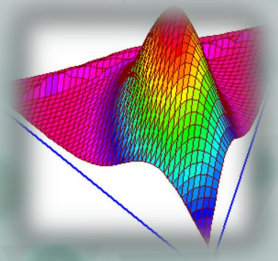
Workshop

Plans d'expériences



4-6 Octobre 2017

Hostellerie La Magnaneraie

VILLENEUVE LÈS AVIGNON

APEX IX

Statistiques appliquées aux plans d'expériences

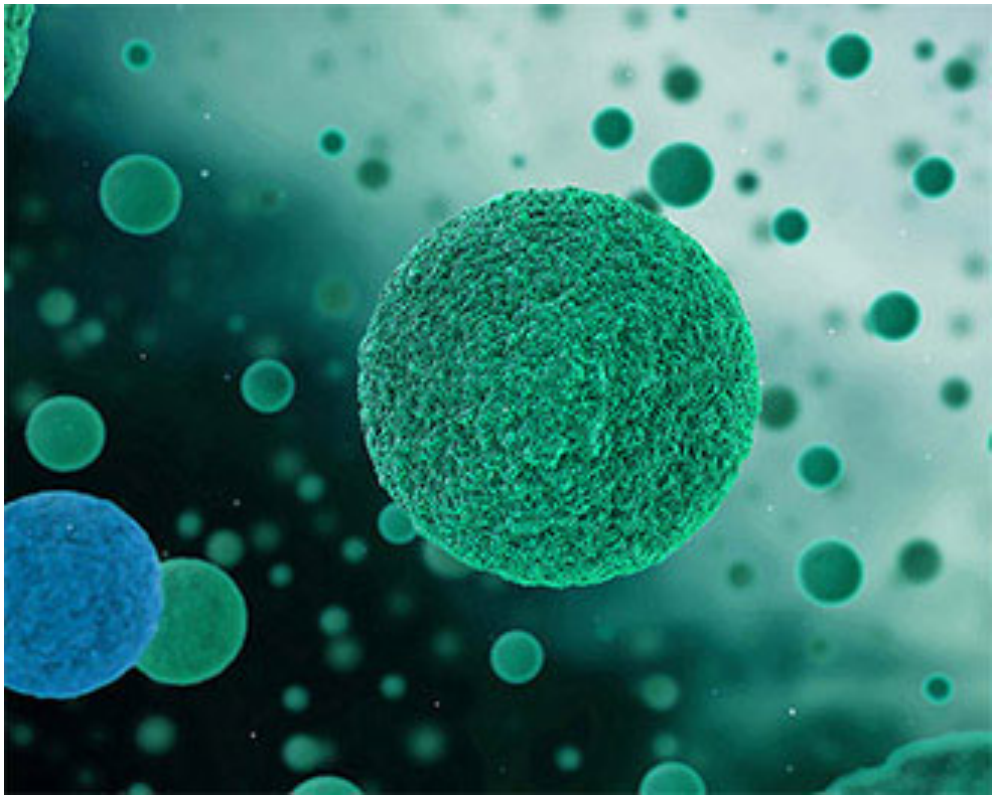
Pierre Lebrun, Pharmalex / Arlenda
pierre.lebrun@arlenda.com

PHARMALEX

Outline

Links with DoE

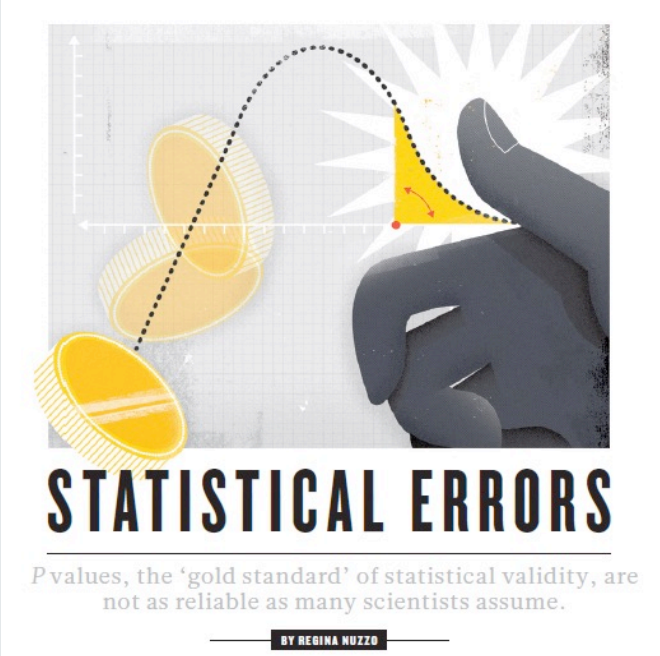
- The Good, the Bad and the Ugly
- A little help from my friends: Design Space
- The Multivariate Regression
- Example 1: a Spray-Dryer
- Example 2: a Pharmaceutical Formulation
- Conclusion



The Good, the Bad and the Ugly

P-value crisis. A small selection.

- Nature, 2014



STATISTICAL ERRORS

P values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume.

BY REGINA NUZZO

- ASA, March 2016

REPRODUCIBILITY

Statisticians issue warning on *P* values

Statement aims to halt missteps in the quest for certainty.

BY MONYA BAKER

Misuse of the *P* value — a common test for judging the strength of scientific evidence — is contributing to the number of research findings that cannot be reproduced, the American Statistical Association (ASA) warned on 8 March. The group has taken the unusual step of issuing principles to guide use of the *P* value, which it says cannot determine whether a hypothesis is true or whether results are important.

This is the first time that the 177-year-old ASA has made explicit recommendations on such a foundational matter, says executive director Ron Wasserstein. The society's members had become increasingly concerned that the *P* value was being misapplied, in ways that cast doubt on statistics generally, he adds.

cannot indicate the importance of a finding; for instance, a drug can have a statistically significant effect on patients' blood glucose levels without having a therapeutic effect.

Giovanni Parmigiani, a biostatistician at the Dana Farber Cancer Institute in Boston, Massachusetts, says that misunderstandings about what information a *P* value provides often crop up in textbooks and practice manuals. A course correction is long overdue, he adds. "Surely if this happened twenty years ago, biomedical research could be in a better place now."

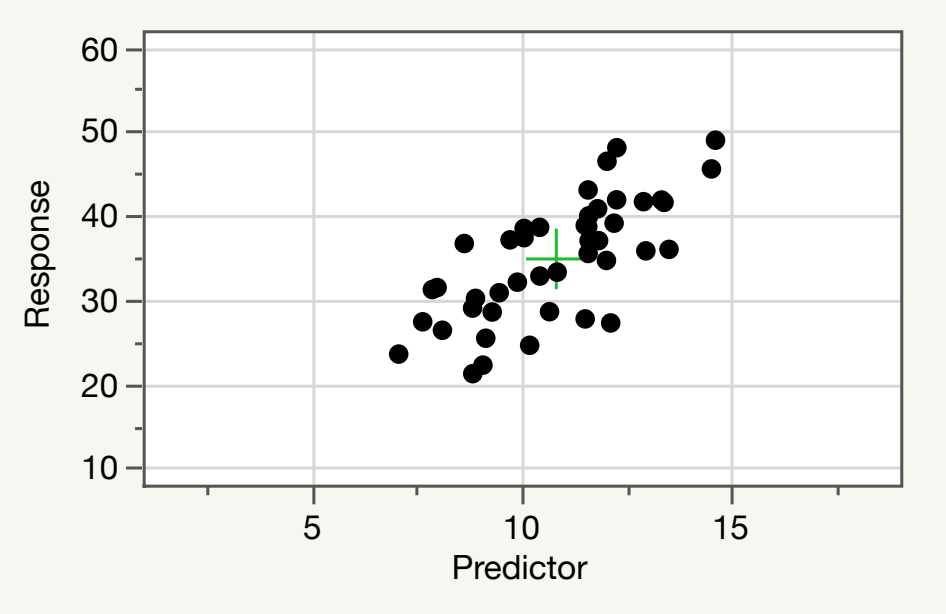
FRUSTRATION ABOUNDS

Criticism of the *P* value is nothing new. In 2011, researchers trying to raise awareness about false positives gamed an analysis to reach a statistically significant finding: that listening to music by the Beatles makes undergraduates younger

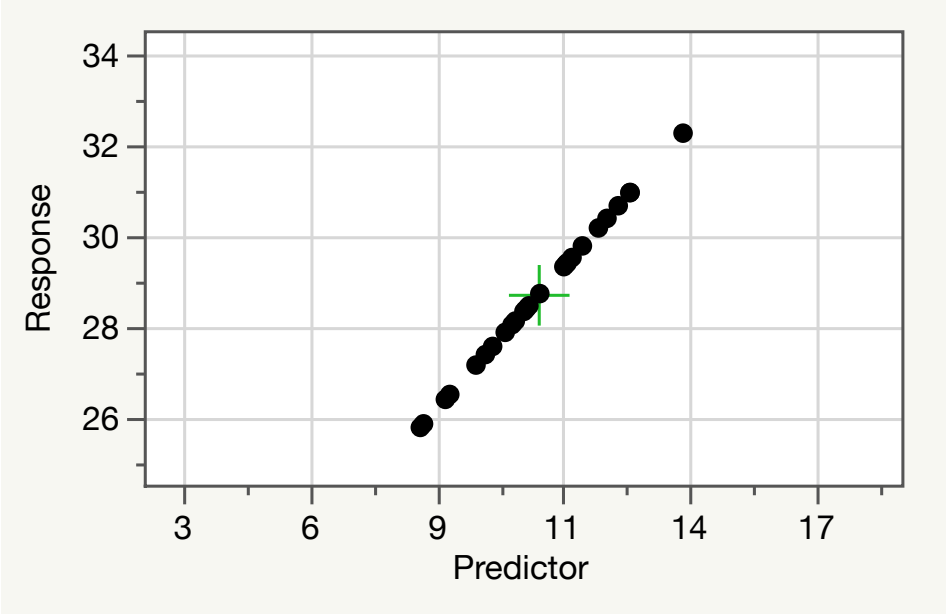
- "*The most important task before us in developing statistical science is to **demolish the P-value culture**, which has taken root to a frightening extent in many areas of both pure and applied science and technology.*" Nelder, J. A. **1999. Statistics for the millennium**. Statistician 48:257–269 (page 261)

e.g. p-values for factor screening

Slope p-value: 0.45



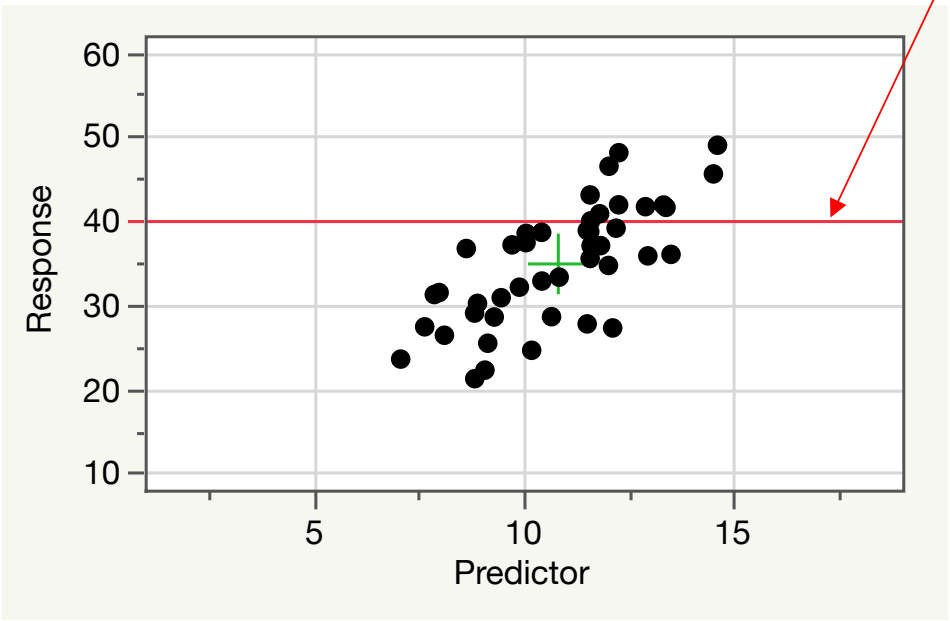
Slope p-value: <0.0001



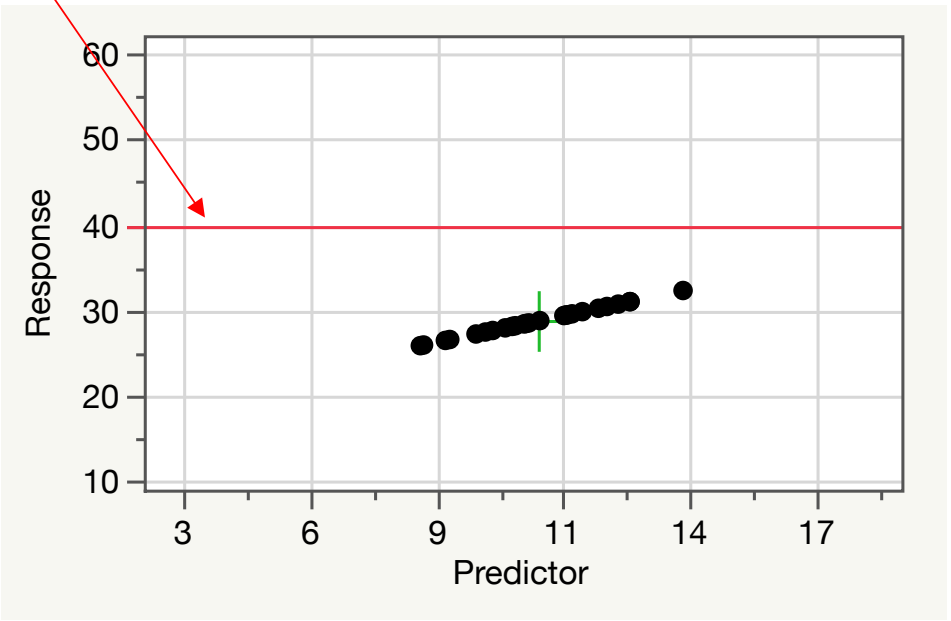
e.g. p-values for factor screening

Specification
(e.g. max tolerated)

Slope p-value: 0.45



Slope p-value: <0.0001

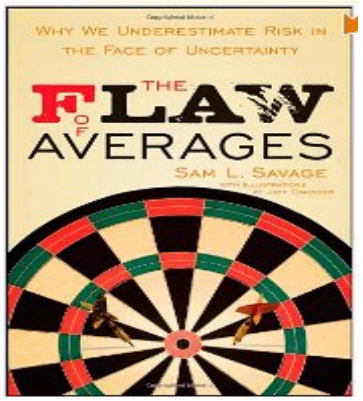


Significant (with respect to spec)

Not significant (with respect to spec)

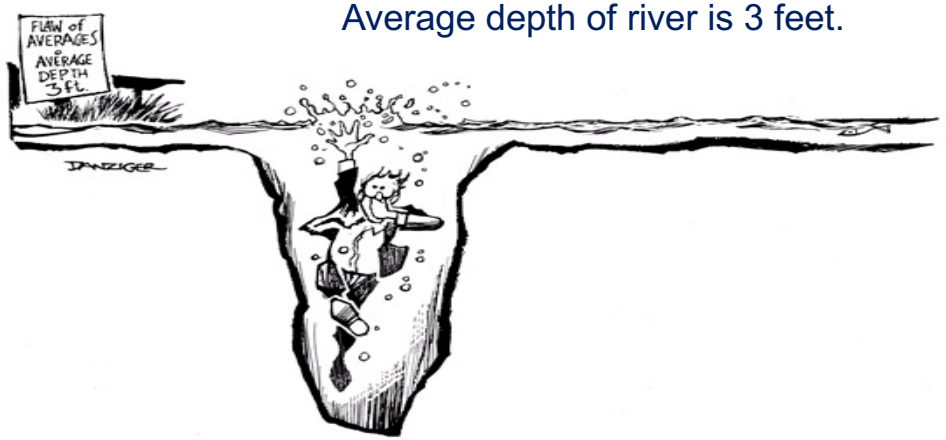
Still, many people use p-value or similar methodologies (e.g. stepwise)

Solution: Prediction of Individual Results



The Flaw of Averages:
Why We Underestimate Risk in the Face of Uncertainty
by Dr. Sam Savage

Average depth of river is 3 feet.



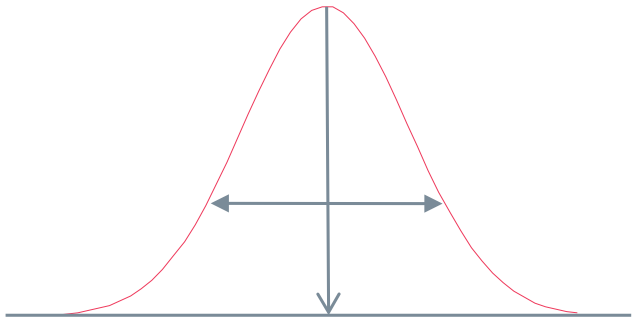
Or… with my process running at X1=30 and X2=10, will it give a product with an Attr1 (e.g. (yield) >95% and an Attr2 (e.g. moisture) <5%?
What are the guarantee this happens ?

From John Peterson, 2012

Implementation: a Bayesian approach, for we want to predict !

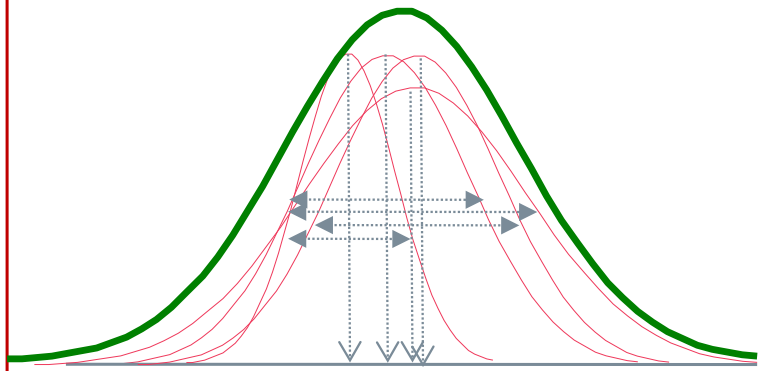
Monte-Carlo Simulations

where the “new observations” are drawn from distribution “centered” on estimated location and dispersion parameters (treated *wrongly* as “true values”). Some use CI limits instead.



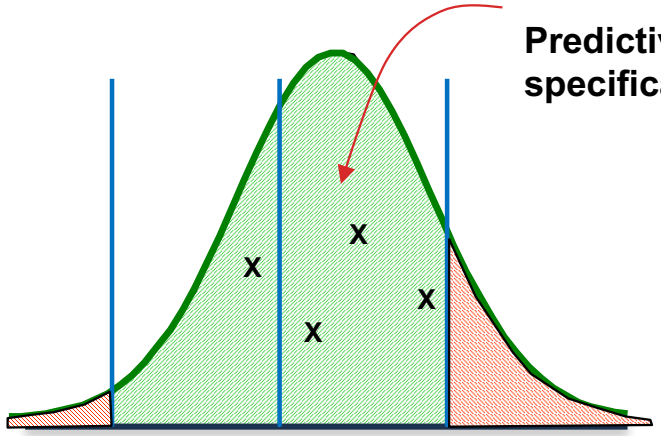
Predictions

Account for uncertainty in mean and in… variability estimates !



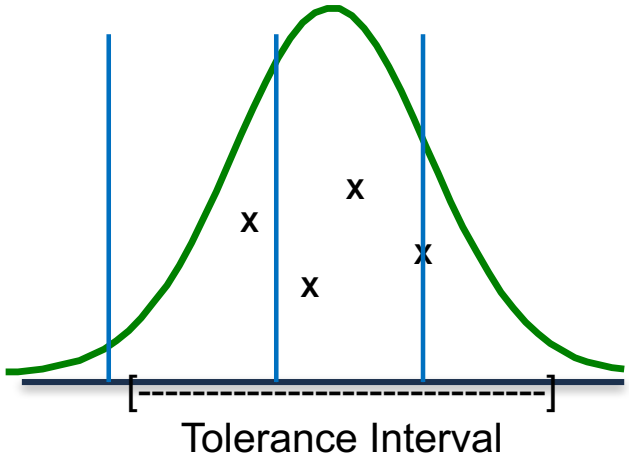
Probability being in specifications vs. Tolerance intervals

- Use the Predictive distribution to compute the probability to be within specifications.



- Bayesian statistics allows computing a probability instead of a Tolerance Interval only.

What's the risk ?

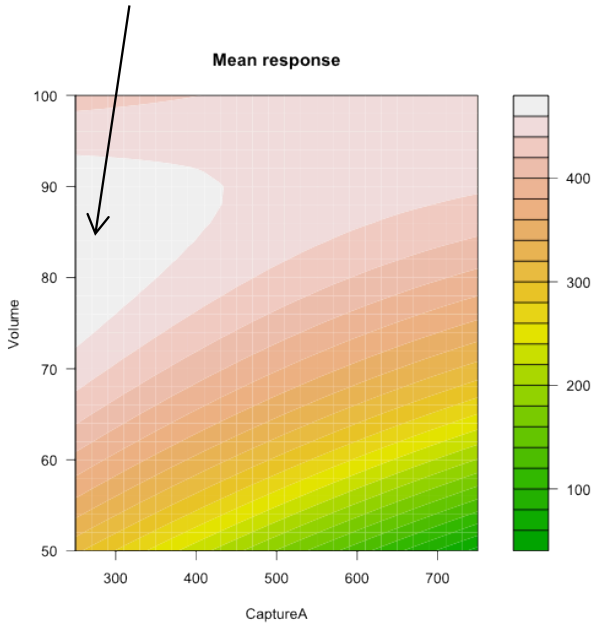


Last but not least…

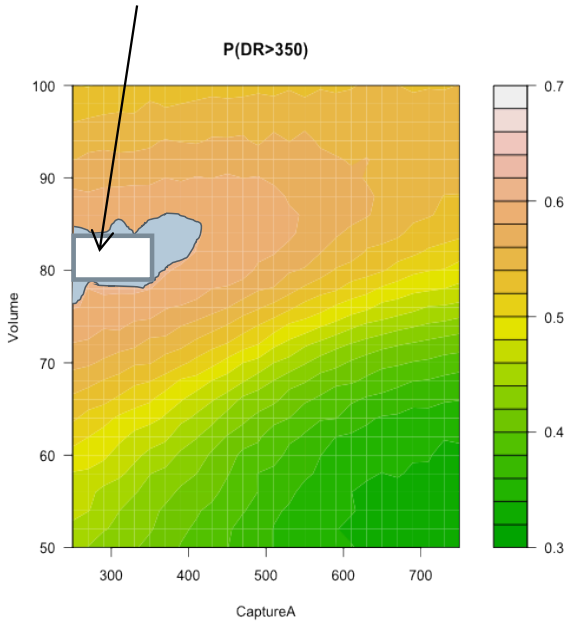
Take into account the uncertainty about future run for defining a region of acceptable process (Design Space) for the parameters.

Think risk, instead of mean.

Mean based

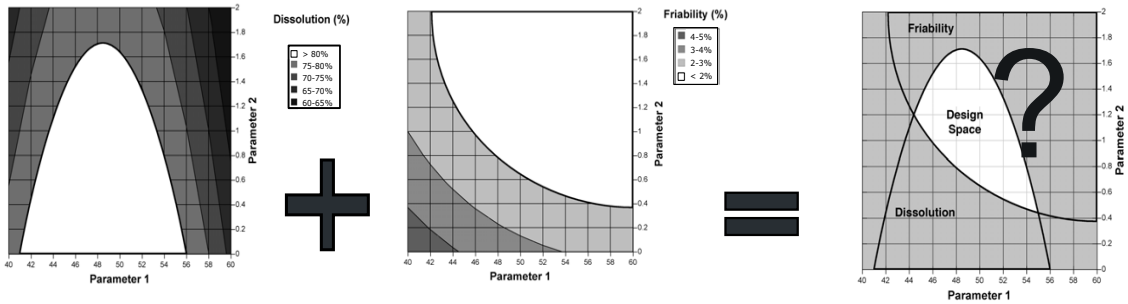


Risk based.



At the edge of the white area, ~50% chance not to achieve claimed quality !

Still ! Mean responses dangers !



(ICH Q8 appendix)

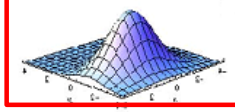
→Generally, mean responses are used for optimization

✗ do not provide any clue about process reliability / capability. You won't observe the mean !!

✗ fail to give any information on how the process will perform in the future

✗ will certainly give disappointing and unexplained results for the future use of the method

✗ The same disappointment with DoE ?



"Distribution of predicted responses" = "stochastic process" + "measurement error"

(courtesy of J.J. Peterson)

Desirability concern

- Most of multi-criteria decision do not tolerate trade-off
  - Quality attributes (responses) must in general achieve pre-defined specification
    - If the product attribute(s) is below specification, it is sub-standard and thrown away…
  - Desirability: *what is the optimal condition that make most of my attributes desirable?*
    - But… Will it work tomorrow ?
  - Probability: *what is the (joint) probability that (all) my attributes will meet those specifications?*
    - *The joint probability measure is actually a good global desirability index*
      - *P(success) = 1 is the most desirable situation and P(success) = 0 the less enviable one*

A little help from my friends: Design Space

Facts

THE WALL STREET JOURNAL.

WEDNESDAY, SEPTEMBER 3, 2003 · VOL. CCXLII NO. 45 · ★★★★ $1.00

Factory Shift

New Prescription For Drug Makers: Update the Plants

After Years of Neglect, Industry Focuses on Manufacturing; FDA Acts as a Catalyst

The Three-Story Blender

By Leila Abboud
And Scott Hensley

BusinessWeek

MAY 3, 2004

SPECIAL REPORT -- QUALITY MANUFACTURING

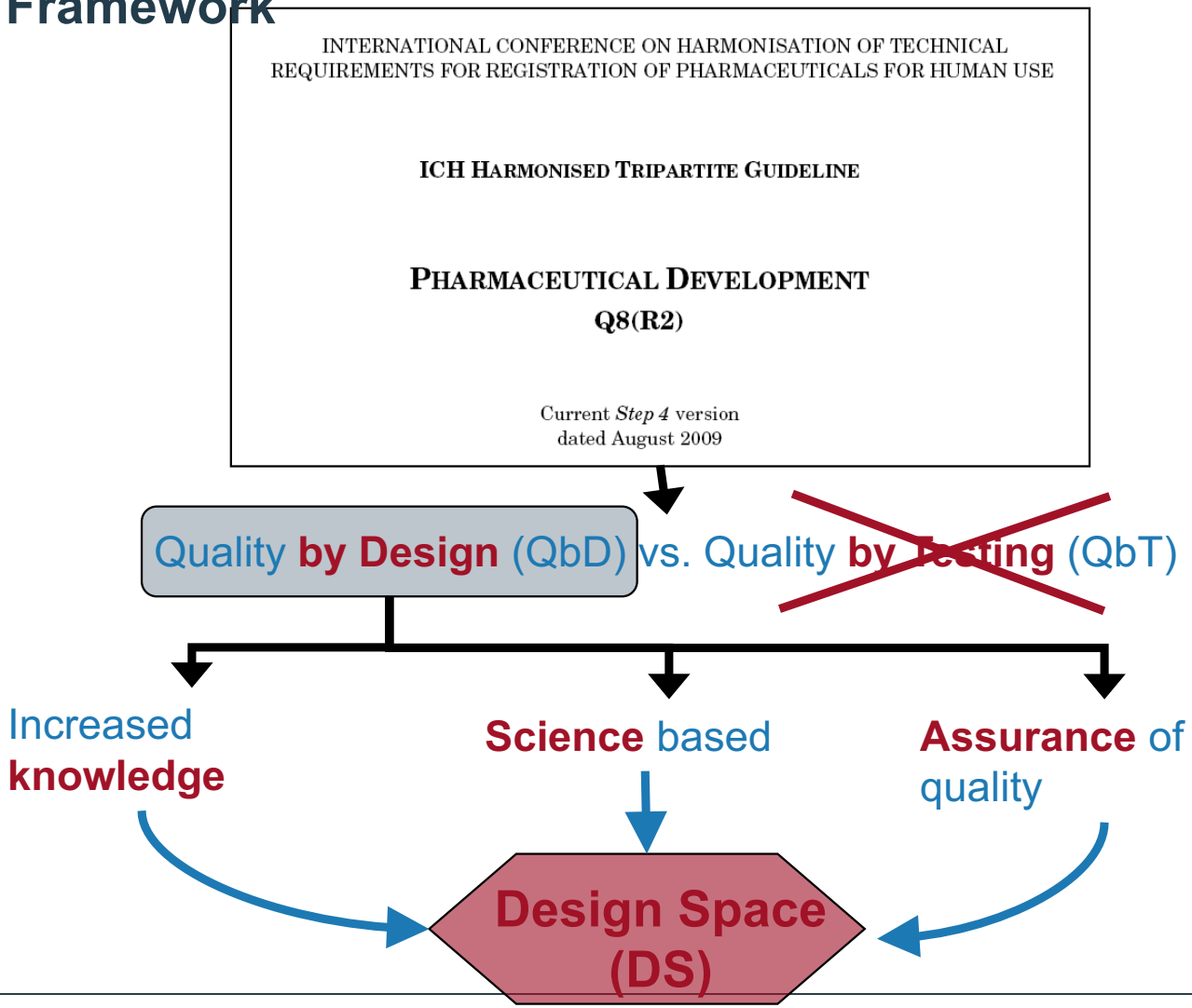
Making Pills The Smart Way

Drugmakers are revamping factories to save money and avoid production mishaps

Desired state

- Product quality and performance achieved and **assured by design** of effective and efficient manufacturing processes
- Product specifications based on mechanistic **understanding** of how formulation and process factors impact product performance
- Ability for **continuous improvement** and **assurance of quality**

Regulatory Framework



Regulatory Framework

- ICH Q8: Design Space (DS):

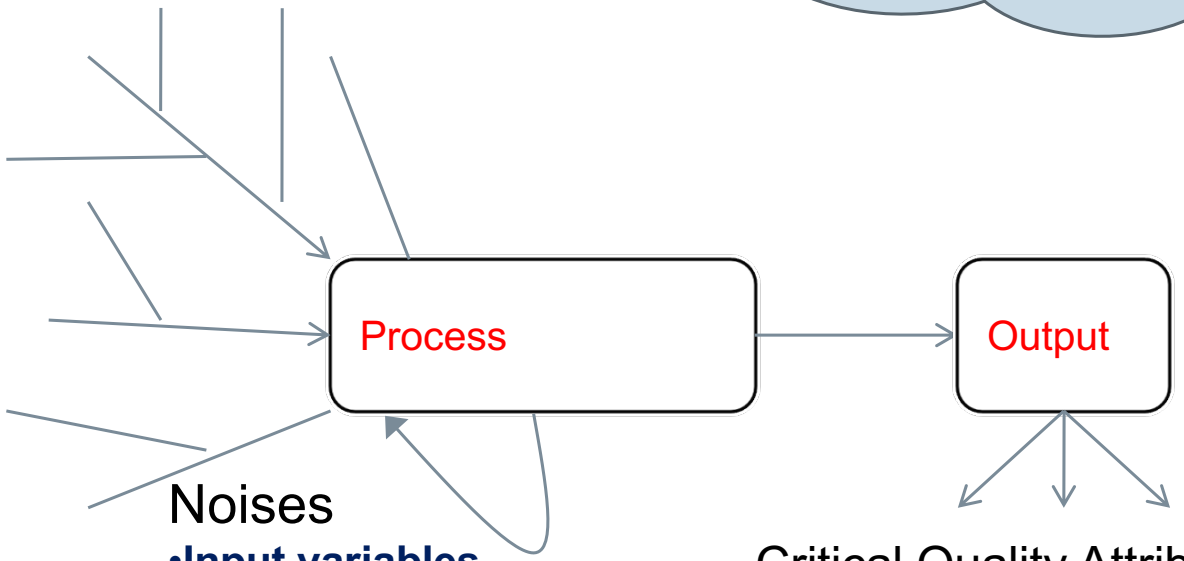
- "*the multidimensional combination and interaction of input variables and process parameters that have been demonstrated to provide assurance of quality* »
- "*working within the DS is not considered as a change* »
- "*Understand and gain knowledge about a process to find a parametric region of reliable robustness for future performance of this process*"

Process characterisation

Running two times the process with X unchanged will not provide two times the same output

Critical Process Parameters (X) :

- **Quantitative**
- **Qualitative**
- **In-process**



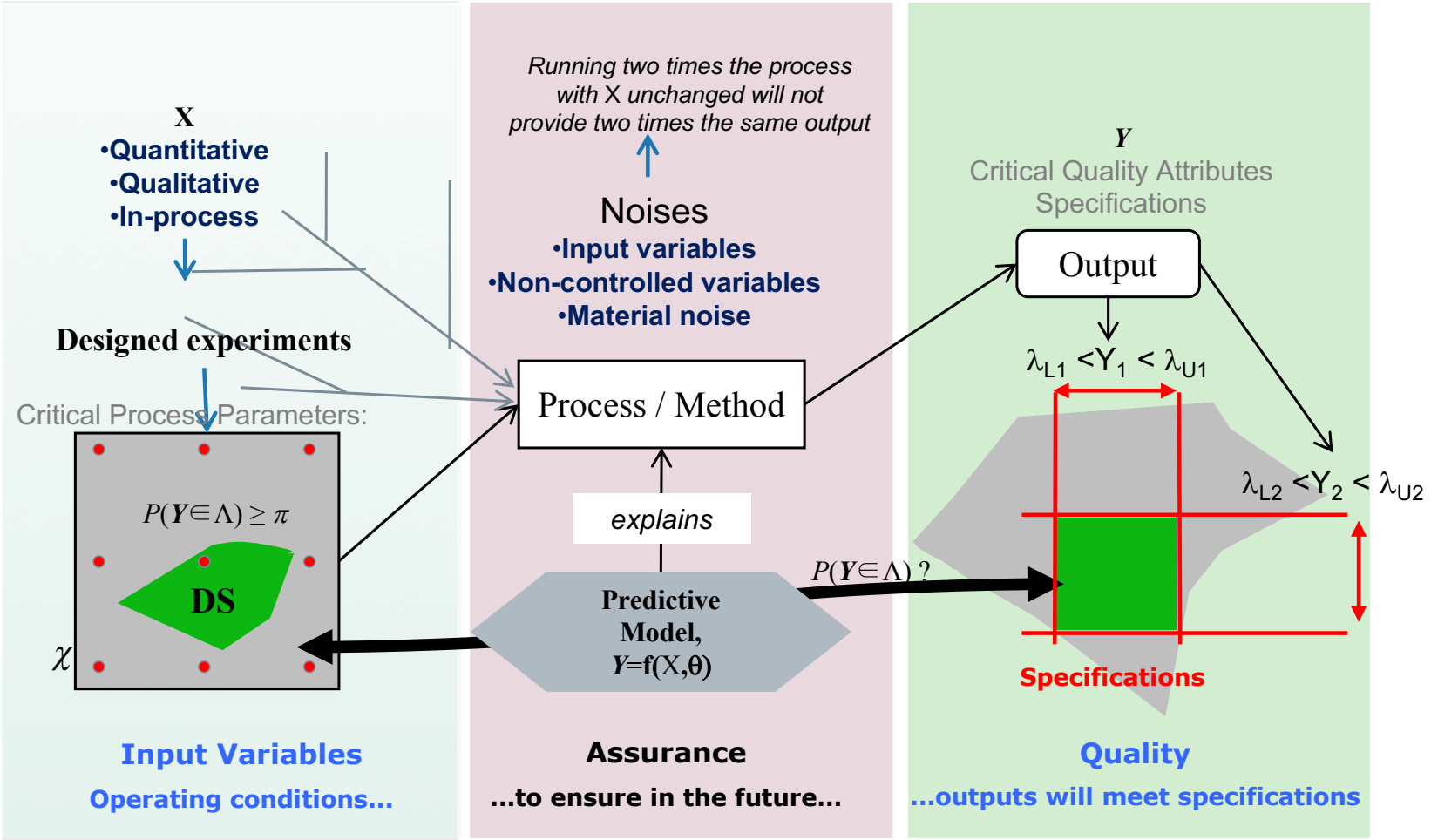
Noises

- **Input variables**
- **Non-controlled variables**
- **Material noise**

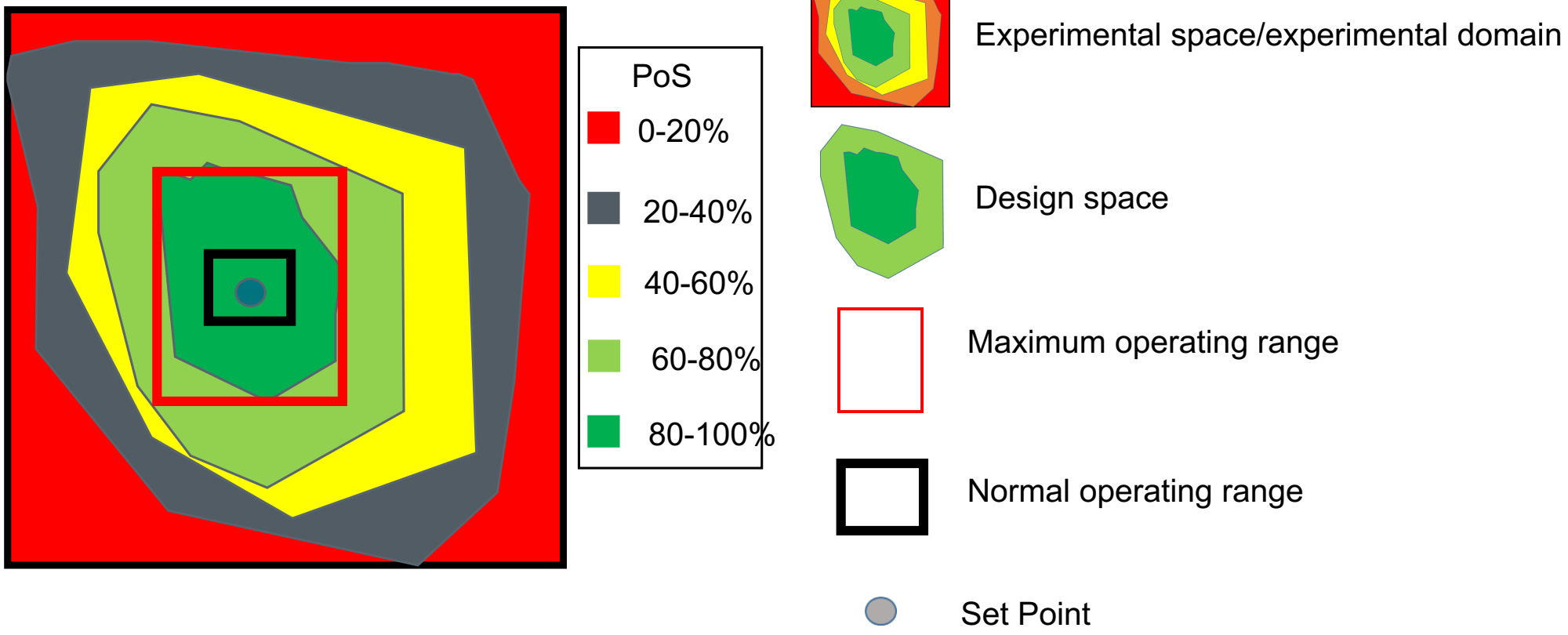
Critical Quality Attributes (Y)

=> specifications

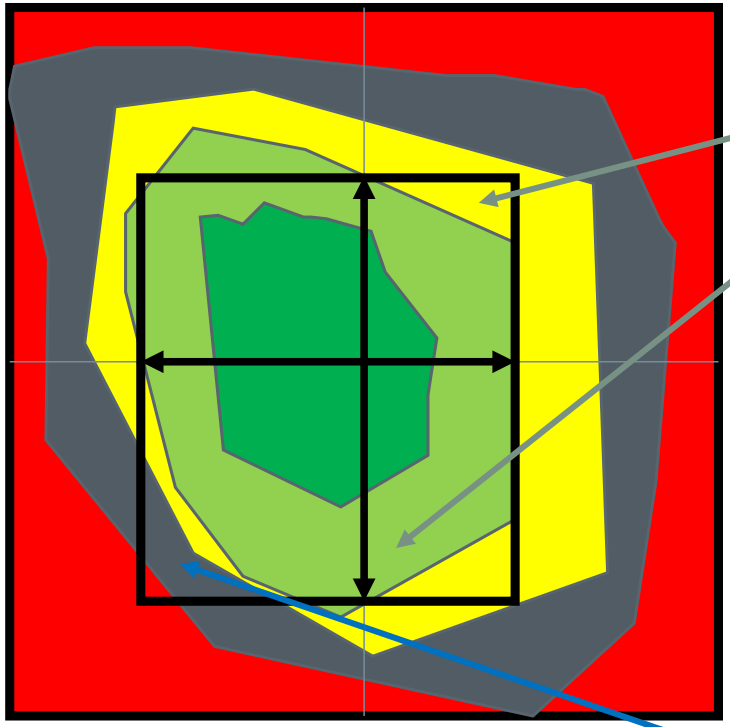
Design Space



Terminology from ICH Q8



The Ugly (again): PAR: univariate ranges



- Risk: over-optimistic process operating ranges
- Using PAR as MOR is not a good idea

Bayesian manifesto

- Why a Bayesian approach ?
  - Because we want to predict (outcome of the process steps)
  - Because we want to make probabilistic statements of an outcome
    - -> P(success) or P(OOS)
  - Because we may (sometimes) have prior knowledge
  - Because, thanks to MCMC simulations, we can handle simple to very complex models in a unified framework (yes, speed of implementation matters more than running speed of the samplers)
    - *In general, models are pretty simple. e.g. two-way random ANOVA models… but with unbalanced data, prediction as a frequentist is already not a good option…*
  - Because, thanks to Monte-Carlo methods, I can pool and propagate all uncertainties from the beginning to the end of the process
    - *Why focus on maximum likelihood when we can play with all the posterior distribution ?*
  - Because we want to predict

The Multivariate Regression

Desirable Modelling Method Criteria

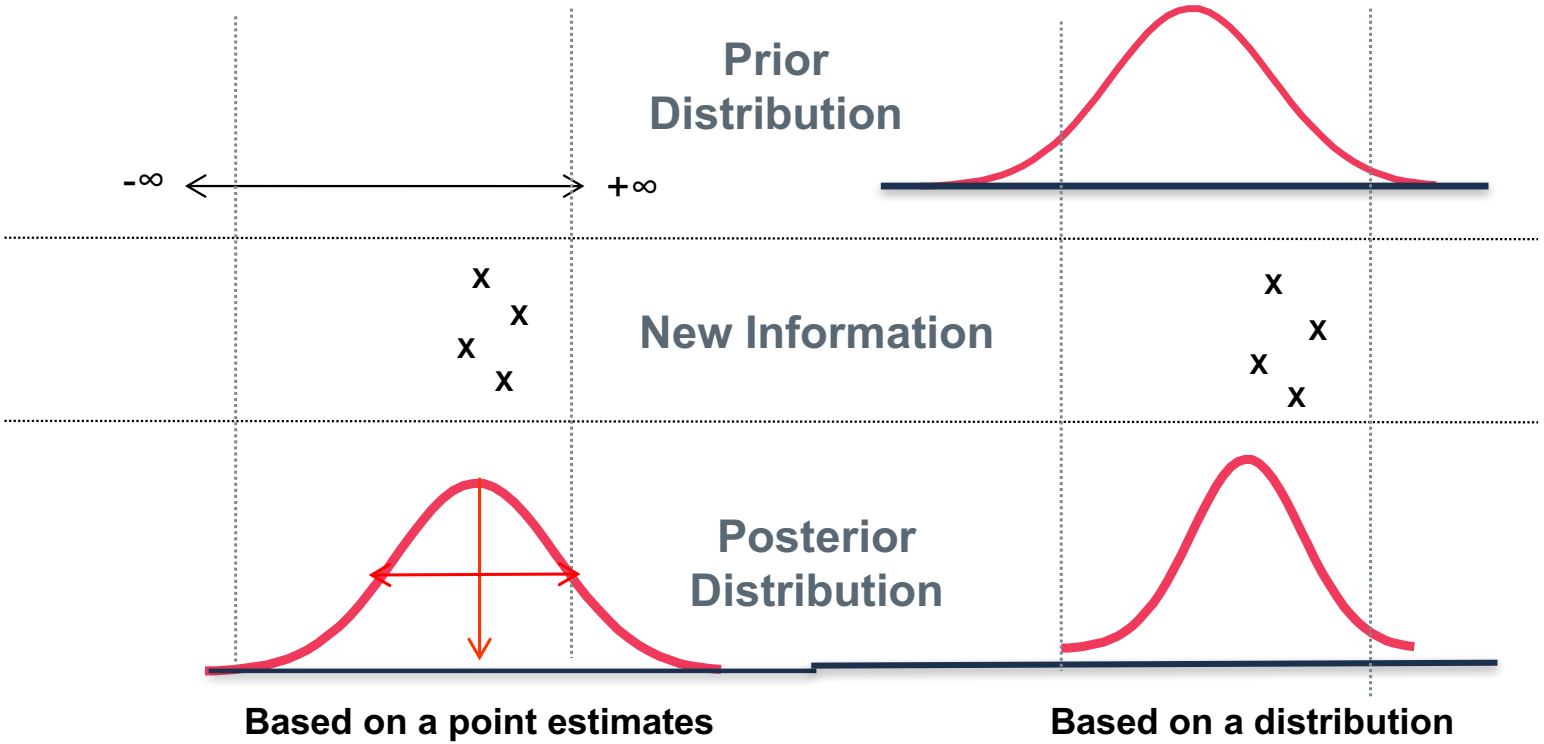
1. Provide prediction of individual outcome (design space)
2. Leverage knowledge and prior information to provide better accuracy/precision and minimize experiments
3. Multivariate (joint output and systems estimation)
4. Flexible (unbalanced experiments/sampling, hierarchical data, does not require Gaussian data)

1) Prediction of individual outcome

- To circumvent the danger of the use of the mean responses, it is sufficient to deal with the predictive distribution instead (more on that later)

- Why would you create a response surface design, an I-optimal design or a definitive screening design to stop with the mean?
  - All these designs are good to minimize the overall (or maximum) variance of prediction
  - So, at the modeling step, don’t forget to include the variance of prediction !

2) Prior knowledge: Frequentist vs. Bayesian Methods

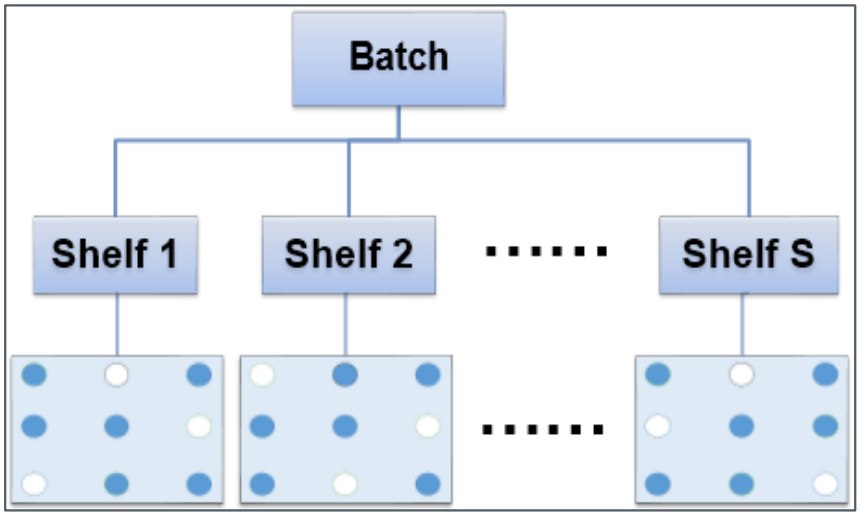


3) Multivariate

- Reality is **multivariate**
  - Several CQAs must *jointly* fall within the specifications
  - Output of unit operation affects downstream results
  - Interactions are not rare
- May be strong dependencies between the CQAs
- When using reduced DOEs (fractional factorials,etc.), degrees of freedom can become a challenge

4) Flexibility

- Multiple levels of hierarchy
- Unbalanced Sampling
- Flexible Probability Distribution



Bayesian Methods Meet These Criteria

- Bayesian methods provide a true prediction of individual unit future performance, i.e., the probability of meeting specification
- Leverage prior knowledge and experimentation, leading to better estimates, and fewer experiments and samples
  - addresses the shrinking degrees of freedom problem
- Complicated hierarchy/ sampling plan not a problem
- Bayesian modelling easily allows multivariate models
  - Joint prediction of multiple CQAs
  - Systems approach to unit operations
- Uncertainty of parameters included, thus improving prediction and reducing risk
- Not affected by non-centering within specification range

- To avoid using mean responses, one need a little bit more:
- The Predictive distribution
  - Function of the data uncertainty
  - Account for parameter uncertainty
  - Possibly account for prior knowledge
- Bayes' theorem (~1763)

$$p(\boldsymbol{\theta} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta}) \;\; p(\boldsymbol{\theta})}{p(\mathbf{y})}$$

$$p(\boldsymbol{\theta} \mid \mathbf{y}) \propto \mathcal{L}(\boldsymbol{\theta} \mid \mathbf{y}) \;.\; p(\boldsymbol{\theta})$$

$$\text{Posterior} \propto \text{Likelihood} \;\times\; \text{Prior}$$

Parameters distribution: the model
+ the prior knowledge

A Little bit of background

➤Prediction

- Achieved by replacing/integrating the parameters, including their (posterior) uncertainty, within the model

$$p(\tilde{y} \mid \mathbf{y}) = \int_{\boldsymbol{\theta}} p(\tilde{y} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta} \mid \mathbf{y})\, d\boldsymbol{\theta}$$

- Density of the prediction given a particular value of the parameters (the likelihood function)
- Posterior distribution of the model parameters
- Predictive distribution of a new response, integrating out the parameter distribution

...

- Let's skip the math

- See e.g.
  - J.J. Peterson, A posterior predictive approach to multiple response surface optimization, J. Qual. Technol. 36(2) (2004) 139-153.
  - J.J. Peterson, A Bayesian approach to the ICH Q8 definition of design space. J. Biopharm. Stat. 18 (2008) 959-975.
  - J.J. Peterson and K. Lief. The ICH Q8 Definition of Design Space: A Comparison of the Overlapping Means and the Bayesian Predictive Approaches. Statistics in Biopharmaceutical Research, 2:249–259, 2010.
  - P. Lebrun, B. Boulanger, B. Debrus, Ph. Lambert, Ph. Hubert, A Bayesian Design Space for analytical methods based on multivariate models and predictions, J. Biopharm. Stat. (2012) http://hdl.handle.net/2268/128222
  - P. Lebrun, F. Krier, J. Mantanus, H. Grohganz, M. Yang, E. Rozet, B. Boulanger, B. Evrard, J. Rantanen, and P. Hubert. Design Space Approach in The Optimization of The Spray-Drying Process. European Journal of Pharmaceutics and Biopharmaceutics, 80(1):226–234, 2012b.
  - Lebrun, P., Giacoletti, K., Scherder, T., Rozet, E., Boulanger, B., 2015. A quality by design approach for longitudinal quality attributes. J. Biopharm. Stat. 25, 247–259.
  - My PhD thesis freely available on http://bictel.ulg.ac.be/ETD-db/collection/available/ULgetd-12192012-155142/

Predictive distribution (summary)

- Model for multivariate regression

$$\underset{(n\times m)}{\mathbf{Y}} = \underset{(n\times p)}{\mathbf{X}}\ \underset{(p\times m)}{\mathbf{B}} + \underset{(n\times m)}{\mathbf{E}}$$

$$\mathbf{y}_i \sim N_m\left(\mathbf{x}_i\mathbf{B}, \mathbf{\Sigma}\right), \quad i = 1, ..., n,$$
$$\mathcal{L}\left(\mathbf{B}, \mathbf{\Sigma} \mid \mathbf{Y}\right) \propto |\mathbf{\Sigma}|^{\frac{-n}{2}} . \exp\left(-\frac{1}{2} tr\left[\mathbf{\Sigma}^{-1}\left(\mathbf{Y} - \mathbf{XB}\right)'\left(\mathbf{Y} - \mathbf{XB}\right)\right]\right)$$

- Non-informative priors (~classical results)

$$p\left(\mathbf{B}, \mathbf{\Sigma}\right) \propto |\mathbf{\Sigma}|^{-\frac{1}{2}(m+1)}$$

Informative priors

$$\mathbf{B} \mid \mathbf{\Sigma} \sim N_{p\times m}\left(\mathbf{B}_0, \mathbf{\Sigma}, \mathbf{\Sigma}_0\right)$$
$$\mathbf{\Sigma} \sim W_1^{-1}(\mathbf{\Omega}, \nu_0)$$

- Posterior: $p\left(\mathbf{B}, \mathbf{\Sigma} \mid \text{data}\right) \propto \mathcal{L}\left(\mathbf{B}, \mathbf{\Sigma} \mid \mathbf{Y}\right) . p\left(\mathbf{B} \mid \mathbf{\Sigma}\right) . p\left(\mathbf{\Sigma}\right)$
- Predictive: $p\left(\tilde{\mathbf{y}} \mid \tilde{\mathbf{x}}, \text{data}\right) = \int_{\mathbf{\Sigma}} \int_{\mathbf{B}} p\left(\tilde{\mathbf{y}} \mid \tilde{\mathbf{x}}, \mathbf{B}, \mathbf{\Sigma}\right) . p\left(\mathbf{B}, \mathbf{\Sigma} \mid \text{data}\right) . d\mathbf{B} . d\mathbf{\Sigma}$

$$\tilde{\mathbf{y}} \mid \tilde{\mathbf{x}}, \text{data} \sim T_m\left(\tilde{\mathbf{x}}\hat{\mathbf{B}}, \left(1 + \tilde{\mathbf{x}}'(\mathbf{X}'\mathbf{X})^{-1}\tilde{\mathbf{x}}\right)\mathbf{A}, \nu\right)$$

$$\tilde{\mathbf{y}} \mid \tilde{\mathbf{x}}, \text{data} \sim T_m\left(\tilde{\mathbf{x}}\mathbf{M}_{\mathbf{B}\text{post}}, \left(1 + \tilde{\mathbf{x}}'(\mathbf{X}'\mathbf{X} + \mathbf{\Sigma}_0^{-1})^{-1}\tilde{\mathbf{x}}\right)(\mathbf{\Omega} + \mathbf{A}^*), \nu + n_0\right)$$

$$\hat{\mathbf{B}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y}$$
$$\mathbf{A} = \left(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}\right)'\left(\mathbf{Y} - \grave{\mathbf{X}}\hat{\mathbf{B}}\right)$$

$$\mathbf{M}_{\mathbf{B}\text{post}} = \left(\mathbf{X}'\mathbf{X} + \mathbf{\Sigma}_0^{-1}\right)^{-1}\left(\mathbf{X}'\mathbf{X}\hat{\mathbf{B}} + \mathbf{\Sigma}_0^{-1}\mathbf{B}_0\right)$$
$$\mathbf{A}^* = \mathbf{Y}'\mathbf{Y} + \mathbf{B}_0'\mathbf{\Sigma}_0^{-1}\mathbf{B}_0 - (\mathbf{X}'\mathbf{X}\hat{\mathbf{B}} + \mathbf{\Sigma}_0^{-1}\mathbf{B}_0)'(\mathbf{X}'\mathbf{X} + \mathbf{\Sigma}_0^{-1})^{-1}(\mathbf{X}'\mathbf{X}\hat{\mathbf{B}} + \mathbf{\Sigma}_0^{-1}\mathbf{B}_0)$$

One last word : Degrees of Freedom

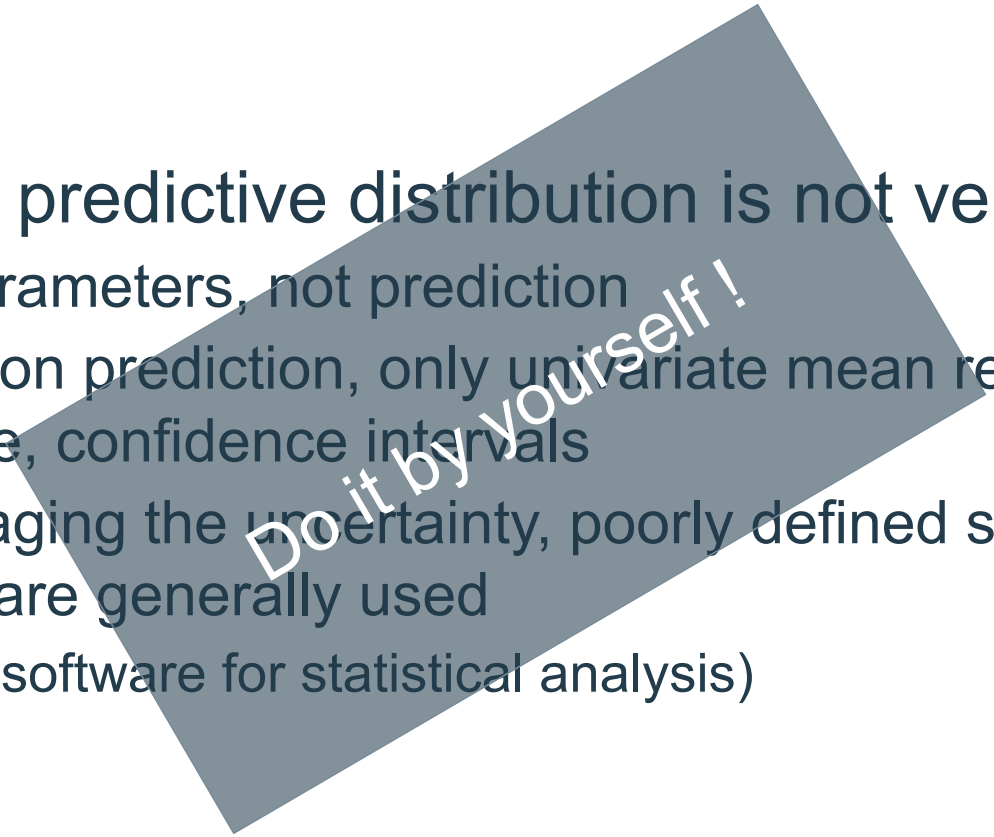
- In multivariate regression, $\nu = n-(m+p)+1$ (compare to the classical $n-p$)
- Hence, some additionnal d.f. are lost (m), du to the fact that one must
  - 1) estimate p different parameters for each response
  - 2) estimate the correlation/covariance between the responses.
- When creating a design, most *D*-optimal and *I*-optimal algorithm makes sure that some degrees of freedom are left after estimation of the model
- In multi-response problems however, this is not taken into account

Implementation

- The choice of R ?
- Basic existing functionalities (lm(),...)
- Extending to multivariate responses
  - Predictive distribution
  - Normal and Student's *t*
- Monte-Carlo simulations using R

Why R

- Unfortunately, predictive distribution is not very used
  - Focus on parameters, not prediction
  - When focus on prediction, only univariate mean responses and in the best case, confidence intervals
  - When envisaging the uncertainty, poorly defined statistical frameworks are generally used
    - (See GUI software for statistical analysis)



Why R

- Several tons of available packages
  - Often, a complete solution to the problem already exists
- A language built for statistics and mathematics
- Infinite flexibility... powerful language
- Beautiful graphics
- Free

R introduction

- How to create a linear model object... from a designed set of experiments

X Y

```
> head(data)
  Inlet.Temperature Feed.Rate Spray.Flow.Rate Yield Tapped.Density Moisture.Content Bulk.density Fraction.Resp
1               165       5.0              45 82.60         0.4613              0.9        0.419        58.710
2               110       2.5              30 73.76         0.4777              0.5        0.407        45.565
3               110       2.5              60 83.77         0.4811              0.7        0.436        66.970
4               110       7.5              30 42.88         0.4934              0.8        0.406        22.440
5               110       7.5              60 65.13         0.3213              1.1        0.288        30.440
6               165       5.0              45 88.50         0.4611              1.0        0.419        53.810

> # data transformation : ensure good modeling properties
> data["yield"]  = log(data["Yield"]/(100-data["Yield"]))    # percentage (not negative or > 100)
> data["tapped"] = log(data["Tapped.Density"])               # positivity
> data["moisture"] = log(data["Moisture.Content"])           # positivity
> data["bulk"] = log(data["Bulk.density"])                   # positivity
> data["fraction"] = log(data["Fraction.Resp"]/(100-data["Fraction.Resp"])) # percentage

> model =  lm(yield ~ Inlet.Temperature + Feed.Rate + Spray.Flow.Rate +
              I(Inlet.Temperature^2) + I(Feed.Rate^2) +
              Inlet.Temperature:Spray.Flow.Rate + Inlet.Temperature:Feed.Rate:Spray.Flow.Rate, data)
```

Bayesian predictive model

$$\underset{(n\times m)}{\mathbf{Y}} = \underset{(n\times p)}{\mathbf{X}}\ \underset{(p\times m)}{\mathbf{B}} + \underset{(n\times m)}{\mathbf{E}}$$

$$\mathbf{y}_i \sim N_m\left(\mathbf{x}_i\mathbf{B}, \mathbf{\Sigma}\right), \quad i = 1, ..., n,$$

$$\mathbf{B} \mid \mathbf{\Sigma}, \text{data} \sim N_{p\times m}\left(\hat{\mathbf{B}}, \mathbf{\Sigma}, (\mathbf{X}'\mathbf{X})^{-1}\right)$$

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

```
> Y = as.matrix(datas[4:ncol(data)])
> # If only the X matrix could be obtained easily...
> #I am sure
> # lm() is computing it for me using the formula !
```

R

```
> X =  model.matrix(delete.response(terms(model)), data, model$contrasts)

> # Recreate the X matrix from the factors in data, including the intercept,
> # the squared terms, the interactions, etc.
```

- `contrasts` also play an important role
  - Ever wonder why results are not the same in SAS and in R when using qualitative factors ?
  - ex : `Run` is a qualitative factor

- In our example, no qualitative factor... ok then !

```
> contrasts(data$Run,length(unique(data$Run))) = contr.SAS(length(unique(data$Run)),contrasts=FALSE)

> # Now, use you can use Run in a R formula and confirm a SAS result !
> # ?contr.SAS will give you all the possibilities
```

```
> head(X)
>   (Intercept) Inlet.Temperature Feed.Rate Spray.Flow.Rate I(Inlet.Temperature^2) I(Feed.Rate^2)  ...
> 1           1               165       5.0              45                  27225          25.00  ...
> 2           1               110       2.5              30                  12100           6.25  ...
> 3           1               110       2.5              60                  12100           6.25
> 4           1               110       7.5              30                  12100          56.25
```

$$\mathbf{B} \mid \mathbf{\Sigma}, \text{data} \sim N_{p\times m}\left(\hat{\mathbf{B}}, \mathbf{\Sigma}, (\mathbf{X}'\mathbf{X})^{-1}\right)$$

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

$$\mathbf{\Sigma} \mid \text{data} \sim W_1^{-1}(\mathbf{A}, \nu)$$

$$\mathbf{A} = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})$$

```
> Y = as.matrix(datas[4:ncol(data)])
> X =  model.matrix(delete.response(terms(model)),
                                                    data, model$contrasts)
> XprimeX  = t(X) %*% X
> XprimeXinv = solve(xprimex)
> hatB =  xprimexinv %*% t(X) %*% Y
> A = t(Y - X%*%hatB)%*% (Y-X%*%hatB)
> F = ncol(X)
> M = ncol(Y)
> N = nrow(Y)
> nu= N-(M+F)+1
```

- All posterior parameters are now available
  - Samples from the posterior of B can now be obtained by sampling B from a matrix-variate Normal, conditional to Σ being an inverse Wishart
    - A matrix-variate Norm... what ?

$$\mathbf{B} \mid \mathbf{\Sigma}, \text{data} \sim N_{p\times m}\left(\hat{\mathbf{B}}, \mathbf{\Sigma}, (\mathbf{X}'\mathbf{X})^{-1}\right),$$

$$vec(\mathbf{B} \mid \mathbf{\Sigma}, \text{data}) \sim N_{pm}\left(vec(\hat{\mathbf{B}}), \mathbf{\Sigma} \otimes (\mathbf{X}'\mathbf{X})^{-1}\right)$$

Direct sampling from the marginal posterior

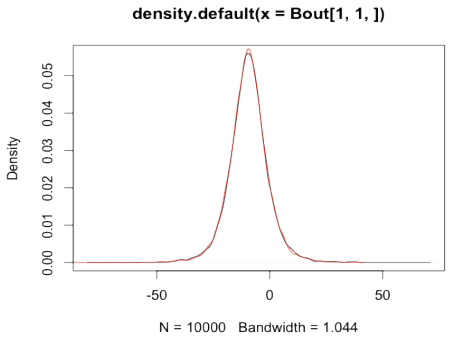
- Thanks to, among other, Box, Tiao, Zellner, Geisser, etc.

$$\mathbf{B} \mid \text{data} \sim T_{p\times m}\left(\hat{\mathbf{B}}, \mathbf{A}, \left(\mathbf{X}'\mathbf{X}\right)^{-1}, \nu\right)$$

- Again, a matrix-variate distribution : the Student's
- Not available in R nor in any language or software of my knowledge... but let's try that:

```
> rmatrixt2<- function(n=1,mean,Sigma,Omega,df){
                                          array(t(rmvt(n=n,delta=c(mean),sigma=kronecker(Sigma,Omega),
                                                                  df=df,type="shifted », method="svd")
                                              ),dim=c(nrow(mean),ncol(mean),n))
                                  }
> Bout2 = rmatrixt2(nsim,hatB,A/(nu),xprimexinv,nu)

> plot(density(Bout[1,1,]))
> lines(density(na.omit(Bout2[1,1,])),col="red")
```



Multivariate prediction

- To obtain a predictive distribution, one needs to solve:

$$p\left(\tilde{\mathbf{y}} \mid \tilde{\mathbf{x}}, \text{data}\right) = \int_{\mathbf{\Sigma}} \int_{\mathbf{B}} p\left(\tilde{\mathbf{y}} \mid \tilde{\mathbf{x}}, \mathbf{B}, \mathbf{\Sigma}\right) . p\left(\mathbf{B}, \mathbf{\Sigma} \mid \text{data}\right) . d\mathbf{B} . d\mathbf{\Sigma}$$

- This may take several pages of mathematical scribbles... but for the linear (fixed) case, the solution exists !
  - Predicting the multivariate response at condition $\tilde{\mathbf{x}}$ yields:

$$\tilde{\mathbf{y}} \mid \tilde{\mathbf{x}}, \text{data} \sim T_m \left(\tilde{\mathbf{x}}\hat{\mathbf{B}}, \left(1 + \tilde{\mathbf{x}}'(\mathbf{X}'\mathbf{X})^{-1}\tilde{\mathbf{x}}\right) \mathbf{A}, \nu \right)$$

  - What is great is that everything is already computed ! But care must be taken in the parameter definition of the *t* (scale vs. covariance)

```
> C_1 =  c(1  + x0 %*% xprimexinv %*% t(x0))
> varY = A/(nu)
> postmean = x0 %*% hatB
> ysim = rmvt(n=nsim,delta=postmux0,C_1*varY,df=nu) #nsim = 2000
```

Original scale

- Monte-Carlo simulations allow simply dealing with response transformations by propagating the uncertainty

```
> head(ysim)
          yield      tapped     moisture      bulk    fraction
[1,] -0.5975110 -0.75475915 -0.903692200 -0.9420145 -2.8076583
[2,] -2.8643233 -0.78628798 -1.472805905 -1.0787624 -7.3316669
[3,] -0.9052849 -0.09001934 -0.008912572 -0.2713737  3.7640472
[4,] -0.4235003 -0.56467757 -0.411099144 -0.7308374  0.1136780
[5,]  0.1094017 -0.52959732 -0.363973128 -0.6825366 -0.7368115
[6,]  4.1056835 -0.58245258 -0.496938347 -0.6553363  4.3186270

> ysim3[,c(1,5)] =100 /(1+exp(-ysim[,c(1,5)])) #logit
> ysim3[,2:4] = exp(ysim[,2:4]) #log

> hausner = ysim[,2]/ysim[,4] #compute important CQAs from responses
> hausner[hausner<1] = NA       #manage constraints during prediction
> ...
```

Original scale

- Now it is pretty simple to obtain mean responses or prediction intervals

```
> meanysim = apply(ysim,2,mean)
> beta= 0.95
> quantilemean =  apply(ysim,c(2),quantile,probs=c((1-beta)/2,(1+beta)/2),names = F)
```

  - Notice the mean is not especially relevant for non-Normal variables
  - HPD intervals might be preferred over quantiles

```
> library(MCMCpack)
> quantilemean =  apply(ysim,2,function(m) HPDinterval(as.mcmc(m),prob=beta))
> quantilehausner =  HPDinterval(as.mcmc(hausner),prob=beta)
```

Predictive risk-based results

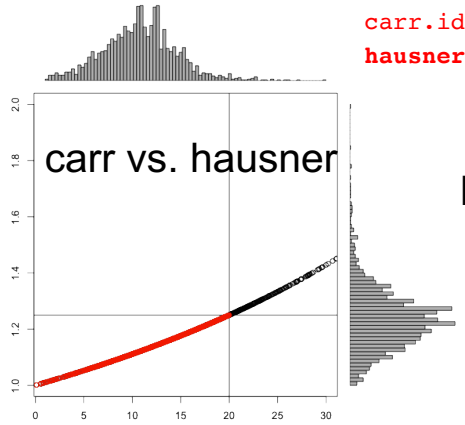
- Neither mean responses nor intervals indicates information about process capability
- Here comes the specifications
  - Assume we want to know the probability the yield is within the following specifications (NLT 70 to 90%)

```
> res["yield>70"] = sum(ysim[,1]>70)/nsim
> res["yield>80"] = sum(ysim[,1]>80)/nsim
> res["yield>90"] = sum(ysim[,1]>90)/nsim
```

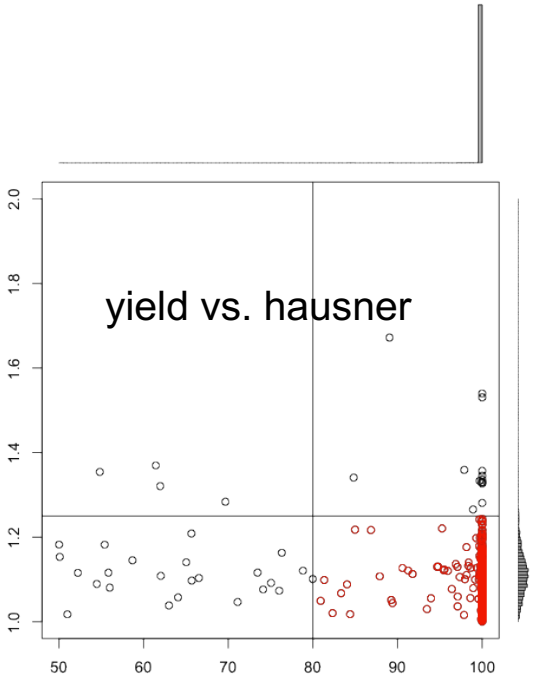
Predictive risk-based results

- The beauty of MC simulations is to let the correlations/dependencies speak without effort
  - e.g. below: check the five specifications jointly

```
> res["Joint.Proba"] = sum( ysim[,1]>80 &     # Yield>80%
                            ysim[,3]<1  &     # Moisture<1%
                            ysim[,5]>50 &     # Inhalable fraction>60%
                            carr.idx <20  &    # ...
                            hausner <1.25, na.rm=TRUE)/nsim
```



It is the MC estimate of the posterior probability $P(CQAs \in \Lambda)$



Example 1: a Spray-Dryer

Spray-drying process

- Spray-drying is intended to create a powder with small and controlled particle's size for pulmonary delivery of a drug substance
- Several Critical Process Parameters (CPP) have an influence on several Critical Quality Attributes (CQA)
  - CPP: inlet temperature, spray flow-rate, feed rate

  (other process parameters are kept constant)
  - CQA: yield, moisture, inhalable fraction, flowability
- Specifications on CQA defined as minimal satisfactory quality
  - yield > 80%
  - moisture < 1%
  - Inhalable fraction > 60%
  - …

How to follow QbD ? Start with the end !

- The process must provide, in its future use, **quality outputs**
  - e.g. during routine
- According to specifications derived from safety, efficacy, economical reasons
  - Whatever future conditions of use, that are not always perfectly controlled
  - Then, outputs should be **not sensitive** to minor changes
- This is **Quality by Design**
  - The way the process is developed leads to the product quality
  - This quality and the associated risks are assessed
  - Achieved using Design Space methodologies

Spray-drying process

- Design Space, Risk and ICH Q8
  - ICH Q8 proposes to use the Design Space (DS) risk-based methodology to fulfil these objectives

Target : "*Understand and gain knowledge about a process to find a parametric region of **reliable robustness** for **future performance** of this process*"

→ Assurance of quality
→ Assessment of the risk not to achieve quality

  - Eases all business decisions

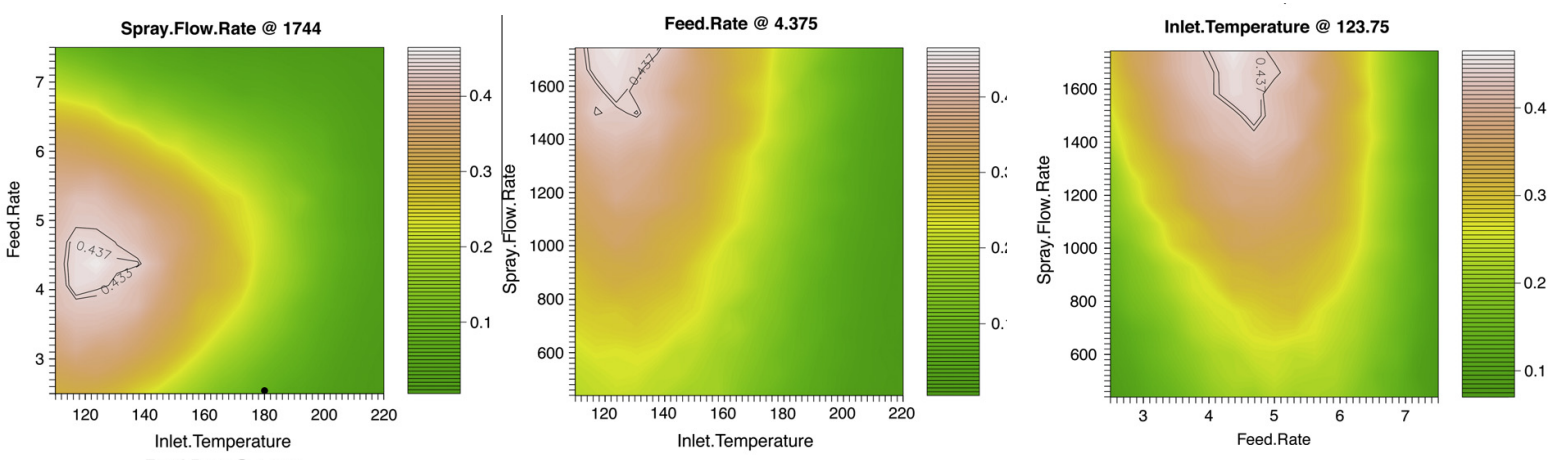
Computation

- This implies to know the behavior of the CQAs in the future
  - How they change when CPPs change
  - How they are statistically *distributed*
  - How they are dependent
- Fortunately, solutions exist in the Bayesian statistical framework for every problem !

- (See previous R codes)

Spray-drying process

- Risk-based design space: predicted P(CQAs$\in \lambda$)
- NOT A RESPONSE SURFACE !!



- In the Design Space, there is merely 45% of chance to observe each CQA within specification, jointly
- There is also 100-45% = 55% of risk not to observe the CQAs within specification (jointly) !

Spray-drying process

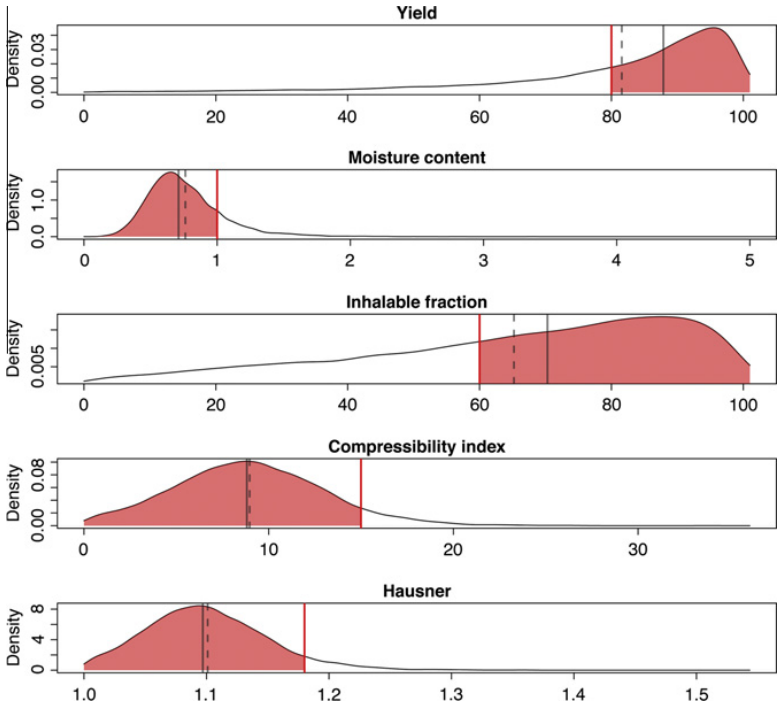
- Validation
  - Experiments have been repeated 3 times independently at optimal condition, i.e.
    - Inlet Temperature: 123.75°C
    - Spray Flow Rate: 1744 L/h
    - Feed Rate: 4.69 ml/min

    - Jointly, 2 out of the 3 runs within specification

Batches	Yield (%)	Moisture content (%)	Inhalable fraction (%)	Compressibility index	Hausner ratio
1	88	<0.2	63	11.6	1.13
2	89	<0.2	62	12	1.14
3	88	<0.2	59	11.5	1.13
Mean	**88.7**	**<0.2**	**61.18**	**11.76**	**1.13**
Standard deviation	0.61	NA	1.82	0.22	0.01

Spray-drying process

- Post-analysis (« How they are statistically *distributed* »)
  - Marginal predictive densities of the CQAs



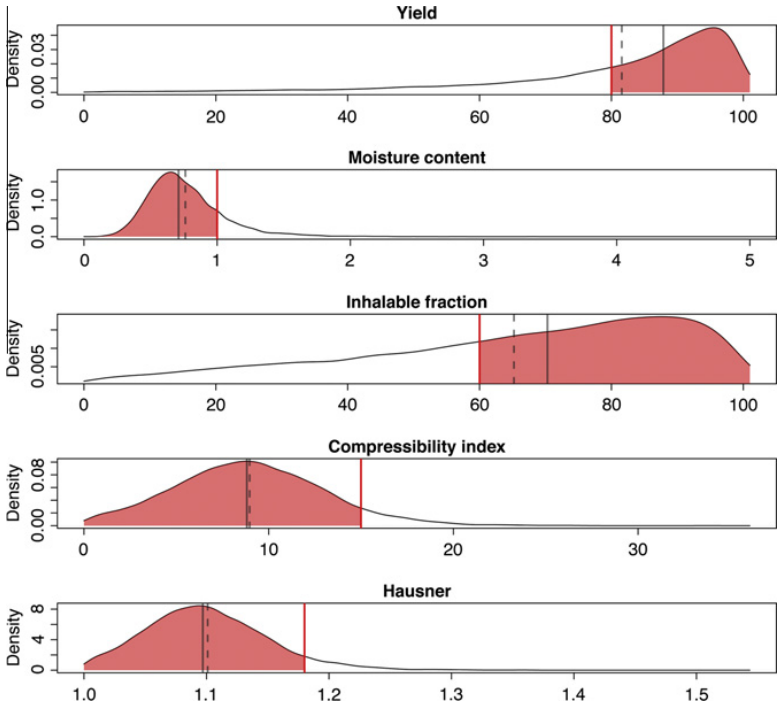
Compared with validation SD,
these uncertainties seems huge

Poor model fit !
Need to increase knowledge !

Predictive uncertainty =
data uncertainty + model uncertainty

Spray-drying process

- Post-analysis (« How they are statistically *distributed* »)
  - Marginal predictive densities of the CQAs



Remember the danger taking only into account mean responses !

So critical when making DoE, as the minimal number of experiments is searched...

...while often, poor knowledge on factor effects misleads the choice of the design type !

Spray-drying process

- Conclusion
  - Effective Design Space is the ultimate tool to optimize a process or a method while concurrently assessed its robustness
    - To provide guarantee that future runs will be on specifications
  - Even in presence of poor model fit…
    - Here, due to a poorly designed set of experiments
  - … it allows providing risk-based results
    - But guarantee is kept low (45%)

Gain

- What are the benefits for industry ?
  - Classical benefits due to DOE
    - The time to run experiments before obtaining results is controlled
    - This time is generally reduced in comparison to “handmade” optimization. Costs are reduced as well
  - Benefits due to risk-based Design Space
    - Guarantee and risk to be on specification are controlled
    - Process/method knowledge leads to quality product and robustness
    - Robustness generally eases transfer between manufacturing sites, for instance
    - Better quality products also allows reducing costs
      - Less batches out-of-specification
      - Improvement of process reliability

Example 2: a Pharmaceutical Formulation

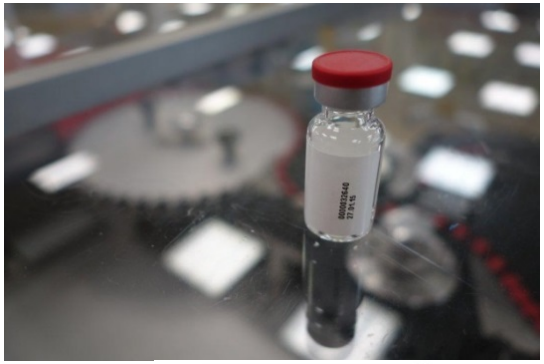
Acknowledgment: Renske Hesselink, Xavier Lories

Context

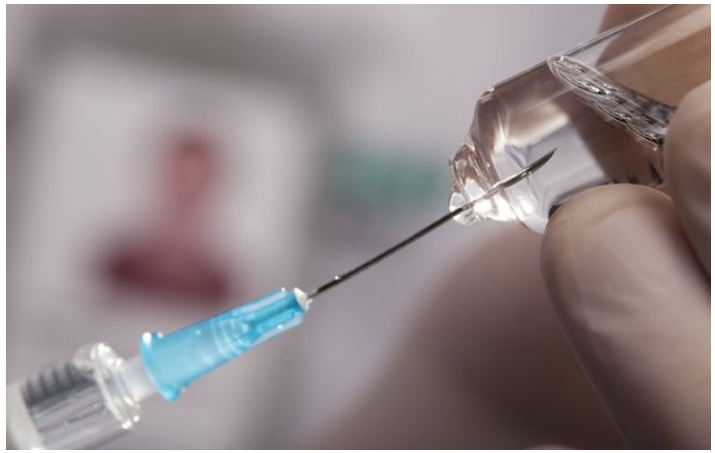
- Formulation Stability study to optimize the shelf-life, stress, and accelarated stability of a vaccine

- Potency (log IU/mL), API concentration, aggregation, and visual appearance are evaluated with potency assays, QPCR, and other analytical assays, at t_0 and t_{lt} using 3 replicates
- Several classical stress conditions are assessed
  - classical storage (1 year)
  - 10x freeze-thaw + agitation
  - accelerated stability (1 month at 25°C)
- Objective: Find *stable* formulation ranges out of 8 identified formulation factors $X_1...X_8$

Remain stable under challenging conditions



Vaccine formulation development



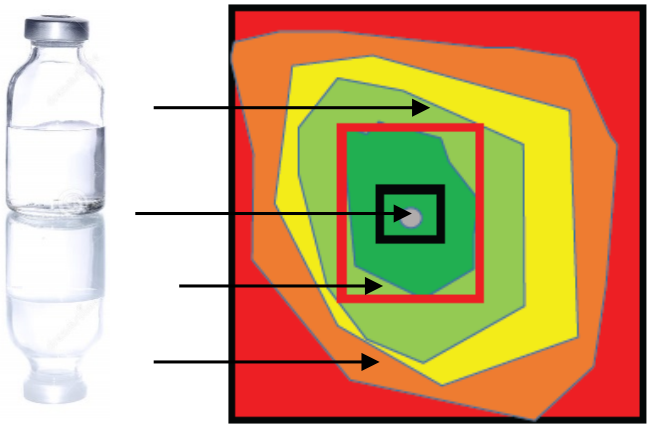
Therefore, a formulation is developed that ensures:

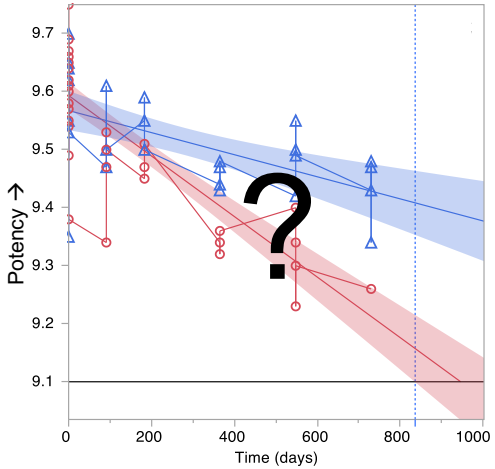
- Efficacy
- Physical, chemical and biological stability
- Easy administration procedure
- Optimal release, delivery and presentation of the molecule at the target site
- Manufacturability
- Low cost of goods
- Minimum side effects
- Ideally without cold chain

Why formulation robustness?

Sources of variability

- Errors in buffer preparation
- Variability in raw materials
- Adsorption or retention of excipients during manufacturing process
- pH changes
- Evaporation
- Excipient degradation

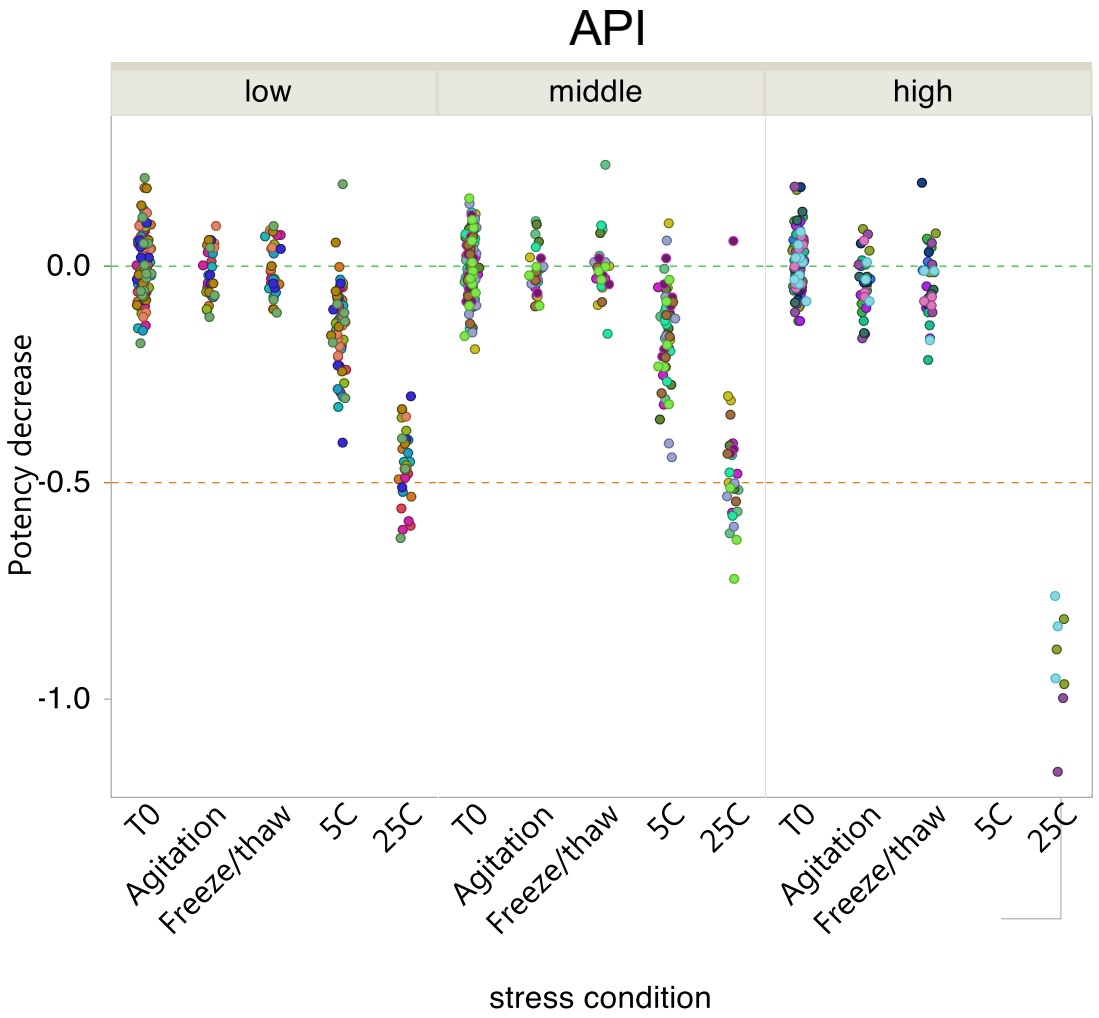




2 years stable at 2-8°C

Results

- Thermal stress has most impact on potency
- Still, all formulations within specification after 2 year at 5°C, right?
- No: after thermal stress, many formulations at high titer were < LLOQ and could not be measured



Critical quality attributes and specifications

- Responses (each modeling one critical quality attributes) must lie within pre-defined specifications

Assay	Response	Acceptance criterion
QPA	Potency	ΔPotency limit = -0.3 log IU/mL
vp-Q-PCR	Virus titer	ΔTiter limit = -0.3 log vp/mL
DCS	Aggregation	Aggregation score < 1
RP-HPLC	Relative peak height (weighted geomean)	Peak height score > 0.80
Visual appearance	Particles observed	Visual appearance score ≤ 1

- Difficulty: the responses are not the critical quality attributes, but their kinetics of degradation

Problem formalization

- Critical Quality Attributes
  - E.g.: difference of potency $\delta_{0\text{-}lt}$: Y
  - Format : a reportable result is the mean of three replicates
- Specifications
  - reportable results of $\delta_{0\text{-}lt}$ > -0.3 IU/mL

- Factors
  - 8 formulations factors have been identified as Critical Process Parameters (CPP)

pH	Surfactant	Cryoprotectant	Salt	Buffer	API	Anti-oxydant	…

  - An experimental design comprising 20 experiments has been conducted for every stress condition

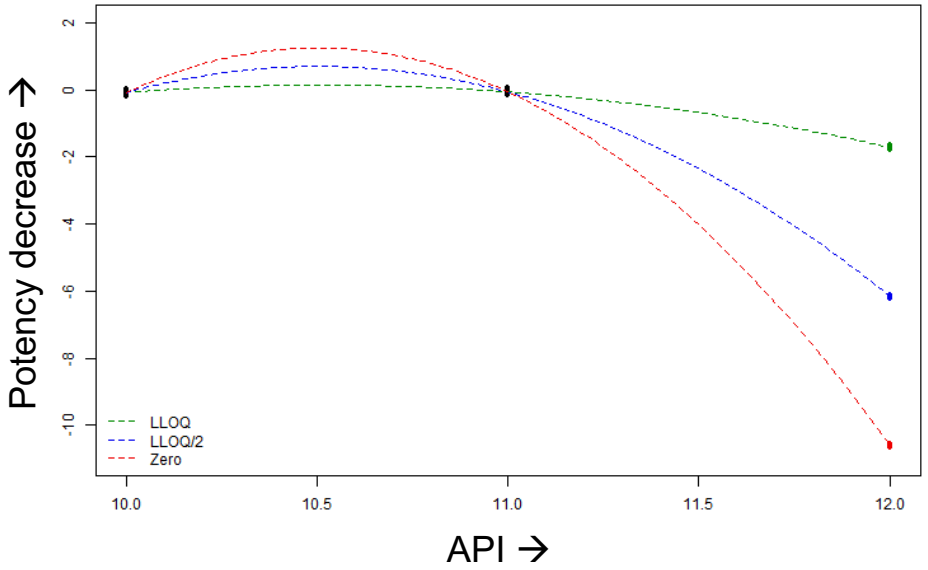
Designed experiments

- Over the 8 factors, a 20 experiments fractional factorial design with resolution IV

- Not a “screening” design, but a “robustness” design
  - Interested in predictions and (less) in parameters
- Factor ranges chosen as normal variation around a target value
- Is all the experimental domain providing a satisfactory stability ?
- Problem: can you trust (mean) predictions of such a design ?
  - Hint: No !

Which model to select?

- Statistics does not provide clarity… use scientific rationale
  - All main effects
  - Quadratic effect for titer → not perfect but better than linear
- Optional 2FI:
  - Interactions for pH × API and surfactant × API
  - Because degradation / aggregation at high API concentration is likely influenced by charge (pH) and surfactant concentration



Predictive Bayesian Model

- Individual predictions will be drawn and the reportable results will be derived using simulations
- Take a lot of time to adjust your model
  - All your decisions are based upon it !
  - “Bad” model leads to very high predictive uncertainty
  - Take care not to overfit your data
- A multiple regression is adjusted
  - Will the attribute(s) be well explained by a (Normal) linear model ?
  - Do you need combination of variables, to transform them ?
  - Y = **Xb** + e
  - A little bit trickier than previously, as missing data are common (not at random, but <LOQ)
    - Uses a regularized horsehoe prior multivariate regression with censored data imputation

Design Space computation

- One simulation for one factor setting
  - From the predictive distribution, sample individual response predictions (reportable results)
  - compare to specification
- From many *n** simulations
  - Compute the MC estimate of the posterior probability of success

$$P(\mathbf{C}\tilde{\mathbf{Q}}\mathbf{A} \in \mathbf{\Lambda} \mid \tilde{\mathbf{x}}, \text{data}) \simeq \frac{1}{n^*} \sum_{s=1}^{n^*} I(\mathbf{C}\tilde{\mathbf{Q}}\mathbf{A}^{(s)} \in \mathbf{\Lambda})$$

- For a grid over the factor setting
  - Draw maps of the posterior probabilities P(success)
  - Identify Design Space: $\{\tilde{\mathbf{x}} \in \chi \mid P(\tilde{\delta}_{0-lt} > -0.3 \mid \tilde{\mathbf{x}}, \text{data}) \geq \pi\}$
- For all the CQA jointly:
  - Use the joint distribution to account for correlations

Design Space representation: DoE considerations

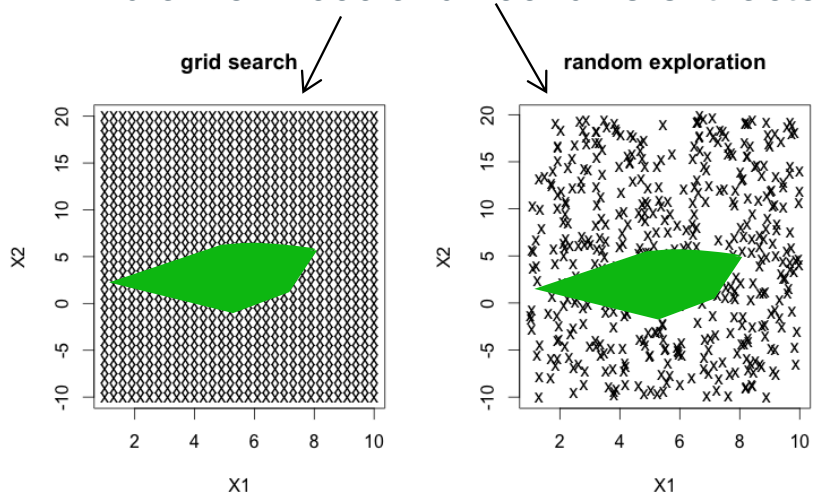
- Unfortunately, not possible to explore every factor setting
  - DoE to analyze only the Critical Process Parameters
- Obviously, the analyst often believes that a lot of factors will impact his/her quality... and might be right about it !

- Computationally, there is a problem to represent high-dimensional spaces of factors
  - Assume we want to explore a grid made from 10 points per factor...
  - 8 factors.....10^8 conditions to explore !

- Parallelization, computer clusters, etc., are of no help in this case

Design Space representation: curse of dimensionality

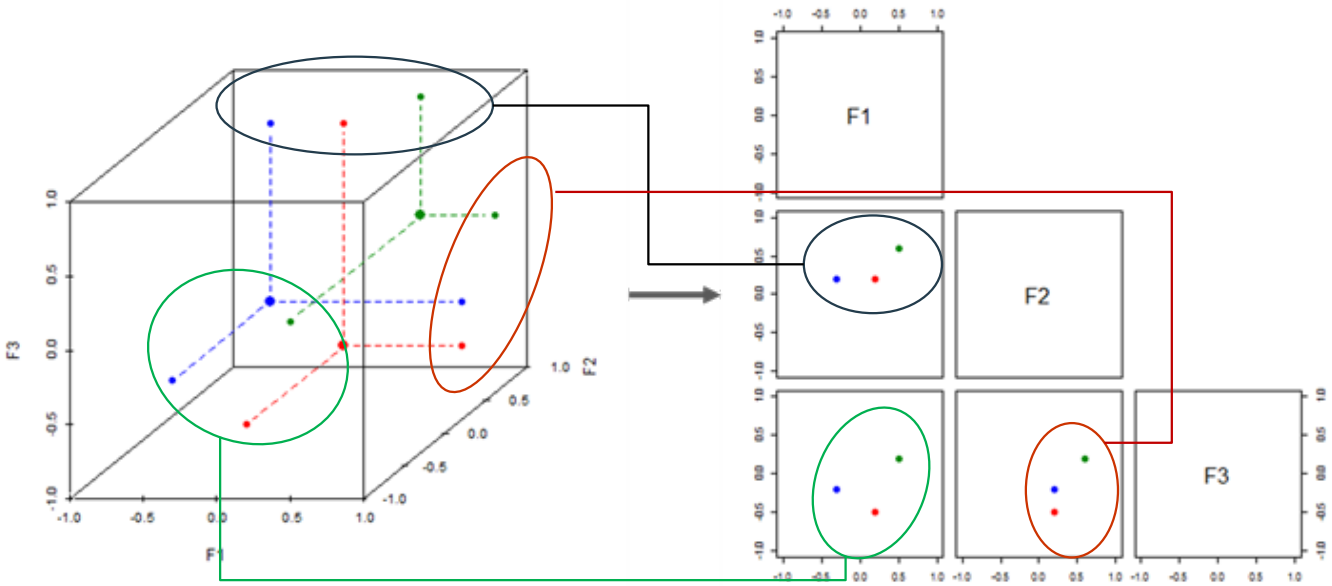
- A possibility is to explore the experimental domain by drawing randomly from a multivariate uniform distribution covering the space of factors (space-filling design for computer simulations)
  - Ex : draw of 1000 and 400 different factor settings



- On each point, compute the (posterior) probability of success
- Then, create bivariate pair-plots of the factors

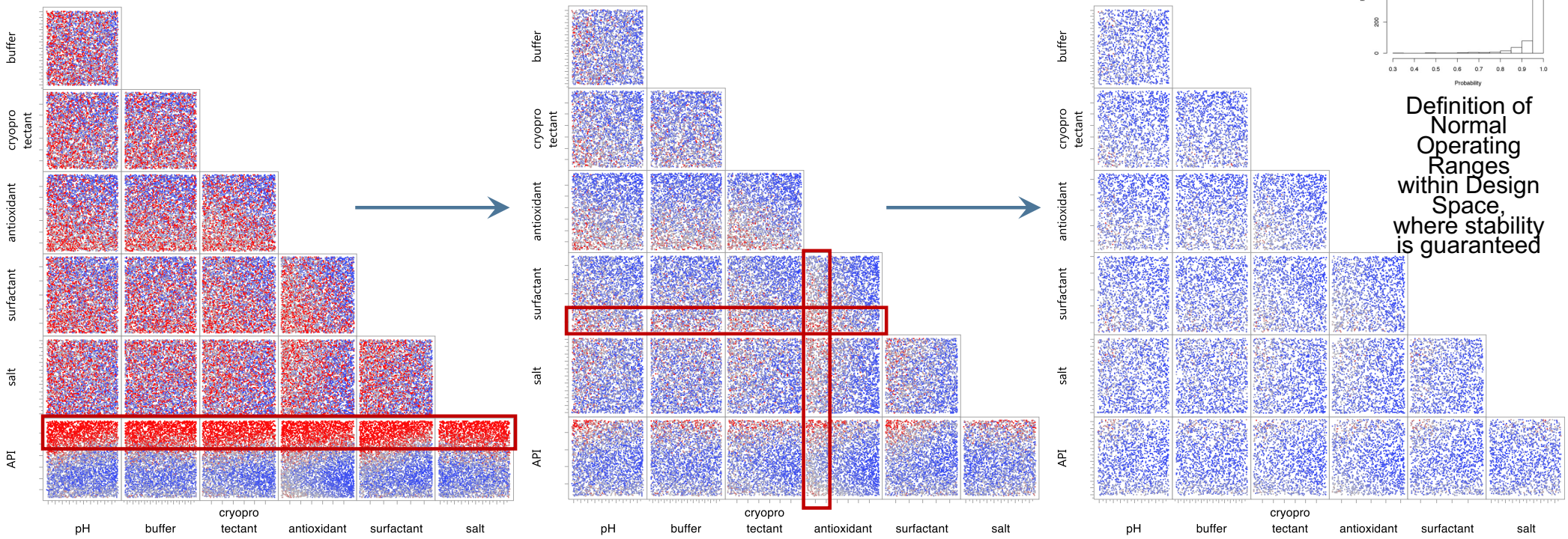
Viewing the results: projections

- Predictions are made for factor setting in the computer simulation, for the probabilities of success to meet the specifications
- The 8-dimensional experimental space is projected as pair plots / scatter plot matrix
- Each simpler pairs plot is a view of the total number of (projected) simulations



Determining the design space

- Find a subspace of the experimental domain where P(success is sufficient (~blue))
- Balance between highest P(success), and what is feasible in process



Conclusions

- Design Space is a tool build over DoE
  - The advantages of DoE are kept...
  - ...while fully taking into account all uncertainties and dependencies to make sure the decision and the associated risks are controlled
- From the 8 process parameters, most of them were found “not so critical” and the risk-based optimization over only some of them allowed to improve and control the drug formulation to obtain satisfying stability given pre-defined specifications
- Pairs plots/scatterplot matrix with space filling designs can help when dimensionality is too high

Conclusion

Conclusions

- Use of Design Space methodologies at early stage of development allows
  - characterizing the chances of success based on a strong rationale as designed experiments are made
  - determining robust optimal factor ranges allowing easier post-approval changes of formulation
    - “working within Design Space is not considered as a change” (Q8)
- DS is not DoE
- DS is not the mean, it is the probability of individual success !
- Don’t limit yourself to response surfaces
  - The data to compute Design Spaces and predict chances of success is already there !
  - There is no guarantee that all will run smoothly with mean response surfaces !