Quantitative Analysis of Lifestyle Impacts on Climate Change in US

Aanchal Adhikari, Mauricio Camacho, Aaron Kang, Dushica Milosavljevikj, James Sinclair, Shi Wang
**All team members contributed a similar amount of effort.*

Global climate change is a paramount issue facing us all and we are at a critical moment. The drastically changing weather can affect everything from human health to food production. Regardless of who you are, impacts of climate change pose a threat to the way we all live. Unless we collectively and individually take action in our daily lives and activities, adjusting to these changes in the future will be costly.

We want to test and visualize the quantitative links between lifestyle and climate change, in the United States. We proxy lifestyle by six types of consumption: consumer goods, housing, leisure, nutrition, mobility, and services. Our proxy for climate change is carbon dioxide (CO_2) output, which is the majority greenhouse gas emission of the US. We aim to establish such a quantitative analysis and corresponding visualization for assessing lifestyle carbon footprint changes in different lifestyle change scenarios.

The Paris agreement advocates adopting sustainable lifestyle across countries to meet the 1.5C target, on a global scale. Two recent studies demonstrated that changes in consumption patterns and dominant lifestyles are a critical and integral part of the solutions package to address climate change[1,2]. Koide, R. et al represents one of the first of its kind in terms of proposing per capita footprint targets with explicit linkages to the Paris Agreement and assessing the solutions based on the physical amount of consumption across fields including nutrition, consumer goods, housing, mobility, leisure and services[1]. A study led by Thogersen points out that not all industries and types of consumption contribute equally to larger GHG and further expands on possible lifestyle changes that may lead to improved levels of GHG [3]. These papers provide tangible measures on various types of consumption and their corresponding carbon footprints.

In looking at nutrition, the most impactful factor one can have on their carbon footprint is their meat consumption, due to the agricultural production process and land use that farming animals incur[4]. With meats like beef requiring 20x more land and emitting 20x more GHG than plant proteins, such as beans[5]. However, eating meat doesn't need to be completely stopped, if the world's highest consuming countries reduce just their beef to about 50 calories per day, it would nearly eliminate the need for expanding agricultural lands and processes[6].

For the mobility portion of one's carbon footprint, the most impactful factor would be fuel consumption. In the US alone, transportation makes up about 27% of all total GHG emissions[8]. Additionally, the US has the world's largest commercial airline system, accounting for 10% of the US's transportation emissions[9]. Flying accounts for around 2.5% of global CO_2 emissions, but 3.5% when taking non-CO_2 impacts on climate into account.[10-11].

While Koide et al. provide thorough study on the relation between lifestyle and climate change on individual levels, it only brings to light data for Finland, Japan, and China, Brazil, and

India[1]. We will conduct a similar study on US data to assess lifestyle carbon footprints and lifestyle change options since, to our knowledge, no study has been conducted on this.

To obtain a more informed analysis of the US, Jones and Kammen provide a link between composition of carbon footprint, geographical regions, and income in the US. Income appears in further studies and comes up as a significant, positively correlated factor with the carbon footprint[12-13]. However, the effect of high income on increased GHG is smaller than in low-income countries[14].

Based on this, if our project is applied in low-income countries, we will need additional adjustment for the low-income country effect (we consider the US, a high-income country). Given the various discussions on income and its role in determining the carbon footprint, we will include it as a variable in our project. We will also utilize Roy and Pal's study which provides a scientific and visual way in depicting the CO_2 emissions and income level (as GNP per capita)[15].

Although some authors limit their research to fewer factors that could impact climate change, we still see their value-add for our project. Since we aim to analyze and visualize the relationship between consumption in different industries and GHG (a proxy for climate change), we consider extracting learnings and guidance from multiple areas.[16–19]

Two studies assess the impact of personal beliefs on personal carbon footprint, where they find a dominant relation between positive environmental attitude and energy saving behavior, especially after receiving feedback on their interview sessions[20-21] This gives us a ground to believe that once exposed to more information about the link between personal lifestyle and climate change, the individuals will begin building more environmentally friendly lifestyles. Hence, our project can contribute to better insights on environmental footprint by providing analytical and visual representation of the relation between consumer lifestyle and climate change.

Innovation ideas

1. We will conduct a quantitative study on US data to assess the relation between lifestyle and climate change on individual levels. The outcome of the project is an interactive website providing a personalized user experience to see the impact of their lifestyle on the climate intuitively, which can result in more actively engaged users. In turn, our product would increase eco-accountability and encourage environmentally responsible lifestyles. To our knowledge, no similar product is available on the market.
2. We will utilize various consumption variables data with an MTS model to predict how changes in each of those variables will impact the total CO_2 emissions in the future.
3. We will have interactive visualizations to let users input lifestyle changes and observe the predicted CO_2 emissions based on those changes.

Methods

Our intention is to enable our users to easily visualize the impact of their consumer behavior on GHG emissions, and thus, understand their impact on the climate. Literature argues that the more directly involved the people with the topic of climate change, the more responsibly they act with respect to their consumption decisions. The tool will allow the user to input their intended personal monthly consumption of different product and service categories. The tool will visualize the effect they will have on CO_2 up to 2026.

The tool will work under the assumption that their state populace makes the same decision as them. That means, the person imputing the values of their intended consumption of factor "Xi", will be able to see their impact on CO_2, in a scenario where everyone else makes their same choices. The user will also be able to define the time horizon for which they would like to see the forecasted CO_2. The time horizon is limited to the year 2026.

These data points represent the planned user consumption of the defined consumer item (X1,… Xn), over the following years.

The user is presented with the ability to reduce any of their 6 lifestyle measures in half in order to visualize how much their decisions could potentially affect their future CO_2 output.

Datasets

The data we have gathered comes from multiple sources including the Bureau of Economic Analysis[22](BEA), the U.S. Energy Information Administration[23](EIA), the Federal Reserve Economic Data (FRED), and the United States Environmental Protection Agency[24](EPA).

We gathered datasets that contained information on the following:

1. Datasets containing the variables or predictors to be used in our model. This dataset has information about each state in the United States, for 1997 to 2021 and contains information like:
    a. Personal consumption expenditures, which is spending on goods and services purchased by, and on behalf of, households. These serve as our proxies for consumer goods and services.
    b. Purchases of food and drink by US households for consumption within the household, our proxy for nutrition.
    c. Spending for housing and utilities which serves as our proxy for housing.
    d. Spending on transportation services which includes long distance motor vehicles, personal motor vehicles, and public transportation services. This serves as our proxy for mobility.
    e. Spending on recreation goods, services, and food accommodations which serves as our proxy for leisure.
    f. Gross Domestic Product (GDP). This data is broken by industry.
    g. Personal Income. Consisting of the income that persons receive in return for their provision of labor, land, and capital used in current production activities. The data contains information by state and per capita.
2. Datasets containing the target variables with information like CO_2, which is the most prominent GHG emission. The data is broken by State, Industry and Facility for 2011 to 2021.

The first dataset contains around 20K records/rows, and the second dataset contains 220K records/rows. In order to clean our data and compile it into a single usable dataset we had to take a few steps. Our GDP data from the BEA only went to 2005, so it had to be merged with data from FRED in order to go back to 1997, which the rest of our consumer data reaches. Additionally, our datasets from FRED typically were in billions of dollars, so we converted those to millions of dollars in order to match data from our other sources. Finally, our CO_2 emissions by state were in mole fractions so we converted them to metric tons in order to match up with the data we got for the entire US.

Models

We want to project the impact of consumption on GHG over time for several periods in the future, using simple time series models built on autocorrelation of a single variable with its historical values. However, we want to estimate the impact of multiple consumption factors and their historical values on CO_2 emissions. For this reason, we will use a multivariate time series model (MTS), where each of the factors is assumed to have a correlation with the target, and possibly, one or several other factors included in the model. The MTS model unifies the capabilities of both, multiple / multivariate regression and time series models. We will consider two options: Vector Auto Regression (VAR) and Facebook Prophet[25-27].

1. The VAR model specifies that each variable is a function of its own historical values and of the past values of all other variables included in the model. That means, VAR is a multivariate model, where any of the variables in the model can be considered both a factor or a dependent variable. To use the VAR model, we will first run a Granger test for endogeneity of the variables. Our chosen set of factors (nutrition, mobility, services, housing, goods, and leisure) might show multicollinearity (for instance, if we have an upward macroeconomic cycle, then we expect more spending in the housing sphere).
2. Facebook Prophet is a univariate series, which gives the possibility of adding independent regressors (time series). Each of the time series are treated as an independent and non-correlated variable with the rest of the variables included in the model. Since we would like to capture the impact of the factor variables only on the CO_2 (our target), Facebook Prophet might be a more appropriate model than VAR.

To predict the future values of the CO_2, the model needs future values of the factor variables for a defined number of periods in the future. We meet the requirements in the following way:

1. Historical data in the time series is supplied by existing datasets, obtained from the mentioned data sources. The past data contains all the data points for both, dependent and independent series (Y and Xi's).
2. Future data for the independent (factor variable) time series is automatically generated, assuming the historical trends continue.
3. Future data for the dependent time series (CO_2 up to the year 2026) is predicted based on the past data and the user-imputed future values for the factor variables. The user imputed values will steer the direction of the timeseries, if they choose to consume lower the amount for one or several of the specified consumer categories.

Experiments/Evaluation

Since we decided to test our data with two different models, we needed to figure out which of the two would produce more accurate predictions. Once we decided which model to use, we then had to determine the realistic effects that reducing consumption of the factors would realistically have on our predictions. Additionally, once we had our data, model, and predictions we needed to test our visualizations in order to assess our output to see if they aligned with our expectations, or presented us with contrary views.

Once we cleaned, reformatted, and brought together our datasets, they were ready for testing with our two models. We used the metric median absolute percentage error (MDAPE) to compare the performance of the two models. MDAPE is an error metric used to measure the performance of regression machine learning models. It is the median of all absolute percentage errors calculated between the predictions and their corresponding actual values. Each model's performance on the test set decided which model would be more appropriate for our visualization. MDAPE is a metric that is not sensitive to outliers, and that can be used to compare different models, which was needed in our case for comparison between Prophet and VAR. MDAPE metric should not be used with data containing values close to zero, which is not the case with our dataset. The resulting value is returned as a percentage.

To run the cross validation, we took out the last 5 years of data on CO2 and ran the model including full data from 1997 – 2016. For the years 2017-2021 we fed the model with the actual values of the regressors (and deleted CO2 values). We then compared the predicted CO2 values for years 2017-2021 with the actual CO2 values, included in the original data set. The differences for the Prophet model were minor. MDAPE showed < 0.01, for overall US data on monthly level and for most of the individual states, MDAPE was < 0.1, on an annual and per state basis. MDAPE score of < 0.1 shows that there are low deviations between predicted and actual values. The resulting value for VAR was > 0.7. The reason for this is that VAR asks for stationary values, and requires additional data preprocessing, manipulation, and additional data search to meet the stationarity requirements. Given the time restrictions and the positive Prophet outcome, we decided to use Prophet modeling output for our visualization.

When running Prophet for the neutral option (no intervention in consumption, that is, running Prophet with no changes in the factor variables), the model produced stable results. However, when adding the negative consumption factors, which would indicate extraordinary reduction in consumption of one or more of the six categories in the years 2022-2026, we noticed that the model produced negative projections. While it is possible for CO_2 emissions to reach negative values, such as the case of the country of Bhutan, we did not think it realistic to predict so many US states would have negative CO_2 emissions by the year 2026. In an effort to curb this effect, we tried converting the historical CO_2 emissions to natural log before inputting them into the model. However, this created abnormally high CO_2 emissions for some of the states that seemed not to follow the historical trend. We suspect that the reason for this is the fact that we used unreasonably high consumption decrease factors. In addition, we did not control for CO2-neutral technology advancement over the years, which can lead to significant reduction in CO2, despite increasing levels of consumption.

Thus, we took a second look at our method, in particular, using the predicted values for the reduced consumption of the factor variables. Since the historical data shows that CO_2 emissions decreased even though consumption was still increasing, this may have trained the model to predict higher values for CO_2 emissions as consumption decreased. Thus, we performed additional literature review in order to determine the effects of reducing consumption for each factor. Upon review of a paper by Jones and Kammen[26] that used national household surveys to develop econometric models of demand for the factors. With those models, they then derived the average household carbon footprint. From their findings we decided upon the following effects: reducing mobility by 1/2 results in a -0.2 change in CO_2 emissions, reducing housing by 1/2 results in a -0.083 change in CO_2 emissions, reducing nutrition by 1/2 results in -0.25 change in CO_2 emissions, reducing shopping for consumer goods by 1/2 results in -0.23 change in CO_2 emissions, reducing purchases of services by 1/2 results in -0.125 change in CO_2 emissions, and reducing leisure activities by 1/2 results in -0.092 change in CO_2 emissions. Adjusting the consumption intervention factors to more reasonable, literature-suggested levels, the model reported more stable projections.

Once the model had produced our final dataset and we had calculated and added data according to the factor reductions, we were ready to set up our visualization. This was done in Tableau and can be accessed through Tableau Public. We used three different visualizations to illustrate the impact of a user's lifestyle on CO_2 emissions. The first visualization is an interactive heat map of the US in which the user can see the CO_2 emissions of each state in metric tons. Hovering over a state will display its name and its change in predicted CO_2 emissions for that year. Additionally, the user can select any combination of factors they would like to reduce to view its effect on their selected year's prediction. Below the map, there are two-line charts that correspond to selections on the map. If no specific state is selected, the information displayed pertains to the whole US. However, when the user selects a state, the line charts will display information for that state alone. The first line chart depicts the CO_2 emissions by year, in metric tons. If a user selects factors to reduce, a second orange line will appear on this chart to demonstrate the effect the reduction has on the predicted emissions. The second line chart depicts the historical spending in millions of current dollars for each of the consumption categories. This chart only adjusts when the user either selects a state or is viewing the entire US.

Link to the visualization:
https://public.tableau.com/views/final_visualization_16700925381540/USACO2Emmisions?:language=en-US&:display_count=n&:origin=viz_share_link

Discussion

In our efforts to create an effective visualization to show the quantitative links between lifestyle and climate change, we certainly saw value in what we created in addition to learning more about how it could be improved. In comparing ours to other tools and visualizations that try to engage users in understanding how their actions affect the total CO_2 emissions, a couple of things are evident. We noticed that things such as carbon footprint calculators, such as one made by the EPA, require quite a bit of time and effort from the user. While this may be more personalized to the user, sometimes complexity may discourage use entirely. We feel we can improve our own visualization by perhaps integrating more factors "under the hood" of the visualization in order to create more dynamic predictions by reducing a major consumption factor. This could involve more research about the technology, taxation, laws, and other factors under the umbrella of each of the 6 consumptions and how they affect the overall CO_2 emissions. We see value in our visualization for its ability to give users a quick insight into what effect their actions have. Perhaps paired with a more complex and time intensive assessment such as the carbon footprint calculator, one can more effectively draw in and capture the attention of users and engage them on more personal levels. This can more effectively encourage users to begin building more environmentally friendly lifestyles.

References:

1. Koide, R. *et al.* Lifestyle carbon footprints and changes in lifestyles to limit global warming to 1.5 °C, and ways forward for related research. *Sustainability Sci.* **16**, 2087–2099 (2021).

2. Moran, D. *et al.* Quantifying the potential for consumer-oriented policy to reduce European and foreign carbon emissions. *Clim. Policy* **20**, S28–S38 (2020).

3. JohnThøgersen. Consumer behavior and climate change: consumers need considerable assistance. *Current Opinion in Behavioral Sciences* **42**, 9–14 (2021).

4. Gregory M. Peters et. al, Red Meat Production in Australia: Life Cycle Assessment and Comparison with Overseas Studies, Environmental Science & Technology **44**, 1327-1332 (2010).

5. Li Xue, Neele Prass, Sebastian Gollnow, Jennifer Davis, Silvia Scherhaufer, Karin Östergren, Shengkui Cheng, Gang Liu. Efficiency and Carbon Footprint of the German Meat Supply Chain. Environmental Science & Technology **53** , 5133-5142 (2019).

6. Cusack DF, Kazanski CE, Hedgpeth A, Chow K, Cordeiro AL, Karpman J, Ryals R. Reducing climate impacts of beef production: A synthesis of life cycle assessments across management systems and global regions. Glob Chang Biol. 2021 May;27(9):1721-1736.

7. Frey HC. Trends in onroad transportation energy and emissions. J Air Waste Manag Assoc. 2018 Jun; 68(6):514-563.

8. Alan Jenn, Inês M. L. Azevedo, and Jeremy J. Michalek, Alternative Fuel Vehicle Adoption Increases Fleet Gasoline Consumption and Greenhouse Gas Emissions under United States Corporate Average Fuel Economy Policy and Greenhouse Gas Emissions Standards, Environmental Science & Technology **50**, 2165-2174 (2016).

9. Graver, Brandon and Rutherford, Dan. U.S. Passenger Jets under ICAO's CO2 Standard, 2018 – 2038. International Council on Clean Transportation (ICCT). 20-18-25 (2018).

10. Lee, D. S., Fahey, D. W., Skowron, A., Allen, M. R., Burkhardt, U., Chen, Q., … & Gettelman, A. (2020). The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric Environment*, 117834.

11. Sausen, R., & Schumann, U. (2000). Estimates of the climate response to aircraft CO2 and NOx emissions scenarios. *Climatic Change*, *44*(1-2), 27-58.

13. Jones, C. M. & Kammen, D. M. Quantifying carbon footprint reduction opportunities for U.S. households and communities. *Environ. Sci. Technol.* **45**, 4088–4095 (2011).

14. Hertwich, E. G. & Peters, G. P. Carbon Footprint of Nations: A Global, Trade-Linked Analysis. *Environmental Science & Technology* vol. 43 6414–6420 Preprint at https://doi.org/10.1021/es803496a (2009).

15. Roy, J. & Pal, S. Lifestyles and climate change: link awaiting activation. *Current Opinion in Environmental Sustainability* **1**, 192–200 (2009).

16. Perch-Nielsen, S., Sesartic, A. & Stucki, M. The greenhouse gas intensity of the tourism sector: The case of Switzerland. *Environ. Sci. Policy* **13**, 131–140 (2010).

17. González, A. D., Frostell, B. & Carlsson-Kanyama, A. Protein efficiency per unit energy and per unit greenhouse gas emissions: Potential contribution of diet choices to climate change mitigation. *Food Policy* **36**, 562–570 (2011).

18. Stoll-Kleemann, S. & Schmidt, U. J. Reducing meat consumption in developed and transition countries to counter climate change and biodiversity loss: a review of influence factors. *Regional Environ. Change* **17**, 1261–1277 (2017).

19. Chuvieco, E., Burgui-Burgui, M., Orellano, A., Otón, G. & Ruíz-Benito, P. Links between Climate Change Knowledge, Perception and Action: Impacts on Personal Carbon Footprint. *Sustain. Sci. Pract. Policy* **13**, 8088 (2021).

20. Boucher, J. L. Culture, Carbon, and Climate Change: A Class Analysis of Climate Change Belief, Lifestyle Lock-in, and Personal Carbon Footprint. *Socijalna ekologija : časopis za ekološku misao i sociologijska istraživanja okoline* **25**, 53–80 (2016).

21. Brandon, G. & Lewis, A. REDUCING HOUSEHOLD ENERGY CONSUMPTION: A QUALITATIVE AND QUANTITATIVE FIELD STUDY. *J. Environ. Psychol.* **19**, 75–85 (1999).

22.https://www.bea.gov/

23.https://www.eia.gov/

24.https://www.epa.gov/

25. https://www.sciencedirect.com/topics/nursing-and-health-professions/exogenous-variable

26.https://facebook.github.io/prophet/docs/seasonality,_holiday_effects,_and_regressors.html#additional-regressors

27.https://stats.stackexchange.com/questions/481112/what-are-the-underlying-statistical-differences-between-the-vector-autoregression

28.Spatial Distribution of U.S. Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density, Christopher Jones and Daniel M. Kammen, Environmental Science & Technology 2014 48 (2), 895-902, DOI: 10.1021/es4034364