Calgary • Houston • Denver

RYDER SCOTT

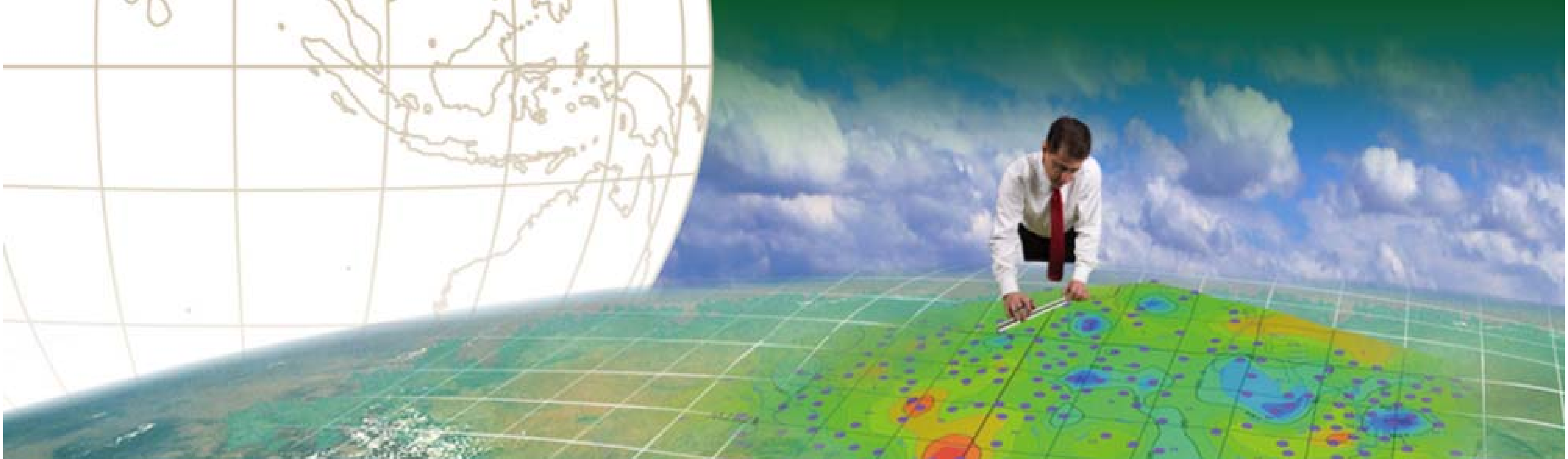
Application of Maximum Likelihood to the Gas Material Balance

Contrast to COGEH Guideline

Ray Dupuis

NIKO RESOURCES LTD

Sept. 5, 2013



COGEH (paraphrased)

- P_{90} … at least a 90% probability that the quantities … will equal or exceed …
- P_{50} … at least a 50% probability …
- P_{10} … at least a 10% probability …

COGEH (direct extraction)

"A quantitative measure of the certainty levels pertaining to estimates prepared for the various reserves categories is desirable to provide a clearer understanding of the associated risks and uncertainties. However, the majority of reserves estimates will be prepared using deterministic methods that do not provide a mathematically derived quantitative measure of probability. In principle, there should be no difference between estimates prepared using probabilistic or deterministic methods."

COGEH Volume 2, page 3-8

COGEH (the problem)

*" … However, the **majority of reserves estimates will be prepared using deterministic methods that do not provide a mathematically derived quantitative measure of probability**."*

- ***This statement is only partly true. Statistical analysis of data fitting a deterministic model can quantify probability.***

COGEH (the solution)

- ***Make estimates based on the method of maximum likelihood.***

- Maximum likelihood …

  sounds exotic,

  looks hard,

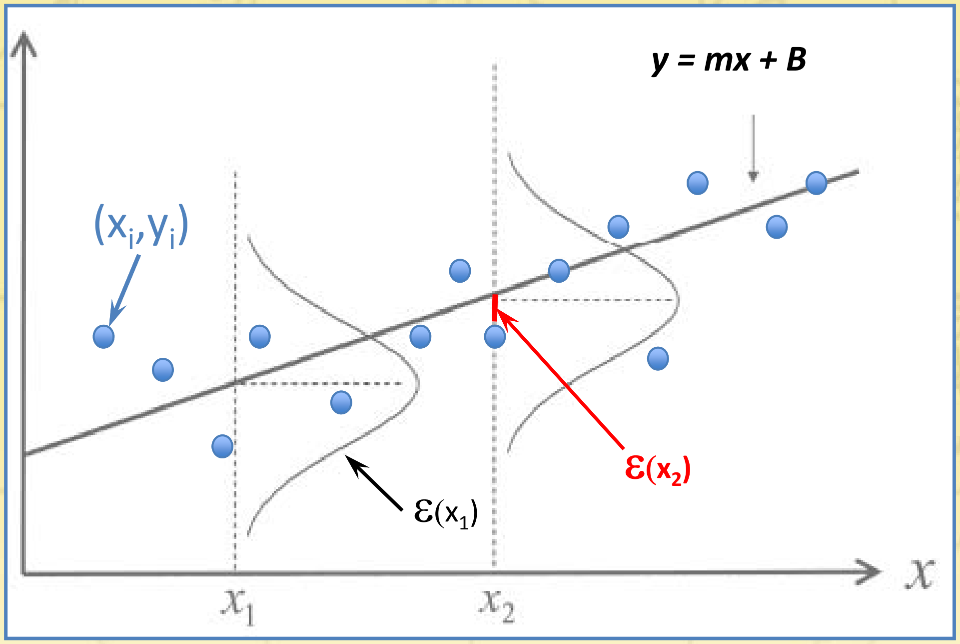
  yet, is quite easy.

Maximum Likelihood?

- Maximum likelihood estimation (MLE) is a strategy for obtaining statistically efficient parameter estimates.

- In layman's terms MLE provides a method which maximizes the likelihood (probability) that measured data fit a model, say for example the equation of a straight line.

Maximum Likelihood?



Model ……… y = mx + B

***Parameters*:**

- **m** = slope of line
- **B** = line's y-intercept

- $\varepsilon(x_i)$ = error at $x = x_i$

  $= y_i - (mx_i + B)$

  $= N(0, \sigma_i^2)$

Statistical Likelihood (looks hard)

The ***Likelihood*** function is defined as:

$$L = \prod_{i=1}^{n} \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{\varepsilon_i^{\,2}}{2\sigma_i^{\,2}}}$$

It is much more convenient for us to use ln(*L*)

$$\ln(L) = \sum_{i=1}^{n} \left[\ln\left(\frac{1}{\sigma_i \sqrt{2\pi}} \right) - \frac{\varepsilon_i^{\,2}}{2\sigma_i^{\,2}} \right]$$

Statistical Likelihood (looks hard)

ln(*L*) is more convenient to use because

- It is easily differentiated with respect to model parameters
- Maxima in ln(*L*) and *L* are at the same places

Gas Material Balance (model)

The conventional volumetric material balance equation for a single-phase gas reservoir is:

$$(P/Z) = (P/Z)_o\left(1 - \frac{G_p}{G}\right) = (P/Z)_o - \frac{(P/Z)_o}{G}G_p$$

Which has the form y = B – mx

Where:

$x_i = Gp_i$, $y_i = (P/Z)_i$

$B = (P/Z)_o$ and $G = B/m = (P/Z)_o/m$

MLE (only looks hard)

The intrinsic error at x_i for the material balance is:

$$\varepsilon_{(x_i)} = y_i - (B - mx_i)$$

$\ln(L)$ then becomes:

$$\ln(L) = \sum_{i=1}^{n} \left[\ln\left(\frac{1}{\sigma_i \sqrt{2\pi}} \right) - \frac{(y_i - B + mx_i)^2}{2\sigma_i^{\,2}} \right]$$

Basic Calculus (refresher)

For $Y = f(X)$

If a maxima exists it is located where $dY/dX = 0$

and

d^2Y/dX^2 must < 0

MLE (starting to look easy)

The maximum likelihood estimates for B and m occur where the following conditions are met:

$$\frac{\partial \ln(L)}{\partial m} = 0 = \sum_{i=1}^{n} \left[\frac{-2mx_i^{\,2} - 2x_i y_i + 2Bx_i}{2\sigma_i^{\,2}} \right]$$

$$\frac{\partial \ln(L)}{\partial B} = 0 = \sum_{i=1}^{n} \left[\frac{-2B + 2y_i + 2mx_i}{2\sigma_i^{\,2}} \right]$$

… since the respective 2nd derivatives are < 0, these are maximum likelihood estimates!

MLE for P/Z Mat. Balance (easy!)

RYDER SCOTT

We have 2 equations with 2 unknowns (B & m) …

Solving for m, B and most importantly G …

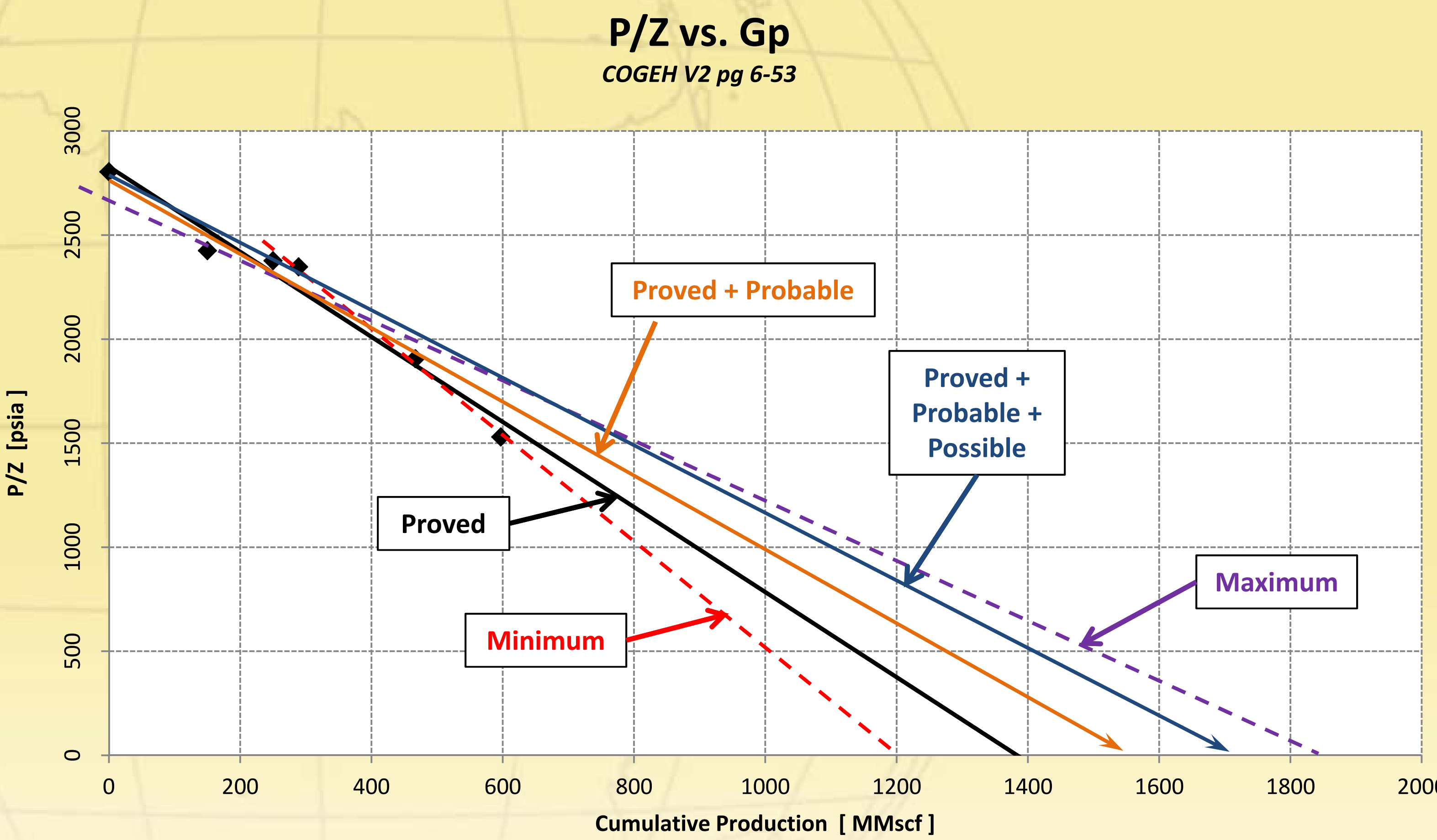
$$m = \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n\sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^{\,2} - \left[\sum_{i=1}^{n} x_i\right]^2}$$

$$B = \left(\frac{P}{Z}\right)_o = \frac{\sum_{i=1}^{n} y_i + m\sum_{i=1}^{n} x_i}{n}$$

$$G = \frac{B}{m} = \frac{(P/Z)_o}{m}$$

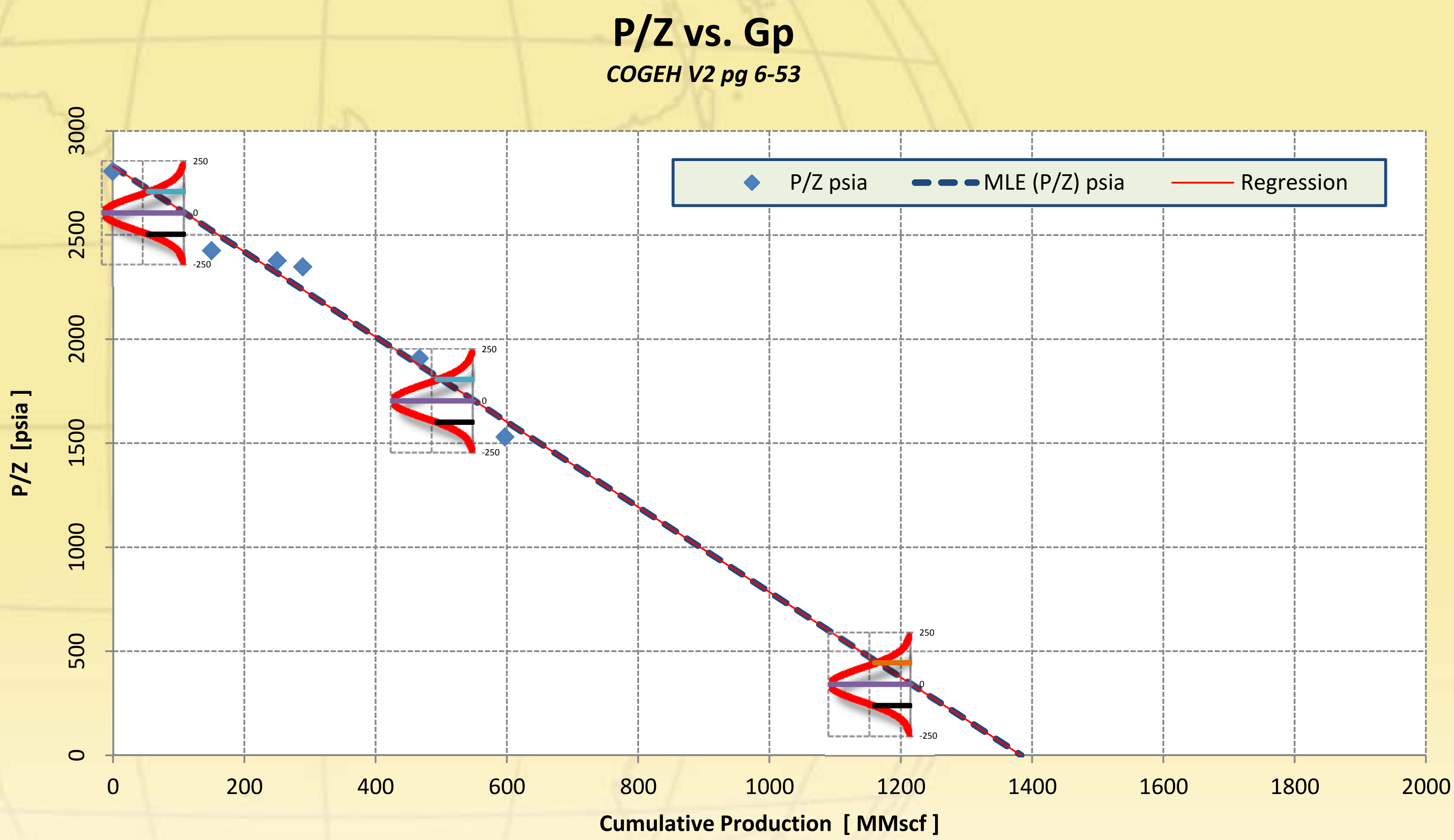
Where: $x_i = Gp_i$, $y_i = (P/Z)_i$

COGEH Mat. Balance (arbitrary)



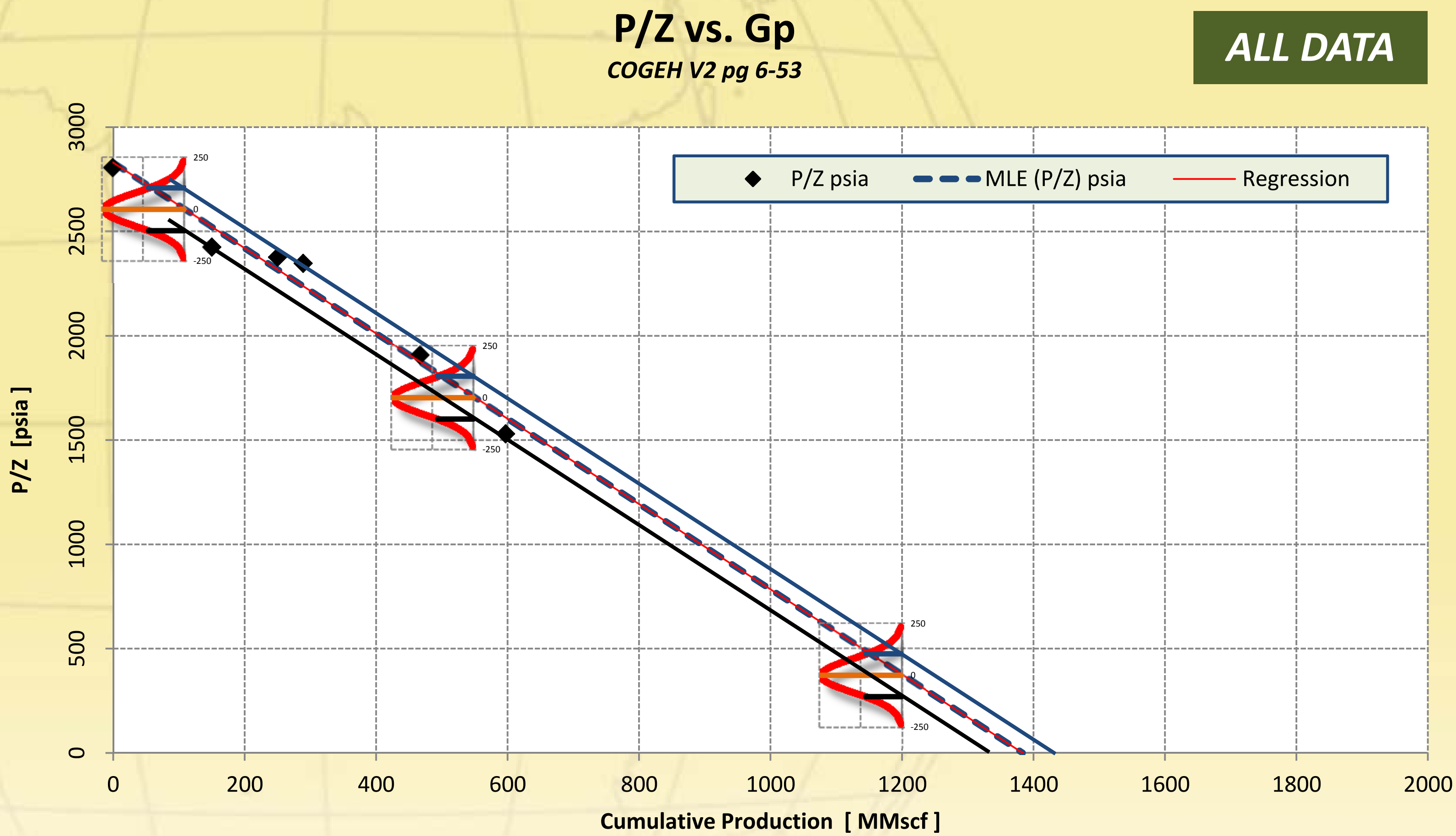
MLE Mat. Balance (statistical)

RYDER SCOTT



MLE Mat. Balance (statistical)

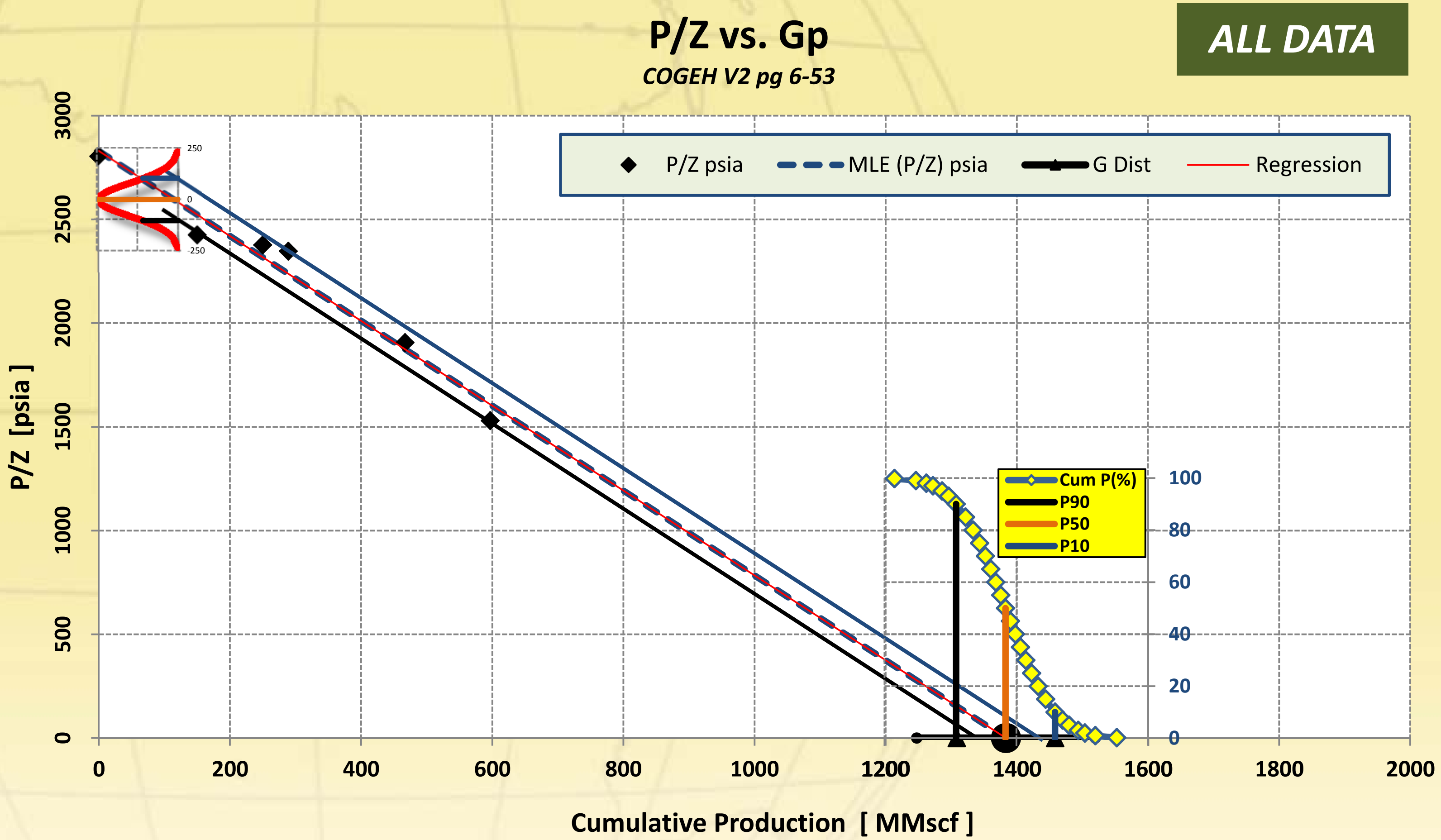
RYDER SCOTT



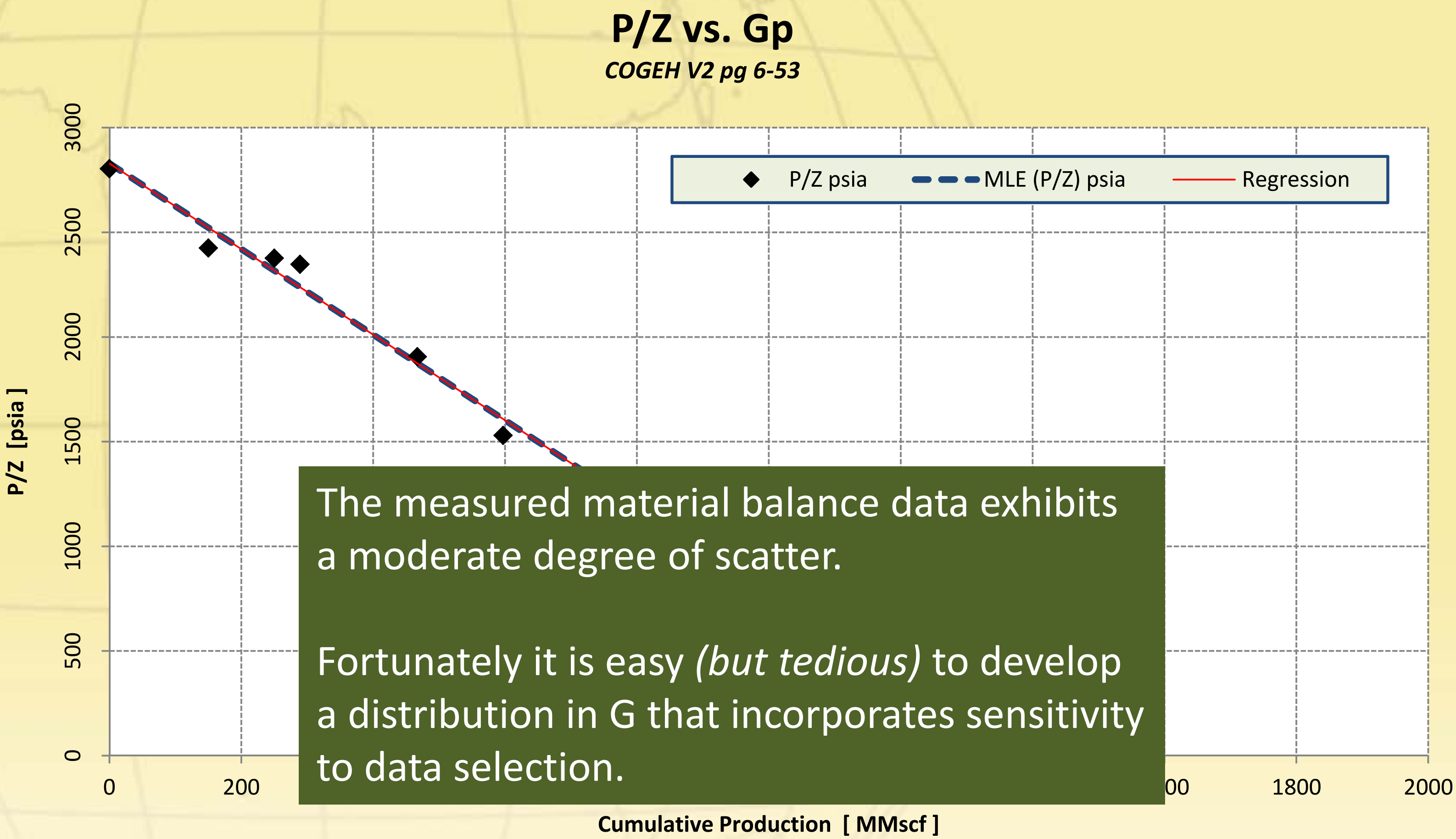
MLE Mat. Balance (statistical)

RYDER SCOTT

ALL DATA



MLE Mat. Balance (data sensitivity)



The measured material balance data exhibits a moderate degree of scatter.

Fortunately it is easy *(but tedious)* to develop a distribution in G that incorporates sensitivity to data selection.

MLE Mat. Balance (data sensitivity)

"n" normal distributions representing estimates of G can be derived by applying MLE to data subsets.

The overall distribution of G by simple averaging is:

$$G(\mu_G, \sigma_G{}^2) = \frac{1}{n}\sum_{i=1}^{n} G_i(\mu_i, \sigma_i{}^2)$$

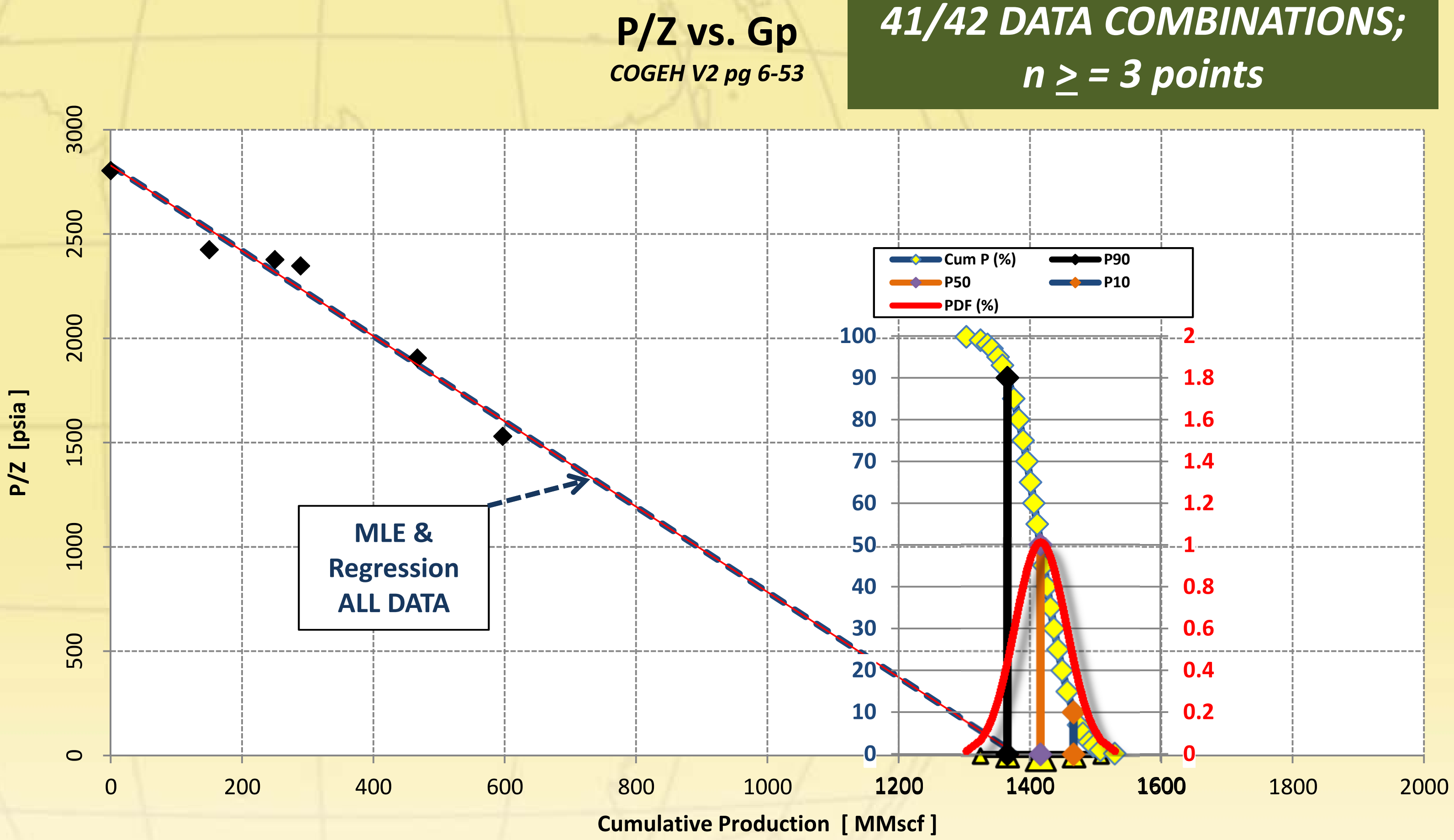
$$\mu_G = \frac{1}{n}\sum_{i=1}^{n} \mu_i$$

$$\sigma_G{}^2 = \frac{1}{n}\sum_{i=1}^{n} \sigma_i{}^2$$

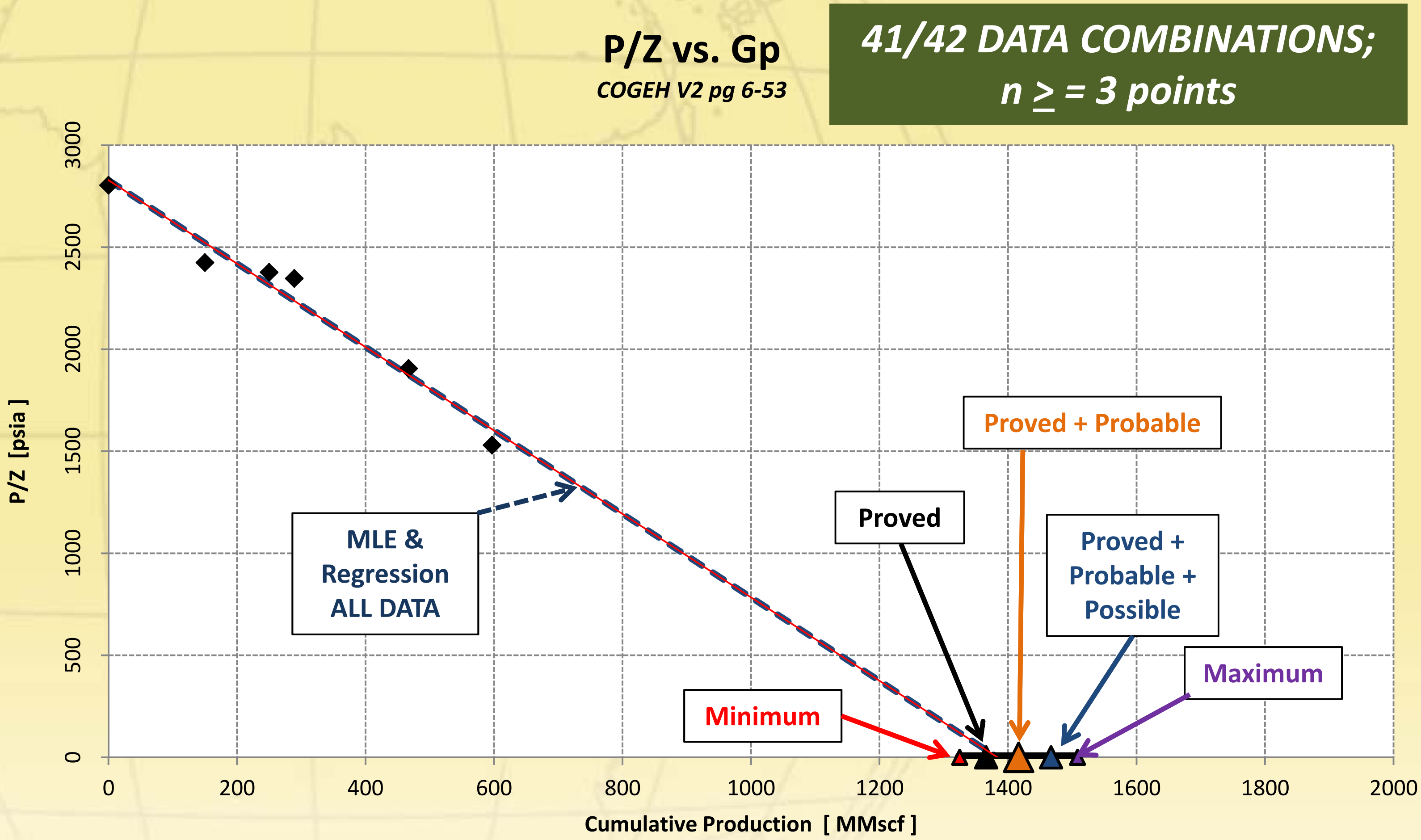
NOTE: **A weighted average formula for G, if desired, has an analogous form.**

MLE Mat. Balance (data sensitivity)

41/42 DATA COMBINATIONS;
n > = 3 points

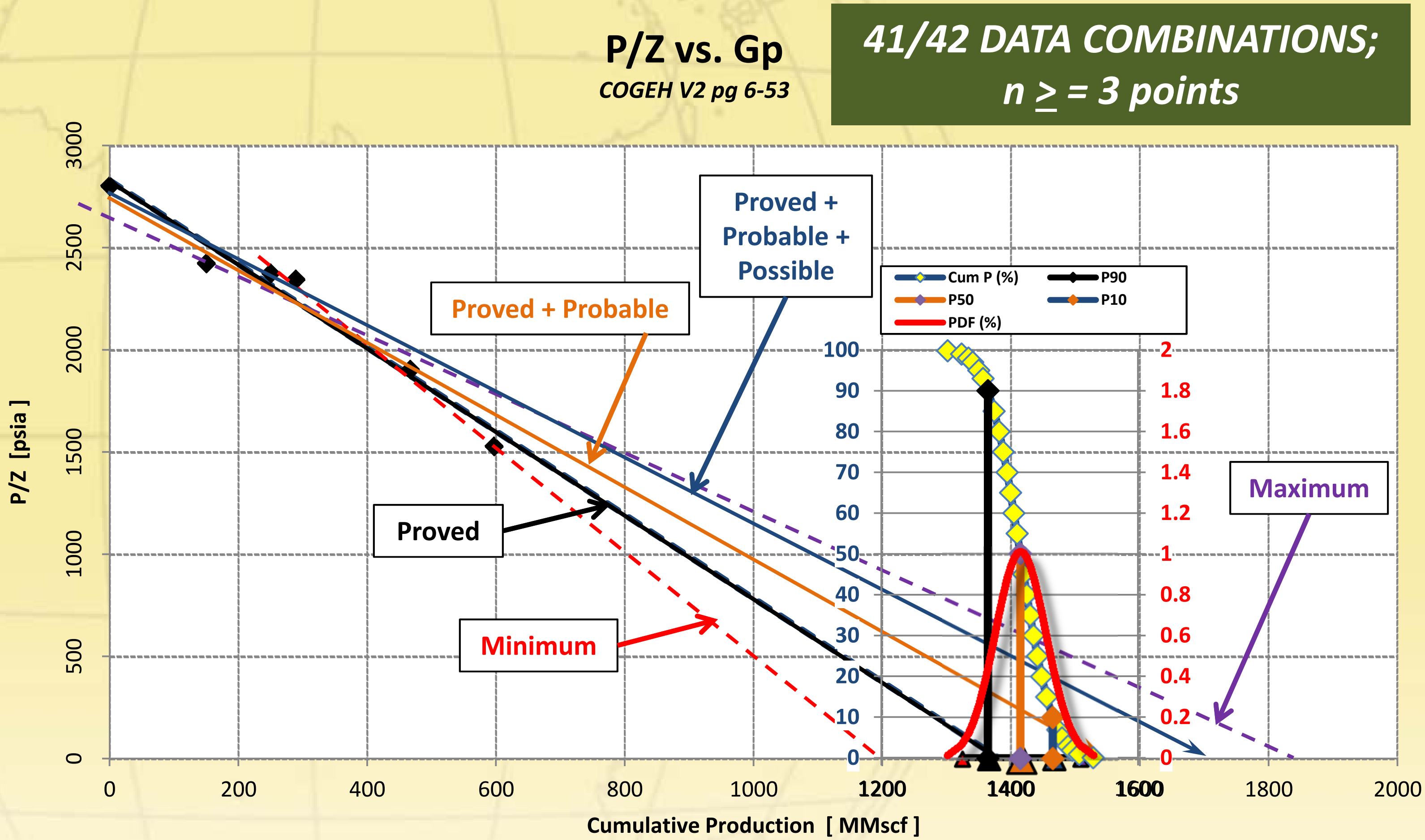


MLE Mat. Balance Results



MLE vs. COGEH

RYDER SCOTT



MLE vs. COGEH

	G (MMscf)		Δ
	COGEH	MLE*	(MLE-COGEH)
Minimum	1200	1181	-1.62%
P_{99}		1325	
Pv (P_{90})	1384	1366	-1.26%
Pv+Pb (P_{50})	1550	1417	-8.61%
Pv+Pb+Ps (P_{10})	1710	1467	-14.21%
P_{01}		1508	
Maximum	1840	1760	-4.36%

* 41/42 DATA COMBINATIONS; n $\geq$ 3 points

Conclusions

- The results of deterministic analyses can be assigned quantitative measures of probability.

- In this simple material balance example, the differences between COGEH P_{90}, P_{50} and P_{10} values of G and those estimated using probabilistic analysis of the data are large enough to be of concern. The differences are more pronounced at the P_{50} and P_{10} levels.