Société d'Etudes et de Conseils Mathématiques SAS
*Outils d'aide à la décision*
*depuis 1995*



# Méthodes robustes pour l'aide à la décision :

## *l'apport des probabilités*

### Résumé

La mode est à l'optimisation ; on veut tout optimiser : les stocks, les trajets, les portefeuilles, les campagnes, les traitements, etc. Dans la réalité, on ne dispose jamais de toutes les données nécessaires et celles dont on dispose sont sujettes à caution : incorrectes, périmées, etc. On ne connaît pas non plus les lois qui régissent le phénomène. Dans ces conditions, rechercher un optimum est une aberration :

- L'optimum dépend de manière essentielle des données : une toute petite modification des données risque de modifier l'optimum de manière catastrophique ;

- On est obligé de faire des hypothèses sur les lois qui gouvernent le système, et ces hypothèses sont toujours artificielles ;

- Il faut un temps de calcul assez long pour parvenir à trouver l'optimum ; souvent, on n'y parvient que grâce à des routines numériques approximatives.

Donc, il est bien préférable d'avoir recours à des méthodes probabilistes. Toutes les données en entrée seront considérées comme probabilistes (on n'est sûr de rien !), et on ne recherche plus un optimum, mais une solution acceptable avec une bonne probabilité, obtenue rapidement : ce que nous appelons une "Quick Acceptable Solution".

Les méthodes probabilistes permettent ainsi de répondre rapidement et à peu de frais aux questions parfaitement déterministes que se pose la société civile : par où faire passer les camions ? quelle est la meilleure répartition des stocks ? à qui donner tel médicament? etc.

*Les méthodes probabilistes sont les seules appropriées pour prendre une décision déterministe en environnement incertain. Il ne faut pas faire semblant de croire que les données dont on dispose sont précises et que l'on connaît les lois qui régissent le processus.*

## 1. Présentation générale

On doit souvent prendre une décision sans avoir toute l'information nécessaire. C'est vrai pour des tâches d'organisation (livraisons, plannings), de prospective (évolution des consommations, des marchés). Dans d'autres circonstances, les données disponibles sont très pauvres et souvent approximatives : c'est le cas pour les préoccupations liées à l'environnement et l'épidémiologie ; les lois physiques sont en outre mal connues.

La modélisation n'est pas une science exacte, mais un art assez récent : il a pris sa place industrielle avec les développements de l'informatique, qui ont fait croire que l'on pouvait tout modéliser, tout simuler, tout optimiser et obtenir des résultats précis à partir de données insuffisantes et imprécises. Dans la réalité, il y a des incertitudes sur les données, sur les lois et sur les objectifs. Il est vain de vouloir obtenir des résultats précis.

Pour tenir compte de ces difficultés, il faut des modèles, essentiellement probabilistes, tenant compte des incertitudes. Il sont robustes, parce qu'ils incorporent les incertitudes dès la conception ; leur sortie n'est pas déterministe, mais probabiliste.

Les méthodes probabilistes permettent une première analyse grossière, qui conduit à une hiérarchie des ordres de grandeur : tel risque mérite d'être pris en compte ; tel autre est négligeable. On peut ensuite affiner la connaissance de ceux qui émergent. Elles permettent en outre de déterminer des solutions satisfaisantes : pas optimales, mais satisfaisantes.

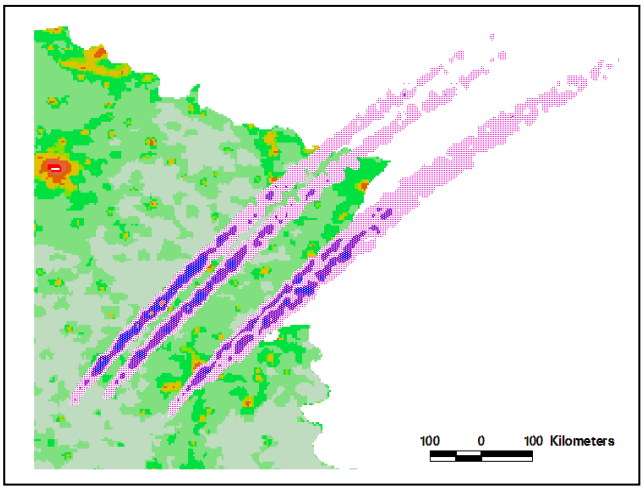
Lorsqu'on met en œuvre des méthodes probabilistes, il n'est pas nécessaire que le phénomène étudié dépende du hasard : rien ne dépend jamais du hasard (sauf peut-être au niveau de la mécanique quantique). Simplement, on décide qu'on "fait comme si" ; on ne recherche pas les causes intimes du phénomène. La plupart du temps, elles dépendent de lois physiques mal connues, sur lesquelles les données sont rares.

## 2. Un principe simple et robuste

Le principe des méthodes probabilistes est simple : on considère que tout ce que l'on ne connaît pas de manière précise est régi par une loi de probabilité. Cela peut-être :

- Une donnée imprécise : on dira par exemple que la consommation de gaz, en France, est liée à la température, mais non pas de manière déterministe.

- Une loi mal connue. En 2004-2005, le CNES nous avait demandé d'étudier les risques associés à la chute des débris provenant de la désintégration de satellites. Il y a très peu de données et les lois sont mal connues. En particulier, dans la formule donnant la résistance de l'air, nous avons considéré que la densité de l'air, aux différentes altitudes, était probabiliste ; de même pour la forme du débris, sa masse, le coefficient de traînée, et même l'exposant de la vitesse : personne n'est sûr que la résistance soit proportionnelle au carré de la vitesse, pour des mobiles à 7 km/s en atmosphère raréfiée.




Le résultat n'a pas été un point de chute précis, mais une "carte probabiliste" : voici où les débris peuvent tomber, avec quelle probabilité. On voit une telle carte sur l'image ci-contre ; elle correspond à la chute de plusieurs objets issus du même satellite ; les zones foncées ont une probabilité plus élevée que les zones claires.

### 3. Amélioration des mesures

Les méthodes probabilistes permettent l'amélioration des mesures, grâce à des tables de calibration appropriées ; ce sont des tables de probabilités conditionnelles. L'idée est simple : voici ce qui a été prédit par le modèle et voici ce qu'on a enregistré en réalité. La différence des deux sert à réaliser une loi de probabilité "erreur de modèle", dont on se sert pour réaliser une correction.

### 4. Différences avec les statistiques

Les statistiques permettent de multiples traitements des données : ajustements, régressions, tests, mais toujours en faisant des hypothèses sur les lois des échantillons recueillis. Or, dans les situations où nous travaillons, ces lois ne sont pas connues. Les statistiques sont une forme raffinée des probabilités (lorsque la loi est connue) et les probabilités sont une forme préliminaire des statistiques (avant que la loi soit connue). Bien entendu, si un historique abondant existe (exemple : ventes d'un produit), il n'y a aucune raison de ne pas utiliser les outils statistiques usuels.

### 5. Modélisation robuste

Une méthode robuste prend en compte, dès le début, les incertitudes sur les lois, sur les données et sur les objectifs. Les programmes industriels n'ont jamais un seul objectif : il y a le court terme et le long terme, les problèmes de production, de maintenance, d'organisation, etc. L'ensemble est toujours complexe, et ramener la question posée à un simple optimum (souvent sur le coût) est réducteur et conduit à des solutions inadaptées.

Il faut abandonner la recherche d'un optimum, qui est toujours fragile : une très petite modification des contraintes risque de modifier considérablement l'optimum. Il faut traiter l'ensemble des objectifs comme des contraintes ; par exemple : "dépenser 10 % de moins que l'année dernière", "réduire les retards de 5 %", etc. Ceci doit être obtenu rapidement : c'est ce que nous appelons une "*Quick Acceptable Solution*". Une fois celle-ci trouvée, rien n'empêche de recommencer, en affinant les contraintes.

Le décideur préférera avoir une solution grossière, mais rapide et robuste, permettant de savoir que telle configuration est satisfaisante, et que telle autre l'est moins.

Beaucoup d'exemples récents, notamment dans le domaine de la logistique, ont montré que les méthodes modernes, reposant sur une optimisation fine et un "just in time" se révélaient très fragiles ; le système s'effondre à la première difficulté.

### 6. Hiérarchisation de paramètres

Un process industriel dépend d'un très grand nombre de paramètres (températures, pressions, compositions chimiques, etc.), qu'il faut hiérarchiser en fonction de leur influence sur une variable de sortie (la qualité du process, quelle qu'en soit la définition). Nos méthodes permettent cette hiérarchisation, quel que soit le nombre des données :
voir la fiche spécifique https://scmsa.eu/fiches/SCM_Hierarchisation.pdf.

### 7. Recherche de zones à risque

Ceci vient en complément du paragraphe précédent. Etant donné un process dépendant d'un très grand nombre de paramètres, il est possible de déterminer les configurations de l'ensemble des paramètres qui conduiront à une situation à risque pour la variable de sortie du process, par exemple une qualité inférieure à un seuil fixé. Ceci se fait uniquement à partir des informations recueillies, et sans introduire d'hypothèse factice.

### 8. Echantillonnage

Il arrive souvent que des mesures aient été faites, mais elles sont peu nombreuses et mal disposées et il faut reconstituer une propriété d'ensemble, par exemple la pollution dans des sols, à partir d'un petit nombre de mesures. La méthode EPH (voir notre livre "PIT", référence plus bas) permet cette reconstitution sans faire d'hypothèse de modèle ; voir à ce sujet nos travaux avec le BRGM.

Bien souvent, l'Industriel veut faire le moins de mesures possible ; les méthodes probabilistes permettent la réalisation de plans d'investigation dynamiques : le résultat du plan précédent conditionne les essais réalisés dans le cadre du plan suivant.

### 9. La "Méthode" d'Archimède

Considérée par Archimède comme son chef d'œuvre, elle a été perdue pendant plus de 2000 ans. Elle consiste à comparer l'information inconnue à une information connue, générée pour la circonstance, et cette comparaison est toujours robuste. Les perspectives sont étonnantes ; voir le livre "Archimedes Modern Works", ref. ci-dessous.

## 10. Nos réalisations

**Livres :**

[IEPE] Bernard Beauzamy : Introduction à l'Etude des Probabilités Expérimentales. Ouvrage édité et commercialisé par la Société de Calcul Mathématique SA. ISBN : 979-10-95773-02-3. ISSN : 1767-1175, janvier 2023.

[MPPR] Bernard Beauzamy : Méthodes probabilistes pour l'étude des phénomènes réels, ISBN : 2-9521458-0-6, Editions de la SCM, mars 2004. Seconde édition, juin 2016.

[PIT] Olga Zeydina et Bernard Beauzamy : Probabilistic Information Transfer. ISBN: 978-2-9521458-6-2, ISSN: 1767-1175. Editions de la SCM, mai 2013.

[AMW] Bernard Beauzamy : Archimedes Modern Works. ISBN 978-2-9521458-7-9, ISSN 1767-1175, Editions de la SCM, août 2012.

[NMP] Bernard Beauzamy : Nouvelles méthodes probabilistes pour l'évaluation des risques. ISBN : 978-2-9521458-4-8, ISSN : 1767-1175, Editions de la SCM, avril 2010.

[RDM] Bernard Beauzamy et Olga Zeydina : Méthodes probabilistes pour la reconstruction de données manquantes, ISBN : 2-9521458-2-2, Editions de la SCM, avril 2007.

**Références récentes :**

- Nuclear Energy Agency, 2010, 2012, 2014, 2015, 2016, 2017 : Détection de données aberrantes dans les bases de données
- ArcelorMittal, 2011-2012 : Méthodes probabilistes pour la qualité d'un usinage
- Air Liquide, 2011 : Construction d'un "indice de proximité" entre pipelines
- IFSTTAR, 2012-2015 (dans le cadre d'un contrat avec le Ministère de l'Ecologie) : Méthodes probabilistes pour l'amélioration du positionnement GPS en environnement urbain
- Areva, 2013 : Hiérarchisation des paramètres intervenant dans un process industriel.
- DCNS, site d'Indret, 2013 : Soudage par Faisceau d'Electrons ; hiérarchisation de l'influence des paramètres par analyse probabiliste
- COSEA (Ligne à Grande Vitesse Sud Est Atlantique), 2013 : détermination des durées de retour pour les crues extrêmes sur la Vienne et la Creuse
- Coop de France déshydratation, 2013 : Réalisation d'un outil d'analyse des Composés Organiques Volatils Non Méthaniques
- IRSN, 2014-2015 : Analyse critique du dimensionnement du réseau de mesure "Teleray"
- Direction Générale Energie-Climat (MEDD), Bureau Qualité de l'Air, 2015 : Relations probabilistes entre trafic et émissions de polluants sur le boulevard périphérique autour de Paris
- Solétanche-Bachy, 2015 : Hiérarchisation des paramètres influant sur la déformation d'un ouvrage d'art

- Carrefour, 2016 : Hiérarchisation des paramètres intervenant sur la vente de jouets
- COSEA, 2016 : Etude probabiliste relative à la turbidité de l'eau
- Taxis G7, 2016 : Analyse probabiliste de bases de données et hiérarchisation de paramètres
- L'Oréal, 2016 : Etude des données disponibles pour les accidents de la route entre le domicile et le lieu de travail
- SNCF Réseau, 2016 : Appui scientifique pour l'analyse des scénarios relatifs à une ligne nouvelle
- ANDRA, 2016, 2017, 2018, 2019 : Optimisation de la position des capteurs pour la surveillance d'un site de stockage
- COSEA, 2016 et 2017 : Etudes statistiques relatives à la turbidité de l'eau
- RATP, 2016-2017 et 2018-2021 : Modélisation du comportement des trains en situation de freinage d'urgence
- RATP, 2017 : Réalisation d'un outil de simulation des temps d'acheminement des trains de travaux
- SNCF/Transilien, 2017 : Analyse critique de modèles de représentation des déplacements ; réalisation d'un outil de simulation
- Monceau Assurances, 2017 : Amélioration de la politique commerciale
- Monceau Assurances, 2017 : Modélisation des catastrophes naturelles et de leur impact sur le portefeuille
- Syndicat des Eaux d'Ile de France, 2017 : appui méthodologique
- Taxis G7, 2017 : Etudes liées à la logistique
- Carrefour/Bazar, 2017 : Outil de recommandation pour les Plans d'Achat
- SNCF Mobilités, 2018 : Etude des déplacements au voisinage du bipôle Nanterre-La Défense
- RTE, 2018 : Analyse des maintenances pour une famille d'équipements
- Atlandes (autoroute A63) : Etude du comportement des véhicules sur les bretelles de sortie de l'autoroute
- RATP, 2018 : Etude probabiliste des efforts dus aux tractions et freinages des matériels roulants
- Eramet, 2018-2019 : Amélioration d'un process industriel
- BRGM, 2018-2019 : Outils probabilistes relatifs à la pollution des sols
- RATP, 2018-2019 : Etude probabiliste des efforts dus aux tractions et freinages des matériels roulants sur la structure des viaducs
- SARP Industries, 2019 : Hiérarchisation des paramètres intervenant dans un process industriel
- Industriel, 2019 : Amélioration d'un process de fabrication
- Orano Mining, 2019 : Hiérarchisation de paramètres intervenant dans un process industriel
- CEA, 2019 : Hiérarchisation de paramètres
- Groupe Atlantic, 2019 : Analyse probabiliste des appels au Service Après-Vente
- Arcelor Mittal Research, 2019-2020 : Amélioration d'un process industriel
- Coop de France Luzerne, 2019 : Analyses statistiques et comparaisons entre usines
- PSA, 2020 : Analyse critique des seuils de réassurance
- Coldway Technologies, 2020 : Réalisation d'une démonstration de sûreté

- Framatome, 2020 : Rédaction d'une démonstration de sûreté pour une carte de contrôle commande
- Air Liquide, 2021 : analyse de la durée de vie de certains composants
- SARP Industries, site de Limay, 2021 : Etude des paramètres influant sur la production de CO2
- Eiffage Rail, 2021 : Outils pour l'analyse de la fiabilité des équipements
- Institution financière, 2021 : Mise au point d'une méthodologie probabiliste pour la prévision temporelle de variables financières
- Monceau Assurances, 2021 : Le risque tempête et le portefeuille de Monceau Assurances
- Institution financière, 2021 : Outil pour l'anticipation des prix du Brent
- RATP, 2021 : Modélisation du comportement des trains en situation de freinage d'urgence
- Teréga, 2021 : Méthodes probabilistes pour la vérification de l'intégrité des canalisations
- Bouygues Energies & Services, 2022 : Appui méthodologique à la conception d'un système d'information "Dysfonctionnements et Maintenances"
- Befesa Valéra, 2022 : Hiérarchisation des paramètres intervenant dans le réglage d'un four
- RATP, 2022 : Analyse de la stabilité de talus anciens ; l'approche d'Archimède
- Léon Grosse, 2022 : Analyse du risque "grêle"
- SNCF, 2023 : Réalisation d'un plan d'inspection des rails
- Agence Nationale des Titres Sécurisés, 2023 : Anticipation des demandes en Titres Sécurisés
- Peptinov, 2023 : Traitement probabiliste de données épidémiologiques
- Cristal Union, 2023 : Méthodes probabilistes pour la comparaison d'essais de biocides
- Coopération Agricole "Luzerne de France" : Homogénéisation de bases de données
- SNCF, 2024 : Analyse d'une approche probabiliste de valorisation des risques associés aux coûts des projets
- Bureau de Recherches Géologiques et Minières, 2024 : Outils mathématiques pour la cartographie des pollutions
- Ville de Villiers le Bâcle, Essonne, 2024 : Calcul de la probabilité de retour de pluies extrêmes
- RATP, 2024-2025 : Mission de type second regard sur le traitement du risque de rupture de rail en exploitation dans le cadre de l'automatisation de la ligne 13 du métro de Paris
- Entreprise industrielle, 2025 : Analyse critique d'un plan de production et suggestions d'amélioration
- Atlandes SA, Autoroute A63, 2025 : Détection d'essieux relevés ; analyse critique de rapports disponibles & traitement d'une base de données
- Orano, 2025-2026 : Définition d'un échantillonnage représentatif de l'ensemble d'un bain de fusion
- Toulouse Métropole, 2025-2026 : Calculs d'incertitudes associées à une installation de méthanisation
- SNCF, 2025-2026 : Etudes relatives à la qualité de l'air

- RATP, 2025-2026 : Etudes relatives à la qualité de l'air
- Ardian, 2026 : Méthodes mathématiques pour l'étude de la pertinence de nouvelles acquisitions
- Bouygues Energies & Services, 2026 : Analyse critique des spécifications relatives à certains équipements
- Caisse Centrale de Réassurance, 2026 : Etude des phénomènes climatiques relatifs à l'île de Mayotte