Ontario Greenhouse Gas Emissions Vary Based on Numerous Environmental Metrics*

Alyssa Schleifer

27 April 2022

Abstract

Data from recent years shows that Canada is falling behind other G20 nations in regards to improving its sustainable consumption and environmental performance in the face of a worsening global climate crisis. By analyzing data regarding energy and water usage of large buildings across Ontario, we can gain insight into several different metrics of environmental performance and determine their overall impact on greenhouse gas emissions. In order to achieve this, I used a multiple linear regression model to determine the relationship between variables such as water and electricity usage and overall greenhouse emissions to draw conclusions on the prevalence of these factors regarding their overall environmental impact. The findings of this analysis showed that many of the most commonly used energy sources and environmental metrics each have notable, yet vastly different effects on total emissions.

Keywords: environmental performance, greenhouse gas emissions, multiple linear regression, sustainability metrics, climate change

Contents

*Code and data supporting this analysis are available at: https://github.com/alyssaschleifer/ontario-energy-analysis

1 Introduction

In June of 2021, the Government of Canada adopted the Canadian Net Zero Emissions Accountability Act, which establishes their plan to achieve net-zero greenhouse gas (GHG) emissions by 2050 ("Net-Zero Emissions by 2050" 2022). Current scientific data suggests that in order to prevent the irreversible and disastrous effects of climate change, the global temperature change must be restricted to a 1.5 degrees Celsius increase compared to pre-industrial times (Buis 2020). However, as emissions continue to rise on a global scale, drastic steps must be taken in order to transition to GHG emission neutrality. While working towards net-zero, analyzing how various sustainability metrics contribute to emissions could help drive decision making. For instance, intensive water usage is strongly correlated to increased CO_2 emissions. A large amount of energy is required to treat and supply water. In addition, the worsening effects of climate change have been shown to have countless adverse impacts on freshwater resources, rendering a great deal of current water supplies unusable. As a result, water suppliers are turning to remote and alternative water sources, which come with both an increased energy cost and a greater contribution to carbon emissions (Griffiths-Sattenspiel and Wilson 2009). Electricity use is another key driver in the increase of emissions. Around 60% of our electricity comes from burning fossil fuels such as coal, oil, and natural gas ("Climate Change Indicators: Greenhouse Gases," n.d.). This burning of fossil fuels leads to the release copious amounts of CO_2 and nitrous oxides into the atmosphere. It has been well-established that greenhouse gases have a number of dire effects in terms of their impact on the environment ("Climate Change Indicators: Greenhouse Gases," n.d.). By absorbing solar energy from the sun, the gases trap this heat in the atmosphere which leads to a gradual warming effect known as the greenhouse effect. This greenhouse effect has been consistently cited as the principal driving force of global climate change ("Climate Change Indicators: Greenhouse Gases," n.d.).

While there is no shortage of evidence linking greenhouse gas emissions to climate change, researchers, scientists, and policy makers continue to struggle in designing and implementing effective and feasible solutions that provide lasting, tangible results. In this case, data analysis is an invaluable tool as it allows us to measure and compare metrics of sustainability and observe trends in the data that can be used to establish a framework for novel insight into the heart of the issue. We can do this not only by looking at metrics known to contribute to GHG emissions, but additionally attempting to measure the ways in which they interact and are influenced by other conditions.

In the following paper, we will conduct an analysis of Ontario's energy and greenhouse gas emissions based on public data sourced from large commercial and industrial buildings throughout the province. The central idea we hope to shed light on is uncovering the principal contributors to greenhouse gas emissions, with an emphasis on the varying effects of common energy sources as well as other demographic variables. To achieve this, a multiple linear regression model will be used to assess the relationship between several sustainability factors such as water, electricity, and natural gas usage, as well as other potentially influential variables such as region and building type (residential, school, wholesale/retailer, etc). From this analysis, we hypothesize that a higher usage of energy and water resources will show a strong, positive correlation with a building's total greenhouse gas emissions, and total emissions will vary considerably based on location as well as sector. After running the model on the data, we will discuss the findings as well as the broader implications of these results in the context of their overall environmental impact and relevance to the issue at hand.

2 Data

The following report, including all data processing, modeling, and analysis, will be carried out using the `R` statistical programming language (R Core Team 2021) as well as packages `tidyverse` (Wickham et al. 2019), `dplyr` (Wickham et al. 2021), `knitr` (Xie 2021b), and `bookdown` (Xie 2021a). All figures, tables, and diagrams in the report are generated and formatted using `ggplot2` (Wickham 2016), `kableExtra` (Zhu 2021), `tableone` (Yoshida 2022), `patchwork` (Pederson 2020), `DiagrammeR` (Iannone 2022), and `webshot` (Chang 2022). Packages supporting the multiple linear regression analysis include `car` (Fox 2021) and `modelsummary` (Arel-Bundock 2022).

2.1 Source

The data used in this report is available publicly through the Ontario Data Catalogue, which is published and maintained by the Ontario government ("Energy and Water Usage of Large Buildings in Ontario" 2021). The Ontario Electricity Act (1998) mandates annual reporting of energy consumption and water use for large buildings across the province ("O. Reg. 506/18: REPORTING of Energy Consumption and Water Use" 2018). In this case, a "large building" is defined as a building of 50,000 square feet or larger. This data is not randomized and is reported directly to the Ontario government by building owners or agents, where the data is then published on a yearly basis. The dataset is titled "Energy and water usage of large buildings in Ontario", and was last updated on December 14, 2021 ("Energy and Water Usage of Large Buildings in Ontario" 2021). The data set includes information regarding several metrics of environmental performance for different building types throughout Ontario.

Although the reporting of energy consumption and water use of large buildings is regulated by the government, this data set is not fully comprehensive as it is missing reported data from a few key building types, including agricultural and heavy industrial buildings. The source explicitly states that data from these buildings is omitted from the data set, but does not indicate a reason for this. This is particularly concerning when considering the the role each of these sectors play in global emissions and energy consumption. The agricultural industry is a major contributor to climate change and there is considerable evidence suggesting agriculture is one of the largest contributors to emissions, making up over 10% of all GHG emissions globally (Laborde 2021). Evidence shows a similar trend for heavy industry, which refers to activities such as production and manufacturing of iron, steel, cement, and chemicals. Estimates show that heavy industry is directly responsible for up to a quarter of all global GHG emissions (Ladislaw 2022). As aforementioned, it is unclear why this data was excluded, and the lack of insight into the environmental metrics in these key sectors hinders the overall comprehensiveness and reliability of the data.

There are additional implications associated with the self-reported nature of the data. Buildings that are required to report data are assigned a unique six-digit identifier, known as an EWRB ID. The ID for a given building remains the same every year, regardless of whether a building is bought or sold. In some cases, the exact number or configuration of buildings on a property is not known, in which case a single ID is assigned to the property as a whole regardless of the number of buildings on the property. This has the potential to introduce a number of inconsistencies in the data; if a property has several buildings but the metrics are being reported under a single ID number, this would greatly overestimate the total metrics for a single building. As stated by the source, the general assumption is that there is only a single building on each property and as a result, the ID is unique to a single building. In some cases this has been rectified when the province can confirm the presence of multiple buildings on the same property, however, most cases operate under the assumption of a single building per lot which is not always the case.

2.2 Variables

The original data set provides measurements for different environmental performance metrics, defined by several different units. To preserve consistency in the analysis, I will focus only on metrics reported as per square meter of a building and omit measurements reported in square feet. A complete breakdown of the units of measurement used for each metric in the analysis can be found in Appendix A. The data also contains a "city" variable, which indicates the location of a given building. For the sake of brevity, I will amend this variable, which contains 93 different values, and group each city into one of five different regions: Southwest, Northern, Eastern, Central, and Greater Toronto Area. The regions were defined based on criteria taken from the Construction Forecasts website published by the government of Canada, which are listed in Appendix B. The final data set contained 783 observations across seven variables. These seven variables include Region, Property Type, Electricity Use Intensity, Natural Gas Intensity, Water Use Intensity, Energy Use Intensity, and Greenhouse Gas (GHG) Emissions Intensity. Table 1 provides a breakdown of the variables in the data as well as their values and some corresponding statistics. The first section of the table shows the regional breakdown of the buildings in the data, which indicate that the vast majority of large buildings are located within the GTA, whereas only 1% of these buildings are located in northern Ontario. This result is unsurprising when considering the regional population density throughout the province. In addition, we can also observe that the majority of large buildings are classified as "housing". The breakdown of property type is an important point to consider in order to understand how different building types affect environmental metrics differently, and the varying impact this has on overall GHG emissions. Finally, the table highlights the five environmental metrics we will use in our analysis, as well as the mean and standard deviation of each of these metrics based on the data collected from the 783 buildings.

Table 1: Overview of energy and water usage of large buildings in Ontario

Name	Overall
Total Sample Size	783
Region (%)	
Central Region	71 (9.1)
Eastern Region	90 (11.5)
Greater Toronto Area	568 (72.5)
Northern Region	9 (1.1)
Southwest Region	45 (5.7)
Property Type (%)	
Distribution Center	73 (9.3)
Hotel	17 (2.2)
Housing	149 (19.0)
Mixed Use Property	8 (1.0)
Office	291 (37.2)
Retail	154 (19.7)
Storage Facility	91 (11.6)
Environmental Metric (mean (SD))	
Electricity Use Intensity	0.49 (0.36)
Natural Gas Use Intensity	12.82 (7.45)
Water Use Intensity	0.70 (0.77)
Energy Use Intensity	0.98 (0.50)
GHG Emissions Intensity	25.04 (13.49)

Figure 1 shows how the five environmental metrics vary based on region as well as building type. The patterns displayed in the graphs show the highest relative level of usage among all metrics within buildings from northern regions. During the cleaning process, we omitted all data that was not weather-normalized to account for differences in weather conditions throughout the province that could affect the results (for example, significantly colder temperatures during winter months in northern regions would result in an increased need for heating systems, and in turn, drive up the natural gas and electricity use metrics for this region). Because of this, we can assume that these metrics differ based off of other factors independent from weather. An interesting observation is that metrics from the GTA are relatively low compared to other regions in most cases. One explanation for the low relative metrics in the GTA compared to high relative

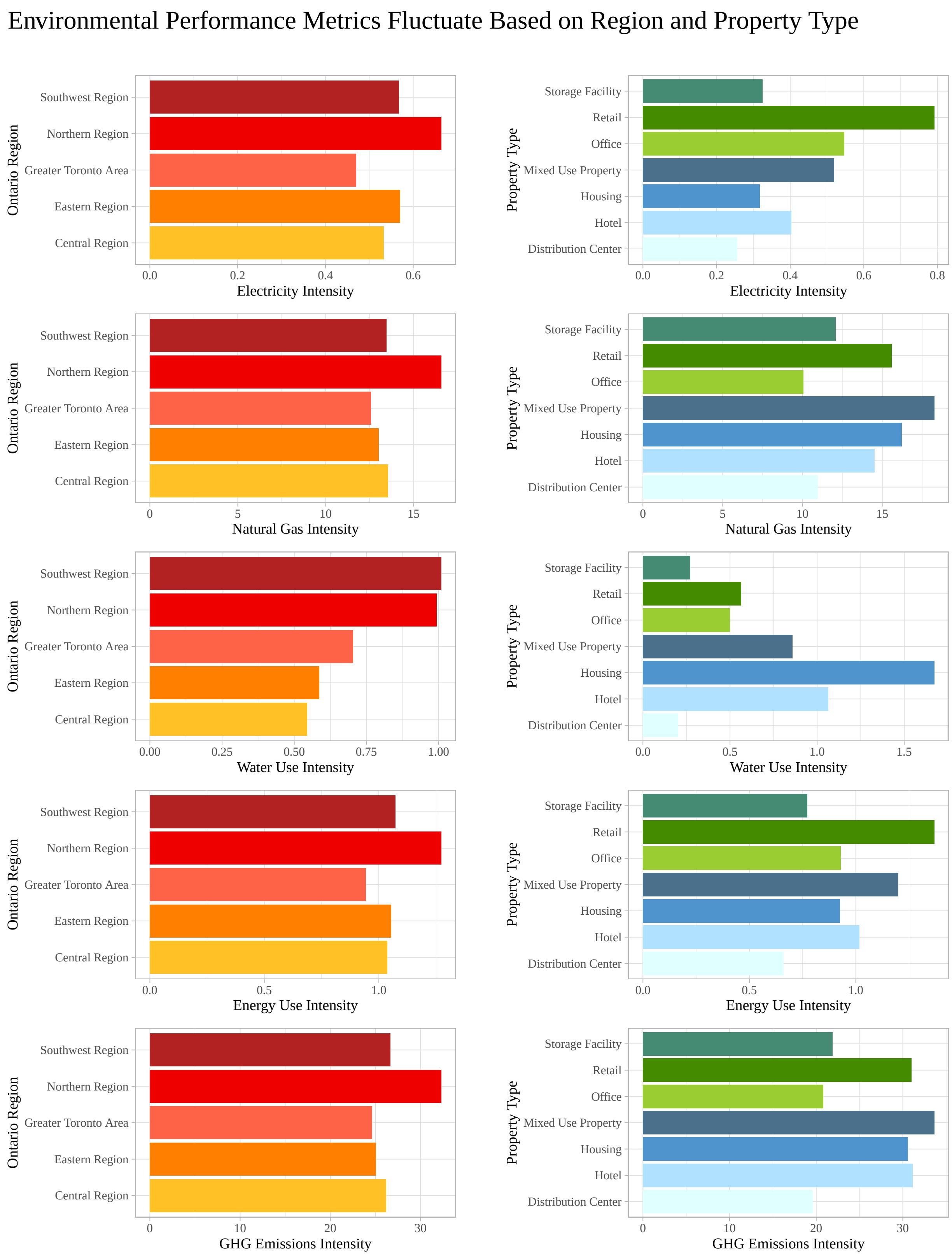

Figure 1: Environmental Metrics by Ontario Region and Primary Property Type

metrics in northern regions could be the respective abundances of these buildings within their regions. For example, large buildings in the northern region, which are found with a much lower degree of density, might have higher usage demands since they are less widespread but serve much larger geographic areas.

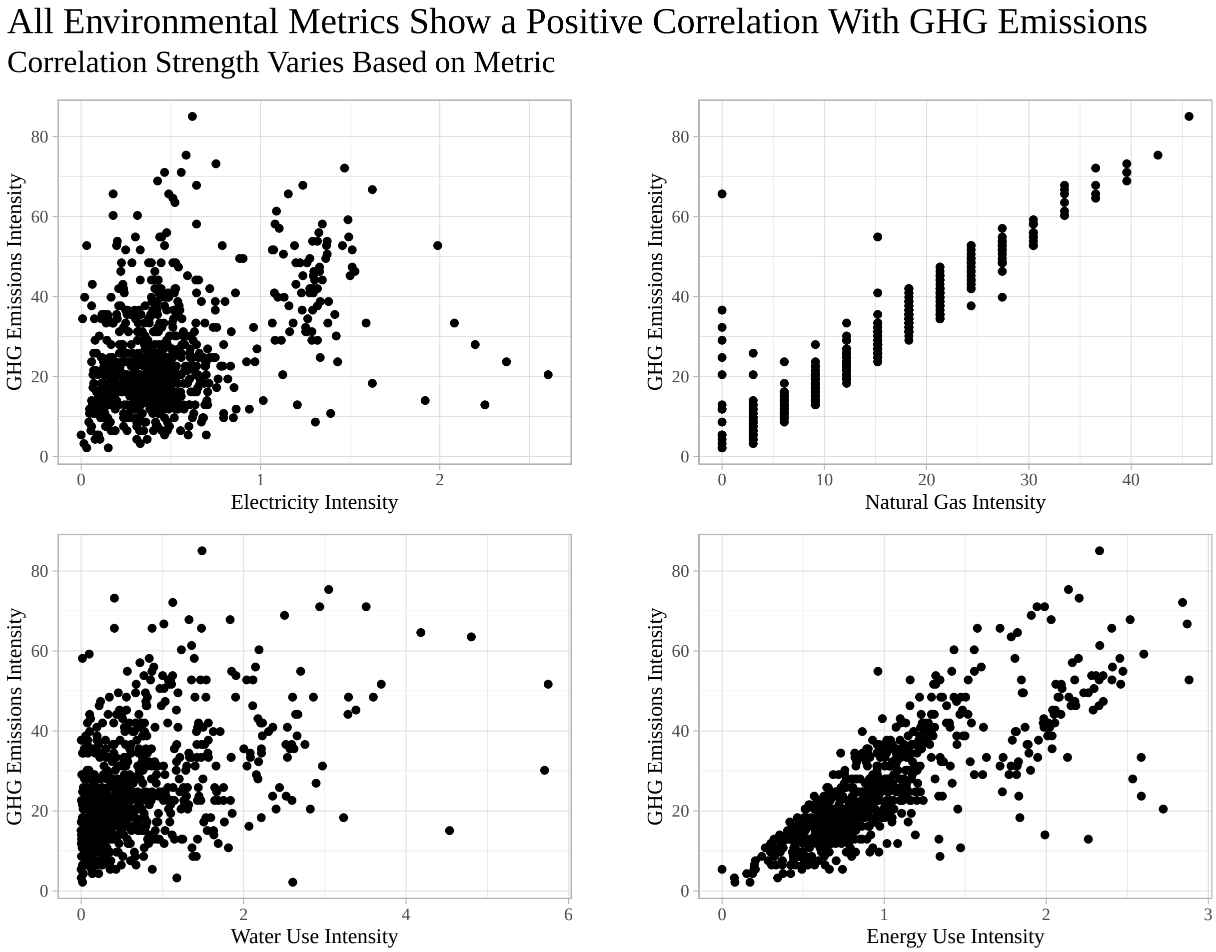

Figure 2: Relationship Between Environmental Performance Metrics and GHG Emissions

Another key relationship to consider between the variables in our dataset is how overall greenhouse gas emissions fluctuate based on different environmental metrics. Electricity, natural gas, water, and energy are among the most widely-used systems with known environmental implications due to their contribution to global emissions. Figure 2 depicts the relationship between GHG emissions and these four different metrics. The first graph shows the electricity use intensity on the x-axis and GHG emissions on the y-axis. From the trend observed in this graph, we can see a weak positive correlation between the two. As electricity use increases, total emissions generally tend to increase as well. We can observe a few outliers that do not follow this trend. For example, there are a few data points that correspond to very high levels of electricity use yet low levels of GHG emissions, and vice versa. A potential explanation for this could be the fact that some buildings might have extremely high levels of electricity use, whereas their other metrics might be very low, resulting in an overall low or moderate degree of greenhouse gas emissions. A similar pattern can be observed for water use intensity, which also shows a weak positive correlation. Natural gas intensity and energy use, on the other hand, show much stronger positive correlations with GHG emissions.

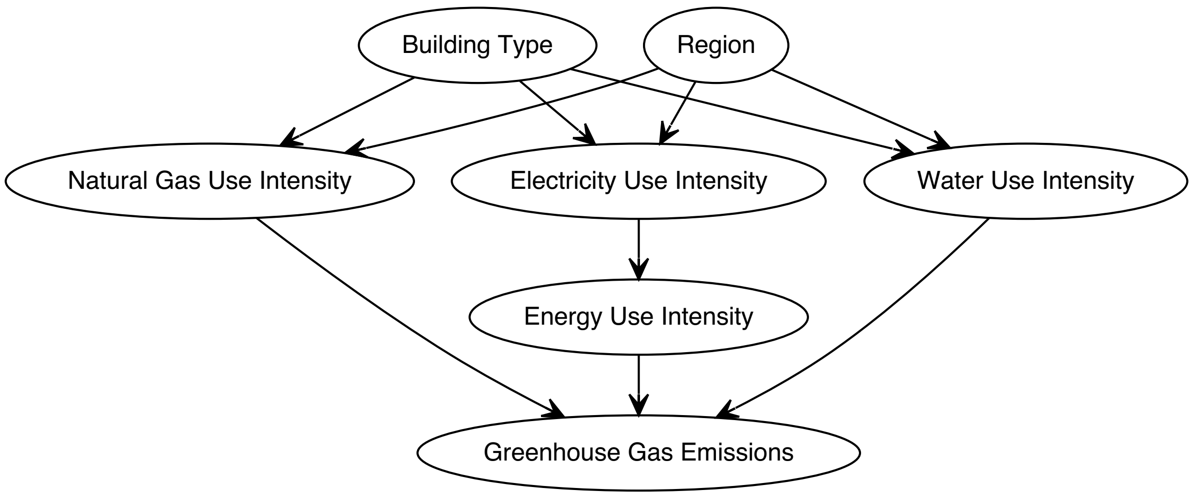

Figure 3: Expected Relationships Between Variables

Figure 3 shows a directed acyclic graph, or a DAG, which depicts our assumptions pertaining to the relationships between the variables. The first assumption is that building type and region primarily affect water use intensity, electricity use intensity, and natural gas use intensity. While we expect the natural gas intensity and water use intensity variables to directly influence total GHG emissions, we assume that electricity use intensity instead influences total energy use, and thus, indirectly influences total emissions. According to the data source, energy use for the buildings is measured as a combination of multiple different sources – one of which is electricity. Because of this, we expect electricity to have the greatest impact on total energy use, which then influences overall emissions.

3 Model

In the following section, we will perform multiple linear regression to estimate the relationship between greenhouse gas emissions of large buildings and six different independent variables: Ontario region, primary property type, electricity use, natural gas use, water use, and energy use.

This linear regression model will be used to measure the extent to which the response variable is dependent on the predictor variables. More specifically, we can use the model to analyze the relationship between GHG emissions and our independent variables to determine this relationship. Equation (1) outlines how we can define this relationship.

$$\begin{aligned} Y_i = \beta_0 &+ \beta_{11}X_{11i} + \beta_{12}X_{12i} + \beta_{13}X_{13i} + \beta_{14}X_{14i} \\ &+ \beta_{21}X_{21i} + \beta_{22}X_{22i} + \beta_{23}X_{23i} + \beta_{24}X_{24i} + \beta_{25}X_{25i} + \beta_{26}X_{26i} \\ &+ \beta_3X_{3i} + \beta_4X_{4i} + \beta_5X_{5i} + \beta_6X_{6i} + \beta_7X_{3i}X_{6i} + \epsilon_i \end{aligned} \tag{1}$$

The key variables of this model are the predictor variables, X_i, and the response variable, Y_i, for $i = 1, 2, ...n$. Since region is a categorical variable with five categories, we treated Central Region as the reference category. In this model, X_{11}, X_{12}, X_{13} and X_{14} refer to the Eastern Region, Greater Toronto Area, Northern Region and Southwest Region, respectively. Property type is another categorical variable and we treated Distribution Center as the reference category. X_{21}, X_{22}, X_{23} and X_{24} refer to the property types "Hotel", "Housing", "Mixed Use Property", "Office", "Retail" and "Storage Facility", respectively. The variables X_3, X_4, X_5 and X_6 are continuous. X_3 refers to the electricity use, X_4 refers to natural gas use, X_5 refers to water use, and X_6 is refers to energy use. Y refers to greenhouse gas emissions, measured as kilograms of carbon dioxide equivalent emitted per square meter ($\mathrm{kgCO_2e/m^2}$). We assume ϵ_i to be independent and identically distributed normal random variables with a mean of 0 and a constant variance σ^2.

The model also includes the parameters β_0 as the intercept, $\beta_{11}, \ldots, \beta_{14}$ as the coefficients for the dummy variables of Region, $\beta_{21}, \ldots, \beta_{26}$ as the coefficients for the dummy variables of Property Type and the coefficients $\beta_3, \ldots, \beta_6$ for the continuous variables electricity, natural gas, water use and energy use intensity. The intercept can be interpreted as the average change in Y if the values of all other predictors are set to zero. In the context of our data, this would refer to the average level of GHG emissions when all other predictors are equal to 0. The other regression coefficients, β_i, where $i > 0$, can be interpreted as the average change in Y for one unit change in X, controlling for other variables in the model. For example, β_4 can be interpreted as the average change in the level of GHG emissions for one unit change in the natural gas use intensity, keeping all other variables fixed.

Another issue to be considered is the potential for interaction among our explanatory variables. In the case of our data, energy use intensity is determined as a function of several different sources. These sources include diesel, propane, district energy, as well as electricity use. Since energy is defined as an aggregate of multiple different metrics, including one of the ones used in our model, we will also include an interaction term in our regression model to account for the probable correlation between the two variables. This interaction between electricity and energy usage is reflected in our model as the X_3X_6 term, with a coefficient of β_7. For every one unit increase in electricity usage, the average change in Y can then be determined to be $\beta_3 + \beta_7X_6$ units. The rationale behind keeping both of these individual metrics rather than eliminating energy use altogether is that this can preserve the interpretability of the model as it will allow us to evaluate the joint effect of each of these metrics rather than just observing the effects of a single metric acting in isolation.

Table 2: Summary Table of Multiple Linear Regression Results

	Model 1
(Intercept)	−0.572 (p = 0.005)*
RegionEastern Region	−0.432 (p = 0.009)*
RegionGreater Toronto Area	0.100 (p = 0.451)
RegionNorthern Region	0.499 (p = 0.178)
RegionSouthwest Region	0.066 (p = 0.739)
Primary_Property_TypeHotel	0.570 (p = 0.050)
Primary_Property_TypeHousing	0.976 (p = 0.000)*
Primary_Property_TypeMixed Use Property	−0.116 (p = 0.766)
Primary_Property_TypeOffice	0.408 (p = 0.007)*
Primary_Property_TypeRetail	0.353 (p = 0.032)*
Primary_Property_TypeStorage Facility	−0.074 (p = 0.648)
Electricity_Intensity	−46.550 (p = 0.000)*
Energy_Use_Intensity	51.776 (p = 0.000)*
Natural_Gas_Intensity	−0.211 (p = 0.000)*
Water_Use_Intensity	0.197 (p = 0.007)*
Electricity_Intensity × Energy_Use_Intensity	0.257 (p = 0.143)
R2	0.994
F	8987.550

4 Results

Our analysis set out to uncover some of the factors that pose the greatest threat to sustainability. In order to do this, we will analyze the results of the model to detect relationships between greenhouse gas emissions and several other variables that might be affecting this outcome. Table 2 shows the results from the regression output.

Although we are only interested in six independent variables, the regression table includes values for 15 variables. Firstly, since "Property Type" is a categorical variable that assumes seven different values, the model has defined six dummy variables to allow us to interpret the results in a meaningful way. Each "type" of property is represented as a dummy variable except for one, which is Distribution Centre in this case. As a result, the coefficients corresponding to each dummy variable allow us to interpret the results by elucidating how the change in GHG emissions differs in a given property type in relation to a reference property type. For example, the coefficient corresponding to Hotel is 0.570. What this tells us is that degree of GHG emissions we would expect to observe in a hotel is 0.570 units greater than what we would expect to observe in a distribution centre, assuming all other variables remain constant between the two. Likewise, "Region" is also a categorical variable with five different values, meaning our model has defined four dummy variables to facilitate the interpretation of this variable. The reference variable in this case is Central Region, and we can again use the coefficients of the dummy variables to observe their relative effects on total emissions. In order to establish the significance of these findings, we have defined a p-value of 0.05. P-values for coefficients below this threshold indicate that the effect of this variable is large as we have a low chance observing this result under the assumption that there is no interaction between the variables. This indicates that the Eastern Region dummy variable has a significant effect on GHG emissions. This variable has a coefficient of -0.432, which tells us that the GHG emission intensity for a building in this region is expected to be 0.432 units lower than what we would observe in a building in the Central Region. For the "Property Type" variables, we can see that the Housing, Office, and Retail dummy variables all have a significant effect on the outcome variable. Using the same rationale, we can expect that these three property types would correspond to increased GHG emissions compared to what would be expected in our reference category.

The other explanatory variables contained in our model are the continuous variables, which include our four environmental metrics. From the results we can see that all four of the metrics are significant in regards to their effect on GHG emissions. Firstly, we can observe that water use has a coefficient of 0.197, which corresponds to the expected change in GHG emissoins for a single unit change in water use intensity. On the other hand, we can observe that natural gas intensity has a negative coefficient, which suggests a low level of natural gas usage could be associated with increased GHG emissions. Furthermore, we can observe the results for the electricity and energy variables. The coefficients appear a bit unusual at first, as we can observe a coefficient value of 51.776 for energy use and a value of -46.550 for energy use. One possible explanation for this could be that a decrease in electricity use could be the result of higher levels of GHG emissions due to the fact that other metrics contributing to total energy use had a greater impact on emissions intensity. This implies that lower electricity use a higher use of other energy sources such as diesel and propane, have a greater effect on emission intensity.

Finally, the intercept can be interpreted as the value we would expect for our outcome when all values of our variables are equal to 0. The findings from this result are not particularly informative as they would require all the environmental metrics to be measured at 0, which is not a realistic outcome that we would expect to observe in practice.

5 Discussion

5.1 Findings and Implications

It has been well-established that greenhouse gases have a numerous catastrophic effects on the environment, primarily due to the gradual warming as a result of the greenhouse effect. This greenhouse effect is a major implication when considering global climate change as a whole. Because of this, climate change policies should aim to tackle the root of the issue. For this to happen, we need a comprehensive and thorough understanding of the key factors that contribute to the exacerbation of greenhouse gas concentrations in the atmosphere. This will aid in establishing a conceptual framework to better understand the primary drivers of climate change. It is important to also consider how current efforts to mitigate the greenhouse effect fall short. According to the annual UN Emissions Gap Report, GHG concentrations surpassed a record high in 2021 and rose at a rate that far exceeded the annual average from the last decade. It seems like we are constantly in the face of continual and worsening reports and statistics from climate change science. As a result, this prompts various initiatives advocating for climate action. In an attempt to emphasize the gravity of the situation, media, news outlets, reports, journals, etc. continue to push the narrative that the world is approaching a tipping point and we are going extinct. However, this approach is clearly ineffective, and there are several issues at play that could explain why. Firstly, risk perception varies on both a temporal and geographical basis. Climate change action requires some degree of sacrifice in return for a payoff sometime the future, however, the general trend is that immediate concerns take precedence over eventual issues. In addition, developed nations like Canada and the United States are better equipped to deal with climate instability and its effects when compared to less developed countries. As a result of this, our perception of risk is skewed. While there are many valid explanations for this ongoing inaction, the most prominent factor at play is the sense of futility. Climate change is arguably one of the biggest challenges we have had to face and in many cases this makes any type of action seem hopeless. As a result, this makes the problem seem overwhelming and insurmountable. The more hopeless the issue becomes, the less people will be willing to invest time and resources into real-world solutions. A better solution is to look at the data, which allows us to gain insight into what is causing the issue, and build policies around this instead. For this to work, we must find technological solutions and effective alternatives that help reduce usage and optimize consumption in a more sustainable manner.

As mentioned, our results indicate a strong correlation between four of the most common energy sources and an increase in greenhouse gas emissions. This aligns with current initiatives to push towards renewable energy. By shifting away from conventional energy sources and emphasizing the importance of renewable resources such as solar, wind, and geothermal power, we could begin to see a gradual reduction in emission on a global scale.

5.2 Limitations and Future Directions

While the results of our model did provide valuble insight, there is certainly aspects of the analysis that could be improved upon. It would be beneficial to our model results to know more information about the other metrics contributing to overall energy use intensity. Based on the results, we assumed that a low level of electricity use was associated with increased GHG emissions potentially because the other metrics involved in calculating total energy use had a greater individual effect on emissions intensity. As a result, lower electricity use, and thus, higher use of other energy sources such as diesel and propane, have a more extreme effect on emission intensity. If we had access to these metrics, it would allow us to verify these assumptions and make more sound inferences.

We've established an explanatory model which offers insight into existing relationships between several metrics and variables and their contribution to GHG emissions. In the future, we can look into creating a model for prediction, however this would require more data. These existing metrics could be used to model greenhouse gas emissions even when concrete emission data is unavailable or impractical to collect. This would allow us to make additional meaningful interpretations about GHG emissions.

6 Appendices

6.1 Appendix A

Table 3: Measurement Units used for Environmental Metrics

Variable	Unit of Measurement	Description
Electricity Use Intensity	GJ/m2	The annual average electricity, measured in gigajoules, used per square meter at a building.
Natural Gas Use Intensity	m3/m2	The annual average natural gas, measured in cubic metres, used per square metre at a building.
Water Use Intensity	m3/m2	The annual average water, measured in cubic metres, used per square metre at a building.
Energy Use Intensity	GJ/m2	The total annual average energy use of all types, measured in gigajoules, per square metre at a building

6.2 Appendix B

Cities in Ontario are grouped into 5 regions: Central, Eastern, Northern, Southwestern, and Greater Toronto Area based on the following regional breakdown ("Definitions of Ontario Regions," n.d.):

Greater Toronto Area:

- Ajax, Aurora, Bolton, Bowmanville, Brampton, Burlington, Concord, Don Mills, East Gwillimbury, East York, Etobicoke, Georgetown, Halton Hills, Markham, Milton, Mississauga, Newmarket, North York, Oakville, Oshawa, Pickering, Richmond Hill, Scarborough, Thornhill, Toronto, Vaughan, Whitby, Woodbridge, York

Central Region:

- Ancaster, Barrie, Brantford, Cambridge, Dundas, Fergus, Guelph, Hamilton, Kitchener, Niagara Falls, Orillia, Paris, Peterborough, Port Colborne, Simcoe, St. Catharines, Stoney Creek, Waterdown, Waterloo, Welland

Southwestern Region:

- Chatham, Corunna, Grey Highlands, Hanover, Kincardine, London, Markdale, Owen Sound, Sarnia, St. Thomas, Stratford, Strathroy, Windsor, Woodstock

Eastern Region:

- Arnprior, Barrhaven, Belleville, Carp, Gloucester, Johnstown, Kanata, Kingston, Nepean, Orleans, Ottawa

Northern Region:

- North Bay, Sault Ste. Marie, Sudbury, Thunder Bay, Timmins

References

Arel-Bundock, Vincent. 2022. *Modelsummary*. https://cran.r-project.org/package=modelsummary.

Buis, Alan. 2020. "A Degree of Concern: Why Global Temperatures Matter – Climate Change: Vital Signs of the Planet." *NASA Global Climate Change*. NASA. https://climate.nasa.gov/news/2865/a-degree-of-concern-why-global-temperatures-matter/.

Chang, Winston. 2022. *Webshot*. https://cran.r-project.org/package=webshot.

"Climate Change Indicators: Greenhouse Gases." n.d. *Environmental Protection Agency*. Environmental Protection Agency. https://www.epa.gov/climate-indicators/greenhouse-gases.

"Definitions of Ontario Regions." n.d. *Construction Forecasts*. https://www.constructionforecasts.ca/en/reference/definitions-ontario-regions.

"Energy and Water Usage of Large Buildings in Ontario." 2021. *Ontario Data Catalogue*. https://data.ontario.ca/dataset/energy-and-water-usage-of-large-buildings-in-ontario.

Fox, John. 2021. *Car*. https://cran.r-project.org/web/packages/car/index.html.

Griffiths-Sattenspiel, Bevan, and Wendy Wilson. 2009. "The Carbon Footprint of Water." *Chicago State University*. Chicago State University. https://www.csu.edu/cerc/researchreports/documents/CarbonFootprintofWater-RiverNetwork-2009.pdf.

Iannone, Richard. 2022. *DiagrammeR: Graph/Network Visualization*. https://cran.r-project.org/package=diagrammeR.

Laborde, David et al. 2021. "Agricultural Subsidies and Global Greenhouse Emissions." *Nature Communications* 12 (2601). https://doi.org/https://doi.org/10.1038/s41467-021-22703-1.

Ladislaw, Sarah. 2022. "Climate Solutions Series: Decarbonizing Heavy Industry." *Center for Strategic and International Studies*. Center for Strategic; International Studies. https://www.csis.org/analysis/climate-solutions-series-decarbonizing-heavy-industry.

"Net-Zero Emissions by 2050." 2022. *Canada*. Government of Canada. https://www.canada.ca/en/services/environment/weather/climatechange/climate-plan/net-zero-emissions-2050.html.

"O. Reg. 506/18: REPORTING of Energy Consumption and Water Use." 2018. *Ontario*. https://www.ontario.ca/laws/regulation/180506.

Pederson, Thomas Lin. 2020. *Patchwork: The Composer of Plots*. https://CRAN.R-project.org/package=patchwork.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2021a. *Bookdown: Authoring Books and Technical Documents with R Markdown*. https://CRAN.R-project.org/package=bookdown.

———. 2021b. *Knitr: A General-Purpose Package for Dynamic Report Generation in R*. https://CRAN.R-project.org/package=knitr.

Yoshida, Kazuki. 2022. *Tableone*. https://cran.r-project.org/package=tableone.

Zhu, Hao. 2021. *KableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.