

## Vers une nouvelle ingénierie combinant physique et données, manipulées et intégrées dans le cadre des mathématiques et l'intelligence artificielle

Journée de la Commission Structures de la 3AF autour du « Big-Data »  
6 décembre 2017 à l'ONERA Châtillon  
Prof. Francisco Chinesta – Arts et Métiers ParisTech

### Le contexte

Initialement, l'industrie a adopté les jumeaux virtuels (*virtual twins*) sous forme d'outils de simulation qui représentaient la physique des matériaux, procédés, structures et systèmes à partir d'un modèle numérique. Cependant, les moyens de calcul plutôt limités au sein des petites et moyennes entreprises, ne permettaient pas de simuler les problèmes industriels aussi vite que les ingénieurs l'auraient souhaité. Malgré cela, ces outils de calcul ont transformé le métier de l'ingénieur pour lui offrir des outils de conception optimisés et sont devenus incontournables dans la presque totalité des industries à la fin du XXe siècle.

C'est au début du XXIe siècle que les données font irruption dans l'ingénierie. Pendant des années, elles ont servi dans d'autres domaines où les modèles étaient moins développés ou restaient plus incertains. Les données collectées massivement ont pu être classifiées, décortiquées, analysées, ... à l'aide des techniques de l'intelligence artificielle. Ainsi les corrélations entre les données ont pu être enlevées, prouvant une certaine simplicité derrière une apparente complexité. En plus, des relations entrée-sortie ont été mises en place avec les données comme seul ingrédient. Les données multidimensionnelles ont trouvé une façon de se montrer à nos yeux et les données massives ont pu nous fournir des clés d'interprétation, nous avertissant d'un événement fortuit imminent, rendant possible une maintenance prédictive, une surveillance et une prise de décision en temps réel. Nous voici dans le royaume des jumeaux digitaux (*digital twin*), où les modèles physiques devant choisir entre précision et rapidité, se sont fait remplacer par des données. L'emploi de ces dernières nécessite un apprentissage adéquat au préalable, comme l'utilisation de modèles a aussi nécessité des longs travaux de découverte et de formulation.

Nous voici avec des jumeaux virtuels hors-ligne et des jumeaux digitaux en-ligne. Cependant ces derniers n'ont pas aujourd'hui la portée des premiers et dans de nombreuses applications, « le tout données » reste hors de portée. C'est là où le jumeau hybride (*hybrid twin<sup>TM</sup>*) rentre en scène, un jumeau combinant modèles et données. Mais pour cela, il a fallu attendre que les simulations numériques des systèmes et des physiques complexes puissent se faire en temps réel.

Au début du XXIe siècle toujours, plusieurs révolutions scientifiques en mathématiques appliquées, en informatique (calculs hautes performances) et en mécanique numérique, sont venues bouleverser les procédures de conception établies. Ainsi, des nouvelles techniques dites de réduction de modèles ont vu le jour. Ces techniques ne réduisent ni ne modifient le modèle, elles se contentent de réduire la complexité de sa résolution et permettent ainsi de transformer une réponse complexe et longue à obtenir, en une réponse en temps réel. De plus, ces nouvelles techniques ont complètement modifié les approches traditionnelles d'utilisation de la simulation, de l'optimisation, de l'analyse inverse (e.g. recalage), du contrôle et de la propagation de l'incertitude, en offrant des réponses temps réel aux problèmes posés. C'est là que le jumeau hybride est né. Le calcul pouvant maintenant être fait en temps réel, les données collectées permettent de calibrer (continuellement) les modèles numériques pour mieux s'adapter à la réalité observée, rendant possible le contrôle temps réel.

Le virtuel avait été enfin combiné avec le digital, la physique avec les données. Cependant, et malgré des énormes succès, certaines difficultés ne tardèrent pas à apparaître : dans de nombreux cas, même un recalage continu ne permettait pas de décrire et prédire la réalité observée. Il semblait que notre conceptualisation de la réalité – nos modèles – contenait une dose non négligeable d'ignorance, la réalité semblant inatteignable à l'échelle du raisonnable. Le jumeau hybride a été ainsi enrichi, en lui associant deux types de modèles : le premier basé sur la physique, connue et calibrée continuellement en temps réel par assimilation des données collectées ; le second est un tout nouveau type de modèle, à caractère plus pragmatique et phénoménologique, construit à la volée à partir (et exclusivement) des données collectées, et servant à expliquer la déviation entre les prédictions du modèle physique calibré et la réalité observée. Là, comme disent les français, on a enfin le beurre et l'argent du beurre !

## La journée

Le programme de la journée organisée par la commission Structures de la 3AF, qui a rempli la salle de conférences au delà de sa capacité et des prévisions, avec des acteurs du monde académique et industriel, a tourné autour de toutes ces questions, avec une double finalité, bâtir l'état des lieux, et puis identifier les tendances actuelles ainsi que les rêves de demain (difficile de se projeter encore au-delà, tellement l'évolution est rapide) sans oublier les verrous à soulever.

Le programme a été en conséquence très riche, combinant fondamental et applicatif. La journée s'est ouverte avec quelques présentations à caractère plutôt fondamental, avec notamment la présentation de la Prof. Catherine Huber - Université Paris 5 René Descartes – qui a passé en revue de nombreuses techniques d'analyses statistique de données, tout en avertissant sur les risques des mauvaises manipulations et en même temps démystifiant certaines procédures obscures mais auxquelles elle a magistralement rendu la lumière.

Puis la présentation du Prof. Francisco Chinesta – ENSAM ParisTech - a établi la différence entre les membres de la famille des jumeaux (virtuels, digitaux et hybrides) et illustré que bien que les données puissent venir, et viennent, enrichir la physique, cette dernière a son mot à dire pour la construction d'une nouvelle science des données, le Smart-Data, plus approprié dans nos métiers d'ingénieur que le Big-Data.

La conférence suivante du duo Bernard Troclet – ARIANEGROUP - et Kambiz Kayvantash – CADLM – s'est plus centrée sur la gestion des données à l'issue des vols des lanceurs ARIANE, avec toute leur richesse, ses dites Vs : volume, variabilité, vitesses, véracité, etc ... Ils ont prouvé que malgré la quantité et la possibilité de les faire parler et s'exprimer (les données), un grand travail reste à faire pour avoir une capacité d'explication et mener une prise de décision fondée sur la compréhension.

Puis le duo du Prof. Joseph Morlier – SUPAERO – et Michele Colombo – Airbus - ont montré une avancée sur le traitement des grands volumes de données grâce aux dites DGP (pour Distributed Gaussian Process), présentation qui a été suivie par celle de M. Gabriel Broux de Dassault-Aviation, qui en plus de nous parler, comme portait le titre de sa présentation, sur les opportunités du Big-Data dans le domaine des structures des avions d'affaires, il a prouvé en plus l'urgence et la nécessité de l'accomplir vite, car seulement une toute petite quantité des données collectées sont dûment traitées et analysées, le reste attendant sur étagère.

Les présentations de l'après midi (celle du duo Jérôme Lacaille et Pierre-Etienne Mosser - Safran Aircraft Engines ; du binôme Christophe Favre - Safran Helicopter Engines et Sébastien Razakarivony - Safran Tech ; et enfin celle de Frédéric Le Corre – MBDA) ont porté sur des applications, comportant analyse de données, extraction de « patterns » pour une maintenance prédictive, pour l'anticipation de comportements fortuits, pour le contrôle non destructif, pour la modélisation basée sur des données (apprentissage profond et renforcé, entre d'autres possibilités comme le sont les techniques de régression) du comportement de composants et de systèmes, l'extraction de paramètres explicatifs, l'enlèvement de redondances et de corrélations via une analyse statistique de données, et enfin l'utilisation des techniques avancées de classification et de visualisation multidimensionnelle, le tout au service d'une prise de décisions temps-réel.

## La discussion

La journée s'est terminée par une table ronde animée par Francisco Chinesta, même si l'animateur n'a pas eu besoin d'animer, l'animation s'est fait toute seule, avec une salle encore bondée en fin de journée et une forêt de mains qui se levaient demandant la parole ! Un régal !

Des sujets majeurs ont été abordés. Pierre Ladevèze (ENS Cachan) a pris la parole pour replacer la vérification et la validation - V&V - dans le nouveau cadre des modèles construits sur la base des données. Même si la V&V est en train de s'imposer dans certains domaine industriel, il est difficile d'imaginer la nouvelle ingénierie l'excluant, elle sera sans doute un protagoniste incontournable. Pierre a abordé aussi les questions de qualité des données, bruits, erreurs de mesure, défaillance de capteurs, biais, ... qui nécessitent des techniques de filtrage adéquats et de modélisation robuste. D'autres collègues ont rebondi sur ces questions, Roger Ohayon (CNAM), Christian Rey (Safran Tech), ... en rajoutant d'autres questions majeures : comment construire des modèles basés sur les données dans un cadre thermodynamique, en accord avec les principes fondamentaux, le rôle des variables internes dans les nouvelles équations de comportement fondées sur les données, avec les questions associées sur l'échantillonnage, le calibrage et le recalage, la prise en compte de couplages, ...

Les questions associées aux métriques ont été soulevées. Comment comparer l'incomparable ? Quelles

métriques ? L'euclidienne s'avère t'elle insuffisante ? Bien que deux arbres avec une distance euclidienne puissent rester très éloignés, un enfant est capable de les associer et exprimer leur ressemblance ! Comment extraire la simplicité cachée dans la complexité. Nous différencions en temps réel un être humain dans la rue, un être humain que nous voyons pour une toute première fois, ... donc en ce qui concerne cette classification, être ou ne pas être humain, peu de paramètres doivent suffire car sinon nous prendrions une demie heure avant de conclure, fait qui nous aurait mis en danger au long de l'évolution des espèces, et certainement nous ne serions pas là à présent, ... si nous le sommes, c'est certainement que nous avons su domestiquer la complexité, extraire sa simplicité cachée, le peu de paramètres essentiels et explicatifs, « goal-oriented ».

Enfin, tous, industriels en tête car pressés, et académiques derrière car paniqués par l'urgence combinée avec la difficulté, réclament un nouveau paradigme, celui du smart-data. Où mesurer (lieu et échelle) ? Quand le faire ? Et quoi mesurer ? Car avoir trop de données peut devenir aussi handicapant que ne pas en avoir. N'oublions pas que dans le domaine de l'ingénierie, contrairement à d'autres domaines, les données sont chères à obtenir (on ne peut pas mettre un capteur là où on veut ou avoir la donnée dont on rêve, il faut faire avec les moyens du bord) et chères à traiter (en temps) surtout si on vise le temps réel imposé par les tâches d'analyse, de surveillance et de contrôle (e.g. voiture autonome, ...).

Les rêves exprimés ont été multiples, et les choses vont tellement vite, que certains des rêves exprimés le 6 décembre 2017 sont au moment où nous écrivons ces lignes (août 2018) déjà réalisés.

La journée a fini en rappelant ce moment, au début du XVIIIe siècle où Alessandro Volta a présenté son invention, la pile électrique, dans une des séances de l'Académie des Sciences à Paris. Les académiciens ont applaudi ce spectacle inouï, où les électrodes appliquées sur une grenouille provoquaient des énormes sauts. A la fin, Napoléon a demandé : M. Volta, pensez vous que ce que vous appelez électricité servira un jour à autre chose qu'à faire sauter des grenouilles ?

Nous voici nous aussi confrontés à un changement de paradigme ! Impossible d'imaginer, la réalité sera bien au-delà de la plus créative des imaginations. Allons-y, construisons cette nouvelle ingénierie pour une nouvelle technologie au service d'une nouvelle société et humanité.