Cahiers de Psychologie Politique

N° 33 | 2018
Quantification et quantité Juillet 2018

Enjeux et limites des modèles

Frédéric PATRAS

Édition électronique :
URL : https://cpp.numerev.com/articles/revue-33/1920-enjeux-et-limites-des-modeles
DOI : numerev_1170
ISSN : 1776-274X
Date de publication : 16/07/2018

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : PATRAS, F. (2018) Enjeux et limites des modèles. *Cahiers de Psychologie Politique*, (33). https://doi.org/10.34745/numerev_1170

DOSSIER : QUANTIFICATION ET QUANTITE

Jusqu'où est-il possible de donner une modélisation mathématique du réel, et de lui faire confiance ? Jusqu'où est-il possible de quantifier le réel – physique ou social – et quelle valeur scientifique, sociale et éthique donner à cette quantification ?

Mots-clés :

DOSSIER : QUANTIFICATION ET QUANTITE

Frédéric Patras est directeur de recherche au CNRS où Il dirige l'équipe algèbre et topologie de l'UMR 6621 (Laboratoire J.-A. Dieudonné) et il enseigne au département de mathématique de l'université de Nice. Ancien élève de l'ENS section mathématique, il est l'auteur de plusieurs ouvrages dont La pensée mathématique contemporaine (PUF, 2001) et plus récemment La possibilité des nombres (PUF 2014).

SOMMAIRE

1. Introduction

2. La tentation pythagoricienne

3. Le problème de la technique

4. L'apprentissage profond

1. Introduction

Jusqu'où est-il possible de donner une modélisation mathématique du réel, et de lui faire confiance ? Jusqu'où est-il possible de quantifier le réel – physique ou social – et quelle valeur scientifique, sociale et éthique donner à cette quantification ?

Les deux questions se ressemblent, et il est tentant de penser qu'au moins en première approximation la seconde ne serait qu'un avatar de la première – la quantification étant,

en fin de compte, l'aboutissement d'une modélisation réussie comme lorsque, en physique des particules, la théorie quantique des champs permet de prédire avec une précision incroyable le moment magnétique de l'électron[1]. La qualité de la prédiction numérique est, de toute évidence, dans de telles situations, l'indice de la qualité du modèle théorique.

Pourtant, dans l'usage quotidien, et surtout dans l'usage sociétal des mathématiques – on pensera surtout ici aux statistiques, mais la remarque vaudrait pour les modèles mathématiques utilisés par exemple en économie –, il est le plus souvent impossible de raisonner comme en physique dès lors que les phénomènes en jeu ne peuvent le plus souvent pas faire l'objet de mesures indépendantes fiables. Tout au plus pourra-t-on constater a posteriori l'inanité d'un sondage électoral, en comparant ses prédictions au résultat des urnes, et pourra-t-on en déduire, soit l'invalidité des modèles qu'il utilise, soit la maladresse ou le manque de professionnalisme du sondeur.

Dans certains cas, un tel recours à un jugement de valeur a posteriori est cependant impossible car il est difficile et parfois impossible de donner un sens objectif à telle ou telle quantité en dehors du contexte que met en place le modèle qui la calcule. Un exemple intéressant de ces phénomènes est donné par les enquêtes de type PISA[2] (Program for international student assessment), dont on sait l'importance qui leur est accordée pour évaluer le système éducatif. Le problème que soulèvent ces évaluations tient à un phénomène trop vite oublié, qui est celui des choix préalables effectués dans la définition même de ce qu'est une « bonne éducation ». Tout rabattre ensuite sur quelques nombres et un classement, en oubliant que ces choix, nécessairement en partie subjectifs, les conditionnent, conduit à un phénomène d'unidimensionnalité proche de celui que dénonçait Marcuse[3] et qui laisse une place limitée à un débat authentique – la tentation des gouvernements étant d'aligner leur politique sur celle préconisée implicitement par ces classements. L'enseignement abonde d'ailleurs d'indicateurs qui, initialement méthodes d'évaluation quantitative, deviennent rapidement normatifs en imposant une certaine grille de lecture sur la qualité et « l'excellence ». Le classement de Shanghaï[4] des établissements d'enseignement supérieur en est un autre exemple célèbre et symptomatique.

La question que cet article voudrait poser, dans cette perspective, est celle du rapport des nombres, mesures, classements, et de leurs différents usages sociétaux aux modèles qui sous-tendent leur calcul. Il voudrait aussi interroger le phénomène – peut-être plus subtil – qui sous-tend la confiance que nous avons spontanément dans la faculté des nombres à décrire, voire à normer le réel – confiance qui conduit souvent à oublier que cette description repose toujours sur une modélisation sous-jacente, et donc sur des hypothèses, des conjectures, des régularités observées, mais dont la portée peut s'avérer incertaine ou d'un domaine de validité restreint.

2. La tentation pythagoricienne

Le pacte apophantique qui lie les nombres[5] au réel a des origines multiples. Peut-être est-il, en fin de compte, une émanation spontanée des règles, des mécanismes qui sous-tendent notre pensée et c'est à ce titre qu'on le retrouverait dans toutes les civilisations et dans tant de formes de mysticisme. Comme souvent, c'est pourtant seulement avec la pensée grecque que ces idées se sont cristallisées sous une forme théorique et, plus précisément, avec le pythagorisme. Rappelons le témoignage d'Aristote :

« Ceux qu'on désigne sous le nom de pythagoriciens se consacrèrent les premiers aux mathématiques. Nourris dans cette discipline, ils estimèrent que les principes des mathématiques sont les principes de tous les êtres. »

Les nombres sont parmi ces principes, premiers ; les pythogoriciens voyaient par ailleurs en eux une multitude d'analogies avec ce qui est et devient. Comme, enfin, tout leur paraissait finalement être formé à la ressemblance des nombres, « dans ces conditions, ils considérèrent que les principes des nombres étaient les principes de tous les êtres, et que le Ciel tout entier est harmonie et nombre ».[6]

Bien entendu, nous ne sommes plus pythagoriciens sous une forme naïve : personne (ou presque) ne penserait aujourd'hui comme eux que l'âme, principe entre tous, s'identifie au 1, l'intelligence, mouvement des prémisses à la conclusion, au 2, et ainsi de suite. Pour autant, le développement même de la science moderne, l'idée galiléenne que le monde est écrit en langue mathématique[7], la possibilité enfin de décrire tant de phénomènes par l'intermédiaire de modèles quantitatifs, ont souvent conduit à une nouvelle forme de prégnance de l'idée que les nombres sont à même de rendre compte de la nature de choses.

Ce que l'on qualifiera ici de « risque pythagoricien » tient à la généralisation abusive et la décontextualisation de certaines modalités mathématiques de connaissance et d'acquisition d'emprise sur le monde. De ce qu'il est possible de quantifier tel ou tel phénomène et d'accéder par-là à une forme authentique, profonde et scientifique de connaissance, il ne faut pas se hâter de conclure, ni que toute connaissance serait nécessairement de cette forme, ni que toute quantification donne accès à l'être des choses – sans même que soit posée la question de sa pertinence et de son domaine de validité.

3. Le problème de la technique

Une des clés de la compréhension des usages déplacés des nombres comme mesures inconditionnelles de l'être des choses tient à l'aura de mystère qui les a toujours entourés – dans le néoplatonisme, l'hermétisme, la kabbale… –, mais également au problème de la technique.

Très souvent, les modèles qui sous-tendent les quantifications du réel nous sont inaccessibles. Fréquemment, parce que les modèles sont trop complexes pour pouvoir être appréhendés – parfois même par des mathématiciens chevronnés. Le seul accès possible à l'intelligibilité est alors au travers de leurs résultats. Ainsi des prévisions météo, qui mettent en jeu des modèles extrêmement sophistiqués, couplés à toute une technologie d'acquisition des données. Parfois, l'absence d'intelligibilité vient, plutôt que de la complexité des modèles, d'un secret technique ou technologique, d'un savoir-faire caché – comme avec les instituts de sondage, dont les méthodes reposent tout autant sur un ensemble de méthodes empiriques de collecte et de retraitement des données que sur des outils statistiques théoriques. Pour autant, dans ces deux cas, un contrôle a posteriori des prévisions et estimations est possible – qui nourrit d'ailleurs une méfiance presque atavique du public à l'égard aussi bien des prévisions météorologiques que des sondages d'opinion.

D'une toute autre nature est le rapport que la société entretient aux quantités issues de modèles qui définissent leurs propres règles d'interprétation et d'intelligibilité. Qui sait, par exemple, pour commencer avec deux exemples pourtant simples et familiers, quel est le panier de produits sur lequel se base le calcul de l'inflation ? Les règles de calcul du taux de chômage ?

Les « subprimes »[8], ces « produits dérivés toxiques » qui ont été à l'origine de la crise financière mondiale de 2008[9], donnent un exemple emblématique de ce phénomène. Il a ceci d'intéressant, outre de par sa dimension historique, que le problème de l'intelligibilité des quantités, de leurs rapports aux modèles sous-jacents, s'est essentiellement posé à l'intérieur d'une communauté – celle de la finance – pourtant aguerrie au maniement des chiffres et tout sauf naïve quant au fonctionnement des modèles.

Beaucoup a été dit sur la crise, sur ses origines en partie frauduleuses, les instituts de crédit ayant émis des prêts hypothécaires sans suivre les règles de prudence les plus élémentaires – puisqu'ils allaient ensuite transférer ces prêts et les risques afférents à d'autres investisseurs. On sait moins que la crise repose pour beaucoup par ailleurs sur plusieurs erreurs de modélisation et la transposition au-delà de leur domaine de validité d'idées mathématiquement et techniquement raisonnables – comme l'utilisation en finance du théorème central limite (qui explique le comportement d'une famille suffisamment nombreuse de quantités aléatoires ayant des comportements indépendants).

Le problème clé posé par les centaines de crédits hypothécaires sous-jacents aux RMBS[10] était précisément que, pour une multitude de raisons, leur comportement ne se prête pas à une modélisation simple. Trop de paramètres entrent en jeu : les taux d'intérêt, les prix de l'immobilier, la solvabilité des ménages. Ce type de quantités suit des dynamiques très complexes qui en rendent la modélisation illusoire sur les périodes très longues typiques pour ce type de prêt (10 ans, 20 ans, parfois plus) alors même qu'à de telles échéances un changement anodin (disons par exemple 1 % sur les taux d'intérêt à 20 ans) affecte très profondément l'ensemble des processus de valorisation.

En deux mots, il s'agissait avec les RMBS de ce que les mathématiciens appellent des systèmes complexes, mais d'une complexité rebelle à la modélisation et à la quantification. Seules des hypothèses extrêmement simplificatrices permettaient de donner un prix aux différents produits financiers issus de ces paniers de crédits hypothécaires, avec les conséquences que l'on sait. Certains des grands dirigeants des banques qui allaient faire faillite ou être sauvées ou rachetées in extremis soutenaient pourtant peu avant la crise que ce type de produits financiers contribuait à stabiliser le système bancaire international[11] ! Le point clé sur lequel il convient d'insister ici est qu'en dépit de défauts de modélisation évidents a posteriori, un large consensus technique soutenait donc leur diffusion à grande échelle. La raison ultime de cet aveuglement tient sans doute en grande partie à la confiance, construite au fil de décennies d'ingénierie financière, dans un certain méta-discours qui accompagnait le développement de la finance quantitative, méta-discours dont les principes mathématiques fondateurs avaient pourtant cessé de valoir dans des configurations nouvelles.

Une telle utilisation de paradigmes mathématiques et techniques au-delà de leur domaine de légitimité ; la confiance qui est accordée à des modèles et les mesures qui en sont issues, dont on ne questionne pas la validité, soit du fait de difficultés techniques ou conceptuelles, soit par le simple jeu de l'inertie et de l'habitude : il s'agit là de situations au fond assez classiques, dont seule la répétition peut surprendre. Ce que cet article voudrait examiner, en conclusion, est de nouvelles modalités de rapport aux quantités et à leur utilisation sociétale, dans le contexte actuel, très particulier, où émergent de nouvelles formes de l'idée même de modèle.

4. L'apprentissage profond

Un changement de paradigme radical est en train de s'opérer dans l'univers de la modélisation – il ne convient d'ailleurs plus de parler de modélisation mathématique proprement dite mais sans doute plutôt de « modélisation digitale ».

Dans toutes les configurations théoriques que nous avons examinées, les nombres, les mesures, étaient issus de modèles, et les questions qui se posaient étaient celles de leur validité et leur intelligibilité. Faute de comprendre un modèle et ses ressorts, notre accès à la signification authentique de telle ou telle quantité qui en était issue nous était refusée, mais, au moins en principe, en accédant au modèle et à ses spécifications, seuls des obstacles cognitifs et techniques pouvaient faire obstruction à une entente authentique des méthodes, des enjeux et des résultats.

Avec l'intelligence artificielle[12] et en particulier les réseaux de neurones et ce que l'on appelle l'apprentissage profond (deep learning), l'essence même de notre rapport à la quantification est en train de changer. Les techniques d'intelligence artificielle reposent en effet sur des mécanismes nouveaux où l'accès au modèle de calcul sous-jacent ne donne plus accès, pour des raisons structurelles et de principe, à la compréhension des résultats. L'idée de départ est assez simple : il s'agit typiquement d'approximer une

fonction à un grand nombre de variables produisant quelques valeurs en sortie. Au hasard, la détection d'une tumeur maligne à partir de photos de peau, les mouvements dans un jeu d'échec, la reconnaissance faciale...

Il s'agit d'un problème on ne peut plus classique qui aurait été résolu traditionnellement en « fittant » la fonction (c'est-à-dire en calibrant de façon optimale ses paramètres) aux données disponibles. Trois idées nouvelles viennent perturber ce schéma, rendues possibles par l'accroissement de puissance de calcul et de mémoire des machines. D'abord, l'utilisation dynamique de masses importantes de données. Ensuite, l'utilisation de compositions itérées de fonctions, qui rend l'espace fonctionnel total à explorer d'une dimension inaccessible à des techniques d'optimisation directes. Enfin, la technique d'optimisation (ou « d'apprentissage ») proprement dite, qui repose sur des algorithmes récursifs permettant de propager l'information d'entrée/sortie des observations disponibles sur les différents niveaux (ou couches) fonctionnels.

Pour ce qui nous intéresse, la grande nouveauté de ces techniques est que leur résultat (en général une succession de compositions de fonctions non linéaires) n'a tout simplement pas de sens, de signification accessible à l'analyse. Aussi bien le processus de calcul des paramètres que le résultat du processus ne peuvent (en général) rien nous dire d'humainement intelligible sur les phénomènes auxquels ils s'appliquent.

Le plus inquiétant n'est peut-être pas pourtant cette perte de contrôle de l'homme sur la machine, sur les algorithmes, que leur extraordinaire efficacité et leur prégnance toujours croissante dans le quotidien : la « pensée numérique » émergente fonctionne dans de nombreuses situations mieux que la pensée humaine, pour des tâches toujours plus variées et parfois requérant traditionnellement une solide formation intellectuelle et scientifique (comme en médecine).

Bien entendu, le problème clé, problème majeur des années qui viennent, est que cette « pensée numérique » n'en est pas une : pensée sans objet, sans contenus, sans conscience, sans éthique, et que son émergence pourrait nous conduire, collectivement, à sacrifier ce qui faisait jusqu'ici l'essence de la pensée scientifique : non pas la seule volonté de maîtriser la nature et les phénomènes, mais également celle d'accéder à leur compréhension, à l'essence des choses.

1 L'ordre de grandeur de l'erreur commise en comparant les prédictions théoriques du modèle dit standard de la théorie quantique des champs aux valeurs expérimentales est extraordinairement faible. De par sa précision incroyable, qui valide le modèle sous-jacent, ce résultat est considéré comme l'un des plus beaux succès de la physique. Voir par exemple B. Odom, D. Hanneke, B. D'Urso, G. Gabrielse, "New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron". Phys. Rev. Lett. 97 (3): 030801 (2006) et Phys. Rev. Lett. 99, 039902 (2007).

2 http://www.oecd.org/pisa/PISA-2015-Brochure-France.pdf

3 H. Marcuse, L'Homme unidimensionnel, Le Seuil, 1968.

4 http://www.shanghairanking.com/

5 Sur l'idée de nombre et son évolution on renvoie à F. Patras, « La Possibilité des nombres », PUF, 2014.

6 Aristote, Métaphysique, trad. J. Tricot, Vrin, Paris, 1986.

7 Sur ces questions, on consultera Husserl, *La crise des sciences européennes et la phénoménologie transcendantale*. Gallimard, 1976, toujours d'actualité sur de nombreux points -en particulier de par ses analyses d'une perte de contact entre la science et le « monde de la vie » mais également de par la méthode d'analyse adoptée, qui s'appliquerait avec profit à diverses questions contemporaines.

8 La crise des subprimes a pour origine des produits financiers complexes structurés à partir de prêts hypothécaires émis aux Etats-Unis. Elle s'est accompagnée d'une crise bancaire, conduisant à la crise financière mondiale de 2007-2008.

9 Bielecki, Tomasz, Damiano Brigo, and Frédéric Patras. *Credit risk frontiers: Subprime crisis, pricing and hedging, CVA, MBS, ratings, and liquidity*. Vol. 138. John Wiley & Sons, 2011.

10 Residential mortgage-backed securities. Il s'agissait de paniers de plusieurs centaines de prêts hypothécaires liés à l'acquisition de résidences individuelles, essentiellement aux Etats-Unis, dont les risques étaient répartis entre différentes catégories d'investisseurs en fonction de leur appétence au risque.

11 Il est d'ailleurs assez intéressant de consulter les rapports techniques des mois qui ont précédé la crise financière, par les grands acteurs de l'époque -à commencer par Lehman Brothers, dont la faillite retentissante allait emporter l'économie et la finance mondiale, brisant le tabou du « too big to fail ».

12 On consultera pour plus de détails sur l'intelligence artificielle les ouvrages de Jean-Gabriel Ganascia, dont : Le mythe de la Singularité : faut-il craindre l'intelligence artificielle ?, éditions du Seuil, Collection Sciences Ouvertes, 2017. L'intelligence artificielle : vers une domination programmée ?, éditions du Cavalier Bleu, 2017.