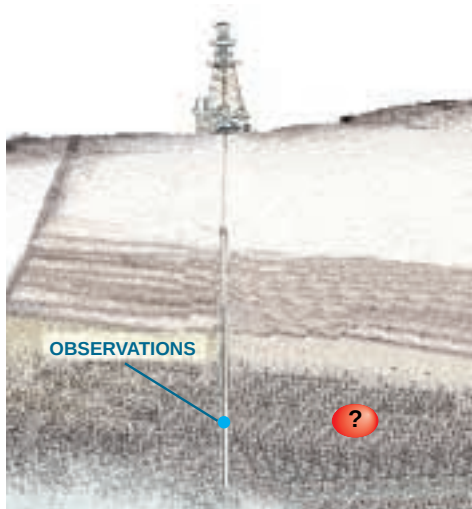
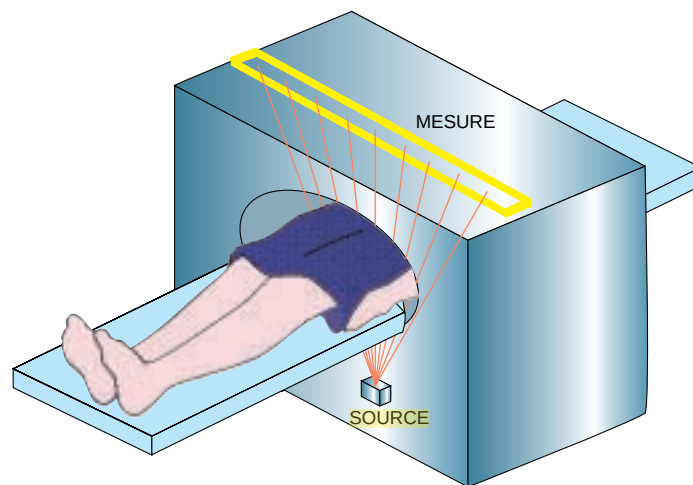




2. COMMENT CONNAÎTRE la quantité de pétrole contenue en un point quelconque d'un gisement ? L'opération est de même nature que la tomographie aux rayons X. Dans les deux cas, les nombres d'in-



connues (la quantité de pétrole en divers points autour du puits, ou l'absorption des rayons X par les différentes zones de la section de corps étudiée) est supérieur au nombre de mesures.

certitude. Nous le verrons, cette démarche se fonde sur la recherche d'une modélisation du problème qui tient compte à la fois des incertitudes et des *a priori* sur les valeurs cherchées.

Le cadre général que nous proposons ne remplace pas le savoir-faire des ingénieurs, mais il les guide pour l'élaboration de méthodes de mesure adaptées à chaque situation. Il leur permet aussi d'évaluer quantitativement la qualité de ces méthodes.

Problèmes non linéaires

La première des difficultés que pose le problème de la mesure tient aux relations entre les inconnues et les observations. Lorsque nous mesurons la vitesse du moteur dans une voiture, nous «remontons» de l'observation de la position angulaire de l'aiguille du compte-tours jusqu'à cette vitesse. On n'a évidemment guère besoin, dans ce cas, de théories ni d'ordinateurs spécialisés. Toutefois, le concepteur du compte-tours a étudié la relation qui existe entre la vitesse (inconnue) et la position angulaire de l'aiguille (observation), puis il a inversé cette relation pour obtenir celle qui existe entre l'observation et la vitesse mesurée : c'est cette dernière relation qu'il a exprimée en traçant la graduation du cadran.

Dans ce cas simple, la relation directe est linéaire : il suffit de multiplier la vitesse par un coefficient constant pour connaître la déviation correspondante de l'aiguille. Pour la mesure de l'épaisseur du zinc, si la relation directe était linéaire, c'est-à-dire si l'impédance à une fréquence donnée était égale à la somme de a fois l'épaisseur, plus b fois

la température, plus c fois la distance, on pourrait écrire trois fois cette relation, pour trois fréquences différentes, et l'on représenterait le problème direct par une matrice carrée, tableau de trois lignes et de trois colonnes contenant les coefficients a , b et c pour chacune des trois fréquences. Afin de résoudre le problème inverse, on calculerait la matrice inverse, et l'on déterminerait les inconnues en multipliant cette matrice inverse par l'ensemble des observations.

Toutefois, en règle générale, on ne dispose pas du même nombre d'observations que d'inconnues. Par exemple, en tomographie aux rayons X, une source ponctuelle émet des rayons X, et l'on souhaite reconstruire l'image d'une coupe (à deux dimensions) d'un objet avec des capteurs disposés en ligne (donc sur une seule dimension). Les inconnues sont l'absorption, par chacun des points de la section corporelle étudiée, du rayonnement émis. Ces absorptions sont à l'évidence plus nombreuses que les observations, l'intensité reçue par les capteurs. *A contrario*, dans le cas des sondages par exemple, on réalise un grand nombre d'observations, alors que l'on ne recherche qu'une inconnue.

Dans tous ces cas, l'inversion n'est pas aussi simple que lorsque le nombre d'observations est égal au nombre d'inconnues. Si la relation directe était linéaire, on la représenterait par une matrice rectangulaire, et on résoudrait le problème inverse en calculant une matrice «inverse», par une méthode mathématique qui généralise la notion d'inverse. Toutefois, les relations directes de la nature physique ne sont

en général pas linéaires, sauf localement (parmi les exemples précédents, les seuls cas linéaires sont ceux du compteur de vitesse et de la tomographie). On ne peut même que rarement exprimer la relation directe sous la forme d'une fonction mathématique explicite qui relie les inconnues aux observations. On est alors contraint de calculer numériquement les valeurs des observations en fonction des inconnues.

Construire un modèle

Dans l'exemple du dépôt de zinc, nous avons utilisé un modèle pour relier les inconnues aux observations. Plus généralement, il existe deux approches pour obtenir un modèle d'un système physique. La première utilise la connaissance des lois internes du système. Dans l'exemple du zinc, on applique les équations fondamentales de l'électromagnétisme à la géométrie de la tôle et de la bobine pour calculer l'impédance de cette dernière. C'est ce que l'on nomme un modèle de connaissance, ou modèle interne. Ces modèles sont utilisés en physique pour la prévision des résultats d'expériences : leur vocation est de représenter au mieux la réalité.

La seconde approche consiste à recenser d'abord les inconnues et les observations ; puis, par des procédés mathématiques appropriés, on choisit la fonction mathématique qui représente au mieux la relation entre les inconnues et les observations. On parle alors de modèles externes, ou de comportement.

Ces deux approches ne sont pas contradictoires ; elles peuvent même être complémentaires. Leurs caractéristiques

sont, en revanche, différentes. Les modèles internes nécessitent non seulement une connaissance fine des phénomènes, qui sera toujours limitée par des approximations, mais aussi la possibilité de déduire de cette connaissance les relations entre les inconnues et les observations. Lorsque l'on veut résoudre le problème direct, on doit intégrer ces lois locales. Sauf dans quelques cas très simples, donc peu réalistes, cette intégration n'est possible que par un calcul numérique. Ces calculs sont longs et complexes, et l'on ne peut donc songer à les utiliser systématiquement dans les mesures, sauf si l'on n'est pas tributaire du temps de calcul.

Les modèles externes ont des avantages et des inconvénients symétriques des modèles internes. Tout d'abord, ils n'ont pas le caractère universel des premiers : ils peuvent être satisfaisants pour un système donné, mais très mauvais pour un autre, à peine différent du premier. De plus, ils ne sont représentatifs que du domaine où l'on a observé les entrées et les sorties du système. Par exemple, un amplificateur peut être modélisé par une constante multiplicative si sa tension d'entrée est inférieure à dix volts, mais des saturations, non prises en compte dans ce modèle, peuvent apparaître au-delà. En revanche, ces modèles ont une très faible complexité calculatoire, ce qui est évidemment avantageux.

Comment alors choisir un modèle ? Un modèle interne est unique, à un degré de finesse donné, tandis qu'il peut exister plusieurs modèles externes, différant par leur structure, c'est-à-dire par la fonction mathématique qui relie les entrées aux sorties. Par exemple, par trois points, on peut choisir de faire passer soit une parabole, soit une cubique (une courbe associée à un polynôme de degré égal à 3), soit encore une fraction rationnelle (le quotient de deux polynômes). Puisque de tels choix existent, on doit tenir compte des objectifs de la mesure.

Imaginons ainsi que nous voulions prédire la population d'un État à partir d'observations passées de celle-ci : il suffit de choisir un modèle de cette évolution et d'extrapoler. Les programmes d'ajustement de courbes proposent souvent, par exemple, d'utiliser des fonctions nommées «polynômes de Lagrange», dont les coefficients sont tels que la courbe représentative passe par tous les points observés (voir la figure 3). Malheureusement, cette courbe chute

rapidement après le dernier point observé : elle ne donne évidemment pas une prévision correcte de la démographie (sauf guerre ou épidémie improbables). En revanche, les polynômes de Lagrange sont utiles pour les estimations, par interpolation, de la population entre deux années d'observation.

Autrement dit, on doit choisir les modèles en fonction de ce que l'on veut mesurer. Un modèle ne sert pas à représenter toute la réalité : il n'est utile que par référence à un objectif donné.

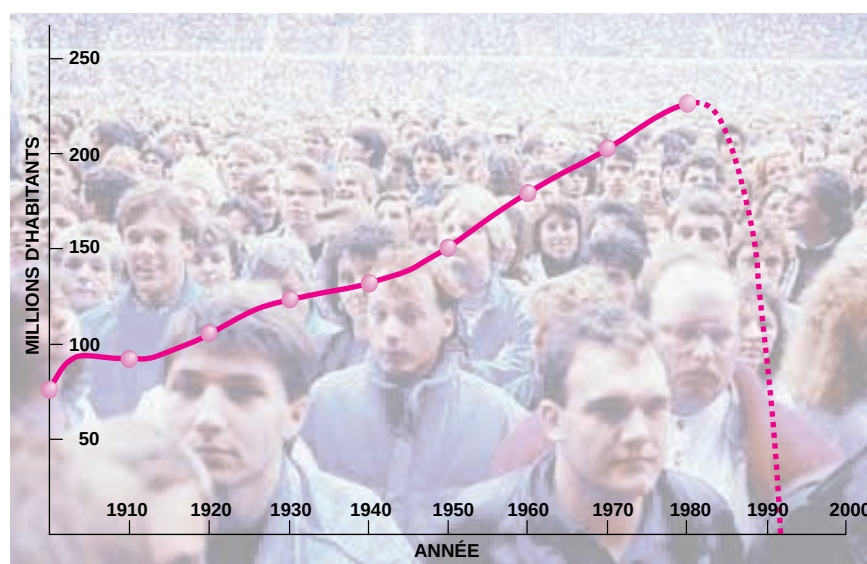
Existe-t-il des critères objectifs de choix de modèles ? Le but d'une mesure est d'obtenir un maximum d'information à partir d'observations. Dans l'exemple de la prédiction de la population, la meilleure fonction n'est pas celle qui passe par tous les points observés, mais celle qui assure la prédiction voulue avec la meilleure précision. De façon générale, le choix du modèle influe sur l'incertitude finale sur les valeurs des inconnues : il est donc logique de choisir le modèle tel que cette incertitude finale soit la plus faible possible. Dans la pratique, on fonde ce choix sur le calcul de la relation entre l'incertitude sur les observations et l'incertitude finale sur les inconnues.

Quantifier l'incertain

On ne parvient toutefois jamais à annuler complètement cette incertitude. S'il est une chose que sait l'homme, ou qu'il devrait savoir, c'est qu'il ne sait pas

tout : sa connaissance de la réalité est imparfaite. La relation directe, aussi simple soit-elle, n'est jamais connue avec une certitude absolue. Dans le cas, pourtant simple, du compteur de vitesse, on ne tient ainsi pas compte du jeu des pièces, ni des vibrations mécaniques, ni du mouvement brownien dans l'air. En pratique, on ne distingue pas l'incertitude des observations des incertitudes liées au choix du modèle : on traite toutes les incertitudes comme s'il s'agissait d'incertitudes sur les observations.

Or, quand on mesure une grandeur, on a autant besoin de connaître sa valeur que la précision avec laquelle elle est obtenue. Nous devons donc, lors de l'inversion, prendre en compte les incertitudes dans nos calculs. Nous utilisons à cette fin une approche probabiliste : à toute grandeur incertaine, nous associons une fonction, nommée fonction de densité de probabilité, dont l'amplitude est proportionnelle à la probabilité de trouver une valeur de la grandeur au voisinage d'un point. Nous pouvons alors appliquer à ces variables incertaines tous les outils mathématiques des probabilités et des statistiques, et nous prenons effectivement en compte l'incertitude dans leur traitement. Cette approche permet aussi de donner une définition quantitative de l'information, par une formule qui exprime la quantité d'information en fonction de la densité de probabilité.



3. LES VARIATIONS DE LA POPULATION des États-Unis sont bien modélisées par une fonction analytique qui passe par tous les points connus (en rose, cet exemple est une mise en garde donnée par le logiciel MATLAB). Ce modèle ne permet manifestement pas de faire des prévisions : il prévoit une diminution très improbable. En revanche, le même modèle est utile pour des interpolations entre des mesures.