

ElectroTempo

Methodology Note: How We Model EV Charging Demand

Why observed session data, raw location signals, and top-down estimates each fall short on their own – and what rigorous modeling actually requires

Reliable EV charging demand forecasts require three independent inputs: calibrated travel behavior, a top-down energy accounting model, and an agent-based charging simulation. No single data stream – observed charging sessions, raw cell phone pings, or telematics logs – is sufficient on its own. This note explains why each piece is necessary and where each falls short without the others.

1. The Three-Input Framework

Charging demand is not directly observable. It must be reconstructed from three complementary signals, each of which corrects for the blind spots of the others.

Data Source	What It Captures	What It Cannot Capture
Observed charging session data	Charger occupancy, session duration, connector type	kWh dispensed; latent demand from unserved drivers; behavior beyond existing supply
Raw cell phone / telematics data	GPS coordinates, stop events, approximate route corridors	Behaviorally correct trip ends; bias-corrected travel patterns; ground-truth trip generation rates
Top-down VMT / fuel economy model	Total fleet energy demand; macro control total for kWh needed in a corridor or market area	Temporal distribution of demand; individual vehicle charging events; site-level load profiles
Agent-based charging simulation	Individual vehicle arrival timing, dwell time, state-of-charge at arrival, stochastic demand peaks	Standalone – requires VMT and calibrated travel behavior as inputs

Observed charging session data measures supply, not demand. A session log records when a charger was occupied and for how long. It does not record the kWh actually dispensed unless the charger is revenue-metered (most are not). More importantly, session data only reflects **drivers who found a charger** — it is structurally blind to drivers who circled and left, charged elsewhere, or did not travel at all because charging infrastructure was unreliable. In high-utilization corridors, session data **systematically understates latent demand** by exactly the margin that matters most for sizing a new site.

Top-down VMT and fuel economy modeling provides the macro control total. By estimating total vehicle miles traveled in a market area, fleet composition, and fleet-average efficiency (MPGe or kWh/mile), we can calculate total electrical energy demand that must be served across all charging touchpoints — home, work, public L2, and DC fast charge. This figure anchors the simulation: the agent-based model's aggregate kWh output must reconcile with the VMT-derived total. Without this step, site-level forecasts are internally consistent but may be **systematically over- or under-scaled** relative to the actual vehicle population.

Agent-based simulation distributes that demand in time and space. Individual synthetic vehicles are assigned home locations, work destinations, dwell time distributions, and state-of-charge (SOC) at trip start, drawn from calibrated behavioral parameters. The simulation produces **hourly load profiles by site type** — the temporal signature that determines peak kW draw, revenue per charger, and grid interconnection requirements. This is the layer that converts a market-level energy forecast into an actionable site investment model.

2. Why Raw Cell Phone and Telematics Data Cannot Replace Calibrated Travel Behavior

The most common misconception we encounter is that large-scale raw location data — cell phone GPS pings, fleet telematics, or connected vehicle logs — is a substitute for the travel behavior modeling that underpins our demand forecasts. It is not, for three structural reasons.

Selection bias does not shrink with sample size. Raw location data over-represents people with specific device types, data-sharing settings, and usage patterns. Certain trip purposes (commuting, freight delivery) and certain population segments (younger, urban, higher-income) are consistently over-captured; others are systematically absent. A dataset with hundreds of millions of pings is not a representative sample of travel behavior — **it is a very large biased sample**. Scale amplifies precision without correcting the underlying misrepresentation.

Trip identification is a behavioral inference problem, not a threshold rule. Raw GPS data does not come labeled with trip starts and ends. Every analyst must define a stopping rule: how long must a vehicle be stationary before the current trip is considered complete? A naive cutoff — say, five minutes — will misclassify brief pauses (waiting for a passenger, stopped at a long traffic

signal) as trip ends, and will incorrectly split single-purpose journeys into multiple trips. The result is a distorted picture of trip length distribution and charging opportunity windows. Our data source applies a **behaviorally calibrated trip identification algorithm** validated against regional household travel surveys, travel demand models, and highway traffic counts — precisely because no threshold rule produces reliable trip ends on its own.

Uncalibrated trip generation cannot be used for infrastructure sizing. The first validation step in rigorous travel demand modeling is confirming that trips are being *produced* in the right locations in the right quantities — what modelers call trip generation. Our data source compares raw trip endpoints at the Census Tract level to established regional models and household surveys before any further analysis is performed. This calibration step is not optional: if trip generation is wrong, every downstream estimate of charging demand, peak load, and revenue is also wrong, regardless of how sophisticated the simulation layer is.

3. Implications for Charging Infrastructure Analysis

These methodological requirements have direct implications for how charging demand forecasts should be evaluated:

- **Demand vs. utilization:** A utilization report based on session data measures how full existing chargers are. It does not measure unmet demand. The two numbers can diverge substantially in underserved markets — which are often the highest-priority investment targets.
- **kWh vs. session-hours:** Revenue and grid impact both depend on energy throughput (kWh), not occupancy time. Session duration is a poor proxy for kWh when charging speeds, vehicle SOC at arrival, and connector mix vary across a site portfolio.
- **Latent demand requires simulation:** Identifying where demand is unserved — the core question for greenfield site selection — requires a model that can simulate travel behavior in the absence of existing infrastructure. Observed data, by definition, cannot answer this question.
- **Ground-truth calibration is non-negotiable:** Any forecast that has not been validated against regional travel surveys and highway count data should be treated as illustrative rather than investment-grade. The calibration process is the mechanism by which analyst-defined assumptions are tested against observed reality.