JAGUAR

LAND ROVER

A Statistical Method for Particle Number Emissions Measurement Variance Analysis

Mike Braisher
Cambridge Particles Meeting
May 2011

Contents

- **Problem outline and motivation**
- **Traditional approach to variance estimation**
- **An alternative approach to variance estimation**
  - > A statistical model
  - > heteroskedasticity
- **Results: Application of alternative approach**
- **Demonstration of instrument bias correction**
- **Summary**

Contents

Problem outline and motivation

- **Particle number measurement:**
  - > European emissions regulations introduce a particle number limit ($6x10^{11}$ per km) for homologation of specific Diesel vehicle types.
- **Experimental error:**
  - > When an experiment is repeated under what are as nearly as possible, the same conditions, the observed results are never quite identical (Box, Hunter and Hunter, 1978).
- **Key questions when undertaking any measurement:**
  - > How robust is the result obtained?
  - > How much variability in repeated results is attributable to the measurement system and how much to the part?
- **Importance of understanding repeatability:**
  - > Tolerances and target setting

Contents

Traditional approach to variance estimation

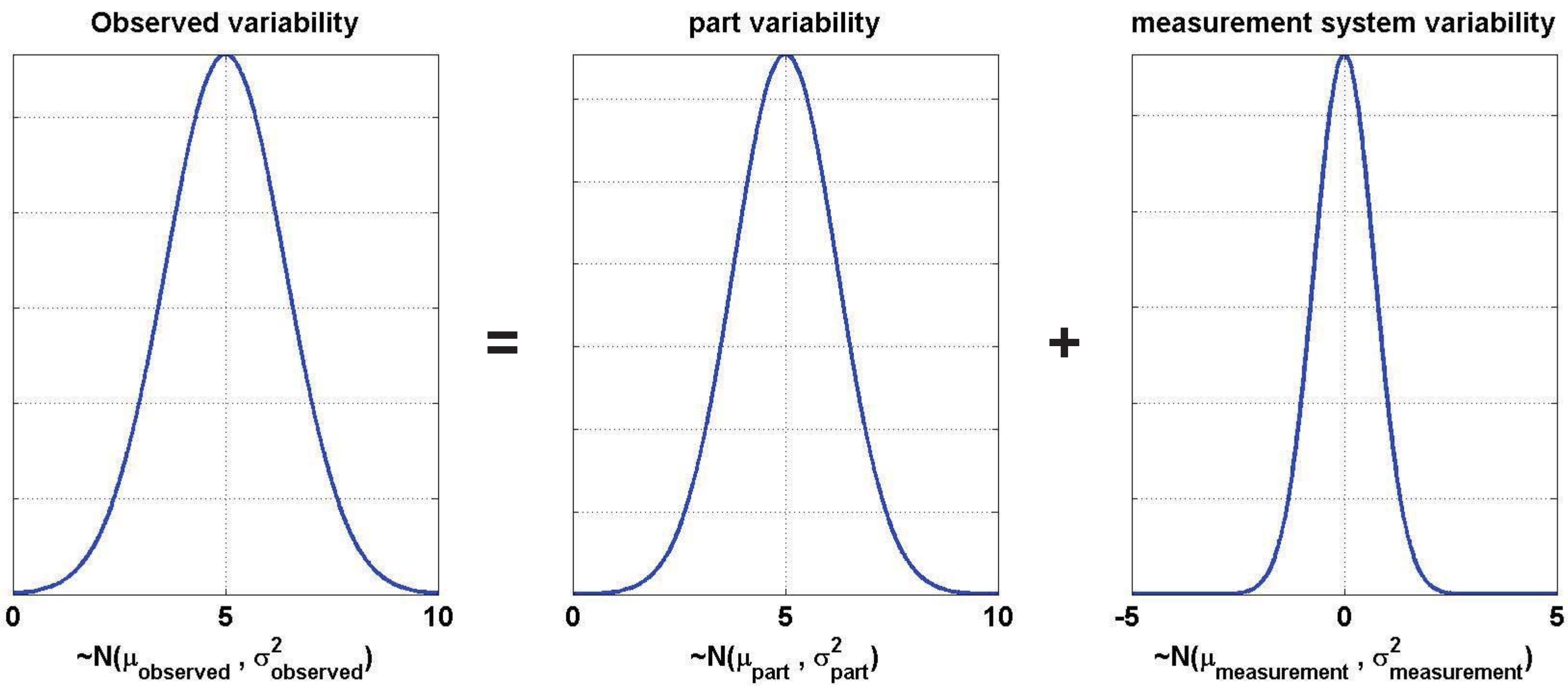

Figure 1

- If the characteristic of the part being measured does not change, then the variability in results is assigned to the measurement system variance.
- **Problem**: Difficult to find a repeatable source of particulate matter.

Contents

- Problem outline and motivation
- Traditional approach to variance estimation
- **An alternative approach to variance estimation**
  - > A statistical model
  - > heteroskedasticity
- Results: Application of alternative approach
- Demonstration of instrument bias correction
- Summary

An alternative approach to variance estimation

- **Grubbs (1948) and Box, Hunter and Hunter (1978) address the problem of estimating instrument variance with destructive testing...**
  - > Taking simultaneous (2 or more) measurements removes the requirement for a repeatable source.
  - > Sample variance is calculated from the simultaneous measurement results.
  - > Sample variance is averaged across repeat experiments.

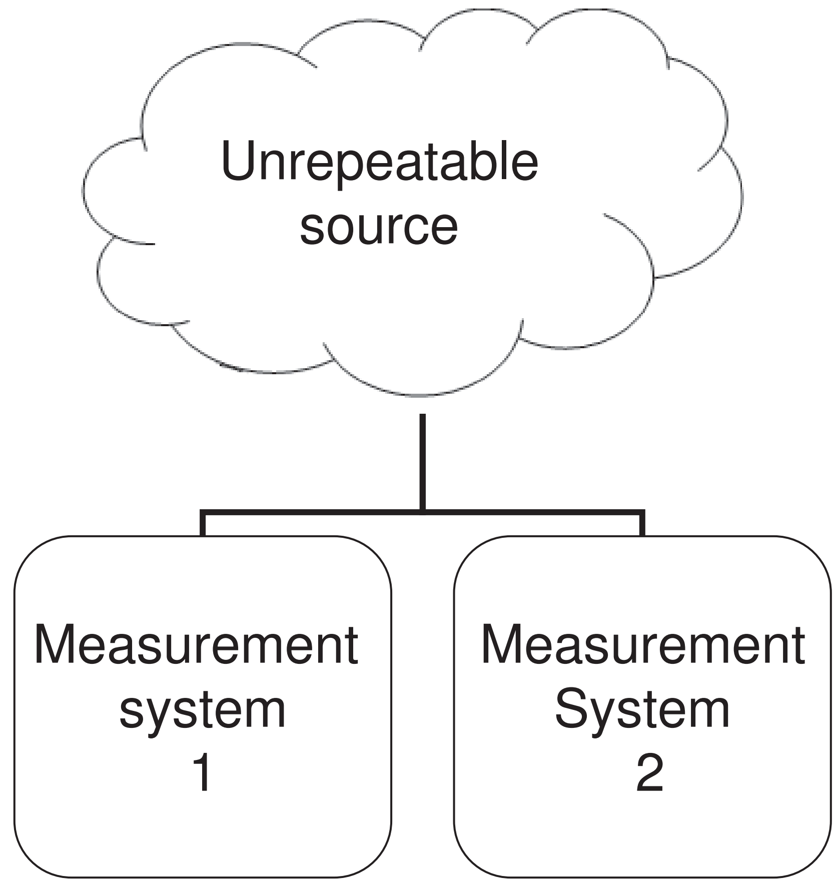

Figure 2

Application to PN measurement

JAGUAR | LAND ROVER

- Particle Number (PN) emissions can be thought of as destructive and not necessarily repeatable.
- Taking simultaneous measurements on the same PN emissions test removes the requirement for a repeatable source.

- **Challenges**
  - > PN emissions span several orders of magnitude (10^9 – 10^{14} #/km)
  - > Variance is not constant across the range of measurements

Statistical model for PN measurements

Model

- Denote measurement results taken with instrument 1 as *PN*1 and results taken with instrument 2 as *PN*2.

Observed values

$$\begin{cases} PN1 = \alpha x + e_1 \\ PN2 = \beta x + e_2 \end{cases}$$

Experimental error terms

Scaling parameter ("bias" term)

True value of PN (not known)

Assumptions:

> error terms for the two counters are sampled from identical distributions (independent and identically distributed random variables).

> Bias scaling factors (α,β) are linear and stable!

Sample statistics

- For a sample size of 2 (i.e. a pair of measurements *PN*1 and *PN*2), **sample variance** is given by equation 1 (Box, Hunter and Hunter 1978).

$$S^2 = \frac{(PN1 - PN2)^2}{2}$$ ……………..Equation 1

- The **sample average** is given by equation 2…

$$X = \frac{PN1 + PN2}{2}$$ ……………..Equation 2

Sample variance (…when α = β)

- *PN1 – PN2* is given by equation 3

$$PN1 - PN2 = (\alpha - \beta)x + (e_1 - e_2)$$ ……Equation 3

- If α = β, then the sample variance reduces to equation 4

$$S^2 = \frac{(e_1 - e_2)^2}{2}$$ ……………..Equation 4

Thus, for the case of α = β, the sample variance for each pair results reduces to an estimate of the measurement variance.

Heteroskedasticity

- **Problem**: As x (true value of PN) increases, the scatter in observations from a PN counter also increases (data is said to be heteroskedastic)…

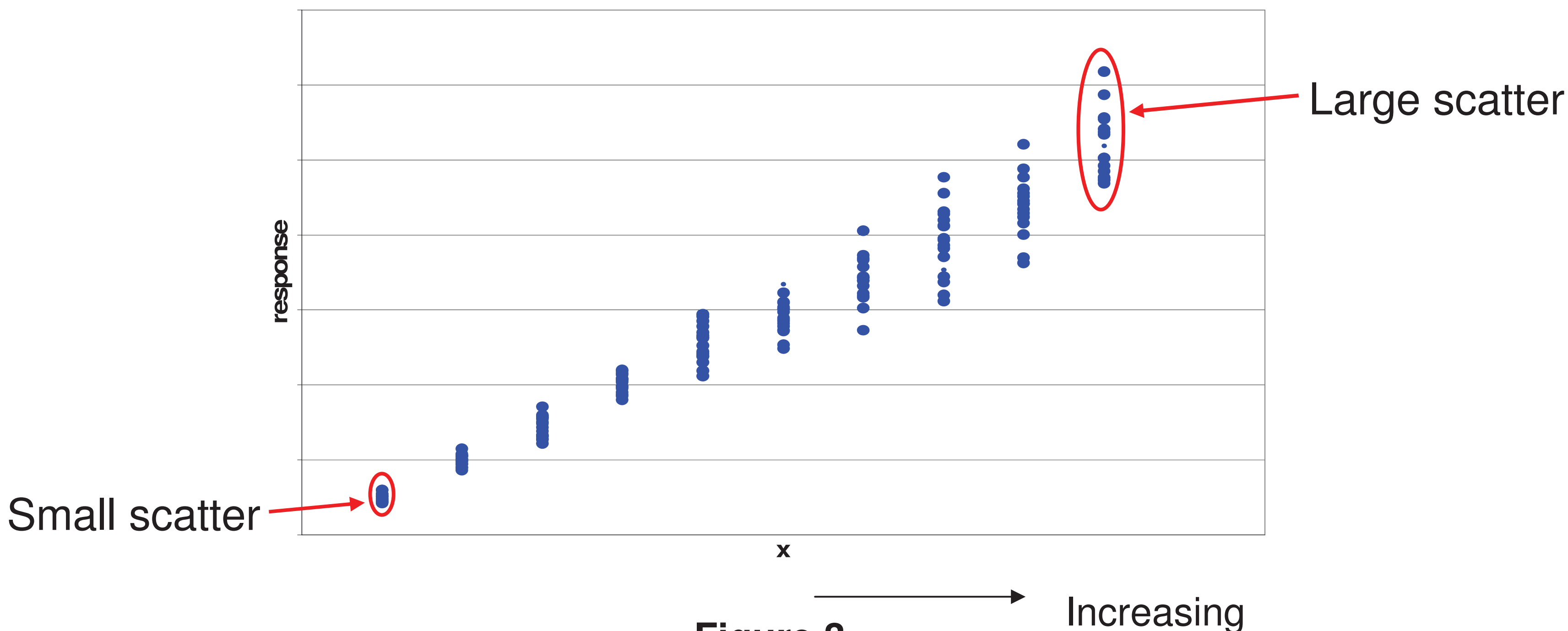

Figure 3

- This indicates a relationship between standard deviation in response (S) and X.

Summary of proposed method…

1. Take multiple measurements from vehicles with two PN counters

Results pairing	Observation counter 1	Observation counter 2
1	*PN1*	*PN2*
2	*PN1*	*PN2*
3	*PN1*	*PN2*
n	*PN1*	*PN2*

2. Calculate X and S for **each pair** (Equations 1 and 2).

3. Perform regression analysis between log(S) and log(X).

Assumption: $\alpha = \beta$

Contents

Application to PN data

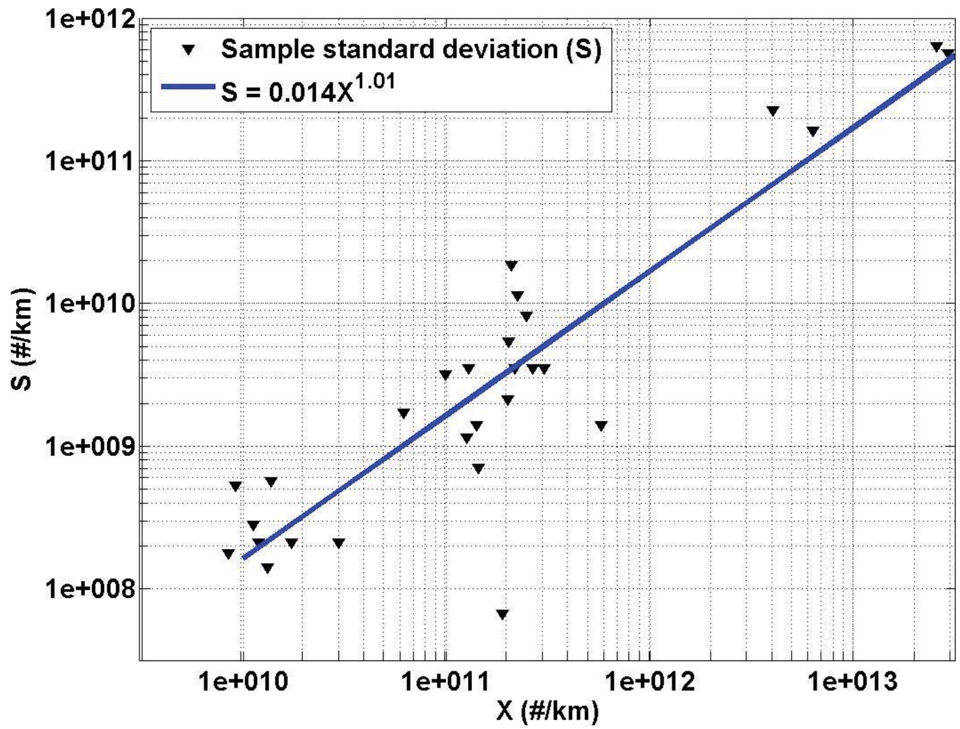

Figure 4

- Series of simultaneous measurements taken from a range of vehicles.
- Regression indicates:
  - > Coefficient of Variance (S/X) of 1.76% - 1.89% across the range 1×10^{10} – 1×10^{13} per km.
  - > At 6×10^{11} per km, S = 1×10^{10} per km.

Assumption: $\alpha = \beta$

Application of variance estimation

- Estimates of variance can be used to determine the capability of the instruments in resolving true changes in PN emissions:
  - > At $6\text{x}10^{11}$ per km, figure 4 indicates $\sigma = 1\text{x}10^{10}$ per km.

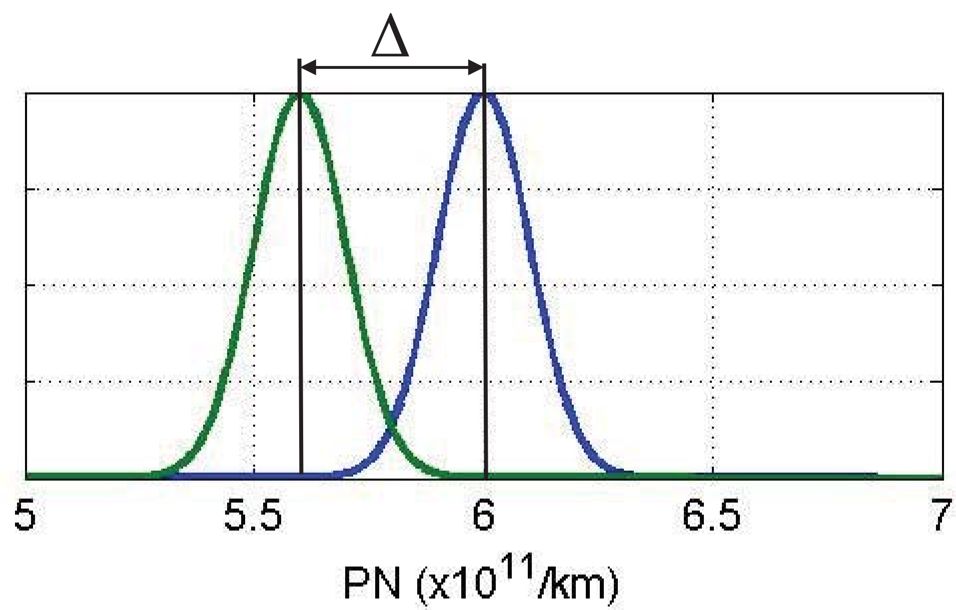

Figure 5

- Difference of 4σ between two peaks accounts for a probability of 95%

$\Delta = 4\sigma = 4\text{x}10^{10}$/km

Assumption: Errors distributed normally.

Instrument detectability

- Extending this 4σ principle across the range, the "detectability" of the instrument is shown in Figure 6…

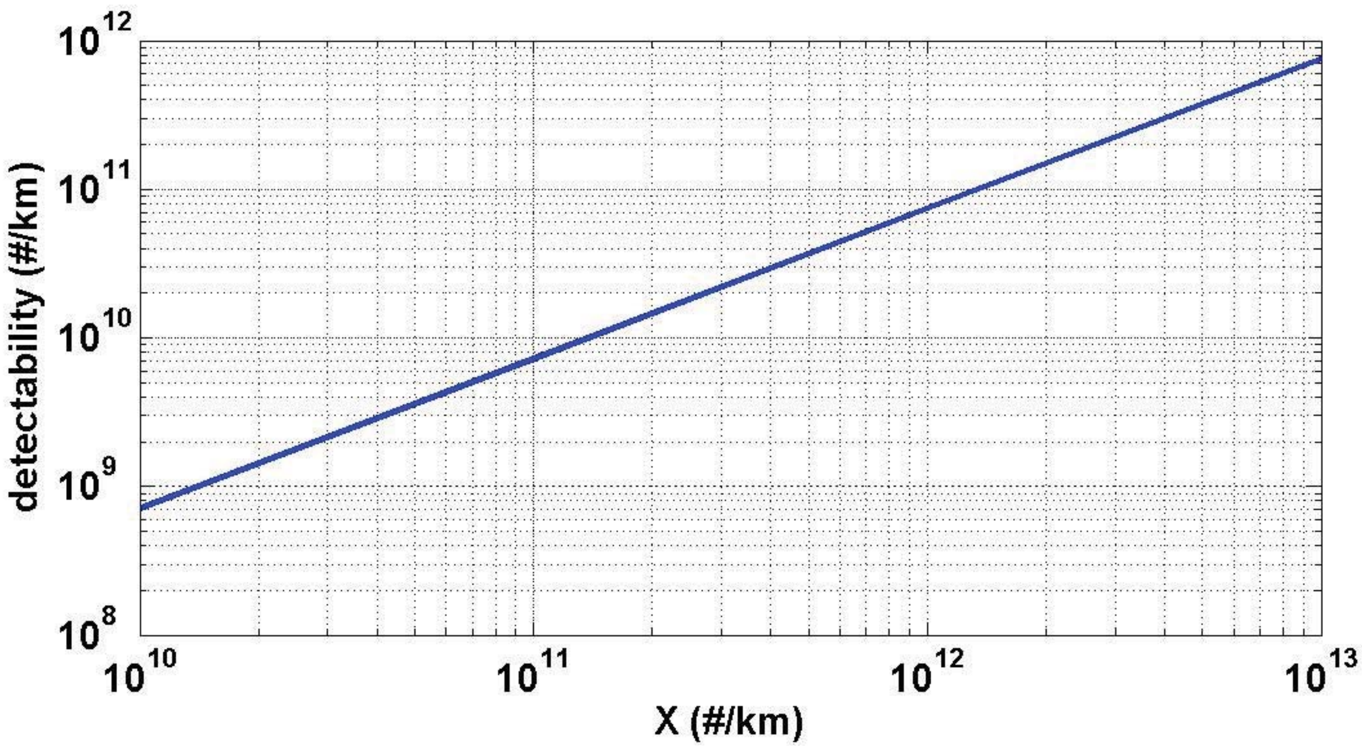

Figure 6

- **Detectability = 4σ = 0.056 $X^{1.01}$**

Contents

- Problem outline and motivation
- Traditional approach to variance estimation
- An alternative approach to variance estimation
  - A statistical model
  - heteroskedasticity
- Results: Application of alternative approach
- **Demonstration of instrument bias correction**
- Summary

Case where α ≠ β (correcting for instrument bias)

- **Problem:** for the case where α ≠ β, estimate of S^2 becomes…

$$S^2 = \frac{\left((\alpha - \beta)x + (e_1 - e_2)\right)^2}{2} \quad \text{................Equation 5}$$

… which acts to <u>over-estimate</u> measurement variance (assuming *x* is bigger than *e*).

- **Solution:** introduce a scaling parameter *c* (Equation 6)…

$$S^2 = \frac{\left(PN1 - cPN2\right)^2}{2} \quad \text{................Equation 6}$$

…and vary *c* to minimise S^2 (case where $c = \alpha/\beta$) by re-running the regression analysis and plotting S as a function of *c*.

Demonstration of correcting for Instrument bias (α ≠ β)

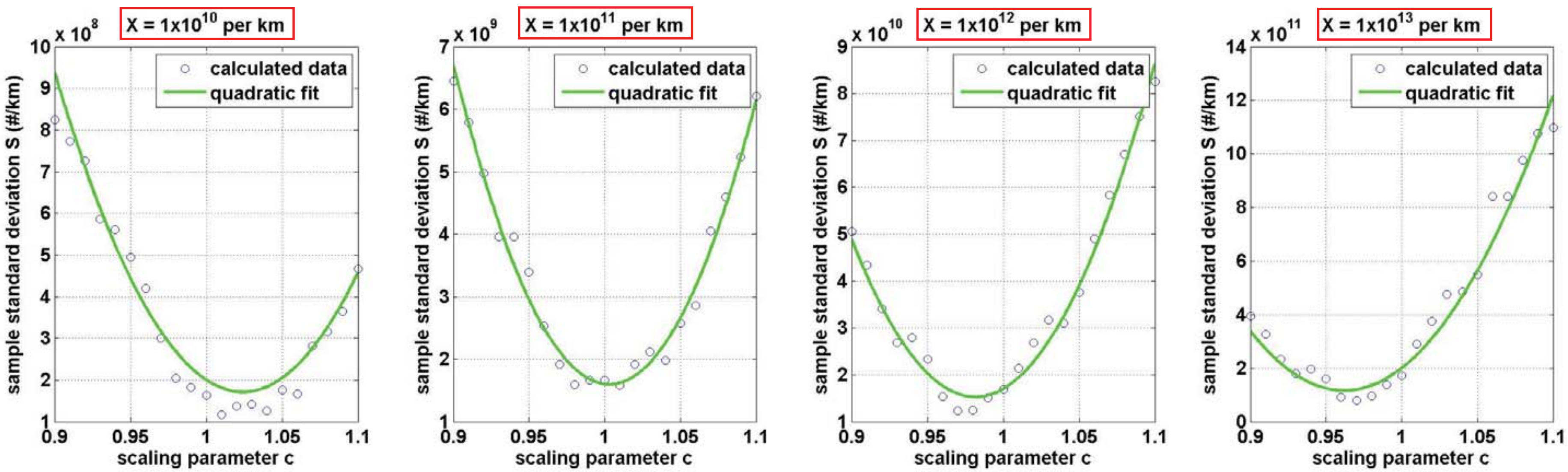

Figure 7

- For all four chosen values of X, S is minimised near a c-value of 1, indicating little or no relative bias between the instruments (α = β).
- Slight differences for minimum c-values across the range indicate either:
  - Non-linear bias factors
  - Non-steady bias factors

Assumptions within the model

- Assumes that the error distributions from two counters are identical.
- Assumes bias scaling factors (α,β) are linear across the range and stable over the course of the data collection.
- Assumes a power relationship between variance (σ^2) and x.
- Ordinary least squares regression performed when there will be errors in the "independent variable" log(X).
- Assumes, for the calculation of instrument discrimination, that the errors are normally distributed.

Contents

Summary

- **A statistical approach to Particle Number measurement repeatability estimation has been presented based on:**
  - > Treating the emissions tests as destructive, removing the requirement for a repeatable source of particle emissions;
  - > Taking simultaneous observations across multiple tests;
  - > Adjusting for relative instrument bias ($\alpha \neq \beta$);
  - > Performing regression analysis to model variance as a function of x (PN data is heteroskedastic);
  - > The assumptions highlighted.
- **An example calculation demonstrates the instrument "detectability" to be $4x10^{10}$ per km at an emissions level of $6x10^{11}$ per km, for a single emissions test.**

References

Box, G.E.P., Hunter, W.G. and Hunter, J.S. (1978), “Statistics for Experimeters”, *Wiley-Interscience*.

Grubbs, F.E. (1948), “On estimating the precision of measuring instruments and product variability”, *Journal of the American Statistical Association*, Vol. 43, No. 242, 243-264.

Thank you for listening. Any comments greatly appreciated.