Bridging Data Gaps in Energy Sector

A Case Study on Utility Bill Verification in Ontario using Digital Data via Green Button and EBT infrastructures

Research Team

Anjali Mudaliar (Researcher) and Francis Palma, PhD (Professor)
SE+AI Research Lab, Faculty of Computer Science, University of New Brunswick, NB, Canada
Zaiyan Khan (Subject Matter Expert, Research Lead), Martin Benum (Subject Matter Expert, Consultant), and Gary Michor (Business Lead, CEO)
Screaming Power Inc, ON, Canada
John Gaby (Subject Matter Expert), Active Business Services

Dated: March 7, 2025

Abstract

Access to standardized, digital energy billing and meter information is crucial for advancing research and innovation in the energy sector. Without granular data, smart grid technology and performance-based algorithms face significant challenges, hindering global efforts to optimize energy costs and develop effective management strategies. Ontario, Canada, exemplifies this complexity with over 50 diverse electric and natural gas utilities, each with unique infrastructures and rate charges.

This case study, conducted over four months, delves into validating bill and meter information from three utilities, offering insights into the intricacies of utility data and its potential for reuse. Building on our previous research published by Energy Central, in the Nov/Dec 2024 Special Issue[1], this study explores bill validation, bill-to-meter verification, and meter-to-meter verification using data from the UEnergyHub portal, a pilot energy research hub. This portal aids in authenticating and authorizing utility account holders to share their energy data with researchers, providing data visibility, transactional auditing, and data translation.

The findings from 16 meters or 16 accounts reveal slight discrepancies in Green Button (GB) data and bill totals, with negligible differences in GB versus Electronic Business Transaction (EBT) data. Key challenges to ensuring 100% verifiable and consistent energy data include addressing data interconnectivities, rate complexities, and inconsistencies due to ad hoc utility system updates. This research underscores the need for regulatory adaptation to technological challenges, ensuring utility data can effectively support the marketplace while maintaining compliance and consumer confidence.

By incorporating insights from subject matter experts and employing triplicate validation methods, this study aims to pave the way for more accurate and reliable energy data management, inspiring further research and innovation in the field.

Keywords: Green Button, EBT, Anomaly Detection, One-Class SVM, Utility Data, Energy Consumption, Line Loss Factor, Billing Discrepancies, Multivariate time series, Regulatory Compliance.

1 Introduction

Energy data is crucial for utility operations, regulatory compliance, consumer transparency, and integrating new solutions to address regional and global issues. However, no single

[1] https://www.flipsnack.com/E58F5ABBDC9/special-issue-digital-transformation-in-the-utility-industry/full-view.html

entity or tool currently manages this data comprehensively, making utilities the primary source of truth. In Ontario, the Green Button (GB) standard and Electronic Business Transaction (EBT) standards have improved data accessibility and management, but there has been no formal review of data accuracy and consistency by utilities or regulators.

The GB standard, regulated for implementation in Ontario by November 2023, allows consumers and authorized third parties to access detailed digital data about energy consumption and related costs, promoting energy efficiency and informed decision-making. Ontario's legacy EBT system has been used since the early 2000s and facilitates standardized data exchange between regulated utilities and energy retailers. EBT is essential for business-to-business operations. Despite these standards, the lack of synergy between data sets and utility systems highlights the need for better meter data handling and billing oversight.

Our research has identified issues such as differences in data recording practices, time misalignments, inconsistent data formats, and incomplete data fields, which undermine billing accuracy and data integrity. These inconsistencies can have significant financial implications, affecting customer and market trust. By incorporating insights from university researchers, subject matter experts, and advanced data processing techniques, this research offers practical recommendations for improving data accuracy and transparency to enhance regulatory compliance, consumer confidence, and innovation in the energy sector.

2 Background - EBT and Green Button

The GB initiative, introduced by the Ministry of Energy, aims to enhance consumer access to utility energy data, promoting transparency and energy efficiency. Properly implemented GB data can provide detailed insights into consumption patterns, helping consumers, innovators, and utilities make informed decisions. In contrast, the legacy EBT system has facilitated the exchange of transactional data between regulated utilities and energy retailers in Ontario's energy market for over two decades, ensuring compliance with Ontario Energy Board (OEB) standards, policies, and market rules.

The OEB oversees the province's energy sector to protect public interests, governing standards, data management, billing practices, and consumer protection. In Ontario, the ratepayer owns the data, and the regulated utility is the custodian. The OEB ensures that energy utilities operate transparently, efficiently, and in the best interest of consumers while enjoying a regulated rate of return for their capital investments. The OEB sets guidelines for market-wide implementations, data accuracy, billing transparency, and enforces public needs in Ontario’s energy marketplace.

Previous research by the University of New Brunswick (UNB) revealed several challenges in Ontario's energy data landscape, including utility discrepancies and the effectiveness of their interpretation of regulations and standards. Issues identified included the lack of readiness and testing by utilities on their GB systems, contrary to their reports to the OEB. Data issues often stem from flaws in system readiness, production activity, and data throughput/delivery practices. This new study builds on these initial findings by incorporating the analysis of detailed billing and meter data, introducing new dimensions to the research. By addressing these issues through validated analytical techniques, this research aims to improve data accuracy and validation, support regulatory compliance, and enhance consumer trust in digital data.

2.1 Green Button Standard

The GB Standard, part of a US White House initiative, provides consumers and authorized third parties with secure, standardized access to energy consumption data. It aims to empower consumers to make informed energy decisions by facilitating data sharing from utilities to third-party service providers. The Ontario Ministry of Energy later adopted this standard, chairing the Green Button Alliance (GBA) alongside Ontario utilities. Ontario uses the 3.3 standard, which was new to the marketplace at its implementation. The GBA, a not-for-profit group, helps evolve the standard and certifies its use through independent utility testing.

GB aims to enhance operational transparency and support regulatory compliance by ensuring that energy data produced through smart technologies is accessible, accurate, and consistent. It fosters innovation and cost savings in the energy sector by enabling the development of new services and applications focused on energy efficiency, demand response, and renewable energy integration. This study uses the GB Standard as a framework for analyzing energy consumption patterns, detecting anomalies, and validating billing accuracy across multiple Ontario utilities, providing new opportunities.

In Ontario, GB is now a regulatory requirement for over 50 electric and natural gas utilities. However, the standard and GBA certification process do not ensure user experience or data/operations synergy across the market. While best practices were developed through thousands of volunteer hours from market specialists in Ontario, there has been no government mandate to follow them. Version 3.3 of the standard includes detailed interval usage data, billing information, and historical consumption records, but it does not ensure sufficient interoperability between market users and implementers.

2.2 Overview of Green Button vs EBT Data

Both GB and EBT systems are central to energy data management and settlement in Ontario, serving distinct purposes. GB data focuses on consumer energy usage, providing detailed interval readings that enable monitoring of energy consumption patterns over time. The XML data structure in GB supports granular insights, useful for anomaly detection. GB emphasizes consumer empowerment by making detailed energy usage information readily accessible, facilitating better energy management, cost savings, and informed decision-making.

In contrast, EBT data designed around 25 years ago, supports the business-to-business (B2B) operations of utilities and regulated energy retailers. It facilitates transactional activities between OEB-regulated parties, focusing on meter readings, settlement, and data collection for utility customer compliance and costing needs. EBT XML data structures and transactions are integrated into standardized business processes, reflecting the nature of energy transactions for retailing energy. There are no known studies on why Ontario elected to use the US-based GB standard over the EBT standard of Ontario to meet new market needs.

Previous research by the University of New Brunswick (UNB) showed a low correlation between GB bills and EBT meter data sets. This discrepancy was not observed through a second small data set review from a different utility. The new evaluation showed a better correlation between the two data sets.

3 Research Methodology

The current research focuses on analyzing energy data discrepancies using extended datasets from the same sources of truth, building on the findings from our previous work with GB and EBT data.

The ongoing primary objective is to identify anomalies in billing and meter data sourced from the UEnergyHub portal by integrating this data to understand discrepancies at a granular level. The methodology involves several key steps, including data collection, preprocessing, integration, analysis, and validation.

As a precursor to the research, the predominant GB accounts subjected to review were residential rate class accounts that had RPPTOU electricity rates applied.

3.1 Data Collection

Data was sourced from the UEnergyHub portal in JSON format, comprising two primary datasets:

- **Billing Files:** Contain customer billing information, including account numbers, billing periods, total charges, and energy usage details.
- **Meter Files:** Provide hourly interval readings for each account, detailing energy consumption patterns over time.

These datasets were obtained for multiple accounts across different Ontario utilities to ensure comprehensive coverage and variability in the analysis. Due to a time limit put on this research, the data and accounts reported in this paper were a subset of the analysis.

3.2 Data Preprocessing

Preprocessing was critical to standardize and prepare the data for analysis. The following steps were undertaken:

- **Data Extraction**: For the manual process: excel spreadsheets were used to import the JSON data for analysis. For researcher and automated processes: Python scripts were developed to extract relevant fields from the JSON files. Extraction included account information, billing details, and meter readings.
- **Time Alignment:** In all data review processes, it was discovered there were timestamp differences in the meter data for different utilities. To correct this, some utility timestamps were shifted backward by one hour to accurately reflect the start of each interval. However, the timestamps for other utilities were already correctly aligned and did not require any adjustment.
- **Data Cleaning**: Due to data irregularities found, duplicate records were removed, missing values were handled through imputation techniques, and inconsistencies in data formats were standardized. When needed, utilities were notified of issues for issue correction in their downstream systems.
- **Feature Engineering**: In all analysis techniques, additional columns were created to enhance the dataset, such as breaking down interval Start into year, month, day, hour, and minute components. These were aggregated into a consolidated timestamp for better analysis.

3.3 Data Integration

For all analyses, the billing and meter data were merged to create a unified dataset that links energy consumption with billing information. This integration involved:

- **Mapping Account Numbers**: Ensuring that billing data matched the corresponding meter data based on account IDs and billing periods.
- **Calculating Interval End**: Since the meter data provided interval start times and durations, an intervalEnd column was created to define the precise end of each reading period.
- **Rate Matching**: For bill validation, the consolidated timestamp was used to match each interval reading with the corresponding TOU rate from the OEB rate table. This allowed the categorization of consumption into off-peak, mid-peak, and on-peak periods.

3.4 Anomaly Detection and Analysis

The integrated dataset was analyzed to detect discrepancies and anomalies through the following methods:

- **Billing vs. Meter Data Comparison:** The total energy consumption calculated from meter readings was compared with the billed usage for each account and billing period. This helped identify overcharges, undercharges, and data mismatches.
- **TOU Rate Analysis:** Energy consumption was categorized based on TOU rates to assess whether the billed amounts aligned with actual usage patterns.
- **Discrepancy Identification:** Cases where the calculated billing amount significantly deviated from the actual billed amount were flagged for further investigation.

4 Research Findings

4.1 Detailed Report: Regulated Price Plan (RPP) Time-of-Use (TOU) Bill Calculation and Green Button (GB) Data Comparison

RPP is the predominant electricity pricing structure in Ontario for approximately 5.4 million residential and small commercial customers. It is designed to provide price stability while ensuring that customers pay a fair share of the cost of electricity. The research completed was complex due to many undocumented bill formulas. The university researcher found it difficult to conclude where the subject matter specialist could complete comparisons of two bills from different utilities for residential customers.

Types of RPP Rates

Customers under the RPP can choose from three pricing structures:

TOU Pricing: Electricity rates vary depending on the time of day, encouraging consumers to shift usage to off-peak hours when demand is lower.

There are three periods: Off-Peak (Lowest Price) – Evenings, overnight, and all day on weekends/holidays. Mid-Peak (Moderate Price) – Daytime hours in spring and fall. On-Peak (Highest Price) – Peak daytime hours in summer and winter. **Tiered Pricing:** Customers pay a set price per kWh up to a specific consumption threshold. If they exceed the threshold, they pay a higher rate for additional usage. Two tiers exist: Tier 1: Lower price for basic usage and Tier 2: Higher price for consumption above the threshold. **Ultra-Low Overnight (ULO) Pricing** *(Introduced in May 2023):* Designed for customers who can shift consumption to overnight hours. Four periods: Ultra-Low Overnight (Lowest Price): 11 p.m. to 7 a.m. Weekend Off-Peak: Daytime weekends and holidays. Mid-Peak: Moderate daytime pricing. On-Peak (Highest Price): Evening weekday hours.

The OEB reviews and updates RPP rates regularly, ensuring they reflect the cost of electricity supply.

Hydro One's Residential RPP TOU bills are detailed financial records that integrate various charge components, reflecting electricity consumption, delivery fees, regulatory charges, and applicable rebates. Beyond the segmentation of electricity consumption into distinct TOU periods, these invoices integrate a complex structure of delivery charges, encompassing fixed service fees, variable distribution tariffs, Retail Transmission Service Rates (RTSR), and governmental subsidy mechanisms such as rural (RRRP) or distribution rate protection (DRP) rebates. Additionally, regulatory fees and statutory taxation measures are embedded, while periodic adjustments—including loss factor corrections and specific rate rider modifications—may further refine the calculation.

In our research, we delineated an exhaustive methodology for calculating an electricity bill for a Hydro One residential R2 (low-density year-round occupation) for one month and systematically juxtaposed the resultant figures against the GB bill dataset for each bill spanning one month.

Overview of Bill Components

Our investigation prepared a sample Hydro One RPP TOU bill composed of the following critical elements: **Electricity Charges:** Quantification of On-Peak, Mid-Peak, and Off-Peak consumption according to the Ontario Energy Board (OEB)-mandated tariff schedule. **Delivery Charges as Prescribed in the OEB-Approved Tariff of Rates and Charges:** Monthly fixed service charge. Variable distribution charge assessed per kWh (unique to this customer group). Fixed and variable deferral/variance account charge allocations, Smart Meter Entity charge. Retail Transmission Service Rates (RTSR) for network and line connection components. Application of Rural or Remote Electricity Rate Protection Plan (RRRP) or Distribution Rate Protection (DRP) adjustments. Computation of TOU consumption adjusted for line loss factors **Regulatory Charges:** Wholesale Market Service Rate (WMS), exclusive of Capacity-Based Recovery (CBR) charges. Capacity-Based Recovery (CBR), applicable exclusively to Class B customers. Rural or Remote Electricity Rate Protection (RRRP) charge. Standard Supply Service Administrative Charge (where relevant) **Rebates and Adjustments:** Ontario Electricity Rebate (OER) alongside any additional jurisdictional rebates. **Taxes:** Harmonized Sales Tax (HST), typically 13%, post-rebate calculation.

General Calculation Approach

To calculate an RPP TOU electricity bill, first, obtain the customer's electricity consumption data and categorize usage into Off-Peak, Mid-Peak, and On-Peak periods. Multiply each by the corresponding TOU rates to determine the total electricity cost. Add delivery and regulatory charges, then apply the Ontario Electricity Rebate (OER) if applicable. Subtract the rebate from the subtotal, then calculate Harmonized Sales Tax (HST) (13%) on the adjusted amount. The final bill is the sum of the post-rebate subtotal and HST. Detailed Calculation Approach is below:

Validation and Preparation of Meter Data: *Extract and structure the raw interval meter data from the GB JSON dataset ("meterData"). Assess temporal consistency, ensuring adherence to established conventions (e.g., "00:00:00" signifies the prior day's final hourly interval). Evaluate the application of loss factors and implement corrective adjustments where necessary.*

Mapping Meter Data to RPP TOU Rate Table: *Assign individual consumption intervals to the appropriate TOU category based on the recorded timestamp. Aggregate hourly consumption figures by TOU classification across the billing period.*

Computation of Electricity Charges: *Multiply the total consumption per TOU period by the corresponding RPP TOU rate. Derive the cumulative electricity cost by summing charges across all TOU classifications.*

Computation of Delivery Charges (including DRP & RTSR Adjustments): *Fixed Charge: Apply the standardized monthly service fee. Variable Distribution Charge: Compute consumption-based charges by multiplying the total kWh by the designated distribution rate. Smart Meter Entity charge. RTSR Charges: Apply per-kWh rates for both network and line connection services. Rural Rate Protection Credit: Deduct the applicable credit amount from the total charges. Line Loss Adjustment: Implement the prescribed loss factor where required. DRP Rebate: Apply the jurisdictional DRP credit. Sum all components to determine the net delivery charge.*

Computation of Regulatory Charges: *Apply regulatory cost factors on a per-kWh basis. Apply SSS charge. Aggregate charges to determine the total regulatory cost.*

Consolidation of Subtotals and Rebate Applications: *Compute the subtotal as follows: Electricity Charges + Delivery Charges + Regulatory Charges. Apply the OER as a proportional reduction.*

Computation of HST and Final Invoice Amount: Calculate HST at 13% on the post-rebate subtotal. Compute the final invoice.

4.2 Comparative Analysis with GB Bill Data

Below is the final comparison between our computed shadow bill and the GB dataset (actual customer bill), confirming the accuracy and reliability of the outlined methodology. Our analysis shows that the GB dataset does not include tiered kWh volumes, only a single final value. We also observe differences in the calculated tiered charges and monetary discrepancies in other bill components. These differences will require further investigation with the utility.

Study Shadow bill calculation	mnemonic	Metric	Total Volume	Total CHARGE
ON-PEAK	ON-PEAK-COMP	RPPTOU-ON-B	487.54	$ 73.62
MID-PEAK	MID-PEAK-COMP	RPPTOU-MID-B	293.57	$ 29.94
OFF-PEAK	OFF-PEAK-COMP	RPPTOU-OFF-B	1,053.90	$ 77.99
ELECTRICITY	**ELECTRICITY-COMP**	**ST-ELEC-RPPTOU-B**	**1,835.01**	**$ 181.55**
DELIVERY	DELIVERY-COMP	ST-DEL-TOT-B		$ 89.14
REGULATORY	REGULATORY-COMP	ST-REG-TOT-B		$ 10.80
OER	OER-COMP	OER-B		-$ 32.93
HST	HST-COMP	HST-B		$ 32.31
TOTAL BILL	**TOTALBILL_COMP**	**TOT-BILL**		**$ 280.87**

Green Button Data (Customer Bill)	mnemonic	Metric	Total Volume	Total CHARGE
ON-PEAK	GB-ON-PEAK	GB-RPPTOU-ON		$ 71.10
MID-PEAK	GB-MID-PEAK	GB-RPPTOU-MID		$ 30.02
OFF-PEAK	GB-OFF-PEAK	GB-RPPTOU-OFF		$ 79.17
ELECTRICITY	**GB-ELECTRICITY**	**GB-ST-ELEC-RPPTOU**	**1,835.01**	**$ 180.29**
DELIVERY	GB-DELIVERY	GB-ST-DEL-TOT		$ 91.39
REGULATORY	GB-REGULATORY	GB-ST-REG-TOT		$ 10.79
OER	GB-OER	GB-OER		-$ 33.05
HST	GB-HST	GB-HST		$ 36.72
TOTAL BILL	**GB-TOTALBILL**	**GB-TOT-BILL**		**$ 286.14**

DIFFERENCE	mnemonic	Metric	Total Volume	Total CHARGE
ON-PEAK				$ 2.52
MID-PEAK				-$ 0.08
OFF-PEAK				-$ 1.18
ELECTRICITY			**0.01**	**$ 1.26**
DELIVERY				-$ 2.25
REGULATORY				$ 0.01
OER				$ 0.12
HST				-$ 4.41
TOTAL BILL				**-$ 5.27**

Concluding Remarks

Our investigation presents a rigorously structured framework for reconstructing Hydro One RPP TOU invoices, validating computational accuracy relative to GB data, and fostering transparency in cost determination methodologies. By adhering to this analytical process, stakeholders—including regulatory bodies, industry professionals, and consumers—can independently authenticate billing accuracy and attain granular insights into cost componentization.

However, discrepancies identified in the comparison highlight potential systemic inconsistencies in bill calculations that warrant further investigation. These errors could indicate differences in rate application, rounding methods, or undisclosed adjustments by the utility. If left unresolved, such discrepancies may affect customer confidence, raise regulatory concerns, and necessitate refinements in billing methodologies to ensure fairness and compliance with OEB regulations.

4.3 Bill to Meter verification GB (Manual)

This section presents the results of bill-to-meter verification based on the GB data obtained through the UEnergyHub portal. The objective of this analysis was to assess whether the energy consumption reported in the billing data aligns with the interval meter readings recorded during the same billing periods. This verification was conducted at a granular level by comparing total billed usage with the sum of hourly meter readings for six utility accounts from two utilities.

Methodology: The analysis followed a structured data processing workflow, integrating billing and meter datasets to ensure accurate comparisons. The key steps involved*:*

Data Extraction and Preprocessing: Billing and meter data were extracted from JSON files using Python scripts. The intervalStart timestamps in meter data were evaluated and adjusted where necessary to ensure alignment with billing periods. Any missing or duplicate records were handled through appropriate data-cleaning techniques.

Time Alignment and Aggregation: Meter readings were summed over each billing period to calculate the total energy consumption per account. The calculated values were then compared against the total usage reported in the corresponding GB billing data.

Anomaly Detection and Discrepancy Analysis: Cases where the total billed energy significantly differed from the sum of meter readings were flagged for further investigation. Differences were categorized based on magnitude and possible causes, including time misalignment, missing data fields, and rounding variations.

Findings and Observations

The results indicated systemic discrepancies between the billed energy usage and actual meter readings across multiple accounts and billing periods. While some variations were expected due to minor rounding errors, several cases exhibited significant differences, raising concerns about the accuracy of the billing data. Key observations included:

Misalignment Between Billing and Meter Data: In some instances, the first meter reading recorded in the dataset did not align with the opening meter value expected for the corresponding billing period. This discrepancy introduced uncertainty in verifying the total consumption reported on the bill.

Inconsistent Handling of TOU Rates: The assignment of TOU pricing tiers was not always consistent with the actual hourly consumption data, leading to variations in the final billed amounts.

This study underscores the importance of robust anomaly detection models and enhanced regulatory oversight in ensuring data integrity within Ontario's energy sector. Future work will explore machine learning approaches to automate bill-to-meter validation and detect anomalies in real-time, contributing to improved billing accuracy and consumer confidence in utility data reporting.

3.3 Bill to Meter verification GB - Manual Our investigation sampled several Utilities GB data, but we have limited our noted findings to two (Hydro One and Toronto Hydro) sample locations for brevity.

Note that the raw GB data presented is in the JSON script converted from the utility XML.

4.3.1 Hydro One

Our first sample is from Hydro One. In the GB billing data file, you will get the following details.

First bill.


```
"bills": [
  {
    "date": "2023-02-02",
    "due_date": "2023-02-22",
    "amount_due": 19335000,
    "account_type": "Unknown",
    "billing_period": "2022-12-29 - 2023-01-27",
    "bill_number": "Unknown",
    "currency": "CAD",
    "notes": [],
    "additional_data": [
      {
        "description": "Commodity Pricing Method"
      },
      {
        "description": "Unit of Measure for Usage"
      },
      {
        "description": "Service Type"
      }
    ],
```


Last Bill

```
{
  "date": "2024-11-01",
  "due_date": "2024-11-21",
  "amount_due": 20874000,
  "account_type": "Unknown",
  "billing_period": "2024-09-26 - 2024-10-26",
  "bill_number": "Unknown",
  "currency": "CAD",
  "notes": [],
  "additional_data": [
    {
      "description": "Commodity Pricing Method"
    },
    {
      "description": "Unit of Measure for Usage"
    },
    {
      "description": "Service Type"
    }
```


The field noted as "date" is the date the bill was issued, noted as the Invoice date below.

The field noted as "billing period" contains two dates, the first noted as "billing from" and the second noted as "billing to". In analysing Hydro One's data, it was determined that the billing to date had to be adjusted back one day. It is our supposition that the timestamp for the billing to date is "2022-11-26 00:00:00" which is midnight for the day ending "2022-11-25".

For our analytical purposes, we create a new field "adj_billing_to" to account for this presentation issue.

Sample Hydro One: note that reported billing to date was found to require backdating by one day.

invoice_date	billing_from	adj_billing_to	billing_to	bill_control_date
2023-02-02	12-29-2022	1-26-2023	1-27-2023	2023-02-02
2023-03-03	1-27-2023	2-24-2023	2-25-2023	2023-03-03
2023-04-03	2-25-2023	3-27-2023	3-28-2023	2023-04-03
2023-05-03	3-28-2023	4-26-2023	4-27-2023	2023-05-03
2023-06-02	4-27-2023	5-26-2023	5-27-2023	2023-06-02
2023-07-04	5-27-2023	6-26-2023	6-27-2023	2023-07-04
2023-08-02	6-27-2023	7-26-2023	7-27-2023	2023-08-02
2023-09-01	7-27-2023	8-25-2023	8-26-2023	2023-09-01
2023-10-03	8-26-2023	9-26-2023	9-27-2023	2023-10-03
2023-11-03	9-27-2023	10-26-2023	10-27-2023	2023-11-03
2023-12-01	10-27-2023	11-24-2023	11-25-2023	2023-12-01
2024-01-03	11-25-2023	12-22-2023	12-23-2023	2024-01-03
2024-02-01	12-23-2023	1-25-2024	1-26-2024	2024-02-01
2024-03-01	1-26-2024	2-23-2024	2-24-2024	2024-03-01
2024-04-03	2-24-2024	3-25-2024	3-26-2024	2024-04-03
2024-05-01	3-26-2024	4-24-2024	4-25-2024	2024-05-01
2024-06-03	4-25-2024	5-27-2024	5-28-2024	2024-06-03
2024-07-03	5-28-2024	6-25-2024	6-26-2024	2024-07-03
2024-08-01	6-26-2024	7-25-2024	7-26-2024	2024-08-01
2024-09-03	7-26-2024	8-26-2024	8-27-2024	2024-09-03
2024-10-02	8-27-2024	9-25-2024	9-26-2024	2024-10-02
2024-11-01	9-26-2024	10-25-2024	10-26-2024	2024-11-01

In the GB billing data file, you will get the following details (Note: “total_kWh” and “usage” both report the total meter read kWh):


```
"services": {
  "electricity": {
    "unit": "kWh",
    "totalCost": 49814000,
    "totalUsage": 1207056,
    "totalAdjustedUsage": 0,
    "reading": [
      {
        "reading_date": "2023-01-27",
        "reading": 22721162.25,
        "previous_reading_date": "2022-12-29",
        "previous_reading": 22660809.45,
        "number_of_days": 29,
        "multiplier": 1,
        "usage": 1207056,
        "adjusted_usage": 0
      }
```

date	totalUsage	reading_date	reading.1	previous_reading_date	previous_reading	number_of_days	multiplier	usage	Subtraction
2023-02-02	1207056	2023-01-27	22721162.25	2022-12-29	22660809.45	29	1	1207056	60352.8
2023-03-03	1089930	2023-02-25	22775658.75	2023-01-27	22721162.25	29	1	1089930	54496.5
2023-04-03	1338304	2023-03-28	22842573.95	2023-02-25	22775658.75	31	1	1338304	66915.2
2023-05-03	1414216	2023-04-27	22913284.75	2023-03-28	22842573.95	30	1	1414216	70710.8
2023-06-02	1421149	2023-05-27	22984342.2	2023-04-27	22913284.75	30	1	1421149	71057.45
2023-07-04	1835005	2023-06-27	23076092.45	2023-05-27	22984342.2	31	1	1835005	91750.25

GB meter data is presented as shown below. This is the very first record. This highlights the delay in the billing start date, which should be 12-29-2022'.

```
"meterData": [
  {
    "intervalStart": "2023-01-17 00:00:00",
    "intervalDuration": 60,
    "value": 1.515
```

This is the very last record. This highlights the overlap in the billing end date, which is 10-25-2024'.

```
{
  "intervalStart": "2024-11-16 23:00:00",
  "intervalDuration": 60,
  "value": 1.23
}
```

The following is the result of assigning the GB data to the relevant time allocation. For further clarity and later assignment of RPPTOU Tier and rates, we added an hour allocation of hour 1 to 24 and created our own timestamp.

meterDate	hour	value	timestamp
1-16-2023	24	1.515	202301162400
1-17-2023	1	1.291	202301170100
1-17-2023	2	1.946	202301170200
1-17-2023	3	1.572	202301170300
1-17-2023	4	1.442	202301170400
1-17-2023	5	1.563	202301170500

Then, we assigned hourly meter reads to the bill date. As you will note, the first hourly meter read is “2023-01-17” which we backed to the last hour (midnight) read on “2023-01-16”

meterDate	hour	value	bill_Date
1-16-2023	24	1.515	2-2-2023
1-17-2023	1	1.291	2-2-2023
1-17-2023	2	1.946	2-2-2023
1-17-2023	3	1.572	2-2-2023
1-17-2023	4	1.442	2-2-2023
1-17-2023	5	1.563	2-2-2023

Then, we compared the sum of meter reads by bill date with the total usage. The following issues arise from this comparison. First, from the billing data, we note the billing from date is 2022-12-19 whereas the first meter read is 2023-01-17. This creates the inability to verify the first bill, as we do not have the opening meter reads. Conversely, the last hourly meter read is 2024-11-16 where the last billing is 2024-10-17, creating excess meter reads. The reason for this has not been explored with the presenting utility. For the remaining billing periods, there were some noted differences in total usage not of a material nature.

bill_control	Sum of meter reads	Bill Total kWh	Difference
#N/A	855.50		
2023-02-02	437.92	1,207.06	- 769.14
2023-03-03	1089.76	1,089.93	- 0.17
2023-04-03	1337.10	1,338.30	- 1.20
2023-05-03	1414.22	1,414.22	-
2023-06-02	1421.15	1,421.15	-
2023-07-04	1835.01	1,835.01	-
2023-08-02	1915.90	1,915.90	-
2023-09-01	1689.07	1,689.07	-
2023-10-03	1255.95	1,255.95	-
2023-11-03	1067.54	1,067.54	-
2023-12-01	1145.14	1,144.02	1.12
2024-01-03	1179.24	1,179.11	0.13
2024-02-01	1631.98	1,632.02	- 0.03
2024-03-01	762.25	762.64	- 0.39
2024-04-03	906.18	907.01	- 0.83
2024-05-01	1321.89	1,321.89	-
2024-06-03	1040.24	1,040.24	-
2024-07-03	1531.29	1,531.29	-
2024-08-01	1756.93	1,756.93	-
2024-09-03	1720.50	1,720.50	-
2024-10-02	1470.86	1,470.86	-
2024-11-01	1216.74		

4.3.2 Toronto Hydro

Our second sample is from Toronto Hydro. In the GB billing data file, you will get the following details.

This is the first bill in the GB billing data file

```
"bills": [
  {
    "date": "2022-12-15",
    "due_date": "2023-01-08",
    "amount_due": 165.24,
    "account_type": "Unknown",
    "billing_period": "2022-11-10 - 2022-12-09",
    "start_date": "2022-11-10",
    "end_date": "2022-12-09",
    "bill_number": 554429121270,
    "currency": "CAD",
    "notes": [],
    "additional_data": [
      {
        "description": "Distributor Loss Factor",
        "amount": 1.03
      }
```


This is the last bill in the GB billing data file


```
{
  "date": "2024-09-16",
  "due_date": "2024-10-10",
  "amount_due": 202.56,
  "account_type": "Unknown",
  "billing_period": "2024-08-12 - 2024-09-10",
  "start_date": "2024-08-12",
  "end_date": "2024-09-10",
  "bill_number": 554534722740,
  "currency": "CAD",
  "notes": [],
  "additional_data": [
    {
      "description": "Distributor Loss Factor",
      "amount": 1.03
    }
```


The field noted as “date” is the date the bill was issued, noted as the Invoice date below.

The field noted as “billing period” contains two dates, the first noted as “billing from” and the second noted as “billing to”. In analysing Hydro One’s data, it was determined that the billing to date had to be adjusted back one day. It is our supposition that the timestamp for the billing to date is “2022-11-26 00:00:00” which is midnight for the day ending “2022-11-25". However, in analysing Toronto Hydro, we found that this adjustment was not required. Hence “billing to” as presented was left alone.

invoice_date	billing_from	adj_billing_to	billing_to	bill_control_date
2022-12-15	11-10-2022	12-8-2022	12-9-2022	2022-12-15
2023-01-17	12-9-2022	1-10-2023	1-11-2023	2023-01-17
2023-02-15	1-11-2023	2-8-2023	2-9-2023	2023-02-15
2023-03-15	2-9-2023	3-8-2023	3-9-2023	2023-03-15
2023-04-18	3-9-2023	4-11-2023	4-12-2023	2023-04-18
2023-05-16	4-12-2023	5-9-2023	5-10-2023	2023-05-16
2023-06-15	5-10-2023	6-8-2023	6-9-2023	2023-06-15
2023-07-18	6-9-2023	7-11-2023	7-12-2023	2023-07-18
2023-08-16	7-12-2023	8-9-2023	8-10-2023	2023-08-16
2023-09-18	8-10-2023	9-11-2023	9-12-2023	2023-09-18
2023-10-16	9-12-2023	10-9-2023	10-10-2023	2023-10-16
2023-11-16	10-10-2023	11-9-2023	11-10-2023	2023-11-16
2023-12-15	11-10-2023	12-10-2023	12-11-2023	2023-12-15
2024-01-16	12-11-2023	1-9-2024	1-10-2024	2024-01-16
2024-02-15	1-10-2024	2-8-2024	2-9-2024	2024-02-15
2024-03-15	2-9-2024	3-10-2024	3-11-2024	2024-03-15
2024-04-15	3-11-2024	4-8-2024	4-9-2024	2024-04-15
2024-05-16	4-9-2024	5-9-2024	5-10-2024	2024-05-16
2024-06-14	5-10-2024	6-9-2024	6-10-2024	2024-06-14
2024-07-16	6-10-2024	7-9-2024	7-10-2024	2024-07-16
2024-08-16	7-10-2024	8-11-2024	8-12-2024	2024-08-16
2024-09-16	8-12-2024	9-9-2024	9-10-2024	2024-09-16
2024-10-16	9-10-2024	10-8-2024	10-9-2024	2024-10-16

In the GB billing data file, you will get the following detail (Note “total_kWh” and “usage” both report the total meter read kWh.):


```
"services": {
  "electricity": {
    "unit": "kWh",
    "totalCost": 203.39999999999998,
    "totalUsage": 1063,
    "totalAdjustedUsage": 1094.36,
    "reading": [
      {
        "reading_date": "2022-12-10",
        "reading": 89986,
        "previous_reading_date": "2022-11-11",
        "previous_reading": 88923,
        "number_of_days": 2505600,
        "multiplier": 1,
        "usage": 1063,
        "adjusted_usage": 1094.36
      }
```

date	totalUsage	reading_date	reading.1	previous_reading_date	previous_reading	number_of_days	multiplier	usage	Subtraction
2022-12-15	1063	2022-12-10	89986	2022-11-11	88923	2505600	1	1063	1063
2023-01-17	1052	2023-01-12	91038	2022-12-10	89986	2851200	1	1052	1052
2023-02-15	1175	2023-02-10	92213	2023-01-12	91038	2505600	1	1175	1175
2023-03-15	1060	2023-03-10	93273	2023-02-10	92213	2419200	1	1060	1060
2023-04-18	938	2023-04-13	94211	2023-03-10	93273	2937600	1	938	938
2023-05-16	741	2023-05-11	94952	2023-04-13	94211	2419200	1	741	741
2023-06-15	970	2023-06-10	95922	2023-05-11	94952	2592000	1	970	970
2023-07-18	1182	2023-07-13	97104	2023-06-10	95922	2851200	1	1182	1182
2023-08-16	830	2023-08-11	97934	2023-07-13	97104	2505600	1	830	830
2023-09-18	1224	2023-09-13	99158	2023-08-11	97934	2851200	1	1224	1224
2023-10-16	778	2023-10-11	99936	2023-09-13	99158	2419200	1	778	778
2023-11-16	1054	2023-11-11	100990	2023-10-11	99936	2678400	1	1054	1054
2023-12-15	1069	2023-12-12	102059	2023-11-11	100990	2678400	1	1069	1069
2024-01-16	1058	2024-01-11	103117	2023-12-12	102059	2592000	1	1058	1058
2024-02-15	1073	2024-02-10	104190	2024-01-11	103117	2592000	1	1073	1073
2024-03-15	1196	2024-03-12	105386	2024-02-10	104190	2678400	1	1196	1196
2024-04-15	998	2024-04-10	106384	2024-03-12	105386	2505600	1	998	998
2024-05-16	915	2024-05-11	107299	2024-04-10	106384	2678400	1	915	915
2024-06-14	840	2024-06-11	108139	2024-05-11	107299	2678400	1	840	840
2024-07-16	1548	2024-07-11	109687	2024-06-11	108139	2592000	1	1548	1548
2024-08-16	1510	2024-08-13	111197	2024-07-11	109687	2851200	1	1510	1510
2024-09-16	1232	2024-09-11	112429	2024-08-13	111197	2505600	1	1232	1232
2024-10-16	1007	2024-10-10	113436	2024-09-11	112429	2505600	1	1007	1007

GB meter data is presented as shown below. This is the first record.

```
"meterData": [
  {
    "intervalStart": "2022-12-05 00:00:00",
    "intervalDuration": 60,
    "value": 0.86
  },
```

This is the last hourly meter read.

```
  {
    "intervalStart": "2024-12-03 23:00:00",
    "intervalDuration": 60,
    "value": 0.83
  }
]
```

meterDate	hour	value	timestamp
12-4-2022	24	0.86	202212042400
12-5-2022	1	0.92	202212050100
12-5-2022	2	0.38	202212050200
12-5-2022	3	0.88	202212050300
12-5-2022	4	0.93	202212050400
12-5-2022	5	0.99	202212050500
12-5-2022	6	1.44	202212050600
12-5-2022	7	2.58	202212050700
12-5-2022	8	1.37	202212050800
12-5-2022	9	1.13	202212050900

When comparing the sum of meter reads (by bill date) to the total usage, discrepancies arise due to misaligned data. Specifically, the billing start date of November 10, 2022, precedes the earliest available meter read (November 28, 2023), making it impossible to verify the opening meter readings for the first bill. In addition, the final available hourly meter read, which ends on November 27, 2024, postdates the required final billing date of November 11, 2024, causing extra meter reads to fall outside the billing period. The reason for this misalignment has not yet been explored with the presenting utility. Additionally, for the remaining billing periods, differences in total usage were observed, partly attributable to rounding adjustments. These inconsistencies highlight potential challenges in reconciling actual consumption data with billed amounts and may require further clarification from the utility.

Meter Data		Billing kWh	Diff
Row Labels	**Sum of value**		
#N/A	550.64		
2022-12-15	392.15	1,063.00	(670.85)
2023-01-17	1054.93	1,052.00	2.93
2023-02-15	1172.62	1,175.00	(2.38)
2023-03-15	1059.21	1,060.00	(0.79)
2023-04-18	937.16	938.00	(0.84)
2023-05-16	741.41	741.00	0.41
2023-06-15	969.53	970.00	(0.47)
2023-07-18	1181.98	1,182.00	(0.02)
2023-08-16	830.45	830.00	0.45
2023-09-18	1223.78	1,224.00	(0.22)
2023-10-16	778.46	778.00	0.46
2023-11-16	1054.87	1,054.00	0.87
2023-12-15	1068.66	1,069.00	(0.34)
2024-01-16	1057.62	1,058.00	(0.38)
2024-02-15	1072.91	1,073.00	(0.09)
2024-03-15	1195.52	1,196.00	(0.48)
2024-04-15	998.23	998.00	0.23
2024-05-16	915.38	915.00	0.38
2024-06-14	839.09	840.00	(0.91)
2024-07-16	1547.99	1,548.00	(0.01)
2024-08-16	1510.49	1,510.00	0.49
2024-09-16	1231.61	1,232.00	(0.39)
2024-10-16	1006.94	1,007.00	(0.06)
2024-11-18	987.94	1,026.00	(38.06)

Issue		
Per Billing	Per Meter Data	
First Read Date	First Read Date	
2022-11-10	2022-11-28	Mismatch
Last Read Date	Last Read Date	
2024-11-11	2024-11-27	Mismatch

4.4 Bill to Meter verification GB (Automated)

The automated approach focuses solely on validating the energy consumption recorded in the GB meter data against the usage values provided in the billing data (both derived from the GB standard). Two sample utilities are provided for this research, with three accounts in total. This section explains the process step by step:

Data Parsing and Conversion

XML Data Source: The raw GB meter data is originally provided in XML format. This XML contains detailed interval readings (typically recorded hourly) that document energy consumption over time. **Conversion to JSON:** Through UEnergyHub, the use of a custom parser infrastructure that reads the XML files and converts the data into a standardized JSON format and CSV format (formats more commonly used today in business services). This step helps to ensure that all energy readings, timestamps, and relevant attributes are organized consistently, which makes subsequent reuse of the data and analysis much easier.

Data Extraction and Alignment

Extracting Meter Readings: From the standardized JSON file, the application extracts key pieces of information, including **Timestamp:** When each energy reading was recorded. **Usage Value:** The actual energy consumed during that interval. **Aligning with Billing Periods:** The billing data (also derived from the GB system) includes a field for the reported energy usage for each billing period, typically labeled as totalUsage. UEnergyHub application identifies the start and end dates for each billing period and then selects only those meter readings that fall within these dates. **Summing the Interval Readings:** After filtering the meter readings by the billing period, the application sums all the individual usage values to calculate a total energy consumption figure for that period. This calculated sum is stored as meterSummary calculated.totalValue for validation.

Automated Comparison Process:

Comparison Formula: For each billing period, the application computes the difference using the following formula:

$$\text{Difference} = \text{totalUsage (from billing data)} - \text{meterSummary\_calculated.totalValue (from meter data)}$$

A small difference indicates that the meter readings closely match the reported usage, whereas a larger discrepancy may signal potential data alignment or quality issues.

Logging and Reporting: The application logs the results for each billing period and then aggregates these differences to provide a total difference for the entire data set. This reporting helps pinpoint periods where the values diverge significantly.

Detailed Results by Utility

Sample results from an automated comparison. Each table represents the results for a particular utility. The “Bill Date” column indicates the end date of the billing period, while the two usage columns show the reported usage (totalUsage) and the calculated usage (meterSummary_calculated.totalValue), respectively.

Utility: Hydro One – Account 1 Residential

Bill Date	totalUsage (Billing Data)	meterSummary_calculated.totalValue (Meter Data)	Difference	Percentage Difference (%)
2023-05-15	1818.898	1814.398	4.500	0.25%
2023-06-15	1170.268	1174.286	-4.018	-0.34%
2023-07-17	839.994	841.910	-1.916	-0.23%
2023-08-16	1276.260	1276.524	-0.264	-0.02%
2023-09-15	1262.680	1267.576	-4.896	-0.39%
2023-10-16	1291.270	1289.924	1.346	0.10%
2023-11-15	2764.422	2759.968	4.454	0.16%
2023-12-13	2644.874	2649.762	-4.888	-0.18%
2024-01-15	2764.066	2779.708	-15.642	-0.56%
2024-02-13	2450.036	2448.360	1.676	0.07%
2024-03-14	3279.908	3281.350	-1.442	-0.04%
2024-04-15	2603.602	2615.412	-11.810	-0.45%
2024-05-14	1081.800	1084.668	-2.868	-0.26%
2024-06-14	806.110	805.042	1.068	0.13%
2024-07-16	1152.430	1149.168	3.262	0.28%
2024-08-15	1470.012	1483.228	-13.216	-0.89%
2024-09-16	2021.616	2019.146	2.470	0.12%
2024-10-15	1362.934	1366.114	-3.180	-0.23%
2024-11-14	2313.142	2295.138	18.004	0.78%
2024-12-12	3778.062	3785.610	-7.548	-0.20%
2025-01-15	5917.172	5918.046	-0.874	-0.01%

Utility: Hydro One – Account 2 Residential

Bill Date	totalUsage (Billing Data)	meterSummary_calculated.totalValue (Meter Data)	Difference	Percentage Difference (%)
2022-10-17	1,119	1,126.73	-7.73	-0.69%
2022-11-16	934	938.36	-4.36	-0.46%
2022-12-15	1,063	1,066.67	-3.67	-0.34%
2023-01-17	1,052	1,034.04	17.96	1.74%
2023-02-15	1,175	1,179.99	-4.99	-0.42%

2023-03-15	1,060	1,070.48	-10.48	-0.98%
2023-04-18	938	931.69	6.31	0.68%
2023-05-16	741	762.18	-21.18	-2.78%
2023-06-15	970	954.41	15.59	1.63%
2023-07-18	1,182	1,174.07	7.93	0.68%
2023-08-16	830	863.77	-33.77	-3.91%
2023-09-18	1,224	1,208.94	15.06	1.25%
2023-10-16	778	787.01	-9.01	-1.14%
2023-11-16	1,054	1,054.56	-0.56	-0.05%
2023-12-15	1,069	1,058.50	10.50	0.99%
2024-01-16	1,058	1,050.82	7.18	0.68%
2024-02-15	1,073	1,092.31	-19.31	-1.77%
2024-03-15	1,196	1,191.46	4.54	0.38%
2024-04-15	998	1,005.97	-7.97	-0.79%
2024-05-16	915	924.89	-9.89	-1.07%
2024-06-14	840	824.40	15.60	1.89%
2024-07-16	1,548	1,517.93	30.07	1.98%

Utility: Alectra

Bill Date	totalUsage (Billing Data, kWh)	meterSummary_calculated.totalValue (Meter Data, kWh)	Difference	Percentage Difference (%)
2023-04-13	826.20	813.78	12.42	1.53%
2023-05-09	563.52	572.26	-8.74	-1.53%
2023-06-09	648.15	650.77	-2.62	-0.40%
2023-07-12	591.44	593.02	-1.58	-0.27%
2023-08-10	554.58	555.92	-1.34	-0.24%
2023-09-12	624.85	629.60	-4.75	-0.75%
2023-10-11	467.56	481.32	-13.76	-2.86%
2023-11-09	730.91	725.01	5.90	0.81%
2023-12-11	730.91	1054.56	-323.65	-30.69%
2024-01-10	621.38	640.67	-19.29	-3.01%
2024-02-09	804.67	783.81	20.86	2.66%
2024-03-11	612.71	633.25	-20.54	-3.24%
2024-04-10	580.03	575.25	4.78	0.83%
2024-05-09	506.79	520.94	-14.15	-2.72%
2024-06-11	555.17	535.50	19.67	3.67%
2024-07-10	584.43	587.60	-3.17	-0.54%
2024-08-12	798.87	809.03	-10.16	-1.26%
2024-09-11	624.85	628.97	-4.12	-0.66%
2024-10-10	552.82	558.69	-5.87	-1.05%

5. Interpretation of Results

Hydro One – Account 1: The total difference across billing periods is –35.782 kWh, which is very minor when compared to the overall usage values. This indicates that the automated process reconstructs the meter data for this set of accounts.

Hydro One – Account 2: The cumulative difference is only –2.18 kWh, demonstrating an almost but not perfect match between the meter data and the billing usage values. Such a small difference provides validation of the data alignment and calculation process.

Alectra: The Alectra account shows a larger cumulative difference of –370.11 kWh. Although the discrepancy is more significant than in the Hydro One cases, this result flags the account for further review. It may indicate issues such as data misalignment, possible errors in the XML data conversion process, or inconsistencies in meter readings.

Automated Comparison Conclusion

Our automated comparison method—using a Node.js application that parses GB XML data into JSON, extracts and aligns meter readings, and sums them over defined billing periods—has proven effective for validating energy usage. In most cases, the differences between the reported billing usage and the calculated meter usage are minimal, reinforcing the reliability of the automated process. Larger discrepancies, as seen in the Alectra account, highlight the need for further investigation and possibly enhanced data consistency checks. This automated verification framework is crucial for ensuring accurate energy billing and improving overall data quality.

4.5 Meter to Meter verification comparing GB vs EBT

In this section, a comparison of the GB meter data against the EBT usage data, both sourced from an Ontario utility, is performed by a market expert. The objective was to determine whether any significant discrepancies exist between the researcher's data from previous research and assess potential reasons for any differences observed using a sample EBT data set from another Utility.

Line Loss Factor Clarification: The EBT Loss Factor seen in the comparison below is the Line Loss Factor explicitly defined within the XML document containing EBT usage data. It is a flat percentage calculated and approved for application by OEB at infrequent intervals, not necessarily the exact line loss factor that existed during that service period.

Data precision: All volume data and line loss factors conformed to numbers defined to four decimal places. The comparison below only shows two decimal places for clarity and represents the loss factor as a percentage to two decimal places to maintain the original data precision. The EBT Relative to GB Meter percentages were not rounded and were only displayed to two decimal places for visual consistency with the EBT Loss Factor precision.

Data Comparison

The table below provides a comparison between the GB meter data and EBT usage data, only where both were concurrently available in the sourced datasets, to arrive at a percentage difference:

Month & Year	GB Meter Data	EBT Usage Data	EBT Relative to GB Meter (%)	EBT Loss Factor (%)	Discrepancy (%)
Feb-24	881,920.93	911,906.24	103.40%	3.40%	0.0000%
Mar-24	848,762.82	878,911.52	103.55%	3.40%	0.1521%
Apr-24	831,524.59	859,827.06	103.40%	3.40%	0.0037%
May-24	865,288.35	893,774.78	103.29%	3.40%	-0.1079%
Jun-24	746,287.81	771,655.53	103.40%	3.40%	-0.0008%
Jul-24	709,364.35	734,589.86	103.56%	3.40%	0.1561%
Aug-24	829,878.53	857,002.59	103.27%	3.40%	-0.1316%
Sep-24	775,719.32	802,768.53	103.49%	3.40%	0.0870%
Oct-24	778,836.73	804,601.20	103.31%	3.40%	-0.0919%
Nov-24	565,417.83	584,285.99	103.34%	3.40%	-0.0630%
Dec-24	590,193.25	610,259.82	103.40%	3.40%	0.0000%

Key Findings and Considerations

Minimal Discrepancy Observed: Across the overlap period where we have both EBT and GB data for full calendar months, after accounting for the line loss factor provided by the distributor within the XML EBT transactions, the discrepancy observed was **0.0003%**.

Billing Data Consideration: The customer's bill is calculated based on EBT data including line losses, and the GB standard requires representing the billing accurately. Therefore, it is expected that GB billing data and EBT usage data will reflect volumes higher than those in the GB meter data, with the difference being (ideally entirely) attributed to the line loss factor. Further investigations should assess whether GB meter data should distinctly include line losses for a more direct comparison with billing data.

Previous Researcher Investigation: A previous EBT vs GB meter data comparison conducted by a researcher showed significant differences in the values between EBT and GB meter data, which was not consistent with this analysis.

The recommendation would be for further research to understand the data structure differences between EBT and GB meter data, variations across different utilities, and their individual interpretations of the GB standard compliance. Further standardization and validation methods may enhance the usability of these datasets for research and regulatory oversight.

4.6 Key Discrepancies Identified

Findings from Previous Research: Our previous study highlighted systemic issues related to data inconsistencies between GB and EBT datasets across multiple Ontario utilities, which was also observed in this research. Key findings included: Significant discrepancies in reported energy consumption, with EBT data frequently showing higher usage figures. Variability in data quality and completeness, particularly in EBT datasets where estimation and adjustment practices varied by utility. Challenges in aligning data due to inconsistent timestamp formats and differing billing cycle definitions.

Findings From New Research:

Data Granularity: GB data offers high-resolution interval readings, while EBT data aggregates consumption over longer billing cycles.

Focus Areas: GB emphasizes consumer-facing data for energy management, whereas EBT focuses on transactional data for billing and compliance.

Data Accuracy: Discrepancies in consumption values often stem from estimation practices in EBT data versus actual meter readings in GB data.

Time Alignment: Differences in how billing periods and interval data are timestamped can lead to misalignments when comparing GB and EBT datasets.

These findings underscore the need for improved data integration and validation processes to ensure consistency and accuracy in energy reporting. They also suggest the importance of regulatory oversight in standardizing data practices across utilities.

Implications for Current Research: Understanding these differences is critical as we integrate data for analysis. By establishing a clear baseline of how GB and EBT data differ, we can better identify anomalies, validate data integrity, and propose strategies for enhancing data accuracy in utility operations. This comparative overview contextualizes our current findings and sets the stage for deeper investigations into the systemic factors affecting energy data quality in Ontario.

5 Validation and Expert Review

To ensure our findings were accurate and reliable, we employed cross-validation—comparing the integrated dataset with historical billing data—and collaborated with SMEs. These steps revealed critical insights into energy data discrepancies that regulators have not fully documented or recognized, emphasizing the need for robust validation to enhance data accuracy and billing integrity.

5.1 Issues and Suggested Remediation

Issue 1: Misalignment in Meter Data Timestamps

Problem: Some utilities' meter data recorded the end of each interval (leading to a one-hour offset), while others were correctly aligned. This discrepancy can confuse consumers and analysts and complicate comparisons across utilities.

Recommendation:

- **Standardize Timestamp Reporting:** We believe that the regulator (OEB) and Green Button Alliance (GBA) may wish to consider defining a universal standard indicating whether intervalStart represents the beginning or end of each interval and clarifying handling of time zones and daylight savings.

Issue 2: Misalignment Between Meter Data and Billing Start Date

Problem: Meter data often fails to align with the official billing start date, effectively shortening the mandated two-year data availability and complicating bill verification.

Recommendation:

- **Align Meter Data with Billing Periods:** The OEB considers requiring utilities to ensure that the first meter interval exactly matches the official billing start date.

- **Refine GB Data Specs:** The OEB is considering collaborating with GBA to mandate explicit alignment of meter intervals and billing cycles.
- **Transparency & Audits:** We believe utilities may wish to inform customers whenever meter data does not align with billing periods, and the OEB may wish to conduct regular audits to confirm proper alignment.

Issue 3: Discrepancies Between Summarized Billing Data and Interval Meter Reads

Problem: The total billed consumption often does not match the sum of interval reads, causing confusion and undermining consumer trust.

Potential Causes:

- Inconsistent or missing data validation
- Different aggregation or rounding methods
- Misalignment with time zone/daylight savings changes
- Gaps in meter data or transmission errors
- Manual or automated billing adjustments

Recommendation:

- **Mandatory Data Validation:** The OEB considers requiring utilities to verify interval meter data consistency before presenting summarized billing information to consumers.
- **Standardize Aggregation & Rounding:** The OEB considers guiding utilities to adopt a uniform methodology for aggregating intervals, clearly documenting any rounding or estimations.
- **Time Consistency:** The OEB considers defining how time zones and daylight savings adjustments are handled across all platforms.
- **Regular Compliance Checks:** The OEB considers conducting regular audits to ensure that interval meter reads align accurately with billed totals.

Issue 4: Inconsistent Application of RPP TOU Rates Across Utilities

Problem: Although the OEB sets RPP TOU rates, each utility or billing agent applies them independently, leading to variation in how intervals are classified as on-, mid-, or off-peak.

Recommendation:

- **OEB Standardization of TOU Rate Application:** The OEB considers publishing a universal "Hourly RPP TOU Application Table" on its open data site, detailing on-/mid-/off-peak hours and handling daylight savings uniformly.

Consistent Utility Adoption: The OEB considers requiring all utilities to use this table to eliminate discrepancies in TOU billing assignments and improve transparency for consumers.

6 Justification for Standardization

Prevents Billing Inconsistencies: Even minor per-customer discrepancies can add up significantly when aggregated across a large customer base.

Enhances Customer Trust: Publicly accessible TOU rate application details allow consumers and analysts to verify billing accuracy.

Reduces Compliance Risks: A unified approach to TOU rates helps avoid violations caused by inconsistent interpretations.

Facilitates Automation: Integrating standardized rates into billing systems eliminates errors from manual or varied rate structures.

Streamlines System Upgrades: Standardization lowers vendor lock-in and replacement costs, ultimately reducing customer rates.

Issue 1: Automated Validation Before Billing

We believe that utilities and their third-party billing agents may wish to consider implementing automated validation to check TOU rate assignments against interval meter data before finalizing bills. Embedding the assigned TOU rate in hourly meter data would allow direct comparison between consumption, assigned TOU tier, and resulting charges.

Justification for Embedding TOU Rates:

Improves Transparency & Validation: Third-party analysts and consumers can independently confirm whether TOU billing matches metered usage.

Enhances Billing Accuracy: Linking each hour's consumption to its actual TOU rate reduces errors in monthly bill summaries.

Reduces Consumer Disputes: Clear data on how usage is billed streamlines dispute resolution.

Facilitates Compliance Audits: Regulators can easily verify correct TOU pricing.

Implementation Recommendations: The OEB considers requiring each utility to include the TOU rate for every interval in its GB data exports. Utilities may wish to consider validating that hourly-based charges match the final billed amount before issuing statements. The OEB considers periodic audits to confirm consistent TOU applications.

Issue 2: Lack of Clear Guidance from the OEB on Billing Algorithms

The OEB approves rates but defers actual billing formulas to each utility or billing provider, which can lead to inconsistent applications and alignment issues.

Recommendations

Standardized Billing Algorithms: The OEB considers publishing official formulas for calculating fixed, kWh, and kW charges—along with proration guidelines for mid-cycle starts/stops.

OEB-Hosted Hourly RPP TOU Table: The OEB considers maintaining a public, authoritative table defining on-/mid-/off-peak hours.

Mandatory Integration of TOU Rates into GB Data: OEB consider requiring every hourly interval to include its assigned TOU rate, making it possible to validate bills before finalization.

Automated Pre-Bill Validation & Audits: OEB consider mandating that utilities validate total charges against interval data.

Clear Dispute Resolution: OEB consider requiring utilities to provide transparent billing breakdowns and to report unresolved disputes for independent review.

Issue 3: Limited Availability of Up-to-Date OEB-Approved Tariff Rates

The OEB's annual database can lag by a year, and limited updates on its open data site hinder real-time bill verification.

Recommendation: OEB Tariff Rate Database: OEB consider maintaining a live, machine-readable repository of all approved rates—updated as soon as they take effect.

Utilities may wish to consider verifying new rates 30 days in advance, ensuring public availability before implementation.

The OEB considers quarterly audits to confirm accurate rate usage.

Issue 4: Lack of Granular Billing Detail in GB Data

GB billing data is provided only at a high-level summary, requiring consumers and analysts to reverse-engineer charges down to a granular level to verify accuracy.

Recommendation: Granular Billing Detail: The OEB considers requiring utilities to embed exact billing formulas for each charge (fixed, kWh, TOU, demand, etc.) directly in GB data. Include actual TOU rate assignments, proration details, and loss factor adjustments for each hour. Utilities may wish to consider validating all GB billing data against final bills to reduce discrepancies and simplify third-party verification.

Issue 5: Inconsistent Identification of Rate Classes and Zones in GB Data

Some utilities do not clearly identify subclasses or rate zones, making it difficult to map usage to the correct tariff.

Recommendation: Standardized Rate Class and Zone Identification: The OEB considers clarifying all rate classes, subclasses, and zones in official tariff sheets. Utilities may wish to consider aligning GB data fields to these standardized classes and zones. The OEB considers maintaining a public, machine-readable repository of these classifications, ensuring utilities use uniform naming.

Justification: Ensures consistency across utilities. Simplifies bill verification for consumers. Enhances regulatory oversight. Lowers costs associated with system upgrades and data inconsistencies.

7 Conclusion

This study builds on previous research comparing GB and EBT data through the UEnergyHub portal and OEB EBT-certified systems. The analysis highlighted systemic issues, including mismatches between calculated and billed energy consumption, inconsistencies in data granularity, and discrepancies related to time alignment.

Lack of access to standardized, digital energy billing and meter information significantly hampers research and innovation in the energy sector. Smart grid technology and performance-based algorithms are hindered without accurate, granular data. This creates challenges for local and global markets to interconnect systems.

Everyone is looking for a “silver bullet” to optimize energy usage and develop effective energy management strategies. Ontario, Canada, is a “perfect storm” of complexity in managing the “source of truth” for energy data. With over 50 electric and natural gas utilities, varying in size from thousands to millions of meter points, no utility uses the exact same infrastructure, as legacy Customer Information Systems (CIS) and Meter Data Stores (ODS/MDM or MDM/R) have evolved for specific utility use cases over time.

This study provides a data-driven foundation for further exploration into Ontario's energy data landscape and emphasizes the necessity for collaboration between utilities, regulators, and third-party data providers. These findings highlight the need for further research to standardize data processing and improve transparency in energy billing practices. Given the observed discrepancies, additional validation measures, including independent verification of utility billing calculations, may be required to ensure fair and accurate charges for consumers.

Acknowledgement

We thank Active Business Services and Screaming Power Inc for their financial support. We also thank the many industries, subject matter specialists, and utility account holders for providing their data through the UEnergyHub and EBT network.

Reference

[1] https://www.missiondata.io/news/2024/5/14/missiondata-announces-university-partnership-on-green-button-scorecard

[2] IESO. (2024). Ontario's Energy System Overview. Available at: https://www.ieso.ca/en

[3] Green Button Alliance. (2022). Green Button information. Available at: https://www.greenbuttonalliance.org

[4] Ministry of Energy: Regulatory proposal for province-wide implementation of Green Button, Last updated: September 16, 2021. Online: https://ero.ontario.ca/notice/019-2564

[5] Screaming Power Inc. (2024). *Energy Data Solutions*. Available at: https://www.screamingpower.ca/uenergyhub-empowering-ontario-for-a-greener-future-through-unbiased-research/

[6] The ElementTree XML API: https://docs.python.org/3/library/xml.etree.elementtree.html

[7] Sulaiman Kagumire: Anomaly Detection with Machine Learning, Available at: https://sulaimank.medium.com/anomaly-detection-with-machine-learning-8fa942fb5adc. May 12, 2021

[8] Scholkopf, B., et al. (1999). Support Vector Method for Novelty Detection. IEEE Transactions on Neural Networks, 12(3), 582-588. Available at: https://ieeexplore.ieee.org

[9] Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825–2830. Available at: https://scikit-learn.org

[10] The Ontario EBT Working Group: Electronic Business Transactions (EBT) Standards Document for Retail Settlement in the Electric Retail Open Access Industry, Version 4.0, January 21, 2008. Online: www.oeb.ca/oeb/_Documents/Regulatory/Ontario_EBT_Standards_v4.pdf

[12] https://www.oeb.ca/sites/default/files/2024-04/Energy-at-a-glance-2022-2023-en.pdf