UC San Diego UC San Diego Previously Published Works

Title

Hybrid Machine Learning Forecasting for Online MPC of Work Place Electric Vehicle Charging

Permalink <https://escholarship.org/uc/item/4463c0cx>

Journal IEEE Transactions on Smart Grid, 15(2)

ISSN

1949-3053

Authors

McClone, Graham Ghosh, Avik Khurram, Adil [et al.](https://escholarship.org/uc/item/4463c0cx#author)

Publication Date

2024-03-01

DOI

10.1109/tsg.2023.3296014

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Hybrid Machine Learning Forecasting for Online MPC of Work Place Electric Vehicle Charging

Graham McClone, Avik Ghosh, Adil Khurram, Byron Washom and Jan Kleissl

Abstract—This work proposes a novel EV forecasting technique that predicts each EV's arrival time (AT), energy demand (ED) and plug duration (PD) over the course of a calendar day using a hybrid machine learning (ML) forecast. The ML forecasts as well as persistence forecasts are then input in a model predictive control (MPC) algorithm that minimizes the electricity costs incurred by the charging provider. The MPC with the hybrid ML forecast reduced peak loads and monthly electricity costs over a base case scenario that determined costs for uncontrolled L2 charging: Reductions in weekday mean peak load during a 30 day summer time case study were 47.0% and 3.3% from the base case to ML MPC and persistence to ML MPC, respectively. Reductions in utility costs during the summer case study were 22.0% and 1.4% from base case to ML MPC and persistence to ML MPC respectively. Results are similar for a 30 day winter case study.

Index Terms—Energy Resources, Forecasting, Learning Systems, Model Predictive Control, Neural Network Applications, Optimal Control.

NOMENCLATURE

 α Weight applied to Arrival Time vector β Weight applied to Plug Duration vector γ_{nc} Non coincident demand peak $\gamma_{\rm op}$ On-peak period demand peak δ Weight applied to Energy Demand vector Δt Time step duration ζforec Unassigned weighted sum of over-forecasted EVs ζ_{real} Unassigned weighted sum of real under-forecasted EVs η Vector of $[\alpha \beta \delta]^T$ $cost$ Matrix of squared errors y_{real} and y_{force} E Total energy dispatched to an EV g_{dc} Demand charges gec Energy charges $g_{\rm oc}$ Other charges J Daily cost for charging all EVs k Time index L Aggregated net load dispatched each time step M Linear assignment array of the *cost* matrix N Number of time steps in MPC horizon Number of time steps in MPC horizon N_{EV} Number of EVs $N_{\rm T}$ Number of time steps in forecasts P Individual net load for each EV r_{ec} Energy charges rate r_{nc} Non-coincident demand charge rate

- $\hat{\boldsymbol{x}}_{\text{force}}$ Normalized $\boldsymbol{x}_{\text{force}}$
- \hat{x}_{real} Normalized x_{real}
- y_{real} Weighted sum of real AT, PD, and ED
- y_{force} Weighted sum of forecasted AT, PD and ED
- ADED Aggregated Daily Energy Demand
- ALM Adaptive Load Management
- AT Arrival Time
- DT Departure Time
- ED Energy Demanded by each EV
- EV Electric Vehicle
- kNN k Nearest Neighbors
- L₂ Level 2 charging
- LSTM Long Short Term Memory
- MIMO Multi-input multi-output
- ML Machine Learning
- MPC Model Predictive Control
- NAPM Number of EV Arrivals Per Minute
- NL Net load dispatched at a given time step
- PD Plug Duration: Time duration each EV is plugged in
- PV Photo voltaic solar power
- RNN Recurrent Neural Network
- SOC State of Charge
- TOU Time of use

I. INTRODUCTION

A. Motivation

The number of electric vehicle (EV) charge ports in the US has steadily increased over the past few years. This growth is expected to continue due to state and national mandates for zero emission transportation [1]. Frequently, drivers utilize EVs for their commute to and from work, making their work place a convenient location to charge. If this charging were to be controlled, it could help mitigate grid operation challenges associated with frequency regulation [2], peak demand reduction [3] and shifting charging load to hours during which solar and wind generation are high [4].

A large percentage of EV charge station growth is from work place Level 2 (L2) charge stations [5]. L2 charge stations are able to distribute between 6.2-7.6 kW of maximum charging power depending on the EV charging capabilities. The University of California, San Diego (UCSD) offers work place L2 EV charging to students, faculty, and staff through a variety of different companies [6]. One of these companies, PowerFlex, specifically requests a driver's preferred energy demand (ED) in addition to a driver's plug duration (PD). Then, PowerFlex utilizes an Adaptive Load Management

G. McClone, A. Ghosh and A. Khurram are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, 92037 USA (UCSD) email:{gmcclone,avghosh,akhurram}@eng.ucsd.edu.

J. Kleissl is the Director of the Center for Energy Research in the Department of Mechanical and Aerospace Engineering at UCSD, email: jkleissl@eng.ucsd.edu.

B. Washom is the Director of Strategic Energy Initiatives at UCSD, email: bwashom@ucsd.edu.

(ALM) algorithm that ensures EV users are satisfied upon departure while reducing peak loads incurred by the utility customer [7]. ALM algorithms would benefit from a forecast technique that predicts the arrival time (AT), ED and PD to optimally schedule EVs as to to reduce peak load and costs.

B. Previous Work

There are a small number of papers that address forecasting of EV charging for short-term dispatch. Papers that forecast EV charging tend to focus on aggregate EV utilization impacts to the grid, aggregated charging demand, forecasting load or forecast specific EVs assuming perfect knowledge of which EVs will arrive as explained next.

A method for forecasting day-ahead aggregated daily energy demand (ADED) based on previous measurements is presented in [8]. In [9] the authors predicted aggregated EV load to schedule EV charging; in [10] the authors do the same but take user convenience into account by prioritizing low SOC and earlier departure times [10]. In [11], cluster analysis was used to forecast aggregate traffic patterns, charging start times, and initial battery SOC. The authors of [12] implemented an adaptive load forecasting algorithm, based on historical load data and meteorological conditions, to determine optimal scheduling using vehicle to grid (V2G) charging. The authors used predetermined initial and final SOC values to obtain their results. In [8]–[13], aggregate forecasts are considered for the purpose of understanding the aggregate impact of large number of EVs. But only forecasting individual EV characteristics allows each EV's constraints to be incorporated into the charging controller to guarantee user satisfaction.

AT, ED, and PD are the most relevant variables to forecast for scheduling EVs as these variables are sufficient to optimize dispatch schedules. The authors of [14] used machine learning models to predict PD and ED for specified EVs with perfect knowledge of each EV's AT. Assuming perfect knowledge of the number of EVs and AT is unrealistic yet forecasting AT, ED, and PD together has not been considered. These variables are important for optimizing charging schedules.

In the literature, EV error analysis is generally performed in one of two ways. (i) In aggregate level forecasting, forecasted behavior such as a load curve or ADED is compared to corresponding real quantities. In this case, the forecasted number of EVs can be different from the actual number of EVs. However, since the error analysis is performed using aggregate quantities, the difference between forecasted and actual number of EVs does not impact the analysis. (ii) In agent based forecasting, forecasted behavior of each EV is compared to the real behavior. In this case, each forecasted EV must be mapped to its corresponding real EV. In either case, there is no need to compare the number of EVs that arrive or how closely any two unknown EVs are related.

Optimization of EV scheduling for peak reduction or for providing other ancillary services has been considered extensively in literature. Specifically, EV charge scheduling can be divided into two broad categories: offline and online charging. Offline charge scheduling assumes perfect knowledge of ED or SOC, AT, and PD to optimize charge scheduling, while online charge scheduling copes with unknowns, i.e., imperfect forecast. EV charge scheduling using perfect forecasting is unrealistic as perfect knowledge of AT, ED, and PD is not available. Thus, the following review focuses only on online EV scheduling techniques.

In [7], the authors presented the online model predictive control (MPC) algorithm currently utilized by PowerFlex to perform online dispatch of EVs. A regularization term in the objective function encourages dispatch in the morning hours, to reduce charging costs during afternoon peak hours. The authors considered several MPC formulations including a time-of-use (TOU) cost structure with demand charges. However, forecast of EV behavior is not included in the MPC formulation. In [15], an online MPC optimally scheduled EV charging using a convex quadratic cost function which is formulated as the square of the net-load (summation of EV charging and building load) over the MPC horizon. The problem formulation in [15] assumes "earliest departure first to charge", to determine individual EV charging schedules. This assumption can result in individual EV dispatch power constraint violations, as an EV may be required to charge at a rate higher than the dispatch capacity of the L2 charge station. The authors of [16] presented an online MPC technique to obtain EV charge schedules that minimize total system energy cost. In [17], the authors presented an MPC technique for online charge scheduling incorporating ED uncertainty using a Markov Chain Monte Carlo technique. A multi-objective MPC smooths the load curve and minimizes costs associated with TOU energy rates.

The authors of [18] minimize the mean waiting time for EVs with a long term constraint on cost. In [19], the authors formulate a stochastic optimization problem that schedules EV charging assuming uncertain departure times in the presence of hourly time-of-use pricing tariffs. In [20], an optimal EV scheduling algorithm is demonstrated for offline global scheduling and online local scheduling to minimize the costs that customers pay.

There are gaps in the literature pertaining to effective forecasting of AT, PD, and ED of EVs for the purpose of implementing online MPC to minimize electricity costs based on realistic energy and demand charges. Specific research gaps include:

- 1) Forecasting of individual AT, ED, and PD for an unknown number of unidentified EVs has not been considered. This information is important to ensure that individual ED requirements are satisfied while determining charging schedules.
- 2) MPC based methods usually consider initial and final SOC for scheduling EV charging. This requires two variables to be forecasted, whereas MPC with ED requires forecast of only a single variable. Reducing the number of forecast variables can reduce forecast error and improve performance.

C. Present Work and Novelty

The present work proposes a novel forecasting and control scheme. First, a hybrid ML technique is developed that constructs a day-ahead forecast for individual EV AT, ED, and PD. Then, finite time horizon MPC is used to determine the optimal charging profile at each time step in real time. The MPC objective function is convex and incorporates both demand charges and TOU energy costs. The proposed MPC formulation initially generates charging profiles using forecasted AT, ED and PD. These charging profiles are then updated with the actual values of AT, ED and PD as the EVs plug in throughout the day. A modular architecture enables implementation across a variety of different charge networks independent of location or charging cost structure.

The hybrid ML technique developed in this work forecasts both the number of EVs as well as the AT, ED and PD of each EV for which both the aggregate error analysis and agent based error analysis described above are not applicable, because:

- 1) The number of forecasted EVs can vary from the number of EVs that actually arrive, creating a challenge of determining error between values at different indices within arrays of differing length.
- 2) The order in which forecasted EVs arrive may not be the same as the real order of EV arrival. For example, an EV that is forecasted to arrive at a given time with specified ED and PD may actually arrive earlier or later than expected, likely changing the EV's order within the array of daily EVs. The best way to determine error between a forecast and reality in this case is not by comparing the first EV that arrives with the first EV that is forecasted, but rather by comparing a forecasted EV that most closely matches a real EV counterpart.

To account for these difficulties, a novel error analysis technique is presented. This technique compares distances between forecasted and actual EVs AT, PD, and ED to construct a matrix of errors from which the linear assignment algorithm [21] is used to determine which forecasted EVs are closest to real EVs.

The main novelties of this paper are:

- 1) A hybrid ML forecasting method is presented that forecasts ED, AT and PD for individual EVs instead of aggregate EV load at work place charge stations.
- 2) Forecasting an undetermined number of unidentified individual EVs rather than forecasting a known number of EVs or EVs as known agents.
- 3) An error analysis technique that allows EV charging data (AT, PD, ED) of differing indices within vectors of differing length to be compared individually.

The rest of this paper is organized as follows. Section II discusses the data utilized in this study. Section III describes the utility customer cost structure. Section IV discusses the mathematical formulation for the MPC optimization problem. Section V describes the linear assignment and error analysis for individual EVs. Section VI discusses the process for forecasting AT, ED and PD and presents the novel error analysis technique for forecasted EVs. Section VII presents the results and discussion, including a sensitivity study. Section VIII is the conclusion.

TABLE I: Time of use energy charge rates: SDG&E AL-TOU tariff [24].

Rate Type	Commodity Rate (S/kWh)	Utility Distribution Company Rate (\$/kWh)
Summer On-Peak	0.11957	0.00671
Summer Off-Peak	0.10008	0.00671
Winter On-Peak	0.09955	0.00671
Winter Off-Peak	0.08835	0.00671

II. DATA

The data used in this work comes from EVs that charged with ChargePoint [22], but it is assumed that there is an ALM system in place based on the ALM developed by Power-Flex [7] [23]. To increase the sample size, the ChargePoint data is associated with EVs that charged at 24 different charge stations with a combined 48 charge ports (two per station) from two different parking plazas on the UCSD campus.

Once an EV plugs into a ChargePoint station, it begins charging and continues to do so until fully charged. Most EV charging commences at a given power setting which is reduced progressively after the battery reaches a large state of charge . Charge session data does not include a time series of dispatched power, only the AT, PD, and ED are included. For "dumb charging" it is therefore assumed that the power for each uncontrolled charging session is constant at a charge rate of 7.2kW until the ED is satisfied.

The raw data for this analysis included 23, 545 charge sessions between the dates of November 10, 2018 and March 30, 2020. Charge sessions with less than 15 min PD or ED equal to 0 kWh were removed from the raw data since they were likely accidental. It is assumed that each charge session ends prior to 00:00 h of the following calendar day. This removes overnight charging from the control algorithm and focuses on work place daytime charging. To effect this assumption, overnight charging sessions were removed from the data set. The remaining data set used in the analysis consisted of 21, 756 charge sessions.

III. COST STRUCTURE

The electricity cost incurred by the utility customer as a result of EV charging is a function of energy and demand charges. Energy charge rates (\$/kWh) consist of TOU commodity rates, and a uniform utility distribution company rate, as depicted in Table I, as obtained from San Diego Gas & Electric (SDG&E). Each year is divided into summer and winter seasons: summer lasts from June 1 through October 31, with the rest of the year being winter. Each day has an onpeak period covering the time period from 16:00 h to 21:00 h, with the others being off-peak hours.

The second major component of electricity costs are demand charges, whose rates (\$/kW) are depicted in Table II. Demand charges are applied to monthly peak loads and are of two types: non-coincident and on-peak demand charges. The noncoincident demand charge is the price owed by the utility customer due to its peak demand in a month. Added to noncoincident demand charges are on-peak demand charges which determine the price owed by the utility customer due to its peak demand during the on-peak hours of a month.

TABLE II: Seasonal and time of day demand charge rates: SDG&E AL-TOU tariff.

	Commodity Rate	Utility Distribution	
Rate Type	(S/kW)	Company Rate (\$/kW)	
On-peak summer	9.78	19.14	
On-peak winter	N/A	19.23	
Non-Coincident	N/A	24.48	

We present the MPC problem formulation in this section in which all pertinent EV charging session data is assumed known. The pertinent data for a day includes the PD, ED, and AT for all EVs charging at a plaza. The forecasting methodology for AT, PD and ED is described in the next section. The problem formulation in this work follows [25] with the SOC formulation in [25] adjusted to the ED formulation. The MPC problem is solved with a finite time horizon with control input only applied to the EVs that are plugged in at the current time step.

A. Optimization Problem

The MPC cost function J provides the total cost in dollars consisting of demand charges (g_{dc}) , energy charges (g_{ec}) and other utility specific charges (g_{oc}^{-1}) and is given by,

$$
J(k, L) = g_{dc}(k, L) + g_{ec}(k, L) + g_{oc}(k, L)
$$
 (1)

where $k \in \{1, \ldots, N\}$ is the starting time index of the MPC horizon, $N = \frac{24h}{\Delta t}$ is the final time index of the MPC horizon, Δt is the time resolution of the control action, $L(j)$ is the net load at the j-th time step and $L = (L(1), \ldots, L(N))^T$. In this work, $\Delta t = 0.25$ h (15 minutes) resulting in $N = 96$. The demand charges and energy charges are given by,

$$
g_{\rm dc}(k,L) = r_{\rm nc} \gamma_{\rm nc}(k) + r_{\rm op} \gamma_{\rm op}(k),\tag{2}
$$

$$
g_{\rm ec}(k,L) = \Delta t \sum_{j=k}^{N} r_{\rm ec}(j)L(j),\tag{3}
$$

where $r_{\text{ec}} \in \mathbb{R}^N$ is the vector of TOU rates of energy. The non-coincident demand charges are calculated from the net load L starting from the current time step k to N as follows,

$$
\gamma_{\rm nc}(k) = \max\{L(m)\}_{m=k}^N.
$$
 (4)

In (4) r_{nc} is the non-coincident demand charge rate. Similarly, the on-peak demand charge rate is r_{op} and the on-peak demand charges are computed from L but only between $T_{\text{start}} = \frac{16h}{\Delta t} + 1$ and $T_{\text{end}} = \frac{21h}{\Delta t}$ as follows,

$$
\gamma_{\rm op}(k) = \max\{L(m)\}_{m \in \mathbb{I}_{\rm op}(k)},\tag{5}
$$

where $\mathbb{I}_{op}(k)$ is the set of indices corresponding to the on-peak time given by,

$$
\mathbb{I}_{\text{op}}(k) = \begin{cases} \{\max\{k, T_{\text{start}}\}, \dots, T_{\text{end}}\}, & \text{if } k < T_{\text{end}},\\ \phi, & \text{otherwise} \end{cases} \tag{6}
$$

¹Other charges consist of the DWR bond charge (\$0.00580 \times total energy usage in a month), the City of San Diego Franchise fee (\$0.0578 \times ${r_{\text{nc}}\gamma_{\text{nc}}(k) + r_{\text{op}}\gamma_{\text{op}}(k) + \Delta t \sum_{j=k}^{N} r_{\text{ec}}(j)L(j)}$, the DWR bond franchise fee (\$0.0688×DWR bond charge), the CA State Surcharge (\$0.00030×total energy usage in a month), and the CA state regulatory charge (\$0.00058×total energy usage in a month) [25].

$$
\min_{L,P,E} J(k,L) \tag{7}
$$

$$
0 \le P_i(j) \le P_{\text{max}},\tag{8}
$$

$$
E_i(j) = E_i(j-1) + P_i(j)\Delta t
$$
 (9)

$$
L(j) = \sum_{m=1}^{N_{\text{EV}}} P_m(j), \quad \forall j \in \{k, ..., N\},
$$
 (10)

$$
P_i(j) = 0, \quad \forall j < \mathbf{AT}_i \quad \text{or} \quad j > \mathbf{DT}_i,\tag{11}
$$

$$
E_i(DT_i) = ED_i, \ \forall i \in \{1, \dots, N_{EV}\}.
$$
 (12)

Equation (8) provides limits on charging power. Equation (9) is the energy balance equation. Equation (10) states that the net load at each time step is the sum of individual loads of each EV, given by P_i where N_{EV} denotes the number of EVs that are plugged in. Equation (11) states that no EV can receive dispatched energy prior to arriving at the charge station or after departing the charge station. Departure time is expressed as the sum of AT and PD (i.e., $DT_i = AT_i + PD_i$). Equation (12) states terminal constraints.

The optimization problem minimizes the daily total electricity cost associated with EV charging. Due to the cost of energy being orders of magnitude lower than the cost of demand, and with the ability to shift energy between on-peak and off-peak periods limited by the PD and the overlap of the layover period with both on and off-peak periods, the most effective EV charging scheme will flatten the load associated with charging to reduce demand peaks. While the demand charge is only applied to the single highest monthly load value, the MPC optimization is run daily. This ensures that the daily peak load and the TOU energy charges are minimized.

In this work, the energy and demand charges in the objective function correspond to SDG&E's cost structure. However the various rates and corresponding time periods can be adjusted to any other utility's cost structure. Finally, the MPC problem is solved using CVX [26], [27], a package for solving convex programs in the MATLAB environment.

B. Assumptions

First, it is assumed that once an EV arrives, the EV user populates accurate values for PD and ED, consistent with how PowerFlex currently operates. Second, all charging operates using a 15 minute update. This creates 96 discrete time steps each calendar day during which any change in the number of EVs that are plugged in must be updated. This work assumes that an EV cannot start charging until the 15 minute time step after the EV arrives. For example, if an EV arrives at 08:05 h, the EV cannot commence charging until 08:15 h. It is also assumed that an EV's ED must be satisfied by the end of the 15 minute interval prior to its departure time. This assumption effectively shortens the PD for each EV from what actually occurred.

When the time of day reaches the AT of a forecasted EV, that EV will be removed from the remaining forecasted EVs. The scheduling algorithm will only update every time a real EVs plug in as the associated ED and PD update at that time.

Energy cannot be dispatched to forecasted EVs, only to real EVs upon their arrival.

V. ERROR ANALYSIS FOR INDIVIDUAL EVS

This work considers EV behavior on an individual basis rather than on an aggregate scale and also relaxes the assumption of prior knowledge of specific EVs and the exact number of EVs that will arrive, thereby simulating a more realistic case. However, the present process also creates challenges in identifying error between a forecasted EV and a real EV. The order in which EVs arrive in the forecast may not be the same as the order in which EVs arrive in reality; while traditional error metrics would assign large errors to an outof-order forecast, the error in the order does not necessarily determine forecast quality for the purposes of implementing MPC. Also a forecast may predict a different number of EVs than those that actually arrive. If the number of forecasted EVs is equal to or fewer than the number of real EVs, each forecasted EV should be matched with its most similar real EV counterpart with leftover real EVs accounting purely for under forecasting error. Conversely, when the forecast predicts more EVs will arrive than actually do; each real EV should be matched with the most similar forecasted EV with leftover forecasted EVs accounting purely for over forecasting error.

Algorithm 1 was developed to match forecasted EVs to real EVs. As described in Algorithm 1, each of the AT, PD and ED are first normalized based on the maximum and minimum values. These normalized quantities are used to compute a single quantity for each EV that consists of the weighted sum of AT, PD and ED. For this work the weights are 10, 1, 7 for AT, PD and ED, respectively. The squared differences between the weighted metrics of all forecasted and real EVs are used to populate a cost matrix. Entries in the cost matrix quantify the similarity between real and forecasted EVs. Then, the linear assignment algorithm [21] is used to uniquely match forecasted and real EVs.

The linear assignment determines the minimum cost of selecting one value from each row and column of the cost matrix. The output of the linear assignment is an array of matching pairs (forecasted to real) and the associated cost for the difference between each EV. If the number of forecasted EVs is greater (or less) than the number of real EVs, the cost of each of the unmatched forecasted (or real) EVs is set equal to the worst case cost of the matched EVs.

VI. FORECASTING

A. Smart Persistence Forecasting

A commonly used benchmark for forecasting is persistence forecasting. 24 hour persistence forecasting assumes that a given behavior is periodic on a daily basis, therefore the number of EVs that are charging at a given moment and their associated PD and ED will be the same as observed on the previous day. Persistence forecasts offer low computational costs, making them useful in situations where a high level of accuracy is not needed. Due to the extreme difference in work place charging behavior between weekdays, weekends, and holidays, this smart persistence forecast assumes that Algorithm 1 Process for determining the distance between forecasted EVs and real EVs for one day.

Input: Daily forecasted and real EV arrays for each EV given by $x_{\text{real}} = [AT_{\text{real}} \, PD_{\text{real}} \, ED_{\text{real}}]$ and $x_{\text{force}} =$ [$AT_{\rm force}$ $PD_{\rm force}$ $ED_{\rm force}$]. Weights: α , β , δ .

Output: Array $M(\min(\mathrm{length}(x_{\mathrm{force}}, x_{\mathrm{real}}), 2))$, which matches each forecasted EV (number in column 1) to the most similar real EV (number in column 2). Arrays ζ_{real} and ζ_{force} contain the additional unmatched EVs if the forecast either under- or over- predicted the number of EVs, respectively.

1: $\hat{\boldsymbol{x}}_{\text{real}} = \text{norm}(x_{\text{real}})$ 2: $\hat{\boldsymbol{x}}_{\text{force}} = \text{norm}(x_{\text{force}})$ 3: $\boldsymbol{\eta} = [\alpha \beta \delta]^T$ $\triangleright \alpha$, β , δ are weights 4: $\bm{y}_{\text{real}} = \hat{\bm{x}}_{\text{real}} \bm{\eta}$ 5: $y_{\text{force}} = \hat{x}_{\text{force}} \eta$ 6: for $i = 1$: length(y_{force}) do 7: **for** $j = 1$: length(y_{real}) **do** 8: cost $(i, j) = (\mathbf{y}_{\text{real}}(j) - \mathbf{y}_{\text{force}}(i))^2$ 9: end for 10: end for 11: $M = linear_assignment(cost)$ \triangleright Reference [21] 12: if length $(y_{\text{real}}) >$ length (y_{force}) then 13: $\zeta_{\text{real}} = y_{\text{real}}$ 14: $\zeta_{\text{real}}(M(:,2)) = [$; 15: else if $\text{length}(\boldsymbol{y}_{\text{force}}) > \text{length}(\boldsymbol{y}_{\text{real}})$ then 16: $\zeta_{\text{force}} = y_{\text{force}}$ 17: $\zeta_{\text{force}}(M(:, 1)) = [$; 18: end if

weekdays will persist from weekdays, weekends will persist from weekends, and holidays will persist from holidays, i.e.

$$
x_p(t) = x_{p-1}(t) \quad \forall \quad t = 1, 2, \dots, N_T,
$$
 (13)

where x is any one-dimensional variable to be forecasted (such as AT, ED or PD), t indexes the time of the day, N_T is the total number of time indices within a day for a given δt . In this work, forecasts are performed with minute resolution, therefore $N_T = 1440$. p refers to the type of day, including, weekdays, weekends, and holidays. For example, charge behavior on Mondays persist from Fridays, Sundays persist from Saturdays, and a holiday persists from the previous holiday.

B. Machine Learning Forecast Overview

The flowchart in figure 1 depicts the process for completing the hybrid forecast. An overview of the process is given in this section, while each block on the flowchart is described in more detail in the following sections including the error analysis. In each block, the choice of the particular forecast methodology was made after comparing its performance with other forecast methodologies as described next. An example of the training data used to forecast for January 10, 2020 is described in Table III. The observed EV AT, PD and ED on January 10, 2020 are then used in the forecast for the next day.

Forecasting AT, PD and ED is a multi-step process. Initially, a forecast for ADED is performed using Matlab's TreeBagger function. TreeBagger was selected to forecast ADED because

Fig. 1: Flowchart depicting the hybrid forecast method for forecasting individual EVs.

TABLE III: Example of respective forecast methods and associated training data with Jan 10, 2020 as the testing day. Forecast methods in (parenthesis) are used for comparison.

it performed better than LSTM and Persistence forecasting techniques. The ADED value becomes a feature for a Long Short Term Memory (LSTM) algorithm that forecasts NAPM from a given starting point to the end of the calendar day. LSTM was selected to forecast NAPM because it performed better than TreeBagger and Persistence forecasting techniques. The LSTM output is a vector of real numbers associated with the cumulative number of EVs that have plugged in at each minute during the calendar day. These values are then rounded to the nearest integer as it is not possible to have a non integer number of EVs, and corrected to be monotonically increasing.

The output of the LSTM forecast is input to a k-nearest neighbor (kNN) algorithm that compares the LSTM output with the NAPM from previous days in the data set. The kNN algorithm determines which previous day best matches the forecasted NAPM and populates the forecasted AT, PD, and ED for each EV forecasted to arrive that day. The result from the kNN algorithm is the forecast utilized by the MPC. The kNN algorithm was selected because it finds the NAPM vector in the past that most closely matches the LSTM forecasted NAPM in the previous step and then provides realistic PD and ED values for all EVs that are expected to arrive.

C. TreeBagger for ADED

Matlab's TreeBagger function was used to forecast ADED [28]. The TreeBagger algorithm trains on two features: ADED from all prior days and the corresponding number of EVs that charged on all prior days. The training data begins on November 10, 2018 and ends on the day prior to that which is forecasted. The value for the number of EVs that arrive on the day to be forecasted is determined using the same smart persistence technique described in Section VI-A.

Figure 2 compares TreeBagger, LSTM (trained with the same data as TreeBagger), and persistence forecast results

against real measured ADED for 50 days from January 10, 2020 to February 28, 2020. TreeBagger outperforms the other forecast techniques in mean absolute error (MAE) and root mean squared error (RMSE) as depicted in Table IV. The mean bias error (MBE) in Table IV suggests that on average, each of the techniques forecasts below the actual value. While there is greater bias in the two machine learning methods, the TreeBagger forecast strongly outperforms the other two techniques in MAE and RMSE, demonstrating that it is more accurate. The result from the TreeBagger forecast is then used as a feature in the LSTM forecast for NAPM.

Fig. 2: Forecasted and measured ADED from January 10 to February 28, 2020.

D. Long Short Term Memory for NAPM

LSTM is a form of recurrent neural network (RNN) [29]. It can be thought of as a RNN with a hidden layer that is replaced by a memory cell. LSTM helps to overcome the challenge of vanishing gradients in RNNs. A full description of LSTM can be found in [30].

The LSTM forecast is trained using features depicted in Table V with data spanning from November 10, 2018 up until the day prior to that which will be forecasted. The validation set is composed of data from the day being forecasted. A quadruplet of time parameters can be used to describe the LSTM forecast: the forecast horizon is 24 h, the resolution is minutely, the lead time is 0 minutes, and the update rate is daily. The lead time of 0 minutes means that the forecast occurs at the start of the calendar day being forecasted. The LSTM forecast estimates NAPM one calendar day at a time. The ADED and the forecasted total daily number of EVs are provided from the TreeBagger and smart persistence forecasts respectively as discussed in Section VI-C. Each of the other features help classify temporal information. The output of the LSTM forecast is a vector in which each value represents the cumulative sum of EVs that have arrived by a corresponding

TABLE V: List of the features used in the LSTM forecast for NAPM.

Feature	Binary	Feature	Binary
ADED	No	In academic session	Yes
Aggregate # EVs per day	No.	Minute of the hour	No
Day of the month	No.	Month	No
Day of the week	No.	# EVs at a given time	N ₀
Holiday	Yes	Weekend	Yes
Hour of the day	No		

Fig. 3: (a) Measured cumulative NAPM on Jan 27, 2020 compared to forecasted cumulative NAPM from kNN, LSTM and smart persistence. (b) Measured number of EVs plugged in throughout the day compared to kNN and smart persistence forecasts.

minute of a day. An example of the NAPM forecasted by both persistence and LSTM techniques compared against the measured NAPM vector is shown in Figure 3(a).

In testing over the same 50 days from January 10, 2020 to February 28, 2020, LSTM outperformed the persistence forecast, as shown in Table VI.

E. kNN for EV AT, PD, and ED

The output of the LSTM NAPM forecast is provided as the input vector for a kNN algorithm that finds the nearest neighbor NAPM to the LSTM output. Using kNN to determine the nearest historical NAPM vector to the LSTM output also yields associated PD and ED values. Using a previous day's EV data automatically ensures that ED does not exceed the maximum power rating of 7.2 kW multiplied by the respective EV's PD. The LSTM forecast is trained using features depicted

TABLE VI: Errors in NAPM for different forecast methods. Errors cover the time period from January 10 to February 28, 2020. The percentage error for each of these values was determined by normalizing using the actual total daily number of EVs that arrived.

	LSTM	kNN	Persistence
MAE $(\%)$	7.12	5.47	12.81
RMSE $(\%)$	9.59	8.43	24.52
MBE $(\%)$	-4.42	-0.99	-0.17

TABLE VII: Errors for ADED (kWh) and total daily # of EVs plugged using kNN and smart persistence forecasting compared to historical data for January 10 to February 28, 2020. The percentage error for total daily # of EVs was determined by normalizing using the actual maximum daily number of EVs that plugged in during a time step.

in Table V with data spanning from November 10, 2018 up until the day prior to that which will be forecasted.

The NAPM errors for the kNN algorithm and LSTM are depicted in Table VI. The errors for the kNN forecasted NAPM improves upon the error from LSTM.

Errors for ADED and total daily number of EVs plugged in for kNN and persistence forecasts are depicted in Table VII. The kNN forecast perfromed better in ADED error but incurred a slightly larger error in total daily number of EVs.

The individual error analysis for EVs was executed on a daily basis per Algorithm 1. The hybrid ML forecast outperformed the persistence forecast in mean AT error and mean PD error, with slightly worse performance in ED, as depicted in Table VIII.

To account for cases in which a forecast had either more or fewer EVs than actually arrived, a cost equal to the single largest daily cost of all linearly assigned pairs is allocated to each unmatched EV. The total daily cost for all paired and unpaired EVs is the sum of the individual costs of all forecasted EVs plus costs allocated to each remaining real EV in cases where under forecasting occurred. The sum of all 50 daily costs from Jan 10 to Feb 28, 2020 for the hybrid ML forecast is 3, 551 and for persistence it is 5, 077. Persistence forecasting results in a larger distance metric as the number of EVs that are either over- or under-forecasted is 354 versus 112 for hybrid ML.

F. Implementation and Computational Cost

Figure 4 depicts the implementation strategy. The ML algorithm trains once daily at 00:00 h which consumes on

TABLE VIII: Mean individual EV forecast errors for hybrid ML and persistence forecasting compared to historical data for January 10 to February 28, 2020.

	AT (minutes)	PD. (minutes)	ED (kWh)	Cumulative #EVs Over/Under Forecasted
Hybrid ML	1.25	1.75	-1.64	112
Persistence	15.8	-13.03	-1.36	354

TABLE IX: Weekday mean peak loads and total monthly cost.

	Summer (kW)	Winter (kW)	Summer (\$)	Winter (\$)
Dumb Charging	137.4	128.3	8.447	7,686
No Forecast MPC	80.8	77.4	7.281	6,049
Persistence Forecast MPC	75.3	71.0	6,686	5,566
Hybrid ML Forecast MPC	73.2	70.0	6,680	5,512
Perfect Forecast MPC	67.0	65.1	6.073	5,069

average 9 minutes of run time. The run time for a single 24 hour MPC is 34 seconds or less. The overall ML MPC algorithm has a run time of 36 minutes for forecasting and controlling EV charging that takes place in a single day.

Fig. 4: Block diagram depicting the implementation of the forecasting and control scheme.

VII. RESULTS AND DISCUSSION

A. Peak and Cost Reduction

Two 30-day case studies were executed with a summer (S) case study ranging from September 23 to October 23, 2019, and a winter (W) case study ranging from January 20 to February 19, 2020. While the periods cover two calendar months, for the purpose of demand charges we assume that only one demand charge is assessed over the 30 days. Both time periods are in academic session. Figures 5 and 6 depict the EV load profiles for each of the techniques over the 30 day period of case study S and W, respectively.

Dumb charging in Figs. 5 and 6 depict what actually occurred with uncontrolled charging. Dumb charging results in large morning peaks and declining charging power in the afternoons due to many customers arriving during morning hours. Occasionally there is a minor afternoon peak that is associated with new customers arriving shortly after 12:00 h. Additionally, the MPC was executed for a no forecast case which optimizes only over EVs that are plugged in at any given moment with known ED and PD (while ignoring EVs that plug in in the future) and a perfect forecast case (assuming future charging demand to be known), demonstrating the worst and best case scenario for this MPC formulation. The ML MPC significantly reduces the peaks experienced by both dumb charging and no forecast MPC, reducing monthly electricity costs. Hybrid ML MPC also outperforms Persistence MPC.

Average peak load values for the 22 weekdays during the two case studies are depicted in Table IX. For case study S, the MPC with ML forecasts reduced the 22-day weekday average peak load by 46.7% from dumb charging, 9.4% from MPC with no forecast and 2.8% from MPC with persistence forecasts. For case study W, the MPC with ML reduced the 22 day weekday average peak load value by 45.4% from dumb charging, 9.6% from MPC with no forecast and 1.4% from

Fig. 5: (a): 30 day EV dispatch load curve for case study S. Non Coincident Demand Charge event values for each dispatch are labeled with arrows depicting the corresponding dispatch color. (b): Zoomed in dispatch load curve of 3 days of case study S.

MPC with persistence. The 22-day weekday average peak loads demonstrate consistent performance, but demand costs are only associated with the single highest peak experienced during the month for each MPC implementation. The monthly peak for persistence MPC in winter was 88.2 kW versus 89.2 kW for hybrid ML MPC and 97.1 kW versus 90.5 kW for case study S. The total cost for both case studies was lower for the hybrid ML forecast MPC than the Persistence MPC.

In order for the LSTM NAPM forecast to accurately predict EV behavior the LSTM learning parameters must be properly tuned, otherwise the Persistence forecast may outperform the Hybrid ML MPC method. For this work case study W used 150 hidden untis and case study S used 200 hidden units. Future research will look into the impact of training parameters on forecast performance.

Case study S costs for the ML MPC method decreased 20.9% from dumb charging cost, 8.3% from the no forecast MPC cost and 0.1% from the persistence MPC cost. Case study W costs for the ML MPC method decreased 28.3% from the dumb charging cost, 8.9% from the no forecast MPC cost and 1% from the persistence MPC cost. These improvements suggest that the hybrid ML forecasting paired with MPC is an effective charge dispatch method for reducing electricity

Fig. 6: (a): 30 day EV dispatch load curve for case study W. Non Coincident Demand Charge event values for each dispatch are labeled with arrows depicting the corresponding dispatch color. (b): Zoomed in dispatch load curve of 3 days of case study W.

costs. MPC with hybrid ML forecasts is most useful during the summer because of the increased summer commodity peak demand charge cost.

B. MPC Sensitivity to Forecast Error

Two 30-day sensitivity studies were executed for the same date ranges of case studies S and W. The sensitivity studies demonstrate how the MPC algorithm performs under varying forecast errors. Each sensitivity study is composed of four error analysis tests that determine how the MPC performs under 1) a normally distributed error with a standard deviation of 10% relative to each value of PD and ED with perfect forecast for AT, 2) a normally distributed error with a standard deviation of 50% on PD and ED with perfect forecast for AT, 3) an over forecasting scenario that assumes perfect forecast for all real EVs plus additional EVs drawn randomly from all previous EV charging data equaling 30% of the total that arrived on a given day and, 4) an under forecasting scenario that assumes perfect forecast for all real EVs minus a number of EVs equal to 30% of the total that arrived on a given day which are removed randomly from the total.

Fig. 7: Sensitivity study dispatch load for Sept 27, 2019. TABLE X: Sensitivity study weekday mean peak loads and

total monthly cost.

Figure 7 depicts an example dispatch of each case on Sept 27, 2019. Over forecasting schedules charging earlier in the day because of anticipated future EV arrivals. This leads to large morning peaks, but reduces afternoon charging during on-peak hours. Under forecasting reduces morning dispatch which results in higher afternoon costs during the on-peak period. The worst case scenario of under forecasting is MPC with no forecast, which we demonstrated in the primary S and W case studies in Table IX and Figures 5-6.

Table X depicts the average peak loads for the 22 weekdays and the total costs incurred during the summer and winter sensitivity studies. While MPC with 30% over forecasting experiences larger peak loads, the costs incurred by over and under forecasting cases are within 0.6% of each other, due to the increased on-peak period energy and demand charge costs incurred by the under forecasting study.

VIII. CONCLUSIONS

This work developed a hybrid machine learning model to forecast individual EV ATs, EDs, and PDs for 48 charge ports at UC San Diego. These forecasted values were input to an MPC based charge scheduler that reduced utility customer costs associated with demand charges and TOU energy rates.

This work addresses a gap in the literature associated with using advanced forecasts to predict individual EV charging characteristics for real time MPC. This work is novel in the following ways:

- 1) It utilized a hybrid ML forecast to estimate AT, ED, and PD for individual EVs.
- 2) It utilized ED rather than SOC for forecasting and implementing MPC, thereby reducing the number of forecasted variables and simplifying the MPC problem.
- 3) It implemented a novel comparison technique for variables at differing indices within vectors of differing length for individual error analysis.

The results demonstrate that forecasting and controlling individual EV charging at a plaza effectively reduces utility customer electricity costs. The novelties enumerated above led to reduced peak loads and reduced costs. Weekday average peak loads using the ML MPC method were 47.0% lower than the dumb charging scenario, 10.0% lower than the no forecast MPC charging scenario and 3.3% lower than the persistence forecast MPC for case study S. ML MPC monthly costs were 22.0% lower than dumb charging scenario, 9.5% lower than the no forecast MPC and 1.4% lower than the persistence forecast case for the S case study. Results for a winter case study were similar.

A sensitivity study was carried out to determine the impact of forecast accuracy on the MPC algorithm. While the sensitivity study forecast with an error having a standard deviation of 10% to PD and ED outperformed the other sensitivity studies, the MPC successfully mitigated the risk of significantly increased cost due to poor forecasting. For the S and W sensitivity studies, the costs for each of the four error analysis tests decreased between $10.5\% - 13.1\%$ and 10.3% − 14.5% from the no forecast MPC, respectively.

While this work focused on solving an objective function with the primary purpose of peak reduction through demand charge management, the objective and constraints could be adjusted slightly to encourage the bulk of charging to occur during periods of high solar PV production, low building demand, or low market prices. The hybrid ML forecasting technique could be used on its own for determining charging demand at any level of granularity. The problem formulation presented in this work only requires knowledge of AT, PD, and ED rather than AT, PD, and EV SOC. Updated forecasts throughout the day could better assist the MPC implementation. Future work will look into how the forecast can be updated most effectively.

REFERENCES

- [1] L. Friedman, "California reveals its plan to phase out new gas-powered cars by 2035," The New York Times, Apr 2022. [Online]. Available: https://www.nytimes.com/2022/04/13/climate/california-electricvehicles.html.
- [2] U. Datta, A. Kalam, J. Shi, and J. Li, "Electric vehicle charging station for providing primary frequency control in microgrid," in *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2019, pp. 2440–2444.
- [3] V.-L. Nguyen, T. Tran-Quoc, S. Bacha, and B. Nguyen, "Charging strategies to minimize the peak load for an electric vehicle fleet, in *IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society*, 2014, pp. 3522–3528.
- [4] A. S. Al-Ogaili, T. J. Tengku Hashim, N. A. Rahmat, A. K. Ramasamy, M. B. Marsadek, M. Faisal, and M. A. Hannan, "Review on scheduling, clustering, and forecasting strategies for controlling electric vehicle charging: Challenges and recommendations," *IEEE Access*, vol. 7, pp. 128 353–128 371, 2019.
- [5] A. Brown, J. Levene, A. Schayowitz, and E. Klotz, "Electric vehicle charging infrastructure trends from the alternative fueling station locator: Second quarter 2021." Golden, CO: National Renewable Energy Laboratory. [Online]. Available: NREL/TP-5400- 81153. www.nrel.gov/docs/fy22osti/81153.pdf.
- [6] B. Washom, "Ucsd strategic energy initiatives." [Online]. Available: https://rmp.ucsd.edu/strategic-energy/index.html
- [7] Z. J. Lee, G. Lee, T. Lee, C. Jin, R. Lee, Z. Low, D. Chang, C. Ortega, and S. H. Low, "Adaptive charging networks: A framework for smart electric vehicle charging," *IEEE Transactions on Smart Grid*, vol. 12, no. 5, pp. 4339–4350, 2021.
- [8] M. Majidpour, C. Qiu, P. Chu, H. R. Pota, and R. Gadh, "Forecasting the ev charging load based on customer profile or station measurement, *Applied Energy*, vol. 163, pp. 134–141, 2016.
- [9] E. S. Xydas, C. E. Marmaras, L. M. Cipcigan, A. S. Hassan, and N. Jenkins, "Forecasting electric vehicle charging demand using support vector machines," in *2013 48th International Universities' Power Engineering Conference (UPEC)*, 2013, pp. 1–6.
- [10] H. Chung, B. Alinia, N. Crespi, and C. Wen, "An EV charging scheduling mechanism to maximize user convenience and cost efficiency," *CoRR*, vol. abs/1606.00998, 2016.
- [11] M. B. Arias and S. Bae, "Electric vehicle charging demand forecasting model based on big data technologies," *Applied Energy*, vol. 183, pp. 327–339, 2016.
- [12] A. Gautam, A. K. Verma, and M. Srivastava, "A novel algorithm for scheduling of electric vehicle using adaptive load forecasting with vehicle-to-grid integration," in *2019 8th International Conference on Power Systems (ICPS)*, 2019, pp. 1–6.
- [13] B. Chokkalingam, S. Padmanaban, P. Siano, R. Krishnamoorthy, and R. Selvaraj, "Real-time forecasting of ev charging station scheduling for smart energy systems," *Energies*, vol. 10, no. 3, 2017.
- [14] S. Shahriar, A. R. Al-Ali, A. H. Osman, S. Dhou, and M. Nijim, "Prediction of ev charging behavior using machine learning," *IEEE Access*, vol. 9, pp. 111 576–111 586, 2021.
- [15] W. Tang and Y. J. Zhang, "A model predictive control approach for low-complexity electric vehicle charging scheduling: Optimality and scalability," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1050–1063, 2017.
- [16] Y. Zheng, Y. Song, D. J. Hill, and K. Meng, "Online distributed mpc-based optimal scheduling for ev charging stations in distribution systems," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 2, pp. 638–649, 2019.
- [17] M. Tahmasebi, A. Ghadiri, M. Haghifam, and S. Miri-Larimi, "Mpcbased approach for online coordination of evs considering ev usage uncertainty," *International Journal of Electrical Power & Energy Systems*, vol. 130, p. 106931, 2021.
- [18] T. Zhang, W. Chen, Z. Han, and Z. Cao, "Charging scheduling of electric vehicles with local renewable energy under uncertain electric vehicle arrival and grid power price," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2600–2612, 2014.
- [19] H. Mohsenian-Rad and M. ghamkhari, "Optimal charging of electric vehicles with uncertain departure times: A closed-form solution," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 940–942, 2015.
- [20] Y. He, B. Venkatesh, and L. Guan, "Optimal scheduling for charging and discharging of electric vehicles," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1095–1105, 2012.
- [21] I. S. Duff and J. Koster, "On algorithms for permuting large entries to the diagonal of a sparse matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 22, no. 4, pp. 973–996, 2001. [Online]. Available: https://doi.org/10.1137/S0895479899358443
- [22] "Evse: Electric vehicle (ev) charging stations," ChargePoint. [Online]. Available: https://www.chargepoint.com/
- [23] "Ev charging." PowerFlex, 2021. [Online]. Available: https://www.powerflex.com/products/ev-charging/.
- [24] "Historical tariffs," Historical Tariffs | San Diego Gas & Electric. [Online]. Available: https://www.sdge.com/rates-andregulations/historical-tariffs
- [25] A. Ghosh, M. Z. Zapata, S. Silwal, A. Khurram, and J. Kleissl, "Effects of number of electric vehicles charging/discharging on total electricity costs in commercial buildings with time-of-use energy and demand charges," *Journal of Renewable and Sustainable Energy*, vol. 14, no. 3, p. 035701, 2022.
- [26] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.
- [27] ——, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/ boyd/graph_dcp.html.
- [28] T. M. Inc., "Statistics and machine learning toolbox," Natick, Massachusetts, United States, 2022. [Online]. Available: https://www.mathworks.com/help/stats/index.html
- [29] S. Hochreiter, S. Hochreiter, and J. Schmidhuber, "Long short-term memory," *Neural computation.*, vol. 9, no. 8, 1997-11-01.
- [30] Z. C. Lipton, "A critical review of recurrent neural networks for sequence learning," *CoRR*, vol. abs/1506.00019, 2015.