Reproductibilité

Anne Cadiou

Laboratoire de Mécanique des Fluides et d'Acoustique

Ateliers et séminaires pour l'informatique et le calcul scientifique
PMCS2I - LMFA
Vendredi 5 octobre 2018

Définition

Wikidépia

La **reproductibilité** *d'une expérience scientifique est* **une des conditions** *qui permettent d'***inclure les observations** *réalisées durant cette expérience dans le processus d'amélioration perpétuelle des connaissances scientifiques. Cette condition part du principe qu'on ne peut tirer de conclusions que d'un événement bien décrit, qui est apparu plusieurs fois, provoqué par des personnes différentes.*

Partie intégrante de la démarche expérimentale

- Notions connexes : répétabilité, réplicabilité, ...
- Critère associé au **caractère scientifique** d'une étude
- **Variabilité** des résultats évaluée par les statistiques et la probabilité

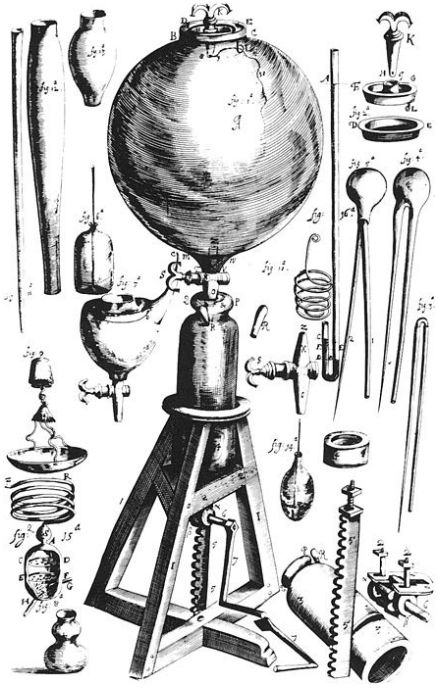
Toutefois

- **Réfuter des hypothèses** nécessite de **varier les expériences** !

Historique

Concept né de l'étude du vide en physique, dont l'existence est discutée au XVII^{e} siècle par les philosophes René Descartes et **Thomas Hobbes**.

- **Robert Boyle** et **Robert Hooke** construisent en 1659 une expérience afin de le mettre en évidence et de l'étudier.
- De son côté **Christiaan Huygens** met également au point en 1662 une pompe à air et produit un fait, en observant une *suspension anormale* de l'eau purgée.
- Invité en 1663 en Angleterre, Christian Huygens pu guider Robbert Hooke pour reproduire cette observation dans leur expérience.



Expériences réalisées en public, protocoles et observations étaient consignés dans des cahiers afin de faciliter leur reproduction.

Historique

- Science expérimentale formalisée au XIX^{e} siècle par le chimiste **Michel-Eugène Chevreul** et le médecin **Claude Bernard**.
- Le philosophe des sciences **Karl Popper** explicite en 1934 le critère de *réfutabilité* comme partie intégrante de la démarche expérimentale. La démarche expérimentale visera donc essentiellement à réfuter une théorie, pas à la vérifier.
- Le statisticien **Ronald Fisher** décrit en 1935 le principe du *plan d'expériences* suivant lequel, un phénomène est expérimentalement démontrable lorsqu'il peut être reproduit de façon statistiquement significative.

Cela pose les bases d'un dogme répandu en science moderne, selon lequel la reproductibilité d'une expérience est une condition nécessaire (mais pas suffisante) pour établir un fait scientifique.

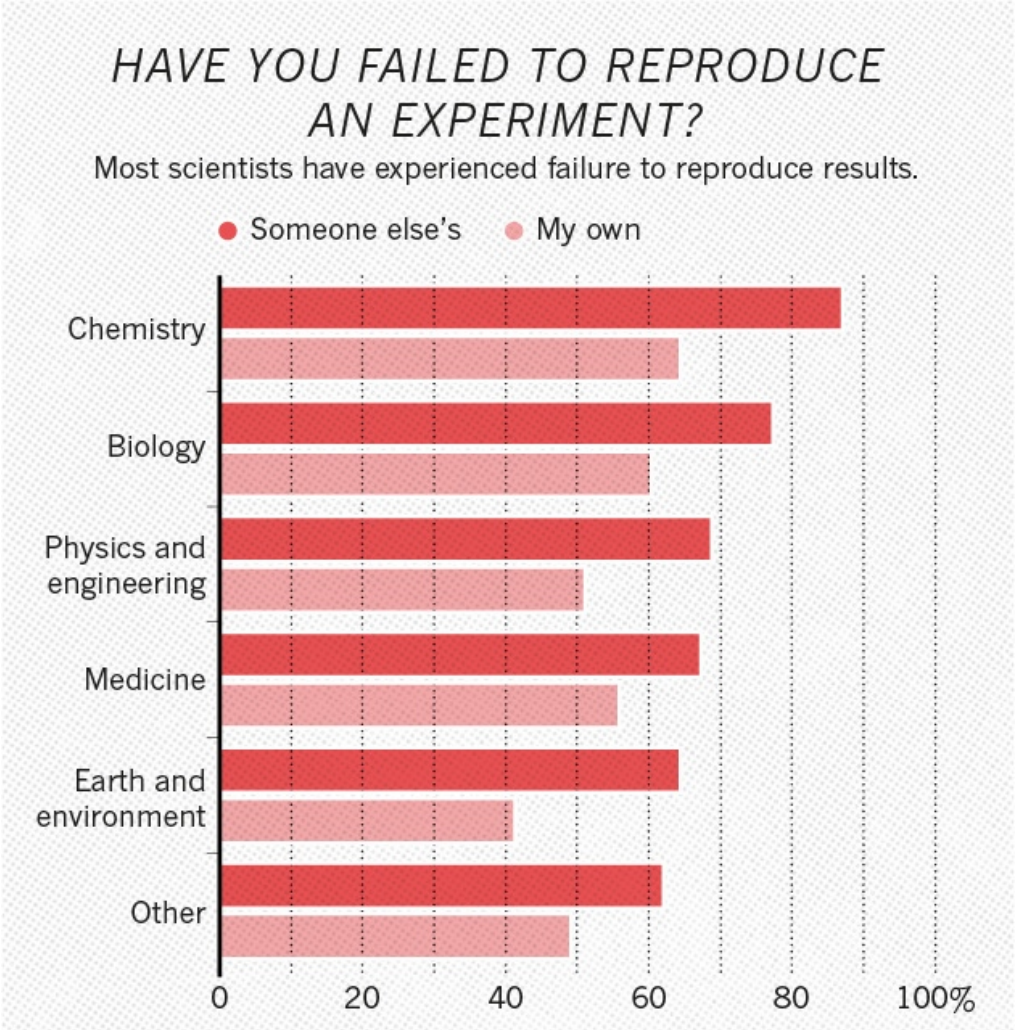
Mais en pratique ?

Difficultés à reproduire les expériences de façon statistiquement significative :

- trop longues (expérience goutte de poix de Thomas Parnell, 1927)
- trop chères (toxicité à long terme du maïs génétiquement modifié roundup ready de Gilles-Éric Séralini, 2012)
- trop dépendantes d'une technicité métier (contribution de la cristallographe Roslind Franklin à la découverte de la structure en double hélice de l'ADN, 1953)
- etc.

Accentué par **la complexité croissante** des objets d'études (multiples interactions) et des bancs d'expérimentation ; ... par la course à la publication ?

Crise de reproductibilité ?



(M. Baker, Nature, vol. 533, 2016)

Mouvement vers la Recherche Reproductible

Basé sur l'idée que le produit de la recherche académique est la publication scientifique, augmentée de tout ce qui a été nécessaire pour la réaliser (cahier d'expérimentation, données, script et codes de calcul, environnement informatique, etc.) (replicabilité)

Initiatives pour encourager la réplication :
The ReScience Journal (K. Hinsen, N.P. Rougier, ed.)

Émergence de Workshops
testant la reproductibilité dans différentes communautés de recherche

- Reproductibility Project, psychologie (2015) ;
- ReproHackathon, bioinformatique (2017) ;
- Reproducible, Customizable and Portable Workflows for HPC (2018) ;
- etc.

Actualité

Regain d'intérêt dans *plusieurs disciplines*, en association avec une réflexion plus générale sur les objets, les méthodes et les pratiques de recherche.

- Reproductible Research
- Mouvements Open Science / Open Data
- Développement de plateformes et d'outils pour l'expérimentation
- Archivage et pérennisation (figshare, Zenodo, RunMyCode, etc.)
- Principes FAIR (Findable, Accessible, Interoperable, Reusable) (H2020)

Challenges pour la recherche reproductible

- **Contrôler** chaque étape de l'expérimentation, de l'analyse
- Systématiser et automatiser la **description** de toutes les étapes (sources, environnement, traces, etc.)
- **Évaluer** la reproductibilité (partielle ou totale), l'exhaustivité des données de l'expérimentation
- **Assurer, pérenniser** l'accès aux données

Challenges associés au calcul numérique

La **reproductibilité numérique** implique la notion de résultat identique en virgule flottante (avec la même performance)

Difficultés :

- Erreurs d'arrondi dans la représentation des nombres en virgule flottante (standard IEEE 754)
- Accumulation des erreurs de troncatures
- Opérations en virgule flottantes $(+, \times)$ commutatives mais pas associatives $x + (y + z) \neq (x + y) + z$
- HPC calcul exaflopique : 10^{18} opérations par seconde (10^{18} erreurs par seconde)
- Erreur dépend de l'ordre des opérations (problème de réduction en parallèle)
- Dépend de la précision de la machine
- Dépend du sysème d'exploitation, de l'environnement, des versions de bibliothèques, du réseau, etc.

Pourquoi travailler de façon reproductible ?

En pratique, cela sert à

- Valider scientifiquement son expérience numérique

  sans en garantir l'exactitude, la justesse ou la pertinence
- Mettre à l'épreuve ses hypothèses
- Faciliter la publication de ses résultats
- Assurer une continuité dans son travail

Concernant le calcul numérique et les simulations,
À la **reproductibilité** et la **replicabilité** d'ajoute la **réutilisabilité**

Deviendra une obligation ?

- Politique de la National Science Foundation (US)
  - Règles de diffusion et partage des résultats de recherche
  - Exigences de DMP (Data Management Plan)
- Politique de la Commission Européenne
  - H2020 Open Research Data Pilot
  - *Guidelines on Open Access to Scientific Publications and Research Data in Horizon 2020* (2013)

R^5

(K. Hinsen, ReScience, 2017)

- Rerunnable (robust)
  - Can you re-run your program ?
  - One day, one week, one month, one year (just kidding) apart ?
- Repeatable (defend)
  - Can you re-run your program and get same results ?
  - Did you save everything, including random seed ?
- Replicable (certify)
  - Can someone reimplement your model and get same results ?
  - Did you describe everything ?
- Reproducible (compare)
  - Can someone re-run your program and get same results ?
  - Did you save the software stack ?
- Reusable (transfer)
  - Can someone reuse your program using different data ?
  - Is your software data-dependent ?

Comment faire ?

Maîtriser les **variabilités numériques** en **identifiant** des différentes sources d'incertitudes et pouvoir les **quantifier**

Améliorer la **description et le contrôle** de l'expérimentation

Expérimentation numérique ?

- Protocole
  - Ensemble d'instructions pour produire les résultats de simulations numériques
- Entrée
  - Données
  - Codes
  - Environnement
- Sortie
  - Données
  - Tracés, post-traitements, codes
  - Article

Environnement ?

Définition

Tout ce qu'il faut pour réaliser un protocole d'expérimentation numérique

Software

- Code
- Système d'exploitation
- Bibliothèques scientifiques

Hardware

- Machine (physique ou virtuelle)
- Réseau
- Stockage

Interactif

À éviter (sans filet).

En ligne de commande

Tout peut être facilement consigné
À gérer (environnements multiples, interconnectés, etc.)

Interfaces graphiques

Plus difficile à consigner
Indissociable du logiciel utilisé

Solutions :
Exploiter les nombreux outils existants pour automatiser chaque étape de l'expérimentation numérique

Comment choisir ?

Qu'est ce qui évolue le plus au cours de l'expérimentation ?

- Les paramètres du calcul
- Les données d'entrée
- Le code

Comment sont les données ?

- Des petits fichiers locaux
- Des grands fichiers locaux
- Des données sur une base distante partagée
- Des données réparties

De quel type de simulation s'agit-il ?

- Rapide, réalisable partout, pouvant être répété aisément
- Long et lourd
- Ne pouvant être fait que très rarement (ressources limitées ou dédiées)

Exemples d'outils

Le code évolue en même temps qu'avance l'expérience

`make`, `cmake`, `ctest`, `bash`, `python`, etc. } ⇒ intégration continue

Expérimentation exploratoire

- Notebook interactifs qui enregistrent les résultats (`Jupyter`)
- Cahier de notes (`Org-mode`)

Enregistrer et tracer les étapes

- `Vistrails`, `Galaxy`, `Taverna` : gestionnaire de flux de travail (workflow)
- `VSR` : enregistrement en temps-réel des actions effectuées
- `git` et `Org-mode`, `Sweace`, `knitR` : programmation lettrée
- conteneurs (`Puppet`, `Docker`, `VirtualBox`, `Kameleon`, `Guix`, etc.)
- plateformes (`RunMyCode.org`, `ResearchCompendia.org`, etc.)
- paramètre et préserve la trace d'exécutions multiples (in-)dépendantes
  - `Sumatra`, `Jube`, `OpenMole`
  - `Execo`, `XPFlow`, `ActivePapers`

DSL/API

Nombreux outils, différentes motivations, différentes approches
Deux grandes catégories :

DSL **D**omain **S**pecific **L**anguage (par opposition au **GPL** (**G**eneral **P**urpose **L**anguage)) Outils exploitant un langage spécifique pour une application ciblée ; apprentissage aisé ; usage restreint aux fonctions prévues

API **A**pplication **P**rogramming **I**nterface : ensemble de classes, méthodes ou fonctions qui sert d'interface à d'autres applications (bibliothèques, service web, etc.) ; prise en main plus longue ; usage plus facilement extensible

Archivage

Article : ArXiv, HAL (doi)

Data : Zenodo (CERN), ISAAC (CINES), figshare (cloud) (doi)

Software : encore à dévlopper (forges découplées des environnements d'exécution)

Références

Largement inspiré des présentations de l'École Thématique PRECIS (2017).

- `https://en.wikipedia.org/wiki/Reproducibility`
- S. Sapin, S. Schaffer, *Leviathan and the Air-Pump : Hobbes, Boyle, and the Experimental Life*, 1985
- M. Baker, *Is there a reproducibility crisis ?*, Nature vol. 533, 2016
- École PRECIS (2017)

  `https://precis.sciencesconf.org/`
- K. Hinsen, *ReScience*, 2017
- G. Varoquaux, `http://gael-varoquaux.info/`
- L. Nussbaum, *Reproducible Research in Computer Science*, 2015