

Notions de Modélisation et de Simulation

I-1- La modélisation

La modélisation est une représentation logique et mathématique du comportement d'un système réel dans un contexte donné. Le modèle est alors issu d'une série d'hypothèses concernant le fonctionnement du système exprimées à l'aide de relations mathématiques, logiques ou symboliques entre ses éléments. Ce qui produit un modèle dit analytique. Pour les systèmes trop complexes on peut utiliser un modèle descriptif qui détaille les différents événements et actions décrivant son comportement.

Suivant le système étudié, un modèle peut être statique (indépendant du temps), dynamique (dépendant du temps ou de l'espace), déterministe (dépend de facteurs connus), stochastique (dépend de facteurs aléatoires), discret, continu ou hybride (combinaison de plusieurs types).

Définir la modélisation en génie des procédés est une tâche délicate tant elle recouvre d'actions et d'objectifs divers. On pourrait néanmoins lui donner la définition synthétique suivante :

La modélisation est une démarche qui permet, à partir de faits expérimentaux, de construire un outil mathématique reliant les sorties d'un système à ses entrées.

On entend par système une structure physique prise de manière isolée; un cristalliseur, un réacteur chimique, un atelier, représentent autant d'exemples. Ses entrées sont les paramètres qui agissent sur son comportement. Elles peuvent être contrôlées ou non. Ses sorties sont les paramètres résultant des valeurs données aux entrées.

Pour illustrer cette notion, considérons un réacteur agité adiabatique fonctionnant en régime permanent dans lequel se produit une réaction chimique $A \rightarrow B$. Les variables d'entrée du système sont le débit de charge, la concentration de A, la température du réacteur, le volume du réacteur; les variables de sortie sont le taux de conversion de A en B ou bien les concentrations de chaque constituant et la température de sortie.

On distingue dans un procédé (figure 1.2) :

- le vecteur des entrées $\{u_1, u_2, \dots, u_k\}$;
- le vecteur des sorties $\{y_1, y_2, \dots, y_m\}$;
- le vecteur des perturbations $\{p_1, p_2, \dots, p_\ell\}$

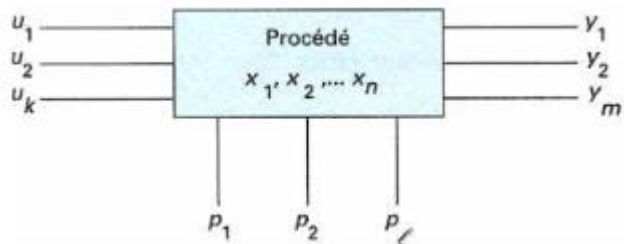


Figure 1.2

Ce dernier ensemble de valeurs (les perturbations) correspond aux paramètres influant sur le comportement statique ou dynamique du procédé sans que l'on puisse les maîtriser ou que l'on veuille les prendre en compte dans le modèle. Ce concept de perturbations est très important pour les problèmes de modélisation dynamique et/ou d'automatisation.

En fait, la description précédente restreint le problème aux bornes du système, c'est-à-dire aux entrées et sorties. Le plus souvent, une succession de phénomènes internes intervient dans la relation entre les entrées et les sorties. Par exemple, dans une colonne à distiller pour laquelle sont définis débit de reflux et puissance de rebouillage, les caractéristiques (débit et composition) du distillât et du résidu sont la conséquence des phénomènes de transfert de matière et d'équilibre intervenant sur chaque plateau.

Ces variables internes sont souvent appelées variables d'état du procédé (x_1, x_2, \dots, x_n) . Elles peuvent dépendre du temps (lorsqu'elles décrivent un procédé non stationnaire) ou non (lorsqu'elles représentent un procédé en régime permanent).

Très généralement, toute description mathématique d'une opération unitaire ou d'un procédé peut se ramener au système d'équations suivantes :

$$\begin{cases} f\left(\frac{dx(t)}{dt}, x(t), u(t), p(t), a(t), t\right) = 0 \\ g\left(\frac{dy(t)}{dt}, y(t), x(t)\right) = 0 \end{cases}$$

t représente le temps et $a(t)$ un vecteur de paramètres permettant d'ajuster le modèle à la réalité expérimentale. Les valeurs de ces paramètres ne sont pas a priori constantes au cours du temps. Un exemple typique est la décroissance de l'activité d'un catalyseur au cours du temps ou l'encrassement des échangeurs de chaleur.

Un tel ensemble d'équations n'est pas toujours accessible ou se révèle impossible à résoudre. Le problème de base lié à une démarche de modélisation est de faire des choix :

- relatifs à la structure de base du modèle (définition des entrées, des sorties et des perturbations) ;
- relatifs à la structure des équations.

Le choix des entrées d'un modèle est extrêmement arbitraire. En fait, ce sont les variables dont on est obligé de fixer les valeurs pour rendre le système soluble.

I-1-1- Modélisation de type boîte noire

Un modèle de type boîte noire se réduit à une description du procédé dénuée de toute signification physique. De nombreuses formes peuvent lui être données depuis la formulation mathématique pure reliant les entrées et les sorties jusqu'à une description qualitative comme les techniques modernes d'intelligence artificielle l'autorisent. Quel que soit le type d'approche adopté, la priorité est d'atteindre la meilleure qualité de représentation des réponses du système en fonction des entrées, ou plus exactement le rapport optimal qualité/complexité du modèle. En effet, il est préférable d'utiliser un modèle moins performant mais plus simple, de manière à avoir une garantie sur sa robustesse.

a) Approche mathématique

Deux problèmes sont à résoudre lorsque l'on cherche à créer des relations mathématiques entre entrées et sorties. Le premier consiste à sélectionner les variables influant le plus sur la sortie étudiée. Le second consiste à chercher la forme mathématique la plus appropriée.

Bien que le premier puisse être traité directement par identification, des outils statistiques peuvent aider, à partir de l'analyse des valeurs expérimentales, à déterminer les variables ayant le plus d'influence (composantes principales).

La recherche des relations est généralement empirique et se limite, le plus souvent, à l'emploi de relations polynomiales. Pour identifier le nombre optimal de paramètres (qui doit être le plus faible possible), des tests statistiques sur le résidu entre valeurs expérimentales et valeurs calculées sont nécessaires. Très généralement, on utilise le test de Student. Mais, pour pallier les inconvénients de ces lois polynomiales (comportements non physiques en dehors des plages de calage, valeurs non bornées lorsque les variables deviennent trop grandes), il est souvent intéressant d'utiliser d'autres fonctions mathématiques sélectionnées pour leurs caractéristiques géométriques. Citons la fonction sigmoïde définie par la relation :

$$\frac{\exp(ax) - 1}{\exp(ax) + 1}$$

'a' est le paramètre de calage. Lorsque celui-ci est grand, l'évolution de cette fonction est brutale et tend vers la fonction de Heaviside.

b) Approche par intelligence artificielle

Considérée par certains comme une technique d'intelligence artificielle et par d'autres comme une technique de modélisation de type boîte noire, la modélisation par réseaux neuronaux a pris une extension récente justifiée par les performances de cette approche. Ces réseaux sont composés d'un certain nombre de *nœuds* (ou neurones) qui sont répartis en couche et rassemblent des informations provenant des couches antérieures. Ces informations sont additionnées et transformées par la fonction de réponse du neurone.

La figure 1.3 est un exemple de représentation d'un système à trois entrées et deux sorties par un réseau neuronal à cinq neurones répartis en deux couches. Pour chacun des nœuds de ce réseau, la sortie est définie par : $W = f(\sum_{i=1}^N A_i U_i)$

Les A_i sont des paramètres de pondération (ou poids ou coefficients synaptiques) de chacune des entrées u_i . f est une fonction de transition choisie par l'utilisateur parmi toute une gamme de fonctions à seuil {fonction *échelon unité* de Heaviside, fonction linéaire à seuil, fonction sigmoïde...}. Un exemple de ce type de fonction est donné par la figure 1.4.

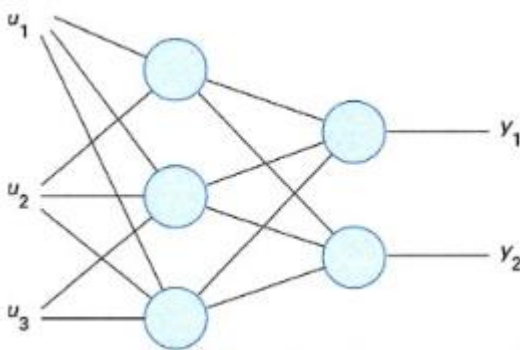


Figure 1.3

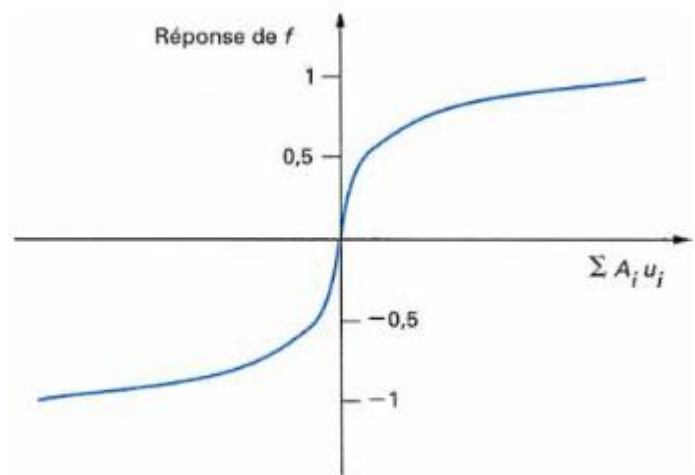


Figure 1.4

Une telle approche est extrêmement intéressante car, en entreprenant un faible effort en termes d'identification (essentiellement les poids A_i), des modèles non linéaires multi variables sont construits et conduisent à un excellent accord entre valeurs expérimentales et valeurs calculées pour des processus fortement non linéaires. On considère souvent que trois niveaux sont suffisants pour y parvenir. Toutefois, ces modèles restent des modèles de régression dont l'utilisation en extrapolation est périlleuse sans précaution. La représentation par réseau neuronal reste une description strictement algébrique des relations entrée/sortie.

D'autres outils sont disponibles comme les systèmes experts, la logique floue, la physique qualitative... dont la caractéristique de base est de fournir des réponses de type décisionnel à partir des informations fournies. Les principes de tous ces outils sont très proches. À partir de connaissances obtenues sur le procédé, des règles logiques reliant des faits et des événements sont établies ; lorsque cette relation est stricte, la technique est appelée système expert. Lorsque les conséquences ne le sont pas, elle est appelée logique floue. Lorsque la définition des événements n'est pas précise mais repérée par de simples niveaux de valeur, on aborde le principe de la physique qualitative. À partir de cette base de règles, un outil (moteur d'inférence pour les systèmes experts) explore toutes les possibilités et donne une solution. Ces techniques, bien adaptées pour identifier des décisions, le sont beaucoup moins lorsqu'une approche plus quantitative des interactions entrée/sortie est recherchée.

c) Cas particulier des systèmes dynamiques

Pour les procédés fonctionnant en régime dynamique, le problème se complique dans la mesure où la dimension temporelle s'ajoute aux autres variables. Toutefois, ce qui a été dit auparavant sur la recherche de structure et de dimension du modèle reste valable.

Il faut rappeler que l'utilisation des fonctions de transfert est intéressante. Les processus linéaires mono variables peuvent être représentés par des équations différentielles simples. Pour un système du premier ordre, cette équation s'écrit : $\frac{dY(t)}{dt} = \frac{Y(t)}{\tau} + b U(t)$ où b est une constante et τ est la constante de temps du procédé.

Pour des systèmes d'ordre supérieur, le nombre de dérivations de y par rapport au temps est augmenté d'autant. Pour éviter de manipuler des équations différentielles, celles-ci peuvent être avantageusement transformées en équations algébriques à l'aide de la transformation de Laplace (ou en z pour les systèmes échantillonnés) avec tous les avantages que cela apporte en termes de manipulation. Ces équations donnent aussi très rapidement accès à des principes de réglage de régulateurs.

Lorsque les processus mis en œuvre ne sont pas linéaires, ils peuvent être approchés de la même manière autour d'un point de fonctionnement. Les paramètres « τ », intervenant dans les équations, représentent alors les principales constantes de temps du procédé, mais il est difficile de remonter à leurs origines.

Très souvent pour des processus échantillonnés, les utilisateurs préfèrent employer des régressions liant la valeur actuelle d'une sortie aux valeurs précédentes et à l'entrée correspondante. Ce type de modèles, décrit par l'équation suivante, est souvent dénommé modèle ARMA :

$$Y[k \Delta t] = \sum_{i=1}^{l} a_i Y[(k-i)\Delta t] + \sum_{j=1}^{m} b_j U[(k-j-r)\Delta t]$$

Dans cette équation, Δt représente le pas de temps, k est l'indice du temps courant, l et m les *horizons de temps* influant sur le comportement du procédé. r représente le retard.

Cette modélisation a l'avantage de se réduire à une expression linéaire facilement exploitable. Seuls les paramètres a, b, l, m, r sont à identifier, ce qui impose une connaissance du passé.

d) Conclusions

Quelles que soient les techniques utilisées, les modèles de type *boîte noire* sont des modèles de régression, c'est-à-dire qu'ils ne représentent de manière fiable que le domaine expérimental sur lequel ils ont été calés. Les utiliser en extrapolation est un risque qu'il faut savoir maîtriser.

Cette approche est souvent la seule possible quand peu de connaissances physiques sont disponibles.

I-1-2- Modélisation de type physique (chimique)

La modélisation de type physique permet d'établir un modèle de procédé à partir des lois fondamentales de la physique et de la chimie telles que la conservation de la masse, le premier principe de la thermodynamique, les équilibres entre phases, les lois de transfert... Dans leur formulation la plus générale, ceux-ci s'écrivent : (ce qui rentre) + (ce qui se crée) = (ce qui sort) + (ce qui s'accumule)

Le terme de création correspond par exemple à la production d'espèce chimique dans un réacteur ou à l'augmentation d'entropie dans un bilan entropique. Il peut être négatif pour traduire une consommation d'espèce chimique dans un réacteur.

Le terme d'accumulation n'intervient que dans les systèmes dynamiques; il traduit les fluctuations de la quantité de produit en réponse à un changement de concentration ou de rétention. Les deux autres termes représentent la contribution des flux convectifs et des phénomènes de transport.

Dans ce type d'approche, les transferts de matière ou de chaleur et les pertes de charge sont définis par des corrélations se référant à des critères hydrodynamiques macroscopiques (nombre de Reynolds) et des constantes physico-chimiques.

Cette modélisation est la plus répandue ; on la retrouve principalement dans les logiciels commerciaux de simulation statique ou dynamique. Ces logiciels extrêmement performants s'appuient sur des modèles thermodynamiques calculant les propriétés de mélange à partir de celles des corps purs (équilibre entre phases, enthalpie, entropie) et sur des variables intensives (température, pression et composition).

Les outils de type génie chimique offrent des ressources importantes pour aider à la compréhension d'un procédé et à la définition d'une structure de procédé. Mais elle impose une parfaite maîtrise des méthodes thermodynamiques employées pour prédire les enthalpies, les entropies (nécessaires pour les compresseurs) et les coefficients d'équilibre.

I-2- La simulation

La simulation est perçue comme un outil d'aide à la décision dont le but principal est d'étudier les performances d'un système complexe. Elle est basée sur la modélisation qui permet de décrire au mieux le fonctionnement réel du système. La simulation consiste alors à conduire des expériences sur le modèle pour en déduire les performances du système, prédire son comportement, tester sa conformité aux normes etc.... Il s'agira souvent d'utiliser des logiciels pour la programmation des modèles analytiques ou descriptifs et leur manipulation afin de réaliser et visualiser des expérimentations.

I-2-1- Principes et limites des logiciels généraux de simulation statique

Les premiers outils de simulation statique de procédés sont apparus au cours des années 1970. Depuis, de nombreux logiciels commerciaux ont été développés. On peut en citer quelques-uns des plus connus :

- ASPEN, commercialisé par Aspen Tech il est extrêmement complet ; il possède une base de données très importante et orientée vers la chimie;
- PRO II, distribué par SimSci (Simulation Science Inc.), il a les mêmes potentiels qu'ASPEN; il est plus orienté vers l'industrie pétrolière ;
- HYSIM, distribué par Hyprotech, il a la particularité d'avoir été développé spécifiquement pour micro-ordinateur; il bénéficie d'une interface très conviviale et il est assez orienté vers l'industrie pétrolière.
- PROSIM, développé à l'ENSIGC de Toulouse et commercialisé par Prosim S.A. ; ce simulateur, moins connu que les premiers cités, est un excellent produit;
- BELSIM, développé par l'Université de Liège, est distribué par Belsim S.A. et présente les mêmes caractéristiques que le précédent.

Tous ces logiciels, par le biais d'une programmation spécifique (*macro-commande*) ou d'une interface graphique permettent le codage d'un problème et exploitent une bibliothèque de sous-programmes et de base de données.

La bibliothèque de sous programmes se décompose en deux parties. La première contient les méthodes de calcul des propriétés physico-chimiques des mélanges ; elles font appel à la base de données qui rassemble les propriétés physiques des corps purs et les paramètres d'interaction caractérisant la non-idéalité des mélanges ; son contenu est essentiel pour que le logiciel soit adapté aux besoins de l'utilisateur. La deuxième rassemble les modèles des opérations unitaires (distillation, équilibre entre phases, échangeurs de chaleur, pompe, compresseur...). La richesse de cette bibliothèque de modèles ou « modules » est un autre critère de comparaison entre les différents produits commerciaux.

Au moment de la résolution du problème, deux approches sont possibles ; elles définissent deux classes de logiciels.

Les plus nombreux résolvent le problème séquentiellement ; les opérations unitaires, appelées «unités» sont calculées les unes après les autres selon un ordre choisi automatiquement ou imposé par l'utilisateur. Dans l'exemple de la (figure 1.5) la simulation du procédé hors recyclage se ferait en calculant le réacteur puis la colonne à distiller.

Lorsque des recyclages existent dans le schéma, et c'est le cas de l'exemple de la figure une procédure doit être mise en place pour les caractériser en débit, composition, température, pression. Les deux opérations unitaires du schéma sont calculées en faisant une hypothèse sur les caractéristiques du recyclage. Après calcul, celles-ci ne seront pas égales, sauf à convergence, avec les valeurs des hypothèses. Il est donc nécessaire de les réactualiser pour tendre vers la solution. Cette décomposition du recyclage en deux flux, l'un permettant de définir une hypothèse, l'autre de la vérifier, est communément appelée coupure. Pour des procédés complexes contenant beaucoup de recyclages, la localisation des coupures n'est pas unique ; leur choix a une forte incidence sur la robustesse et la rapidité des calculs. Celles-ci dépendent aussi des procédures choisies pour accélérer la convergence par le biais de modules adaptés.

Les autres simulateurs résolvent simultanément toutes les équations de bilan à l'aide d'une méthode de type Newton. Les flux de recyclage n'ont plus besoin d'être *coupés* et les convergences sont beaucoup plus rapides. Cette approche est avantageuse pour des procédés comportant beaucoup de recyclages. Inversement, si le choix des hypothèses de calcul est mauvais, l'utilisateur dispose de peu de moyens pour analyser le problème et localiser l'origine de la divergence.

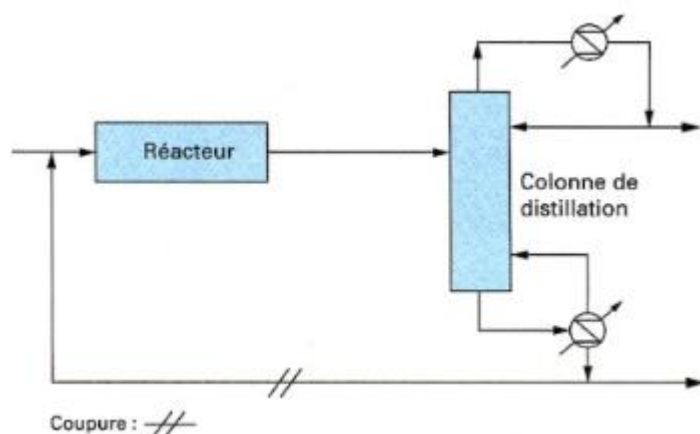


Figure 1.5

Quelle que soit leur origine, ces logiciels ont la caractéristique fondamentale d'être dédiés aux calculs de bilan matière et enthalpique ainsi qu'aux calculs d'équilibre entre phases. Tous les aspects transfert de matière ou de chaleur ne sont pas vraiment pris en compte. En dehors de quelques cas particuliers tels les échangeurs de chaleur, les performances des appareils sont estimées à partir de valeurs fournies par l'utilisateur.

Par contre, ils ont tous la même faiblesse : la modélisation des réacteurs. Tous proposent des modules de réacteur parfaitement agité ou de réacteur piston. Ceux-ci ne sont en fait que des *entités d'accueil* pour y introduire tant le schéma réactionnel que le modèle cinétique. Cette limitation impose donc à l'utilisateur de se procurer son propre modèle en l'achetant spécifiquement à des fournisseurs ou en le développant lui-même.

La faiblesse de la représentation hydrodynamique limite aussi leur domaine d'application. Pour atteindre une représentation plus réaliste, il devient nécessaire d'identifier la distribution des temps de séjour à l'aide des analogies classiques du *génie de la réaction chimique* ou de développer son propre module intégrant toutes les caractéristiques voulues.

Une autre fonction fréquemment rencontrée mérite d'être citée ; presque tous les logiciels offrent des modules d'optimisation exploitant des fonctions créées pour chaque application. Cette option est très importante pour rechercher les meilleures conditions opératoires, minimiser la présence de composés dans les recyclages, optimiser économiquement la conception des unités...

I-2-2- Principes et limites des logiciels généraux de simulation dynamique

Les simulateurs dynamiques calculent l'évolution d'un système en régime transitoire. Contrairement aux simulateurs précédents, ces logiciels tiennent compte des phénomènes d'accumulation dans le procédé. Pris au sens le plus large, ils vont résoudre un système d'équations du type :

$$\begin{cases} g\left[x(t), t, \frac{dx(t)}{dt}, m(t)\right] = 0 \\ h[x(t), t] = 0 \end{cases}$$

Où $x(t)$ représente le vecteur des variables d'état (température, pression, composition, débit...) et t le temps. $m(t)$ est le vecteur des rétentions dans le procédé ; celles-ci sont soit de type matière (volume), soit d'origine thermique (capacité thermique des produits et des parois).

On remarque que ces équations ne tiennent pas compte d'une éventuelle évolution des variables d'état par rapport à une variable spatiale comme on la rencontre dans un réacteur piston, une colonne à garnissage, un échangeur de chaleur. Cette caractéristique est fondamentale et s'énonce de la façon suivante :

Les simulateurs dynamiques ne traitent que des systèmes à paramètres localisés, c'est-à-dire ceux où une seule valeur décrit l'état d'une variable dans un volume.

Un réacteur agité en est bien sur l'exemple typique. Cette limitation est extrêmement importante au niveau pratique car tout système n'appartenant pas à cette classe doit être décomposé en un certain nombre de cellules interconnectées. Cette méthode de discrétisation de la variable spatiale est une approximation satisfaisante dans la plupart des cas (les échangeurs de chaleur, par exemple).

Le choix de la structure du modèle est en fait l'étape déterminante pour obtenir des résultats quantitatifs. Ainsi, dans une colonne à distiller où condenseur et rebouilleur sont les appareils ayant les rétentions les plus importantes, la qualité des résultats sera largement conditionnée par le choix du modèle. Réduire ces appareils à un volume parfaitement agité est souvent insuffisant.

Contrairement au cas précédent où l'offre est pléthorique, peu de simulateurs dynamiques existent commercialement. A ce jour, le seul produit largement diffusé est SPEED UP, développé initialement par l'Impérial Collège de Londres et distribué par AspenTech.

Il faut toutefois noter que les principales fonctions sont les mêmes que celles des simulateurs statiques (établissement des bilans matière et d'enthalpie, absence de calcul des coefficients de transfert, modules de réacteur limités au rôle d'entité d'accueil).

Les principales applications de la simulation dynamique réalisables à l'aide de ce type de logiciels sont :

- l'étude de procédés discontinus (optimisation de conditions opératoires) ;
- les tests de protocoles d'automatisation pour comparer leurs efficacités ;
- les études de sécurité où les conséquences d'événements (fermeture d'une vanne, apparition de fuites) sont quantifiées ;
- les simulateurs d'entraînement. Ceux-ci requièrent un environnement important en termes d'interface.