A Bayesian model for filling of a product to reduce risk of being OOS in presence of uncertainty.

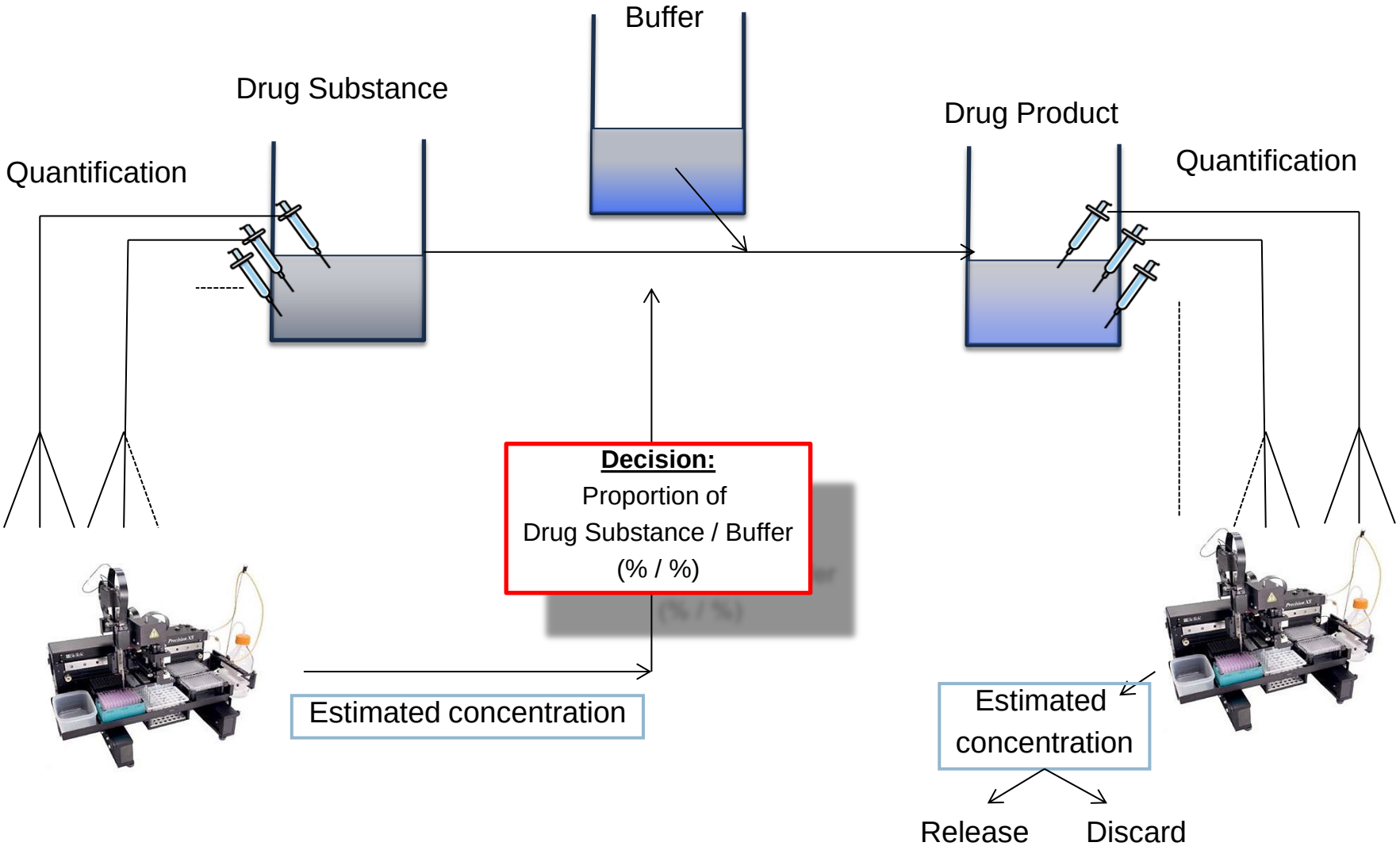
Bruno Boulanger - Pierre Lebrun –, BAYES2013, Rotterdam

Example: Dilution protocol

- Context
  - A manufacturing site receive a (concentrated) bulk Drug Substance
  - Dilution with a buffer has to be made to obtain the Drug Product ready to be filled in vials
- Problem
  - The quantification of both the Drug Substance and Drug Product is subject to uncertainty
  - Assume the specification release limit for Drug Product is LSL=2 mg/mL

    If the concentration>LSL, Drug Product -> discarded
- Question
  - What is the (safe) dilution to ensure the Drug Product is within specification ?
  - What is the format of the assays to obtain satisfactory reportable results about content ?

Dilution


Buffer
Drug Substance
Drug Product
Quantification
Quantification
Decision:
Proportion of
Drug Substance / Buffer
(% / %)
Estimated concentration
Estimated
concentration
Release
Discard

Purpose of the dilution protocol

- Start with the end

  What is the very objective of the assays ?

- To provide results used to make important decisions
  - Release of a production batch after dilution
  - Optimization of a process (dilution protocol)
  - Etc…
- What matters are the results produced by the assays, not the assays!
  - *E.g. dilution will be decided based on the results obtained*
  - *E.g. batch will be released based on the results obtained…*

Purpose of the dilution protocol

- It must provide, in its future use, quality product
  - e.g. during routine
- According to specifications
- Whatever future conditions of use, that are not always perfectly controlled
  - Then, results should be not sensitive to minor changes

    e.g. dilution not perfect, failed assay
- This is Quality by Design
  - The way the assays are developed leads to know quality & risks

Dilution

- Simulations
  - Idea: test the dilution with different formats, at different levels of (mean) concentration for the Drug Substance
  - Remember that neither the Drug Substance nor the Drug Product concentrations are known with certainty
  - Thus, rely on the estimated posterior predictive distribution of the concentrations
- Question
  - What are the guarantees that, from an estimated concentration of the Drug Substance over a certain number of series and replicates, the resulting diluted Drug Product is within specification given an estimated concentration over a certain number of series and replicates ?

→ Design Space problem

→ Over the **dilution** factor and the **format**

Dilution

- Viewed as a Design Space problem
  - Optimization over the following factors

    J_{DS} : the number of (independent) series for Drug Substance measurements

    n_{DS} : the number of replicates/series for Drug Substance measurements

    J_{DP} : the number of (independent) series for Drug Product measurements

    n_{DP} : the number of replicates/series for Drug Product measurements

    d : the dilution to apply (% of Drug Substance)

    $conc_{DS}$: the true concentration of Drug Substance (not to be optimized)
  - CQA

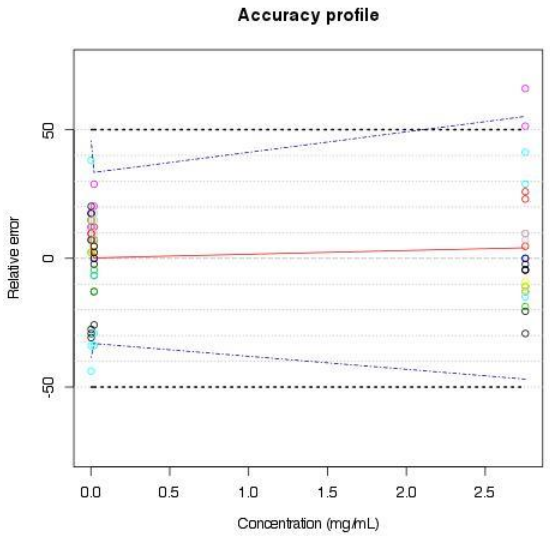
    $\bar{y}_{DP}$: Reportable results of the concentration of Drug Product
  - Specifications

    $\bar{y}_{DP} > 2 \ \ (mg/\mathrm{mL})$

    $2 < \bar{y}_{DP} < 2.4 \ \ (mg/\mathrm{mL})$, if possible

1st Data: prior information

- Precision of the assay is first provided with qualification data



9 series
3 replicates
3 concentration levels

- Assuming the level 2.7 mg/mL is the closest to the targeted concentration, and precision is homogenous among levels
  - The same precision will be used for any concentration levels of Bulk and Drug product
  - This assumption is useful as these are so far the only available data

One-way ANOVA random model

- Model description (for one concentration level)

$$y_{ij} = \mu + \alpha_j + \varepsilon_{ij}, \quad \text{with } \varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon) \text{ and } \alpha_j \sim N(0, \sigma^2_\alpha)$$

$$j = 1, ..., m, \quad i = 1, ..., n_j, \quad n = \sum_{j=1}^{m} n_j.$$

- Prior distributions for μ, σ^2_ε, σ^2_α
  - p(μ) = N(μ_0, τ_0) p(σ^2_ε) = gamma(a,b) p(σ^2_α) = gamma(c,d)
- Example of BUGS model for unbalanced data

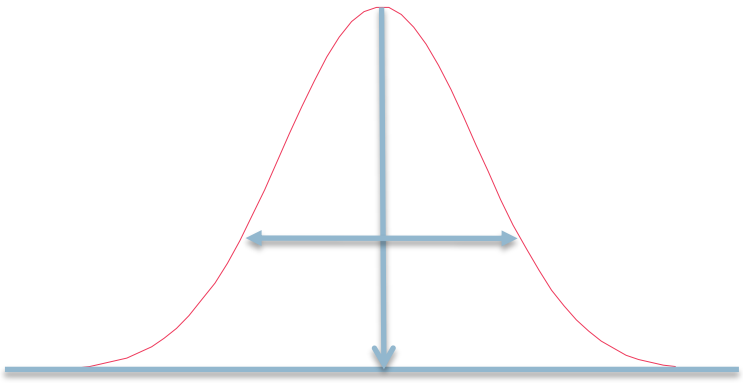
```
model{
        for(j in 1:m){
                for(i in n[j]:(n[j+1]-1)){
                    y[i] ~ dnorm(a[j], tau.e) #likelihood
                }
                a[j] ~ dnorm(mu, tau.a)
        }

        #flat prior distributions
        mu ~ dnorm(0, 0.0001)
        tau.a ~ dgamma(0.0001,0.0001)
        tau.e ~ dgamma(0.0001,0.0001)

        #convert precision into variance
        sigma2.e <- 1/tau.e
        sigma2.a <- 1/tau.a
}
```

Output:

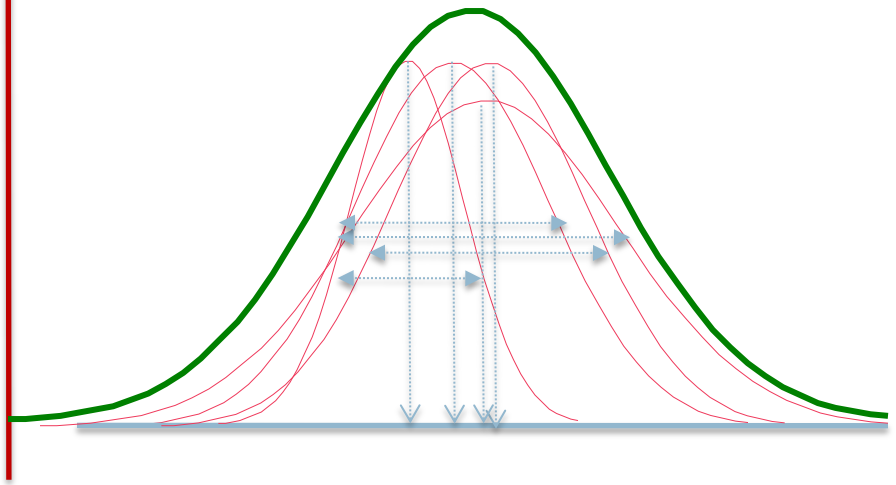
$$p(\mu, \tau_\alpha, \tau_\varepsilon \mid \text{data})$$

Difference Simulations/Predictions

the “new observations” are drawn from distribution “centered” on estimated location and dispersion parameters (treated as “true values”).



the uncertainty of parameter estimates (location and dispersion) is taken into account before drawing “new observations” from relevant distribution



To account for uncertainty of prediction under the variation of series and replicates

→ One way ANOVA random model

→ accounting for uncertainty of (mean and variances) parameter estimates

One-way ANOVA random model

- Prediction applied to this model

$$p(\tilde{y} \mid \text{data}) = \int_{\mu} \int_{\tau_\alpha} \int_{\tau_\varepsilon} p(\tilde{y} \mid \mu, \tau_\alpha, \tau_\varepsilon).p(\mu, \tau_\alpha, \tau_\varepsilon \mid \text{data}).d\tau_\varepsilon.d\tau_\alpha.d\mu$$

  - Not solvable (see Mee's approximation, 1984)

- Sampling scheme to obtain samples from the predictive distribution

For $s = 1$ to n^*

1. sample $(\mu^{(s)}, \tau_\alpha^{(s)}, \tau_\varepsilon^{(s)})$ from $p(\mu, \tau_\alpha, \tau_\epsilon \mid \text{data})$, (from BUGS output)
2. sample $\tilde{\alpha}^{(s)}$ from $N(\mu^{(s)}, \sigma_\alpha^{2(s)})$, or from $N(\mu^{(s)}, 1/\tau_\alpha^{(s)})$, (in R or SAS)
3. sample $\tilde{y}^{(s)}$ from $N(\tilde{\alpha}^{(s)}, \sigma_\varepsilon^{2(s)})$, or from $N(\tilde{\alpha}^{(s)}, 1/\tau_\varepsilon^{(s)})$.

End

Dilution

- one simulation at one “operating condition”

1. **Measure of the Drug Substance**

uncertainty of the mean scaled to target concentration

Sample J_{DS} series ($\tilde{a}_{DS}^{(s)}$ | data) following N($conc_{DS}$ | data, σ_a| data)

- For each series s, sample n_{DS} measurements N($\tilde{a}_{DS}^{(s)}$ | data, $\sigma_e^{(s)}$ | data)
- Compute the grand mean (rep. result) of the n_{DS}*J_{DS} measurements: $\bar{y}_{DS}$

2. **For a given dilution d**

- Compute the true but unknown concentration of Drug Product: $conc_{DS}$*d
- Compute the concentration one would obtain if $mean_{DS}$ was true: $\bar{y}_{DS}$ * d

3. **Measure of the Drug Product**

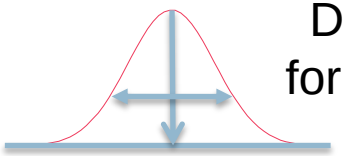
Sample J_{DP} series ($\tilde{a}_{DP}^{(s)}$ | data) following N($conc_{DS}$ *d | data, σ_a| data)

- For each series s, sample n_{DP} measurements N($\tilde{a}_{DP}^{(s)}$ | data, $\sigma_e^{(s)}$ | data)
- Compute the grand mean of the n_{DP}*J_{DP} measurements: $\bar{y}_{DP}$

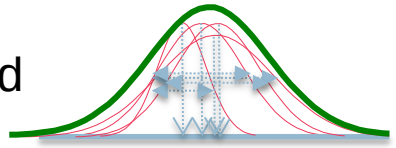
- 1000 simulations this “operating condition”

Compute $P(\bar{y}_{DP} > 2 \ \& \ conc_{DS}*d > 2 \mid data)$ using Monte-Carlo

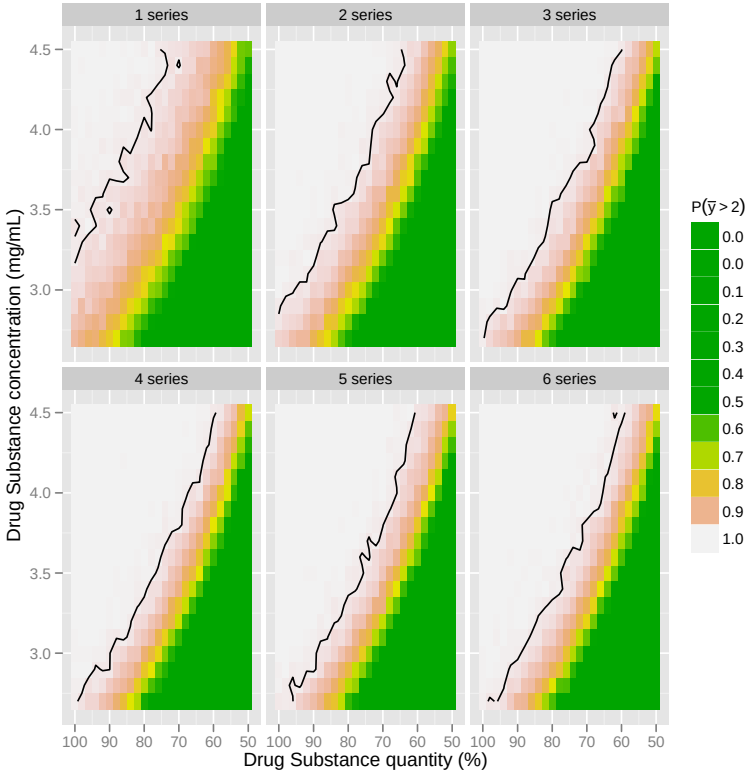
First results (process Design Space)



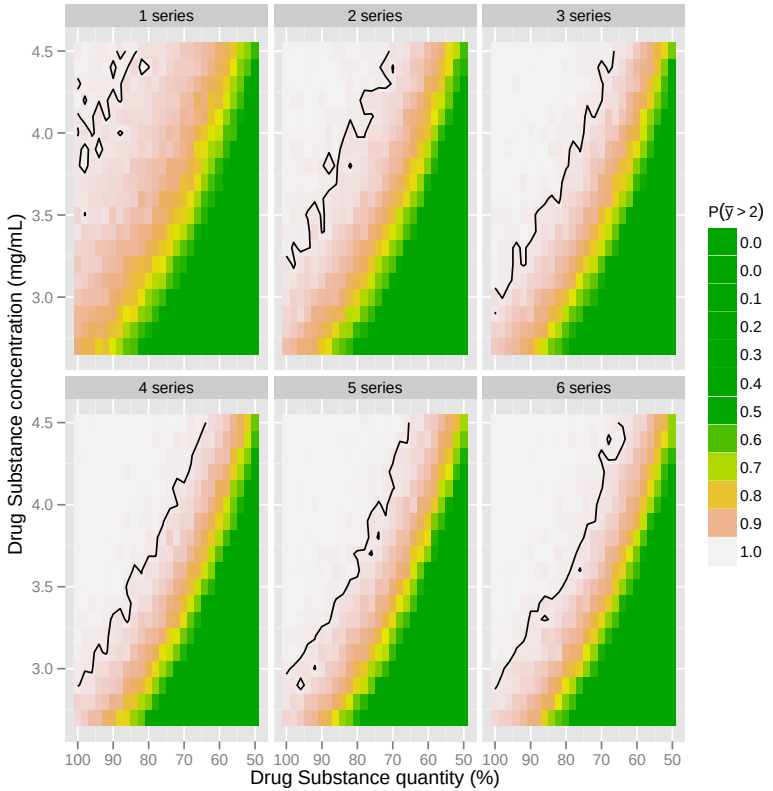
Draw a map of the probability
for every factor setting explored



Without parameter uncertainties



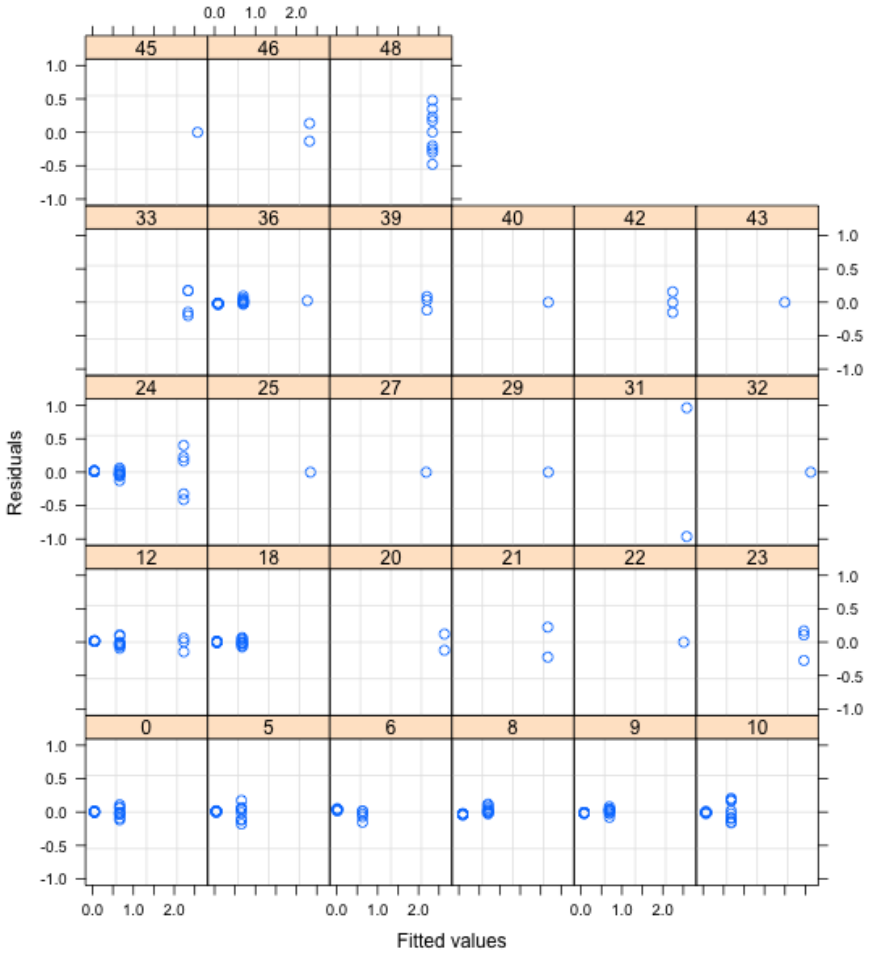
With parameter uncertainties



Black lines shows 99% quality level (posterior probability of CQA within specifications)

First prior was found poor to answer the question

- New data were available lately
  - stability data to assess uncertainty
    - 3 concentration levels (mg/mL)
    - 27 days of study
    - 3 runs per day
    - 3 repetitions per run
    - Sparse stability design
    - Many missing values
  - Interest in 2 variability sources
    - Residual variability
    - Effect of the runs
  - Mixed modeling



Dilution

- Model
  - $\text{Log-conc}_{ijkl} = \text{conc\_target}_l + \text{time}_k + a_j + e_{ijkl}$

    - conc_target and time are used as qualitative factors

    - a is a random effect for the series: $a_j \sim N(\mu_a, \sigma_a)$

    - e is the residual error: $e_{ijkl} \sim N(0, \sigma_e)$

    (fitted using the `MCMChregress` function of MCMCpack for R)

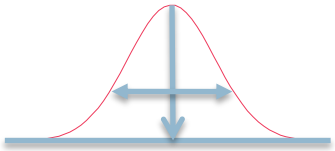
  - Allows drawing new individual measurements, whatever the series

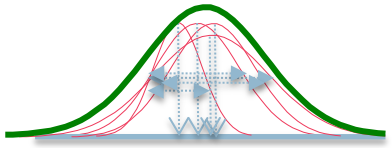
    1./ Sample a new series $(\tilde{a}^{(s)} \mid \text{data})$ from $N(\mu_a^{(s)} \mid \text{data}, \sigma_a \mid \text{data})$

    2/. Sample $(y^{(s)} \mid \text{data})$ from $N(\tilde{a}^{(s)} \mid \text{data}, \sigma_e^{(s)} \mid \text{data})$, s=1...#sampled elements

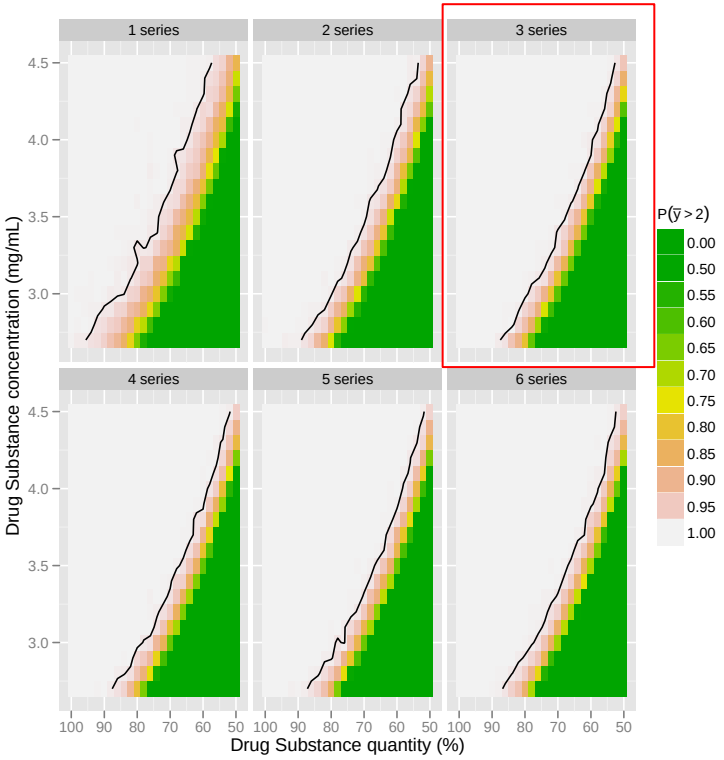
  - Again, it is assumed that the computed total variability is appropriate for the range of concentration of interest

New results (process DS - dilution decision)



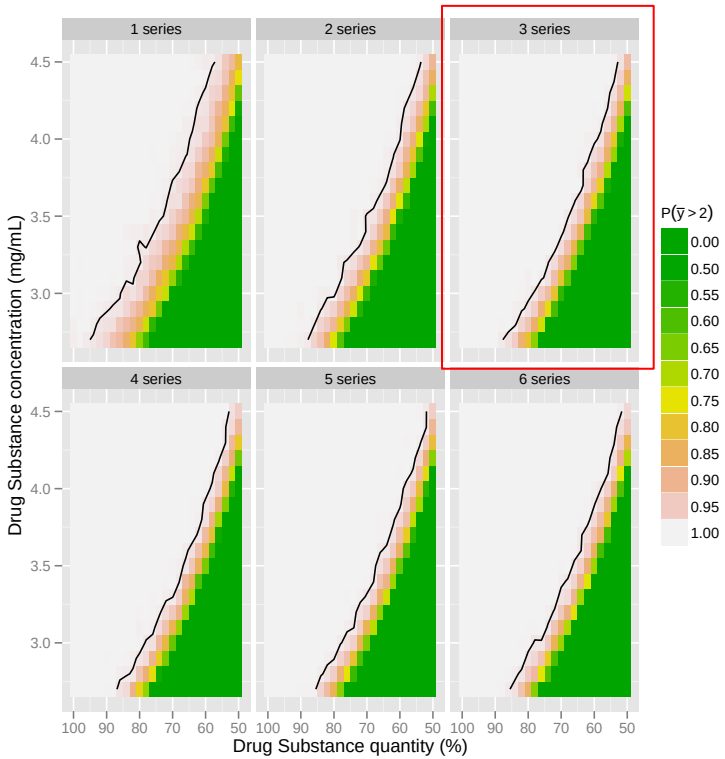


Without parameter uncertainties



Safe dilution easier to achieve

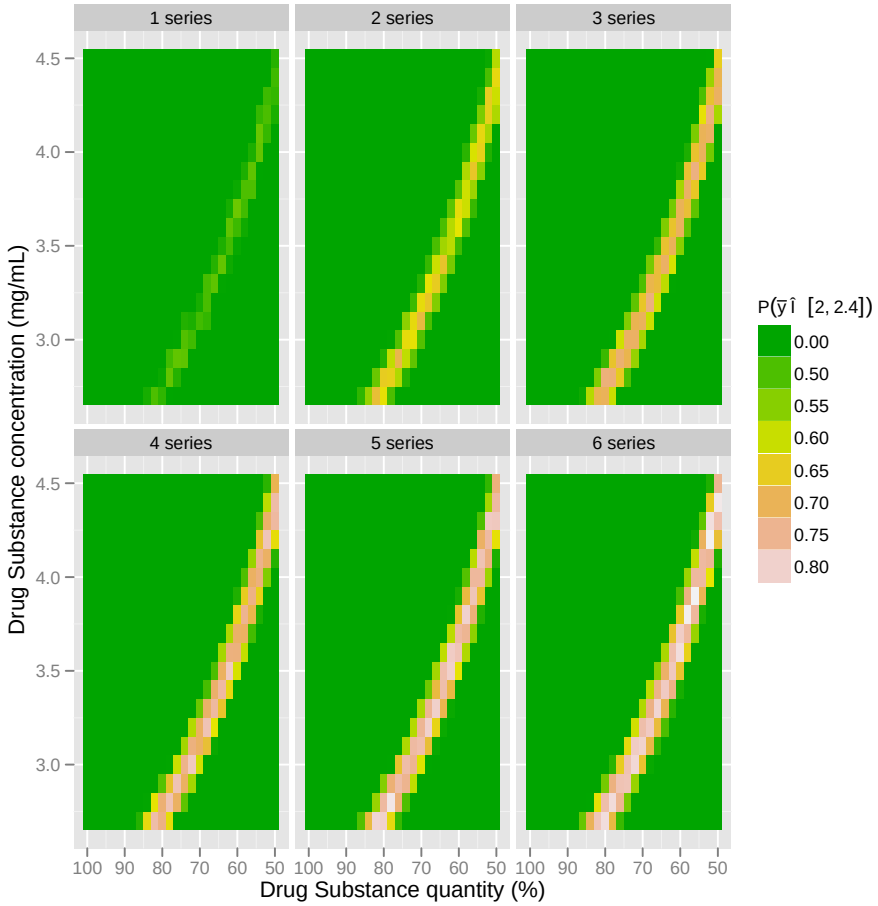
With parameter uncertainties



This time, **no noticeable difference** between the simulations using **maximum likelihood estimators** and the ones using complete **posterior parameter uncertainties**
→ due to the higher amount of data, parameter uncertainties are very limited

New results (process DS - dilution decision)

- Secondary objective (reportable drug product result in [2, 2.4] mg/mL)



In this case, quality level never exceed 80%

Furthermore, the small “high quality” area gives low confidence in robustness against dilution error

Possibility to improve reportable result precision using 7, 8, 9, etc. series for both Bulk and Drug products, but costs increase as well

Now, the management can take a decision knowing the risks

Conclusion

- Effective Design Space is the tool to optimize a process/assay while concurrently assess its robustness
- Design Space allows providing guarantee that future runs will be on specifications
- It may be used even when available data are not perfect
  - To provide risk-based results
  - To allow efficient and knowingly decision
- The industrial impact of such an approach is important
  - ➔ More companies are looking for this type of methodologies

Arlenda
Modeling Your Drug Development

bruno@arlenda.com

THANK YOU

- Arlenda is recruiting statisticians with knowledge in Bayesian statistics
  - In Europe in Brussel's area
  - In the USA, in Philadelphia area.

Contact Bruno.boulanger@arlenda.com