
A User Study of Perceived Carbon Footprint

Victor Kristof
EPFL

Valentin Quelquejay-Leclère
EPFL

Robin Zbinden
EPFL

Lucas Maystre
Spotify

Matthias Grossglauser
EPFL

Patrick Thiran
EPFL

Abstract

We propose a statistical model to understand people’s perception of their carbon footprint. Driven by the observation that few people think of CO₂ impact in absolute terms, we design a system to probe people’s perception from simple pairwise comparisons of the relative carbon footprint of their actions. The formulation of the model enables us to take an active-learning approach to selecting the pairs of actions that are maximally informative about the model parameters. We define a set of 18 actions and collect a dataset of 2183 comparisons from 176 users on a university campus. The early results reveal promising directions to improve climate communication and enhance climate mitigation.

1 Introduction

To put the focus on actions that have high potential for emission reduction, we must first understand whether people have an accurate perception of the carbon footprint of these actions. If they do not, their efforts might be wasted. As an example, recent work by Wynes and Nicholas [6] shows that Canadian high-school textbooks encourage daily actions that yield negligible emission reduction. Actions with a higher potential of emission reduction are poorly documented. In this work, we model how people perceive the carbon footprint of their actions, which could guide educators and policy-makers.

In their daily life, consumers repeatedly face multiple options with varying environmental effects. Except for a handful of experts, no one is able to estimate the absolute quantity of CO₂ emitted by their actions of say, flying from Paris to London. Most people, however, are aware that taking the train for the same trip would release less CO₂. Hence, in the spirit of Thurstone [5] and Salganik and Levy [4] (among many others), we posit that the perception of a population can be probed by simple pairwise comparisons. By doing so, we shift the complexity from the probing system to the model: Instead of asking difficult questions about each action and simply averaging the answers, we ask simple questions in the form of comparisons and design a non-trivial model to estimate the perception. *In fine*, human behaviour boils down to making choices: For example, we choose between eating local food and eating imported food; we do not choose between eating or not eating. Our awareness of relative emissions between actions (of the same purpose) is often sufficient to improve our carbon footprint.

Our contributions are as follows. First, we cast the problem of inferring a population’s global perception from pairwise comparisons as a linear regression. Second, we adapt a well-known active-learning method to maximize the information gained from each comparison. We describe the model and the active-learning algorithm in Section 2. We design an interactive platform to collect real data for an experiment on our university campus, and we show early results in Section 3. Our approach could help climate scientists, sociologists, journalists, governments, and individuals improve climate communication and enhance climate mitigation.

2 Model

Let \mathcal{A} be a set of M actions. For instance, "flying from London to New York" or "eating meat for a year" are both actions in \mathcal{A} . Let (i, j, y) be a triplet encoding that action $i \in \mathcal{A}$ has an *impact ratio* of $y \in \mathbf{R}_{>0}$ over action $j \in \mathcal{A}$. Said otherwise, if $y > 1$, action i has a carbon footprint y times *greater* than action j , and if $y < 1$, action i has a carbon footprint $1/y$ times *smaller* than action j .

Given some parameters $w_i, w_j \in \mathbf{R}$ representing the perceived (log-)carbon footprint in CO₂-equivalent of action i and action j , we posit

$$y = \frac{\exp w_i}{\exp w_j}.$$

We gather the parameters in a vector $\mathbf{w} \in \mathbf{R}^M$. Assuming a centered Gaussian noise $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$, $\sigma_n^2 \in \mathbf{R}$, we model the (log-)impact ratio

$$\log y = w_i - w_j + \epsilon = \mathbf{x}^\top \mathbf{w} + \epsilon, \quad (1)$$

where the comparison vector $\mathbf{x} \in \mathbf{R}^M$ is zero everywhere except in entry i where it is $+1$ and in entry j where it is -1 . Vector \mathbf{x} "selects" the pair of actions to compare. For a dataset $\mathcal{D} = \{(i_n, j_n, y_n) : n = 1, \dots, N\}$ of N independent triplets and since $\log y \sim \mathcal{N}(\mathbf{x}^\top \mathbf{w}, \sigma_n^2)$, the likelihood of the model is

$$p(\mathbf{y} | \mathbf{X}, \mathbf{w}) = \prod_{i=1}^N p(y_i | \mathbf{x}_i^\top \mathbf{w}, \sigma_n^2) = \mathcal{N}(\mathbf{X}\mathbf{w}, \sigma_n^2 \mathbf{I}),$$

where $\mathbf{y} \in \mathbf{R}^N$ is the vector of observed (log-)impact ratios, and $\mathbf{X} \in \mathbf{R}^{N \times M}$ is a matrix of N comparison vectors.

We assume a Gaussian prior for the weight parameters $\mathbf{w} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}_p)$, where $\boldsymbol{\mu} \in \mathbf{R}^M$ is the prior mean and $\boldsymbol{\Sigma}_p \in \mathbf{R}^{M \times M}$ is the prior covariance matrix. To obtain the global perceived carbon footprint of each action in \mathcal{A} and to enable active learning, we compute the posterior distribution of the weight parameters given the data,

$$\begin{aligned} p(\mathbf{w} | \mathbf{X}, \mathbf{y}) &= \frac{p(\mathbf{y} | \mathbf{X}, \mathbf{w})p(\mathbf{w})}{p(\mathbf{y} | \mathbf{X})} \\ &= \mathcal{N}\left(\bar{\mathbf{w}} = \boldsymbol{\Sigma}(\sigma_n^{-2} \mathbf{X}^\top \mathbf{y} + \boldsymbol{\Sigma}_p^{-1} \boldsymbol{\mu}), \boldsymbol{\Sigma} = (\sigma_n^{-2} \mathbf{X}^\top \mathbf{X} + \boldsymbol{\Sigma}_p^{-1})^{-1}\right). \end{aligned} \quad (2)$$

The noise variance σ_n^2 , the prior mean $\boldsymbol{\mu}$, and the prior covariance matrix $\boldsymbol{\Sigma}_p$ are hyperparameters to be tuned. The global perceived carbon footprint is given by the posterior mean as $\exp \bar{\mathbf{w}}$. We use the posterior covariance matrix $\boldsymbol{\Sigma}$ to select the next pair of actions, as described in the following section.

Active Learning We collect the triplets in \mathcal{D} from multiple users who take a quiz. During one session of the quiz, a user sequentially answers comparison questions and decides when to stop to see their overall results. Active learning enables us to maximize the information extracted from a session.

Let $\boldsymbol{\Sigma}_N$ and $\boldsymbol{\Sigma}_{N+1}$ be the covariance matrices of the posterior distribution in Equation (2) when N and $N+1$ comparisons have been respectively collected. Let \mathbf{x} be the new $(N+1)$ -th comparison vector. As proposed by MacKay [3], we want to select the pair of actions to compare that is maximally informative about the values that the model parameters \mathbf{w} should take [1, 2]. For our linear Gaussian model, this is obtained by maximizing the total information gain

$$\Delta S = S_N - S_{N+1} = \frac{1}{2} \log(1 + \sigma_n^{-2} \mathbf{x}^\top \boldsymbol{\Sigma}_N \mathbf{x}), \quad (3)$$

where $S_N = \frac{M}{2}(1 + \log 2\pi) + \frac{1}{2} \log \det \boldsymbol{\Sigma}_N$ is the entropy of a multivariate Gaussian distribution with covariance matrix $\boldsymbol{\Sigma}_N$. The full derivation of the total information gain can be found in Appendix A.1. To maximize ΔS , we maximize $\mathbf{x}^\top \boldsymbol{\Sigma}_N \mathbf{x}$ for all possible \mathbf{x} in our dataset. Recall that comparison vectors \mathbf{x} are zero everywhere except in entry i (+1) and in entry j (-1). By denoting $\boldsymbol{\Sigma}_N = [\sigma_{ij}^2]_{i,j=1}^M$, we seek, therefore, to find the pair of actions

$$(i^*, j^*) = \arg \max_{i,j} \{\sigma_{ii}^2 + \sigma_{jj}^2 - 2\sigma_{ij}^2\}.$$

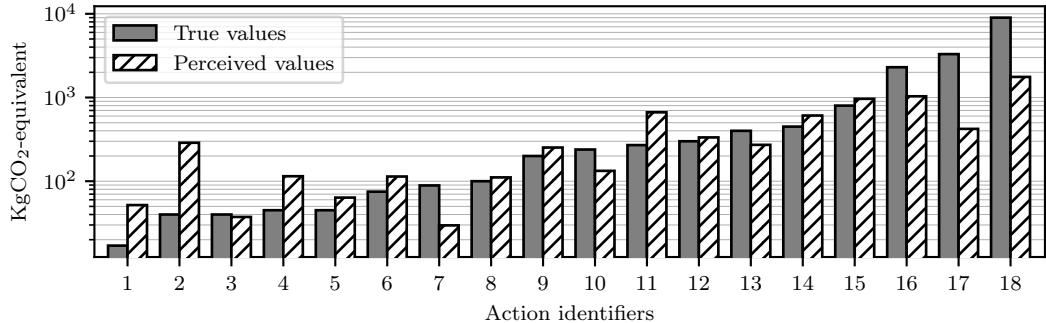


Figure 1: Global perceived carbon footprint of 18 actions in kgCO₂-equivalent and their true values (log scale). The list of actions is provided in Appendix A.2.

The prior covariance matrix Σ_p could capture the prior knowledge about the typical user perception of relative carbon footprint. In future work, we intend to further reduce the number of questions asked during one session by a judicious choice of Σ_p . In our experiments so far, we simply initialize it to a spherical covariance, as explained in the next section.

3 Results

Starting with no information at all, we arbitrarily set the prior noise $\sigma_n^2 = 1$ and the prior covariance matrix to a spherical covariance $\Sigma_p = \sigma_p^2 \mathbf{I}$, with $\sigma_p^2 = 10$. Our results are qualitatively robust to a large range of values for σ_p^2 . In order to compare the perceived carbon footprint $\exp \bar{w}$ with its true value $\exp v$, we set the prior mean to $\mu = c\mathbf{1}$, where $c = \frac{1}{M} \sum_{i=1}^M v_i$ is the mean of the (log-)true values. This guarantees that the perceived carbon footprint estimated from the model parameters have the same scale as the true values.

We compile a set \mathcal{A} of $M = 18$ individual actions about transportation, food, and household (the full list of actions is provided in Appendix A.2). We deploy an online quiz¹ to collect pairwise comparisons of actions from real users on a university campus. We collect $N = 2183$ triplets from 176 users, mostly students between 16 and 25 years old. We show in Figure 1 the true carbon footprint, together with the global perception of the population, i.e., the values $\exp \bar{w}_i$ for each action $i \in \mathcal{A}$.

The users in our population have a globally accurate perception. Among the actions showing the most discrepancy, the carbon footprint of short-haul flights is *overestimated* (Action 11), whereas the carbon footprint of long-haul flights (16) is highly *underestimated* (the scale is logarithmic). Similarly, the carbon footprint of first-class flights (18) is also *underestimated*. The users tend to *overestimate* the carbon footprint of more ecological transports, such as the train, the bus, and car-sharing (1, 4, and 6). The users have an accurate perception of actions related to diet (8, 14, and 15) and of actions related to domestic lighting (3 and 10). They *overestimate*, however, the carbon footprint of a dryer (2). Finally, they highly *underestimate* the carbon footprint of oil heating (17). Switzerland, where the users live, is one of the European countries whose consumption of oil for heating houses is the highest. There is, therefore, a high potential for raising awareness around this issue.

4 Conclusion

In this work, we proposed a statistical model for understanding people’s global perception of their carbon footprint. The Bayesian formulation of the model enables us to take an active-learning approach to selecting the pairs of actions that maximize the gain of information. We deployed an online platform to collect real data from users. The estimated perception of the users gives us insight into this population and reveals interesting directions for improving climate communication.

For future work, we will open and deploy our platform to a wider audience. We plan to collaborate with domain experts to further analyze people’s estimated perception of their carbon footprint and to translate the conclusions of the results into concrete actions.

¹ Accessible at <http://www.climpact.ch>

References

- [1] W. Chu and Z. Ghahramani. Extensions of Gaussian processes for ranking: Semi-supervised and active learning. In *Proceedings of the NIPS 2005 Workshop on Learning to Rank*, Whistler, BC, Canada, Dec. 2005.
- [2] N. Houlsby, J. M. Hernández-Lobato, F. Huszár, and Z. Ghahramani. Collaborative Gaussian processes for preference learning. In *Advances in Neural Information Processing Systems 25*, Lake Tahoe, CA, Dec. 2012.
- [3] D. J. MacKay. Information-based objective functions for active data selection. *Neural computation*, 4(4):590–604, 1992.
- [4] M. J. Salganik and K. E. C. Levy. Wiki surveys: Open and quantifiable social data collection. *PLoS ONE*, 10(5):1–17, 2015.
- [5] L. L. Thurstone. The method of paired comparisons for social values. *The Journal of Abnormal and Social Psychology*, 21(4):384–400, 1927.
- [6] S. Wynes and K. A. Nicholas. The climate mitigation gap: education and government recommendations miss the most effective individual actions. *Environmental Research Letters*, 12(7), 2017.

A Appendix

A.1 Total Information Gain for Multivariate Gaussian Distributions

Recall that the entropy of a multivariate Gaussian distribution $\mathcal{N}(\mu, \Sigma)$, $\mu \in \mathbf{R}^M$, $\Sigma \in \mathbf{R}^{M \times M}$, is given by

$$S = \frac{M}{2}(1 + \log 2\pi) + \frac{1}{2} \log \det \Sigma. \quad (4)$$

Let Σ_N and Σ_{N+1} be the covariance matrices of the posterior distribution in Equation (2) when N and $N + 1$ data points have been collected, respectively. Let \mathbf{x} be the new $(N + 1)$ -th data point. The total information gain is

$$\begin{aligned} \Delta S &= S_N - S_{N+1} \\ &= \frac{1}{2} \log \frac{\det \Sigma_{N+1}^{-1}}{\det \Sigma_N^{-1}} \\ &= \frac{1}{2} \log \frac{\det[\Sigma_N^{-1} + \sigma_n^{-2} \mathbf{x} \mathbf{x}^\top]}{\det \Sigma_N^{-1}} \end{aligned} \quad (5)$$

$$= \frac{1}{2} \log \frac{(\det \Sigma_N^{-1})(1 + \sigma_n^{-2} \mathbf{x}^\top \Sigma_N \mathbf{x})}{\det \Sigma_N^{-1}} \quad (6)$$

$$= \frac{1}{2} \log(1 + \sigma_n^{-2} \mathbf{x}^\top \Sigma_N \mathbf{x}).$$

We obtain Equation (5) by observing that $\Sigma_N^{-1} + \sigma_n^{-2} \mathbf{x} \mathbf{x}^\top = \sigma_n^{-2} \mathbf{X}^\top \mathbf{X} + \Sigma_p^{-1} + \sigma_n^{-2} \mathbf{x} \mathbf{x}^\top = \Sigma_{N+1}^{-1}$. We obtain Equation (6) by the matrix determinant lemma.

A.2 List of Actions

We provide here the full list of actions, together with the true carbon footprint associated with each of them. Because different countries use different sources of energy, we calculate the carbon footprint *relative* to the country where our university is located. The actions are ordered according to their true carbon footprint.

1. Take the train in economy class on a 1000-km round-trip.

The train is a high-speed train with 360 seats. The seat-occupancy rate is 55% (average rate for these types of trains). We count the CO₂ emissions per passenger.

Carbon footprint: 17 kgCO₂-equivalent.

2. Dry your clothes with a dryer for one year.

A dryer emits CO₂ because it consumes electricity. We consider a dryer of average quality. The electricity is consumed from a grid with average CO₂ rate.

Carbon footprint: 40 kgCO₂-equivalent.

3. Light your house with LED bulbs.

LED bulbs emit CO₂ because they consume electricity to generate light. The electricity is consumed from a grid with average CO₂ rate.

Carbon footprint: 40 kgCO₂-equivalent.

4. Take the bus on a 1000-km round-trip.

The bus is a standard-size bus with 60 seats. The seat-occupancy rate is 50% (average rate for buses). We count the CO₂ emissions per passenger.

Carbon footprint: 45 kgCO₂-equivalent.

5. Drive an electric car alone on a 1000-km round-trip.

The car is a compact electric car that consumes 15 kWh/100km. The electricity is consumed from a grid with average CO₂ rate. There are no other passengers in the car. We count the CO₂ emissions per passenger.

Carbon footprint: 45 kgCO₂-equivalent.

6. Car-share with three other persons on a 1000-km round-trip.

The car is a mid-sized gasoline car that consumes 7 l/100km. There are four persons in the car. We count the CO₂ emissions per passenger.

Carbon footprint: 75 kgCO₂-equivalent.

7. Eat local and seasonal fruits and vegetables for one year.

Growing food emits CO₂ because it requires fertilizing and driving agricultural machines. The goods are then transported to grocery shops and to your home.

Carbon footprint: 89 kgCO₂-equivalent.

8. Eat eggs and dairy products for one year.

The production of eggs and dairy products (milk, cheese, etc.) emits CO₂ because of water and land consumption, animal methane, and fossil fuel consumption for transportation and heating. We consider an average citizen consuming 50 kg of eggs and dairy products per year.

Carbon footprint: 100 kgCO₂-equivalent.

9. Throw all waste in the same trash for one year.

Throwing all waste (PET, glass, cardboard, etc.) in the same trash, i.e., without recycling, emits CO₂ because more energy is needed to extract, transport, and process raw materials. Incinerators also burn more waste, and organic waste decomposition generates methane.

Carbon footprint: 200 kgCO₂-equivalent.

10. Light your house with incandescent bulbs.

Incandescent bulbs emit CO₂ because they consume electricity to generate light. The electricity is consumed from a grid with average CO₂ rate.

Carbon footprint: 239 kgCO₂-equivalent.

11. Fly in economy class for a 800-km round-trip.

The plane is a standard aircraft for short-distance flights with 180 seats. The seat-occupancy rate is 80%. We count the CO₂ emissions per passenger.

Carbon footprint: 270 kgCO₂-equivalent.

12. Drive alone for a 1000-km round-trip.

The car is a mid-sized gasoline car that consumes 7 l/100km. There are no other passengers in the car. We count the CO₂ emissions per passenger.

Carbon footprint: 300 kgCO₂-equivalent.

13. Heat your house with a heat pump for one year.

A heat pump emits CO₂ because it consumes electricity to generate heat. The house is of average size. The electricity is consumed from a grid with average CO₂ rate.

Carbon footprint: 400 kgCO₂-equivalent.

14. Eat imported and out-of-season fruits and vegetables for one year.

Growing food emits CO₂ because it requires fertilizing and driving agricultural machines. Importing food emits CO₂ because of fossil fuel consumption for transportation. Out-of-season food emits CO₂ because it grows in greenhouse that needs to be heated. The goods are then transported to grocery shops and to your home.

Carbon footprint: 449 kgCO₂-equivalent.

15. Eat meat for one year.

Meat production emits CO₂ because of water and land consumption, animal methane, and fossil fuel consumption for transportation and heating. We consider an average citizen consuming 50 kg of meat per year.

Carbon footprint: 800 kgCO₂-equivalent.

16. Fly in economy class for a 12000-km round-trip.

The plane is a standard aircraft for long-distance flights with 390 seats. The seat-occupancy rate is close to 100%. We count the CO₂ emissions per passenger.

Carbon footprint: 2300 kgCO₂-equivalent.

17. Heat your house with an oil furnace for one year.

An oil furnace emits CO₂ because it burns fuel to generate heat. The house is of average size.

Carbon footprint: 3300 kgCO₂-equivalent.

18. Fly in first class for a 12000-km round-trip.

The plane is a standard aircraft for long-distance flights with 390 seats. The seat-occupancy rate is close to 100%. We count the CO₂ emissions per passenger. Passengers flying in first class use more space than passengers in economy.

Carbon footprint: 9000 kgCO₂-equivalent.