

## COMPARAISON DE TECHNIQUES DE VALIDATION DYNAMIQUE DE DONNEES

---

G.HEYEN(1), C. GERKENS(2), C. ULLRICH(3), M. MATEUS(4)

- ◆ (1) Laboratoire d'Analyse et de Synthèse des Systèmes Chimiques  
Université de Liège, Sart-Tilman, B6A, B-4000 Liège, Belgium  
email : [G.Heyen@ulg.ac.be](mailto:G.Heyen@ulg.ac.be)
  - ◆ (2) Laboratoire d'Analyse et de Synthèse des Systèmes Chimiques  
Université de Liège, Sart-Tilman, B6A, B-4000 Liège, Belgium  
email : [C.Gerkens@ulg.ac.be](mailto:C.Gerkens@ulg.ac.be)
  - ◆ (3) Laboratoire d'Analyse et de Synthèse des Systèmes Chimiques  
Université de Liège, Sart-Tilman, B6A, B-4000 Liège, Belgium  
email : [cullrich@student.ulg.ac.be](mailto:cullrich@student.ulg.ac.be)
  - ◆ (4) Belsim s.a., Rue Georges Berotte 29A  
B-4470 Saint-Georges-sur-Meuse, Belgium  
e-mail : [miguel.mateus@belsim.com](mailto:miguel.mateus@belsim.com)
- 

**Abstract.** Three methods of dynamic data validation will be studied : the extended Kalman filter and two moving horizon techniques : a Runge-Kutta integration method and a orthogonal collocation discretisation method. Two studied cases will allow to compare those three techniques : two connected tanks and Liebman's reactor.

**Résumé.** Trois méthodes de validation dynamique de données seront étudiées : le filtre de Kalman étendu, et deux techniques de fenêtre mobile : une méthode d'intégration par Runge-Kutta et une méthode de discrétisation au moyen de collocations orthogonales. Deux exemples simples permettront de comparer ces trois techniques : deux cuves connectés et le réacteur de Liebman.

**Mots Clés.** Validation dynamique de données, collocation orthogonale, fenêtre mobile, filtre de Kalman

### INTRODUCTION

Le contrôle efficace des procédés est un problème clé concernant la bonne conduite des usines chimiques. En effet, toute mesure est entachée d'erreurs. Un bon contrôle ne sera dès lors réalisable que si un réseau de capteurs adéquat est installé et si les erreurs de mesure peuvent être détectées et corrigées. Les méthodes de validation de données permettent grâce aux redondances des mesures et aux lois de conservation de réduire l'incertitude des mesures et de calculer toutes les variables non mesurées du système (telles que l'efficacité de turbines ou de compresseur, la conversion des réactifs ...). Ces techniques sont utilisées couramment de nos jours pour les procédés fonctionnant à l'état stationnaire alors que les méthodes de validation des procédés fonctionnant en dynamique sont toujours en cours d'élaboration.

Dans les procédés stationnaires, seules les redondances spatiales peuvent être exploitées, alors que les redondances temporelles seront également utilisées pour la validation de données dynamique. Plusieurs méthodes de validation dynamique de données existent, mais aucune d'entre elles n'a pu à l'heure actuelle démontrer sa supériorité sur les autres pour tous les procédés dynamiques. Les procédés chimiques impliquent de nombreuses équations non linéaires, ce qui rend encore plus ardue la réconciliation des données.

Dans cette étude, trois méthodes de validation dynamiques seront étudiées :

- une méthode de filtrage : le filtre de Kalman étendu ;
- deux méthodes de fenêtres mobiles : ces deux méthodes implique l'intégration des ODE soit en utilisant l'intégrateur de Runge-Kutta, soit par discrétisation au moyen de collocation orthogonales. L'optimisation des conditions initiales et des inputs est réalisée au moyen de différents algorithmes NLP.

Deux études de cas permettront de comparer les résultats obtenus avec les différentes techniques : deux cuves se déversant l'une dans l'autre et le réacteur de Liebman.

## FILTRE DE KALMAN ETENDU<sup>1 2</sup>

Le filtre de Kalman étendu à pour objectif l'estimation des variables d'état  $x$  de procédés gouvernés par des équations différentielles stochastiques non linéaires :

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) \approx \tilde{x}_k + A_k (x_{k-1} - \hat{x}_{k-1}) + W_k w_{k-1}$$

Et ayant pour mesures :

$$z_k = h(x_k, v_k) \approx \tilde{z}_k + H (x_k - \tilde{x}_k) + V v_k$$

Dans ces équations,

- $x_k$  et  $z_k$  sont les variables d'état et les mesures au temps  $k$  ;
- $f$  est une fonction permettant de relier les variables d'état  $x$ , les variables de contrôles  $u$  et le bruit du procédé  $w$  au temps  $k-1$  au variables d'état au temps  $k$  ;
- $h$  est une fonction permettant de relier les variables d'état  $x$ , et le bruit des mesures  $v$  au temps  $k$  aux valeurs des mesures au temps  $k$  ;
- Les bruits de procédé  $w$  et des mesures  $v$  sont des bruits de moyenne nulle. En pratique, les variables d'état et les mesures sont approximés en considérant ces deux vecteurs nuls ( $\tilde{x}_k$  et  $\tilde{z}_k$ ) ;
- $\hat{x}_k$  est l'estimation à posteriori des variables d'état au temps  $k$  ;
- $A_k$  est la matrice jacobienne des dérivées partielles de  $f$  par rapport à  $x$  au temps  $k$  ;
- $W_k$  est la matrice jacobienne des dérivées partielles de  $f$  par rapport à  $w$  au temps  $k$  ;
- $H_k$  est la matrice jacobienne des dérivées partielles de  $h$  par rapport à  $x$  au temps  $k$  ;
- $V_k$  est la matrice jacobienne des dérivées partielles de  $h$  par rapport à  $v$  au temps  $k$ .

Le filtre de Kalman procède en deux étapes pour chaque estimation :

- La prédiction ou mise à jour des temps : lors de cette étape, les variables d'état et la matrice de covariance des erreurs de prédiction sont estimées de la manière suivante :

$$\hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1}, 0)$$

$$P_k^- = A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T$$

avec  $Q$  la matrice de covariance du bruit du procédé.

- La correction ou mise à jour des mesures : cette étape commence par le calcul de la matrice des gains qui a pour objectif la minimisation de la matrice de covariance de erreurs à posteriori :

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1}$$

avec  $R$  la matrice de covariance du bruit des mesures.

Les variables d'état ainsi que la matrice de covariance des erreurs d'estimation sont ensuite évaluées :

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, 0))$$

$$P_k = (I - K_k H_k) P_k^-$$

Une analyse de la matrice des gains montre que si les erreurs de mesures sont faibles ( $R$  est quasiment nulle), la pondération des résidus sera plus importante, tandis que si le modèle est presque parfait, la pondération de l'estimation des variables d'état sera plus importante.

Un des désavantages de la méthode du filtre de Kalman est qu'elle considère les variables de contrôle comme constantes : elle ne permet donc ni de les corriger, ni de prendre en compte leurs incertitudes. Un second problème de cette méthode est qu'elle nécessite le calcul de dérivées qui ne sont pas toujours faciles à obtenir. Pour palier à cet inconvénient, M. Noorgard et ses coauteurs<sup>3</sup> ont créé une méthode de filtrage basée sur les différences centrales du second ordre. Nous ne la détaillerons pas dans cet article.

## METHODES DE FENETRES MOBILES

### Intégration par Runge-Kutta

Dans cette méthode de fenêtre mobile, le problème de validation dynamique de données est exprimé sous la forme d'un problème général de programmation non linéaire. Le modèle dynamique est intégré sur la fenêtre de validation par la méthode d'intégration de Runge-Kutta du quatrième ordre. La fonction objectif est calculée comme suit sur l'horizon de validation :

$$F_{obj} = \sum_{\text{Horizon de validation}} \sum_{\text{Variables mesurées}} \frac{(x_{\text{mesuré } j,i} - x_{\text{estimé } j,i})^2}{\sigma_j^2} + \sum_{\substack{\text{Variables état} \\ \text{Variables indépendantes}}} \frac{(x_{Cl j} - x_{\text{estimé } j,0})^2}{\text{relaxation}_j^2}$$

Dans cette équation,

- $x_{\text{mesuré } j,i}$  est la valeur de la variable j mesurée au temps i ;
- $x_{\text{estimé } j,i}$  est la valeur de la variable j estimée au temps i ;
- $\sigma_j$  est l'écart-type de la variable j ;
- $x_{Cl j}$  est la valeur de la condition initiale (temps 0) sur la variable j. Cette valeur est constante au cours de l'optimisation ;
- $\text{relaxation}_j$  est le facteur de relaxation. Le terme de relaxation de la fonction objectif a pour objectif de s'assurer que les valeurs des variables d'état et indépendantes en début d'intervalle ne sont pas trop différentes de celles obtenues pour ce même temps lors de la validation précédente.

La fonction objectif est optimisée en utilisant une méthode d'optimisation (Davidon<sup>4</sup> ou SQP<sup>5</sup>). La taille optimale de l'horizon de validation ainsi que le déplacement optimum de la fenêtre sont des paramètres qui dépendent du problème étudié et de la fréquence des mesures. Ils doivent être choisis de manière à assurer une redondance temporelle suffisante en veillant à garder un problème à optimiser de taille convenable.

Cette méthode présente l'inconvénient de ne pas permettre dans l'état actuel des connaissances de réaliser le calcul des variances des variables à posteriori. Pour utiliser efficacement la méthode, il faudrait trouver un moyen de les estimer.

### Discrétisation par collocations orthogonales<sup>6,7</sup>

Cette méthode permet de représenter les variables d'état et leurs dérivées au moyen de polynômes et ainsi de transformer les équations différentielles en équations algébriques. Etant donné que dans le cas des équations algébriques, il existe des formules qui permettent de calculer les variances à posteriori. La méthode s'articule en cinq points :

- Classer les variables suivant qu'il s'agit de variables d'état, de variables dépendantes, de variables indépendantes, de variables mesurées ou non mesurées ;
- Ecrire les équations différentielles pour les variables d'état et les équations de liaison ;
- Réaliser les collocations orthogonales pour les équations d'état. Pour ce faire, les temps de collocations sont déterminés comme étant les zéros de polynômes de Jacobi. Pour illustrer cette méthode, considérons par exemple l'équation différentielle :

$$\frac{dx}{dt} = g x$$

En utilisant les interpolation de Lagrange, la dérivée de la variable x s'écrit :

$$\frac{dx}{dt} = \frac{dl_0}{dt} x_0 + \frac{dl_1}{dt} x_1 + \frac{dl_2}{dt} x_2$$

avec  $dl_i$ , les dérivées des facteurs d'interpolation de Lagrange.

Il faut alors résoudre le système d'équations linéaires suivant :

$$\frac{dl_{0,t_1}}{dt} x_0 + \frac{dl_{1,t_1}}{dt} x_1 + \frac{dl_{2,t_1}}{dt} x_2 = g x_1$$

$$\frac{dl_{0,t_2}}{dt} x_0 + \frac{dl_{1,t_2}}{dt} x_1 + \frac{dl_{2,t_2}}{dt} x_2 = g x_2$$

$$x_0 = x_{0, \text{conditions initiales}}$$

Ce système est résolu par la méthode de Broyden pour de petit système et par « soldog » pour des systèmes plus importants. Les variables d'état sont alors évaluées aux temps des mesures en utilisant les interpolation de Lagrange :

$$x_t = l_{0,t} x_0 + l_{1,t} x_1 + l_{2,t} x_2$$

- Estimer la fonction objectif définie comme dans le cas de la méthode d'intégration de Runge-Kutta ;
- Réaliser l'optimisation de la fonction objectif. Cette optimisation peut être obtenue, pour de petits problèmes, grâce à la méthode de Davidon sans direction de recherche. Pour des problèmes plus importants, la méthode « SQP » est utilisée.

## ETUDES DE CAS

### 2 cuves connectées

Dans ce premier cas, la dynamique de deux cuves se déversant linéairement l'une dans l'autre sera étudiée. Ce système est décrit par les équations suivantes :

$$\begin{aligned} \frac{dH_1}{dt} &= \frac{F_0}{A_1} - \frac{\beta_1 H_1}{A_1} \\ \frac{dH_2}{dt} &= \frac{\beta_1 H_1}{A_2} - \frac{\beta_2 H_2}{A_2} \\ F_1 &= \beta_1 H_1 \\ F_2 &= \beta_2 H_2 \end{aligned}$$

Dans ces équations, les paramètres ont été fixé comme suit :  $\beta_1 = 1$ ,  $\beta_2 = 2$ ,  $A_1 = 10$  et  $A_2 = 12$ .

Les résultats présentés ci-dessous ont été obtenus dans les conditions suivantes

- Le débit d'alimentation  $F_0$  vaut 50 au départ, il passe à 100 ; au temps 31 et à 60 ; au temps 81.
- Au temps  $t = 0$ ,  $H_1 = 58.358$ ,  $H_2 = 56.417$ ,  $F_1 = 59$ ,  $F_2 = 110$
- Les mesures sont réalisées toutes les secondes ;
- Durée totale de simulation : 110 s ;
- Un bruit gaussien de moyenne égale à 2% est ajouté aux valeurs simulées avant la validation ;
- Les variable d'état sont  $H_1$  et  $H_2$  et les variables dépendantes  $F_1$  et  $F_2$  ;
- La fenêtre de temps contient 12 intervalles de mesure et son déplacement est de deux intervalles ;

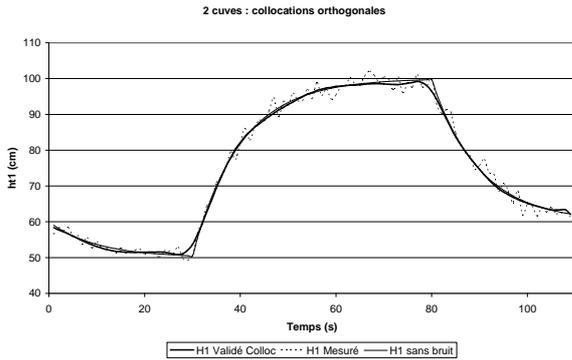


Figure 1

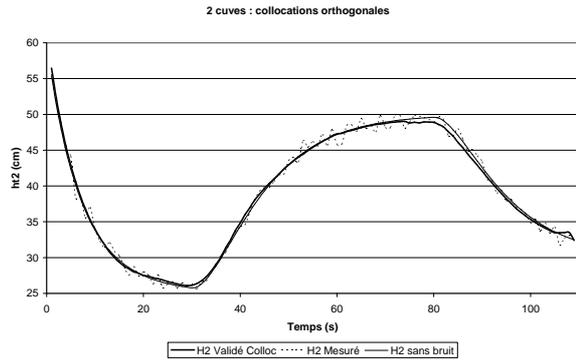


Figure 2

Les trois méthodes réduisent toutes plus ou moins bien l'incertitude des mesures. La méthode la moins efficace est celle de Runge-Kutta, vient ensuite la méthode du filtre de Kalman. La méthode des collocations est celle qui fournit les meilleurs résultats. Ces derniers sont présentés sur les figures 1 et 2 pour les deux variables d'état H1 et H2 respectivement.

### Réacteur de Lieberman

Dans cette seconde étude de cas, un réacteur adiabatique comprenant une réaction exothermique du premier ordre sera traité<sup>7</sup>. Les équations dynamiques de ce système s'écrivent de la manière suivante en utilisant une température  $T_r$  et une concentration  $A_r$  de référence :

$$\begin{aligned} \frac{dA}{dt} &= \frac{q}{V} (A_0 - A) - \alpha_d k A \\ \frac{dT}{dt} &= \frac{q}{V} (T_0 - T) - \alpha_d \frac{-\Delta H_r A_r}{\rho C_p T_r} k A - \frac{U A_R}{\rho C_p V} (T - T_c) \\ k &= k_0 \exp\left(\frac{-E_A}{T T_r}\right) \end{aligned}$$

Dans ces équations,

- $A$  est la concentration dans le réacteur ;

- $T$  est la température dans le réacteur ;
- $A_0$  est la concentration du flux d'alimentation et  $A_0 = 6.5$  ;
- $T_0$  est la température du flux d'alimentation et  $T_0 = 3.5$  ;
- $A_r$  est la concentration de référence et  $A_r = 10^{-6} \text{ mol/cm}^3$  ;
- $T_r$  est la température de référence et  $T_r = 100 \text{ K}$  ;
- $k$  est la constante de vitesse avec  $k_0 = 7.86 * 10^{12} \text{ s}^{-1}$  et  $E_A = 14090 \text{ K}$  ;

Les autres paramètres sont fixés comme suit :  $q = 10 \text{ cm}^3/\text{s}$ ,  $V = 1000 \text{ cm}^3$ ,  $\Delta H_r = -27000 \text{ cal/mol}$ ,  $\rho = 0.001 \text{ g/cm}^3$ ,  $C_p = 1 \text{ cal/mol K}$ ,  $U = 5 * 10^{-4} \text{ cal/cm}^2 \text{ s K}$ ,  $A_R = 10 \text{ cm}^2$ ,  $T_c = 3.4$ ,  $\alpha_d = 1$ .

Les résultats présentés ci-dessous ont été obtenus dans les conditions suivantes :

- Concentration à l'état stationnaire = 0.155934 ;
- Température à l'état stationnaire = 4.605732 ;
- Les mesures sont réalisées toutes les 0.025 s ;
- Au temps 0.75,  $A_0$  passe de 6.5 à 7.5 ;
- Durée totale de simulation : 25 s ;
- Un bruit gaussien de moyenne égale à 1% est ajouté aux valeurs simulées des concentrations et de 1K aux températures avant la validation.
- Les variables d'état sont  $A$  et  $T$  et les variables de contrôle  $A_0$ ,  $T_0$  et  $k$ .
- La fenêtre de temps contient 12 intervalles de mesure et son déplacement est de deux intervalles ;

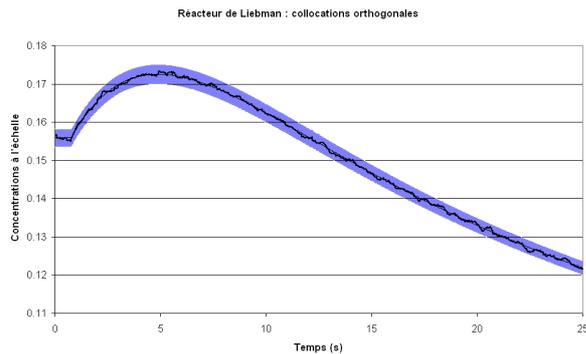


Figure 3

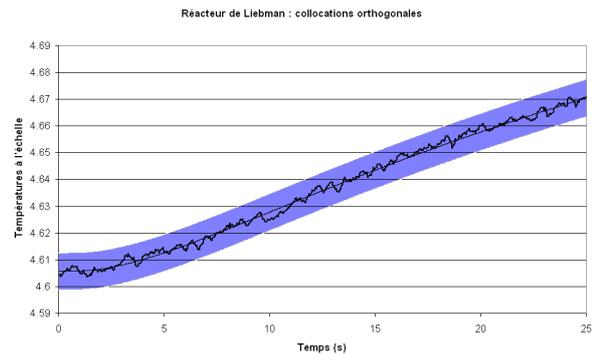


Figure 4

Les résultats sont similaires à ceux obtenus pour le premier exemple : les trois méthodes réduisent toutes plus ou moins bien l'incertitude des mesures. La méthode la moins efficace est celle de Runge-Kutta, vient ensuite la méthode du filtre de Kalman. La méthode des collocations est fournie les meilleurs résultats. Ces derniers sont présentés sur les figures 3 et 4 pour la concentration et la température dans le réacteur respectivement.

## CONCLUSIONS ET PERSPECTIVES FUTURES

Les trois méthodes de validation dynamique de données que nous avons étudiées permettent toutes trois de réduire les erreurs de mesure dans les cas étudiés. Cependant, dans les cas que nous avons traités, il apparaît que la méthode de discrétisation au moyen des collocations orthogonales fournit de meilleurs résultats que le filtre de Kalman étendu et la méthode utilisant l'intégrateur de Runge-Kutta. De plus, cette méthode permettra par la suite de calculer les variances des variables à posteriori et permet contrairement au filtre de Kalman de valider les variables de contrôle. Les autres développements futurs seront d'étendre la méthode à de plus grands systèmes et de créer un module permettant de vérifier l'observabilité du système avant d'en valider les mesures.

## REMERCIEMENTS

Les auteurs désirent remercier la Région Wallonne et le Fonds Social Européen qui ont financé cette recherche dans le cadre d'un programme « First Europe objectif 3 » ainsi que Ministère de la culture de

l'enseignement supérieur et de la recherche du Grand-Duché du Luxembourg qui a financé cette recherche dans le cadre d'une bourse formation-recherche.

#### REFERENCES

- [1] ROUSSEAUX-PEIGNEUX P., Filtre de Kalman hiérarchisé pour l'estimation d'état dynamique de grands systèmes électriques, Université de Liège, (1988).
- [2] WELCH G. et BISHOP G., Annual conference on computer graphics and interactive techniques, An introduction to Kalman Filter, ACM Press, Addison-Wesley, Los Angeles, Californie, USA, (2001).
- [3] NORGAARD M., POULSEN N. K. et RAVN O., New développements in state estimation for nonlinear systems, Automatica, Vol. 36, 1627-1638, (2000).
- [4] DAVIDON W. C., Optimally conditioned optimization algorithm without line search, Math. Programming, Vol. 9, 1-30, (1975).
- [5] KYRIAKOPOULOU D., Development and implementation of an interior point optimization algorithm for process engineering, Laboratoire d'Analyse et Synthèse des Systèmes Chimiques, Université de Liège, (1997).
- [6] VILLADSEN J. et MICHELSEN M. L., Solution of differential equation models by polynomial approximation, Prentice-Hall, Englewood Cliffs, New Jersey, (1978).
- [7] LIEBMAN M. J., EDGAR T. F. et LASDON L. S., Efficient data reconciliation and estimation for dynamic processes using nonlinear programming techniques, Computers & Chemical Engineering, Vol. 16, N° 10/11, 963-986, (1992).