


Empreinte environnementale des outils bioinformatiques

The Carbon Footprint of Bioinformatics

Jason Grealey,^{*,†,1,2} Loïc Lannelongue ^{†,3,4,5} Woei-Yuh Saw,¹ Jonathan Marten,^{‡,4} Guillaume Méric,^{1,6}
Sergio Ruiz-Carmona,¹ and Michael Inouye^{*,1,3,4,5,7,8}

© The Author(s) 2022. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

Mol. Biol. Evol. 39(3):msac034 doi:10.1093/molbev/msac034 Advance Access publication February 10, 2022

1

Green Algorithms

How green are your computations?

Details about your algorithm

To understand how each parameter impacts your carbon footprint, check out the formula below and the [methods article](#)

Runtime (HH:MM)

12

0

Type of cores

CPU

Number of cores

12

Model

Xeon E5-2683 v4

Memory available (in GB)

64

Select the platform used for the computations

Local server



253.64 g CO₂e

Carbon footprint



2.28 kWh

Energy needed



0.28 tree-months

Carbon sequestration



1.45 km

in a passenger car



0.51%

of a flight Paris-London

Share your results with [this link!](#)

Select the platform used for the computations

Local server

Select location

Europe

Austria

Do you know the real usage factor of your CPU?

Yes No

Do you know the Power Usage Efficiency (PUE) of your local data centre?

Yes No

Do you want to use a Pragmatic Scaling Factor?

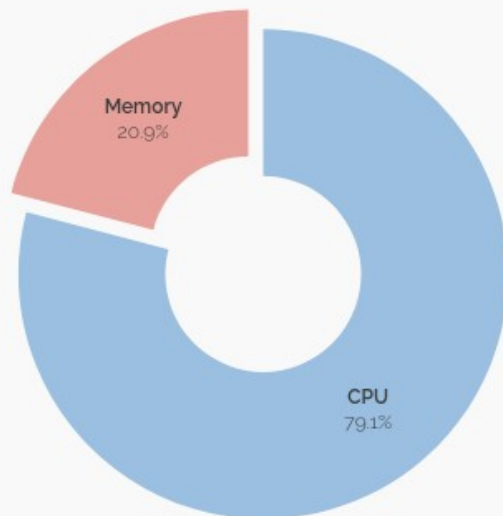
Yes No

[Reset](#)

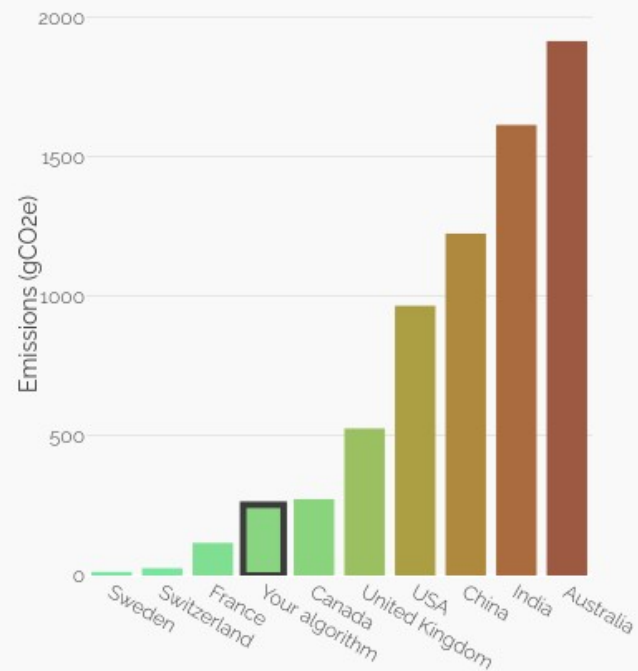
[Change app version](#)

Share your results with [this link!](#)

Computing cores VS Memory

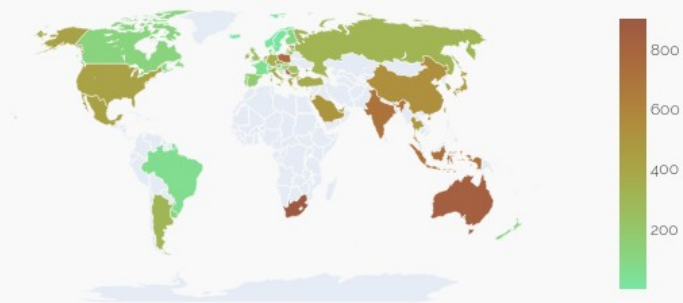


How the location impacts your footprint



More details about the methodology in the [methods paper](#)

Carbon Intensity across the world



About CO₂e

"Carbon dioxide equivalent" (CO₂e) measures the global warming potential of a mixture of greenhouse gases. **It represents the quantity of CO₂ that would have the same impact on global warming** as the mix of interest and is used as a standardised unit to assess the environmental impact of human activities.

What is a tree-month?

It's the amount of CO₂ sequestered by a tree in a month. **We use it to measure how long it would take to a mature tree to absorb the CO₂ emitted by an algorithm.** We use the value of 11 kg CO₂/year, which is roughly 1kg CO₂/month.

What can you do about it?

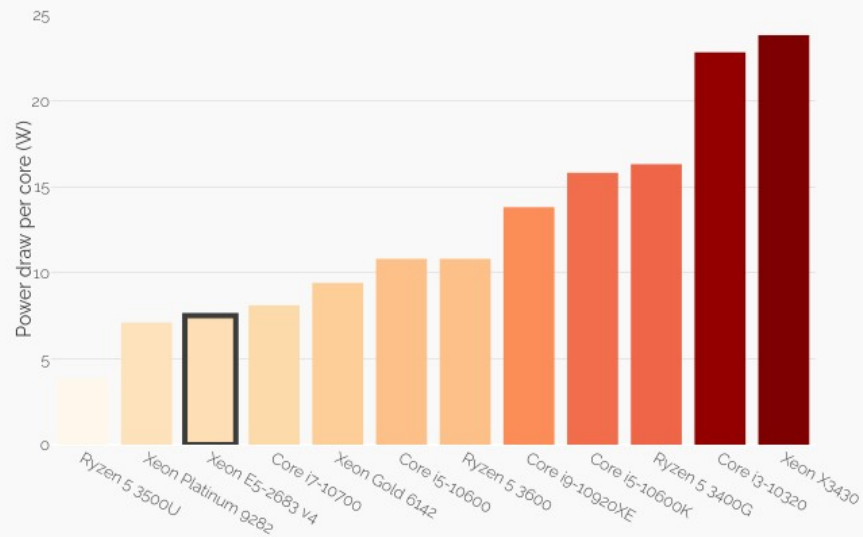
The main factor impacting your footprint is the location of your servers: the same algorithm will emit **74 times more** CO₂e if ran in Australia compared to Switzerland. Although it's not always the case, many cloud providers offer the option to select a data centre.

Memory power draw is a huge source of waste, because **the energy consumption depends on the memory available, not the actual usage**. Only requesting the needed memory is a painless way to reduce greenhouse gas emissions.

Generally, taking the time to write optimised code that runs faster with fewer resources saves both money and the planet.

And above all, **only run jobs that you need!**

Power draw of different processors



The formula

The carbon footprint is calculated by estimating the energy draw of the algorithm and the carbon intensity of producing this energy at a given location:

$$\text{carbon footprint} = \text{energy needed} * \text{carbon intensity}$$

Where the energy needed is:

$$\text{runtime} * (\text{power draw for cores} * \text{usage} + \text{power draw for memory}) * \text{PUE} * \text{PSF}$$

The power draw for the computing cores depends on the model and number of cores, while the memory power draw only depends on the size of memory available. The usage factor corrects for the real core usage (default is 1, i.e. full usage). The PUE (Power Usage Effectiveness) measures how much extra energy is needed to operate the data centre (cooling, lighting etc.). The PSF (Pragmatic Scaling Factor) is used to take into account multiple identical runs (e.g. for testing or optimisation).

The Carbon Intensity depends on the location and the technologies used to produce electricity. But note that **the "energy needed" indicated at the top of this page is independent of the location.**

How to report it?

It's important to track the impact of computational research on climate change in order to stimulate greener algorithms. For that, **we believe that the carbon footprint of a project should be reported on publications alongside other performance metrics.**

Here is a text you can include in your paper:

This algorithm runs in 12h on 12 CPUs Xeon E5-2683 v4, and draws 2.28 kWh. Based in Austria, this has a carbon footprint of 253.64 g CO₂e, which is equivalent to 0.28 tree-months (calculated using green-algorithms.org v2.2 [1]).

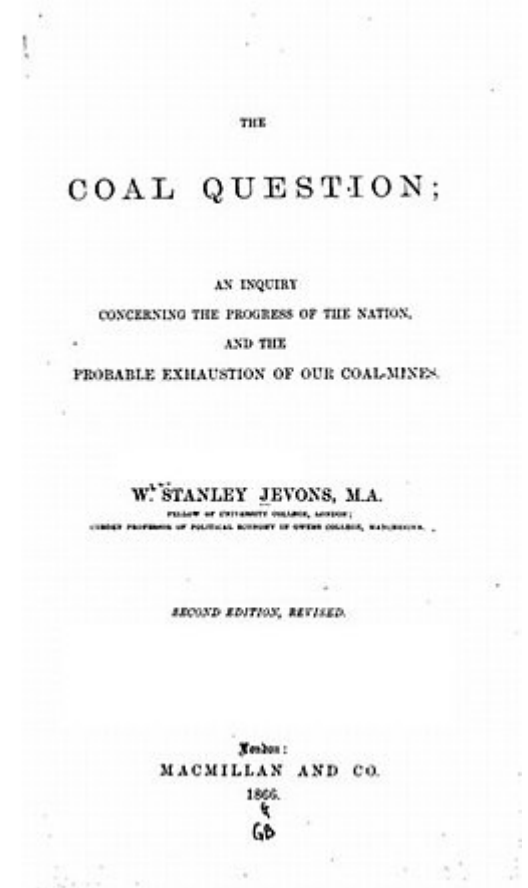
[1] Lannelongue, L., Grealey, J., Inouye, M., Green Algorithms: Quantifying the Carbon Footprint of Computation. Adv. Sci. 2021, 2100707.

Including the version of the tool is useful to keep track of the version of the data used.

Les effets rebond de l'optimisation

Un effet « paradoxal » : le « paradoxe de Jevons »
(économiste britannique du XIXe)

« L'idée selon laquelle un usage plus économe du combustible équivaldrait à une moindre consommation est une confusion totale. C'est l'exact contraire qui est vrai. »



Des paradoxes à expliquer :

- En 2008, le rapport « SMART 2020 » du Global e-Sustainability Initiative (GESI) estimait que les technologies numériques pouvaient permettre une réduction de 15 à 30% des émissions de gaz à effet de serre mondiaux d'ici 2020. Il ne mentionne pas les effets rebonds.

2008 : 32 10⁹ t CO2 numérique 0.6

2020 : 35 10⁹ t CO2 numérique 1.4

- De 1990 à 2012,

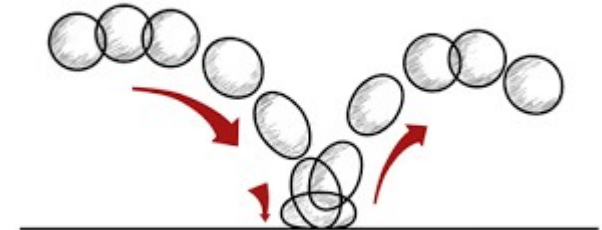
consommation d'une voiture aux 100km : 8.3l → 6.7l

consommation par habitant en France : 551l → 554l

- De 1990 à 2005,

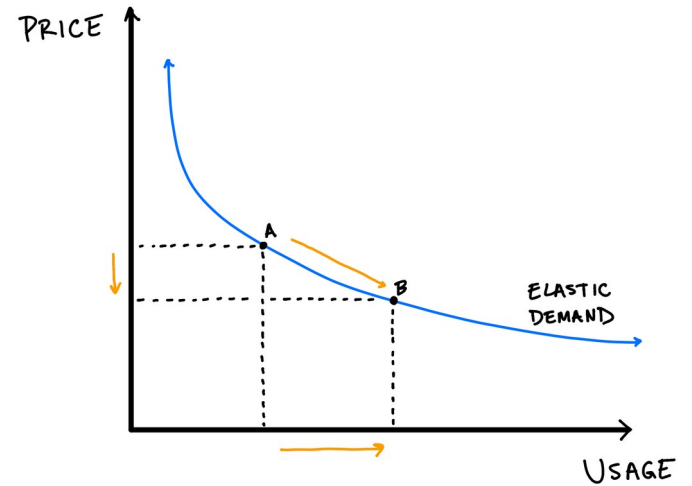
la masse d'un téléphone est divisée par 4

la masse des téléphones est multipliée par 8



De multiples variantes :

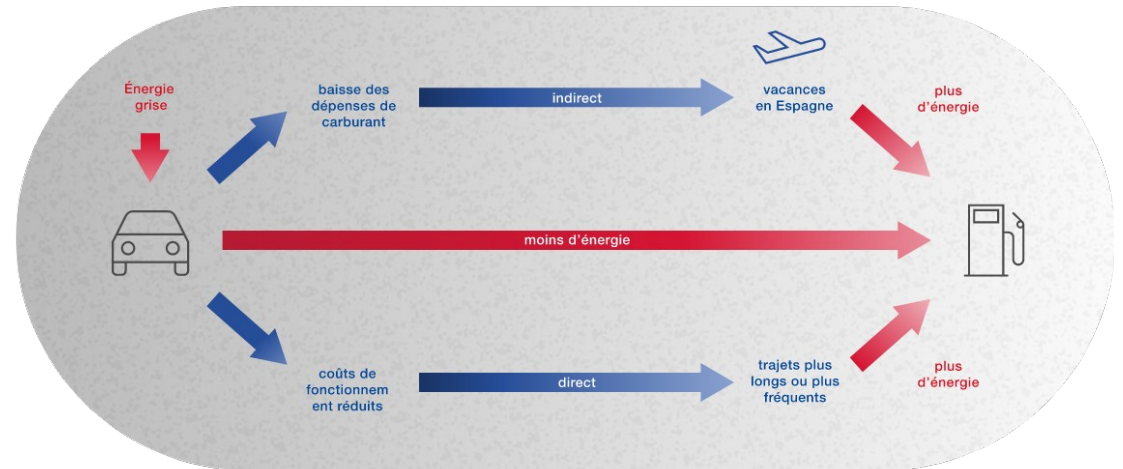
- On a accès à un article scientifique beaucoup plus rapidement, mais on passe plus de temps à faire la veille scientifique sur un sujet
- Plus les transports sont rapides, plus on passe de temps dans les transports
- Les écrans plats prennent moins de place, mais plus de place est consacrée aux écrans
- Le prix d'un produit baisse, et son vendeur gagne plus
- Remplacer prématurément son équipement par un plus récent et plus performant peut augmenter la pression sur les ressources



Trois types d'effet rebond

- Directs : la baisse du coût d'une ressource déclenche une augmentation de la demande. Par exemple, si une machine à laver consomme moins d'énergie, les consommateurs peuvent se permettre de laver leur linge plus souvent.

- Indirects : les économies faites sur un secteur induisent des dépenses sur un autre. Par exemple, prendre son vélo tous les jours au lieu de la voiture fait réaliser une économie, qu'on peut dépenser en prenant des vacances plus loin.



- Structurels : la baisse du coût d'une ressource modifie les paramètres du système. Par exemple, un carburant moins cher permet d'habiter plus loin de son lieu de travail.

On recherche les effets rebond :

- Si on développe une nouvelle technologie, ça peut être dans le but de gagner de nouveaux utilisateurs, et pas seulement de capter les utilisateurs de l'ancienne technologie qu'elle remplace. Par exemple, un logiciel de phylogénie rendu efficace et convivial.
- La taille des données est virtuellement infinie, c'est une demande élastique maximale

Mesurer et prévoir les effets rebond

L'effet rebond se mesure :

Si suite à une baisse du prix de l'énergie de 10 %, la consommation augmente de 2 %, l'effet rebond est de 20 %

Les effets rebond inférieurs à 100 % génèrent une vraie économie d'énergie

Les mesures sont parfois réalisées par des économistes dans des cas particuliers