

Optimal design of experiments (DoE) for the modelling of food processes

D. Goujot^{1,2}, X. Meyer³, F. Courtois^{2,1}

¹INRA, UMR1145 Ingénierie Procédés Aliments, Massy, France

²AgroParisTech, UMR1145 Ingénierie Procédés Aliments, Massy, France

³Université de Toulouse, Laboratoire de Génie Chimique, UMR CNRS/INPT/UPS
5503, Toulouse, France

May 17, 2011

je presente ma thematique de recherche en collaboration avec FC et avec Xuan du IGC de Toulouse. Il s'agit de stratégie optimale pour planifier des expériences sur pilote process en conditions variables, connu en anglais comme design of experiments.

je vais d'abord présenter le contexte et l'objectif d'utilisation de cette stratégie, ensuite des exemples de conditions variables et de leur efficacité sur l'identification, puis faire un bilan rapide de ce qui est nécessaire pour pouvoir appliquer cette stratégie, et conclure sur ma contribution.

Context: Simulation for another product

- We have a simulation model for corn drying
- Objective: use this model for **rice** drying
- Hence we need to update model parameters
 - C_p
 - geometry
 - $D_w(X, T)$
 - $h(T_{air})$
 - inner structure
- The model has nonlinear ODEs and is non-linear in parameters: DoE is difficult.

Objective: parameter determination

Two types of parameters to adjust:

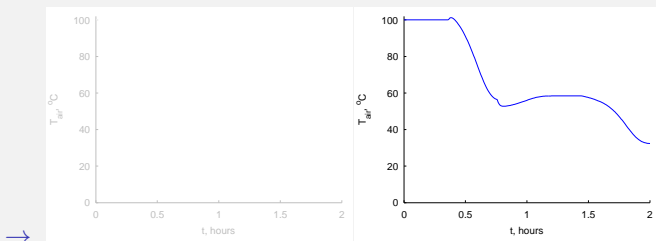
- Will be measured off-site:
 - C_p
 - ...
- Needs costly experiments effort:
 - D_w , as a function of X, T
 - ...

We need to spare on-site experiments.

Intuitive versus optimized strategy

Two strategies are available to choose T_{air} and HR_{air} .

- Intuitive strategy: two-factor three-level grid of 9 parallel runs:
 - $T_{air} = 50^{\circ}\text{C}, 70^{\circ}\text{C}, 90^{\circ}\text{C}$
- Model-based optimal strategy with **varying** optimized drying conditions:



The best strategy is the one with minimal effort for 20% confidence intervals.

2011-05-17

Optimal design of experiments (DoE) for the modelling of food processes

└ Objective, material and methods

└ Intuitive versus optimized strategy

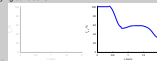
Intuitive versus optimized strategy

Two strategies are available to choose T_{in} and HR_{in} .

- o Intuitive strategy: two-factor three-level grid of 9 parallel runs:

→ T_{in} = 50°C, 70°C, 90°C

- o Model-based optimal strategy with **varying** optimized drying conditions:



The best strategy is the one with minimal effort for 20% confidence intervals.

I only show the prescribed inlet air temperature for the sake of simplicity

Intuitive strategy

Two-factor, three-level grid of 9 experiments of 2 hours each:

HR_{air}	50°C	70°C	90°C
0%	2h	2h	2h
20%	2h	2h	2h
40%	2h	2h	2h

Transfert coefficient identified so far: $h(100^{\circ}\text{C}) = 100 \pm 100$
 101 ± 99 100 ± 100 125 ± 5 14.5 ± 0.5 10.8 ± 0.5 14.5 ± 0.6
 12.1 ± 0.4 13.0 ± 0.5 $\text{W.K}^{-1}.\text{m}^{-2}$.

Four other parameters are also identified.

Optimal design of experiments (DoE) for the modelling of food processes

└ Intuitive experiment conditions of parameter identification

└ Intuitive strategy

Two-factor, three-level grid of 9 experiments of 2 hours each:

RH ₀	50°C	70°C	90°C
20%	2h	2h	2h
40%	2h	2h	2h

transfer coefficient identified so far: $h(100^\circ\text{C}) = 100 \pm 100$
 125 ± 5 14.5 ± 0.5 10.8 ± 0.5 14.5 ± 0.6
 12.1 ± 0.4 13.9 ± 0.5 $\text{W.K}^{-1}.\text{m}^{-2}$.

Four other parameters are also identified.

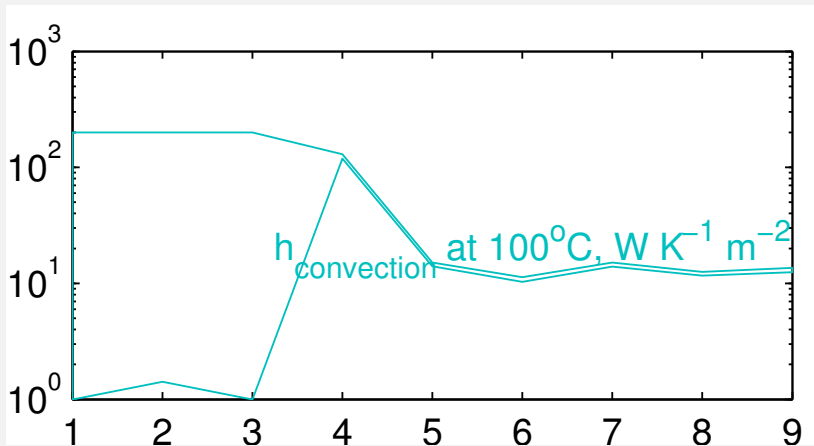
The intuitive strategy is based on 9 experiments and 18 hours of experiments. It is a two

The parameter may be identified from the drying kinetics of the nine experiment.

We show here, for simplicity, only the first experiment, which is the minimal of diffusivity of water.

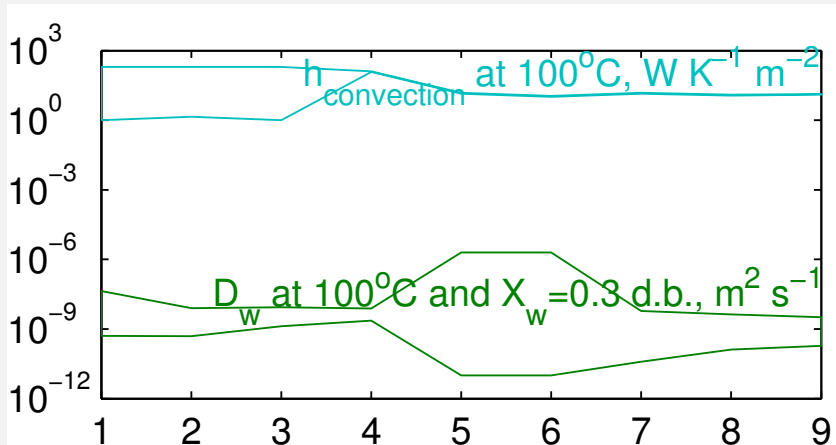
Dispo sur demande:

Les valeurs de D pour moins de 8 experiences sont maintenant "bonnes" sauf pour $i=8$, et eventuellement $i=9$ qui a un mauvais θ_0 ; et je passe aux valeurs de $h(T_{\max})$ qui évolue plus correctement.



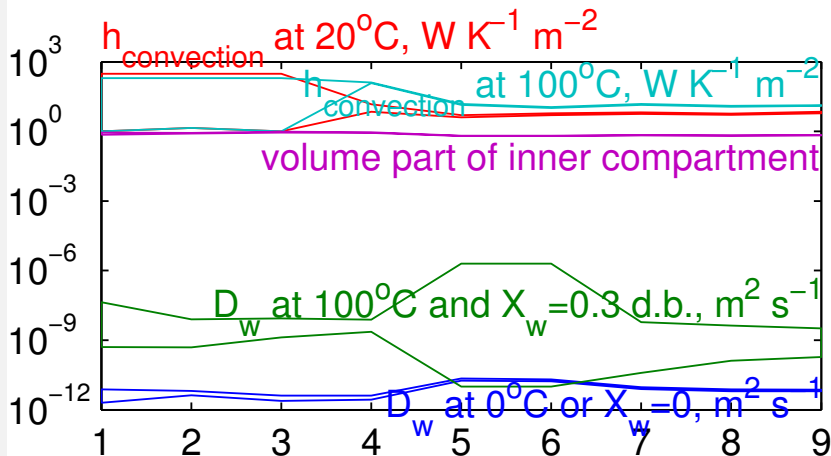
Parameter confidence gabarit.

Confidence for parameter $h(100^\circ\text{C})$ is good after 5 experiments.



Parameter confidence gabarit. One may enlarge, but not all at the time.

Confidence for parameter $D_w(100^\circ\text{C}, 0.3 \text{ d.b.})$ is good after 8 experiments.

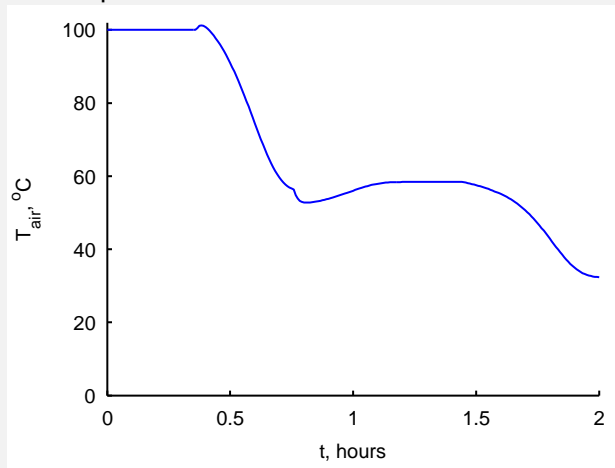


Parameter confidence gabarit. One may enlarge, but not all at the time.

Confidence for other three parameter is good after 8 experiments.

Optimal "A-optimal" strategy

One experiment of 2 hours:

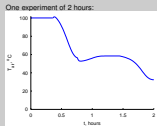


$h(100^{\circ}\text{C}) = 15 \pm 9 \text{ W.K}^{-1}.\text{m}^{-2}$. HR_{air} is optimized similarly.

Optimal design of experiments (DoE) for the modelling of food processes

- Dynamic experiment condition and optimal parameter identification

- Optimal "A-optimal" strategy



The optimal strategy results in non-constant air drying conditions. The optimized temperature profile for the first experiment is shown here.

The result of identified

Dispo sur demande:

$$D_w(0, T_{min}) = 20 \pm 5 \mu\text{m}^2/\text{s}$$

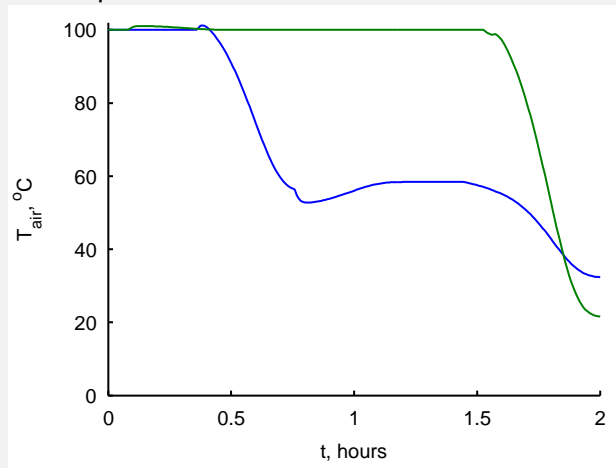
$$D_w(X_0, T_{max}) = 100 \pm \infty (\mu\text{m})^2/\text{s}$$

$$3 \text{ W/Km}^2 < h(T_{min}) < 20 \text{ W/Km}^2$$

volume fraction of inner compartment : 0.55 ± 0.02

Efficient "A-optimal" strategy

Two experiments of 2 hours:

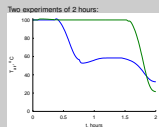


$$h(100^{\circ}\text{C}) = 100 \pm 100 \text{ W.K}^{-1}.\text{m}^{-2}.$$

Optimal design of experiments (DoE) for the modelling of food processes

- Dynamic experiment condition and optimal parameter identification

- Efficient "A-optimal" strategy



$h(100^\circ\text{C}) = 100 \pm 100 \text{ W.K}^{-1}.\text{m}^{-2}$.

Dispo sur demande:

$$D_w(0, T_{min}) = 7.66 \pm 0.9 \mu\text{m}^2/\text{s}$$

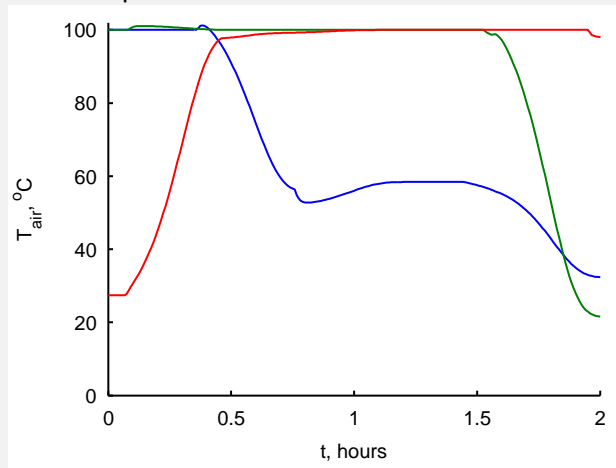
$$130 < D_w(X_0, T_{max}) < 5700 \mu\text{m}^2/\text{s}$$

$$6 \text{ W/Km}^2 < h(T_{min}) < 300 \text{ W/Km}^2$$

volume fraction of inner compartment : 0.68 ± 0.03

Efficient "A-optimal" strategy

Three experiments of 2 hours:

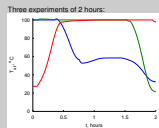


$$h(100^{\circ}C) = 11.5 \pm 0.8 \text{ W.K}^{-1}.\text{m}^{-2}.$$

Optimal design of experiments (DoE) for the modelling of food processes

- Dynamic experiment condition and optimal parameter identification

- Efficient "A-optimal" strategy



Three experiments of 2 hours:
 $h(100^\circ\text{C}) = 11.5 \pm 0.8 \text{ W.K}^{-1}.\text{m}^{-2}$.

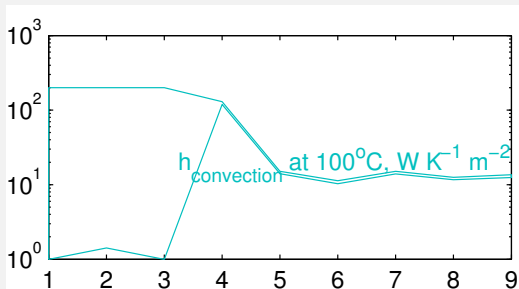
Dispo sur demande:

$$D_w(0, T_{min}) = 6.6 \pm 0.9 \mu\text{m}^2/\text{s}$$

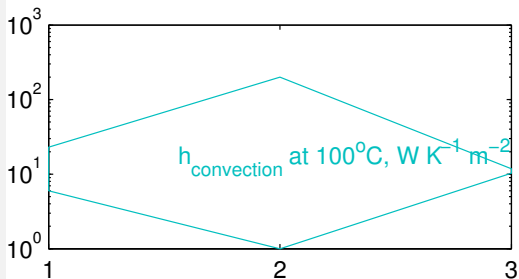
$$137 < D_w(X_0, T_{max}) < 9700 (\mu\text{m})^2/\text{s}$$

$$h(T_{min}) = 14.1 \pm 0.9 \text{ W/Km}^2$$

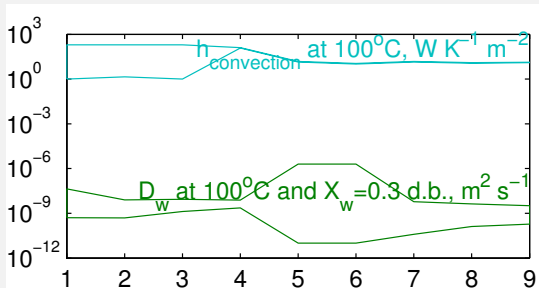
volume fraction of inner compartment : 0.57 ± 0.01



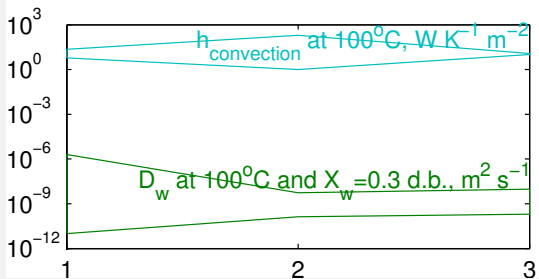
Parameter confidence gabarit for **intuitive** experiment conditions.



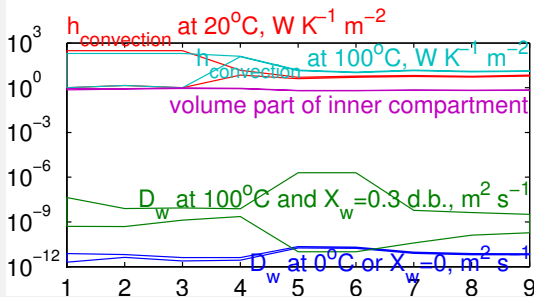
For **optimal** experiment conditions. Convergence is not clear.



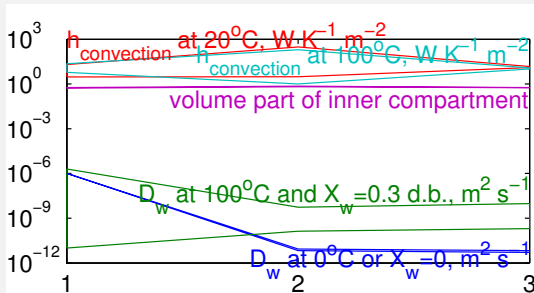
Parameter confidence gabarit for **intuitive** experiment conditions.



Parameter confidence gabarit for **optimal** experiment conditions. Optimized = less time, less used product.



All parameters are given.



Conclusion remain the same: less time, less used product.

Framework of the Optimized strategy

I implemented a MATLAB Toolbox, requiring:

- Simulation model for the product
 - At least initial guess for parameters
- Pilot plant:
 - transitory regimes
 - of limited inertia
 - of known inertia
 - online instrumentation
 - minimal uncertainties on mesurment
 - towards structural identification
- High-demanding in computation

These are requirements common to all model-based optimal DoE.

└ Requirements

└ Framework of the Optimized strategy

I implemented a MATLAB Toolbox, requiring:

- Simulation model for the product
 - At least initial guess for parameters
- Pilot plant:
 - transitory regimes
 - of limited inertia
 - of known inertia
 - online instrumentation
 - minimal uncertainties on measurement
 - towards structural identification
- High-demanding in computation

These are requirements common to all model-based optimal DoE.

Ma contribution dans la Matlab Toolbox:
utilisation de la reparamétrisation de sorte à rendre la planif possible
pour des modèles sans solution analytique avec optimiseurs de
Matlab

This Matlab toolbox is validated over 3 other applications, 2 of which being taken from literature:

- Experimentation on rice drying.
- Reproduction of two papers (pharmacokinetic results from Pronzato 2008, fermentation results from van Derlinden 2008.)

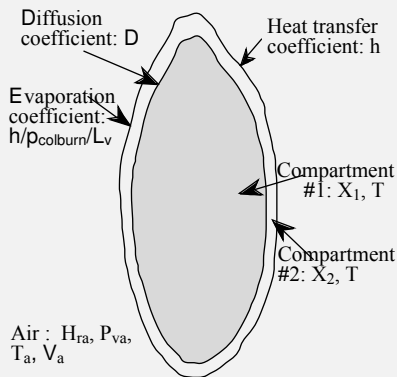
Results

- Available as a MATLAB Toolbox.
- Double-checked DoE of van Impe (2008, E. Coli) et Pronzato (2008, pharmacology).
- Experimental validation on application of rice drying.
- Versatility: tested on a wide range of applications.

Thank you

Thank you for your attention.

Simulation model



$$\frac{dX_1}{dt} = \frac{D_w S_{\text{spec}}}{l_1} (X_2 - X_1) \quad (1)$$

$$\frac{dX_2}{dt} = \frac{D_w S_{\text{spec}} \tau_1}{l_1 \tau_2} (X_1 - X_2) + \frac{h S_{\text{spec}}}{c_{\text{colburn}} L_v \rho_{dm} \tau_2} (P_{va} - P_{vp}) \quad (2)$$

$$\frac{dT}{dt} = h S_{\text{spec}} \frac{T_a - T + \frac{(P_{va} - P_{vp})}{c_{\text{colburn}}}}{\rho_{dm} (c_{pdm} + \bar{X} c_{pw})} \quad (3)$$

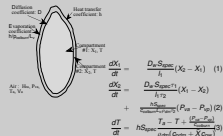
2011-05-17

Optimal design of experiments (DoE) for the modelling of food processes

└ Questions

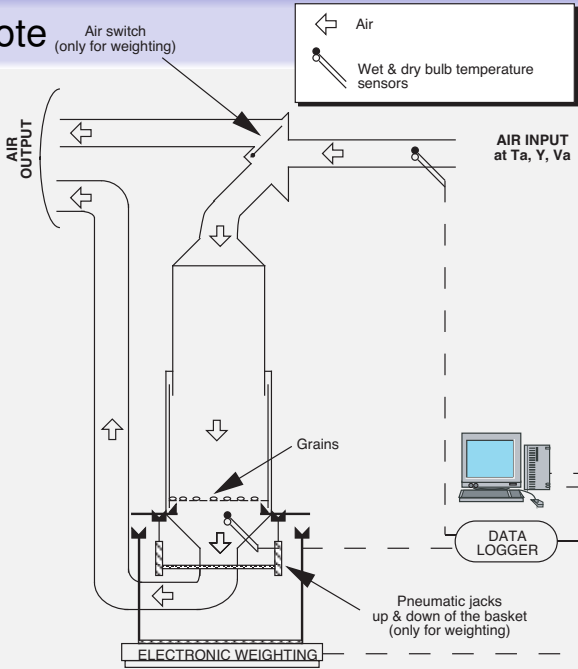
└ Simulation model

Simulation model



TODO mettre les parametres en rouge avec le usepackagecolor qui est implicite dans beamer.

Pilote



Pilote rodé,
 H_{Ra} et T_a
commandables.
Modèle
intéressant
rodé,
mesures
fiables

$$M_{\text{Jacobian}} = \begin{pmatrix} \frac{d\bar{X}}{d\theta_1}(t_{\tilde{\varphi}_1,1}), & \frac{d\bar{X}}{d\theta_2}(t_{\tilde{\varphi}_1,1}), & \dots & \frac{d\bar{X}}{d\theta_5}(t_{\tilde{\varphi}_1,1}), \\ \frac{d\bar{X}}{d\theta_1}(t_{\tilde{\varphi}_1,2}), & \frac{d\bar{X}}{d\theta_2}(t_{\tilde{\varphi}_1,2}), & & \frac{d\bar{X}}{d\theta_5}(t_{\tilde{\varphi}_1,2}), \\ \vdots, & \vdots, & & \vdots, \\ \frac{d\bar{X}}{d\theta_1}(t_{\tilde{\varphi}_1,N_{\tilde{\varphi}_1}}), & \frac{d\bar{X}}{d\theta_2}(t_{\tilde{\varphi}_1,N_{\tilde{\varphi}_1}}), & \dots & \frac{d\bar{X}}{d\theta_5}(t_{\tilde{\varphi}_1,N_{\tilde{\varphi}_1}}), \\ \frac{d\bar{X}}{d\theta_1}(t_{\tilde{\varphi}_2,1}), & \frac{d\bar{X}}{d\theta_2}(t_{\tilde{\varphi}_2,1}), & \dots & \frac{d\bar{X}}{d\theta_5}(t_{\tilde{\varphi}_2,1}), \\ \frac{d\bar{X}}{d\theta_1}(t_{\tilde{\varphi}_2,2}), & \frac{d\bar{X}}{d\theta_2}(t_{\tilde{\varphi}_2,2}), & & \frac{d\bar{X}}{d\theta_5}(t_{\tilde{\varphi}_2,2}), \\ \vdots, & \vdots, & \dots & \vdots, \\ \frac{d\bar{X}}{d\theta_1}(t_{\tilde{\varphi}_2,N_{\tilde{\varphi}_2}}), & \frac{d\bar{X}}{d\theta_2}(t_{\tilde{\varphi}_2,N_{\tilde{\varphi}_2}}), & \dots & \frac{d\bar{X}}{d\theta_5}(t_{\tilde{\varphi}_2,N_{\tilde{\varphi}_2}}), \\ \frac{d\bar{X}}{d\theta_1}(t_{\varphi_3,1}), & \frac{d\bar{X}}{d\theta_2}(t_{\varphi_3,1}), & \dots & \frac{d\bar{X}}{d\theta_5}(t_{\varphi_3,1}), \\ \vdots, & \vdots, & & \vdots, \\ \frac{d\bar{X}}{d\theta_1}(t_{\varphi_3,N_{\varphi_3}}), & \frac{d\bar{X}}{d\theta_2}(t_{\varphi_3,N_{\varphi_3}}), & \dots & \frac{d\bar{X}}{d\theta_5}(t_{\varphi_3,N_{\varphi_3}}) \end{pmatrix}$$

Estimation

Maximum de vraisemblance:

$$\theta = \mathit{arg} \min_{\phi} \sum_i \left(X_i - \tilde{X}_i \right)^2$$

i index of available measures (actually recorded in past experiments).

This is for linear model, for measurements with identical normal uncorrelated measure noise.

Model-based Optimal design of experiment: MBDOE

A-optimality = minimization of trace of Fisher information matrix:

$$\phi = \mathit{arg} \min_{\phi} \mathit{trace} \left(M_{\mathit{Jacobian}}^T \times M_{\mathit{Jacobian}} \right)^{-1}$$

In the line i column j , derivative of the i -th available measures (actually recorded in past experiments, or currently planned)
This is for linear model, for measurements with identical normal uncorrelated measure noise.

Transparents disponibles faits le 22/6/2010

Rappels sur l'approche classique

- Analyse des 16 plans à deux facteurs dans Drying 2009, 10-11.
- Deux facteurs, trois niveaux: très représentatif.
- $3^2 = 9$ expériences, c'est beaucoup.
- Ne prend pas le modèle en compte.

Les références faisant mieux sont rares.

Problématique: Maximiser l'information expérimentale

- On dispose d'un modèle rodé.
- On dispose d'un pilote expérimental instrumenté et commandable.
- Chaque expérience donne une multitude de mesures peu incertaines.
- On souhaite identifier et valider un modèle.
 - Coût expérimental élevé.
 - Grand nombre de paramètres.

Limites des outils existants

- Applicables à des modèles simples (linéaires en les paramètres et polynomiaux en les variables).
- Couplent deux sources importantes de perte de précision:
 - Intégration numérique des modèles de systèmes dynamiques non-linéaires.
 - Sensibilité approximées par différences finies.
- Logiciels essentiellement pour les modèles polynomiaux.
- Logiciels ne permettant aucune liberté ("customisation").

Pourquoi autant d'incertitude ?

Sont pris en compte à la fin de toutes les expériences:

- Phénoménologie
-

Injecter plus d'information a priori pour gagner plus d'information a posteriori ?

Plans expérimentaux D- ou A-optimaux: l'exception

- Les D- et A-optimalité combinent modèle, simulation et optimisation.
- Les références faisant des plans D-optimaux ou A-optimaux sont rares.
- Elles concernent des cas simples, avec solution analytique: van Impe (2008, fermentation) et Pronzato (2008, pharmacologie).
- Leur appropriation est un objectif de Réactial.

PLUS d'information a priori donne PLUS d'information a posteriori

- Utilise au mieux l'aspect commandable du pilote.
- Connaissance de l'inertie du pilote.
- Connaissance du Modèle.
- Connaissance a priori des paramètres par la littérature.
- Convergence pas à pas.
- Planifier l'expérience pour affiner ces informations.