, b_{E1} = 40, \sigma_{E2} = 85 \text{ (15\%)}$$

SOLUTION: $a_{S1} = 10$ and $b_{S1} = 40$.

UMC analysis $\langle \sigma_1 \rangle = 25.00$ (34.6%)

$$\langle \sigma_2 \rangle = 89.33 \text{ (12.0\%)}$$

GLS analysis $\langle \sigma_1 \rangle = 25.00$ (24.5%)

$$\langle \sigma_2 \rangle = 89.33 \text{ (12.0\%)}$$

Case 2 (strong overlap):

$$a_{C1} = 2, b_{C1} = 40$$

$$a_{E1} = 5, b_{E1} = 60$$

SOLUTION: $a_{S1} = 5$ and $b_{S1} = 40$.

UMC analysis $\langle \sigma_1 \rangle = 22.50$ (44.9%)

GLS analysis $\langle \sigma_1 \rangle = 24.72$ (36.5%)

Case 3 (weak overlap):

$$a_{C1} = 10, b_{C1} = 20$$

$$a_{E1} = 16, b_{E1} = 40$$

SOLUTION: $a_{S1} = 16$ and $b_{S1} = 20$.

UMC analysis $\langle \sigma_1 \rangle = 18.00$ (6.4%)

GLS analysis $\langle \sigma_1 \rangle = 16.92$ (15.7%)



PDF 3. exponential– normal (no correl)

A cross section estimate is provided but no uncertainty is given

Exponential distribution

$$q(z) = C \exp(-z/\langle z \rangle) \quad \text{where } 0 < z < \infty \text{ and } \langle z \rangle > 0 ;$$

$$\text{Mean value} = \langle z \rangle, \quad \text{Variance } V_z = \langle z^2 \rangle - \langle z \rangle^2 = E_z^2 = \langle z \rangle^2.$$

Case 1:

UMC

GLS

$$\sigma_{C1} = 210.0 \text{ (30.00\%)}, \sigma_{C2} = 22 \text{ (no uncert) (100\%)}$$

$$\sigma_{E1} = 205.6 \text{ (61.68\%)}, \sigma_{E2} = 34 \text{ (no uncert) (100\%)}$$

SOLUTION:

$$\text{UMC analysis} \quad \langle \sigma_1 \rangle = 208.1 \text{ (21.1\%)}; \quad \langle \sigma_2 \rangle = \mathbf{14.90 \text{ (9.9\%)}}$$

$$\text{GLS analysis} \quad \langle \sigma_1 \rangle = 207.8 \text{ (21.1\%)}; \quad \langle \sigma_2 \rangle = \mathbf{25.54 \text{ (72.3\%)}}$$

The posterior function p includes the product of two simple exponentials. This product equals the exponential factor $\exp(-\sigma_2/\sigma_{m2})$, where $(1/\sigma_{m2}) = (1/\sigma_{C2}) + (1/\sigma_{E2})$. Thus, for the input values given, one obtains $\sigma_{m2} = 13.4$



CONCLUDING REMARKS

Covariance Workshop, Port Jefferson 2008

- ❑ UMC is a viable tool for cross-section data evaluation.
- ❑ UMC Metropolis sampling scheme is recommended.
- ❑ When data values are cross sections or cross sections and integral (spectrum-averaged) cross sections, GLS and UMC are equivalent so GLS is recommended.
- ❑ **If ratio or other explicitly non-linear data are introduced UMC may be preferable to GLS.**

- ❑ GLS is equivalent to UMC for PDFs closely resembling symmetric normal distributions with small to modest uncertainties.
- ❑ UMC results are more reliable for broad or strongly asymmetric distributions.
- ❑ The lognormal distribution can be used to good advantage in certain situations involving large uncertainties.
- ❑ **Use of the uniform and exponential distributions discouraged.**