Data-driven Targeting in Commercial Energy Programs using AMI Data and Digital Twin Modeling

Pranav Gupta, ICF
Mitch Brown, ICF

ABSTRACT

This paper presents a digital twin modeling framework that integrates utility Advanced Metering Infrastructure (AMI) data, firmographic data, and physics-based building energy simulation to support customer targeting in commercial energy efficiency programs. By linking interval energy use data with simulated commercial building archetypes, the framework provides account-level estimates of electricity savings potential across multiple energy efficiency measure types.

The analysis covers 57,214 AMI accounts across 58 business types, which are mapped to 6,174 building archetypes to evaluate the potential impacts of 38 energy efficiency, beneficial electrification, and demand response upgrades. This enables consistent estimation of upgrade-level savings potential across commercial building types under standardized modeling assumptions.

A targeted application explored in this paper focuses on customers in the K–12 school, hospital, and hospitality sectors, where over 12,000 accounts were analyzed to characterize the distribution of technical savings potential. Results show that savings potential is highly concentrated within each sector, with a relatively small subset of accounts accounting for a disproportionate share of total modeled savings.

The paper describes the digital twin methodology and demonstrates how modeled savings outputs can be used to support customer recruitment and lead prioritization for program outreach. While the results are based on ex ante modeled technical potential and do not reflect realized program outcomes, they provide a data-driven basis for identifying high-opportunity customers and informing targeting decisions in commercial sector energy efficiency programs. The digital twin framework in this study was implemented within a utility commercial energy efficiency program; however, the same methodology is scalable and can be extended to beneficial electrification and demand response programs across multiple sectors.

INTRODUCTION

Commercial buildings play a central role in achieving energy efficiency, affordability, and grid reliability objectives. With U.S. electricity demand projected to increase by nearly 50% by 2050 and the commercial sector expected to account for approximately 38% of total electricity demand, utilities face mounting pressure to increase participation across commercial energy efficiency and demand response programs (EIA 2025). Despite substantial investments in customer data, targeting tools, and digital engagement platforms, commercial program participation remains uneven, even among customers that are technically eligible, cost-effective, and aligned with program objectives. At the same time, accelerating electricity demand driven by electrification and data-intensive end uses, along with growing exposure to grid reliability risks, has increased the importance of more targeted and effective commercial sector program strategies.

Conventional targeting approaches, often based on broad segmentation by business type, building size, or historical participation can struggle to identify customers with the greatest potential for energy savings or grid benefits. This challenge is particularly pronounced within the small and medium-sized business (SMB) segment, where large numbers of heterogeneous customers and limited site-specific data make it difficult to differentiate opportunities across accounts. As a result, high-potential customers within this segment are often overlooked despite representing a meaningful share of aggregate savings potential. Moreover, commercial and industrial customers do not make participation decisions based solely on observable business characteristics. Instead, adoption is shaped by internal organizational factors such as budget approval processes, competing capital priorities, stakeholder alignment, operational risk tolerance, and limited staff time (Gillis and Johnson 2026). These internal dynamics determine whether a business moves from awareness to action, and whether that participation is sustained over time. As a result, there is a need for data-driven approaches that can better differentiate technical savings potential across individual customer accounts and provide a basis for prioritization in program targeting and lead generation.

Recent advances in AMI deployment and building energy modeling create new opportunities to improve program design and implementation. High-frequency interval data provides insight into operational patterns and seasonal load drivers, while physics-based building simulations enable estimation of measure-specific savings. When combined, these tools can support a "digital twin" representation of individual customer buildings, enabling utilities to move from broad segmentation to account-level targeting.

This paper presents the application of a digital twin modeling framework to enhance customer targeting for commercial energy efficiency programs implemented by a large investor-owned utility. The primary objective is to characterize variation in savings potential across commercial customers and demonstrate how these estimates can be used to inform customer prioritization for program outreach.

DATA SOURCES

The analysis is based on data provided by a large investor-owned electric utility in the southern region of the U.S.A. with a fully deployed AMI network and a portfolio of commercial energy efficiency incentive programs. The datasets include AMI interval data, Customer Information System (CIS) data, and historical program participation records. Together, these sources enable characterization of customer load profiles, building and business attributes, and prior program participation, forming the basis for digital twin matching and subsequent targeting analysis.

AMI Data

AMI data is high-resolution electricity consumption data recorded at regular intervals (e.g., 15-minute). In this analysis, the data was recorded at the meter level at 1-hour intervals and spanned multiple years, with most of the usable data covering the 2022–2025 period.

The interval data was used to characterize customer load profiles, including temporal patterns such as daily usage, seasonality, and variability in consumption. This information served as the primary input for the digital twin matching process, where observed load profiles were compared against simulated building archetype profiles to identify representative matches at the account level.

Prior to use, the AMI data was screened for completeness and quality. Accounts with insufficient temporal coverage, extended data gaps, prolonged periods of zero consumption, or anomalous patterns indicative of metering or data integrity issues were excluded from the analysis.

CIS Data

CIS data is account-level data that includes attributes such as account identifiers, account status, business type/sector classifications, and geographic information. These records were used to identify commercial accounts, associate AMI meters with accounts, and support sector-level segmentation.

To improve representation of building characteristics relevant for energy modeling, the CIS data was supplemented with commercially available third-party datasets comprising firmographic and property attributes, including estimated floor area, building use type, and heating and cooling system indicators where available. These attributes were used to constrain the set of candidate building archetypes during the digital twin matching process.

Program Participation Data

The historical program participation data captures customer engagement across more than 4,000 commercial accounts with past participation activity. These records reflect customer involvement in programs that include standard prescriptive, custom, and small business offerings. The dataset includes detailed program and measure identifiers, along with participation and installation dates. This information enables analysis of baseline participation patterns, including the types of measures installed (e.g., lighting, HVAC, controls), the timing and frequency of participation, and repeat engagement across program years.

Figure 1 presents the distribution of historical program participation records by high-level measure categories (e.g., lighting, HVAC, controls). Each non-null record was grouped into broader categories to improve interpretability and align with common commercial program design conventions. As shown in the figure, lighting measures account for the majority of historical participation, while HVAC, controls, and other measures represent a smaller share.

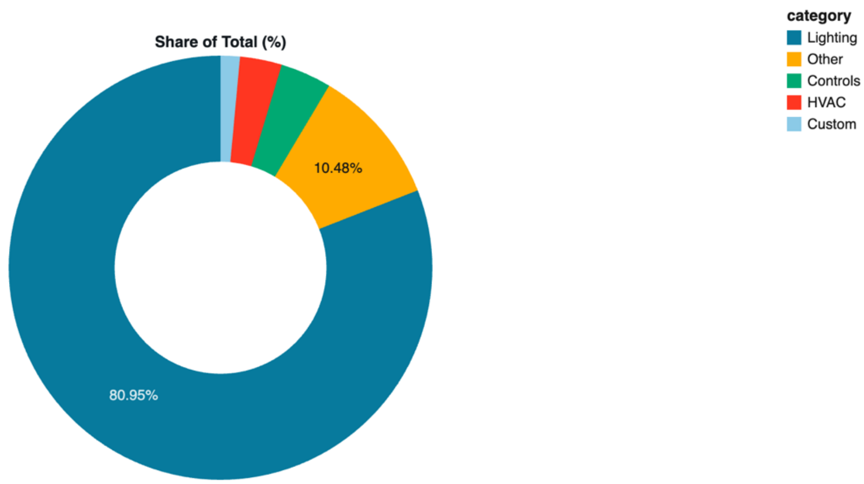

Figure 1. Distribution of historical program participation grouped by measure categories.

METHODOLOGY

The digital twin methodology integrates high-resolution AMI data with CIS data and physics-based building energy simulations to estimate upgrade-level electricity savings potential for commercial customers. The framework is designed to support scalable, account-level analysis across diverse commercial building types without requiring site-specific audits or field data collection.

In this paper, "digital twin" refers to the physics-based building energy model most closely resembling the commercial building. Each digital twin is created by matching observed AMI load profiles for an individual meter to a library of pre-simulated building energy models (archetypes). This digital twin framework is implemented as a matching and inference process, in which each account's observed interval consumption is compared against simulated baseline load profiles to identify a representative building archetype. While the matching algorithm is proprietary, its broader structure can be described in terms of the below core components and workflow steps.

At the core of the approach is a library of simulated commercial building archetypes constructed to represent the diversity of building stock within the utility's service territory. The archetype library comprises over 6,000 distinct configurations generated using OpenStudio based commercial building simulations informed by national building stock datasets and regional construction characteristics (NREL 2025). Archetypes are parameterized across key dimensions that drive electricity consumption, including building use type, conditioned floor area, construction vintage, envelope performance, internal loads and schedules, and HVAC system configuration.

Each archetype produces a baseline whole-building electric load profile at hourly resolution, along with end-use disaggregation for major components such as cooling and heating. To evaluate energy efficiency, beneficial electrification, and demand response opportunities, a portfolio of 38 *upgrades* was simulated for each archetype. In this paper, the term "upgrade" is used to refer to a modeled intervention, which may represent either an individual measure or a bundle of measures implemented together.

Upgrade-level energy impacts were calculated as the difference between baseline and post-upgrade simulated electricity consumption, enabling consistent estimation of annual savings across all modeled buildings. These estimates represent ex ante technical savings potential and are not calibrated to realized program savings at the individual measure or account level. As such, results should be interpreted as indicative of relative savings potential rather than precise predictions of realized outcomes.

Observed commercial customer AMI data were mapped to the archetype library using a proprietary matching algorithm. The matching process primarily relies on statistical comparison between hourly AMI load profiles and archetype baseline profiles. CIS data was incorporated as constraints to restrict the set of candidate archetypes, thereby improving matching accuracy and reducing ambiguity across similar load shapes. This matching process captures key operational characteristics reflected in observed load shapes, such as base-load behavior, seasonality, and daily usage patterns. Multiple customers may be mapped to the same archetype, reflecting common construction and operational characteristics across the commercial building stock.

Figure 2 compares observed monthly electricity consumption from AMI data with the corresponding digital twin–modeled consumption for a representative commercial account. The comparison highlights how the matched archetype (or digital twin) closely reproduces the seasonal pattern and magnitude of observed usage, capturing both the summer peak and lower winter loads.

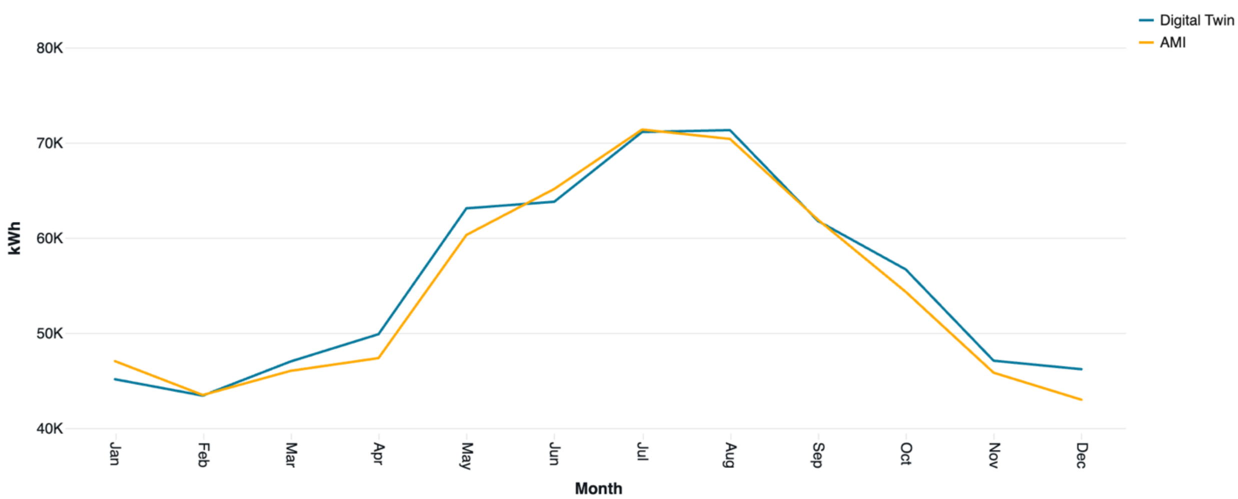

Figure 2. Comparison of observed AMI load profile and matched archetype baseline (or digital twin) for a representative account.

Once matched, customer-specific savings potential for each upgrade was derived directly from the simulated performance of the corresponding archetype. This structure enables parallel evaluation of multiple upgrades across a large number of commercial accounts and supports consistent comparison of savings potential across building types. Outputs from the digital twin analysis include account-level estimates of baseline consumption, upgrade-level savings potential, and relative performance rankings across accounts. The results represent ex ante estimates of relative savings potential and are used to support customer prioritization by highlighting accounts with comparatively higher technical potential, rather than replicating ex post evaluation of realized program savings. Figure 3 illustrates the end-to-end workflow, including data ingestion, digital twin matching using AMI and CIS inputs, and downstream estimation of savings potential for targeted outreach.

By integrating simulated building archetypes directly within the digital twin workflow, the methodology moves beyond broad sector level assumptions and enables account-level insights at scale. This structure supports scalable analysis and consistent comparison of potential impacts across multiple upgrades and building types. The results presented in the following section illustrate how this approach can be used to identify variation in savings potential across customers and support prioritization of higher-impact opportunities.

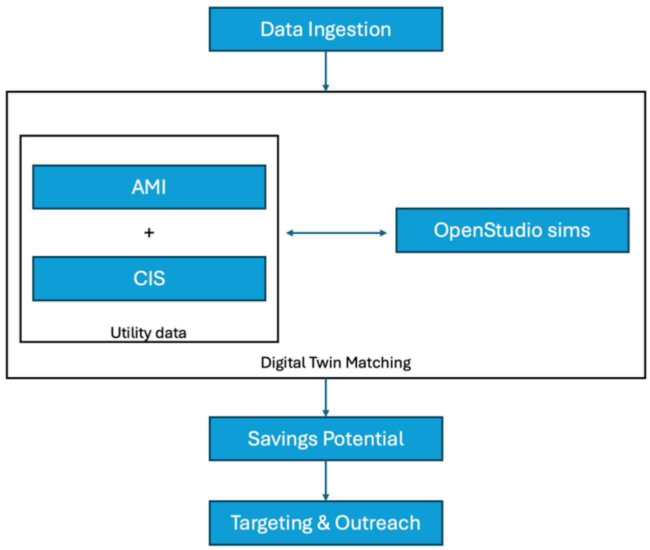

Figure 3. End-to-end targeting workflow.

RESULTS

Sector Coverage and Digital Twin Sample

A customer targeting application of the digital twin framework focused on three commercial business-type segments: hospital, hospitality, and K–12 schools. These segments were selected due to their operational diversity, energy intensity, and historical underrepresentation in some commercial energy efficiency programs.

In total, over 12,000 commercial customer digital twins were created across these sectors by mapping AMI load profiles and CIS attributes to simulated building archetypes.

Table 1 summarizes the distribution of modeled accounts by business type. Hospitality and K–12 schools' customers account for the majority of the modeled sample, while hospitals represent a smaller share of accounts but are typically larger and more energy-intensive facilities. Each digital twin represents an individual commercial customer's baseline operating characteristics and serves as the analytical unit for estimating upgrade-level electricity savings potential.

Table 1. Distribution of commercial digital twins by sector.

Sector	Digital Twins created	Share of Total (%)
Hospital	1,905	15%
Hospitality	5,143	41%

Sector	Digital Twins created	Share of Total (%)
K-12 School	5,573	44%

Distribution of Savings Potential Across Digital Twins

Across all three sectors, modeled electricity savings potential exhibits a strongly right-skewed distribution when evaluated at the digital twin level. Rather than being evenly distributed across customers, savings potential is concentrated among a relatively small subset of digital twins within each sector. Figures 4–6 illustrate this pattern by ranking digital twins from highest to lowest estimated annual electricity savings, where savings for each digital twin are defined as the maximum single upgrade savings potential across all simulated upgrades.

In each sector, cumulative savings accrue rapidly among the highest-ranked digital twins, with diminishing marginal returns beyond the top opportunity tier. This indicates that digital twins with the greatest technical potential account for a disproportionate share of total modeled savings, while a long tail of digital twins contribute relatively modest incremental savings.

Sector-Specific Concentration Patterns

While the overall concentration pattern is consistent across sectors, the magnitude and drivers of savings concentration vary by business type.

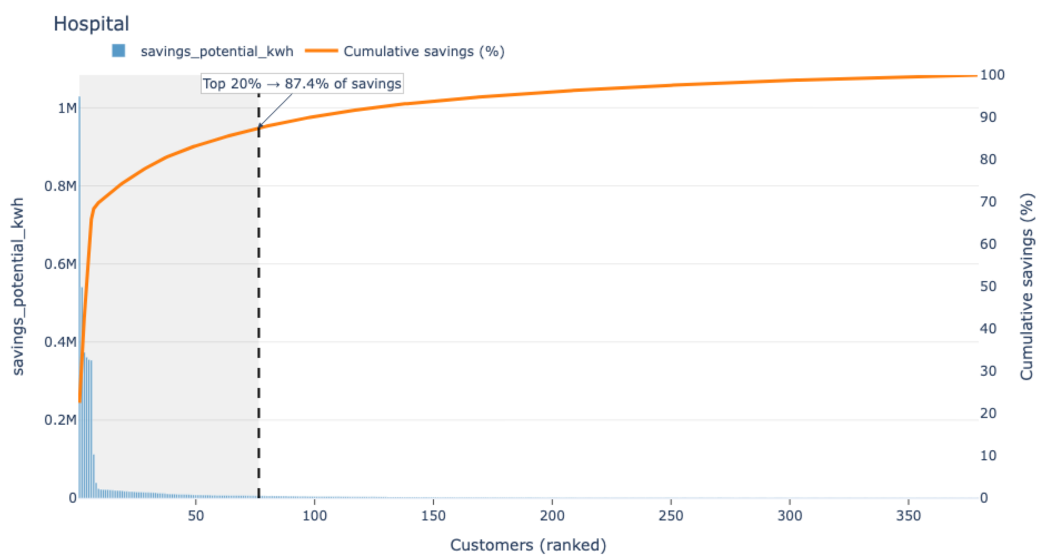

Figure 4. Pareto chart of modeled electricity savings potential for hospital digital twins.

The **hospital sector** exhibits a pronounced concentration of modeled savings potential among a relatively small number of high-impact digital twins. As shown in Figure 4, the top 20% of hospital digital twins account for 87.4% of total modeled savings. The bar distribution reveals a steep initial decline, indicating that a limited number of large, energy-intensive facilities dominate sector-level savings potential. This pattern is consistent with hospitals' round-the-clock operations, high ventilation requirements, and reliance on large central plant systems, which amplify the savings impact of HVAC, controls, and system optimization measures.

Across the hospital sector, the most frequently occurring upgrade with the highest potential savings was a bundled configuration combining envelope improvements (wall and roof insulation and new windows) with LED lighting and high-efficiency HVAC system upgrades (heat pump RTU and heat pump boiler). The upgrade led to higher modeled savings because improvements to the building envelope reduce heating and cooling demand, enabling HVAC systems to operate more efficiently across operating conditions.

Overall, the hospital results highlight both a high concentration of modeled savings potential and significant variation across facilities, with a relatively small subset of buildings accounting for a large share of estimated savings.

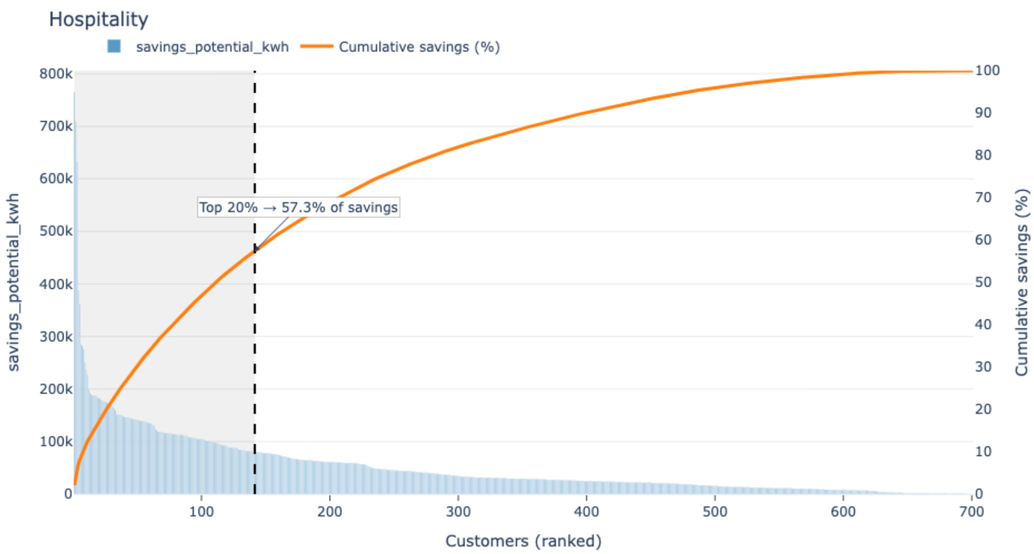

Figure 5. Pareto chart of modeled electricity savings potential for hospitality digital twins.

The **hospitality sector** displays a more gradual concentration of modeled savings potential relative to hospitals. As illustrated in Figure 5, the top 20% of hospitality digital twins account for 57.3% of total modeled savings, indicating meaningful yet less extreme concentration.

smoother decline in individual savings contributions reflects heterogeneity in building size, service level, and occupancy patterns across hotels and lodging facilities. Differences between limited-service and full-service properties, as well as variability in operating schedules and equipment configurations, drive this distribution.

Similar to the hospital sector, the most frequently occurring upgrade with the highest potential savings combined lighting, envelope, and HVAC system improvements. Hotels and lodging facilities operate across extended hours with varying occupancy patterns, resulting in consistently high baseline energy use. In this context, bundled upgrade configurations are particularly effective because they address multiple sources of energy consumption simultaneously, including internal loads, thermal losses, and system inefficiencies, leading to larger aggregate savings than isolated interventions.

While fewer hospitality facilities dominate sector-level savings in the same way as hospitals, digital twin ranking still reveals a clear opportunity to prioritize engagement with higher-potential properties rather than applying uniform outreach across the sector.

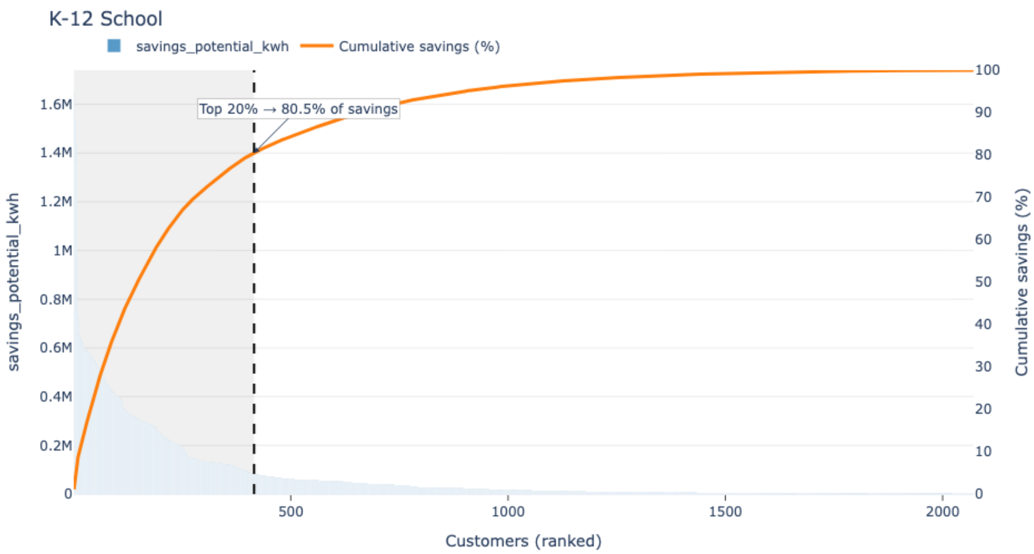

Figure 6. Pareto chart of modeled electricity savings potential for K-12 school digital twins.

Among the three sectors analyzed, the **K–12 school sector** exhibits the strongest proportional concentration of modeled savings potential. As shown in Figure 6, the top 20% of K–12 school digital twins account for 80.5% of total modeled electricity savings, substantially higher than in either hospital or hospitality sector. This pronounced concentration reflects wide variation in school size, campus configuration, and operating schedules.

In contrast to the hospital and hospitality sectors, K–12 school digital twins most frequently achieve their highest savings under the advanced rooftop unit (RTU) control upgrade. These buildings are characterized by well-defined occupancy schedules and existing HVAC infrastructure, where a significant portion of energy consumption is driven by operational inefficiencies such as unnecessary runtime outside occupied hours. Control-oriented upgrades improve system scheduling and responsiveness, enabling meaningful savings without requiring major equipment replacement, which makes them particularly effective in this sector.

Implications for Program Targeting

Table 2 quantifies the concentration of modeled electricity savings potential across sectors. In all cases, the top 20% of digital twins account for a majority of total estimated electricity savings, with the strongest concentration observed in the Hospital and K-12 School sectors.

Table 2. Concentration of modeled electricity savings potential by sector.

Sector	Top 20% digital twins account for (%) of total savings
Hospital	87.4%
Hospitality	57.3%
K-12 School	80.5%

These results indicate that modeled savings potential is not uniformly distributed across customers but instead concentrated among a subset of higher potential accounts. As a result, ranking digital twins based on estimated technical savings potential provides a basis for differentiating customers according to relative opportunity.

While the results are based on ex ante modeled savings and do not reflect realized program outcomes, they illustrate how digital twin outputs can be used to support prioritization of customers for further screening or outreach. This approach enables comparison of potential impacts across customers using a consistent, model-based framework.

FUTURE

The digital twin framework provides a structured approach for estimating pre-installation savings potential and can serve as a complementary input to existing commercial program evaluation approaches. While modeled savings do not replace formal evaluation, measurement, and verification (EM&V), they offer an additional perspective for understanding potential variation in technical opportunity across customer accounts.

As part of ongoing program implementation, digital twin outputs are being used to support targeted marketing and program outreach efforts through customer site visits, email and other customer engagement channels. Future work will focus on linking these targeted outreach efforts with observed program participation and installation data to evaluate how modeled savings potential aligns with realized outcomes. This includes assessing whether customers with higher modeled technical potential are more likely to engage with programs and implement upgrades after receiving targeted outreach.

Additional analysis may explore expansion of the archetype library to better represent under-modeled building types and operational configurations. Improving coverage of building characteristics could reduce uncertainty in savings estimates, particularly for complex or less common facility types.

Further work could also examine how digital twin-based prioritization compares with existing targeting approaches in terms of customer engagement and program outcomes. Such analysis would help characterize the applicability and limitations of digital twin–based methods in supporting commercial program design and delivery.

CONCLUSION

This paper demonstrates the application of a digital twin modeling framework that integrates Advanced Metering Infrastructure (AMI) data, Customer Information System (CIS) data, and physics-based building energy simulations to estimate account-level electricity savings potential of multiple upgrades across commercial customers.

Application of the digital twin framework to hospital, hospitality, and K–12 school customers illustrate that modeled savings potential is highly concentrated within each sector, though the structure of that concentration varies by sector. Pareto analyses show that a relatively small subset of digital twins account for a disproportionate share of total estimated savings, highlighting the uneven distribution of technical opportunity across customers.

The results suggest that digital twin outputs can be used to support prioritization of customers based on modeled savings potential, providing a structured way to compare customers based on modeled energy savings potential. However, the results are based on ex ante modeled estimates and do not reflect realized program outcomes or customer participation behavior.

Importantly, digital twin modeling is intended to complement traditional evaluation, measurement, and verification (EM&V) practices by providing additional context for program planning and customer selection. As utilities continue to expand the use of AMI data and advanced analytics, this approach offers a scalable method for estimating and comparing technical savings potential across diverse commercial customer segments.

REFERENCES

EIA (U.S. Energy Information Administration). 2025. Annual Energy Outlook 2025: Reference Case Tables. Washington, DC: U.S. Department of Energy. www.eia.gov/outlooks/aeo/tables_ref.php

Gillis, A., and A. Johnson. 2026. "How to Improve C&I Energy Efficiency Program Participation." Washington, DC: ICF. www.icf.com/insights/energy/commercial-industrial-energy-efficiency-program-participation

NREL (National Renewable Energy Laboratory). 2025. OpenStudio®: Open-Source Software for Whole-Building Energy Modeling. Golden, CO: National Renewable Energy Laboratory. https://openstudio.net/

AI-ASSISTED TOOLS: DISCLOSURES AND RESPONSIBILITY

AI-assisted tools used: Microsoft Copilot.
Used for: Drafting and revising sections of the manuscript; improving clarity and organization of the technical narratives.
Extent: Drafting and revising the manuscript.
Human review: The authors reviewed and verified all content, calculations, and citations.