Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

EVIMAS - DIGITAL TWIN-BASED ELECTRIC VEHICLE INFRASTRUCTURE MODELING AND ANALYTICS SYSTEM

Aparna Kishore, Kazi Ashik Islam, and Madhav Marathe

Dept. of Computer Science, University of Virginia, Charlottesville, VA, USA
Biocomplexity Institute, University of Virginia, Charlottesville, VA, USA

ABSTRACT

The growing shift to electric vehicles (EVs) presents significant challenges due to the complexities in spatial, temporal, and behavioral aspects of adoption and infrastructure development. To address these challenges, we present the EV Infrastructure Modeling and Analytics System (EVIMAS), a modular and extensible software system built using microservices principles. The system comprises three loosely coupled components: (*i*) a data processing pipeline that constructs a static digital model using diverse inputs, (*ii*) a modeling and simulation pipeline for simulating dynamic, multi-layered interactions, and (*iii*) an analytics pipeline that supports task execution and the analysis of results. We demonstrate the utility of the EVIMAS via three case studies. Our studies show that such analysis can be done efficiently under varying constraints and objectives, including geographic regions, analytical goals, and input configurations. EVIMAS supports fine-grained, agent-based EV simulations, facilitating the integration of new components, data, and models for EV infrastructure development.

1 INTRODUCTION

1.1 Background and Motivation

The transition to electric vehicles (EVs) significantly affects communities, influencing energy systems, infrastructure development, the economy, technology, supply chains, and human behavior. Research is rapidly advancing across these areas to tackle various challenges, particularly those related to charging infrastructure, grid adaptability, battery performance optimization, the factors driving adoption, and equity issues. As we look ahead to the future of EVs over the coming decades, it is important to note that their adoption is not uniform across different regions. This uneven distribution adds complexity to the models and simulations because it introduces variability in outcomes and requires careful calibration. While simulations can tackle individual issues, the integration of spatial complexity and uncertainties makes these models increasingly intricate.

To meet these demands, we need an architecture and the associated software system that is reusable, highly modular, and extensible. One approach is to build a pipeline-based system composed of modular and interconnected components. Informally, a pipeline refers to a series of operations on software components that take one or more inputs to produce one or more outputs, serving a specific functionality of that module (Cedeno 2019). Each pipeline can be further made up of smaller pipelines. This combination enables the creation of various loosely coupled components. However, this alone does not solve the issue of processing diverse and evolving input data. It is essential to have structured, standardized, and easily replaceable inputs to account for varying spatial and temporal resolutions. Digital twins representing EV infrastructure play a crucial role in addressing these issues. They capture detailed components of society across multiple layers, providing a coherent and layered view that supports flexible integration and dynamic simulation.

Our Contributions. This work presents the EVIMAS (Electric Vehicle Infrastructure Modeling and Analytics System) – a software system based on pipeline-based architecture to study and analyze EV infrastructure. EVIMAS consists of three primary pipelines. The first pipeline generates a digital model from various unstructured inputs to create a static representation of the real world. The simulation and modeling pipeline introduces complexity and allows for dynamic interactions between the layers, enabling the representation of a system that evolves over time. The analytics pipeline helps to formulate questions and analyze the outputs of these models, providing valuable insights. Finally, we adapt the EV Charging Simulation Framework (Islam et al. 2025) as a representative use case to demonstrate the pipeline's ability to support diverse analytical tasks and objectives. Specific contributions include the following:

- We present EVIMAS – a flexible and reusable software system for simulation, modeling, and analytics of EV Infrastructure. The three main components are represented as software pipelines, mainly (*i*) data processing pipeline (DPP) to build a static digital model, (*ii*) modeling & simulation pipeline (MSP) to study the dynamic interactions between the different layers of the digital twin, and (*iii*) an analytics pipeline (AP) includes various analytical modules designed for task execution and result analysis.
- Next, as an illustration of EVIMAS 's modular design, we integrate a high-resolution (person-level) agent-based model using a bottom-up approach to build charging load profiles and assess both residential and charging station (CS) charging loads. The framework is flexible, enabling it to operate under various conditions and integrate new advances as technology evolves.
- Finally, to illustrate the utility of EVIMAS we describe three case studies. The case studies show how EVIMAS can be useful in: (*i*) minimizing human effort and time, (*ii*) enhancing computational efficiency, and (*iii*) offering flexibility and extensibility for improved analysis and decision-making. The first case study demonstrates how the pipeline facilitates dataset substitution across regions in the DPP and allows seamless replacement of EV models in the MSP, enabling the AP to generate insights and guide modifications to EV charging infrastructure that are specific to each region. The second case study utilizes different modules within the AP to assign various tasks, allowing us to understand the simulation outcomes at different resolutions. The third case study explores the pipeline's flexibility in integrating new datasets within the DPP to support a new charging station placement strategy.

Policy insights. From a policy perspective, the three case studies emphasize the need for (*i*) expanding and upgrading CS infrastructure, (*ii*) tailored strategies to shift charging demand from peak to off-peak hours, and (*iii*) equitable access to charging facilities in both urban and rural areas. The first case study identifies counties where charging infrastructure is insufficient to serve the majority of EV users, helping prioritize new station placements. The second case study highlights the temporal patterns of charging peaks and identifies locations requiring infrastructure expansion at a finer resolution. The third case study demonstrates how strategic placement of new charging stations can reduce access violations and improve overall system performance.

Paper organization. Section 2 contains related work. In Section 3, we introduce our reusable and extensible pipeline-based architecture for EVIMAS. In Section 4, we present our pipeline-based EV charging Simulation Framework forming the three-layer digital twin. Section 5 provides the performance evaluation in terms of scalability and flexibility of the framework and discusses three case studies and a summary in Section 6.

2 RELATED WORK

There is a significant amount of literature on EV simulation, covering various aspects such as EV charging and load management (Ni and Lo 2020; Yan et al. 2020), the placement of EV charging stations (EVCS) (Wu et al. 2024; Xi et al. 2013), interactions with the power grid (Xiang et al. 2019), and user behavior for charging (Pagani et al. 2019; Liao et al. 2023). For research involving EV charging, the load is estimated using historical data from specific regions (Liao et al. 2023) or national surveys (Xiang et al. 2019). The parameters used in EV simulation models are either deterministic (Lee et al. 2019) or probabilistic in nature (Ul-Haq et al. 2017; Ni and Lo 2020).

Recent studies have shown potential for the use of digital twins in various domains of EV infrastructure, such as battery diagnostics (Eaty and Bagade 2023), charging station scheduling and management (Francisco et al. 2023), behavior and mobility modeling (Zhang et al. 2019), and smart cognitive charging stations (Yu et al. 2021). However, these models are often developed catering to specific applications. They are not developed as multi-layered systems, which limits their ability to integrate additional components, thus lacking extensibility. Consequently, these models are not reusable, limiting their effectiveness in EV infrastructure planning and operations.

In contrast, research in domains such as manufacturing (Redelinghuys et al. 2020) and healthcare (Noeikham et al. 2024) has emphasized the importance of designing multi-layered digital twins that are capable of evolution and modular integration. These architectures support scalability and adaptability, which are underexplored in EV infrastructure. Furthermore, studies across diverse domains have highlighted the benefits of pipeline-based architectures. These include building energy modeling (Khalilnejad et al. 2020), bioinformatics workflows (Di Tommaso et al. 2017), and machine learning pipeline orchestration (Bisong 2019). While these coexist independently and have advantages, there is limited research on integrating digital twins, simulation models, and analytics pipelines to build an evolving, reusable, modular EV infrastructure.

3 PIPELINE FOR EVIMAS

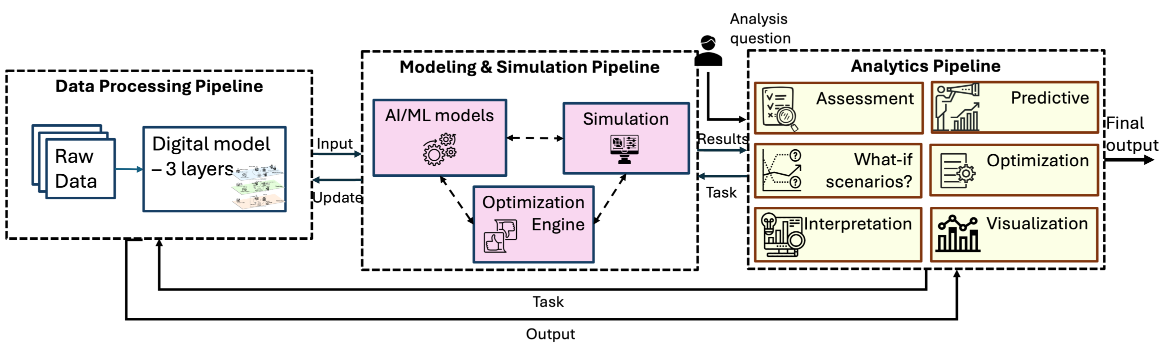


Figure 1: EVIMAS comprises three pipelines: (*i*) The Data Processing Pipeline, which ingests and processes raw data across various layers of the digital model. (*ii*) The Modeling and simulation Pipeline utilizes input from the Data Processing Pipeline to model the interactions between the layers and create real-world dynamics, forming a digital twin. (*iii*) The Analytics Pipeline is responsible for designing tasks and interacting with the other two pipelines to develop final results based on those tasks.

Our pipeline for EV infrastructure is inspired by the design of modular pipelines for residential energy modeling (Thorve et al. 2022). This design takes inspiration from two concepts: (*i*) Microservice-oriented architecture (MSA) (Thönes 2015; Richards 2015) that decomposes the system into several loosely coupled, modular, and scalable components, and (*ii*) Pipes and Filters (Bass et al. 2021), where each processing stage acts as a filter that communicates through pipes. These interconnected designing approaches facilitate the development of reusable and scalable independent modules for various downstream tasks. EVIMAS is mainly composed of three pipeline modules: (*i*) Data Processing Pipeline (DPP), (*ii*) Modeling & Simulation Pipeline (MSP), and (*iii*) Analytics Pipeline (AP) as shown in Figure 1.

3.1 Data Processing Pipeline

The DPP takes raw input from various discrete data sources, including open-source datasets and surveys designed for different tasks and applications. These datasets are processed, combined, and modeled to create verified and usable data. In addition, data augmentation can be performed in this layer, where multiple data are fused together to produce verified and usable data. The data can be forwarded to the MSP for additional modeling and then subsequently passed to the DPP as input for the next tasks. Next, these data are added as components of the digital model in the EV infrastructure, as shown in Figure 2.

The digital model is a static model made up of three layers: the infrastructure layer, the people layer, and the vehicle layer. The infrastructure layer includes all components of the EV infrastructure ecosystem. This encompasses the locations of residences, various activity sites, residential chargers (including their types, such as L1 and L2, and their specific locations), rooftop PV panels, batteries, and their locations, as well as the locations of charging stations and the number and types of chargers available at those stations. Additionally, it encompasses the power grid along with its supporting infrastructure and its connections to the homes. The people layer includes the individuals, their demographics, EV and PV adoption, charging preferences, battery states (if EV adopters), and activity patterns. For instance, an EV adopter's daily activity sequence might start with them leaving home at 8:00 AM to go to work, arriving at 8:25 AM. After finishing work at 5:00 PM, they head to the supermarket before returning home at 6:00 PM, departing from the supermarket at 5:45 PM. When this activity sequence is combined with the user's charging preferences and battery status, the user can choose to charge their vehicle while at work, near the supermarket, or wait until they return home to charge based on the state of charge (SoC) of the battery. The third layer is the vehicle layer, which encompasses the electric vehicle itself and its specific parameters, including SoC, battery capacity, and energy consumption. It also considers specific car trips resulting from an individual's commuting schedule. The interconnection between these layers, which models real-world dynamics, results in the formation of a digital twin.

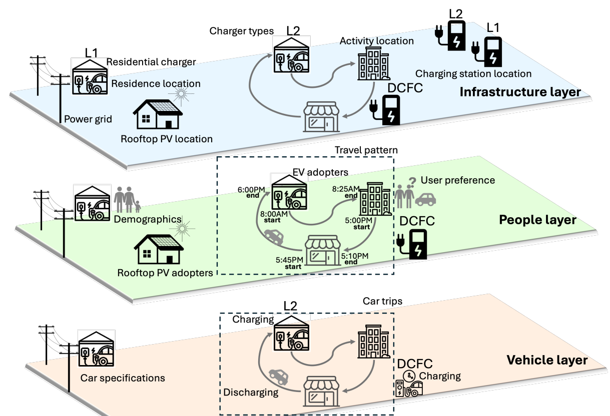


Figure 2: The static digital model is composed of three interconnected layers: the infrastructure layer, the people layer, and the vehicle layer. The infrastructure layer consists of EV infrastructure components and their respective locations. The people layer includes individuals and relevant personal details. The vehicle layer contains EVs, their specifications, and trip data.

The output from the DPP is verified and usable data structured across the three layers of the EV infrastructure digital model. They are continuously updated through interactions with the MSP and AP, forming the digital twin.

3.2 Modeling & Simulation Pipeline

The input data from different layers of the DPP is fed to the MSP. The MSP consists of different simulation and modeling techniques. The AI/ML models include the model learning from the input given to them. These models can couple with the simulation model, where the simulation model can use the input from the AI/ML model, or the simulation can be used to recalibrate the AI/ML model as a part of the feedback loop. For example, historical data fed to the AI/ML module can predict the EV charging energy consumption per session, which can be utilized in the simulation model to develop charging profiles in the entire community.

The Simulation module can be of any type of simulation such as agent-based, micro-simulations, discrete-event based simulation, or their combinations, simulation tools such as MATSim (Horni et al. 2016), AnyLogic (Borshchev and Grigoryev 2013) and so on. The optimization engine works together with the simulation module to recalibrate its input iteratively.

The MSP's output consists of write operations and is modeled output data based on the AP task. The MSP's results can be sent to the DPP to update its inputs or forwarded to the AP for further analysis, interpretation, or visualization.

3.3 Analytics Pipeline

The AP is responsible for developing scenarios or interventions, which can then be assigned as tasks to other pipelines. The AP takes the analysis question from the user. This pipeline includes several modules, such as assessment for detailed analysis and summary statistics, predictive forecasting tasks, what-if scenarios for creating intervention scenarios, and optimization for enhancing tasks with specific objectives and constraints. Based on the user's requirements, the AP communicates with the DPP for the input data. Additionally, this module can interpret and visualize results received from the MSP. The AP can also assign tasks to the DPP to update the inputs or to the MSP for iterative task refinement.

The output from the AP can be directed to either DPP or MSP. Alternatively, it may result in a final output that involves interpreted data or a visualization.

4 EVIMAS FOR EV CHARGING INFRASTRUCTURE SIMULATION FRAMEWORK

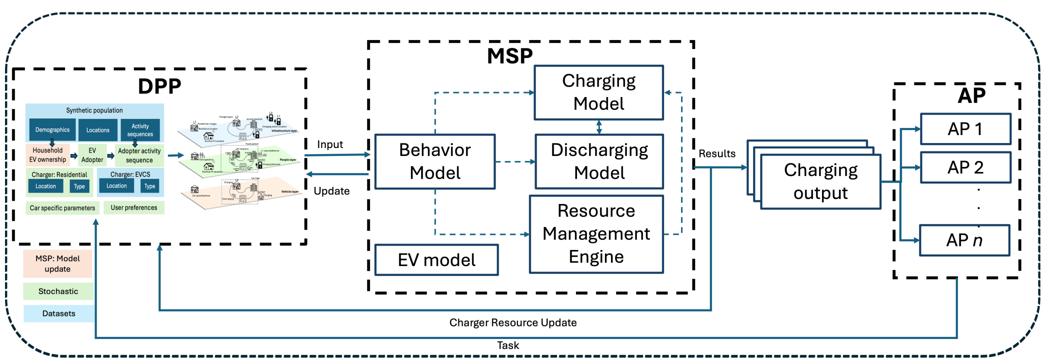


Figure 3: The figure provides an overview of the three pipeline interactions used for modeling and generating synthetic EV charging output using EVIMAS framework. The Analytics Pipeline (AP) assigns tasks, the Data Processing Pipeline (DPP) transforms and processes inputs, and the Modeling and Simulation Pipeline (MSP) carries out the modeling and simulation tasks.

The EV Simulation Framework (Islam et al. 2025), utilizing EVIMAS, is a modular, scalable, and flexible tool designed to create a digital twin of EV charging infrastructure as shown in Figure 3. This framework assists in modeling and gaining insights into various downstream tasks, such as analyzing the impact of policies on EV adoption and user behavior, understanding trends in EV adoption, assessing current limitations of the infrastructure, and examining energy consumption across different levels.

The framework employs synthetic population data to model agent-based simulations. Based on the analysis question described in the case study, the AP updates the task to the DPP. The raw datasets in the DPP include the synthetic population, as referenced in (Adiga et al. 2015), and the EVCS dataset (Deptartment of Energy 2023). The synthetic population data encompasses households, their demographic information, activity sequences, and geographical locations. Additionally, this dataset is enhanced with information on solar adoption, as detailed in existing literature (Kishore et al. 2024). Additionally, the EV model within the MSP identifies EV adopters and their home chargers. The stochastic inputs concerning user preferences and specific car parameters are also included as inputs. The digital model is updated for the next modeling task. The MSP takes in various processed inputs from the DPP's digital model. The behavior module processes each agent's respective inputs based on their activity sequence. The behavior module includes various types of behavior functions that an EV agent can exhibit, such as home charging, convenience charging (charging when a station is available nearby), and plan-ahead charging (planning ahead of time) (Liao et al. 2023).

Depending on the charging decision, the EV travels to the chosen CS. The resource management engine allocates the charger to the user based on their preferences, wait times, and availability. If a charger is successfully allocated, the EV begins charging based on the charger type and duration and updates its SoC. Whenever the vehicle travels from one location to another, the discharging module activates, updating the SoC based on vehicle parameters, the distance traveled, and other environmental conditions. Each of these events is logged in the system's output, including any unsuccessful attempts by the EV agent to charge, along with the corresponding charging output. The charger resources are updated in the DPP. The MSP forwards the charging input to the AP for interpretation, visualization, or subsequent tasks.

5 PERFORMANCE EVALUATION AND CASE STUDIES

Table 1: Table for performance evaluation and case studies. The scope column denotes the area's resolution. The activity duration can represent either a day-long or week-long activity for the simulation. Population refers to the total number of households included in the analysis. The EV adoption column shows the total number of adopters taken into account.

Location	Scope	Activity Duration	Population	EV adoption	Simulation time (s)
Tazewell	County	Day	17,278	68	3.51
Richmond	City	Day	89,153	2071	35
Cumberland	County	Week	118,413	1,170	129.69
Maine	State	Week	552,335	8,102	365.97
Fairfax	County	Day	392,061	41,901	434.29
Virginia	State	Day	3,094,255	154,492	2310.17

We evaluated the scalability and flexibility of the framework by conducting multiple tests at different resolutions using various input sets in the DPP. We utilized different EV models, including machine learning-based approaches and weighted random sampling, in the MSP to identify EV adopters and residential charging locations. The locations used for performance evaluation and case studies are detailed in Table 1. Evaluations were carried out at the city, county, and state levels, with activity sequences that spanned both a day and a week. A state with high EV adoption was able to complete a day-long simulation in just 39 minutes.

Next, we conducted three case studies to highlight the significance of the specified pipeline and the results based on the objectives. The first case study illustrates the substitution of data sources within the DPP for a what-if scenario task from AP. The task aims to analyze the impact of increased EV adoption on the charging infrastructure. The second case study involves an assessment task that evaluates the energy and equity dimensions of EV charging infrastructure, highlighting the flexibility and computational efficiency of the pipeline. The third case study utilizes the what-if scenario module in the AP to assign a task, highlighting the ease of adding a new dataset in DPP. Based on the violations identified in the first iteration, this case study necessitates additional infrastructure placement for charging.

5.1 Case Study 1: Data Substitution: Effect of Increased EV Adoption

This study aims to understand the impacts of increased EV Adoption on total charging events and types of violations in Maine. Transportation yields 50% of the carbon emission from the state. Additionally, due to the colder climate in this region, EV batteries deplete more quickly, highlighting the need for a robust charging infrastructure as EV adoption increases. Our analysis utilizes EV adoption predictions for

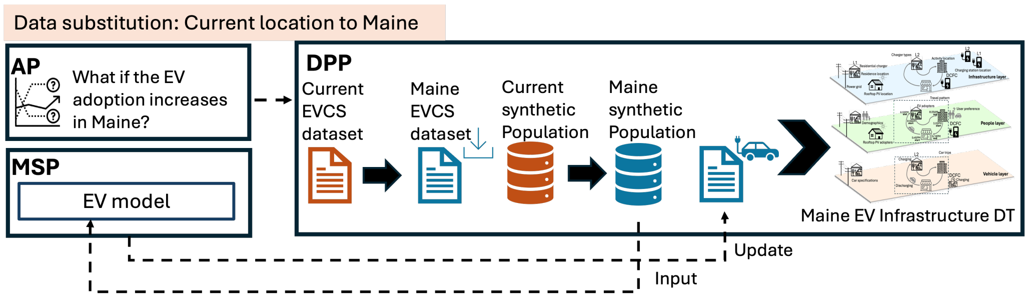
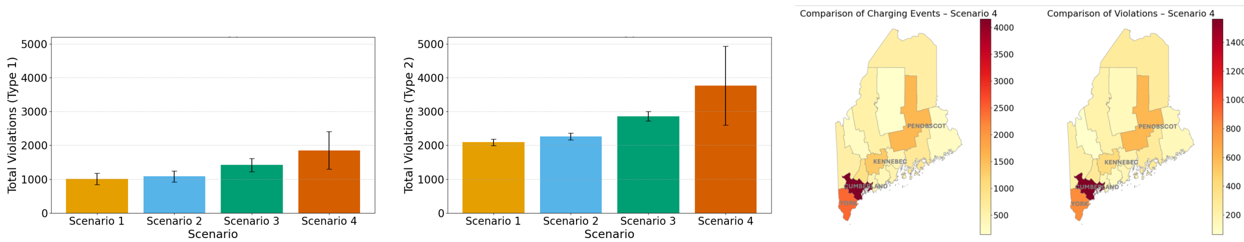


Figure 4: Data substitution in DPP of EVIMAS. The current EVCS dataset is replaced with the Maine EVCS dataset. The current synthetic population is replaced with Maine's, and the EV model predicts the EV adopters for the Maine population.

2026 and 2030 found in existing literature (Efficiency Maine Trust 2024). The scenarios are based on household income and include the Efficiency Maine Rebates, along with Federal Tax Credits for EVs. For the year 2026, we examine three scenarios of EV adoption: *Scenario 1*, assumes there are no rebates for EV adoption. *Scenario 2*, includes rebates only for low-to-medium income (LMI) households for EV adoption until 2029. *Scenario 3*, which reflects the current rebate system (Status Quo), provides rebates for all income groups until 2028, with lower-income households receiving higher rebates compared to moderate and high-income households, which receive the least. Additionally, we present *Scenario 4*, which explores the effects of the Status Quo rebate system on increased EV adoption for the year 2030. The projected EV adoption numbers for each scenario are as follows: 5,916, 6,468, 8,102, and 14,975.

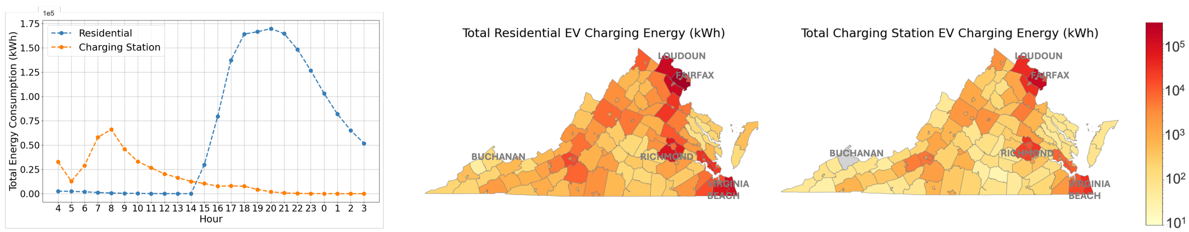
The raw data in the DPP is updated from the current population data to reflect the Maine synthetic population. Additionally, the existing EVCS infrastructure data is refreshed to represent the Maine EVCS infrastructure. The DPP interacts with the MSP to update the EV adopters using the EV model. The input for this model includes the Maine synthetic database and the count of EV adopters (Efficiency Maine Trust 2024). The adoption of EVs in Maine is modeled by assigning probabilities to non-adopters based on the number of vehicles in a household, income level, and ownership of PV systems. A weighted scoring function is then applied and normalized to reflect the relative likelihood of adoption. The updated dataset is then utilized to create a digital model of Maine's EV infrastructure, as illustrated in Figure 4. The digital model transmits the input to the simulation module of the MSP in the EV charging infrastructure simulation framework. The simulation module is implemented using the SimPy discrete-event simulation framework (Matloff 2008), which models the dynamic allocation of EVs to chargers based on real-world constraints such as station capacity, arrival times, and dwell times. Within the MSP, each travel sequence is processed to determine an EV user's charging behavior. Based on user preferences, car-specific parameters (e.g., battery size, SoC), and charger availability at the location, a specific behavior (e.g., Home, Opportunistic, Plan-ahead) is selected from the behavior model. The charging or discharging model is then triggered accordingly, and the SoC is updated after each activity. The resulting charging output is then analyzed and visualized using the interpretation and visualization modules in the AP.

For the analysis, the inability to charge due to a lack of infrastructure is categorized as a violation. *Type 1 Violation* occurs when a CS is unavailable, preventing the user from charging their vehicle. *Type 2 Violation* happens when the user cannot charge their vehicle due to longer wait times than their personal threshold in the existing CS. The results indicate that the current infrastructure exhibits Type 1 and Type 2 violations across all scenarios as the number of EV adoptions increases, as illustrated in Figures 5a and 5b. These violations are primarily observed in Cumberland County, followed by York and Penobscot counties, as shown in Figure 5c. The case study highlights the pipeline's flexibility, robustness, and its ability to support its extensibility across different geographic regions. The analysis emphasizes the need for new CSs and the upgrading of existing CS infrastructures in specific regions to accommodate the growing number of EV adopters.



(a) Type 1 Violations (b) Type 2 Violations (c) Number of Charging Events and Violations

Figure 5: Comparison of Type 1, Type 2 violations and Charging Events for Varying EV Adoption in Maine. Scenario 1 is EV adoption prediction in Maine for the year 2026 with no rebates; scenario 2 corresponds to the year 2026 with only LMI rebates; scenario 3 corresponds to the year 2026 with Status Quo rebates, and scenario 4 corresponds to the year 2030 with Status Quo rebates. Figures (a) and (b) show the respective violation counts. Figure (c) displays the county-wise heatmap distribution of average charging events and violations. All evaluations are based on 30 runs for the state of Maine over a week-long activity simulation.



(a) Hourly Energy Consumption (b) County-Wise Energy Comparison

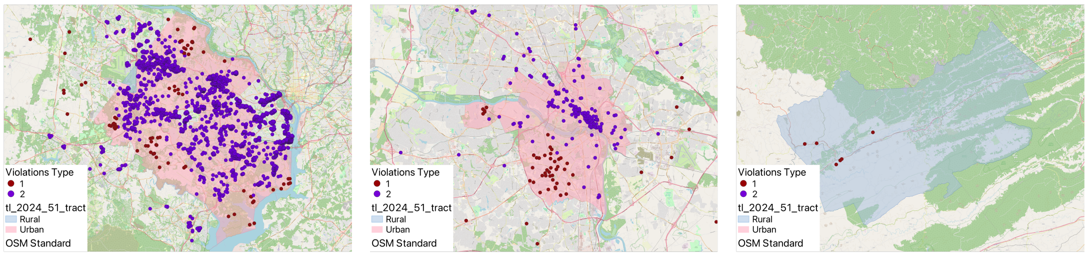
Figure 6: Energy Consumption at Residential and CSs due to charging. Figure (a) shows a line plot with hour in the x-axis and total energy consumption in kWh in y-axis. Figure (b) provides a county-wise comparison of total energy consumption from both residential and CS charging. Both figures are based on 30 evaluation runs conducted for the state of Virginia, focusing on a single day's activity.

5.2 Case Study 2: Energy and Equity Dimensions of EV Charging Infrastructure

This case study has two main objectives: first, to highlight the pipeline's flexibility and computational efficiency, and second, to examine energy consumption and equity dimensions at various resolutions in Virginia. This case study is divided into two parts. First, we aim to understand the energy consumption patterns in Virginia using the task assigned by the AP assessment module. Next, we will examine different regions in Virginia at various resolutions and analyze the violations and their connections to socio-economic factors. To achieve this, we will need to replace data in the DPP, similar to what was done in case study 1, and select regions at various spatial resolutions. The EV adoption model for this case study utilizes transfer learning to draw insights from the synthetic solar adoption data modeled by Kishore et al. (Kishore et al. 2024) in order to predict EV adoption. Following this, we estimate the number of EV adopters for the year 2026 using a logistic growth function. To identify these adopters, we use a weighted probabilistic sampling method based on income levels and the number of vehicles owned.

5.2.1 Analysis on Energy Consumption For EV Charging

First, we provide insights into when and where energy demand peaks due to EV charging at both residential and CSs, considering different behavior patterns of EV users. The data sources in DPP have been updated to reflect Virginia's synthetic population and its EVCS infrastructure, synthesizing EV infrastructure digital



(a) Fairfax County (b) Richmond City (c) Tazewell County

Figure 7: Types of violations in the choropleth map of Fairfax County, Richmond City, and Tazewell County, along with urban-rural classifications. Figures (a) and (b) show predominantly urban areas (shaded pink), while Figure (c) illustrates rural areas (shaded blue). Urban regions are experiencing both types of violations. Tazewell County has only Type 1 violations.

model of Virginia. The hourly energy consumption patterns at CSs and residential areas are illustrated in Figure 6a. Energy consumption at the CSs peaks during the day, while residential energy consumption increases at night. This rise in residential energy use begins around 3 PM when people return home from work. The highest energy consumption levels are primarily found in urban areas of Virginia, as shown in Figure 6b. Additionally, residential charging across various counties contributes to the overall high energy consumption. It is also noteworthy that Buchanan County lacks a CS, even though residential charging is still present in the area.

5.2.2 Socio-economic Analysis of Charging Infrastructure

In this part of the case study, we delve deeper into the location resolutions. We update the input sources in DPP to reflect the digital model for each region of various spatial resolutions. We compare three locations: (*i*) Fairfax County, which is primarily a densely populated urban area; (*ii*) Richmond City, the capital of the state and an independent city, also predominantly urban; and (*iii*) Tazewell County, which is mostly rural, adjacent to Buchanan County, and lacks essential CS infrastructure.

The charging outputs are analyzed using AP's visualization module. In the urban areas depicted in Figure 7, Type 2 violations (indicated in violet) outnumbered Type 1 violations (shown in brown). In Fairfax County, the clustering of Type 2 violations occurs primarily near the airport and the urban center, highlighting the need to upgrade the existing infrastructure. In contrast, Type 1 violations are mainly found in low-density residential areas (Fairfax County Department of Planning and Development 2024). In Richmond City, the violation types are spread across either side of the city, indicating a lack of CS infrastructure in one area, while the other half requires infrastructure upgrades. In the rural county of Tazewell, only Type 1 violations were observed. Although these violations are fewer in number, they signify inadequate CS infrastructure in the region.

Case study 2 highlighted the pipeline's flexibility, computational efficiency, and ease of switching between various resolutions and tasks. The initial analysis showed charging peaks during off-peak and peak hours, demonstrating the need for tailored recommendations for load management (Wu et al. 2024). Urban areas exhibit a mix of gaps in EV charging infrastructure, with Type 2 violations clustering near high-demand zones. In contrast, rural areas show limited but critical Type 1 violations, indicating insufficient coverage.

5.3 Case Study 3: Data Addition: Designing New Charging Stations

The objective of this case study is to demonstrate the extensibility of the pipeline and design new CS placements based on the current infrastructure violations, as shown in Figure 8. We use the state of Virginia, which is considered as "Current". To analyze the impact of building new CSs to support EV adopters, we

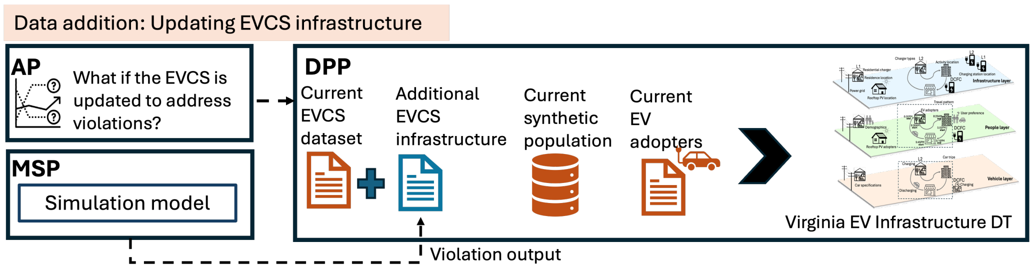
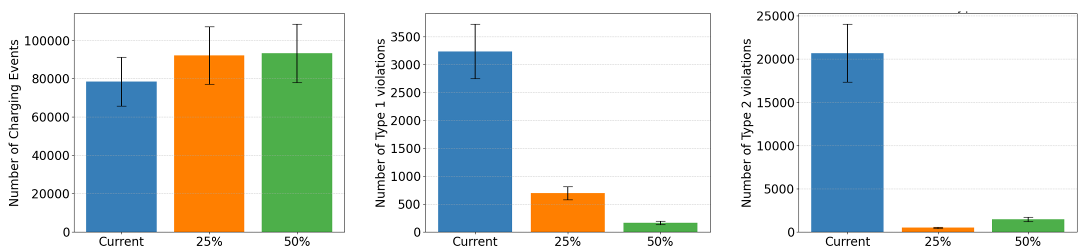


Figure 8: Data addition in DPP of EVIMAS. Additional EVCS infrastructure, based on the violation output, is added to the current EVCS dataset in DPP.



(a) Charging Events (b) Type 1 violations (c) Type 2 violations

Figure 9: Comparison of charging events and violations between the current CS infrastructure and the updated infrastructures. The x-axis bar plots are the current infrastructure, infrastructure built to resolve 25% of Type 1 violations, and infrastructure built to resolve 50% of Type 1 violations. Updated infrastructure in all three figures performs better than the current infrastructure.

conduct two sets of experiments: (*i*) New infrastructure to address 25% of Type 1 violations, and (*ii*) new infrastructure to reduce 50% of Type 1 violations based on the evaluation runs on the existing infrastructure. Type 2 violations are used to improve the chargers within the existing infrastructure. We identified a total of 2,887 CS and 5,909 CS, which includes 1,825 existing CS.

The number and type of chargers were determined based on the violations and the duration of the activities. DCFCs are allocated for sessions lasting 30 minutes or less, L2 chargers for sessions between 30 minutes and 8 hours, and L1 chargers for sessions longer than 8 hours. Approximately 10–20% of total sessions are allocated to chargers, with a maximum of 20 chargers allowed per location. The allocation of chargers is based on the ratio of each charger type to the total number of chargers. The existing infrastructure is updated in DPP to meet the predicted energy consumption due to Type 2 violations.

The increase in infrastructure by 58% and 223% led to a rise in charging events of 17% and 19%, as illustrated in Figure 9a. Additionally, this expansion resulted in a significant reduction of Type 1 violations by 79% and 95% (Figure 9b) and Type 2 violations by 98% and 93% (Figure 9c). However, it is important to note that the scenario with a 223% increase in the number of CS saw an increase in Type 2 violations compared to the previous scenario with fewer CS. One reason for this shift is that some Type 1 violations were converted into Type 2 violations, highlighting the importance of further optimization.

This case study demonstrates the pipeline's modularity and extensibility. The analysis addressed the lack of charging infrastructure by strategically placing CSs to meet the increasing demand for EVs.

6 SUMMARY AND FUTURE WORK

We introduced a pipeline-based system, called EVIMAS, for EV infrastructure consisting of three loosely coupled data construction, modeling, and analytics pipelines. EVIMAS is designed to be modular and loosely coupled, allowing for easy extensions to accommodate various applications beyond EV infrastructure planning and expansion. For example, we can integrate rooftop solar energy through a PV profile generation module (Kishore et al. 2024) in the MSP. Doing this would allow us to study the role of renewables in future energy scenarios. As another example, we can simulate retrofitting scenarios (Kishore et al. 2023) by modifying household components in the DPP and incorporating energy demand models in the MSP. Furthermore, this pipeline can be utilized for various tasks related to EV infrastructure. For example, it can offer tailored recommendations for load management strategies by shifting charging demand from peak hours to off-peak hours (Wu et al. 2024). This can be achieved by adjusting the DPP to incorporate pricing signal data and modifying the MSP to model charging behavior that responds to pricing fluctuations. A key feature of EVIMAS is its ability to accommodate different spatial and temporal resolutions while considering EV adopters' demographics, behaviors, preferences, infrastructure, and vehicle-specific parameters. The MSP pipeline models and simulates interactions across three system layers based on the task from AP, forming the core of a digital twin representation. We adapted a framework for EV charging infrastructure utilizing these pipelines (Islam et al. 2025). Its scalability and modularity are highlighted through performance evaluations and case studies. Three case studies demonstrated the pipeline's adaptability in ease of modifying the datasets, providing insights on the EV charging infrastructure, and integrating new data to improve strategies for CS placement. Currently, EVIMAS lacks a formal verification and validation (V&V) pipeline, which restricts our ability to evaluate the simulation results rigorously. Future efforts will focus on addressing these limitations by (*i*) developing a robust V&V framework and (*ii*) expanding the modular architecture to address a broader range of applications, thus creating a comprehensive digital twin for energy systems.

ACKNOWLEDGEMENTS

This work is partially supported by the University of Virginia Strategic Investment Funds SIF160, SIF186, and the Contagion Science P2PE fund SIF176A.

REFERENCES

Adiga, A., A. Agashe, S. Arifuzzaman, C. L. Barrett, R. J. Beckman, K. R. Bisset, *et al*. 2015. "Generating a Synthetic Population of the United States". Technical Report NDSSL 15-009, Network Dynamics and Simulation Science Laboratory.

Bass, L., P. Clements, and R. Kazman. 2021. *Software Architecture in Practice*. Addison-Wesley Professional.

Bisong, E. 2019. "Kubeflow and Kubeflow Pipelines". In *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*, 671–685. Springer.

Borshchev, A., and I. Grigoryev. 2013. "AnyLogic 7: New Release Presentation". In *Proceedings of the Winter Simulation Conference (WSC)*, 3625–3626 https://doi.org/10.5555/2675983.2675980.

Cedeno, V. I. 2019. *Pipelines for Computational Social Science Experiments and Model Building*. Ph. D. thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. accessed 11th April 2025.

Deptartment of Energy 2023. "Alternative Fuels Data Center". https://afdc.energy.gov/fuels/electricity-locations#/find/nearest?fuel=ELEC&location=virginia&country=US. accessed 20th March 2025.

Di Tommaso, P., M. Chatzou, E. W. Floden, P. P. Barja, E. Palumbo, and C. Notredame. 2017. "Nextflow Enables Reproducible Computational Workflows". *Nature Biotechnology* 35(4):316–319.

Eaty, N. D. K. M., and P. Bagade. 2023. "Digital twin for Electric Vehicle Battery Management With Incremental Learning". *Expert Systems with Applications* 229:120444.

Efficiency Maine Trust 2024. "Electric Vehicle Market Assessment: Historical and Forecasted Adoption and Influence of Rebates". Technical Report TPVI Appendix L3, Efficiency Maine Trust.

Fairfax County Department of Planning and Development 2024. "Comprehensive Plan – Special Planning Areas". https://www.fairfaxcounty.gov/planning-development/comprehensive-plan/special-planning-areas. accessed 26th March 2025.

Francisco, A. M., J. Monteiro, and P. J. Cardoso. 2023. "A Digital Twin of Charging Stations for Fleets of Electric Vehicles". *IEEE Access* 11:125664–125683.

Horni, A., K. Nagel, and K. W. Axhausen. 2016. "Introducing MATSim". In *The Multi-Agent Transport Simulation MATSim*, 3–7. Ubiquity Press.

Islam, K. A., A. Kishore, R. Meyur, S. Thorve, D. Q. Chen, H. V. Poor *et al*. 2025. "CHARGE-MAP: An Integrated Framework to Study Multi-criteria EV Charging Infrastructure Expansion Problem". Technical Report BI-2025-8, Biocomplexity Institute, University of Virginia.

Khalilnejad, A., A. M. Karimi, S. Kamath, R. Haddadian, R. H. French, and A. R. Abramson. 2020. "Automated Pipeline Framework for Processing of Large-Scale Building Energy Time Series Data". *PloS one* 15(12):e0240461.

Kishore, A., S. Thorve, and M. Marathe. 2023. "Budget-Constrained Optimal and Equitable Retrofitting Problems for Achieving Energy Efficiency". In *Proceedings of the 14th ACM International Conference on Future Energy Systems*, 473–484.

Kishore, A., S. Thorve, and M. Marathe. 2024. "A Generative AI Technique for Synthesizing a Digital Twin for U.S. Residential Solar Adoption and Generation". *arXiv preprint arXiv:2410.08098*.

Lee, Z., D. Johansson, and S. H. Low. 2019. "ACN-Sim: An Open-Source Simulator for Data-Driven Electric Vehicle Charging Research". In *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, 411–412.

Liao, Y., Ç. Tozluoğlu, F. Sprei, S. Yeh, and S. Dhamal. 2023. "Impacts of Charging Behavior on BEV Charging Infrastructure Needs and Energy Use". *Transportation Research Part D: Transport and Environment* 116:103645.

Matloff, N. 2008. "Introduction to Discrete-Event Simulation and the Simpy Language". *Davis, CA. Dept of Computer Science. University of California at Davis. Retrieved on August* 2(2009):1–33.

Ni, X., and K. L. Lo. 2020. "A Methodology to Model Daily Charging Load in the EV Charging Stations Based on Monte Carlo Simulation". In *2020 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*, 125–130. IEEE.

Noeikham, P., D. Buakum, and N. Sirivongpaisal. 2024. "Architecture Designing of Digital Twin in a Healthcare Unit". *Health Informatics Journal* 30(4):14604582241296792.

Pagani, M., W. Korosec, N. Chokani, and R. S. Abhari. 2019. "User Behaviour and Electric Vehicle Charging Infrastructure: An Agent-Based Model Assessment". *Applied Energy* 254:113680.

Redelinghuys, A., A. H. Basson, and K. Kruger. 2020. "A Six-Layer Architecture for the Digital Twin: A Manufacturing Case Study Implementation". *Journal of Intelligent Manufacturing* 31(6):1383–1402.

Richards, M. 2015. *Microservices vs. Service-Oriented Architecture*. O'Reilly Media Sebastopol.

Thönes, J. 2015. "Microservices". *IEEE Software* 32(1):116–116.

Thorve, S., A. Vullikanti, S. Swarup, H. Mortveit, and M. Marathe. 2022. "Modular and Extensible Pipelines for Residential Energy Demand Modeling and Simulation". In *2022 Winter Simulation Conference (WSC)*, 855–866 https://doi.org/10.1109/WSC57314.2022.10015339.

Ul-Haq, A., M. Azhar, Y. Mahmoud, A. Perwaiz, and E. A. Al-Ammar. 2017. "Probabilistic Modeling of Electric Vehicle Charging Pattern Associated With Residential Load for Voltage Unbalance Assessment". *Energies* 10(9):1351.

Wu, J., S. Powell, Y. Xu, R. Rajagopal, and M. C. Gonzalez. 2024. "Planning Charging Stations for 2050 to Support Flexible Electric Vehicle Demand Considering Individual Mobility Patterns". *Cell Reports Sustainability* 1(1):100006.

Xi, X., R. Sioshansi, and V. Marano. 2013. "Simulation–Optimization Model for Location of a Public Electric Vehicle Charging Infrastructure". *Transportation Research Part D: Transport and Environment* 22:60–69.

Xiang, Y., Z. Jiang, C. Gu, F. Teng, X. Wei, and Y. Wang. 2019. "Electric Vehicle Charging in Smart Grid: A Spatial-Temporal Simulation Method". *Energy* 189:116221.

Yan, J., J. Zhang, Y. Liu, G. Lv, S. Han, and I. E. G. Alfonzo. 2020. "EV Charging Load Simulation and Forecasting Considering Traffic Jam and Weather to Support the Integration of Renewables and EVs". *Renewable Energy* 159:623–641.

Yu, G., X. Ye, X. Xia, and Y. Chen. 2021. "Towards Cognitive EV Charging Stations Enabled by Digital Twin and Parallel Intelligence". In *2021 IEEE 1st International conference on Digital Twins and Parallel Intelligence (DTPI)*, 290–293. IEEE.

Zhang, T., X. Liu, Z. Luo, F. Dong, and Y. Jiang. 2019. "Time Series Behavior Modeling With Digital Twin for Internet of Vehicles". *EURASIP Journal on Wireless Communications and Networking* 2019:1–11.

AUTHOR BIOGRAPHIES

APARNA KISHORE is a PhD candidate in the Department of Computer Science at the University of Virginia. Her current research interests include AI and large-scale agent-based simulations for digital twin modeling and developing equity-aware optimized models towards sustainable and green energy. Her email address is ak8mj@virginia.edu.

KAZI ASHIK ISLAM is a PhD candidate in the Department of Computer Science at the University of Virginia. His research interests include developing computational tools for public policy and decision-making as it relates to the complex challenges in disaster management, infrastructure development, and transportation engineering. His email address is ki5hd@virginia.edu.

MADHAV MARATHE is a Professor of Computer Science at the University of Virginia and serves as Executive Director & a Distinguished Professor at BII. He is a fellow of ACM, SIAM, IEEE, and AAAS. His research interests include AI, modeling and simulation, network science and sustainability. His email address is marathe@virginia.edu.