

DOCTORAT

Histoire et philosophie des sciences

La valeur de l'incertitude: l'évaluation de la précision des mesures physiques et les limites de la connaissance expérimentale

Fabien GRÉGIS

Univ. Paris Diderot, Sorbonne Paris Cité,
laboratoire Sphere UMR 7219, F-75205 Paris, France
fabien.gregis@etu.univ-paris-diderot.fr

Thèse dirigée par Nadine de Courtenay et Olivier Darrigol
Soutenue publiquement le 25 mars 2016 à l'Université Paris Diderot

Résumé

Cette thèse est structurée en trois parties, précédées d'un bref chapitre d'exposition, le chapitre 1 : « Préambule : erreur et incertitude ». Ce préambule présente la nature de mon objet d'étude en introduisant les concepts qui reviendront de façon récurrente dans les différentes parties : erreur et incertitude de mesure, grandeur physique et « mesurande », modèle et fonction de mesure. Ce chapitre explicite également le contexte dans lequel je me place, celui des développements récents que la métrologie a connus sous l'impulsion des institutions nationales et internationales de métrologie.

La première partie est consacrée au traitement probabiliste de l'incertitude de mesure. Elle se situe d'emblée dans une perspective contemporaine et sous l'angle des fondements techniques de l'analyse d'incertitude, et montre que les développements techniques des métrologues ont peu à peu fait émerger des problématiques philosophiques. Dans la seconde partie, j'examine et je remets partiellement en question l'élaboration philosophique des métrologues dans leurs ouvrages techniques. Cette partie est centrée autour d'un concept particulièrement controversé en métrologie : celui de « valeur vraie » d'une grandeur physique. La troisième partie se tourne vers un contexte plus spécifique qui permet d'illustrer certains usages de l'incertitude de mesure. Ce faisant, elle sort du domaine exclusif de la métrologie et étend le cadre des deux premières parties, pour aborder la question de l'incertitude de mesure sous l'angle de

la physique de précision, et plus précisément de l'activité des « ajustements » des constantes de la physique. Ce changement de perspective me permet d'enrichir le spectre des réponses que l'on peut apporter aux problématiques développées jusqu'alors.

En addition des trois parties qui constituent le cœur de ma thèse, un dernier chapitre, le chapitre 12 « Épilogue : la genèse du GUM », retrace brièvement, dans une perspective plus historique, les étapes de la rédaction du *Guide pour l'expression de l'incertitude de mesure* (GUM) et, à travers elles, les questions institutionnelles et conceptuelles qui se sont posées aux métrologues dans le courant des dernières décennies du XX^e siècle. Elle montre en particulier que la rédaction d'un tel document international est guidée par des influences à deux niveaux. D'une part, le document porte la marque des institutions internationales qui pilotent sa conception, comme l'ISO ou le BIPM. D'autre part, sa rédaction a été débloquée à plusieurs reprises grâce à des initiatives à l'échelle individuelle.

Première partie : l'interprétation probabiliste de l'incertitude de mesure

La première partie vise à aborder le concept d'incertitude de mesure et ses ramifications philosophiques, à partir de la façon dont le GUM tente de le rationaliser. L'objectif est d'inspecter la structure technique du concept, pour identifier la façon dont certains questionnements techniques font émerger des problématiques d'ordre philosophique. Mon étude se focalise sur la nature et la fonction des modèles statistiques d'analyse d'incertitude. Je montre en particulier que la métrologie contemporaine est traversée par des questionnements très ouverts sur la nature des probabilités à employer dans ces modèles, et que l'adhésion traditionnelle à une interprétation fréquentiste des probabilités est sérieusement remise en question depuis le début des années 1970. À contre-courant de la conception traditionnelle, les textes contemporains de métrologie proposent de s'appuyer sur une interprétation des probabilités dite « épistémique ». Un nombre grandissant d'acteurs proposent même d'aller plus loin encore et de développer une approche pleinement bayésienne de la mesure. Ce mouvement accompagne une évolution notable dans la conception de la mesure et de la science au XX^e siècle, que l'on peut qualifier de « tournant épistémique » en métrologie. Certains métrologues soulignent tout particulièrement la subjectivité selon eux inhérente à l'activité de mesure, ce qui va à rebours de la conception classique de la mesure comme outil objectif par excellence d'enquête sur le monde physique. Ce point a engendré des discussions nourries en métrologie, et je m'attelle à comprendre les raisons pour lesquelles les métrologues en sont venus à critiquer la conception de la mesure qui était en vigueur au milieu du XX^e siècle, allant jusqu'à chercher à développer des alternatives qui leurs semblent préférables.

Le chapitre 2 : « Fondements généraux de l'analyse d'incertitude » présente les éléments communs qui servent de fondement à l'analyse d'incertitude et à l'introduction des probabilités et des modèles statistiques pour son calcul. Le chapitre discute en particulier la dichotomie entre erreurs dites « aléatoires » et erreurs dites « systématiques » de mesure qui en est l'un des fondements les plus classiques.

Cela me permet d'attaquer ensuite de front le modèle fréquentiste de l'analyse d'incertitude, objet central du chapitre 3 : « L'approche fréquentiste de la mesure ». L'un des points essentiels de mon analyse consiste à identifier quel est *l'objet* des probabilités au sein du mo-

dèle, et à comprendre sur quelles hypothèses fondamentales ce modèle s'appuie. Je décris alors pourquoi le calcul fréquentiste des incertitudes de mesure ne permet pas de rendre compte de toutes les composantes d'erreur de mesure, ce qui a été perçu comme une limite sérieuse par les métrologues à la fin des années 1970.

L'utilisation de probabilités épistémiques en métrologie vise à pallier ces limites, et pousse au développement d'une conceptualisation bayésienne de la mesure. C'est ce que je décris dans le chapitre 4 : « Des probabilités épistémiques à une modélisation bayésienne », qui procède en deux étapes. Je montre d'abord comment les probabilités épistémiques viennent répondre à un problème spécifique posé par l'approche traditionnelle, à savoir que cette dernière ne permet pas d'effectuer un traitement probabiliste des erreurs « systématiques » de mesure. Dans un deuxième temps, j'explique comment certains métrologues et statisticiens ont proposé d'étendre cette solution aux erreurs « aléatoires », de façon à aboutir à une approche entièrement bayésienne de la mesure, dont certaines formulations assument explicitement une épistémologie subjectiviste.

Les chapitres 3 et 4 montrent que les deux approches archétypales qui y sont étudiées, fréquentiste et bayésienne, amènent à interpréter l'incertitude de mesure de façon sensiblement différente. L'approche fréquentiste insiste sur l'objectivité de la mesure et est centrée sur un idéal d'« exactitude », l'exactitude étant définie en métrologie comme une indication sur la proximité entre le résultat de mesure obtenu et la « valeur vraie » de la grandeur que l'on souhaite mesurer – c'est-à-dire qu'elle traduit la tendance d'un processus de mesure à produire des erreurs de mesure faibles. L'approche bayésienne insiste au contraire sur la subjectivité de l'activité de mesure, et engage un idéal de rationalité orienté vers la recherche de la meilleure expression possible d'un état présent de connaissance. Cette opposition engendre de nombreuses discussions dans la littérature spécialisée, que j'explore dans le chapitre 5, « Discussion : les ramifications philosophiques du débat statistique ». Cela m'amène à soulever deux questions particulières. La première porte sur le statut de la valeur vraie d'une grandeur dans l'approche contemporaine de la mesure. La seconde porte sur le dialogue entre deux concepts essentiels de la mesure, celui d'« incertitude de mesure » d'une part, et celui d'« exactitude de mesure » d'autre part. Ces deux questions annoncent les deux parties suivantes de mon travail.

Seconde partie : la « valeur vraie » d'une grandeur

La seconde partie est consacrée à la notion de « valeur vraie » d'une grandeur physique, intuitivement définie comme la valeur que l'on obtiendrait d'une mesure parfaite, c'est-à-dire d'une mesure qui n'est affectée d'aucune erreur. Cette définition fait cependant apparaître une circularité, puisque l'erreur de mesure est justement définie comme une déviation à la valeur vraie. De fait, le problème de la « valeur vraie » prend racine dans la question même des fondations de la mesure, et mène à interroger la possibilité de décrire des grandeurs physiques par des valeurs numériques. Ce dernier point est l'objet de développements classiques en philosophie, qui regroupent un large spectre de points de vue, allant d'un réalisme fort qui voit le monde comme quantitatif par nature à des positionnements anti-réalistes ou cohérentistes. L'une des approches les plus classiques est la théorie représentationnelle de la mesure, qui, bien que largement débattue, permet de discuter des conditions auxquelles l'usage des mathématiques dans les sciences empiriques a un sens, tout en rappelant que la notion de valeur

d'une grandeur est un élément de *représentation* des phénomènes physiques. Sans s'engager dans des développements purement philosophiques, les métrologues ont quant à eux critiqué à plusieurs reprises le concept de valeur vraie d'une grandeur physique. Cette critique se retrouve dans la littérature spécialisée où l'on discerne plusieurs tentatives de faire disparaître le concept de valeur vraie hors du formalisme de l'analyse d'incertitude. J'affirme y distinguer une idée sous-jacente qui consiste à affirmer qu'en supprimant l'emploi d'un tel concept, on vient simplifier les conceptions des métrologues – et des scientifiques en général – en évitant les questions métaphysiques et les problèmes considérés comme purement philosophiques. Je distingue deux types d'arguments que j'étudie tour à tour : un argument d'« inconnaissabilité » et un argument de « non-unicité ». Malgré les difficultés réelles qu'introduisent ces deux arguments, je défends qu'il n'y a pas lieu pour autant d'abandonner le concept lui-même.

Dans le chapitre 6 : « L'argument d'« inconnaissabilité » de la valeur vraie », j'aborde la première critique de la valeur vraie. Celle-ci consiste à voir dans la valeur vraie d'une grandeur un idéal à jamais inconnaissable, qui constitue dès lors un objectif illusoire dont il est préférable de se passer. J'oppose deux réponses à cet argument. Dans un premier temps, à partir des développements de la première partie de ma thèse, j'affirme que les modèles principaux de l'analyse d'incertitude, dont l'approche bayésienne, font usage de la notion de valeur vraie dans leur formalisme technique. Par conséquent, s'il existe une critique valide concernant la valeur vraie d'une grandeur dans le cadre de ces modèles statistiques, celle-ci ne peut pas porter sur l'usage du concept mais sur son interprétation. C'est pourquoi, dans un second temps, j'analyse les interprétations possibles que l'on peut faire du concept de valeur vraie et en particulier de son rapport à la vérité scientifique. Je conclus que l'argument d'« inconnaissabilité » n'a pas la force que les métrologues peuvent lui prêter, et je décris la critique de ces mêmes métrologues comme l'expression sous-jacente d'une position essentiellement anti-métaphysique qui n'est pas pour autant anti-réaliste. En fin de compte, le problème philosophique général ne peut pas être résolu simplement, tant il renvoie à des questionnements généraux et déjà anciens en philosophie, qui ne se limitent pas à la question de l'incertitude de mesure ou de la mesure. Il demeure cependant intéressant d'étudier les conséquences que peuvent avoir un tel positionnement philosophique sur les pratiques des scientifiques. À cet égard, l'attachement au concept de « valeur vraie » présente des vertus épistémiques, parmi lesquelles celle de donner un cadre conceptuel au maintien d'un processus permanent de correction qui guide le progrès expérimental.

Le chapitre 7 : « Valeur vraie : non unicité et incertitude définitionnelle » est tourné vers le second argument porté contre le concept de valeur vraie d'une grandeur. Cet argument part du constat que les cibles des mesures, les « mesurandes », sont des concepts vagues qui reposent sur une représentation idéalisée du processus de mesure, laquelle occulte beaucoup de détails dans le comportement des objets physiques et des phénomènes physiques décrits. De fait, pour un mesurande donné, il n'est que très rarement possible de concevoir une valeur vraie unique, parce que la définition du mesurande n'est elle-même pas assez précise pour cela. Mais alors, si une valeur n'est pas unique, comment peut-elle seulement être dite « vraie » ? Ici encore, l'argument renvoie à la question de la possibilité d'utiliser les mathématiques pour représenter des phénomènes physiques, et plus spécifiquement de la possibilité d'utiliser des nombres pour

décrire des grandeurs physiques. Ma réflexion ébauche une réponse qui empreinte à la question du réductionnisme en science et à l'épistémologie de l'approximation. Je défends que la « valeur vraie » d'une grandeur conserve une légitimité, à condition de reconnaître le caractère approché de la connaissance scientifique et des théories scientifiques (et donc les lois formulées à l'intérieur de ces théories). On préférera alors parler de « valeur approximativement vraie ». Dans ce cadre, l'« incertitude définitionnelle », que les métrologues ont introduite pour rendre compte de la non-unicité des valeurs des grandeurs physiques, apparaît non comme une limite de connaissance à propos de la cible de la mesure, mais comme une estimation de la limite de précision des modèles théoriques dans lesquels les grandeurs que l'on souhaite mesurer sont mobilisées.

Troisième partie : les ajustements des constantes de la physique

La troisième partie de ce travail vise à donner une perspective supplémentaire à mon étude, en cherchant dans la physique de précision des réponses nouvelles aux questions qui ont été posées dans les parties précédentes. Je propose une étude de la pratique dite d'« ajustements des constantes de la physique », qui prend forme en 1929 avec les travaux pionniers de Raymond T. Birge, et qui perdure encore aujourd'hui sous l'égide d'une institution internationale, le CODATA (Committee on Data for Science and Technology). Ma réflexion dans cette partie s'appuie sur deux questions de travail : comment peut-on combiner entre eux différents résultats de mesure obtenus dans des conditions différentes ? Comment les scientifiques peuvent-ils s'accorder sur la valeur d'une grandeur ? Cela m'amène à m'interroger sur l'usage qui est fait de l'incertitude de mesure au sein de la pratique spécifique des ajustements. Je conclus en revenant sur la question essentielle du rapport entre incertitude de mesure et exactitude de mesure, qui avait été soulevée dès la première partie de ma thèse.

Le chapitre 8 : « Combinaison des résultats et ajustements aux moindres carrés » commence par rappeler un épisode de l'histoire de l'astronomie au XVIII^e siècle lors duquel les savants ont été amenés à accepter l'idée d'une « combinaison des observations ». Cette expression, due à Stephen Stigler, exprime le fait que des résultats de mesure ne sont plus considérés individuellement, indépendamment les uns des autres, mais que les savants reconnaissent au contraire qu'ils peuvent utiliser conjointement des résultats de mesure obtenus par diverses personnes dans des conditions différentes, et à propos de phénomènes différents. À travers cet épisode, j'identifie les fondements d'une idée de *reproductibilité* d'une mesure, attachée à un examen attentif des sources d'erreurs, ce qui suggère en filigrane le début d'un calcul d'incertitude. Cet épisode marque aussi les débuts d'un traitement statistique des erreurs qui sera ensuite systématisé par Legendre et Gauss avec la « méthode des moindres carrés », essentielle à la pratique des ajustements.

Dans le chapitre 9, « L'ajustement des constantes de la physique : l'initiative de Birge », j'attaque le sujet des ajustements proprement dits au cours du XX^e siècle. Les ajustements des constantes de la physique sont une pratique qui consiste à collecter les meilleurs résultats connus à un moment donné pour les mesures d'un grand nombre de constantes physiques (souvent appelées « constantes fondamentales ») qui sont reliées entre elles par un réseau d'équations qui découlent elles-mêmes du cadre théorique adopté. L'ajustement se conclut par une

mise en commun de ces valeurs empiriques afin de produire un ensemble de « meilleures valeurs » ou de « valeur recommandées » de ces constantes valable jusqu'à ce que de nouvelles mesures soient effectuées. J'étudie dans ce chapitre le travail de Birge dans ses articles de 1929 et 1932, dont la « combinaison des observations » décrite au chapitre 8 est un pivot essentiel. Je montre le rôle qui est accordé à l'incertitude de mesure dans la comparaison et la combinaison des résultats de mesure. L'incertitude de mesure apparaît comme un outil central pour quantifier et traiter mathématiquement le désaccord ou l'accord à l'intérieur d'une collection de résultats expérimentaux caractérisant des grandeurs reliées théoriquement entre elles par un système d'équations. Cependant, elle ne peut pas être simplement interprétée ni comme une estimation de l'amplitude de l'erreur de mesure attachée au résultat de mesure, ni comme une évaluation de la confiance que l'expérimentateur accorde au résultat en question.

Je continue l'examen des ajustements au chapitre 10 : « Le traitement des données discordantes dans les ajustements », en m'intéressant à la période de trente ans qui suit l'initiative de Birge. Cette période s'achève en 1970 environ, lorsque des physiciens, des métrologues et des statisticiens se concertent autour de la question de la pratique des ajustements, de sa légitimité, et du problème plus spécifique du traitement des données discordantes. Cette discussion fait apparaître différentes positions des acteurs quant au statut de l'incertitude de mesure. Une approche « conservatrice » affirme qu'il faut faire en sorte de s'assurer que les résultats de mesure soient les plus exacts possibles à tout instant, même si cela doit se faire au prix d'une incertitude plus grande. Une seconde position, défendue par des acteurs influents du domaine, considère qu'il est au contraire préférable de proposer des résultats qui soient les plus précis possibles, quitte à ce que ces derniers se révèlent inexacts. Cette seconde position repose sur l'idée que seuls des résultats précis peuvent permettre de tester efficacement la cohérence et la validité des théories scientifiques. De fait, cette position est dirigée vers un progrès à *long terme* et se concentre non pas sur l'exactitude immédiate des résultats de mesure, mais sur la possibilité d'un *progrès* scientifique qui s'appuie sur un processus sans fin d'identification et de correction des erreurs de mesure. Le CODATA conserve aujourd'hui encore la trace de cette seconde position, comme le montre un exemple contemporain portant sur un sujet similaire, à propos de la mesure du rayon du proton.

Le chapitre 11 : « Incertitude et exactitude » vient alors conclure la dernière partie de ce travail en revenant de façon plus systématique sur le rapport entre incertitude de mesure et exactitude de mesure. J'affirme que la position de progrès à long terme décrite au chapitre précédent permet de concilier les deux conceptions de la mesure exposées dans la première partie de cette thèse, à savoir l'approche objectiviste-fréquentiste et l'approche épistémique-bayésienne, en défendant l'idée que ces deux conceptions sont partiellement complémentaires, plutôt qu'exclusives l'une de l'autre. Je défends aussi que l'approche à long terme correspond à une perspective de recherche, qui voit l'exactitude de mesure comme un concept non pas statique mais dynamique, tourné vers un progrès futur, mais qui ne couvre pas les besoins de tous les utilisateurs potentiels de l'activité de mesure. En particulier, dans des perspectives plus appliquées (ingénierie, industrie, santé publique, etc.), l'exactitude des mesures devient un enjeu beaucoup plus immédiat. Les acteurs, amenés à réfléchir en termes de risque, ne peuvent alors pas se contenter de la promesse d'un progrès hypothétique dans le futur.