Noble Complexity in *In-Vivo* Spectroscopy

Christopher D. Brown, Ph.D.
chris.brown@inlightsolutions.com

InLight Solutions, Inc.
Albuquerque, NM
www.inlightsolutions.com

Glucose Validation Study (summer 1999)§

{ 40 subjects (validation), 7 week duration }

plot of estimated *in-vivo* glucose vs. YSI reference

{Scatterplot removed at the request of the legal department prior to distribution of copies of the slides via mail-out.}

What does this mean?

A Learning Philosophy

"The … (inductive) [method] derives axioms from the senses and particulars rising by a gradual and unbroken assent, that it arrives at the most general axiom last of all."

Francis Bacon, in *Novum Organum*

{we aspire to avoid the gradual where possible}

Outline

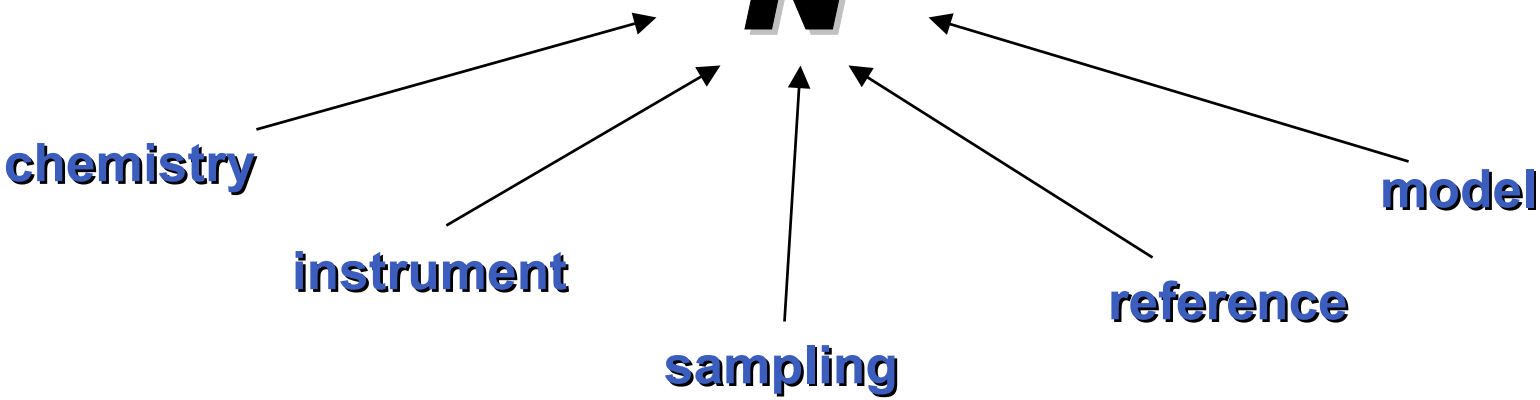
- Reference and noise in multivariate systems
- Signal and noise in multivariate systems
- Quirks of calibration in complex systems
- Verification and validation

Signal and Noise in Multivariate Systems

"S is that S which yields optimal multivariate SNR"

{net analyte signal/response}

$$\frac{S^*}{N^*}$$



Noise Categories

Artificial Noise: *'my fault' noise*

- **Reference Noise:** noise in reference values
  - **calibration error**
  - **validation error**
- **Model Noise:** noise arising from the misspecification of the model

Spectral Noise: *characteristic of the system*

- **Chemical Noise:** noise from any other active chemical species in the matrix (*e.g.*, *in-vivo* ~10^5 potentially absorbing species)
- **Instrumental Noise:** noise from spurious instrumental signals (*e.g.*, dark current, drift)
- **Sampling Noise:** noise arising due to the nature of the sample itself (*e.g.*, heterogeneity)

Noise Categories

Artificial Noise: *'my fault' noise*

- **Reference Noise:** noise in reference values
  - **calibration error**
  - **validation error**
- **Model Noise:** noise arising from the misspecification of the model

Spectral Noise: *characteristic of the system*

- **Chemical Noise:** noise from any other active chemical species in the matrix (*e.g.*, *in-vivo* ~10^5 potentially absorbing species)
- **Instrumental Noise:** noise from spurious instrumental signals (*e.g.*, dark current, drift)
- **Sampling Noise:** noise arising due to the nature of the sample itself (*e.g.*, heterogeneity)

Reference Error: *Calibration*

- **calibration reference error:** *inhibits* the ability of the modeling technique to resolve the *net analyte signal*

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}_y$$

$$\hat{\mathbf{b}} = \mathbf{X}^{\ddagger}\mathbf{y}$$

$\mathbf{X}$ – spectra
$\mathbf{y}$ – concentrations
$\mathbf{e}$ – reference error
$\mathbf{b}$ – regression vector

$$\mathrm{var}\left(\hat{\mathbf{b}}\right) = s_{e_y}^2 \left(\mathbf{X}^{\mathrm{T}}\mathbf{X}\right)^{\ddagger}$$

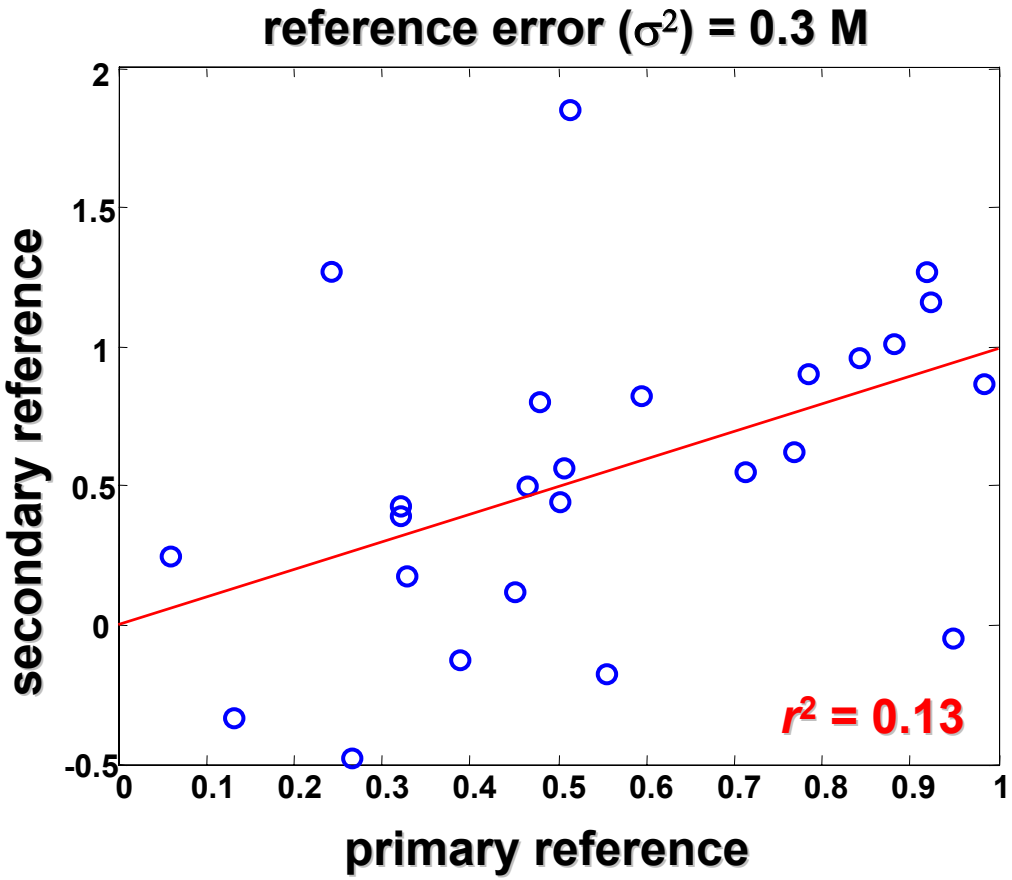
… but the calibration equation ***is waiting*** for reference error

*** error symmetry:** least-squares is reasonably robust to many deviations from *iid* (in $\mathbf{y}$) *if* the error distribution is symmetric

reference method must represent reality *in the limit*

Reference Error: *Calibration*

Is this calibration data of any use whatsoever?



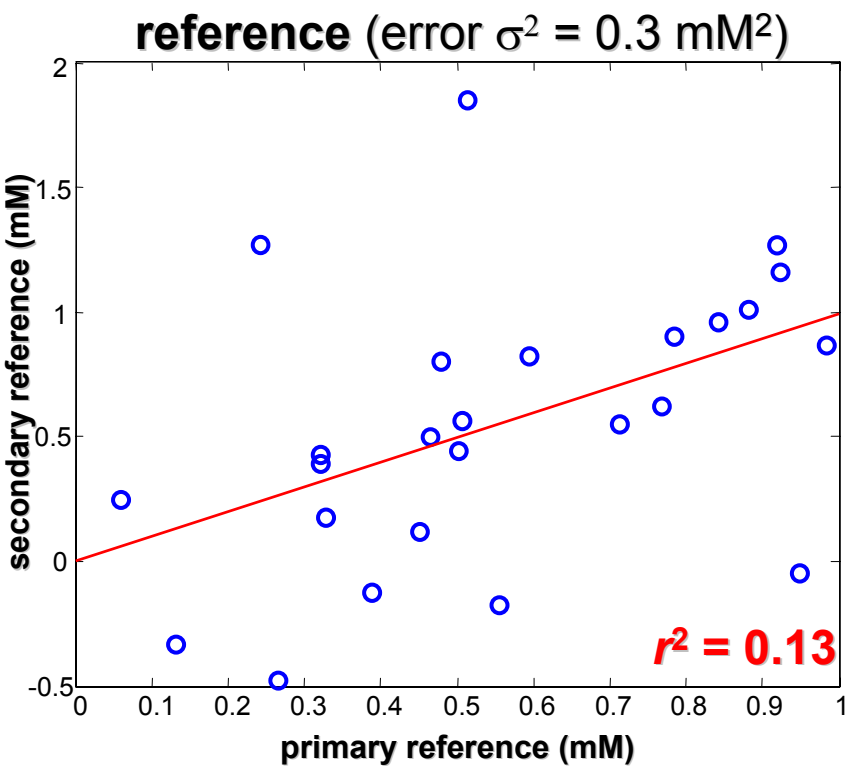
Note: according to this data, *r* is not significantly different from 0.

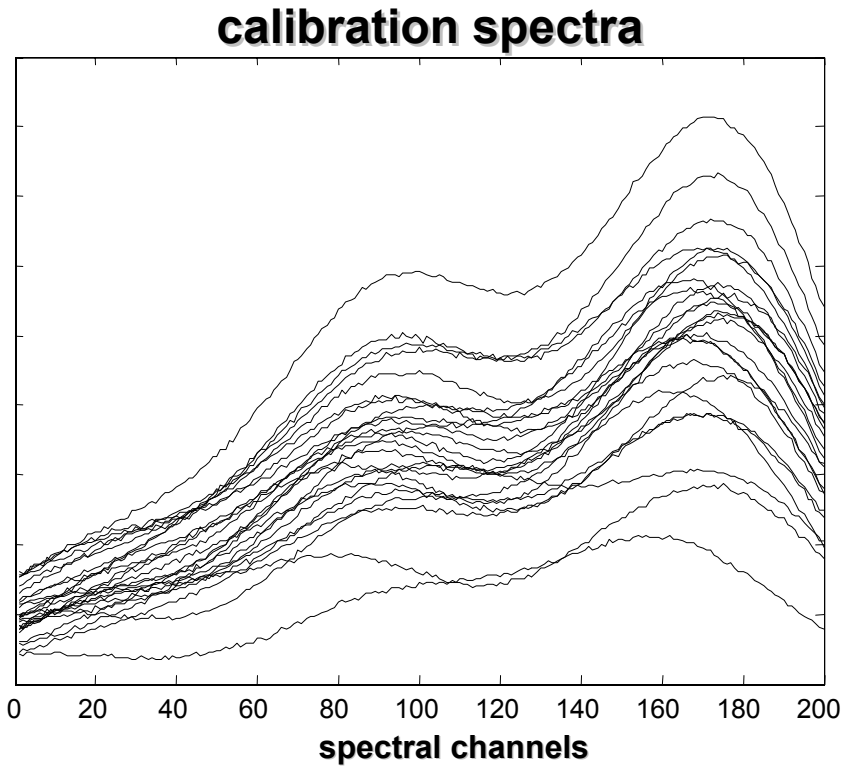
[95% limit: 0.381]

$$\mathbf{SNR_y} \sim 1 \left\{ \frac{\|\mathbf{y}^o\|}{\|\mathbf{e}_{ref}\|} \right\}$$

Reference Error: *Calibration*

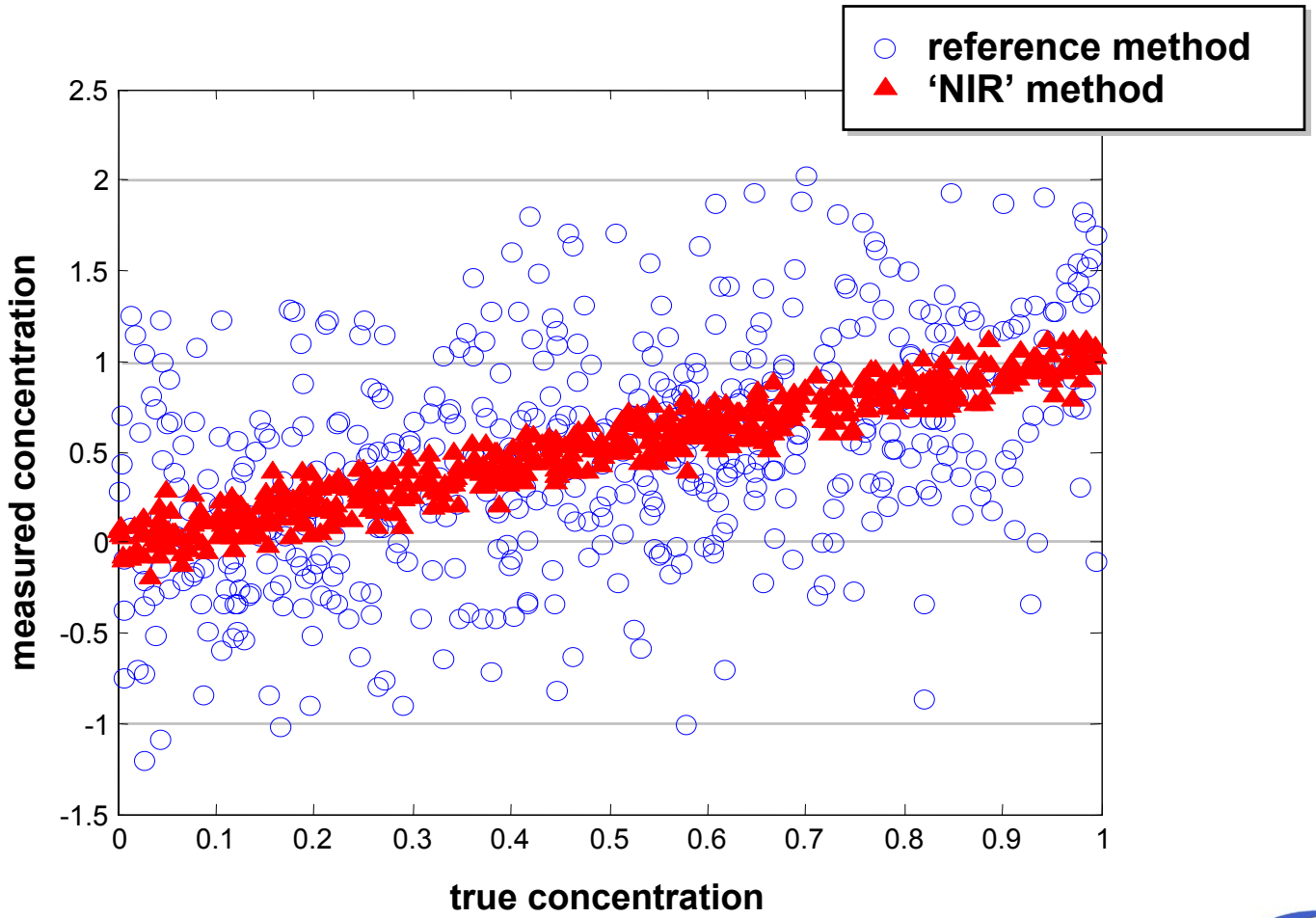
Is this calibration data of any use whatsoever?





Reference Error: *Calibration*

{ **Calibration Data:** 25 samples, reference range 0-1, reference error $\sigma^2_{ref} = 0.3$, classical SNR is about 500 }



Reference Error: *Validation*

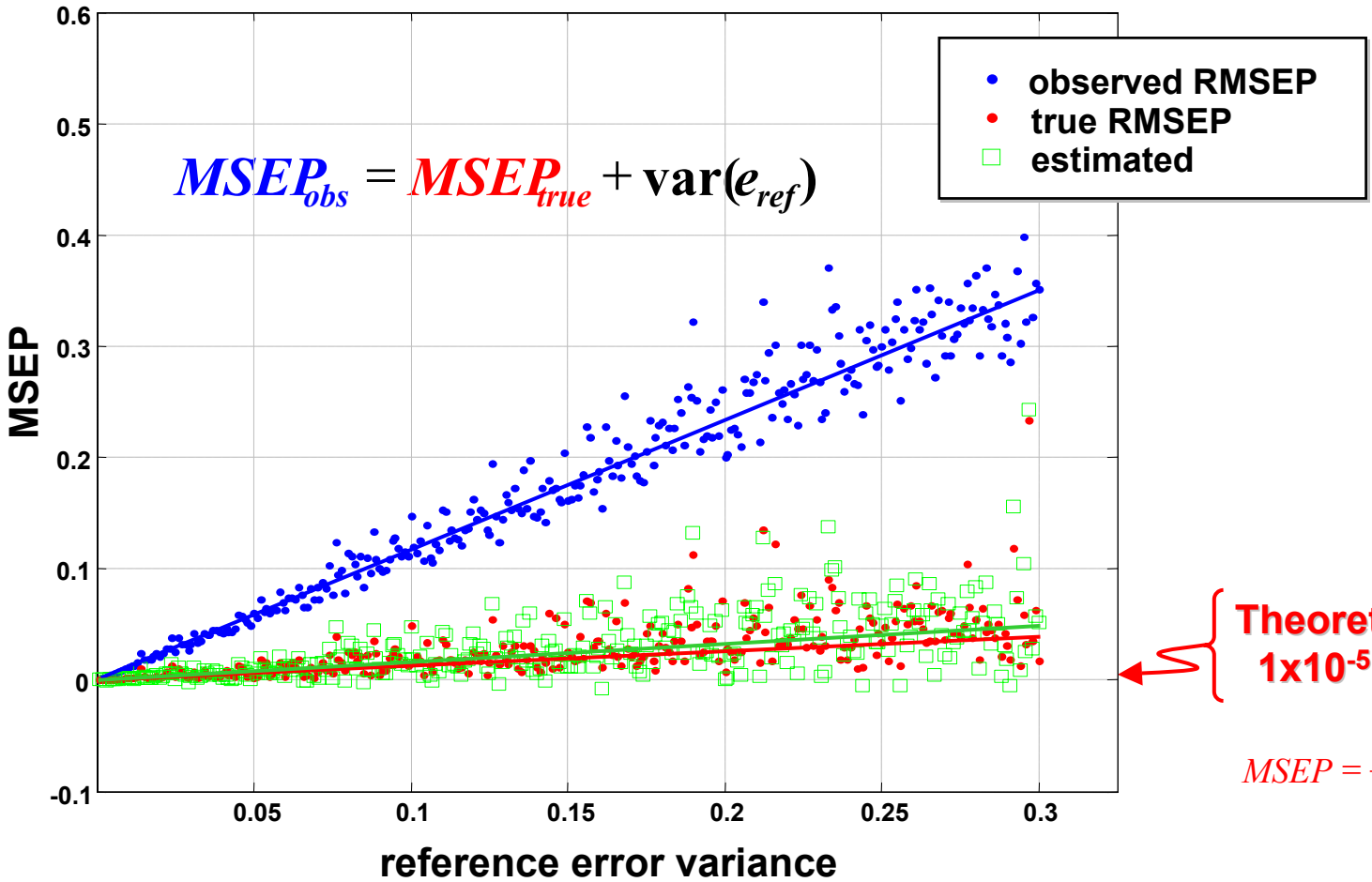
- **calibration reference error:** *inhibits* the ability of the modeling technique to resolve the *net analyte signal*

- **validation reference error:** apparent inflated *RMSEP* due to simple propagation of error

$$MSEP_{obs} = MSEP_{true} + \mathrm{var}(e_{ref}) + 2\,\mathrm{cov}(e_{true}, e_{ref})$$

Reference Error: *Calibration and Validation*

{ **Calibration Data:** 25 samples, reference range 0-1 }



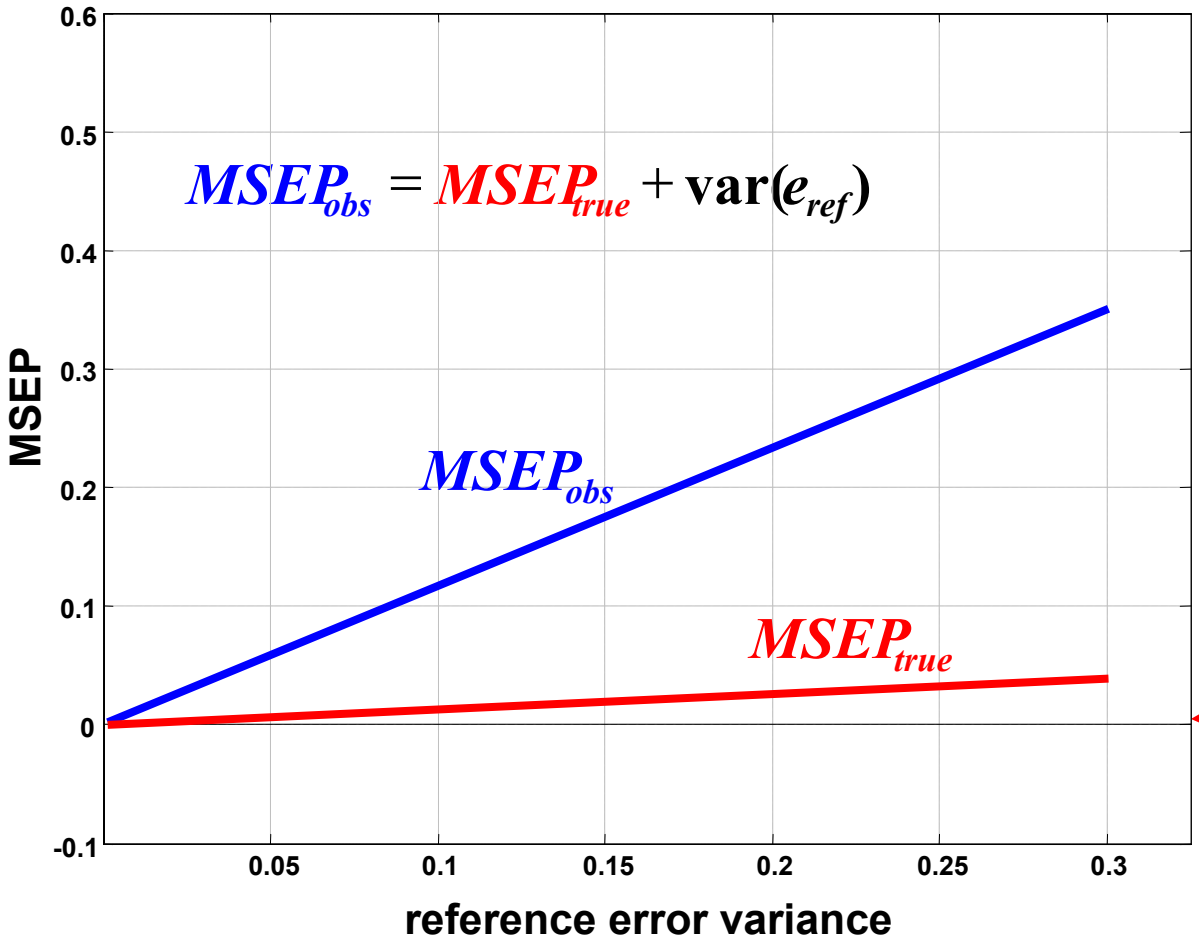
Theoretical Limit:
1×10^{-5} M

$$MSEP = \frac{1}{\|SEN\|^2} \cdot \sigma_x^2$$

Faber and Kowalski *Appl. Spec.*, **51**(5) 660 (1997)
DiFoggio, *Appl. Spec.* **49**(1), 67 (1995)

Reference Error: *Calibration and Validation*

{ **Calibration Data:** 25 samples, reference range 0-1 }



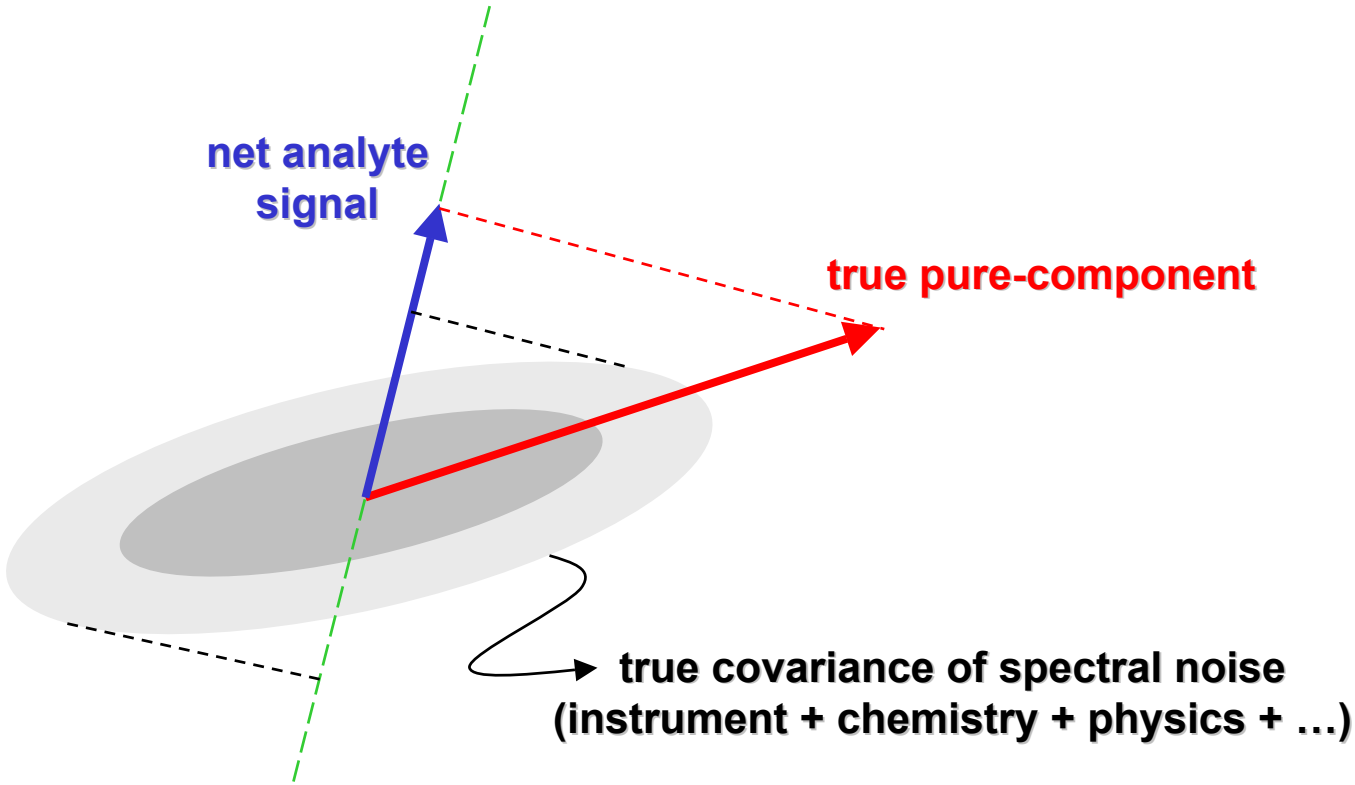
Theoretical Limit:
1×10^{-5} M

$$MSEP = \frac{1}{\|SEN\|^2} \cdot \sigma_x^2$$

Faber and Kowalski *Appl. Spec.*, **51**(5) 660 (1997)
DiFoggio, *Appl. Spec.* **49**(1), 67 (1995)

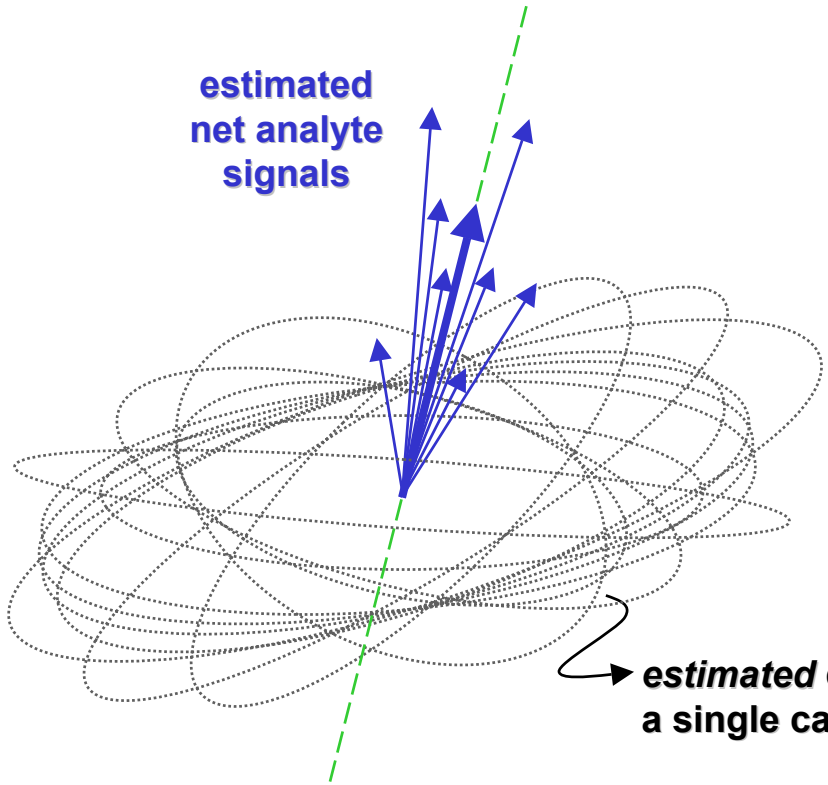
Model Error

- **Net Analyte Signal:** fraction of pure-component spectrum of the target which yields maximum multivariate SNR



Multivariate Signal and Observed Noise

- **Model Noise:** erroneous estimates of the NAS

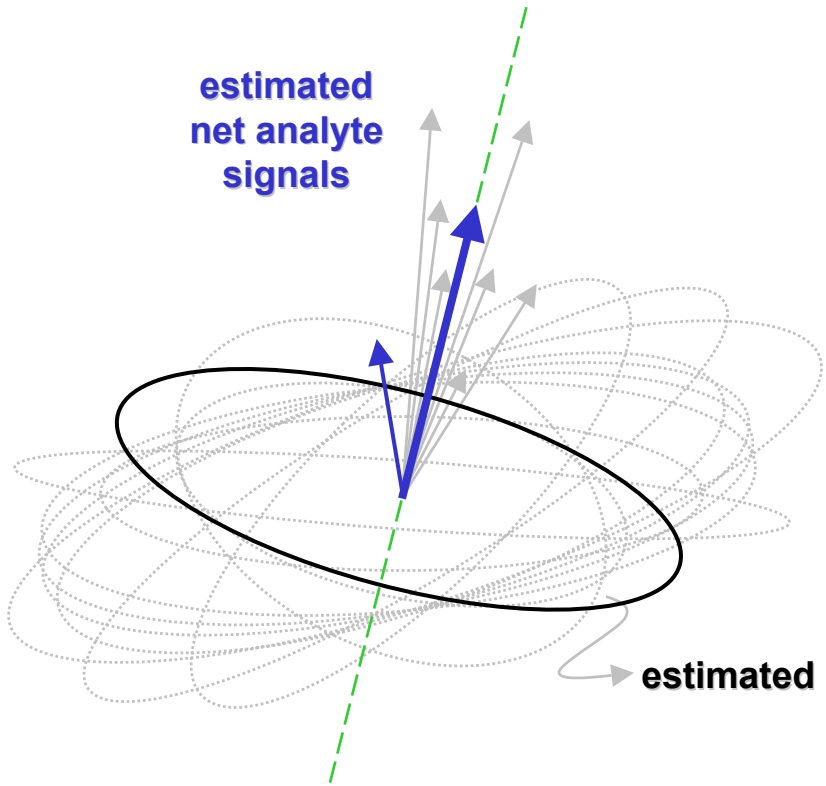


Usual Suspects:

- spectral noise sources
- experimental design
- sampling error (statistical sampling)
- instrumentation
- outliers

Multivariate Signal and Observed Noise

- **Model Noise:** erroneous estimates of the NAS



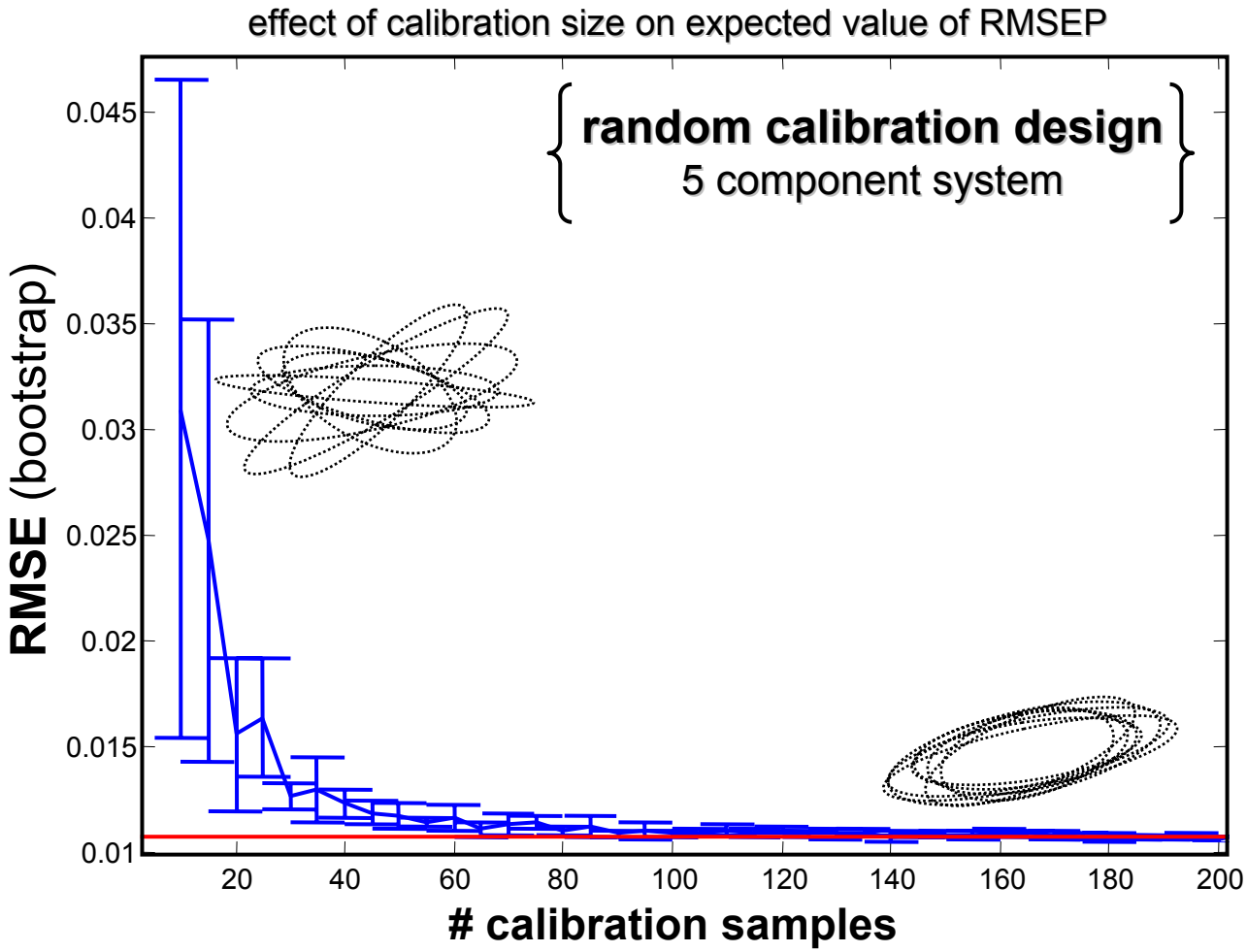
Usual Suspects:

- spectral noise sources
- experimental design
- sampling error (statistical sampling)
- instrumentation
- outliers

How much of the RMSE is due to crummy fitting?

Multivariate Signal and Model Noise

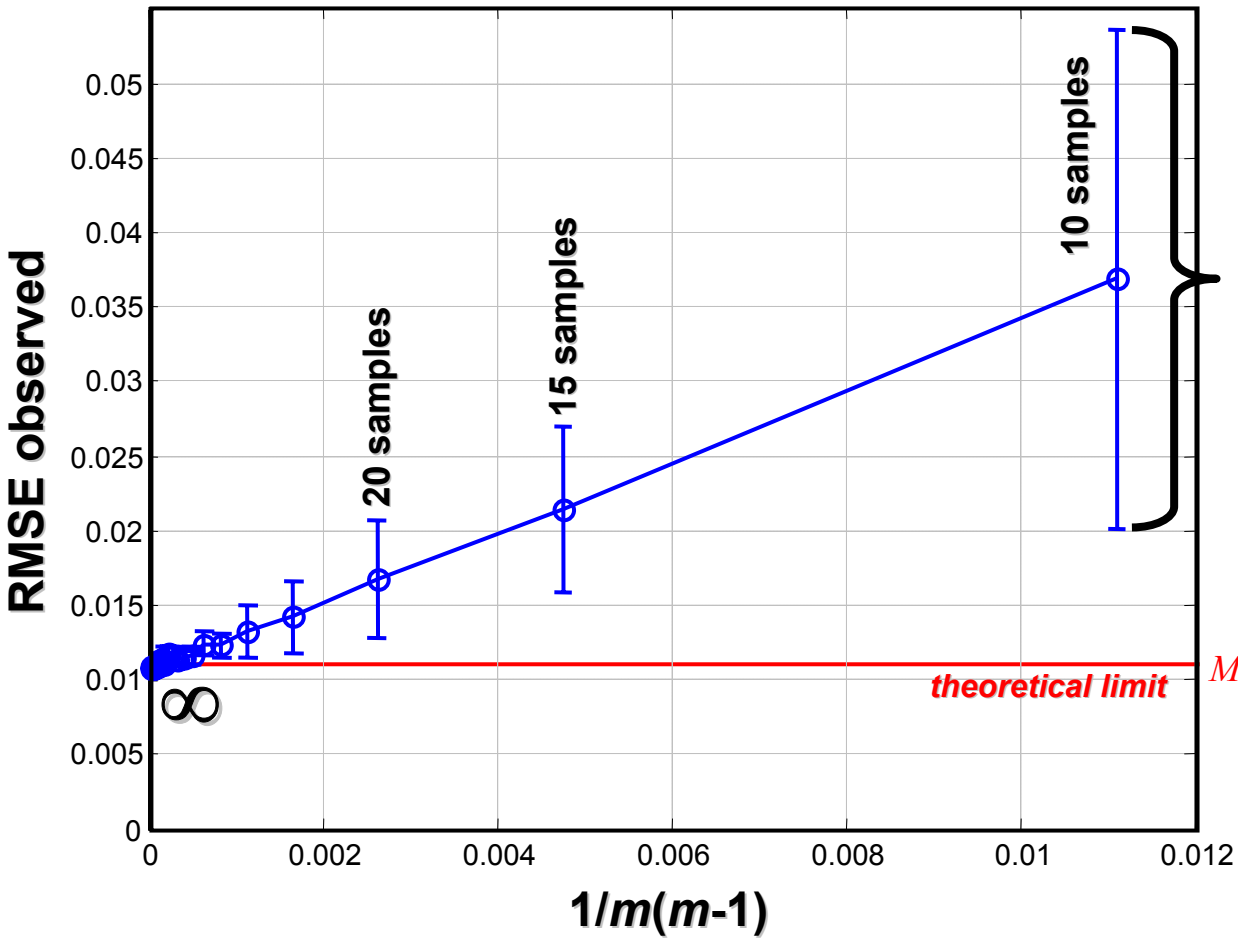
- **Model Noise:** erroneous estimates of the NAS



theoretical limit

$$MSEP = \frac{1}{\|SEN\|^2} \cdot \sigma_x^2$$

Multivariate Signal and Model Noise



Estimates of RMSE under resampling schemes

$$MSEP = \frac{1}{\|SEN\|^2} \cdot \sigma_x^2$$

{m = # calibration samples}

The Impact of Experimental Design

	Spectral Domain	Concentration Domain
chemometrics:	*net analyte signal*	*net analyte concentration*
statistics:	*regression vector*	*'D-optimality'*

N.A.Signal

fundamentally, how much of the analyte spectrum is available for quantitative inference?

Selectivity of Spectral Signal

N.A.Concentration

by the design of your calibration experiment, what portion of your effort has been wasted?

Selectivity of the Exp. Design

A. Lorber, *J. Chemom.*, **2**, 67 (1988)

The Impact of Experimental Design

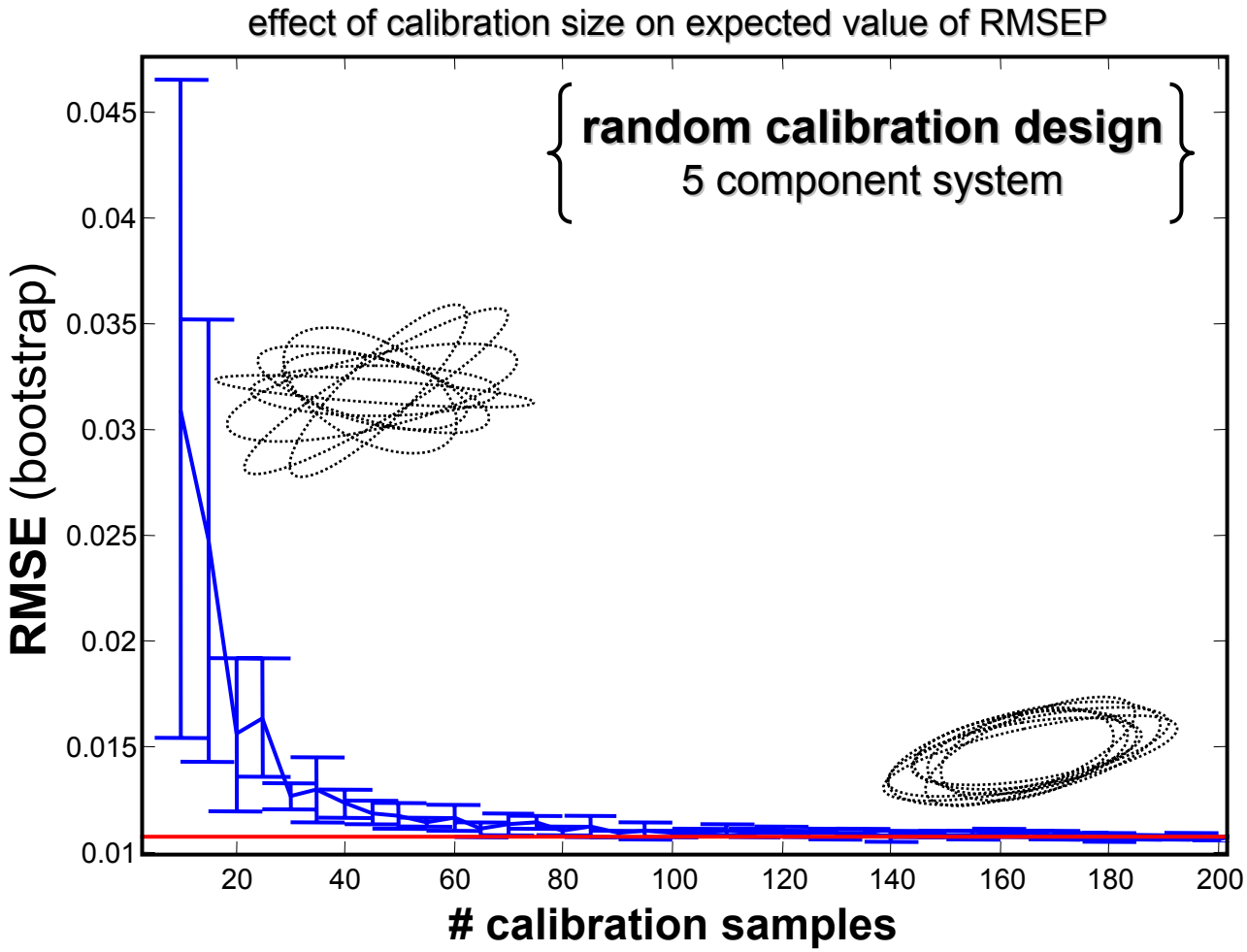
validation **calibration**

(0-1) (0-1) (0-1)

$$SEL_{system} \Leftarrow SEL_{NAS} \cdot SEL_{NAC}$$

Multivariate Signal and Model Noise

- **Model Noise:** erroneous estimates of the NAS



theoretical limit

$$MSEP = \frac{1}{\|SEN\|^2} \cdot \sigma_x^2$$

The Impact of Experimental Design

validation **calibration**

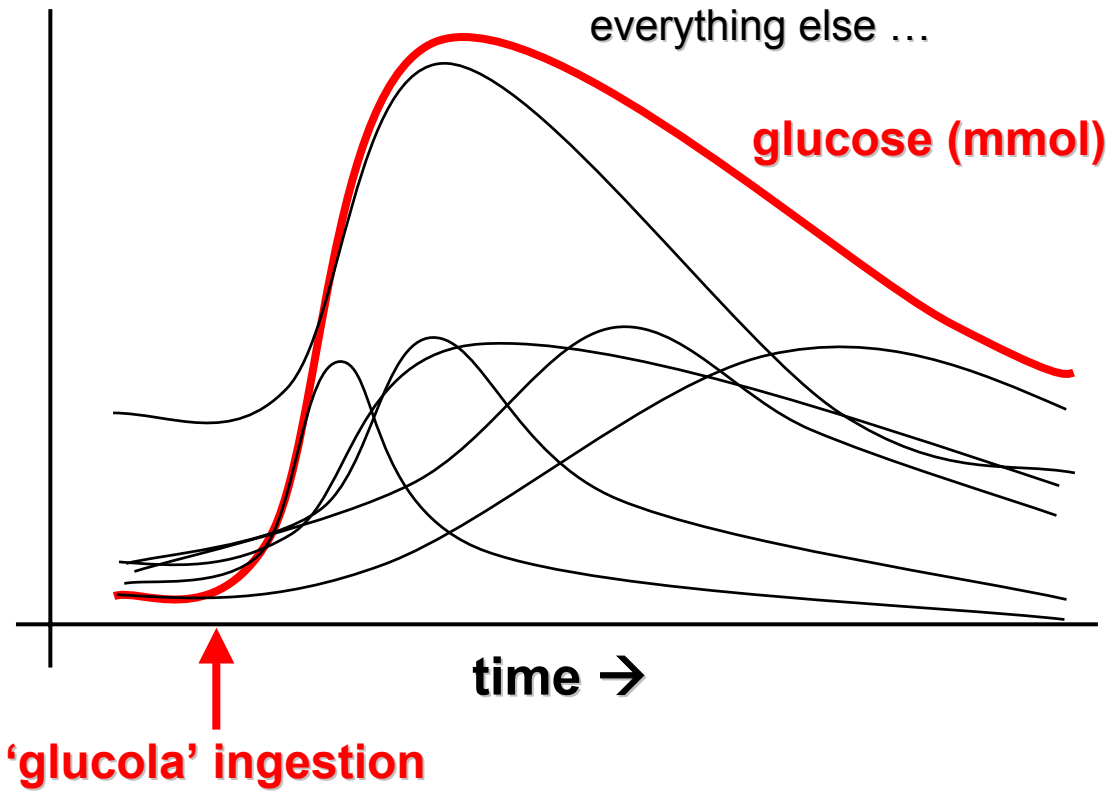
(0-1) (0-1) (0-1)

$$SEL_{system} \Leftarrow SEL_{NAS} \cdot SEL_{NAC}$$

- many of the 'spectral preprocessing methods' actually compensate for unfortunate experimental design (low SEL_{NAC}), and not low SEL_{NAS}
  - smoothing, derivatives, *etc.*
  - wavelength selection
  - PLS/PCA/*etc.*
  - …

Is your signal *really* your signal?

- **The 'OGTT calibration'** (oral glucose tolerance test)



"shock response"

- wild hydration changes
- perfusion/vascularization
- temperature
- scattering
- insulin, glucagon, cortisol, epinephrine …

The body has been terrorized, and responds accordingly.

Is your signal *really* your signal?

Calibration Process Says:

"the net analyte signal is that multivariate response maximally correlated to the reference values"

Is your signal *really* your signal?

Calibration Process Says:

"the net analyte signal is that multivariate response maximally correlated to the reference values"

Analyst Says:

Is my model *specific* for the target analyte?

- does it respond *only* to changes in analyte concentration?
- if not, what perturbs it, and by how much?

Verification and Validation

- Verification: Have you generated the **system right**?
  SPECIFICITY QUESTIONS:

- Validation: Have you generated the **right system**?
  SENSITIVITY QUESTIONS:

Verification and Validation

- **Verification: Have you generated the system right?**
  SPECIFICITY QUESTIONS:
  - How specific is the system for my target analyte?

- **Validation: Have you generated the right system?**
  SENSITIVITY QUESTIONS:

Verification Example: *predicting phantoms*

{ *in-vivo* model predicting glucose in phantom solutions }

plot *in-vivo* model predicting phantom solutions with no correlates

{Scatterplot removed at the request of the legal department prior to distribution of copies of the slides via mail-out.}

Multi-Subject Calibration Model

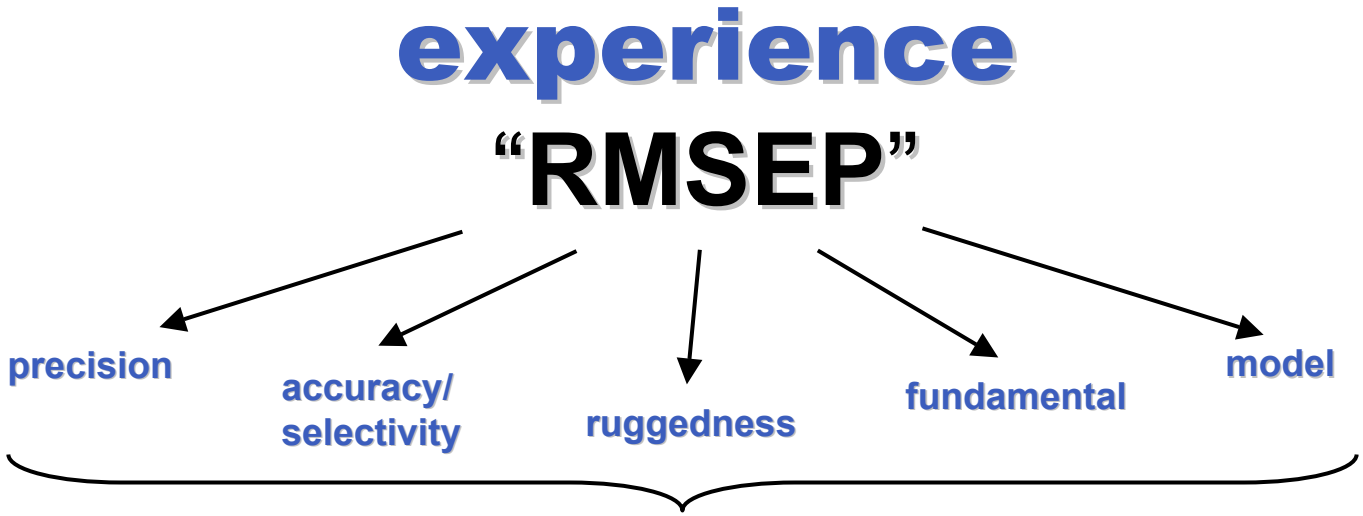
Phantom (Verification) System:
(fract. fact. design with replication)

- glucose
- urea
- creatinine
- water
- scattering beads

Verification and Validation

- **Verification: Have you generated the system right?**
  SPECIFICITY QUESTIONS:
  - How specific is the system for my target analyte?
  - Does the instrument mSNR stack up against target?
  - Does the chemical mSNR stack up?
  - Does the sampling mSNR stack up?
  - Does the outlier/inlier coverage respond as it has been designed?
- **Validation: Have you generated the right system?**
  SENSITIVITY QUESTIONS:

Verification: *Variance Components Analysis*



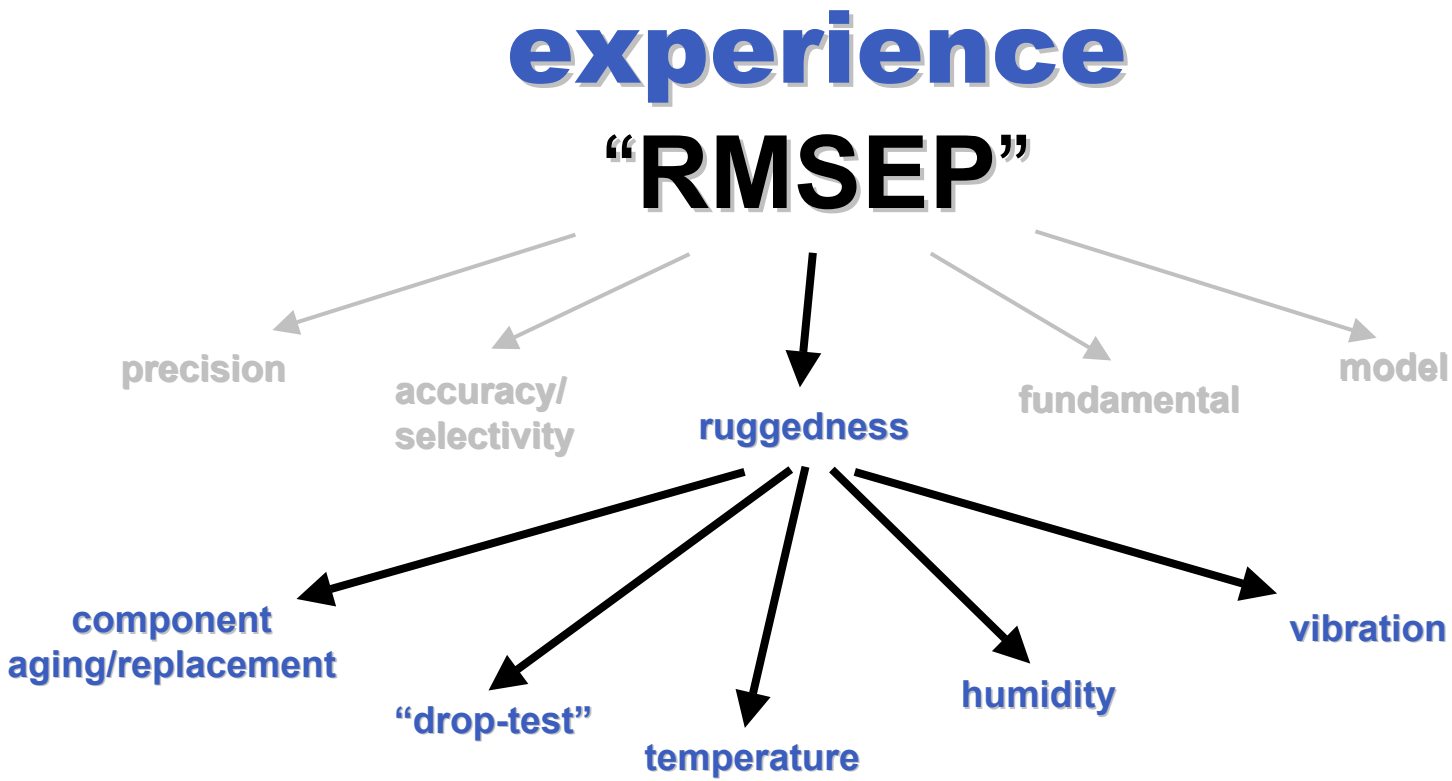
How do each of these terms contribute to that grand "RMSEP"?

Variance Components Model for total *MSEP*:

$$\left\langle \left(y-\hat{y}\right)^2 \right\rangle = \sigma^2_{prec} + \sigma^2_{acc} + \sigma^2_{rugged} + \sigma^2_{fund} + \sigma^2_{ref} + \ldots$$

Searle, Casella, and McCulloch *"Variance Components"* J. Wiley & Sons: New York (1991)

Verification: *Variance Components Analysis*



Variance Components Model for '*ruggedness*':

$$\sigma^2_{rugged} = \sigma^2_{age} + \sigma^2_{drop} + +\sigma^2_{temp} + \sigma^2_{RH} + \ldots$$

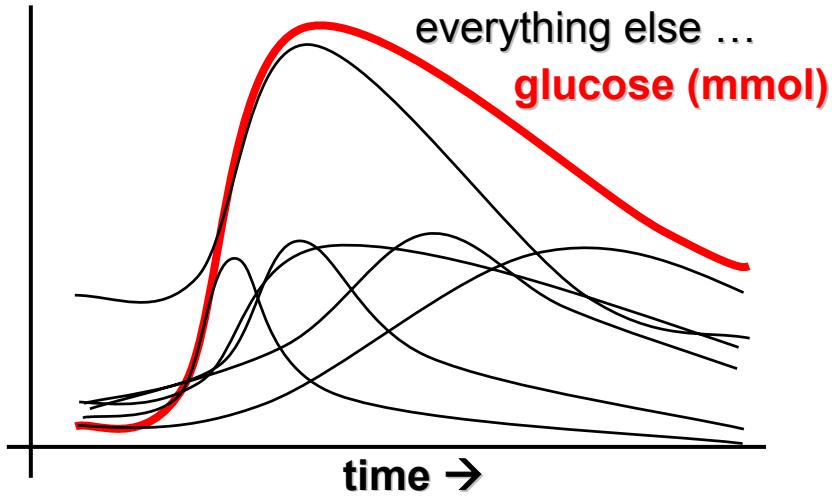
Verification and Validation

- **Verification: Have you generated the system right?**
  SPECIFICITY QUESTIONS:
  - Is the system responding to my target analyte, or something else?
  - Does the instrument mSNR stack up against target?
  - Does the chemical mSNR stack up?
  - Does the sampling mSNR stack up?
  - Does the outlier/inlier coverage respond as it has been designed?
- **Validation: Have you generated the right system?**
  SENSITIVITY QUESTIONS:
  - Do you meet specifications for performance & operational?

We find *verification* both more informative, and decisive than *validation* in complex systems.

Validation: *the final hurdle?*

- **often not as informative as verification**
  - *calibrate on an OGTT*
  - *validate on an OGTT*
  - *?*



- **with highly complex systems, the experimental design for an all-encompassing validation can rapidly become unwieldy**
- **useful for spotting 'unexpected events', and continued variance components analysis**

Glucose Validation Study (summer 1999)§

{ 40 validation subjects, 7 week duration }

plot of estimated *in-vivo* glucose vs. YSI reference

{Scatterplot removed at the request of the legal department prior to distribution of copies of the slides via mail-out.}

Summary

- **Reference Error**
  - rarely a problem for calibration (save your $$); models are incredibly robust
  - very manageable when validating/verifying the performance
- **Model Error**
  - quite dangerous in complex systems
    - use sound experimental design when possible
    - unfortunately, can't design in many spectral noise sources
    - quantify the model error – what is to be gained by collecting more samples, or improving the design by other means?
- **Verification**
  - allows one to *challenge* the assumptions with *specific* tests
    - sensitivity – how responsive is the system to changes in analyte concentration?
    - specificity – how responsive is the system to changes in noise sources?
  - free from many pitfalls of calibration/validation
- **Validation**
  - useful, but realistically … not very informative

Acknowledgements

InLight Solutions: Modeling & Spectroscopy Group

Mark Rohrscheib, M.D.

Clinical Research Team