TITLE

# A Marginal Emissions Greenhouse Gas Inventory Methodology for Buildings

AUTHORS

Bernie McIntyre[1], Erik Janssen[1,] Aidan Brookson[1], Jason Choy[1], Leigh St. Hilaire[1], Dana Fountain[1]

[1]Toronto and Region Conservation Authority (TRCA), Vaughan, ON, L4K 5R6, Canada



## ABSTRACT

Quantifying greenhouse gas (GHG) emissions is a key part of reducing them. However, current GHG emission inventory methodologies underestimate the GHG emission reductions associated with electricity conservation because they rely on annualized emission factors that neglect important dynamic aspects of the electricity grid. Decision makers may therefore not take advantage of opportunities for cost-effective GHG emission reductions through electricity conservation – some of which may include simple no-cost/low-cost changes to building controls – because there is no way to track the reductions using the current inventory approaches. This paper addresses this issue. It consists of two parts. Part A provides a simple and transparent methodology for estimating the hourly marginal emission factor (MEF) for Ontario grid electricity using easily accessible data from the Independent Electricity System Operator (IESO). Part B provides a methodology for incorporating the MEF into annual GHG emission inventories for specific buildings. Part B also provides an example inventory in an actual building using the proposed methodology. Overall, the paper provides an improved methodology for GHG emission inventories in buildings which can be used to identify, quantify, and promote cost-effective GHG emission reduction opportunities through electricity conservation. It is recommended that organizations and building owners consider the proposed methodology in their buildings to identify new opportunities and properly quantify the impacts of current electricity conservation efforts.

## 1. INTRODUCTION

In-line with the international community, different levels of government across Canada have committed to aggressive greenhouse gas (GHG) emission reductions to mitigate the impacts of climate change. The GHG emissions impact of an organization's activities can be determined using different methodologies that vary in complexity and GHG emission calculations results can vary greatly depending on the assumptions about the emission factor (EF) of grid electricity. As an example, Figure 1 displays three different grid EFs for the Ontario electricity grid, derived by the Atmospheric Fund (TAF) using 2018 Independent Electricity System Operator (IESO) data [1]. The differences are large.

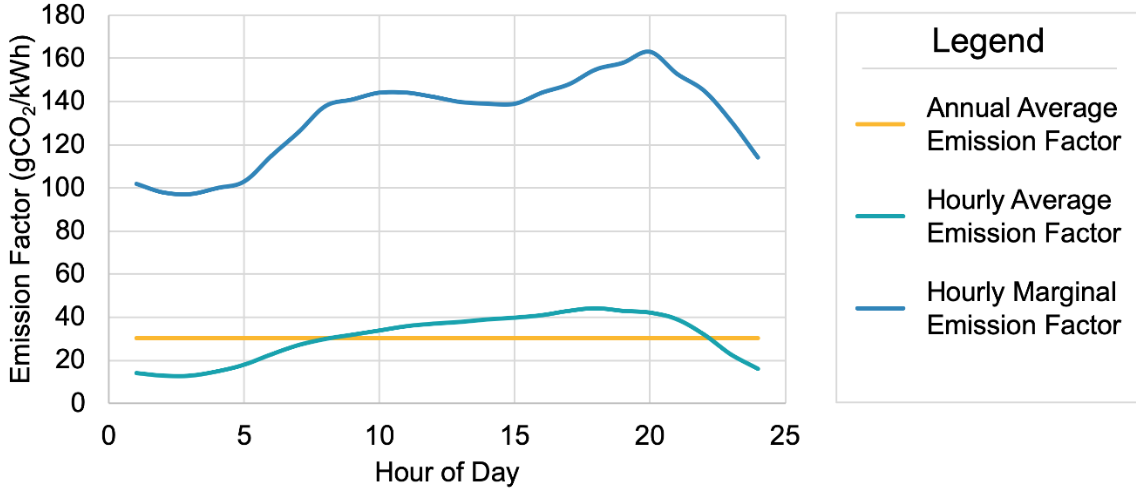


*Figure 1. Comparison of Ontario grid emission factors for the year 2018 based on TAF data.*

The simplest methodology uses an *annualized emission factor* (AEF). This number is commonly obtained from the National Inventory Report (NIR) [2] published every year by The Government of Canada, which summarizes the annual average EF for grid electricity for every province and territory two years prior to the date of publishing. While utilizing the AEF is convenient, it has shortcomings.

Each NIR provides grid EF values for two years prior to the date of publishing, meaning that any emission projections derived using these numbers are at best two years out of date. The 2019 NIR published an Ontario grid EF of approximately 20 $gCO_2/kWh$ for the year 2017 [2], however, an analysis of 2018 generation data indicates that the NIR EF will increase to approximately 30 $gCO_2/kWh$ [1]. This is a significant increase in just one year and can lead to significant differences in GHG emission estimations for the same activity.

Nevertheless, if an annual averaged EF is derived using more recent data, it still neglects the change in the generation profile of the grid on a seasonal, daily, or hourly basis. In Ontario, the IESO publishes hourly generation data for the grid. This data, in conjunction with generator or fuel-type emission factors, can be used to develop an *hourly average emission factor* (HEF) for every hour of the year.

If hourly or sub-hourly metered electricity data is available for a given building, then utilizing an HEF will provide a greater degree of accuracy than using an AEF. Additionally, since the IESO publishes generation data live, HEFs can reflect current values, and in some instances be predicted based on IESO supply and demand forecasts.

However, both AEF and HEF methodologies neglect the impact that changing demand would have on the generation profile of the grid. Since the supply and demand of electricity must balance, a variation in demand will result in a variation in supply. The type of generator(s) that responds to a change in demand (referred to as a generator *on the margin*) will then define the emissions associated with that change in demand.

In Ontario, the two main generator types that respond to changes in demand are hydro and natural gas. Nuclear facilities have minimal to no capability to ramp-up or down, and other renewables such as wind and solar are generally uncontrollable (although newer technologies have the ability to be curtailed). Biofuel facilities can ramp up or down their generation but represent a negligible portion of the overall generation mix, accounting for 0.4% of generation in 2017 [3].

If a proposed activity is to, for example, increase loads on the grid by 100 kW from 9:00 AM to 3:00 PM, an hourly *marginal* emissions analysis will determine the emissions impact of that activity based on the generators that would respond to the proposed increased demand. This can produce a range of results, as generators that are on the margin frequently change. For example, if it was determined that hydro generation was increased to meet an increase in demand, then the proposed action would have a relatively small emissions impact. However, if natural gas generation were to increase, then the emissions impact would be far greater than an annual average or hourly analysis would imply.

This information is captured in marginal emission factors (MEFs). The Greenhouse Gas Protocol, developed by the World Resources Institute, recommends that a marginal approach (referred to as "operating margin emissions", or "OM emissions") is the preferred approach. From Chapter 10 of the Greenhouse Gas Protocol: Guidelines for Quantifying GHG Reductions from Grid-Connected Electricity Projects [4]:

> *The ideal method to estimate operating margin (OM) emissions would be to identify precisely which power plants on a grid are backed down in response to the project activity's operation. In practice, this is difficult if not impossible to do. Various methods can be used to approximate OM emissions, each with advantages and disadvantages concerning accuracy and ease of use.*

As the protocol points out, a marginal analysis can be quite difficult to conduct, particularly when attempting to predict the future state of the grid, as the approach requires detailed knowledge of the grid's operating characteristics. However, many utilities provide both generation and demand data, which can be used to analyse historical grid characteristics for the purpose of performing emission inventories

MEFs have been calculated for the Ontario Grid by TAF. It's important to note that TAFs reporting [1] states that: "… the IESO makes several sets of data available to the public (including a series of public reports and a data directory), but this data is not enough to accurately determine marginal values (confirmed by IESO)." TAF has done their own analysis to estimate hourly MEFs but their methodology is not described in detail and is therefore not easily reproduceable. There could be important benefits if MEF calculations were sufficiently straightforward that any organization was able to take the publicly accessible real-time grid data and make an informed estimate on the MEF for any given hour.

While a marginal emissions analysis can produce more realistic emissions estimates than annual averaged or hourly analyses, it should be noted that a marginal analysis *cannot* be applied to an organization's "business-as-usual" emissions inventory - at least not as is conventionally calculated. By its very nature, a marginal analysis attempts to quantify the impact of an activity or project relative to business-as-usual. MEFs must be applied to proposed or achieved changes in grid consumption, rather than, a building's electricity consumption for a single year. It is for this reason that TAF recommends that AEF values be used in inventories [1].

This creates an obvious problem. On one hand, it is recommended that the impact of a specific ECM be calculated using MEFs to accurately capture real-world grid effects. In Ontario, this means that the calculated *GHG emission reductions will typically increase significantly* in magnitude. On the other hand, it is recommended that inventories use the AEF, but this will *significantly underestimate the actual impacts* of electricity conservation. Inventories are the main tool used by organizations to track GHG emissions and the overall progress towards a GHG reduction target. If GHG reductions cannot be tracked in an inventory, then opportunities for cost-effective GHG reductions through electricity conservation will be missed.

This paper addresses this issue. It consists of two parts. Part A provides a simple and transparent methodology for estimating the hourly MEF for Ontario using easily accessible data from the IESO. Part B provides a methodology for incorporating the MEF into annual inventories for specific buildings. Part B also provides an example inventory in an actual building using the proposed approach. Overall, the paper provides a more accurate methodology for GHG inventories which can be used to identify, quantify, and promote cost-effective GHG reduction opportunities through electricity conservation.

# 2 METHOD

## 2.1 Part A: Methodology for calculating hourly MEF in Ontario

This work is based entirely on publicly available IESO data. The analysis is designed to be performed on one year of data at a time, although the methodology can be applied to smaller datasets as well. All data was obtained from the IESO data directory [5]. The specific data sets are outlined below.

The fuel types of IESO generators have been compiled from two IESO sources:

1. Report on Generator Output and Capability Data [6]
2. IESO Active Generator Contract List [7]

The first source generally contains all the necessary information, however, due to occasional discrepancies in naming conventions, the second source can be used to validate assumptions.

Generator emission factors by fuel type have been obtained from International Panel on Climate Change (IPCC), National Renewable Energy Laboratory (NREL), and NIR reports. Table 1 and Table 2 summarize emission factors by fuel type for combustion-based and life cycle assessment-based (LCA-based) analyses, respectively. The combustion-based EFs are used in this report for greater consistency with current approaches but LCA values could be incorporated in future work.

*Table 1. Electricity generation emission factor by fuel type for combustion-only analysis*

| Fuel Type | Generation Emission Factor ($gCO_2/kWh$) | Source |
|---|---|---|
| Nuclear | 0 | - |
| Hydro | 0 | - |
| Gas | 448 | Derived from historical NIR reports |
| Wind | 0 | - |
| Solar | 0 | - |
| Biofuel | 0 | - |
| Coal | 1103 | Derived from historical NIR reports |

*Table 2. Electricity generation emission factor by fuel type for life cycle assessment (LCA) analysis*

| Fuel Type | Generation Emission Factor ($gCO_2/kWh$) | Source |
|---|---|---|
| Nuclear | 12 | [8] |
| Hydro | 4 | [9] |
| Gas | 469 | [9] |
| Wind | 11 | [10] |
| Solar | 32 | [11] |
| Biofuel | 18 | [9] |
| Coal | 980 | [12] |

Generator Output and Capability (GOC) data contains hourly output and capability data for generating facilities with a maximum output capability of 20 MW or more. It can be found under the *Supply* tab of the IESO data directory. Output data represents the hourly energy output of each generator in the IESO-administered market. Forecast capability values are provided for variable generation facilities (wind and solar).

The Ontario and Market Demand Reports contain "total energy and operating reserve scheduled, and Ontario demand, as established by the constrained run of the IESO's Dispatch Scheduling and Optimization (DSO) algorithm." It can be found under the *Featured Reports* tab of the IESO data directory. Columns in this dataset are defined, by the IESO, as:

Total Energy: Total energy dispatched into the IESO-controlled grid, calculated as Ontario generation plus imports (referred to as "Market Demand" in the Demand Report)

Ontario Demand: Total Ontario electricity demand, calculated as:

*Total Energy + Total Generation Without Offers – Total Exports + Total Off Market +/- Over/Under Generation*

In this work, "Ontario Demand" is used to represent demand for the province. At this time, imports are not explicitly accounted for in the marginal emission factor calculation methodology. In 2017, Ontario supplied 144 TWh of electricity [3] and only 6.6 TWh [13] was from imports, the bulk of which comes from Quebec, a region whose NIR-reported electricity emission factor was merely 1.5 $gCO_2/kWh$ in 2017 [14]. Due to the relatively low fraction of imports, they have not, at this time, been incorporated into the marginal emission calculation procedure.

Using the data described above, the steps to calculating the *HEF* are as follows:

1. Calculate the total energy ($E_{tot,i}$) produced in a given i$^{th}$ hour by summing the output of all generators (output from individual generators is $E_{G,i}$ where "G" is a subscript denoting a given generator) for that hour. See Equation 1.

$$E_{tot,i} = \sum_{G} E_{G,i} \quad (1)$$

2. Calculate the total emissions ($GHG_{tot,i}$) produced in a given i$^{th}$ hour by multiplying the hourly energy produced by every generator with the emission factor of that generator's fuel type. For simplicity, combustion-only emission factors have been assumed in this work (Table 1), although LCA EFs could be used in future work. See Equation 2.

$$GHG_{tot,i} = \sum_{G} E_{G,i} \cdot EF_G \quad (2)$$

3. Divide the total GHGs emitted by all generators on the grid by the total supplied energy for that hour. The result is the hourly emission factor of the grid for that hour. See Equation 3. This is then repeated across all hours of the year.

$$HEF_i = \frac{GHG_{tot,i}}{E_{tot,i}} \tag{3}$$

The steps to calculating the *MEF* in the proposed methodology are as follows:

1. Calculate the *change* in generation ($\Delta E_{G,i}$) for each generator for each i$^{th}$ hour (Equation 4).

$$\Delta E_{G,i} = E_{G,i} - E_{G,i-1} \tag{4}$$

2. Calculate the *change* in Ontario demand ($\Delta D_i$) for the same hour (Equation 5).

$$\Delta D_i = D_i - D_{i-1} \tag{5}$$

3. Identify any generators that should be excluded from the calculation for that hour. For example, if it is known that a given generator is self-scheduling, or completely inflexible, it should be excluded from calculations. At this time, only natural gas, hydro, and coal generators were included in the calculation.
4. Compare the sign (positive or negative, increase or decrease) of each generator's change in generation to the change in Ontario demand. Any change in generation with the *same* sign as the change in demand, that has not been excluded for other reasons, indicates that that particular generator can be assumed to be on the margin for that given hour. Exclude all other generators. The function $\alpha_{G,i}$ (Equation 6) is equal to 1 if a generator is assumed to be on the margin, and 0 if it is not.

$$\alpha_{G,i} = \begin{cases} 1 \; if \; \mathrm{sgn}\left(\Delta E_{G,i}\right) = \mathrm{sgn}(\Delta D_i) \\ 0 \; if \; \mathrm{sgn}\left(\Delta E_{G,i}\right) \neq \mathrm{sgn}(\Delta D_i) \end{cases} \tag{6}$$

5. Calculate the total marginal generation across all generators assumed to be on the margin (Equation 7). Calculate the total marginal GHG emissions associated with that marginal generation (Equation 8). For simplicity, combustion-only emission factors (denoted by $EF_G$) have been assumed in this work (Table 1).

$$E_{marginal,i} = \sum_{G} \Delta E_{G,i} \cdot \alpha_{G,i} \tag{7}$$

$$GHG_{marginal,i} = \sum_{G} \Delta E_{G,i} \cdot \alpha_{G,i} \cdot EF_{G} \tag{8}$$

6. Divide the total marginal GHG emissions by the total marginal generation for that hour (Equation 9). The result is the marginal emission factor of the grid for that hour. Repeat the calculation for all hours.

$$MEF_{i} = \frac{GHG_{marginal,i}}{E_{marginal,i}} \tag{9}$$

The calculation is a weighted average. It considers the GHG emissions associated with the generators that appeared to respond to a change in demand. The weighting for a given generator is the extent to which the generator responded (i.e. how much a given generator's output changed in response to a change in demand). If a given generator responded strongly then it's associated GHG emissions were weighted more strongly, and similarly for weakly responding generators.

Put differently, the two key assumptions of this procedure are:

1. A generator is assumed to be on the margin, in a given hour, if it changed its output in the same direction as the change in demand.
2. The extent to which a given generator is on the margin, relative to all other generators, can be estimated by the magnitude of its change in output in response to the change in demand (i.e. say demand increased and Generator A changed its output by 5 kW while Generator B changed its output by 10 kW; then the EF associated with the electricity produced by Generator B is weighted as two times greater than that from Generator A when calculating the MEF for that hour).

It should be noted that these assumptions are not strictly true at all the time. Generators may increase or decrease in demand for reasons other than changes in demand. This may be due to existing contracts. Some generators may also ramp up in advance of an anticipated peak, and a generator's output may also change during maintenance periods or during testing. A robust defence of the methodology given these factors is not offered here. Rather, the efficacy of the methodology, and the inherent assumptions, is evaluated in the Results section where the MEFs from the proposed methodology are compared against those calculated by TAF in [1].

### 2.2 Part B: Methodology for Incorporating MEF into Annual Inventories for Buildings

MEFs must be applied to changes in electricity consumption. For example, if an ECM reduces energy consumption by a known amount then calculations based on MEFs can be used to calculate the corresponding reduction in emissions. The calculations take into account *what actually occurred* on the electricity grid with *what would have occurred* had there been no ECM - specifically, that the generators on the margin would have increased their output.

In contrast, common practice for annual GHG inventories are to consider only what *actually occurred*. By not considering *what would have occurred* without the ECMs, annual GHG inventories underestimate the benefits of ECMs. Knowledge of the hourly MEF, as calculated according to Section 2.1, is a first step towards addressing this issue. The next step is to incorporate the MEF into GHG inventory calculations. To do this, the energy consumption in any given year (termed the "target year" in this paper) needs to be referenced against a baseline energy consumption to quantify any changes.

The baseline energy consumption represents the energy consumption *that would have occurred* had there been no ECMs, and no changes to the building or it's usage, that would have impacted energy consumption. The baseline energy consumption for any given target year is based on an hourly model of the baseline building energy consumption. The model is developed using data from a baseline year and the baseline year is chosen as the year against which future emissions are referenced. As an example, an organization may have a target of reducing carbon emissions by 50% from 2010 levels by the year 2030 (or similar). In this case, 2010 would be selected as the baseline year.

Continuing the example, the hourly model of the baseline building energy consumption is trained using 2010 data. The model takes as inputs any relevant independent variables that impact energy consumption – for example, this might include the time of day, the day of the week, and the outdoor temperature – and provides the estimated building energy consumption for a given hour as an output. An inventory for 2011 (or any subsequent target year) would then compare the hourly energy consumption *that did occur* to the hourly energy consumption predicted by the baseline model given the independent variables for the 2011 year (i.e. *what would have happened* if there were no changes to the building, or its usage, that impacted energy consumption).

If there were no changes to the building, or its usage, then the actual energy usage and the predictions from the baseline energy model ought to closely align. If there were changes, then they would not align. The resulting GHG emission impacts from the changes can be calculated by applying the hourly MEFs to the hourly difference between actual and baseline consumption. This approach is formalized below.

The model of the baseline building energy consumption for the $i^{th}$ hour in the year is described in Equation 10. It is trained using data from the baseline year and may take into account multiple independent variables ($x_1, x_2, \dots, x_j$) as discussed above.

$$E_{baseline,i} = f(x_1, x_2, \dots, x_j) \tag{10}$$

The change between the actual energy consumption for a given target year ($E_{target,i}$), in a given hour, and the corresponding baseline energy consumption is given in Equation 11.

$$\Delta E_{target,i} = E_{target,i} - E_{baseline,i} \tag{11}$$

For the purpose of inventorying GHG emissions. The actual energy consumption for a given target year can be expressed as in Equation 12 (rearranging Equation 11).

$$E_{target,i} = E_{baseline,i} + \Delta E_{target,i} \tag{12}$$

In Equation 12, the first term *does not* represent a change in electricity so the HEF is applied when calculating GHG emissions. In contrast, the second term *does* represent a change in energy so the MEF is applied when calculating GHG emissions. This is shown in Equation 13.

$$GHG_{target,i} = E_{baseline,i} \cdot HEF_i + \Delta E_{target,i} \cdot MEF_i \tag{13}$$

The electricity GHG emissions inventory for a given year is given in Equation 14.

$$GHG_{target,tot} = \sum_{i=1}^{8760} GHG_{target,i} \tag{14}$$

In comparison to the currently used inventory methodology based on AEFs, the benefit of this approach is that it more accurately incorporates the beneficial impacts of ECMs into emissions inventories. However, it also presents new challenges; like how to determine the baseline model and deal with modelling errors, or the fact that if the change in energy is great enough it can drive the annual GHG emissions to negative values. It also raises questions regarding the limits of the applicability; for example, a newly constructed building would have an energy consumption of zero in the baseline year. These and other issues will be discussed in further detail in Section 4.

# 3 RESULTS

## 3.1 Part A: Hourly MEF in Ontario

The methodology outlined in Section 2.1 was applied to the Ontario grid using generation data from 2010 to 2018. A histogram of the hourly MEF is shown in Figure 2 for each year. The emission factors decrease in magnitude moving away from 2010 because coal was phased out of the grid mix. From 2014 onwards, the maximum MEF for any given hour is 448 g $eCO_2$/kWh due to gas generators being predominately on the margin. The minimum MEF for any given hour is 0 g $eCO_2$/kWh due to hydro being on the margin. The MEF reaches greater values than the HEF because, in the MEF calculation, the clean baseload is removed.

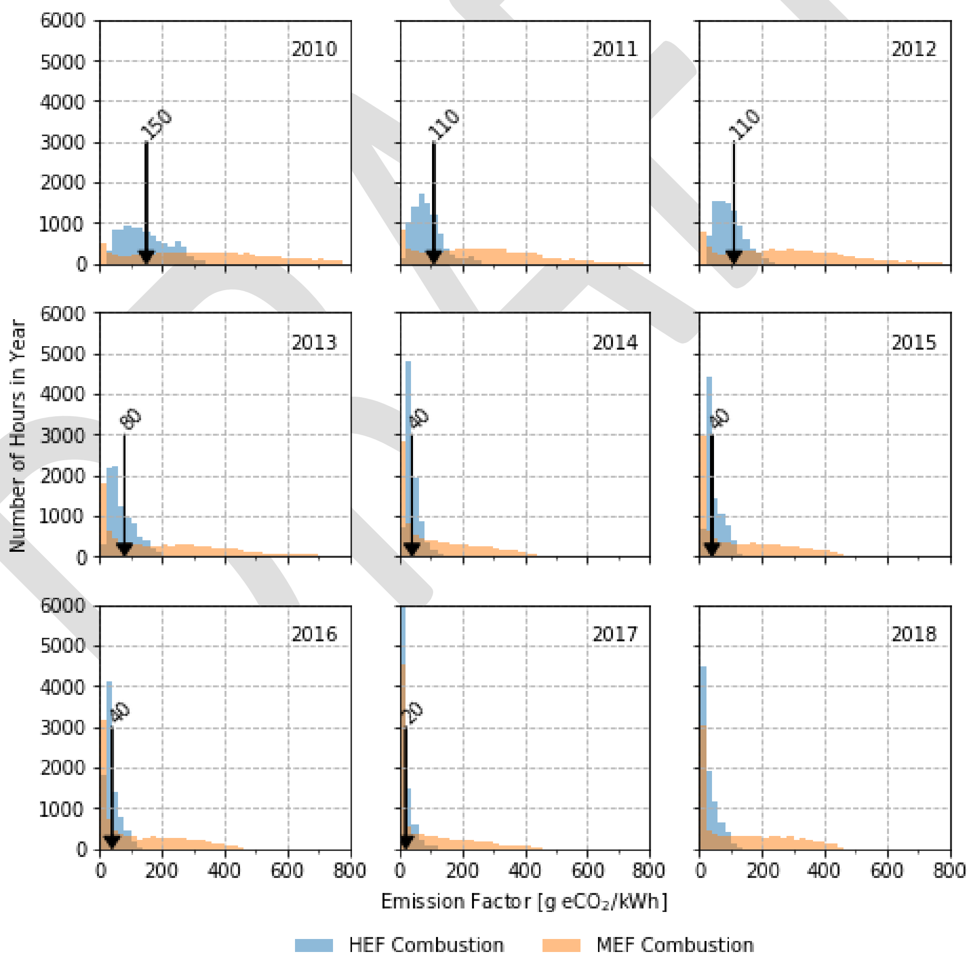


*Figure 2. Histogram of Ontario electricity grid hourly MEF and HEF from 2010 to 2018 using combustion only emission factors (i.e. not life-cycle emission factors). The annualized emission factors from the NIR are indicated by the black arrows.*

Figure 3 shows how the MEF, HEF, and AEF diverge on an hourly, daily, and seasonal basis using data from 2017. Seasonal variations are pronounced. During the shoulder months (April, May, June and October) the MEF is lowest. The MEF is highest in summer but still somewhat high in the winter in relation to the HEF and AEF.

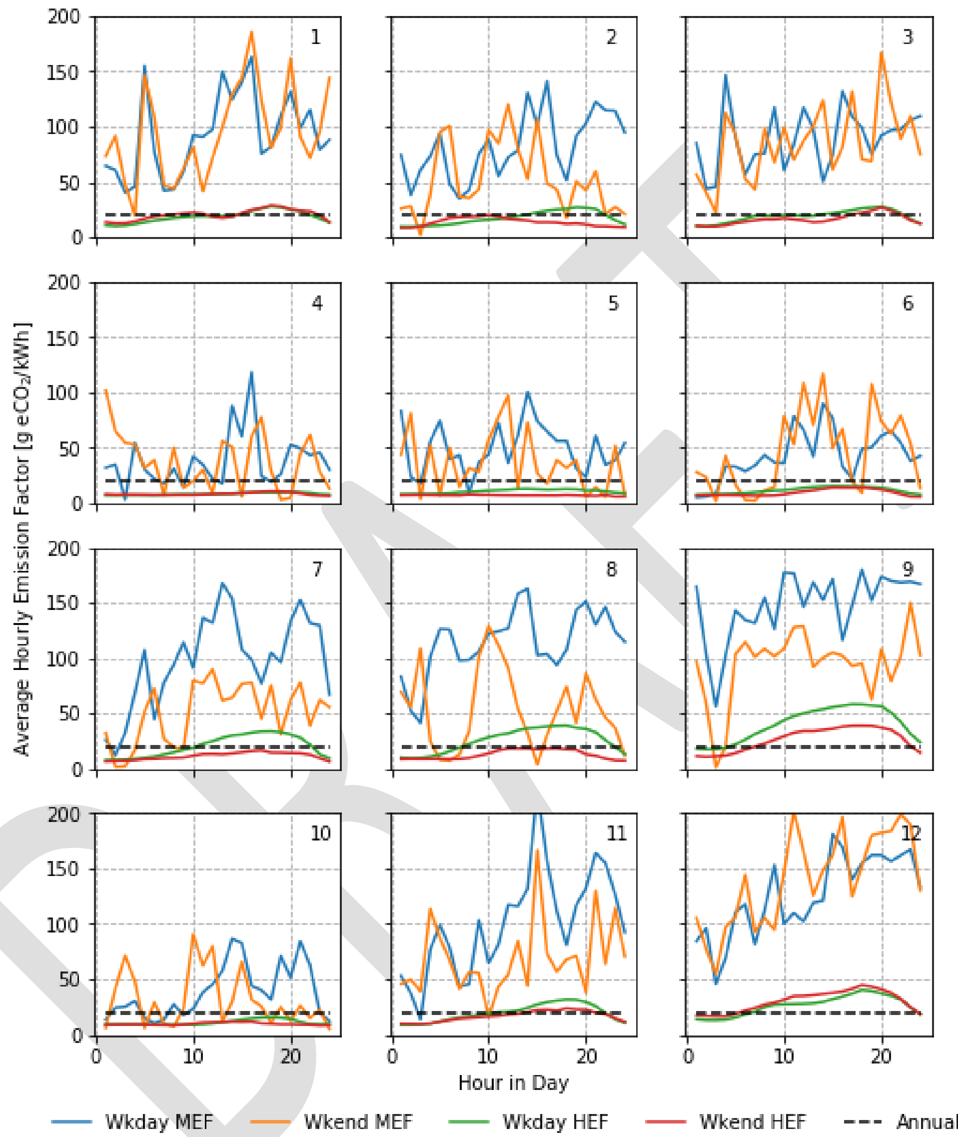


*Figure 3. MEF, HEF and AEF for 2017, aggregated according to hour, weekday/weekend, and month (month is indicated in upper right hand corner of subplots).*

The MEF results from this methodology were then compared against those from TAF available in [1]. The highest granularity data provided by TAF is the average MEF for each hour of the day aggregated according to season for 2018 (i.e. each season has 24 data points, one for each hour of the day). The results from the proposed methodology were aggregated in the same way for the sake of comparison (Figure 4). The MEF aggregated according to season only (i.e. 1 MEF value averaged per season) is also shown for the sake of comparison (Figure 5).

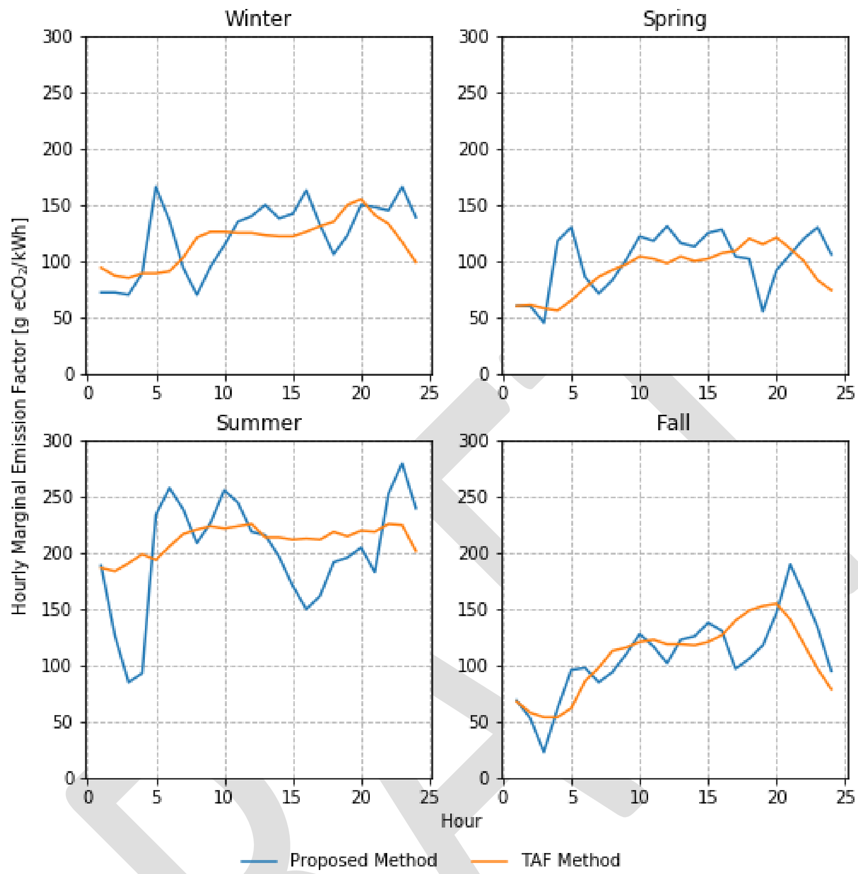


*Figure 4. The calculation results from the proposed methodology are compared against the corresponding results from TAF. Data is from 2018 and aggregated according to both season and hour of day.*

There are two important observations. According to seasonal aggregation of the MEF, both the TAF method and the proposed produce the same values. However, the proposed method has more variability on an *hourly* basis. As an example, in Figure 4 for hour 3 in the summer, the MEF drops to a low value according to the proposed methodology. This is because, during that hour, the change in demand was met primarily by a change in hydro generators. The TAF methodology predicts that the hourly MEF would be relatively unchanged.

There is not enough information to explain why the TAF method would still consider gas to be on the margin in that example scenario. Regardless, it stands that the two methods produce similar results and, to an extent, that ground-truths the proposed methodology. The benefit of the proposed methodology is that it is simple and transparent. Anybody can do it at any time, including in real-time if desired.

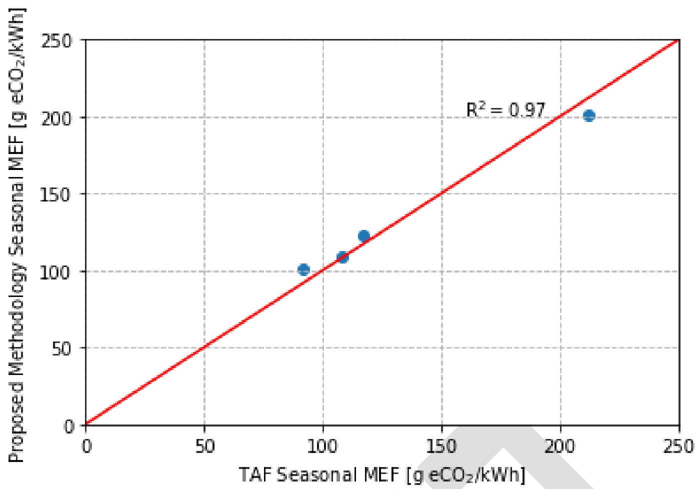


*Figure 5. The emission factor calculation results from the proposed methodology are compared to the corresponding results from TAF on a seasonal basis. Data from 2018 was used. The results are in close agreement.*

Overall, it's clear that the grid emissions change frequently. When looking for opportunities to reduce emissions, assuming a single annualized emission factor for electricity will lead to a much different result than an analysis that adequately took these changes into account. The extent of the difference is related to *when* the electricity is reduced and this is dependent on the nature of different ECMs. An exhaustive evaluation of different ECMs given these MEF results is beyond the scope of this paper. However, it is worthwhile to look at least one ECM with electricity reductions that are straightforward to model – solar photovoltaics (PV).

This example assumed a 50 kWp zero-export PV system installed in Toronto Ontario. It directly offsets the electricity consumption of a building, but it was assumed to be sufficiently large such that all the electricity could be utilized (i.e. at no point was the PV generation greater than the building electricity load). The installation was modelled using System Advisor Model (SAM) [15], developed by the National Renewable Energy Laboratory (NREL) in the United States, using standard parameters. Climate data was available for 2010 to 2017 from the National Solar Radiation Database (NSRDB) [16], also developed by NREL.

The installation was modelled for each year using actual climate data from that year. The result was the hourly energy generation output from the installation. On an hourly basis, the generation data was then multiplied by either the corresponding MEF, HEF or AEF to calculate the total GHG emission reductions using each approach. Results are shown in Figure 6.

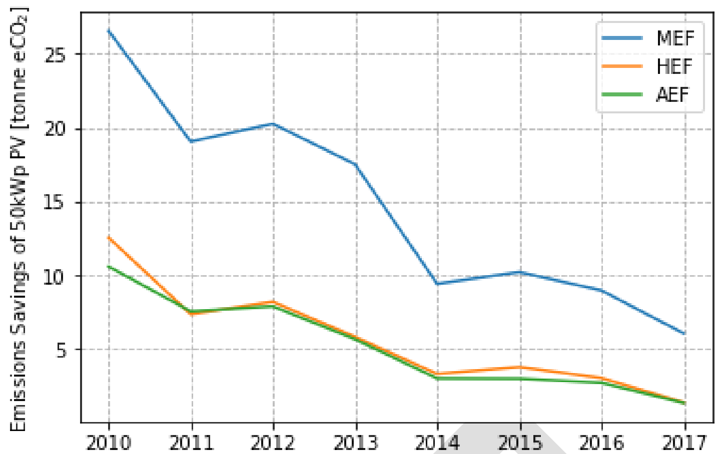


*Figure 6. GHG emissions reductions from a 50kWp PV installation in Toronto using the MEF, HEF or AEF.*

The result from the HEF and AEF closely align because both approaches incorporate the clean baseload generation into the emission factor calculation. The emissions savings is 2 to 4 times greater on an annual basis when using the MEF instead of the HEF or AEF. This is substantial. ECMs that are more targeted to reduce electricity during times of high MEF could yield even better results than PV.

For example, in July or August, an ECM that shifted some of a building's weekday electricity load during the hours of 8 to 12 pm to the hours of 12 pm to 4 am the following day would have substantial GHG reduction benefits. In some cases, this may simply be a no-cost change to the building's control. These sorts of decisions can be made using previous year's data but also in real-time using the freely available data from the IESO. Regardless, this paper presents an easy solution to support data-driven decision-making that can optimize the real-world GHG emissions reduction benefits of electricity conservation. The remaining challenge lies in ensuring that those benefits are captured in annual inventories.

## 3.2 Part B: Incorporating MEF Analysis into Specific Buildings

### *3.2.1 Developing a baseline energy model*

Section 2.2 showed that the MEF can be incorporated into annual inventories by comparing the actual hourly electricity consumption in a given target year to an hourly baseline energy consumption. The actual hourly electricity consumption is *what did happen* and the baseline energy consumption is *what would have happened* had there been no changes to the building, or its usage, that would impact energy consumption.

The baseline energy consumption is normalized to target year conditions. It is a model trained using the environmental data (as well as other independent variables) from a selected base year which is then mapped to the target year conditions. If no changes were made between the base year and the target year then the actual and baseline energy consumption ought to closely align.

The challenge of the method lies in finding a suitable baseline energy model for a given building that can predict energy consumption sufficiently accurately at an hour-level granularity. The model also needs to be simple, such that it can be applied on a widespread basis without requiring a large amount of specialized knowledge.

While building energy modelling software packages likely satisfy the former criterion, they don’t satisfy the latter – in that they require the specialized knowledge of a building energy modelling professional. That knowledge may be available to some organizations, in which case it would be a viable approach, but a simpler route is needed for others.

Free open-source data science software packages make data-driven insights easier than ever before. In the context of building energy, these tools make it easy to develop empirical models that can quantify the relation between inputs (independent variables) and outputs (building energy consumption) to an acceptable degree. Machine learning (ML) models based on artificial neural networks (ANN) are possible, but so are those based on more-straightforward multi-variable non-linear regression. The latter approach was used in this work. It was implemented in the Sci-kit Learn module (details in [17]) available in the Python programming language.

Equation 15 shows the form of a 4th order non-linear multi-variable regression model incorporating 3 independent variables $(x_1, x_2, x_3)$ and one output variable $(y)$. It is a simple polynomial expression. The software is used to optimize the coefficient set ($C_0$ to $C_{12}$) to create the best fit.

$$\begin{aligned} y(x_1, x_2, x_3) = {} & C_0 + C_1 x_1 + C_2 {x_1}^2 + C_3 {x_1}^3 + C_4 {x_1}^4 + C_5 x_2 + C_6 {x_2}^2 \\ & + C_7 {x_2}^3 + C_8 {x_2}^4 + C_9 x_3 + C_{10} {x_3}^2 + C_{11} {x_3}^3 + C_{12} {x_4}^4 \end{aligned} \tag{15}$$

Interval data at 5-min granularity was collected from three office buildings in the Greater Toronto Area (GTA) – referred to as Building A, Building B, and Building C, in this paper. There were no special considerations in the selections of these buildings. They were selected simply because the data was easily available to the research team. For each building, a multi-variable non-linear regression model of building energy consumption was developed for a base year.

Exploratory analysis of the data quickly revealed that the three most important independent variables impacting the building energy consumption in a given hour were: (1) the average outdoor temperature for that hour (dry bulb was used), (2) the hour of the day (from 1 to 24), and (3) the day of the week (each day was assigned a number from 1 to 7). These were the variables used to develop the models of building energy consumption.

Example scatter plots illustrating the impact of the independent variables on energy consumption for Building A in a 2010 base year are provided in Figure 7. They show a single independent variable on the horizontal axis and building energy consumption on the vertical axis. Note that the point clouds are highly scattered simply because energy consumption is impacted by more than one variable.

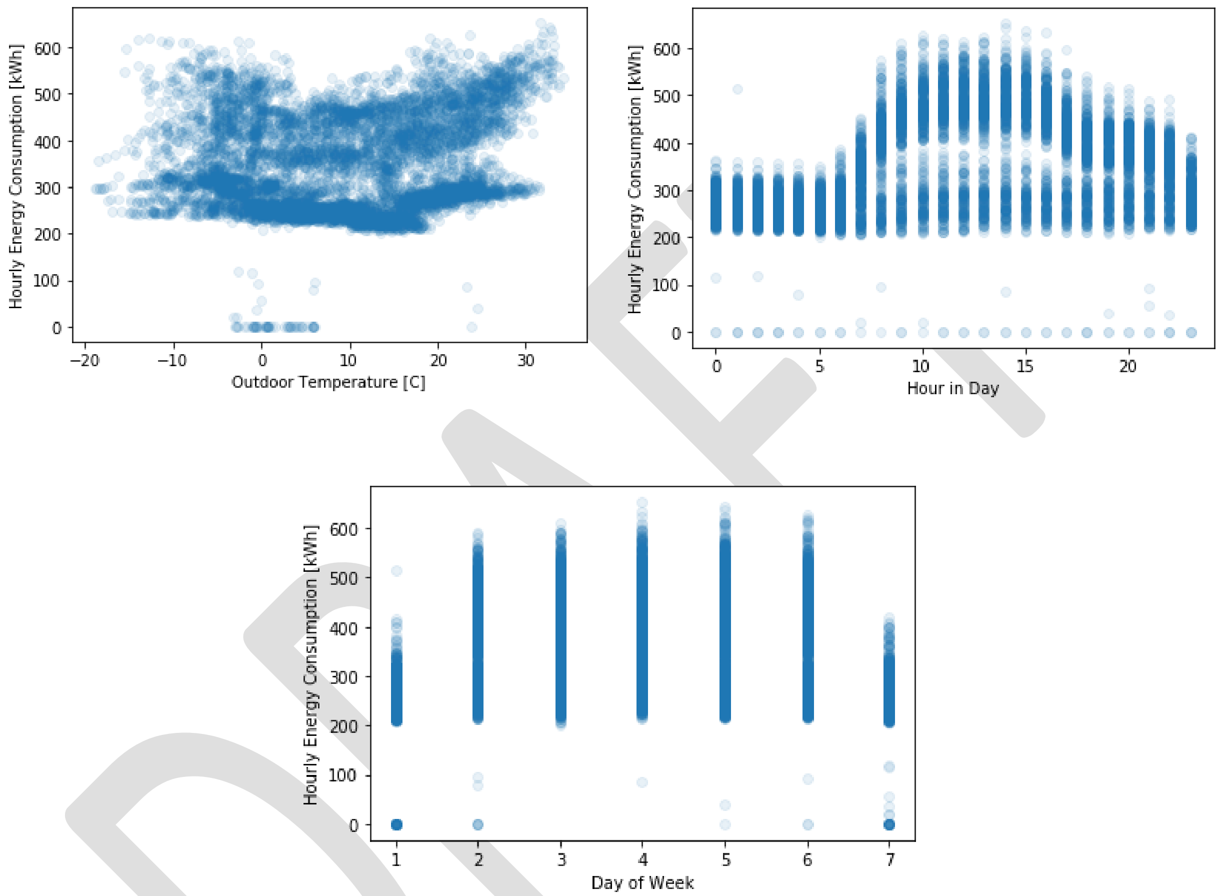


*Figure 7. Impact of outdoor temperature (top left), hour of day (top right) and day of week (bottom) on energy consumption of Building A in the 2010 base year.*

Energy consumption is typically greatest in warmer and colder conditions, it is lowest on weekends (Day 1 and 7), and it peaks in the early and mid-afternoon. It's clear why non-linear regression is required since none of these relationships are linear. In this work, terms up to the fourth order were considered. The method proposed in this work uses these relationships in an empirical model to define the building baseline energy consumption.

For Buildings A to C, the data from a base year was randomly sorted into training (70% of data) and testing (30% of data) sets. The training set was used to develop the model and the testing set provided information on how effective the model is at predicting building energy consumption using data that was not used to train it. The coefficient of determination ($R^2$) of the testing set was used to determine the performance of the models. Note that, in this context, $R^2$ "…represents the proportion of variance (of y) that has been explained by the independent variables in the model. It provides an indication of goodness of fit and therefore a measure of how well unseen samples are likely to be predicted by the model, through the proportion of explained variance." [18]

Figure 8 plots results from the testing sets in each building. The vertical axis is the predicted energy consumption from the model on an hourly basis while the horizontal axis is the actual energy consumption. $R^2$ values for Buildings A, B, and C, are 0.80, 0.80, and, 0.81 respectively. The performance is quite good given the simplicity of the input variables. To further improve the quality of the models for certain buildings, it's possible to add independent variables.

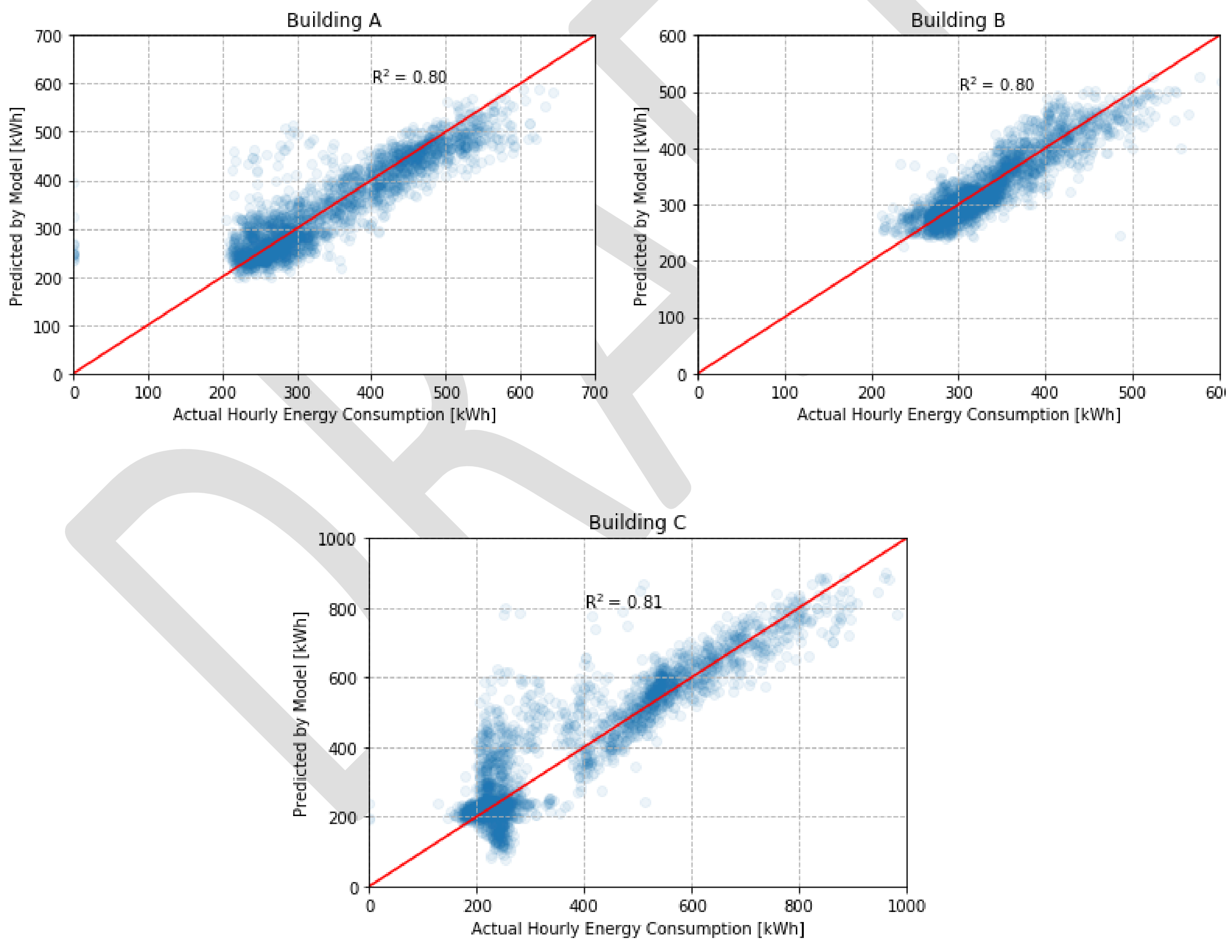


*Figure 8. For Buildings A, B, and C, the hourly energy consumption for a base year was separated into training (70%) and testing (30%) sets. The training set data was used to train a model of hourly energy consumption based on the outdoor dry bulb temperature, the hour of the day, and the day of the week. The model performance was evaluated by applying the model to the testing set data (shown here).*

Future work will incorporate the performance of the baseline energy model into the methodology to estimate the error of the resulting GHG inventory. For now, it stands that *the baseline energy consumption of some buildings can be modelled sufficiently well using simple approaches*.

*3.2.2 Performing an inventory based on the MEF*

The proposed inventory methodology was applied to Building A. Building A was selected because it had a complete set of electricity interval data from 2010 to 2018. Inventories were also conducted using the standard approach with the AEF and also, with a slightly improved approach using the HEF. The results are shown in Figure 9.

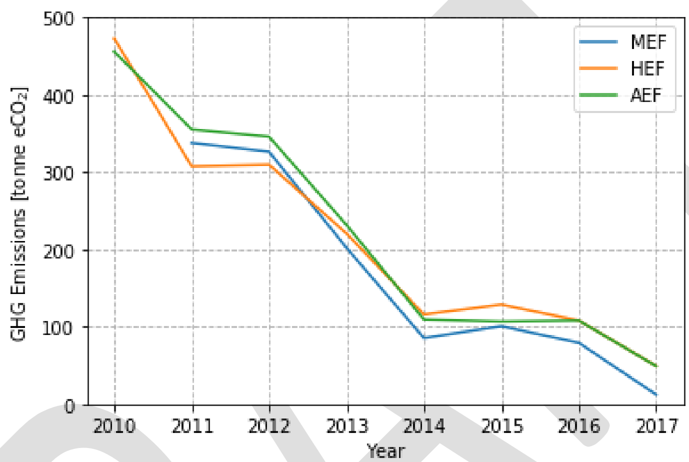


*Figure 9. The GHG inventory for Building A was performed using either the AEF, HEF, or MEF.*

GHG emissions decrease over time due primarily to decreasing grid emission factors as Ontario reduced, and then eliminated, dependency on coal. This trend is clear from Figure 2. The HEF and AEF inventories converge and diverge in relation to how well electricity consumption correlates with a higher HEF in a given year. For most years, the differences are not substantial. It's also the case that, for most years, the MEF inventory produces a result that is lower in magnitude then the other approaches.

To understand why, it is helpful to look at the electricity consumption of Building A year-over-year (Figure 10). It's decreasing. The specific measures and changes to the building motivating the decreasing trend have not been documented in this work. It's less important. The key thing to note is that the MEF inventory identified that the decrease is larger in magnitude then can be accounted for by weather normalization (Figure 11). In the emissions calculation, a larger emission factor is applied to the savings then to the consumption (according to Equation 13) and overall this produces a GHG emissions inventory that is lower in magnitude.

Looking closer at 2017 in Figure 9, the MEF inventory approaches a very small value – near zero. This is a consequence of a few key factors:

- in 2017, there is a net hourly change of -75 kWh on average with respect to the baseline model (Figure 11);
- the energy consumption in 2017 is predominantly between 200 and 400 kWh (Figure 10) which means that net hourly change is significant in proportion to total consumption;
- the AEF (and therefore the HEF) is very low at 20 g $eCO_2$/kWh (Figure 2); and
- the MEF values are high, between 100 and 200 g $eCO_2$/kWh (Figure 6 shows emission for 2018, but results are similar for 2017).

Putting it all together, there is large savings, a large MEF, and a low AEF/HEF. In Equation 13, these factors combine to produce a very low GHG inventory overall.

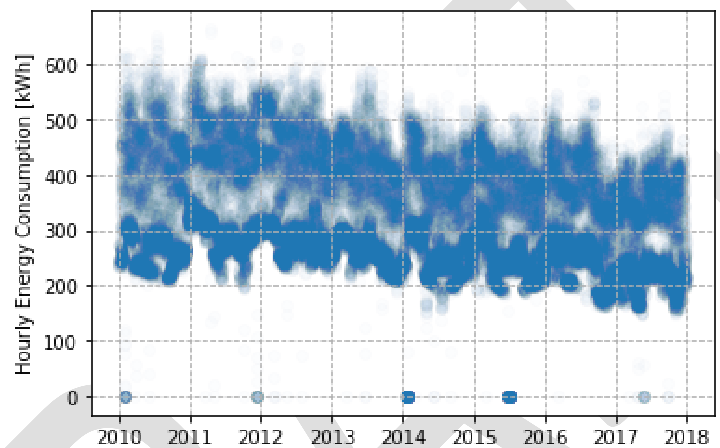


*Figure 10. The hourly energy consumption of Building A decreases year-over-year. Figure 11 shows that the decrease is greater than could be explained by changes in weather.*

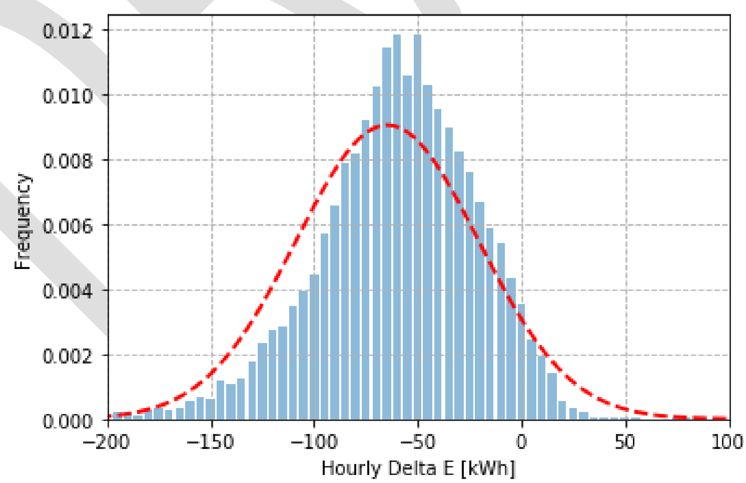


*Figure 11. A frequency histogram of $\Delta E_{target,i}$ for 2017 shows the actual energy consumption is lower than the baseline model of energy consumption (which has been normalized for weather) because the mean of the distribution is a negative value. There is a real reduction to which the MEF can be applied. The red dotted line shows a fit to a normal distribution. The distribution is normal because the baseline model of energy consumption is subject to random errors.*

*3.2.3 Impact of ECMs on the inventory*

As a final step, it was important to look at how an ECM impacts the proposed inventory methodology. The generation data from the PV system simulated in Section 3.1 was subtracted from the consumption of Building A starting in 2012 to generate a new energy consumption profile representative of Building A with a zero export PV system. The three inventory methods based on the AEF, HEF, and MEF were then recalculated.

Figure 12 shows the decrease in emissions caused by the PV system within each methodology. It was determined by completing the inventory both with, and without, the PV and then finding the difference in emissions. The results from Figure 6 are also shown as "Individual ECM." Recall that Figure 6 looked at a PV installation in isolation as a specific ECM and not in the context of an inventory in a building. For each hour, the PV generation was simply multiplied by the corresponding MEF.

The key thing to note in Figure 12 is that MEF inventory developed in this paper calculated the same GHG savings from PV as when the GHG reductions from PV were evaluated in isolation. In other words, *the actual real-world benefits of electricity conservation from PV has now been adequately captured in an inventory*, as this work sought to do, where previously the benefits would be underestimated using the AEF. Since the benefits have been incorporated into an inventory they can be properly quantified and tracked. New opportunities for GHG emissions reductions can now be identified as well.

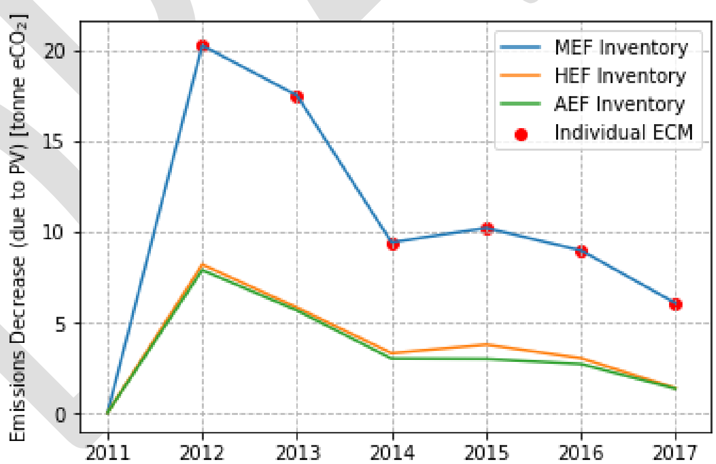


*Figure 12. A 50 kWp PV zero-export installation was simulated for Building A starting in 2012 and the resulting impacts on different inventory methodologies is shown. Also shown are results from PV being considered in isolation as an ECM using the MEF. The proposed MEF methodology developed in this work is able to properly account for the real-world grid impacts of electricity conservation. It is in agreement with the results obtained when an ECM is considered in isolation using a marginal emissions analysis.*

# 4 DISCUSSION

## 4.1 Summary

This work started with the observation that, according to standard practice, GHG emissions inventories significantly underestimate the benefits of electricity conservation. This is a problem because it means that important GHG emissions reduction opportunities may be missed. These opportunities may include low-cost/no-cost measures or measures with a strong business case.

Inventories underestimate the GHG emissions reductions associated with electricity conservation because they rely on annualized emission factors that do not take into account the actual generators that are curtailed when electricity is conserved. It is frequently the case that natural gas driven generators are curtailed in response to decreased demand. It follows that electricity conservation frequently equates to natural gas conservation, and this has a greater GHG emissions impact than is otherwise quantified. Annualized emission factors are used in inventories because they are simple and widely used by others – not because they are correct in this application.

*Marginal* emission factors take into account the actual generators that are curtailed when electricity is conserved and are therefore a more accurate representation of real-world GHG emissions impacts. They are less commonly used because they are more complex to calculate. This work showed how an hourly marginal emission factor can be calculated for the Ontario grid using a simple procedure and freely available data from the IESO.

A wide variation in the marginal emission factor was observed over different time scales (hourly, daily, and monthly). This impacts different ECMs in different ways, depending on *when* they conserve electricity and not just how much they conserve. Analysis of a PV installation showed emissions reductions on the scale of 2 to 4 times greater than is normally calculated using annualized emission factors.

A larger challenge beyond the calculation of marginal emission factors is the integration of them into an annual inventory. Even if an individual ECM can be analyzed with marginal emission factors, if the annual inventory for a building is based on the *annualized* emission factor then the actual benefits will not be properly tracked. This work showed that marginal emissions can be incorporated into annual inventories in a given target year by comparing actual energy consumption to baseline energy consumption.

To reiterate, the actual electricity consumption is *what did happen* and the baseline energy consumption is *what would have happened* had there been no changes to the building, or its usage, that would impact energy consumption. A marginal emission factor can be applied to the difference in energy consumption between *what did happen* and *what would have happened.*

The challenge lies in defining a baseline energy consumption model that can adequately map baseline energy consumption to the target year conditions on an hour-level granularity – normalizing for important independent variables like weather. Building energy modelling software packages could be used but this requires highly specialized expertise that would limit the broadscale applicability of the proposed method.

A promising alternative to building energy modelling software is a simpler empirical model of building energy consumption. This work showed that a multi-variable non-linear regression model based on the outdoor temperature, hour of day, and day of week, can model the energy consumption of some buildings sufficiently well. Such models are straightforward to develop and train using free open-source data science software packages like those available in the Python programming language. Additional independent variables may be needed to improve the accuracy in other buildings.

The proposed inventory approach was applied to data from a real office building in the GTA. A zero-export PV installation was simulated as an ECM for the building and *it was shown that the proposed inventory approach properly accounted for the GHG emission reductions, taking into account the actual marginal impacts of electricity conservation*.



**4.2 Issues**

It's worthwhile to address other important consequences of proposed methodology. Firstly, in Equation 13, it's entirely possible for the second term ($\Delta E_{target,i} \cdot MEF_i$) to be negative and greater in magnitude than the first term ($E_{baseline,i} \cdot HEF_i$). This means that the GHG inventory for a given building in a given target year ($GHG_{target,tot}$) can be negative – despite the fact that the building is still consuming electricity.

This seems strange but it is a real physical phenomenon and not an artifact of the analysis. It's helpful to look at an extreme example to understand why. Say there was a building that was decommissioned, and its electricity consumption was reduced to zero except for some exterior lighting that was left in place for security. In this case, it's clear that a marginal emission factor should be applied to the reduction in electricity.

The real-world grid response is that the generators on the margin would have reduced their output to respond to the decrease in demand. The fact that there is some remaining net consumption does not alter this fact. This example demonstrates that it's possible for a building's GHG emissions impact on the grid to be dominated by the electricity *savings* it produces rather then by it's net consumption and this is why a building's GHG emissions inventory can be negative. This fact holds true for less extreme examples of electricity savings as well.

The prospect of net-negative carbon buildings is beneficial. However, just as the analysis properly accounts for the real-world impacts of conservation, it also accounts for the real-world impacts of new loads – and this means that new electricity consumption after the base year is weighted more heavily in the GHG emissions inventory.

For example, the emissions inventory of any new building constructed after the base year ought to be based entirely on marginal emission factors simply because the energy consumption in the base year for that building is zero (since that building did not exist). In the context of the grid, when the new building is added, it is the generators that are on the margin that will respond to meet the new demand – so it is reasonable that marginal emission factors ought to be applied.

It's also important to acknowledge that, over long time periods, the cumulative changes in demand across many buildings may alter the average baseload. This is somewhat at odds with the assumptions within the analysis that whenever energy consumption is different than it otherwise would have been (had there been no changes to a building or its usage) that the grid adapts by changing the output of the generators on the margin. Over long periods, if individual ECMs in individual buildings had some small part in changing the average baseload then that assumption is not completely correct.

A solution to this specific issue is not offered in this work. It would introduce a level of complexity and analysis that would make this approach potentially intractable. Rather, this paper acknowledges that, while the assumptions within the proposed approach are not completely correct, they are much more correct than the assumptions currently used in GHG emission inventories and the proposed approach is therefore a better alternative.

A final issue to be discussed is the independent variables. The example in this work used the outdoor temperature, hour of the day, and the day of the week, to develop an hourly model of baseline energy consumption. These are variables that are uncontrollable. It's possible that the models could be improved if other variables were incorporated as well but it is important to discuss this further.

A specific example is helpful. Say that, in the base year, a building *was not* occupied on weekends but then, in a later year (i.e. a target year), it *was* occupied on weekends. If the hourly model of baseline energy consumption included occupancy as an independent variable then the baseline model could normalize for the increased occupancy in the target year. This might seem helpful initially but this action would miss an important point. The increase in occupancy represents a real increase in electricity that could otherwise be avoided. It is new consumption that needs to be met by generators that are on the margin. In contrast, the other variables (outdoor temperature, hour of day, and day of week) could not be avoided.

It's helpful to keep in mind that the actual hourly consumption needs to be compared against something that would have feasibly occurred in a business-as-usual scenario. This is the purpose of the hourly model of baseline energy consumption. A change in the usage, building control, or building condition, in a given target year *is* a deviation from business-as-usual. On the other hand, a change in outdoor temperature *is not* a deviation from a business-as-usual scenario so it needs to be normalized for the target year conditions.

The problem remains that the energy consumption of a building may to be strongly dependent on factors like occupancy and, if not taken into account, the hourly model of baseline energy consumption will have poor performance. In these cases, it may be necessary to include additional independent variables that better incorporate the fluctuations in occupancy in the base year. However, when the model is applied to the target year, the base year occupancy would still be used and the model therefore would *not* normalize for differences in occupancy between base year and target year.

The key point is that the independent variables used in the hourly model of baseline energy should not normalize for differences between the base and target years that are within the control of the building owner(s).

### 4.3 Future Work

There are several important next steps that will be the subject of future work:

- The performance of the baseline energy model can be used to define the uncertainty of the overall GHG emissions inventory. This is important because accuracy will vary across different buildings. However, the need to account for uncertainty should not be seen as a major barrier. The current inventory approach based on annualized emission factors is based on incorrect assumptions. This is widely acknowledged, but there are generally no efforts to define the uncertainty of that approach. The proposed method represents a simple and achievable move to an improved methodology than is currently used despite that fact the further refinement is still needed.

- Additional ECMs should be evaluated within the proposed approach. Simple low-cost/no-cost measures like changes to building control could have notable impacts from a marginal perspective but are not currently deployed.

- A simple approach to baseline energy building energy modelling was applied to three random office buildings. Results were positive but additional work should explore the consumption data of additional buildings to further gauge the widescale applicability of the methodology.

## 5 CONCLUSION

Quantifying greenhouse gas (GHG) emissions is a key part of reducing them. However, current GHG inventory methodologies underestimate the GHG savings associated with electricity conservation because they rely on annualized emission factors that neglect important dynamic aspects of the electricity grid. Decision makers may therefore not take advantage of opportunities for cost-effective GHG reduction through electricity conservation – some of which may include simple no-cost/low-cost changes to building controls – because there is no way to track the reductions using the current approaches. This paper addresses this issues. It provides a simple and transparent methodology for estimating the hourly marginal emission factor (MEF) for Ontario and a methodology for incorporating the MEF into annual inventories for specific buildings. Overall, the paper provides an improved methodology for GHG inventories in buildings which can be used to identify, quantify, and promote cost-effective GHG reduction opportunities through electricity conservation. It is recommended that organizations and building owners consider the proposed approach in their buildings to identify new GHG reduction opportunities and properly quantify the impacts of current electricity conservation efforts.

## REFERENCES


[1] The Atmospheric Fund (TAF). A Clearer View on Ontario's Emissions: Electricity emissions factors and guidelines. 2019. Accessed online Jan 2020: taf.ca/wp-content/uploads/2019/06/A-Clearer-View-on-Ontarios-Emissions-June-2019.pdf

[2] Government of Canada. National Inventory Report 1990–2017: Greenhouse Gas Sources and Sinks in Canada. 2019. Accessed online Jan 2020: unfccc.int/documents/194925

[3] IESO. Year End Data: 2017. Accessed online Jan 2020: www.ieso.ca/en/Corporate-IESO/Media/Year-End-Data/2017

[4] World Resource Institute. Greenhouse Gas Protocol: Guidelines for Quantifying GHG Reductions from Grid-Connected Electricity Projects. 2007. Accessed online Jan 2020: www.wri.org/publication/guidelines-quantifying-ghg-reductions-grid-connected-electricity-projects

[5] IESO. Data Directory. Accessed online Jan 2019: www.ieso.ca/en/Power-Data/Data-Directory

[6] IESO. Generators Output and Capability Report. Accessed online Jan 2020: reports.ieso.ca/public/GenOutputCapability/PUB_GenOutputCapability.xml

[7] IESO. Active Contracted Generator List. Accessed online Jan 2020: www.ieso.ca/-/media/Files/IESO/Document-Library/power-data/supply/IESO-Active-Contracted-Generation-List.xlsx

[8] Ethan S. Warner and Gavin A. Heath. Life Cycle Greenhouse Gas Emissions of Nuclear Electricity Generation, Systematic Review and Harmonization. Journal of Industrial

Ecology. Volume 16 Issue s1, p: S73-S92. April 2012. Access online Jan 2020: onlinelibrary.wiley.com/doi/full/10.1111/j.1530-9290.2012.00472.x

[9] International Panel on Climate Change (IPCC). Special Report on Renewable Energy Sources and Climate Change Mitigation. Annex II Methodology. Table A.II.4, p. 982. Accessed online Jan 2020: www.ipcc.ch/site/assets/uploads/2018/03/SRREN_Full_Report-1.pdf

[10] Stacey L. Dolan and Gavin Heath. Life Cycle Greenhouse Gas Emissions of Utility-Scale Wind Power Systematic Review and Harmonization. Journal of Industrial Ecology. Volume 16 Issue s1, p: S136-S154. April 2012. Access online Jan 2020: onlinelibrary.wiley.com/doi/pdf/10.1111/j.1530-9290.2012.00464.x

[11] David D. Hsu, Patrick O'Donoughue., Vasilis Fthenakis, Garvin A. Heath, Hyung Chul Kim, Pamala Sawyer, Jun-Ki Choi and Damon E. Turney. Life Cycle Greenhouse Gas Emissions of Crystalline Silicon Photovoltaic Electricity Generation Systematic Review and Harmonization. Volume 16 Issue s1, p: S122-S135. April 2012. Access online Jan 2020: onlinelibrary.wiley.com/doi/full/10.1111/j.1530-9290.2011.00439.x

[12] Michael Whitaker, Garvin A. Heath, Patrick O'Donoughue and Martin Vorum. Life Cycle Greenhouse Gas Emissions of Coal-Fired Electricity Generation Systematic Review and Harmonization. Volume 16 Issue s1, p: S53-S72. April 2012. Access online Jan 2020: onlinelibrary.wiley.com/doi/full/10.1111/j.1530-9290.2012.00465.x

[13] IESO. Supply Overview: Imports and Exports. Accessed online Jan 2020: www.ieso.ca/en/Power-Data/Supply-Overview/Imports-and-Exports

[14] Government of Canada. National Inventory Report 1990–2017: Greenhouse Gas Sources and Sinks in Canada. 2019. Part 3, Table A13-6, p. 64. Accessed online Jan 2020: unfccc.int/documents/194925

[15] National Renewable Energy Laboratory (NREL). System Advisor Model (SAM) Software. Available online at: sam.nrel.gov

[16] National Renewable Energy Laboratory (NREL). National Solar Radiation Database (NSRDB). Available online at: maps.nrel.gov/nsrdb-viewer/

[17] Sci-kit Learn Module Information. Accessed online Jan 2020: scikit-learn.org

[18] Sci-kit Learn Module Information. Regression Metrics. Accessed online Jan 2020: scikit-learn.org/stable/modules/model_evaluation.html#regression-metrics