

---

# Comment juger les approximations? Pièges des statistiques.

Barolomeo Civalleri<sup>1</sup>, Roberto Dovesi<sup>1</sup>, Erin Johnson<sup>2</sup>, Davide Presti<sup>3</sup>, and Andreas Savin<sup>\*†4</sup>

<sup>1</sup>University of Torino, Torino – Italie

<sup>2</sup>University of California, Merced – États-Unis

<sup>3</sup>University of Modena and Reggio-Emilia, Modena – Italie

<sup>4</sup>Laboratoire de chimie théorique (LCT) – CNRS : UMR7616, Université Pierre et Marie Curie (UPMC) - Paris VI, Université Pierre et Marie Curie [UPMC] - Paris VI – Tour 12-13, 4ème étage, 4 place Jussieu, Case Courrier 137, 75252 PARIS, France

## Résumé

Plusieurs facteurs ont contribué à l'essor rapide des calculs de chimie quantique : le gain de puissance des ordinateurs, l'amélioration des algorithmes, et également la proposition de nouvelles méthodes, telles que les fonctionnelles de la densité approchées. Pour juger la qualité de ces dernières, ainsi que pour leur développement, de grandes bases de données sont utilisées. Dans ce cas, l'utilisation de statistiques devient un outil précieux, voire même irremplaçable. Cette présentation met en évidence quelques écueils de ces analyses statistiques [1,2]. On montre que leurs conclusions peuvent être affectées par des facteurs bien connus dans d'autres domaines scientifiques. Par exemple, le choix de la "meilleure" fonctionnelle peut dépendre du critère utilisé, ou de l'échantillonnage. Même avec de très bonnes approximations, le risque de publier des résultats inexacts augmente de façon naturelle avec leur nombre.

B. Civalleri, D. Presti, R. Dovesi, A. Savin, *Chemical Modelling*. 9, 168–185 (2012).

E.R. Johnson, A. Savin, *Topics in Current Chemistry*, soumis pour publication.

**Mots-Clés:** DFT, bases de données, statistique

---

\*Intervenant

†Auteur correspondant: andreas.savin@lct.jussieu.fr