PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

REAL-TIME STRATEGIES FOR AN ELECTRIC VEHICLE AGGREGATOR TO PROVIDE ANCILLARY SERVICES

GEORGE ALFONS WENZEL ACUÑA

Thesis submitted to the Office of Research and Graduate Studies
in partial fulfillment of the requirements for the degree of
Master of Science in Engineering

Advisors:
MATÍAS NEGRETE PINCETIC
DANIEL OLIVARES QUERO

Santiago de Chile, August 2016

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

REAL-TIME STRATEGIES FOR AN ELECTRIC VEHICLE AGGREGATOR TO PROVIDE ANCILLARY SERVICES

GEORGE ALFONS WENZEL ACUÑA

Members of the Committee:
MATÍAS NEGRETE PINCETIC
DANIEL OLIVARES QUERO
FELIPE NÚÑEZ RETAMAL
DORIS SÁEZ HUEICHAPAN
LUCIANO CHIANG SÁNCHEZ

Thesis submitted to the Office of Research and Graduate Studies
in partial fulfillment of the requirements for the degree of
Master of Science in Engineering

Santiago de Chile, August 2016

Gratefully to my parents

ACKNOWLEDGEMENTS

I would like to dedicate this work to my parents, Susana and Alfonso. Their inconditional support has been fundamental for all my achievements so far, and their help made my studies and this research work possible.

I would like to thank professors Matías Negrete and Daniel Olivares for receiving me in their research group, and for making me part of the founders of OCM Lab at PUC, an amazing group of people that has grown a lot since its foundation during 2014, and which I am sure will continue to do so. From this group, I also want to thank Rodrigo Henríquez for being such a good research partner.

I would also like to thank Matías and Daniel for giving me the opportunities of going with them to the 2015 IEEE Power and Energy Society General Meeting in Denver, and spending three months at the University of California, Berkeley, during 2015, as a visiting student researcher. These two international experiences were of vital importance to my learning process, and have opened many doors for my future.

Additionally, I want to thank professor Duncan Callaway for receiving me at Berkeley with his amazing research group EMAC, and the people I met there who helped me both with my work and with having a great time there: Jason MacDonald, Somil Bansal and Sandro Iacovella.

This work was supported by the Institute in Complex Engineering Systems (Grants ICM:P05-004F and CONICYT:FB016) and the Fondecyt Grant 11140621.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Real-time charging strategies, in the context of vehicle to grid (V2G) technology, are needed to enable the use of electric vehicle (EV) fleets batteries to provide ancillary services (AS). In this work we develop tools to manage charging and discharging in a fleet to track an Automatic Generation Control (AGC) signal when aggregated. We propose a real-time controller based on convex optimization, that considers bidirectional charging efficiency, and extend it to study the effect of looking ahead when implementing Model Predictive Control (MPC). Different complexity levels are introduced to the controller, by using approximations and/or relaxations, which hold under sufficient conditions that are provided. Simulations show that more complex controllers improve tracking error as compared with benchmark scheduling algorithms, as well as regulation capacity and battery cycling impacts. However, they face a trade-off between ease of implementation and performance.

Keywords: Electric Vehicles, Resource Scheduling, Ancillary Services Market, Vehicle to Grid Control, Model Predictive Control, Convex Optimization

RESUMEN

Las estrategias de control en tiempo real, en el contexto de la tecnología *vehicle-to-grid* (V2G), son necesarias para permitir el uso de flotas vehículos eléctricos (EVs), mediate sus baterías, para proveer servicios complementarios a los sistemas eléctricos de potencia. En este trabajo se desarrollan herramientas para administrar la carga y descarga de las baterías de una flota, con el objetivo de seguir una señal de control automático de generación (AGC) cuando los vehículos se usan de forma agregada. Se propone un controlador en tiempo real basado en optimización convexa, que considera la eficiencia bidireccional de carga de las baterías, y se extiende para estudiar el efecto de mirar hacia el futuro con control predictivo (MPC). Se introducen diferentes niveles de complejidad al controlador, mediante el uso de aproximaciones y/o relajación de restricciones, que son válidas bajo condiciones suficientes que son indicadas. Simulaciones muestran que los controladores más complejos mejoran el error en el seguimiento de la señal AGC, al comparar con algoritmos alternativos, así como también la capacidad de regulación y el impacto en el *cycling* de las baterías. Sin embargo, enfrentan una disyuntiva entre facilidad de implementación y rendimiento.

Palabras Clave: Vehículos Eléctricos, Planificación de Recursos, Servicios Complementar, Control de *Vehicle-to-grid*, Control Predictivo, Optimización Convexa

1. INTRODUCTION

1.1. Context

High levels of penetration of renewable energy, such as wind and solar, have brought increased intermittency, volatility and uncertainty to the operation of the electrical grid. These salient characteristics of renewable energy sources create challenges for the achievement of the required equilibrium between supply and demand in energy systems. The system's ability of adapting to achieve such equilibrium is called flexibility in recent technical literature.

Different ways of improving the flexibility of power systems have been studied over the past few years, including flexibility from generation, transmission, demand and system operation strategies/policies. Currently, the operation of electrical power systems is dominated by large thermal generators, with high levels of CO_2 emissions. Being able to harness the potential for flexibility all the aforementioned resources have is the key to achieve larger shares of intermittent, but environmentally friendly generation technologies.

New generation, demand, transmission and storage technologies present opportunities to increase power system flexibility. In particular, distributed resources have emerged as an alternative to traditional technologies. However, in order to capitalize on these opportunities, new algorithms that coordinate distributed resources to provide aggregate response are needed to make proper use of these assets (efficient and reliable). This work focuses on the potential of a subset of flexible storage technologies: namely, plug-in electric vehicles.

The optimal amount of energy storage capacity in a power system scales with the variability in demand and power generation (Gast et al., 2013); however stationary batteries are still too expensive for most grid-tied applications despite their decreasing cost (D. L. Anderson, 2009; Nykvist & Nilsson, 2015). Electric vehicle (EV) batteries can be used during their idle time when parked to extract/inject power from/to the grid in the

same way that stationary batteries might. By creating revenue for EV owners – and lowering the total cost of EV ownership – this vehicle to grid (V2G) framework could provide a cost-effective means to add storage capacity to the grid.

V2G technology is an ongoing research subject. The available energy V2G can provide to the grid at any given time is considerable, since EVs are parked for most of the day, and the high rate at which their batteries can be charged or discharged makes V2G specially valuable to stabilize intermittency in the operation of the grid. However, there are many challenges that need to be addressed before enabiling the large-scale implementation of V2G technology. Since this technology turns EV owners into service providers, market schemes for organizing the aggregate participation of these distributed resources in the stabilization of the grid are being developed. These schemes must be designed in such a way that incentives for aggregators and EV owners are proportional to the benefits they bring to the operation of the grid.

1.2. Literature Review

Multiple sources of variability are currently being integrated into power grids, mainly due to the large-scale integration of renewable energy sources driven by concerns about environmental impacts and supply security issues. As a consequence, new ways of operating and controlling the grid are under development. Renewable generation has succeeded in achieving great environmental benefits; however, resources with such volatility present a great operational challenge for power systems operators. In fact, some authors show that a threshold can be found on the variability of net load[1] beyond which the value of these resources is questionable (Meyn et al., 2010). In order to address these challenges and crystallize the benefits of renewable energy sources, studying the concept of *Flexibility* in power systems is of utmost importance.

[1] Net load is the demand curve minus non-dispatchable generation, such as wind or solar.

1.2.1. Flexibility in Power Systems

Flexibility in power systems may be understood as the ability to respond to changes in load and generation. According to Cochran et al. (2014), it is composed by:

- **Generation flexibility:** determined by minimum up and minimum down times, ramping capabilities, and power output limits.
- **Transmission flexibility:** determined by the bottlenecks and access.
- **Demand flexibility:** determined by the availability of Demand Response (DR), storage and load control capability.
- **System operation flexibility:** determined by institutional market structures to support the deployment of physical flexibility. This may include the ability to make decisions closer to real time and better information access.

Since early 80s, market structures have been developed in power and energy systems. A common characteristic among these designs is the existence of forward markets, in addition to the spot market, for which they are referred to as *multi-settlement markets*. These forward markets behave differently depending on the time scale in which they are cleared: as they get closer to real time, the price is set by a market coordinator and competitors' bids, rather than by bilateral negotiations between generators and consumers (S. C. Anderson, 2004).

Electricity markets today typically consider the interaction between two coupled markets: day-ahead and real-time. While the day-ahead market is cleared one day ahead of production, based on forecasts for the Independent System Operator (ISO) to schedule generation at each hour, real-time market is in charge of fine-tuning this resource allocation to compensate deviations from day-ahead predictions, due to the unpredictable nature of net demand[2]. Thus, resources able to provide flexibility, that helps making the fine-tuning process easier and more efficient, can become specially valuable and are beginning to be compensated.

[2]Net load.

The idea of exploiting the flexibility associated with the demand side has been widely investigated (Albadi & El-Saadany, 2008; Spees & Lave, 2007), and markets aiming for flexible loads to be serviced by zero-marginal cost renewable generation have been designed as well (Nayyar et al., 2014). The consensus is that flexibility is an attribute that needs to be compensated in electricity markets; however, there is no standardized metric for measuring the flexibility of a system. In Ulbig & Andersson (2015), the authors define operational flexibility as the ability of a power system unit to modulate electrical power fed into the grid and/or power fed out from the grid over time, using four metrics from Makarov et al. (2009):

- Power capability π for up and down regulation in [MW].
- Energy storage capability ϵ in [MWh].
- Power ramping capability ρ in [MW/min]
- Power ramping duration δ in [min]. This indicator is dependent of π and ρ, $\delta = \frac{\pi}{\rho}$.

The role of the tuple $\{\rho, \pi, \epsilon\}$ in the flexibility of electrical systems is depicted in Figure 1.1. When a system is capable of supporting steeper ramps, more power or larger amounts of energy, it is considered more flexible than other system with less of those capacities. These three terms are closely related through differentiation and integration.

There are many papers that consider the ramping capabilities of a system to be the principal source of flexibility (M. Chen et al., 2006; Cho & Meyn, 2010). However, considering the capacity of the system to change in terms of power and energy provides a richer concept of flexibility. The presence of three terms enables researchers to plot a *flexibility cube* of maximum available flexibility of a generic power system and study its evolution over time. The illustrations from Ulbig & Andersson (2015) show that aggregation of different types of flexibility cubes can result in units that are even more flexible than the exact sum of their volumes. These studies propose a whole new approach to study flexibility, through the geometry of these units and how they interact with each other.

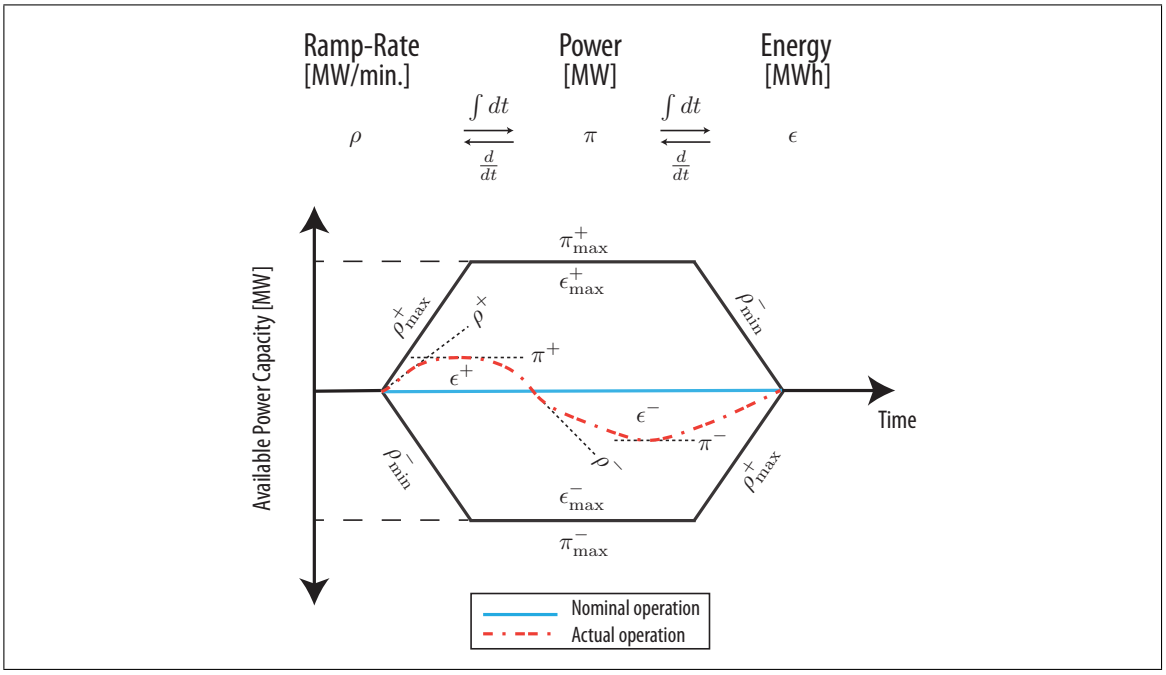


Figure 1.1. Flexibility trinity. Source: Ulbig & Andersson (2015)

A particularly insightful characterization of the flexibility of a system is provided in Petersen et al. (2013). Flexibility is defined as the *ability to deviate from the plan*, and the authors focus on four constraints to characterize a flexible system: 1. Power Capacity, 2. Energy Capacity, 3. Energy level at a specific deadline, and 4. Minimum runtime. Using that set of characteristics, three simple flexibility models are built: *The Bucket*, *The Battery* and *The Bakery*. The first model, *The Bucket*, is a power and energy constrained integrator (1 and 2), and could be used for Thermostatically Controlled Loads (TCLs). The second model, *The Battery*, adds the restriction of a specific deadline with an energy goal (1 to 3), which could be used to describe Electric Vehicles (EVs). The third model, *The Bakery*, adds the constraint that the process must run in one continuous stretch at constant power consumption (1 to 4), which could be the case of a commercial green house.

Finally, knowing what the consequences of a lack of flexibility might be may be relevant to understand its relevance. Signs of inflexibility in the physical operation of power

systems include difficulties balancing demand with supply and significant Renewable Energy Source (RES) curtailments, while in the economic operation it can result in negative market prices and high price volatility.

1.2.2. Ancillary Services

1.2.2.1. General description

Power systems require Ancillary Services (AS) to deliver energy to the consumers in a reliable way, which is achieved by properly managing supply-demand imbalances at various time-scales. A definition of these services is provided by the Federal Energy Regulatory Commission (FERC) (FERC, 2016):

> Ancillary Services are those services necessary to support the transmission of electric power from seller to purchaser, given the obligations of control areas and transmitting utilities within those control areas, to maintain reliable operations of the interconnected transmission system.

Similarly, Kuzle et al. (2007) states that the ISO uses AS with the following objectives:

- Keeping the frequency of the system within certain bounds.
- Controlling the voltage profile of the system.
- Maintaining the stability of the system.
- Preventing overloads in the transmission system.
- Restoring the system or portions of the system after a blackout.

The offer of these services may be managed in many ways, among which the market approach is the one preferred by the most advanced power system operators in the world, such as PJM[3] and CAISO[4]. Therefore, AS markets can be understood as reward schemes

[3]Pennsylvania-New Jersey-Maryland Interconnection, plus 11 States
[4]California ISO

to compensate resource owners for providing operational flexibility, for which different mechanisms have been designed.

As explained in Kirby (2007), an important characteristic about AS is that they are capacity services, not energy services, which means that their price is strongly tied with the temporal volatility of the marginal price for power. Therefore, the costs of AS are mainly due to opportunity costs, and consequently their prices are volatile as well.

There is no unique classification of Ancillary Services in the world, and it is common for a particular service to have different names in different countries. As stated in Kuzle et al. (2007), the European Union uses the European Directive 2003/54/EC (European Parliament, 2003) as a general framework, and in USA FERC's Order 888 (FERC et al., 1996) defines 12 technical and non-technical AS, from which 6 are mandatory (providers and customers are forced to sell and buy them). The main services are:

(i) Scheduling, System Control and Dispatch Service [5]
(ii) Reactive Supply and Voltage Control from Generation Sources Service [5]
(iii) Regulation and Frequency Response Service [6]
(iv) Energy Imbalance Service [6]
(v) Operating Reserve - Spinning Reserve Service [6]
(vi) Operating Reserve - Supplemental Reserve Service [6]

A description of the most common AS is shown in Table 1.1, where they are classified according to the situations where they are required: Normal Conditions, Contingency Conditions or Other Services. However, with the objective of comparing the different names they receive around the world, Ancillary Services can also be classified according to the time scale in which they are required (Table 1.2) (Beck & Scherer, 2010):

[5] Transmission Provider must provide and Transmission Customer must purchase from Transmission Provider.

[6] Transmission Provider must offer to provide only to Transmission Customer serving load in Transmission Provider's control area and Transmission Customer must acquire, but may do so from Transmission Provider, a third party or self supply.

(i) **Primary Control:** Commonly known as Frequency Response, occurs within the first few seconds following a change in system frequency to stabilize the Interconnection. It is provided by Governor Action and Load. It will not return frequency to normal, but only stabilize it.

(ii) **Secondary Control:** Typically includes the balancing services deployed in the minutes time frame. This control is accomplished using Automatic Generation Control (AGC)[7] and the manual actions taken by the dispatcher to provide additional adjustments. Also includes initial reserve deployment for disturbances.

(iii) **Tertiary Control:** It encompasses actions taken to get resources in place to handle current and future contingencies. Reserve deployment and Reserve restoration following a disturbance are common types of Tertiary Control.

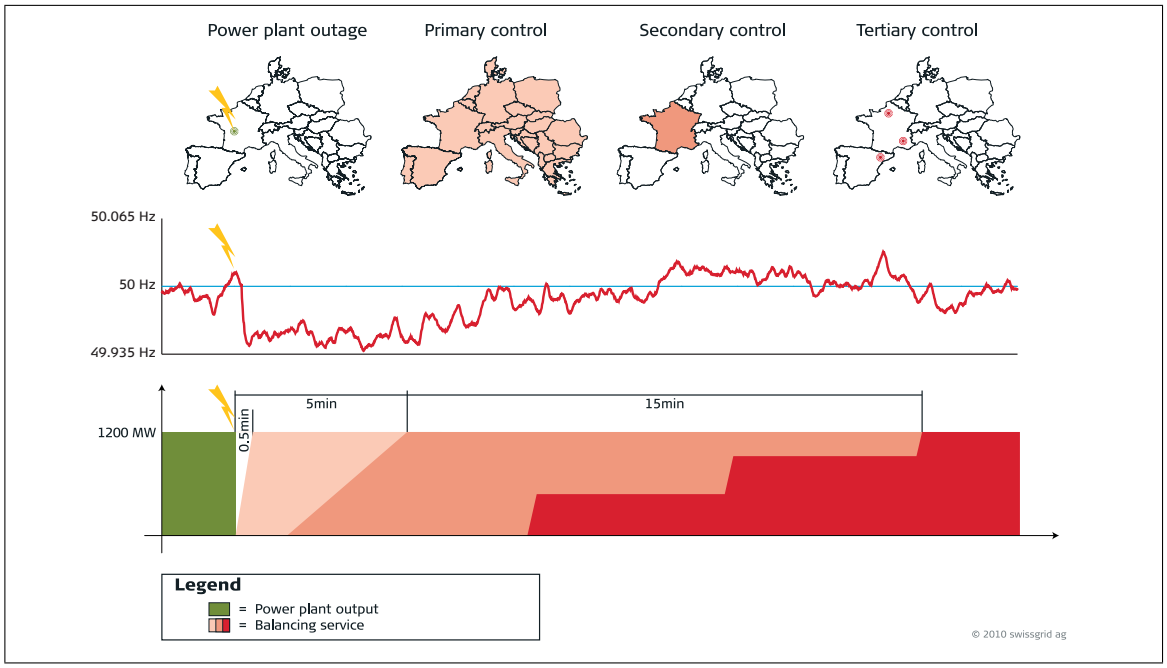


Figure 1.2. Example of a power plant outage in France. Source: Beck & Scherer (2010)

[7]Automatic Generation Control consists in adjusting the power output of multiple generators in response to changes in load. In order to do this, the dispatcher emits a signal to the generators that indicates them how to adjust their power output.

Table 1.1. Properties of key Ancillary Services. Source: Kirby (2007); NERC (2011b).

Category	Service	Service Description	Response Speed	Duration	Cycle Time	Market Cycle
Normal Conditions	Regulating Reserve	Online resources, on automatic generation control, that can respond rapidly to system operator requests for up and down movements; used to track the minute-to-minute fluctuations in system load and to correct for unintended fluctuations in generator output to comply with Control Performance Standards (CPSs) 1 and 2 of the North American Electric Reliability Council (NERC 2006).	1 min	Minutes	Minutes	Hourly
	Load Following or Fast Energy Markets	Similar to regulation but slower. Bridges between the regulation service and the hourly energy markets.	10 min	10 min to hours	10 min to hours	Hourly
Contingency Conditions	Spinning Reserve	Online generation, synchronized to the grid, that can increase output immediately in response to a major generator or transmission outage and can reach full output within 10 min to comply with NERC's Disturbance Control Standard (DCS).	Seconds to <10 min	10 to 120 min	Hours to days	Hourly
	Non-spinning Reserve	Same as spinning reserve, but need not respond immediately; resources can be offline but still must be capable of reaching full output within the required 10 min.	<10 min	10 to 120 min	Hours to days	Hourly
	Replacement or Supplemental Reserve	Same as supplemental reserve, but with a 30-60 min response time; used to restore spinning and non-spinning reserves to their pre-contingency status.	<30 min	2 hours	Hours to days	Hourly
Other Services	Voltage Control	The injection or absorption of reactive power to maintain transmission-system voltages within required ranges.	Seconds	Seconds	Continuous	Year(s)
	Black Start	Generation, in the correct location, that is able to start itself without support from the grid and which has sufficient real and reactive capability and control to be useful in energizing pieces of the transmission system and starting additional generators.	Minutes	Hours	Months to years	Year(s)

Table 1.2. European Reserves, Sample North American and Chilean Designations. Source: NERC (2011b); CDEC-SIC (2016).

<table>
<tr><th></th><th>Primary Frequency Control Reserves</th><th>Secondary Frequency Control Reserves</th><th colspan="4">Tertiary Frequency Control Reserves</th></tr>
<tr><td rowspan="2">PJM</td><td rowspan="2">Frequency Response</td><td colspan="3">Operating Reserve</td><td rowspan="2">Reserve Beyond 30 minutes</td></tr>
<tr><td>Regulation</td><td>Spinning Reserve</td><td>Quick Start Reserve</td></tr>
<tr><td rowspan="3">CAISO</td><td rowspan="3">Spinning Reserve</td><td colspan="3">Operating Reserve</td><td rowspan="3">Replacement Reserve and Supplemental Energy</td></tr>
<tr><td>Regulation</td><td colspan="2">Contingency Reserves</td></tr>
<tr><td colspan="2">Spinning Reserve</td><td>Non-Spinning Reserve</td></tr>
<tr><td>Germany</td><td>Primary Reserve</td><td>Secondary Reserve</td><td colspan="2">Minutes Reserve</td><td>Hours Reserve and Emergency Reserve</td></tr>
<tr><td rowspan="2">France</td><td rowspan="2">Primary Reserve</td><td rowspan="2">Secondary Reserve</td><td colspan="3">Tertiary Reserve</td></tr>
<tr><td colspan="2">Rapid 15 Minute Reserve</td><td>Complementary 30 Minute Reserve</td></tr>
<tr><td>Spain, Netherlands, Belgium</td><td>Primary Reserve</td><td>Secondary Reserve</td><td colspan="4">Tertiary Reserve</td></tr>
<tr><td rowspan="2">Great Britain</td><td>Operating Reserve</td><td rowspan="2">(does not exist)</td><td colspan="2">Operating Reserve</td><td colspan="2">Contingency Reserve</td></tr>
<tr><td>Response Primary/Secondary High Frequency</td><td>Regulating Reserve</td><td>Standing Reserve</td><td>Fast Start</td><td>Warming and Hot Standby</td></tr>
<tr><td>Sweden</td><td>Frequency Reserve and Disturbance Reserve</td><td>(does not exist)</td><td colspan="4">Seven different types of reserves</td></tr>
<tr><td>Australia</td><td>Contingency Reserve Fast/Slow/Delayed</td><td>Regulating Service and Network Loading Control</td><td colspan="4">Short Term Capacity Reserve</td></tr>
<tr><td>New Zealand</td><td>Instantaneous Reserve Fast/Sustained Over Frequency</td><td>Frequency Regulating Reserve</td><td colspan="4">(No name)</td></tr>
<tr><td>Chile</td><td>Primary Reserve (Spinning Reserve)</td><td>Secondary Reserve (Spinning Reserve)</td><td colspan="4">(does not exist)</td></tr>
</table>

1.2.2.2. Regulation and Frequency Response Service

As described in Table 1.1, regulation is the use of on-line generation that is equipped with Automatic Generation Control (AGC) and that can change output quickly to track the moment-to-moment fluctuations in customer loads and to correct for the unintended fluctuations in generation. It is the most expensive Ancillary Service (Kirby, 2007), and all load serving entities have hourly Regulation Obligation, which can be satisfied by self-scheduling own resources, bilateral transactions with other participants, or purchasing from AS market (PJM, 2013; FERC et al., 1996).

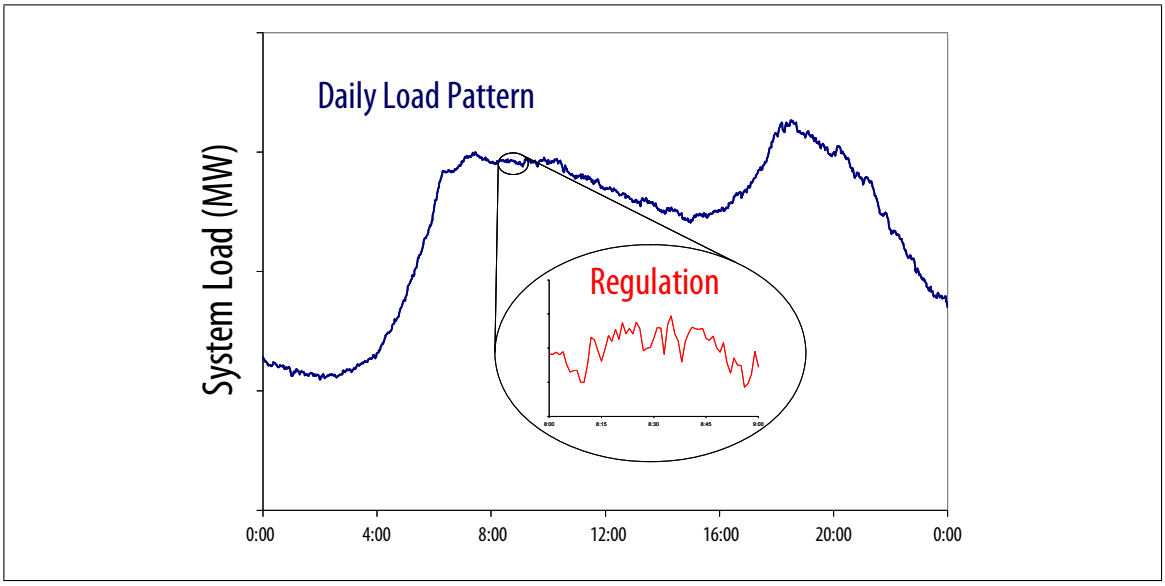


Figure 1.3. Regulation is a zero-energy service that compensates for minute-to-minute fluctuations in total system load and uncontrolled generation while load following compensates for the slower, more predictable changes in load. Source: Kirby (2007)

System frequency is an instantaneous estimator of how well the balance between generation and load is being done (it drops when load exceeds generation and rises when generation exceeds load), so there is a bandwidth around the nominal system frequency (50 Hz for Chile, 60 Hz for the US) within which it must be kept. Frequency can be interpreted as the fundamental indicator of the health of a power system (NERC, 2011a). The ISOs have a high willingness to pay for fast-response AS that will help keeping the frequency within its tolerance, because it can otherwise result in power system collapse and/or equipment damage.

Regulation can be provided by distributed resources, such as the ones mentioned in PJM (2013):

- **Generation:** Steam, Hydroelectric, Combustion Turbines, Combined Cycle.
- **Grid Energy Storage:** Batteries, Flywheels.
- **Behind-the-meter Storage:** Water Heaters, Batteries, Plug-in Hybrid EVs

- **Demand Response:** Variable Speed Pumps, Ceramic Thermal Storage.

These diverse resources have different capabilities of contributing to the regulation of the system. Thus, FERC's Order 755 establishes payment methodologies meant to incentive the use of the resources that contribute the most to the network.

1.2.2.3. FERC's Order 755: Capacity and Performance-based Payments

FERC's Order 755, *Frequency Regulation Compensation in the Organized Wholesale Power Markets*, states that a two-part payment methodology must be used to compensate frequency regulation resources. The first part consists in a capacity payment, that accounts for opportunity costs for holding capacity in reserve, and the second part consists in a performance payment, that accounts for accuracy in responding to a transmission system operator's AGC signal and reflects the amount of work each resource performs in real-time (*mileage*) in response to the system operator's dispatch signal (Y. Chen et al., 2015).

Prior to Order 755, different Regional Transmission Organizations (RTOs) had different mechanisms to compensate for frequency regulation services that considered capacity but not performance. By the end of April 2012, all RTOs in USA (CAISO, ISO-NE[8], MISO[9], NYISO[10] and PJM) submitted compliance fillings to Order 755, which included two-part regulation offers with a regulating capacity offer and a regulating performance offer (that takes mileage into account) (Y. Chen et al., 2015). However, different schemes were proposed for measuring performance by the different RTOs. This review will first focus on PJM's definitions, because it is the RTO that leads the fast-frequency regulation market (Energy Storage Update, 2016) (PJM is also the largest wholesale electricity market in the world, serving 61 million people), and will then describe CAISO's definitions because they are relevant for the Los Angeles Air Force Base (LAAFB) project and for the proposed research, as it will be explained later. Nonetheless, there are recent works

[8]New England ISO
[9]Mid-continent ISO
[10]New York ISO

describing how MISO (Y. Chen et al., 2015) and other authors (Papalexopoulos & Andrianesis, 2014) have designed their own performance-based payments.

PJM's AS market Performance Metrics: PJM proposed to calculate an accuracy score that reflects a regulation resource's accuracy in response to PJM's dispatch signal (Y. Chen et al., 2015). The Performance Score is a weighed average of the *Accuracy Score*, the *Delay Score* and the *Precision Score* (PJM, 2013):

(i) **Accuracy:** Maximum statistical correlation (ρ) or degree of relationship between two time series: the control signal $S(t)$ and regulating unit's response $R(t)$. It is a 5 minute rolling correlation with 10 second granularity, and it is re-calculated with time shift (δ) from 10 seconds up to 5 minutes.

$$\text{Accuracy Score} := \max_{\delta \in [0,5 \text{ min}]} \rho_{S(t\in[0,5 \text{ min}]),R(t\in[\delta,\delta+5 \text{ min}])} \tag{1.1}$$

(ii) **Delay:** Time delay between the control signal $S(t)$ and point of highest correlation from the Accuracy calculation (δ). Calculated together with Accuracy, over a 5 minute period with a 10 second propagation delay.

$$\text{Delay Score} := \frac{5 \text{ min} - \delta}{5 \text{ min}} \tag{1.2}$$

(iii) **Precision:** Difference in the energy provided versus the energy requested by the regulation signal, which is calculated as the difference between the areas under the curve for the control signal $S(t)$ and the regulating unit's response $R(t)$, calculated hourly with a 10 second granularity.

$$\text{Precision Score} = 1 - \left| \frac{\sum_{t\in[0,1 \text{ hour}]} |R(t)| - \sum_{t\in[0,1 \text{ hour}]} |S(t)|}{\sum_{t\in[0,1 \text{ hour}]} |S(t)|} \right| \tag{1.3}$$

Alternatively, it could be calculated as a function of the resource's regulation capacity assignment or Capability C, which is a constant measured in MW. In order to do this, the average deviation of $R(t)$ from $S(t)$ is calculated (both time

series' length is n).

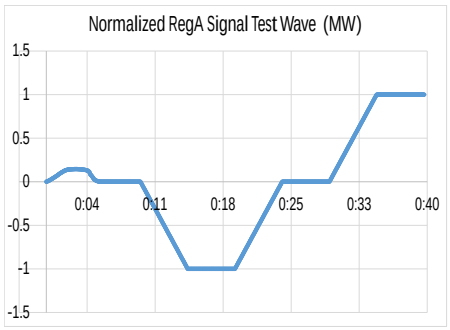
$$\text{Precision Score} = 1 - \frac{1}{n} \sum_{t \in [0,1 \text{ hour}]} \left| \frac{R(t) - S(t)}{\text{C}} \right| \tag{1.4}$$

One of the major innovations PJM introduced into the performance-based payments for AS is the distinction between different kinds of resources when it comes to send an AGC signal, so that each resource receives a signal that is well suited for its physical characteristics. On one hand, RegA signal is a low pass filter of PJM regulation requirements or Area Control Error (ACE)[11] sent to traditional ramp-limited resources, such as steam and combined cycle units (Figure 1.4a), and on the other hand RegD signal is a high pass filter of PJM ACE sent to dynamic or fast-response regulating resources, such as batteries or flywheels, that are only limited by their energy capacity (Figure 1.4b). While RegA signal may remain at full raise or lower for extended periods, RegD signal is balanced around zero in the short term. There are some resources that may qualify for both, such as hydroelectric and combustion engines (Xiao et al., 2014).

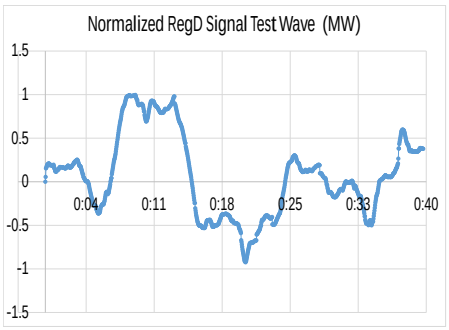
The *Mileage* concept becomes relevant because of the distinction between RegA and RegD. It represents the absolute sum of movement[12] of the regulation signal in a given time period (PJM, 2013), so a resource that follows the dynamic regulation signal (RegD) will show more Mileage than other that follows the traditional signal (RegA). Thus, there is a *Benefit Factor* that increases the incentives for dynamic resources.

[11] The concept of ACE is a real-time value that englobes the Interchange Error and the Frequency Bias of a particular ISO in an interconnected system, and represents to a particular ISO in an interconnection what frequency represents for the whole system.

[12] For a time series x_t, the absolute sum of movement during a time period $\mathcal{T}$ is understood as $\sum_{t \in \mathcal{T}} |x_t - x_{t-1}|$.



(a) PJM Normalized RegA Self Test Signal, sampled every 4 seconds for 40 minutes



(b) PJM Normalized RegD Self Test Signal, sampled every 4 seconds for 40 minutes

Figure 1.4. Different types of PJM regulation signals.

Resources that qualify for providing regulation in PJM must pass the Regulation Qualification Test, which consists of meeting the following criteria:

- Pass three consecutive tests with a Performance Score above 75%.
- Follow RegA or RegD signal for 40 minutes. No more than one test scheduled in a day.
- Resources can qualify for RegA and RegD signal by completing additional tests.

Disqualification threshold is based on a 100 hour rolling average with an average performance score below 50%. Once disqualified, resources can re-qualify, but their Performance Score starts a new rolling average.

In PJM's market clearing process for AS there are three major inputs that determine the Capability Clearing Price and the Performance Clearing Price. These inputs are the Adjusted Capability Cost, the Adjusted Performance Cost and the Adjusted Lost Opportunity Cost, for each resource that participates in the AS market, and are calculated as follows (PJM, 2013):

$$\text{Adjusted Capability Cost } (\$) := \left(\frac{\text{CO}}{\text{BF}}\right) \cdot \left(\frac{\text{C}}{\text{HPS}}\right) \tag{1.5}$$

$$\text{Adjusted Performance Cost } (\$) := \left(\frac{\text{PO}}{\text{BF}}\right) \cdot \left(\frac{\text{C}}{\text{HPS}}\right) \cdot (\text{MR}) \tag{1.6}$$

$$\text{Adjusted Opportunity Cost } (\$) := \left(\frac{\text{ELO}}{\text{BF}}\right) \cdot \left(\frac{\text{C}}{\text{HPS}}\right) \tag{1.7}$$

Where:

- CO: Capability Offer ($/MW)
- BF: Benefit Factor
- C: Capability (MW)
- HPS: Historic Performance Score
- PO: Performance Offer ($/$\Delta$MW)
- MR: Mileage Ratio (ΔMW/MW)
- ELO: Estimated Lost Opportunity ($/MW)

Notes: *Historic Performance Score* is the average of the last 100 hours of performance scores, *Benefit Factor* translates fast moving resources MWs into traditional MWs and *Mileage Ratio* is the 30 day average of historical mileage ratio of the offered resource signal type.

Then, the regulation market clearing works as follows. A Rank Price is calculated for each resource, by summing the costs described in Equations 1.5, 1.6 and 1.7 and dividing that value by the Capability (MW). Afterwards, the resources are in increasing order by their Rank Price, and they provide regulation MWs until the following constraint is met:

$$\text{Regulation Capability Requirement (MW)} \leq \sum_{i \in \text{Resources}} (\text{C}_i) \cdot (\text{BF}_i) \cdot (\text{HPS}_i) \tag{1.8}$$

Once the constraint is met, the last resource that contributed regulating capacity is designed as the *marginal resource*, and the clearing prices can be calculated. The Regulation Market Total Clearing Price (RMTCP) corresponds to the marginal resource's Rank Price, which is the highest among the selected resources. Similarly, the Regulation Market Performance Clearing Price (RMPCP) is equal to the marginal resource's Adjusted Performance Offer, and finally, the Regulation Market Capacity Clearing Price (RMCPC) is equal to RMTCP minus RMPCP. A concrete example of this process is described in PJM (2013).

Prior to performance-based payments, the regulation requirement in PJM was 1% of the peak/valley load forecast. With the implementation of the performance-based regulation, that percentage was reduced to 0.7% by the end of year 2012 while still meeting the performance standards, as measured by CPS1[13] and BAAL[14] (PJM, 2013; Xiao et al., 2014), thus confirming its positive impact.

CAISO's AS Market Performance Metrics: CAISO's AS market is different than PJM's. As a result, different products are traded and different metrics are used to measure performance. In CAISO, there are four types of AS products: Regulation Up, Regulation Down, Spinning Reserve and Non-Spinning Reserve. This review will focus on the resources used to control the system frequency responding to AGC signals, which are Regulation Up and Regulation Down. They are subject to a capacity requirement and a Mileage requirement, given that a minimum performance threshold is met.

CAISO's accuracy measurement is based on a weighted average of 15 minute intervals during a calendar month, using Mileage as the weight (CAISO, 2015a). If that value is less than 25% for Regulation Up or Regulation Down, the performance threshold is not met

[13]Control Performance Standards: CPS1 is a statistical measure of ACE variability and its relationship to frequency error, and is intended to provide a frequency-sensitive evaluation of how well PJM meets its demand requirements with its supply resources (PJM, 2013).

[14]Balancing Authority ACE Limit: BAAL seeks to maintain interconnection frequency by measuring ACE variations up or down that benefit or hurt system control, and adds an additional dimension that represents the amount of time the system operated within the upper and lower boundaries (PJM, 2013).

for that resource, and the resource must be re-certified at the end of the calendar quarter when the failure occurred.

The kind of resource relevant for this review is energy storage. In CAISO, non-generator resources as batteries can be certified for a maximum Regulation Up and Regulation Down capacity of 4 times the maximum energy (MWh) the resource can generate or consume in 15 minutes (1/4 hours), after issuance of a dispatch instruction (CAISO, 2015a).

The accuracy of a resource's response to CAISO's Energy Management System (EMS) signals is calculated on a 15-min basis by comparing the sum of the total deviations of the response $R(t)$ from the AGC signal $S(t)$ with the sum of the resource's AGC signal set-points, which results in a percentage, with data samples every 4 seconds (CAISO, 2015b; Y. Chen et al., 2015; Papalexopoulos & Andrianesis, 2014):

$$\text{Accuracy} := 1 - \frac{\sum_{t\in[0,15\text{ min}]} |R(t) - S(t)|}{\sum_{t\in[0,15\text{ min}]} |S(t)|} \tag{1.9}$$

However, the monthly Regulation Up and Regulation Down performance is not a direct average of the 15-minute Accuracy terms measured in Equation 1.9: each term must be weighted according to the instructed mileage of its respective 15-minute time interval (CAISO, 2014).

$$\text{ACU}_r := \frac{\sum_{i\in\text{MU}_r} \text{IUM}_{r,i} \cdot \text{CU}_{r,i}}{\sum_{i\in\text{MU}_r} \text{IUM}_{r,i}} \tag{1.10}$$

Where:

- ACU_r: Monthly average performance accuracy for Regulation Up for resource r for all market intervals.
- MU_r: the set of 15-minute intervals for which the resource r is scheduled for Regulation Up in each calendar month.

- IUM$_{r,i}$: the resource r instructed Regulation Up Mileage of 15 minute interval i.
- CU$_{r,i}$: the resource r Regulation Up performance accuracy of 15-minute interval i (Calculated as in Eq. 1.9).

Similarly, ACD$_r$ can be calculated in the same way as in Equation 1.10 (with its respective MD$_r$, IDM$_r$, CD$_r$ parameters), and represents the monthly average performance accuracy for Regulation Down for resource r, for all market intervals (CAISO, 2014).

1.2.3. Controllable Loads and Vehicle-to-Grid technology

Highly distributed loads in power systems can be aggregated and controlled to provide power system services, such as AS. In order to balance systemic and local control objectives, the authors in Callaway & Hiskens (2011) state that load control schemes must be both fully responsive and non-disruptive. The former characteristic is defined as enabling high-resolution system-level control across multiple time scales, and the latter is defined as having an imperceptible effect on end-use performance. These goals can be managed by quantifying metrics of load availability and willingness to participate in aggregated control activities, which along with temporal constraints are considered critical for any successful aggregated load control strategy. The authors explore a wide range of control architectures, and state that hierarchical control may hold the most promise, because it allows third parties to organize loads and bid them into energy and AS markets, and it allows the ISO to conceptualize load groups managed by each aggregator as individual resources, which would make it easier to fit these resources into the legacy control paradigm. Two load control applications are discussed in depth in Callaway & Hiskens (2011): TCLs and plug-in EVs. This last application is the main interest of this work.

The Vehicle-to-Grid (V2G) concept is a specific way of increasing the flexibility of power systems, by enabling the charging and discharging of EV batteries during their idle time. Unlike conventional generators, the fast bidirectional capability of EV batteries is particularly well suited for supporting the integration of large amounts of renewable

generation (Pillai & Bak-Jensen, 2009). In literature, both the concept (Mets et al., 2010; Tomić & Kempton, 2007) and impact (Göransson et al., 2010; Pillai & Bak-Jensen, 2009) of using EVs for grid stabilization have been extensively investigated, and it has been shown that providing such services can be profitable. Moreover, simulations are shown in Almeida et al. (2011) for an EV fleet providing primary frequency control during a contingency in an islanded system where the fast response of EV batteries allowed the ramping of hydro units to be smoothed and thus avoided the premature fatigue of their mechanical parts. While Tomić & Kempton (2007) shows that EVs can indeed be used to track frequency regulation signals, it does not propose a tracking mechanism. Though implementation projects have been conducted (Kempton et al., 2008), a small amount of work has been done concerning the distribution of power among aggregated resources in real time.

Recently, Juul et al. (2015) stated a convex-optimization-based method for real-time tracking of AGC signals, able to distribute the charging commands among a fleet of EVs, while achieving State of Charge (SoC) goals when EV owners leave the parking lot (which fits in the *Battery* definition stated in Petersen et al. (2013)). However, it can not handle bidirectional charging with inefficiencies. In contrast, H. Liu et al. (2013) presents a decentralized strategy for primary frequency regulation that considers bidirectional charging with inefficiencies, but it does not guarantee optimality of the charging controls.

As for the economic aspects, bidding strategies in the day-ahead market (González Vayá & Andersson, 2015) and coordination with renewable resources for EV aggregators (Haddadian et al., 2016), as well as game theory strategies for competition between EV aggregators (Wu et al., 2016) have been studied, but none of them takes into account the coordination of the control signals among the vehicles. Furthermore, in González Vayá & Andersson (2016) a fleet of EVs is used to provide balancing services for wind power, but the proposed self-scheduling algorithm works in an aggregated way as well. Regarding the profitability of owning an EV fleet that provides AS, a study case for California was presented where smart charge timing substantially reduced the cost of driving EVs, and

when AS were provided, EV owners obtained a net profit (Rotering & Ilic, 2011). Finally, Sarker et al. (2015) explored the profitability of an aggregator from trading energy and providing regulation services in wholesale markets, while reimbursing EV owners for battery degradation and cycling, concluding that the aggregator would obtain greater benefits from the latter.

1.2.4. Control Techniques

1.2.4.1. Real-Time Scheduling

Real-time scheduling consists in assigning resources to *processes* in a way that takes into account the timing requirements of the processes, and it is critical in the development of real-time systems (Herrtwich, 1990). Before 1990, real-time scheduling techniques were mainly used in computer applications, specifically in Processor Time Allocation (PTA), which involves scheduling a set of computation tasks on a single processor (Subramanian et al., 2012). However, the use of computers for other real-time applications has increased rapidly over the years, specially with the increasing popularity of microcomputers, such as Raspberry Pi or BeagleBone Black. As a consequence, real-time systems have been extensively studied.

Some important definitions must be introduced before explaining common real-time scheduling algorithms. These definitions are taken from Baruah & Goossens (2004).

- **Job**: A real-time job $j = (a, e, d)$ is characterized by three parameters - an arrival time a, an execution requirement e and a deadline d, with the interpretation that this job must receive e units of execution over the interval $[a, d)$. A real time instance J is a finite or infinite set of jobs: $J = \{j_1, j_2, \dots, \}$.
- **Schedule**: For any set of jobs J, a (uniprocessor) schedule S is a mapping from the Cartesian product of the real numbers and the set of jobs to $\{0, 1\}$:

$$S \quad : \quad \mathbb{R} \times J \rightarrow \{0, 1\}$$

with $S(t, j)$ equal to one if schedule S assigns the processor to job j at time-instant t, and zero otherwise. Hence, for all t, there is at most one $j \in J$ for which $S(t, j) = 1$.

- **Active Job**: A job $j = (a, e, d)$ in J is defined to be active in some schedule S of J at time instant t if both $a \leq t \leq d$ and $\sum_{t'=a}^{t} S(j, t) < e$. That is, an active job is one that has arrived, has not yet executed for an amount equal to its execution requirement, and has not yet had its deadline elapse.

By extending these basic definitions, real-time systems can be easily described. The most common kinds of *task* systems are *periodic* or *sporadic*, which involve adding a fourth parameter to the job definition: the *period* for the former (separation between arrival times of successive jobs), and the *minimum period*[15] for the latter. Thus, the set of jobs defined with the period parameter is called a *task* $T_i = (a_i, e_i, d_i, p_i)$, and special kinds of task systems can be described:

- **Implicit-deadline task systems:** Periodic and sporadic task systems in which $d_i = p_i$ for all tasks T_i.
- **Constrained-deadline task systems:** Periodic and sporadic task systems in which $d_i \leq p_i$ for all tasks T_i.
- **Synchronous task systems:** Periodic task systems in which the offset (arrival) of all tasks is equal.

An additional definition that is useful to describe real-time systems is *laxity*. It is defined as the largest time for which the scheduler can safely delay the start of the process before running it to completion without interrupts (Herrtwich, 1990).

According to Baruah & Goossens (2004), most single-processor scheduling algorithms operate by sorting, at each time step, the active jobs by priority and executing the one with the highest priority. Dynamic priority algorithms update the priority list between

[15]The minimum period is also known as *dwell time*.

time steps, while static priority algorithms work with a fixed list. Two dynamic priority algorithms will be reviewed:

(i) **Earliest Deadline First:** At each time-instant t schedule the job j active at time-instant t whose deadline parameter is the smallest.

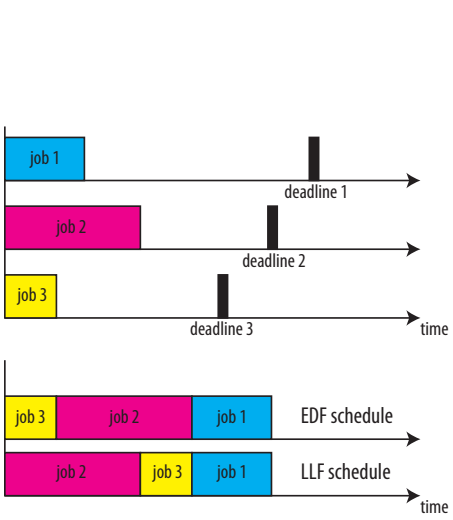
(ii) **Least Laxity First:** At each time-instant t schedule the job j active at time-instant t whose laxity state is the smallest.

Earliest Deadline First (EDF) is an optimal scheduling algorithm which is capable of achieving full processor utilization (C. L. Liu & Layland, 1973; Devi, 2003): if it is possible to schedule a given set of independent jobs such that all of them meet their deadlines, the EDF-generated schedule for that job set will meet all deadlines as well (Baruah & Goossens, 2004). However, when there is more than one processor, EDF is not always optimal. In contrast, Least Laxity First (LLF) is not only optimal when there is one processor, but also for multi-processors provided that all processes have identical *ready times*[16] (Herrtwich, 1990). LLF is superior than EDF because laxity takes into account more information, by considering not only the deadline but how long the job is going to take as well. Examples are shown in Figures 1.5a and 1.5b to compare both scheduling techniques.

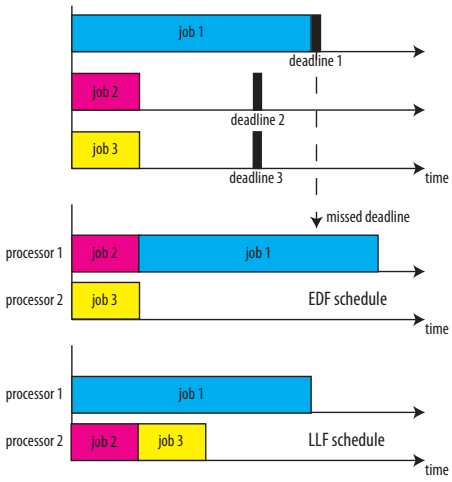
Both EDF and LLF have limitations: they are short-sighted, which means that they take decisions sequentially over time, with the information of currently active jobs but with no knowledge about future arrivals. In practice, this means that a new arrival when jobs are still being executed may cause the EDF or LLF generated schedule to be different than the optimal.

The main advantage these scheduling techniques have is that they are based on simple sorting algorithms, so they are extremely easy to implement (Subramanian et al., 2012). This makes them an ideal choice for real-time control, in which the time the control decision takes to be calculated is a key constraint.

[16]Time when the job is ready to be executed.



(a) Uniprocessor scheduling. EDF and LLF schedules. Source: Herrtwich (1990)



(b) Multiprocessor scheduling. EDF is not optimal, LLF is optimal. Source: Herrtwich (1990)

Figure 1.5. Uniprocessor vs. multiprocessor scheduling.

1.2.4.2. Trajectory Following Control

Trajectory Following (also known as Trajectory Tracking or Reference Tracking) is a form of optimization-based control, which is a control strategy that is particularly useful for constrained systems. This section is based on the formulations made by Stephen Boyd from Stanford University in his lecture notes (Boyd, 2015).

In its general form, the problem consists in describing the dynamics of a system's state x_t and deciding the optimal input u_t to achieve a certain objective $f(x_t, u_t)$, for each time step t. Its simpler form is the linear time-invariant formulation, which is described next for an infinite time horizon $\mathcal{T} = \mathbb{Z}_0^+$.

(i) **Objective Function:** A function that describes the distance between the state and/or the control inputs and a reference. This function is minimized in standard notation, but can also me maximized.

$$\sum_{t \in \mathcal{T}} f(x_t, u_t)$$

(ii) **Dynamic Equation:** An equation that describes how the future state depends on the current state and the inputs.

$$x_{t+1} = Ax_t + Bu_t \quad \forall t \in \mathcal{T}$$

(iii) **Output Equation:** An equation that calculates an output that depends on the current state and the inputs.

$$y_t = Cx_t + Du_t \quad \forall t \in \mathcal{T}$$

(iv) **Feasibility constraints:** A set of constraints that describe the domain of the state x_t and the inputs u_t.

$$x_t \in \mathcal{X}, u_t \in \mathcal{U} \quad \forall t \in \mathcal{T}$$

(v) **Initial condition:** A single constraint for the initial state.

$$x_0 = z$$

The variables of this problem are the states $x_t \in \mathbb{R}^n, \forall t \in \mathcal{T}$, and the inputs $u_t \in \mathcal{R}^n, \forall t \in \mathcal{T}$. The input data of the problem is described next.

(i) Dynamics and input matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$.
(ii) Convex state and input constraint sets $\mathcal{X} \subseteq \mathbb{R}^n, \mathcal{U} \subseteq \mathbb{R}^m$ with $0 \in \mathcal{X}, 0 \in \mathcal{U}$.
(iii) Convex stage cost function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $f(0, 0) = 0$.
(iv) Initial state $z \in \mathcal{X}$.

Note: matrices A and B must describe a controllable system.

The aforementioned problem is time-invariant, which means that the whole time horizon is solved in one problem. However, real-life implementations need to solve the problem in a sequential way. The simpler way to do this is using *greedy control*: just minimize the current stage cost, ignoring the effects of the chosen input u_t for the future states, except for $x_{t+1} \in \mathcal{X}$. This is also called *myopic control* along this work. Using greedy

control typically works very poorly, and can lead to divergent objective functions for future time steps. The formal expression of greedy control is defined as follows:

$$u_t = \text{argmin}_w\{f(x_t, w)|w \in \mathcal{U}, Ax_t + Bw \in \mathcal{X}\} \tag{1.11}$$

A special case of the Trajectory Tracking problem is the *linear quadratic regulator*, that consists in using a quadratic objective function:

$$f(x_t, u_t) = x_t^T Q x_t + u_t^T R u_t, \quad Q \succeq 0, R \succ 0 \tag{1.12}$$

When written in its general form, the evaluated objective function is designed for both the states and controls to converge to 0. This problem can be solved using Dynamic Programming, as the optimal policy has been shown to be a linear state feedback, in which the terms are calculated with an algebraic Ricatti Equation (Willems, 1971).

There are two major modifications that can be made to the Trajectory Tracking problem. The first is to use a reference trajectory for the state x_t^r and/or the inputs u_t^r in the objective function, which is done by offsetting the variables by those values ($x_t - x_t^r$ instead of x_t, and/or $u_t - u_t^r$ instead of u_t). The second is adding an output y_t of the system to the objective function.

1.2.4.3. Model Predictive Control (MPC)

Also known as Receding Horizon Control (RHC). This strategy consists in solving the previously presented time-invariant problem, for a finite horizon of length N beyond the present time step t, and then using only u_t. Then, the new state of the system is calculated, and the same problem of length N is solved again. Thus, this strategy is taking into account the future trajectory of both the state and the inputs of the system when making a decision on the current input. The general form of the problem for time step t is shown next (Camacho & Bordons, 2013; García et al., 1989).

$$
\begin{aligned}
\min \quad & \sum_{\tau=t}^{t+N} f(x_\tau, u_\tau) \\
\text{s.t.} \quad & x_{\tau+1} = Ax_\tau + Bu_\tau \quad \forall \tau \in t, ..., t+N \\
& x_\tau \in \mathcal{X}, u_\tau \in \mathcal{U} \qquad \forall \tau \in t, ..., t+N \\
& x_{t+N} = 0
\end{aligned}
\tag{1.13}
$$

The variables of the problem are x_{t+1}, ..., x_{t+N}, u_t, ..., u_{t+N-1}, and the variables x_t, A, B, f, $\mathcal{X}$,$\mathcal{U}$ are known. The solution of the optimization problem is a *plan of action* for next N steps, denoted by $\tilde{x}_{t+1}$, ..., $\tilde{x}_{t+N}$, $\tilde{u}_t$, ..., $\tilde{u}_{t+N-1}$, but only $\tilde{u}_t$ is implemented. In general, there is a trade-off between the length of N and the performance of the algorithm: choosing a longer N increases the dimensionality and the computation time, but achieves a better value for the objective function. However, it is common to find a threshold for N above which the performance does not improve much.

Some variations of this problem may include replacing the *terminal constraint* $x_{t+N} = 0$ by a *terminal cost*, converting other constraints to *soft constraints* (violation penalties), or using the *plan of action* for more than one step each time. For Trajectory Following problems, the main concern is to ensure the *recursive feasibility* (for all feasible initial states feasibility is guaranteed at every state along the trajectory) and *stability* (the effect of uncertainty is bounded and the objective function's value will always converge) of the system, for which the authors usually seek sufficient conditions. An example of this is presented in Bansal et al. (2014), where the authors ensure those key features by ensuring that the trajectory the controller is following is both *periodic* and *reachable*, as explained in detail in Chu & Chen (2012) and Limon et al. (2012).

MPC is a well studied technique with solid theoretical background, for which conditions can be found so that feasibility and stability are guaranteed, but it is usually not used for high-speed real-time applications due to its computational burden. In general,

there are two main paradigms to solve MPC problems: *online MPC*, in which the optimization is made online; and *offline MPC*, in which the control action is pre-computed for each possible case and stored offline (Explicit MPC). While the former requires large processing capabilities, the latter requires large storage capacity. The gap between MPC theory and practice may be reduced by using methods that combine online and offline appproaches, with other techniques such as *warm-start* (helping the solver with an initial solution) (M. N. Zeilinger et al., 2011).

Real-time MPC is an ongoing research subject (M. Zeilinger, 2011). The main goal of real-time online MPC is to have guarantees that within the real-time constraint, a feasible solution satisfying a stability criteria for any admissible initial state is found. While the real-time constraint requires early termination of the optimization process, feasibility of the solution requires a Robust MPC formulation and stability requires using the so-called *Lyapunov constraints* (for the objective function to behave like a *Lyapunov function*) (M. N. Zeilinger et al., 2009, 2014).

As for the implementation of real-time MPC, efficient optimization methods that exploit the highly structured nature of MPC problems[17] are still being proposed (Wang & Boyd, 2010; Domahidi et al., 2012). It has also been shown that certain computational architectures can achieve high performance, even at sample rates beyond 1 MHz (Jerez et al., 2014).

The application of real-time MPC on Smart Grid is an ongoing research subject as well. Schemes for controlling deferrable loads using real-time MPC have been proposed with different focuses: for example, N. Chen (2014) seeks to reduce the variance of aggregate load with demand response in a high renewable generation context, while the work presented in Zhou & Cai (2015) describes a controller to maximize user comfort level when exploiting the flexibility of a heating ventilation and air-conditioning (HVAC) system.

[17] Most MPC problems can be described using the aforementioned set of constraints: Dynamic Equation, Output Equation, Feasibility Constraints and Initial Condition.

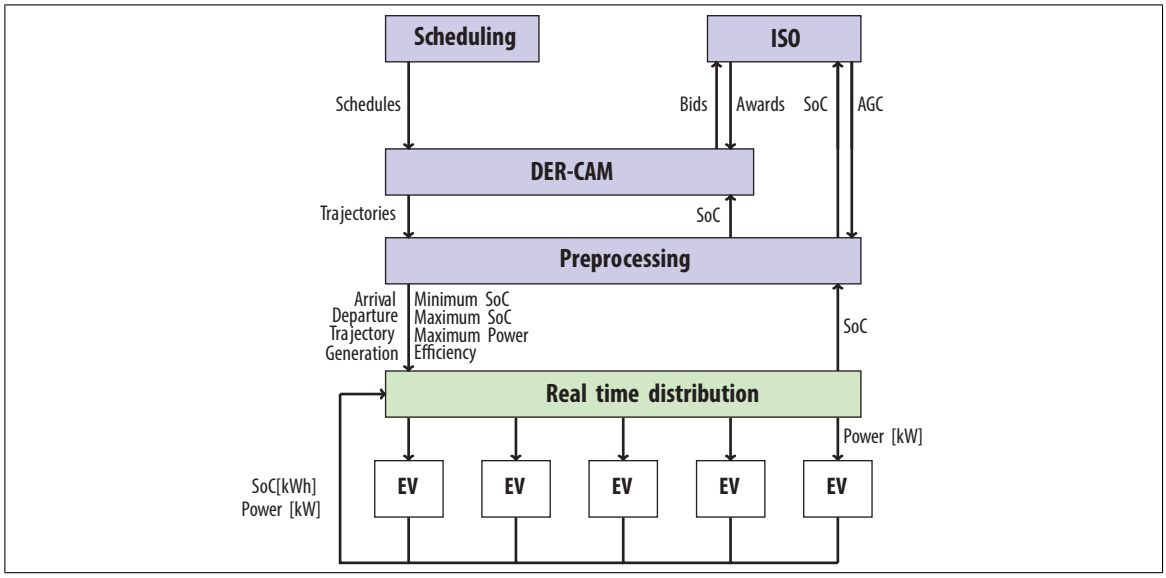


Figure 1.6. Overview of the hierarchical control framework in the LAAFB project.

1.3. Los Angeles Air Force Base (LAAFB) project

The V2G concept is being demonstrated on an operational fleet at the LAAFB. This project uses a hierarchical control framework, in which day-ahead and hour-ahead electricity market participation and charging schedule are handled by an optimization platform called DER-CAM (Distributed Energy Resources Customer Adoption Model) developed at Lawrence Berkeley National Laboratory (LBNL). DER-CAM optimizes distributed energy resources operation over economic and environmental objectives (Marnay et al., 2013). DER-CAM performs a constrained economic optimization to generate bids for bulk energy and AS markets based on the forecasts of vehicle usage by calculating the vehicles' state of charge (SoC).

However, DER-CAM's optimization is not granular nor fast enough to respond to uncertain Automatic Generation Control (AGC) signals within a few seconds, which is key for achieving an accurate response to such signals. Scheduling methods must be designed to allow real-time operation of the fleet, and these methods must distribute power among the vehicles while following an uncertain AGC signal. Figure 1.6 depicts a schematic

diagram of the control hierarchy for the LAAFB V2G project, showing the interaction of the real-time distribution developed for that project with the rest of the project. In previous work (Juul et al., 2015), a real-time controller based on convex optimization was described, and it was shown that better results can be achieved with that controller as compared to alternative suboptimal approaches. This control algorithm is currently integrated into the EV fleet management platform developed by Kisensum, Inc. (Kisensum, 2016) for the project.

1.4. Contributions

This work (1) develops a real-time charging controller to operate a fleet of EVs participating in the AS market and (2) assesses the importance of accurately describing battery efficiency and forecasted EV availability and AGC signals in allocating charging resources. We extend and refine an earlier conference paper (Juul et al., 2015) to first configure the controller to explicitly consider bidirectional efficiency. This improves regulation capacity and accuracy when following regulation signals[18] when compared to simpler approaches. Second, we develop methods to incorporate estimates of batteries' future SoC trajectories and the expectation of AGC signals into a predictive controller. Model Predictive Control (MPC) is used along with forecasts of the AGC signal to do this. We investigate the performance of these innovations (and compare them to other myopic benchmarks) via simulations under a range of scenarios for both the regulation signal and the efficiency of the batteries. More complex controllers achieve better performance, but depending on the scenario, the improvement may not be worth losing simplicity.

This work is organized as follows. Chapter 2 describes the models used for the batteries, the task concept and market participation. Chapter 3 presents the proposed controller and the benchmarks. Chapter 4 describes an MPC version of the controller. Simulation

[18] Accuracy is particularly relevant because of performance payments. The more accurate the response is, the less regulation is required by the system.

results are presented and discussed in Chapter 5, and Chapter 6 explains this work's main conclusions.

2. PROBLEM SETUP

In the proposed framework, the functional unit is the battery, and it must be characterized in terms of a set of parameters. The task concept is used to describe the characteristics of a battery when it is active, and thus can be used within its physical limits by the EV aggregator to provide frequency regulation services, as opposed to an inactive task. The laxity concept is used as well to describe each task's flexibility (Subramanian et al., 2012).

2.1. Tasks and Batteries

An aggregator coordinating a fleet of EVs in real time faces, at each time step $k \in \{1, \ldots, T\}$, the challenge of fulfilling the energy requirements associated with each EV $i \in \{1, \ldots, V\}$, together with the power requirements associated with the AGC signal. While the nominal trajectory, calculated by DER-CAM, is a parameter for vehicle i at time k, a feasible trajectory is defined as any path for the SoC of a particular EV such that both physical and scheduling requirements are fulfilled. This means that for any vehicle i at time k, its SoC must lie within the interval $[\beta_i^-, \beta_i^+]$ to respect the limits of the battery. Likewise, each vehicle arrives at time a_i with a known SoC and is scheduled to leave at time d_i with a minimum level of energy for the EV to be able to operate normally, so $E_{id_i} \in [E_i^-, E_i^+]$. For simplicity, $\beta_i^+ = E_i^+$ is assumed. There are power limits for the operation of each battery as well, so the charging/discharging (positive/negative) rate is $p_{ik} \in [m_i^-, m_i^+]$. This is illustrated in Figure 2.1.

DEFINITION 2.1. *The laxity, ϕ_{ik}, is defined as the amount of time left until vehicle i must charge at its maximum charge rate to reach its minimum SoC goal, E_i^- at time d_i.*

$$\phi_{ik} = d_i - k - \frac{E_i^- - E_{ik}}{m_i^+} \tag{2.1}$$

Table 2.1. Parameter description for task i

Parameter	Description
r_{ik}	Nominal trajectory for time step k
E_{ik}	State of Charge for time step k
β_i^+, β_i^-	State of Charge physical limits
E_i^+, E_i^-	State of Charge goal limits
a_i, d_i	Arrival/departure time
p_{ik}	Charging/discharging rate for time step k
m_i^+, m_i^-	Charging/discharging rate limits
η_i^+, η_i^-	Charging/discharging efficiency
ϕ_{ik}	Laxity for time step k

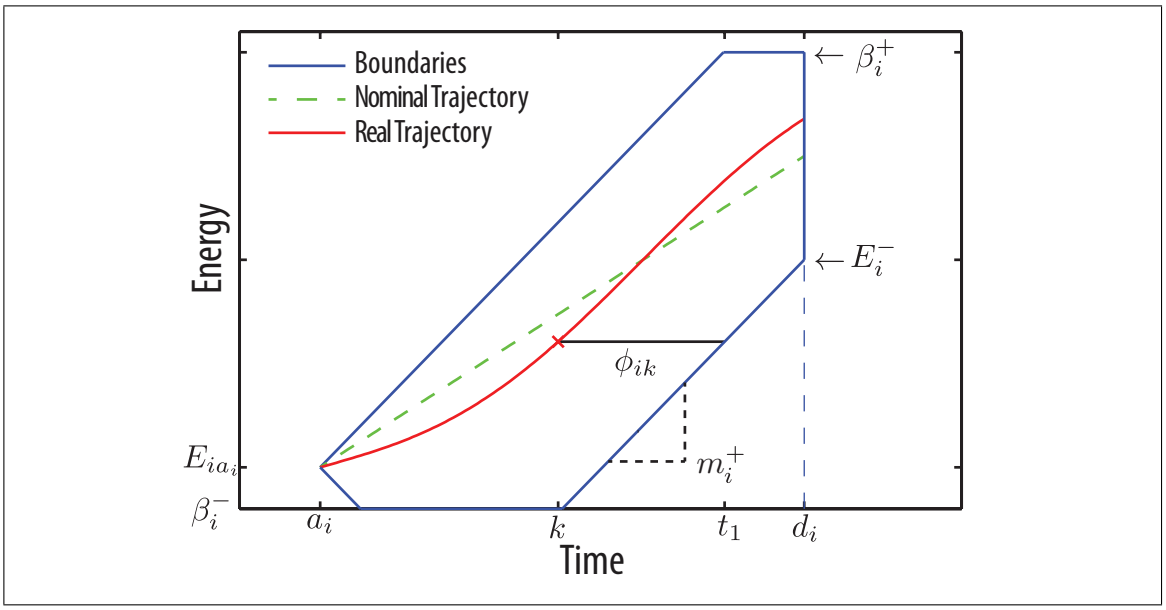


Figure 2.1. Battery model for EV i and feasible trajectories.

It is common for battery models to consider an efficiency scalar $0 < \eta < 1$ to account for the difference between the power they received and the power they were able to transform into energy for storage. In the context of bidirectional charging, this effect must be considered both ways, as it is shown in Figure 2.2.

When the power variable is positive, the batteries are being charged and when it is negative, the batteries are being discharged. Let x_{ik} be the power necessary from a source

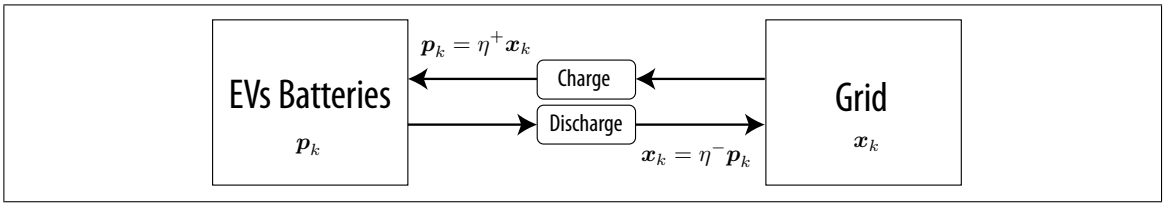


Figure 2.2. Efficiency of the batteries when charging and discharging.

to charge the battery of vehicle i with p_{ik}, and p_{ik} be the power necessary from a battery to provide the grid with x_{ik}. This can be written in a compact way:

$$p_{ik} = \Big(\frac{\eta_i^+}{2} + \frac{1}{2\eta_i^-}\Big)x_{ik} + \Big(\frac{\eta_i^+}{2} - \frac{1}{2\eta_i^-}\Big)|x_{ik}| \tag{2.2}$$

The inverse relationship is defined $F(p_{ik}) = x_{ik}$ and can be easily derived from Equation 2.2.

DEFINITION 2.2. *A task T_i can be represented by its parameters (m_i^-, m_i^+, η_i^-, η_i^+, a_i, d_i, β_i^-, β_i^+, E_i^-, E_i^+), with states E_{ik} and ϕ_{ik}, as described in Table 2.1. The index set $\mathbb{T}_k = \{i | k \in [a_i, d_i)\}$ contains the indexes of all active tasks in time step k. Each active task models an EV that is available to provide regulation.*

For notation simplicity, vectors are defined in bold symbols when referring to their components associated with active tasks: $\boldsymbol{x}_k = \{x_{ik}, i \in \mathbb{T}_k\}$, $\boldsymbol{p}_k = \{p_{ik}, i \in \mathbb{T}_k\}$, $\boldsymbol{E}_k = \{E_{ik}, i \in \mathbb{T}_k\}$, $\boldsymbol{r}_k = \{r_{ik}, i \in \mathbb{T}_k\}$, and $\boldsymbol{\Gamma}_k^{+,-} = \{\Gamma_{ik}^{+,-}, i \in \mathbb{T}_k\}$.

2.2. Limits

If a task is close to its boundaries, depending on the energy state, it is possible that the charging rate may need to be reduced. The limits that guarantee that no upper or lower boundaries are violated for task i at time step k are denoted Γ_{ik}^+ and Γ_{ik}^-, respectively. These limits include all the information needed to ensure that the SoC stays within physical limits and is able to fulfill the task's minimum energy requirements (Juul et al., 2015).

$$\Gamma_{ik}^{+} = \min\left[m_i^{+}, \frac{E_i^{+} - E_{ik}}{\Delta t}\right] \tag{2.3}$$

$$\Gamma_{ik}^{-} = \min\left[\max[\max[m_i^{-}, \frac{\beta_i^{-} - E_{ik}}{\Delta t}], (1-\phi_{ik})m_i^{+}], \Gamma_{ik}^{+}\right] \tag{2.4}$$

The additional terms of Γ_{ik}^{-} are only relevant when $\phi_{ik} < 2$. When laxity is in that range, the lower bound of p_{ik} will be pushed in order to be able to fulfill the minimum energy requirement at the time the task ceases to be active, and saturated by physical limits. Additionally, when aggregating these expressions for the fleet, the regulation capacity of the system can be calculated.

$$R_k^{+} = \sum_{i \in \mathbb{T}_k} F(\Gamma_{ik}^{+}), \quad R_k^{-} = \sum_{i \in \mathbb{T}_k} F(\Gamma_{ik}^{-}) \tag{2.5}$$

Therefore, the *feasible regulation region* for time step k will be defined as the interval $[R_k^{-}, R_k^{+}]$.

2.3. Market Participation

LAAFB project is being developed using CAISO's rules for AS, which take into account both energy and AS when optimizing the system as a whole. Therefore, the EV fleet participates, in aggregate, in two markets: the energy market and the frequency regulation market[1]. From these, every four seconds the aggregate resource receives a frequency regulation signal every four seconds, which that must be distributed among the individual vehicles. This AGC signal g_k for each time step k has two components: a fixed level of generation from the energy market award (g_k^{EM}), and a variable regulation quantity from the hourly AS market award (g_k^{FR}). The generation signal $g_k = g_k^{EM} + g_k^{FR}$ must be followed as accurately as possible for different magnitudes of the variable component.

[1]While the resource settles its electricity costs at the retail price, in order to effectively participate in frequency regulation, the EV fleet must create a baseline electricity consumption on which it will regulate around, which is done in the wholesale energy market.

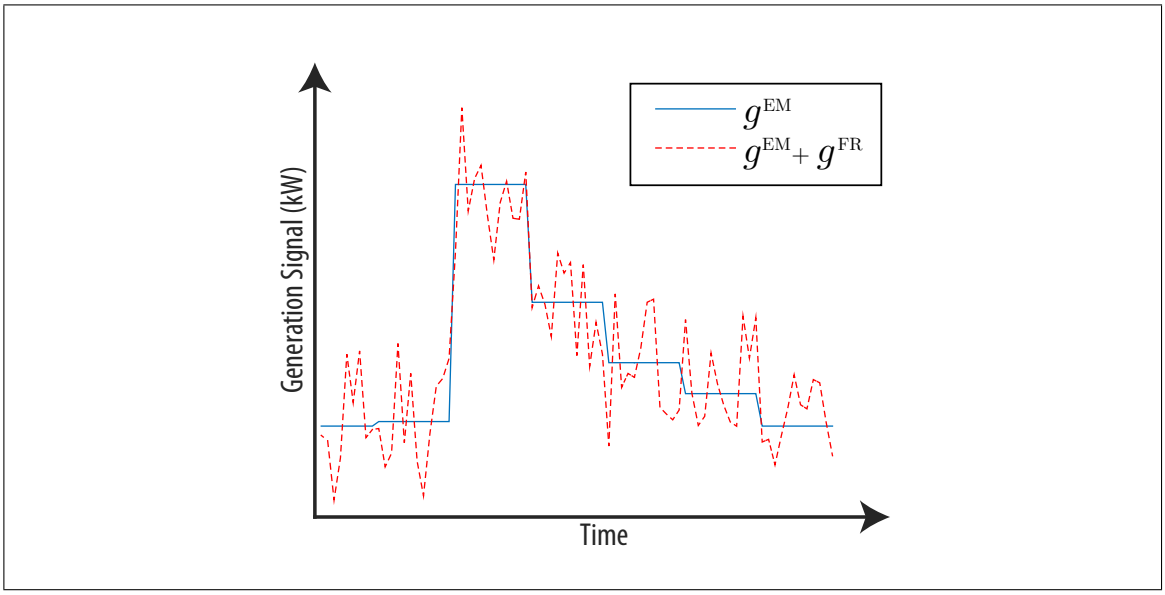


Figure 2.3. AGC signal components

The error is defined as the difference between the loads associated with the vehicles and the generation signal:

$$e_k = \sum_{i \in \mathbb{T}_k} x_{ik} - g_k \tag{2.6}$$

Thus, to maximize the performance, $|e_k|$ must be minimized. The relevant metric is the Accuracy, defined next for a time interval $\mathcal{T}$.

$$\text{Accuracy} := 1 - \sum_{k=1}^{T} |e_k| / \sum_{k=1}^{T} |g_k^{FR}| \tag{2.7}$$

This metric is tied to payments for AS providers, which according to FERC's order 755 (FERC, 2011), are proportional to capacity and performance, which was already adopted by CAISO's AS market (CAISO, 2015a).

3. MYOPIC CONTROL

A myopic or short-sighted controller is presented, along with benchmarks that use simpler approaches. The ease of implementation is a relevant subject, so it will be useful to compare simulation results.

3.1. Trajectory Following (TF)

The proposed TF controller relies on previously calculated reference trajectories $\boldsymbol{r}_k$ for the SoC of each EV in a fleet, considering their departures and arrivals, made by an external optimizer (DER-CAM (Marnay et al., 2013)). It takes as an input the reference SoC trajectories, and reschedules at each operational time the actual charging trajectory for each vehicle in order to achieve the frequency regulation requirements, given by the realization of an AGC signal that tells the fleet which instantaneous aggregated power input/output it should have. These prespecified trajectories are calculated based on optimizing the participation of the fleet in the frequency regulation market and the charging of each EV under a retail electricity tariff. The trajectories can be understood as a nominal path for the SoC of each vehicle, which enables the TF controller to incorporate information about future arrivals and departures.

For each time step k, the following convex optimization problem returns the optimal power vector, given the SoC of the previous time step $\boldsymbol{E}_{k-1}$. The proposed controller is defined as TF$(k) :=$

$$\min_{\boldsymbol{p}_k^c, \boldsymbol{p}_k^d} \quad \alpha_1 ||\boldsymbol{r}_k - \boldsymbol{E}_k||_2 + \alpha_2 |e_k| + ||\boldsymbol{\alpha}_3 \circ \boldsymbol{p}_k^c||_1 + ||\boldsymbol{\alpha}_3 \circ \boldsymbol{p}_k^d||_1$$

$$\text{s.t.} \quad \boldsymbol{E}_k = \boldsymbol{E}_{k-1} + (\boldsymbol{p}_k^c + \boldsymbol{p}_k^d)\Delta t \tag{3.1}$$

$$e_k = \sum_{i \in \mathbb{T}_k} p_{ik}^c / \eta_i^+ + p_{ik}^d \eta_i^- - g_k \tag{3.2}$$

$$\boldsymbol{\Gamma}_{k-1}^- \leq \boldsymbol{p}_k^c + \boldsymbol{p}_k^d \leq \boldsymbol{\Gamma}_{k-1}^+ \tag{3.3}$$

$$\boldsymbol{p}_k^c \geq 0, \quad \boldsymbol{p}_k^d \leq 0 \tag{3.4}$$

Where $||\cdot||_p$ stands for the ℓ_p norm[1], and $\circ$ stands for element-wise vector multiplication, also known as the Hadamard product. In the formulation of TF controller, due to the non-convex relationship between $\boldsymbol{p}_k$ and $\boldsymbol{x}_k$ (Equation 2.2), we split the power variable in the EVs' side into charging ($\boldsymbol{p}^c$) and discharging ($\boldsymbol{p}^d$) power, given that one of them is 0. We relaxed the non-convex constraint $\boldsymbol{p}_k^c \cdot \boldsymbol{p}_k^d = 0$, and added a penalty for the variables so that the solution fulfills that requirement. This penalization also achieves non-aggressive control moves, thus reducing the cycling of the batteries compared to other benchmarks.

The real-time Trajectory Following (TF) controller consists of an objective function that sums three terms with different purposes: (1) tracking the SoC trajectories, (2) following the AGC signal and (3) penalizing the power variables for feasibility, with strictly positive penalties $\alpha_1, \alpha_2, \boldsymbol{\alpha}_3$. As for the constraints, Eq. 3.1 represents the dynamics of the batteries, Eq. 3.2 represents the error in following the AGC signal and Eqs. 3.3 and 3.4 bound the decision variables.

If efficiencies η^+, η^- are 100%, the optimal solution always requires that $\boldsymbol{p}_k^c \cdot \boldsymbol{p}_k^d = 0 \quad \forall k$. When they are lower, it can be shown that the constraint also holds when a sufficient condition is fulfilled. Then, a convex relaxation based on tuning the penalties in the objective function is described in Theorem 3.1.

[1] A class of vector norms, called a p-norm and denoted $||\cdot||_p$, is defined as $||\boldsymbol{x}||_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p} \quad p \geq 1, \boldsymbol{x} \in \mathbb{R}^n$.

THEOREM 3.1. *If the penalties $\alpha_2, \boldsymbol{\alpha}_3$ are such that:*

$$\alpha_{3i} > \alpha_2 \frac{1 - \eta_i^+ \eta_i^-}{2\eta_i^+} \ \forall i,$$

the optimal solution to TF(k) will satisfy $p_{ik}^c \cdot p_{ik}^d = 0 \ \forall i \in \mathbb{T}_k$.

PROOF 3.1. *Assume that the optimal solution for some i, k is given by $p_{ik}^c = t_1 + t_2, p_{ik}^d = -t_2$ such that $t_1 > 0, t_2 > 0$ (so that $(t_1 + t_2)t_2 \neq 0$). For simplicity also assume that $\eta_i^+ = \eta^+, \eta_i^- = \eta^- \ \forall i$. When comparing the cost of the solution $(t_1 + t_2, -t_2)$ with the cost of the solution $(t_1, 0)$, the first term of the objective function, $\alpha_1 ||\boldsymbol{r}_k - \boldsymbol{E}_k||_2$, is the same in both cases. Thus, the rest of the objective function remains to be compared. Condition $p_{ik}^c \cdot p_{ik}^d = 0$ would always hold if a sufficient condition was found so that the value of the objective function is always lower for $(t_1, 0)$ than for $(t_1 + t_2, -t_2)$. When replacing, the following expression is obtained:*

$$\begin{aligned} & f(t_1, 0) < f(t_1 + t_2, -t_2) \\ & \alpha_2 \Big| \frac{t_1}{\eta^+} - g_k \Big| + \alpha_{3i} |t_1| < \alpha_2 \Big| \frac{t_1 + t_2}{\eta^+} - t_2 \eta^- - g_k \Big| + \alpha_{3i} (|t_1 + t_2| + |-t_2|) \\ & \iff \alpha_2 \Big| \frac{t_1}{\eta^+} - g_k \Big| < \alpha_2 \Big| \frac{t_1}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k \Big| + 2\alpha_{3i} t_2 \end{aligned} \tag{3.5}$$

Note that because of the assumption on t_2 and $0 < \eta < 1$, $t_2(\frac{1}{\eta^+} - \eta^-)$ is a strictly positive term. There are some cases that must be analyzed:

(i) $\frac{t_1}{\eta^+} - g_k > 0$*:*

$$\begin{aligned} & \alpha_2 \Big(\frac{t_1}{\eta^+} - g_k \Big) < \alpha_2 \Big(\frac{t_1}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k \Big) + 2\alpha_{3i} t_2 \\ & \iff 0 < \alpha_2 t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) + 2\alpha_{3i} t_2 \end{aligned} \tag{3.6}$$

Which holds because all the terms are positive.

(ii) $\frac{t_1}{\eta^+} - g_k = 0$*:*

$$0 < \alpha_2 t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) + 2\alpha_{3i} t_2 \tag{3.7}$$

Holds as well.

(iii) $\frac{t_1}{\eta^+} - g_k < 0$*:*

This case must be split into three more cases:

(a) $\frac{t_1}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k > 0$

$$\begin{aligned} -\alpha_2 \Big(\frac{t_1}{\eta^+} - g_k \Big) &< \alpha_2 \Big(\frac{t_1}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k \Big) + 2\alpha_{3i} t_2 \\ \iff 0 &< 2\alpha_2 \Big(\frac{t_2}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k \Big) + 2\alpha_{3i} t_2 - \alpha_2 t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) \end{aligned} \tag{3.8}$$

Because of the assumption,

$$2\alpha_2 \Big(\frac{t_2}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k \Big) > 0 \tag{3.9}$$

So, a condition can be obtained if the rest is also positive:

$$\begin{aligned} 2\alpha_{3i} t_2 - \alpha_2 t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) &> 0 \\ \iff \alpha_{3i} &> \alpha_2 \frac{1 - \eta^+ \eta^-}{2\eta^+} \end{aligned} \tag{3.10}$$

(b) $\frac{t_1}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k = 0$

$$-\alpha_2 \Big(\frac{t_1}{\eta^+} - g_k \Big) < 2\alpha_{3i} t_2 \tag{3.11}$$

Because of the assumption,

$$\begin{aligned} \alpha_2 t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) &< 2\alpha_{3i} t_2 \\ \iff \alpha_{3i} &> \alpha_2 \frac{1 - \eta^+ \eta^-}{2\eta^+} \end{aligned} \tag{3.12}$$

The same condition is obtained.

(c) $\frac{t_1}{\eta^+} + t_2 \Big(\frac{1}{\eta^+} - \eta^- \Big) - g_k < 0$

$$
\begin{aligned}
-\alpha_2\Big(\frac{t_1}{\eta^+} - g_k\Big) &< -\alpha_2\Big(\frac{t_1}{\eta^+} + t_2\Big(\frac{1}{\eta^+} - \eta^-\Big) - g_k\Big) + 2\alpha_{3i}t_2 \\
\iff 0 &< -\alpha_2 t_2\Big(\frac{1}{\eta^+} - \eta^-\Big) + 2\alpha_{3i}t_2 \qquad (3.13) \\
\iff \alpha_{3i} &> \alpha_2\frac{1-\eta^+\eta^-}{2\eta^+}
\end{aligned}
$$

Again, the same condition is obtained.

Therefore, if the sufficient condition is respected for every i, there is an analytical guarantee that $p^c_{ik} \cdot p^d_{ik} = 0$ will be satisfied for every i, k in the optimal solution:

$$\boxed{\alpha_{3i} > \alpha_2\frac{1-\eta^+\eta^-}{2\eta^+} \quad \forall i}$$

■

The rationale behind Theorem 3.1 is intuitive: there should be a threshold for the penalties of the control moves above which it is not optimal to charge and discharge simultaneously, because it would imply a higher cost for the objective function while not making the fleet response more accurate. If there was no penalty for $\boldsymbol{p}^c$ and $\boldsymbol{p}^d$, the controller could do double-charging ($\boldsymbol{p}^c_k \cdot \boldsymbol{p}^d_k \neq 0$) if it implied better accuracy.

It should be noted that TF controller will provide a fast and feasible solution due to its convexity. Evidence suggests that if $\boldsymbol{p}^c_k \cdot \boldsymbol{p}^d_k = 0 \quad \forall k$ constraint is included by using binary variables instead of tuning the $\boldsymbol{\alpha}_3$ penalty according to Theorem 3.1, the results of the non-convex problem are the same as in the convex problem, but computation time grows significantly.

3.2. Benchmarks

To benchmark the performance of the TF controller we consider two methods from the Processor Time Allocation (PTA) literature that have been applied to electric load scheduling (Subramanian et al., 2012), as well as version of the TF controller that assumes 100% round trip battery efficiency, all of which are applied in a sequential optimization simulation. In addition, a time-invariant benchmark was used.

3.2.1. Earliest Deadline First (EDF)

EDF creates a priority list based on the departure time of the tasks, and therefore will allow vehicles with the latest deadlines to remain at a low SoC until sufficient resources are available to charge them. We adapt this algorithm for discharging by coordinating vehicles such that those with the latest departure times are discharged first.

3.2.2. Least Laxity First (LLF)

LLF creates a merit order list sorted by laxity (see Eq. 2.1), and therefore will allow vehicles with larger laxity to remain at a low SoC until sufficient resources are available to charge them. Similarly to EDF, we adapt the algorithm for discharging such that the vehicles with the highest laxity are discharged first.

3.2.3. Trajectory Following with Approximate Battery State (TFAPPROX)

We remove the nonconvexity that results from bidirectional charging by making the approximation $p_{ik} \approx x_{ik}$ in the battery dynamics equation. However, the quality of the approximation degrades with declining efficiency.

3.2.4. Time-invariant Trajectory Following (Oracle)

We developed an *Oracle* benchmark that solves the complete run time at once. This additional benchmark provides a best-possible-performance case.

4. MODEL PREDICTIVE CONTROL

This section will explain how to extend the TF myopic controller by implementing a predictive controller. We employ model predictive control, which takes into account not only the present current state of a system, but also its forecasted states over a finite time horizon (of length N), when making a decision. The underlying motivation is that MPC should allow the algorithm to achieve better Accuracy, because it will consider the EV arrivals and departures as well as a forecast of the AGC signal when deciding how to update the SoC of the vehicles.

4.1. Trajectory Following with Model Predictive Control (TFMPC)

A first approach to use MPC with TF would be to sum up the objective function values, while interpreting the bounds $\Gamma_{ik}^{+,-}$ as functions of the SoC. Including the upper bound constraint with the future E_{ik} as a variable is not a problem, due to its concavity, but including the lower bound constraint would imply using a nonconvex expression as a lower bound:

$$p_{ik} \geq \min[\underbrace{\max[\max[m_i^-, \frac{\beta_i^- - E_{ik}}{\Delta t}], (1-\phi_{ik})m_i^+]}_{\text{convex}}, \underbrace{\Gamma_{ik}^+}_{\text{concave}}] \tag{4.1}$$

In the myopic problem, this constraint's objective was to fulfill the task's minimum energy requirements at departure, which can also be achieved by transforming the laxity part of the lower bound for power into an energy constraint, so that the MPC problem is convex. For energy constraints to work, efficiency effects must be fully considered by the MPC controller in the time steps along the forecast horizon, so that the SoC can be properly estimated. Before formulating the MPC problem, some additional definitions are necessary:

- $\hat{g}_{j|k}$ is the forecast for signal g in period j, made in time step k, so $k < j$.

- The set of time steps in the forecast horizon for time step k is $\{k+1, \ldots, k+N\}$.
- Bold symbols must include all tasks that may be active: $\{i \in \cup_{j\in\{k,\ldots,k+N\}} \mathbb{T}_j\}$.

For simulation purposes, the SoC of inactive tasks has to be updated using a previously calculated vector for the power variable. This represents the energy used by the EVs when they are not grid connected, and the values for E_i^- must be consistent with that vector. Thus, the MPC problem solved for each time step k, with feasible region $\mathcal{Z}_k$, is defined as TFMPC$(k) :=$

$$\min_{\boldsymbol{p}_j^c, \boldsymbol{p}_j^d} \quad \sum_{j\in\{k,\ldots,k+N\}} \alpha_1 ||\boldsymbol{r}_j - \boldsymbol{E}_j||_2 + \alpha_2 |e_j| + ||\boldsymbol{\alpha}_3 \circ \boldsymbol{p}_j^c||_1 + ||\boldsymbol{\alpha}_3 \circ \boldsymbol{p}_j^d||_1$$

$$\text{s.t.} \quad \boldsymbol{E}_j = \boldsymbol{E}_{j-1} + (\boldsymbol{p}_j^c + \boldsymbol{p}_j^d)\Delta t \quad \forall j \in \{k, \ldots, k+N\} \tag{4.2}$$

$$e_k = \sum_{i\in\mathbb{T}_k} p_{ik}^c/\eta_i^+ + p_{ik}^d \eta_i^- - g_k \tag{4.3}$$

$$e_j = \sum_{i\in\mathbb{T}_j} p_{ij}^c/\eta_i^+ + p_{ij}^d \eta_i^- - \hat{g}_{j|k} \quad \forall j \in \{k+1, \ldots, k+N\} \tag{4.4}$$

$$\mathcal{Z}_k = \Big\{ m_i^- \leq p_{ij}^c + p_{ij}^d \leq m_i^+, \tag{4.5}$$

$$p_{ij}^c \geq 0, \quad p_{ij}^d \leq 0, \tag{4.6}$$

$$\beta_i^- \leq E_{ij} \leq \beta_i^+, \tag{4.7}$$

$$E_{ij} \geq E_i^- - (d_i - j) m_i^+ \Delta t \Big\} \quad \forall i \in \mathbb{T}_j,\ \forall j \in \{k, \ldots, k+N\} \tag{4.8}$$

The terms in objective function of the real-time MPC controller have the same meaning as in TF. As shown in Theorem 3.1, properly tuning the penalties $\alpha_2, \boldsymbol{\alpha}_3$ is critical for satisfying the $\boldsymbol{p}_k^c \cdot \boldsymbol{p}_k^d = 0$ constraint.

The constraints consider (1) the dynamics of the batteries (Eq. 4.2), (2) the present and future error in following the AGC signal (Eqs. 4.3 and 4.4) and (3) the constraint set $\mathcal{Z}_k$, that ensures feasibility for both the power variables and the energy state (Eqs. 4.5 to 4.8).

The purpose of Eq. 4.8 is twofold. First, it is needed to ensure that the task's minimum energy requirements are fulfilled. Second, it acts as a terminal constraint to ensure the feasibility of the MPC controller, regardless of the length of the forecast horizon.

4.2. Automatic Generation Control Signal Forecast

As mentioned earlier, the generation signal g_k has a random component g_k^{FR}. As this work handles uncertainty with the MPC approach, sequential forecasts are considered for g_k^{FR} by using an ARIMA approach with the information available up to each time step. This assumes that g^{FR} is a zero-mean signal, which may not be true in real applications (when defining drive cycles, non-zero means can be compensated (Hafen et al., 2011)). To compare results, a comparison was made between simulations with (1) an ARIMA forecast and (2) a perfect forecast. More information about the forecasts can be found in Appendix C.

5. NUMERICAL RESULTS

The data set used in our numerical simulations is the same as in Juul et al. (2015), with a fleet of 18 EVs, with a maximum regulation capacity of ± 15 kilowatts [kW] per vehicle, and a run time of two days with time steps every five minutes, during which the number of available vehicles changes according to a fixed task schedule described in Appendix A. TFMPC used a forecast horizon of $N = 10$. Simulations were run using the MATLAB toolbox YALMIP (Löfberg, 2004) along with the solver Gurobi (Gurobi Optimization, Inc., 2015), following the simulation flowchart described in Appendix B. The variable component of the generation signal, g_k^{FR}, was simulated offline as an ARIMA time series, based on historical data for PJM's *regd* test signal, meant to be used for fast regulation resources such as EV batteries (PJM, 2013). For each time step, a forecast was made with the information of past realizations of the AGC signal. These forecasts were done offline as well. Input data were obtained from PJM's AS website (PJM, 2015), where normalized dynamic (*regd*) and traditional (*rega*) regulation signals are provided from seven days in May 2014. The dynamic regulation signal was used for this experiment. In terms of computation time, the simulations were run in on a 2.5 GHz Intel Core i5-3210M processor, and the TFMPC algorithm with $N = 10$ (the most computationally expensive) took always less than 0.1 seconds to be solved.

5.1. Accuracy Results

All of the proposed versions of the algorithms were tested with six different test AGC signals and the Accuracy results were averaged. These represent the performance of each algorithm. Figure 5.1 shows the results for different magnitudes of the AGC signal and different battery efficiencies ($\eta^+ = \eta^- = \eta$). Note that the AGC Magnitude value represents the a higher bound for the signal's absolute value.

Figure 5.2 shows the average Accuracy of each controller, relative to the Oracle case, which represents a higher bound for the controllers' performance.

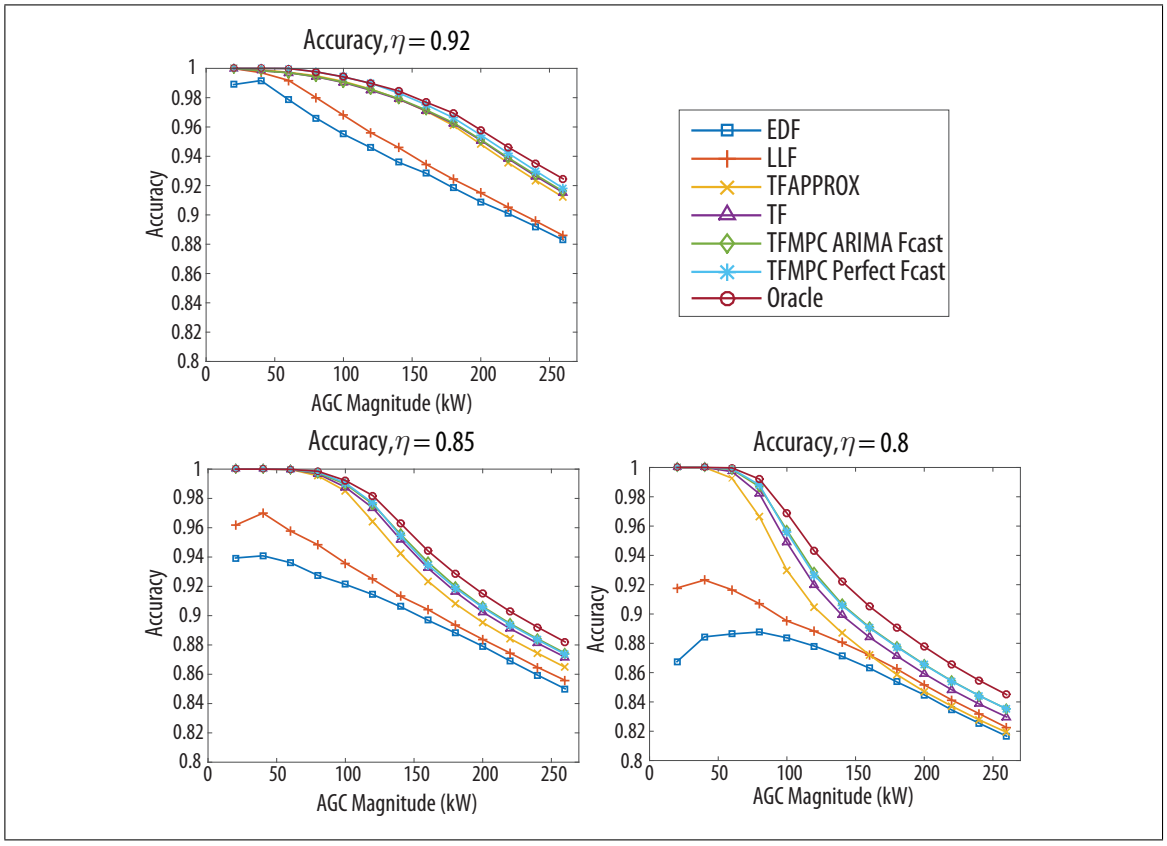


Figure 5.1. Accuracy for different algorithms, AGC signal magnitudes and battery efficiencies $\eta \in \{0.92, 0.85, 0.8\}$.

The results must be discussed separately depending on the efficiency of the batteries. When $\eta = 0.92$, the effective plugged in power capacity of each EV is ± 13.8 [kW], and sorting the performance of the algorithm and its benchmarks gives the following list: (1) Oracle, (2) TFMPC Perfect Forecast, (3) TFMPC ARIMA Forecast, (4) TF, (5) TFAPPROX, (6) LLF, and (7) EDF. The Accuracy results for LLF and EDF, the only algorithms that do not track predefined SoC trajectories, are noticeably worse than for the other algorithms, for all the magnitudes of the AGC signal; for the remainder of the algorithms the performance is similar.

When the efficiency of the batteries is decreased to $\eta = 0.85$, the effective plugged in power capacity of each EV is reduced to ± 12.75 [kW], and the performance of all the algorithms degrades but keeps the same order. In percentage terms, EDF and LLF are

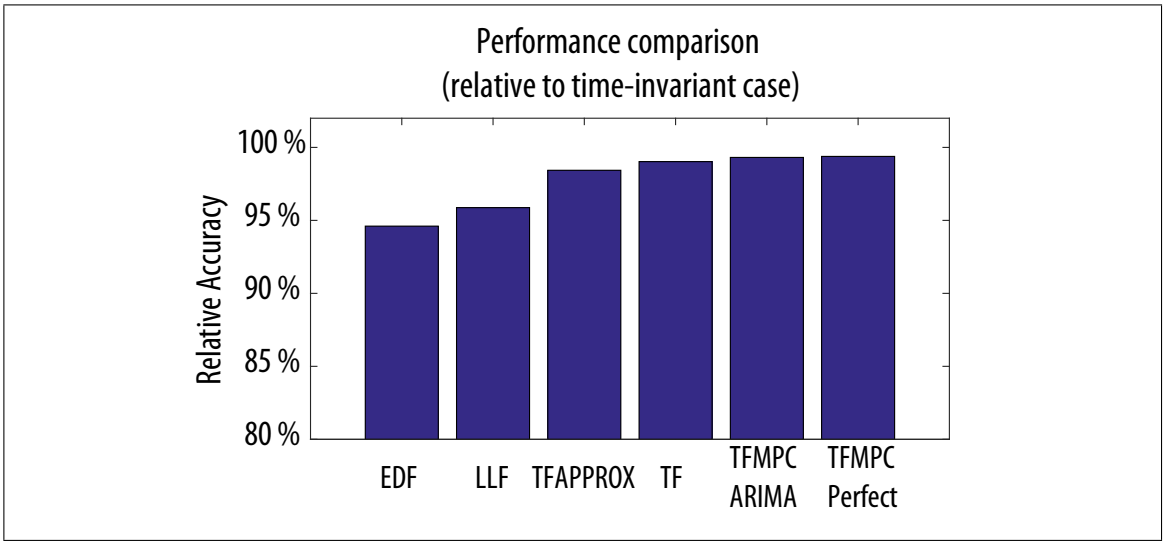


Figure 5.2. Average Accuracy for each algorithm, relative to Oracle

farther from the rest for low magnitudes of the AGC signal, and closer for high magnitudes. This happens because the benefits of trajectory following are greatest when real SoC trajectories are close to the reference trajectories – which happens to be when the AGC magnitude is smallest. On the other hand, when the AGC signal is large, the difference between real and reference trajectories is inevitably large – therefore reference trajectory following provides little benefit relative to EDF and LLF.

Finally, when $\eta = 0.8$, the effective plugged in power capacity of each EV is reduced to ± 12 [kW], and the tendency described in the former paragraph is confirmed: EDF and LLF show bad performance for low AGC magnitudes, but their performance for high magnitudes compared to the other algorithms is similar, due to the limited benefits to reference trajectory following when real trajectories are substantially different (as described in the previous paragraph). Note that both TFMPC options achieve Accuracy results that dominate over all the non-predictive algorithms. These are close to the Oracle's, but there is still some room to improve, which could be done with longer forecast horizons and more accurate forecasts.

5.2. Regulation Capacity Results

We define the feasible regulation region $[R_k^-, R_k^+]$ as the range of power in which a generation signal g_k should lie so that the fleet can provide regulation with no error, while being able to fulfill the minimum energy requirements.

Outside the feasible region, the algorithms with a myopic approach behave differently than TFMPC. For myopic algorithms, their behavior is easy to understand: in that situation, all the difference between g_k and the fleet's capacity to provide regulation results in error. In contrast, the look-ahead characteristic of the latter provides other possibility. TFMPC algorithms may choose to save battery energy in a given time step, resulting in avoidable error in the short run in favor of reducing long-run error to minimize total error. This emphasizes the relevance of properly tuning the parameters $\alpha_1, \alpha_2, \boldsymbol{\alpha}_3$ in a way that the system makes desirable decisions in the face of these trade-offs. The results for R^+ and R^- are shown in Figure 5.3.

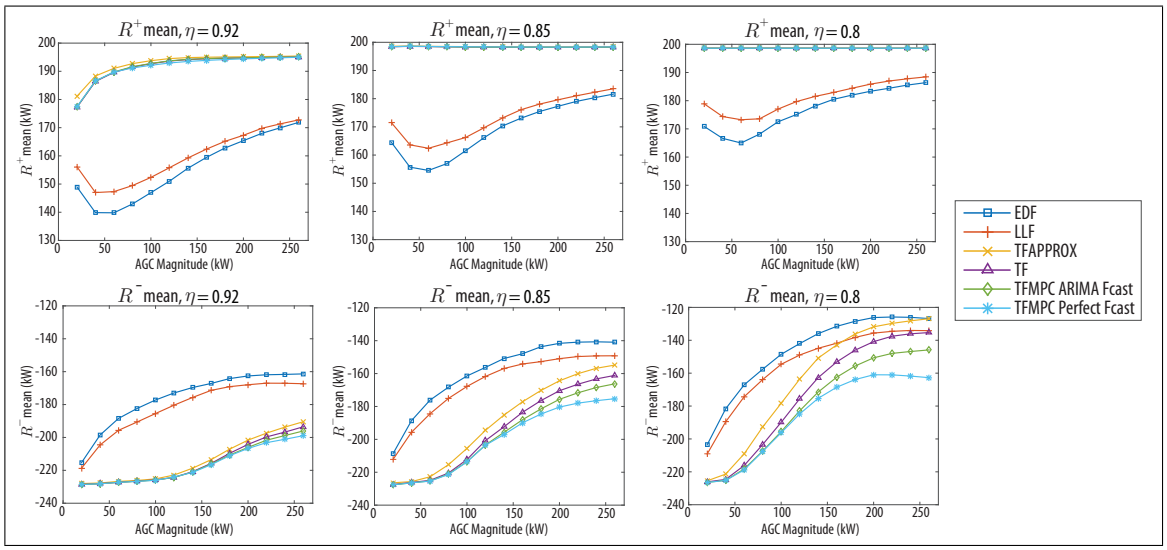


Figure 5.3. Mean R^+ and R^- for different algorithms, AGC signal magnitudes and battery efficiencies $\eta \in \{0.92, 0.85, 0.8\}$.

When sorting the algorithms by the width of the mean feasible regulation region they achieved, in general the order is the same as in the ranking shown for Accuracy (not including Oracle).

As for R^+ results (charging capacity), TF approaches achieve better levels than EDF and LLF in all cases, which is another benefit of tracking SoC trajectories. As the system receives AGC signals with higher magnitude or uses less-efficient batteries, batteries get drained, and therefore R^+ increases because there is more room for the batteries to get charged.

On the other hand, R^- (discharging capacity) is the real bottleneck of the system for large AGC signals, because the battery SoC is typically well below what is required to follow the AGC signals. Furthermore, minimum energy requirements at a task's departure also constrain the discharging capacity of each EV. Results for R^- are similar to the Accuracy results when sorting the performance of the algorithms, except for TFAPPROX when $\eta = 0.8$; the bad quality of the $p_{ik} \approx x_{ik}$ approximation directly impacts the discharging capacity of the batteries. It is clear that implementing MPC improves the capacity of the EV fleet to discharge its batteries, and this effect is magnified and therefore can be seen more clearly when the efficiency of the batteries decreases. Thus, the importance of using MPC as opposed to myopic strategies is greater when the system works closer to its limits.

5.3. Cycling Results

Here we examine the impact of the control approaches on cycling – which is believed to degrade battery state of health – by using the arc length of the SoC curves as a proxy. Figure 5.4 shows the average value of this metric over all the experiments, relative to the results of EDF.

TF approaches lead to less cycling, and we attribute this result to one key factor: by penalizing deviations from an SoC trajectory, each of the TF approaches tend to cycle all batteries in roughly the same way. This results in roughly equal distribution of ramping

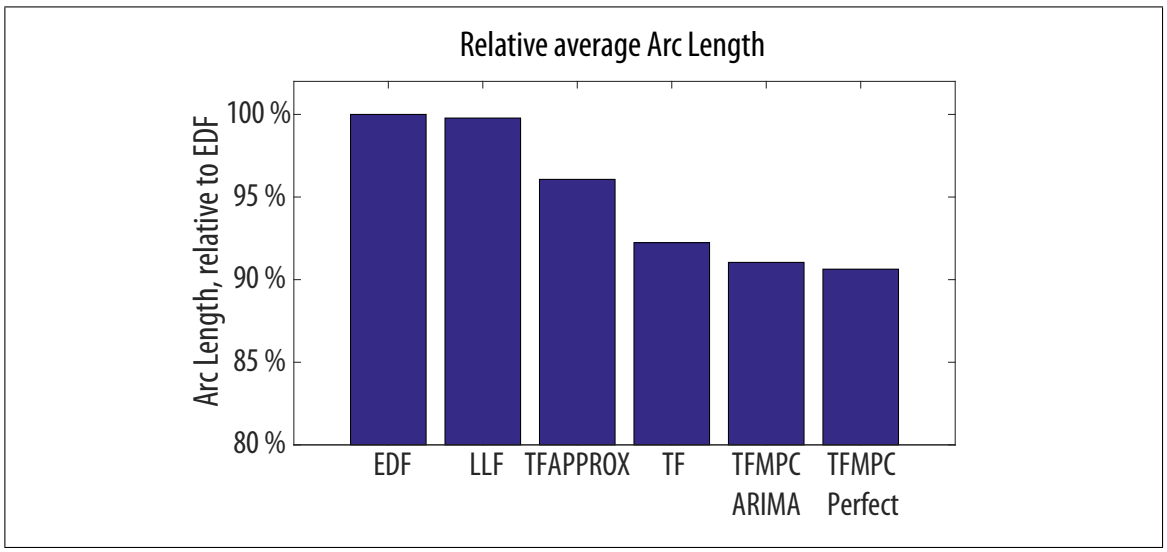


Figure 5.4. Average arc length for each algorithm, relative to EDF

across all the EVs, rather than distributing power changes to EVs in the most extreme states (as with EDF and LLF). We also see that, as one might expect, appropriately penalizing the control variables (as with TF vs TFAPPROX) results in better performance

6. CONCLUSIONS

In this work, we focus on the design of charging schedules of EVs for the provision of frequency regulation services. In particular, we propose several real-time scheduling schemes differentiated by the way of handling future information (myopic/non-myopic), the level of accuracy on following regulation signals and the resulting cycling on the batteries. A method for considering bidirectional efficiency while enabling the estimation of the future state of charge of the EVs in the fleet is provided, which allows the use of model predictive control schemes.

Extensive simulation results show the trade-off between the complexity of the controllers and their accuracy on following regulation signals: for practical implementations, both the ease of use and the performance are relevant. A key insight is that higher accuracy in following regulation signals coincides with less cycling of the batteries and, in most cases, with better regulation capacity. This highlights the importance of keeping the state of charge of the batteries away from their physical boundaries when providing frequency regulation services.

The generality of the approach enables the use of the same framework for any kind of energy battery, such as water reservoirs or HVAC loads. Other stochastic fast-response resources suitable for demand response, such as buildings or industrial processes, can be integrated into the proposed controller as well. Future work should take into account uncertainty in the EVs' arrival and departure, in which MPC along with advanced forecasting techniques can be specially valuable.

REFERENCES

Albadi, M., & El-Saadany, E. (2008). A summary of demand response in electricity markets. *Electric Power Systems Research*, *78*(11), 1989 - 1996.

Almeida, P. M. R., Lopes, J. A. P., Soares, F. J., & Seca, L. (2011, June). Electric vehicles participating in frequency control: Operating islanded systems with large penetration of renewable power sources. In *Powertech, 2011 IEEE Trondheim* (p. 1-6).

Anderson, D. L. (2009). *An evaluation of current and future costs for lithium-ion batteries for use in electrified vehicle powertrains* (Unpublished doctoral dissertation). Duke University.

Anderson, S. C. (2004). *Analyzing strategic interaction in multi-settlement electricity markets: A closed-loop supply function equilibrium model* (Unpublished doctoral dissertation). Harvard University Cambridge, Massachusetts.

Bansal, S., Zeilinger, M. N., & Tomlin, C. J. (2014, Dec). Plug-and-play model predictive control for electric vehicle charging and voltage control in smart grids. In *Proceedings of the 53rd IEEE Conference on Decision and Control (CDC)* (p. 5894-5900).

Baruah, S., & Goossens, J. (2004). Scheduling real-time tasks: Algorithms and complexity. *Handbook of scheduling: Algorithms, models, and performance analysis*, *3*.

Beck, M., & Scherer, M. (2010, April). *Overview of ancillary services.* Retrieved 01-June-2016, from `https://www.swissgrid.ch/dam/swissgrid/experts/ancillary_services/Dokumente/D100412_AS-concept_V1R0_en.pdf`

Boyd, S. (2015). *Model predictive control.* EE364b: Convex Optimization II, Stanford University Lecture. Retrieved 01-June-2016, from `http://web.stanford.edu/class/ee364b/lectures/mpc_slides.pdf`

CAISO. (2014). *Business requirements specification, pay for performance enhancement, version 1.1.* Retrieved 20-April-2016, from https://www.caiso.com/Documents/BusinessRequirementsSpecification-Pay-PerformanceYearOneDesignChanges.pdf

CAISO. (2015a). *Business practice manual for market operations, version 45.* Retrieved 20-April-2016, from https://bpmcm.caiso.com/BPM%20Document%20Library/Market%20Operations/Market%20Operations%20BPM%20V45_redline.pdf

CAISO. (2015b). *Resource performance verification, formerly g-214, operating procedure 5730, version 5.3.* Retrieved 20-April-2016, from http://www.caiso.com/Documents/5370.pdf

Callaway, D. S., & Hiskens, I. A. (2011, Jan). Achieving controllability of electric loads. *Proceedings of the IEEE*, *99*(1), 184-199.

Camacho, E., & Bordons, C. (2013). *Model predictive control*. Springer Science & Business Media.

CDEC-SIC. (2016). *Informe de definición y programación de servicios complementarios [report for defining and scheduling ancillary services].* Retrieved 13-June-2016, from http://www.cdecsic.cl/wp-content/uploads/2016/03/Informe-DPSSCC-CDEC-SIC-v20160415.pdf

Chen, M., Cho, I.-K., & Meyn, S. P. (2006). Reliability by design in distributed power transmission networks. *Automatica*, *42*(8), 1267 - 1281.

Chen, N. (2014). *Model predictive control for deferrable loads scheduling* (Unpublished doctoral dissertation). California Institute of Technology.

Chen, Y., Leonard, R., Keyser, M., & Gardner, J. (2015, Jan). Development of performance-based two-part regulating reserve compensation on MISO energy and ancillary service market. *IEEE Transactions on Power Systems*, *30*(1), 142-155.

Cho, I.-K., & Meyn, S. P. (2010). Efficiency and marginal cost pricing in dynamic competitive markets with friction. *Theoretical Economics*, *5*(2), 215–239.

Chu, Y.-C., & Chen, M. Z. (2012). Efficient model predictive algorithms for tracking of periodic signals. *Journal of Control Science and Engineering*, *2012*, 1.

Cochran, J., Miller, M., Zinaman, O., Milligan, M., Arent, D., Palmintier, B., ... Soonee, S. (2014). *Flexibility in 21st century power systems.* National Renewable Energy Laboratory (NREL). Retrieved 13-January-2016, from `http://www.nrel.gov/docs/fy14osti/61721.pdf`

Devi, U. C. (2003, July). An improved schedulability test for uniprocessor periodic task systems. In *Real-time systems, 2003. proceedings. 15th euromicro conference on* (p. 23-30).

Domahidi, A., Zgraggen, A. U., Zeilinger, M. N., Morari, M., & Jones, C. N. (2012, Dec). Efficient interior point methods for multistage problems arising in receding horizon control. In *Proceedings of the 51st IEEE Conference on Decision and Control (CDC)* (p. 668-674).

Energy Storage Update. (2016, April). *PJM leads the us fast-frequency regulation market.* Retrieved 20-April-2016, from `http://analysis.energystorageupdate.com/market-outlook/pjm-leads-us-fast-frequency-regulation-market`

European Parliament. (2003, June). Directive 2003/54/EC of the European Parliament and of the Council of 26 june 2003, concerning common rules for the internal market in electricity and repealing Directive 96/92/EC. *European Union*. Retrieved 25-June-2016, from `http://data.europa.eu/eli/dir/2003/54/oj`

FERC. (2011). *Order 755: Frequency Regulation Compensation in the Organized Wholesale Power Markets.* Retrieved 01-June-2016, from http://www.ferc.gov/whats-new/comm-meet/2011/102011/E-28.pdf

FERC. (2016). *FERC: Guide to market oversight, glossary.* Retrieved 02-February-2016, from http://www.ferc.gov/market-oversight/guide/glossary.asp

FERC, et al. (1996). *Order 888: Promoting wholesale competition through open access non-discriminatory transmission services by public utilities; recovery of stranded costs by public utilities and transmitting utilities.* Retrieved 01-June-2016, from http://www.ferc.gov/legal/maj-ord-reg/land-docs/order888.asp

García, C. E., Prett, D. M., & Morari, M. (1989). Model predictive control: Theory and practicea survey. *Automatica*, *25*(3), 335 - 348.

Gast, N. G., Le Boudec, J.-Y., Proutiere, A., & Tomozei, D.-C. (2013). *Impact of Storage on the Efficiency and Prices in Real-Time Electricity Markets* (Tech. Rep.). EPFL, Switzerland.

González Vayá, M., & Andersson, G. (2015, Sept). Optimal bidding strategy of a plug-in electric vehicle aggregator in day-ahead electricity markets under uncertainty. *IEEE Transactions on Power Systems*, *30*(5), 2375-2385.

González Vayá, M., & Andersson, G. (2016, April). Self scheduling of plug-in electric vehicle aggregator to provide balancing services for wind power. *IEEE Transactions on Sustainable Energy*, *7*(2), 886-899.

Göransson, L., Karlsson, S., & Johnsson, F. (2010). Integration of plug-in hybrid electric vehicles in a regional wind-thermal power system. *Energy Policy*, *38*(10), 5482 - 5492.

Haddadian, G., Khalili, N., Khodayar, M., & Shahidehpour, M. (2016). Optimal coordination of variable renewable resources and electric vehicles as distributed storage for energy sustainability. *Sustainable Energy, Grids and Networks*, *6*, 14 - 24.

Hafen, R. P., Vishwanathan, V. V., Subbarao, K., & Kintner-Meyer, M. C. (2011). *Requirements for defining utility drive cycles: An exploratory analysis of grid frequency regulation data for establishing battery performance testing standards*. Pacific Northwest National Laboratory. Retrieved 25-June-2016, from http://www.osti.gov/scitech/servlets/purl/1028571

Herrtwich, R. G. (1990). *An introduction to real-time scheduling*. International Computer Science Institute.

Jerez, J. L., Goulart, P. J., Richter, S., Constantinides, G. A., Kerrigan, E. C., & Morari, M. (2014, Dec). Embedded online optimization for model predictive control at megahertz rates. *IEEE Transactions on Automatic Control*, *59*(12), 3238-3251.

Juul, F., Negrete-Pincetic, M., MacDonald, J., & Callaway, D. (2015, July). Real-time scheduling of electric vehicles for ancillary services. In *Proceedings of the 2015 IEEE Power Energy Society General Meeting* (p. 1-5).

Kempton, W., Udo, V., Huber, K., Komara, K., Letendre, S., Baker, S., ... Pearre, N. (2008). *A test of vehicle-to-grid (V2G) for energy storage and frequency regulation in the PJM system* (Vol. 32). Retrieved 17-June-2016, from http://www1.udel.edu/V2G/resources/test-v2g-in-pjm-jan09.pdf

Kirby, B. (2007, July). *Ancillary services: Technical and commercial insights.* Retrieved 01-June-2016, from http://www.consultkirby.com/files/Ancillary_Services_-_Technical_And_Commercial_Insights_EXT_.pdf

Kisensum. (2016). *Connecting electric vehicles and storage to the smart grid.* Retrieved 13-January-2016, from http://www.kisensum.com/#home

Kuzle, I., Bosnjak, D., & Tesnjak, S. (2007, June). An overview of ancillary services in an open market environment. In *Control automation, 2007. MED '07. mediterranean conference on* (p. 1-6).

Limon, D., Alamo, T., de la Pea, D., Zeilinger, M., Jones, C., & Pereira, M. (2012). MPC for tracking periodic reference signals. *IFAC Proceedings Volumes*, *45*(17), 490 - 495.

Liu, C. L., & Layland, J. W. (1973, January). Scheduling algorithms for multiprogramming in a hard-real-time environment. *J. ACM*, *20*(1), 46–61.

Liu, H., Hu, Z., Song, Y., & Lin, J. (2013, Aug). Decentralized vehicle-to-grid control for primary frequency regulation considering charging demands. *IEEE Transactions on Power Systems*, *28*(3), 3480-3489.

Löfberg, J. (2004). Yalmip : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the cacsd conference.* Taipei, Taiwan. Retrieved from `http://users.isy.liu.se/johanl/yalmip`

Makarov, Y. V., Loutan, C., Ma, J., & de Mello, P. (2009, May). Operational impacts of wind generation on california power systems. *IEEE Transactions on Power Systems*, *24*(2), 1039-1050.

Marnay, C., Chan, T. W., DeForest, N., Lai, J., MacDonald, J., Stadler, M., ... Reid, E. (2013). Los angeles air force base vehicle to grid pilot project. ECEEE.

Mets, K., Verschueren, T., Haerick, W., Develder, C., & De Turck, F. (2010, April). Optimizing smart energy control strategies for plug-in hybrid electric vehicle charging. In *Proceedings of the network operations and management symposium workshops (noms wksps), 2010 IEEE/IFIP* (p. 293-299).

Meyn, S., Negrete-Pincetic, M., Wang, G., Kowli, A., & Shafieepoorfard, E. (2010, Dec). The value of volatile resources in electricity markets. In *Proceedings of the 49th IEEE conference on Decision and Control (CDC)* (p. 1029-1036).

Nayyar, A., Negrete-Pincetic, M., Poolla, K., & Varaiya, P. (2014). Duration-differentiated services in electricity. *Computing Research Repository (CoRR)*, *abs/1404.1112*. Retrieved from `http://arxiv.org/abs/1404.1112`

NERC. (2011a). *Balancing and Frequency Control: A Technical Document Prepared by the NERC Resources Subcommittee.* Retrieved 2-December-2013, from http://www.nerc.com/docs/oc/rs/NERC%20Balancing%20and%20Frequency%20Control%20040520111.pdf

NERC. (2011b). *IVGTF task 2.4 report. operating practices, procedures and tools.* Retrieved 01-June-2016, from http://www.nerc.com/files/ivgtf2-4.pdf

Nykvist, B., & Nilsson, M. (2015). Rapidly falling costs of battery packs for electric vehicles. *Nature Climate Change*, *5*(4), 329–332.

Papalexopoulos, A. D., & Andrianesis, P. E. (2014, Jan). Performance-based pricing of frequency regulation in electricity markets. *IEEE Transactions on Power Systems*, *29*(1), 441-449.

Petersen, M. K., Edlund, K., Hansen, L. H., Bendtsen, J., & Stoustrup, J. (2013, June). A taxonomy for modeling flexibility and a computationally efficient algorithm for dispatch in smart grids. In *Proceedings of the 2013 American Control Conference (ACC)* (p. 1150-1156).

Pillai, J., & Bak-Jensen, B. (2009). Electric vehicle based battery storages for future power system regulation services. In *Proceedings of the 5th nordic wind power conference.* Technical University of Denmark (DTU).

PJM. (2013, October). *Performance based regulation: Year one analysis.* Retrieved 20-April-2016, from https://www.pjm.com/~/media/committees-groups/committees/mic/20131009/20131009-item-07-performance-based-regulation-first-year-presentation.ashx

PJM. (2013). *PJM ancillary services - regulation.* Retrieved 2016-Jan-07, from https://pjm.adobeconnect.com/_a16103949/p7ssg501jfn/

PJM. (2015). *PJM ancillary services.* Retrieved 2016-Jan-07, from `http://www.pjm.com/markets-and-operations/ancillary-services.aspx`

Gurobi Optimization, Inc. (2015). *Gurobi optimizer reference manual.* Retrieved 25-June-2016, from `http://www.gurobi.com`

Rotering, N., & Ilic, M. (2011, Aug). Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets. *IEEE Transactions on Power Systems*, *26*(3), 1021-1029.

Sarker, M. R., Dvorkin, Y., & Ortega-Vazquez, M. A. (2015). Optimal participation of an electric vehicle aggregator in day-ahead energy and reserve markets. *IEEE Transactions on Power Systems*, *PP*(99), 1-10.

Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications: with R examples* (Second Edition ed.). Springer Texts in Statistics.

Spees, K., & Lave, L. B. (2007). Demand response and electricity market efficiency. *The Electricity Journal*, *20*(3), 69–85.

Subramanian, A., Garcia, M., Domínguez-García, A., Callaway, D., Poolla, K., & Varaiya, P. (2012, June). Real-time scheduling of deferrable electric loads. In *Proceedings of the 2012 american control conference (ACC)* (p. 3643-3650).

Tomić, J., & Kempton, W. (2007). Using fleets of electric-drive vehicles for grid support. *Journal of Power Sources*, *168*(2), 459 - 468.

Ulbig, A., & Andersson, G. (2015). Analyzing operational flexibility of electric power systems. *International Journal of Electrical Power & Energy Systems*, *72*, 155 - 164.

Wang, Y., & Boyd, S. (2010, March). Fast model predictive control using online optimization. *IEEE Transactions on Control Systems Technology*, *18*(2), 267-278.

Willems, J. (1971, Dec). Least squares stationary optimal control and the algebraic riccati equation. *IEEE Transactions on Automatic Control*, *16*(6), 621-634.

Wu, H., Shahidehpour, M., Alabdulwahab, A., & Abusorrah, A. (2016, Jan). A game theoretic approach to risk-based optimal bidding strategies for electric vehicle aggregators in electricity markets with variable wind energy resources. *IEEE Transactions on Sustainable Energy*, *7*(1), 374-385.

Xiao, Y., Su, Q., Bresler, F. S. S., Carroll, R., Schmitt, J. R., & Olaleye, M. (2014, July). Performance-based regulation model in PJM wholesale markets. In *2014 IEEE PES General Meeting* (p. 1-5).

Zeilinger, M. (2011). *Real-time Model Predictive Control* (Doctoral dissertation, ETH Zurich, ETH Zurich). Retrieved 01-June-2016, from `https://control.ee.ethz.ch/index.cgi?page=publications;action=details;id=3977`

Zeilinger, M. N., Jones, C. N., & Morari, M. (2011). Real-time suboptimal model predictive control using a combination of explicit mpc and online optimization. *IEEE Transactions on Automatic Control*, *56*(7), 1524–1534.

Zeilinger, M. N., Jones, C. N., Raimondo, D. M., & Morari, M. (2009, Dec). Real-time mpc - stability through robust mpc design. In *Proceedings of the 48th IEEE conference on decision and control (CDC)* (p. 3980-3986).

Zeilinger, M. N., Raimondo, D. M., Domahidi, A., Morari, M., & Jones, C. N. (2014). On real-time robust model predictive control. *Automatica*, *50*(3), 683 - 694.

Zhou, K., & Cai, L. (2015, May). A dynamic water-filling method for real-time hvac load control based on model predictive control. *IEEE Transactions on Power Systems*, *30*(3), 1405-1414.

APPENDICES

APPENDIX A. INPUT DATA

Table A.1. Arrival, departure and trip data used for simulations.

	Arrival 1			Arrival 2			Arrival 3			Physical Parameters				
Vehicle	a [×5 min]	d [×5 min]	E^- [kWh]	a [×5 min]	d [×5 min]	E^- [kWh]	a [×5 min]	d [×5 min]	E^- [kWh]	Initial SoC [kWh]	β^- [kWh]	β^+ [kWh]	m^+ [kW]	m^- [kW]
1	0	72	18.9	137	380	18.3	440	576	6	15	6	30	15	-15
2	0	97	11.25	141	397	13.75	450	576	5	12.5	5	25	15	-15
3	0	115	12.6	151	396	18.6	431	576	6	15	6	30	15	-15
4	0	137	14.7	168	388	14.7	428	576	6	15	6	30	15	-15
5	0	98	15.5	148	429	16.5	449	576	5	12.5	5	25	15	-15
6	0	91	16.75	121	447	14.25	460	576	5	12.5	5	25	15	-15
7	0	129	16.25	147	466	14.5	479	576	5	12.5	5	25	15	-15
8	0	114	25.55	166	454	16.8	467	576	7	17.5	7	35	15	-15
9	0	122	13.5	167	410	16.8	431	576	6	15	6	30	15	-15
10	0	129	14.7	160	378	24.15	458	576	7	17.5	7	35	15	-15
11	0	110	14.25	171	393	11.25	455	576	5	12.5	5	25	15	-15
12	0	144	10.5	185	415	12.25	450	576	5	12.5	5	25	15	-15
13	0	107	18.3	141	406	12.9	456	576	6	15	6	30	15	-15
14	0	95	12.5	156	421	10.5	440	576	5	12.5	5	25	15	-15
15	0	125	18.55	153	371	16.8	453	576	7	17.5	7	35	15	-15
16	0	101	20.3	149	412	22.4	435	576	7	17.5	7	35	15	-15
17	0	84	15.3	152	419	14.1	452	576	6	15	6	30	15	-15
18	0	101	10.5	166	421	14	459	576	5	12.5	5	25	15	-15

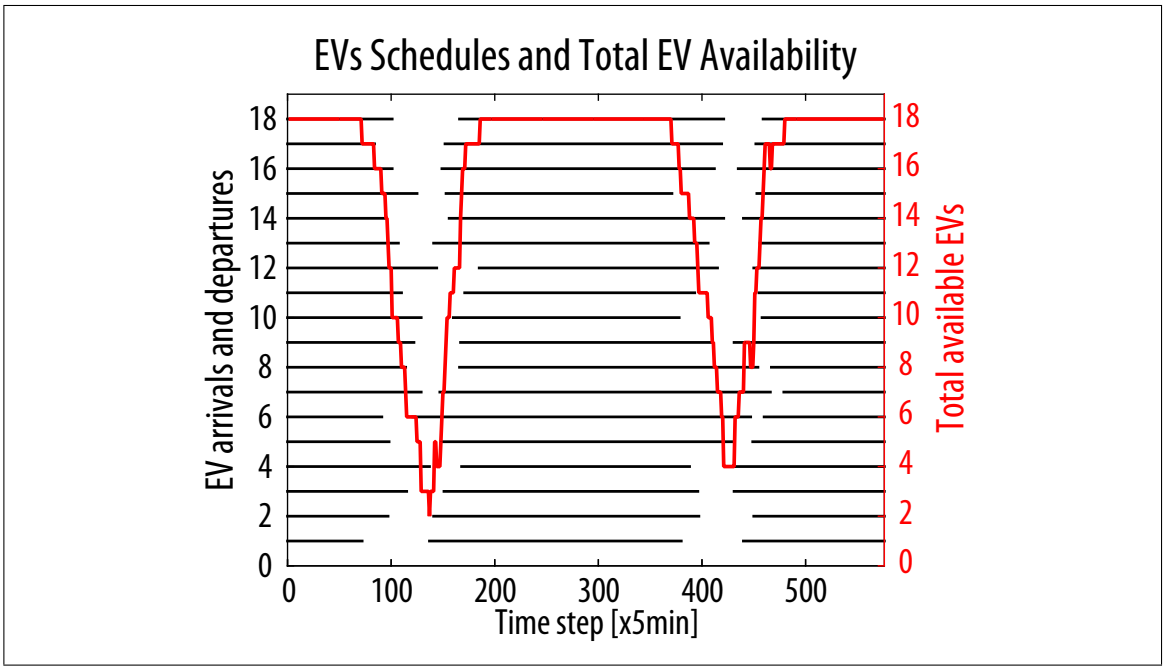


Figure A.1. EVs schedules and total EV availability

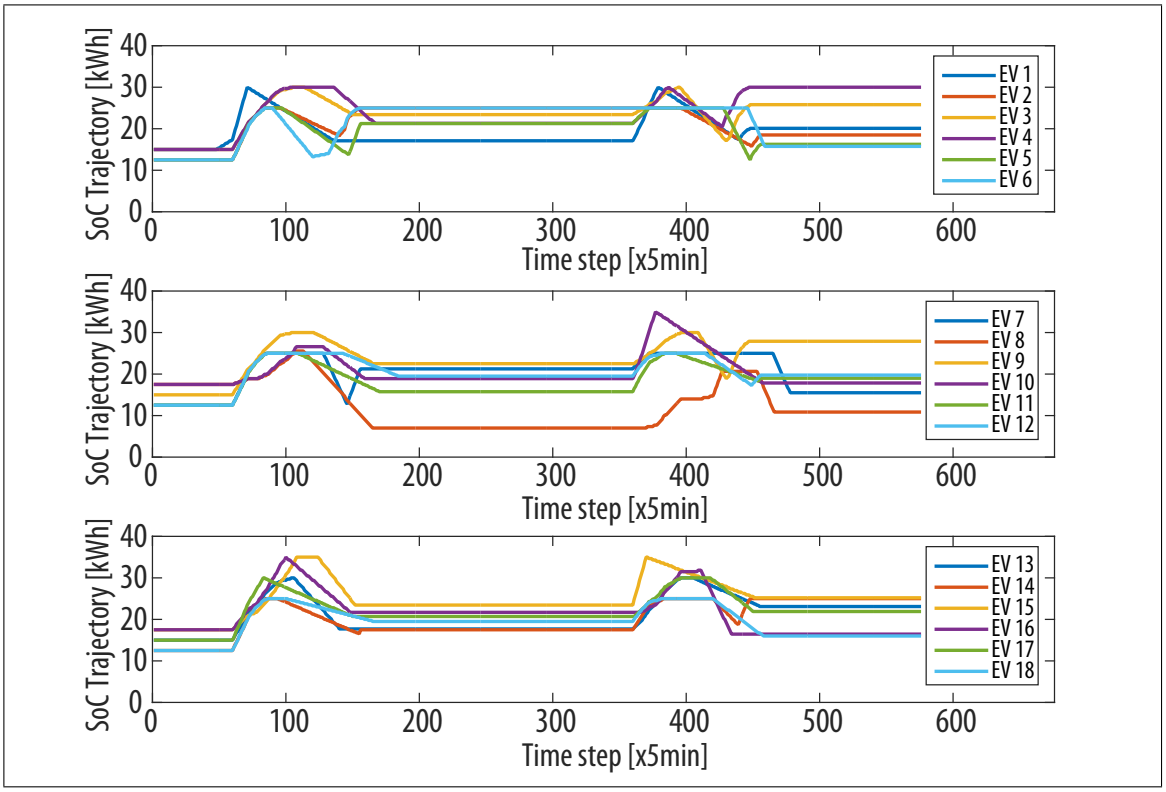


Figure A.2. Input SoC trajectories from DER-CAM

APPENDIX B. SIMULATION FLOWCHART

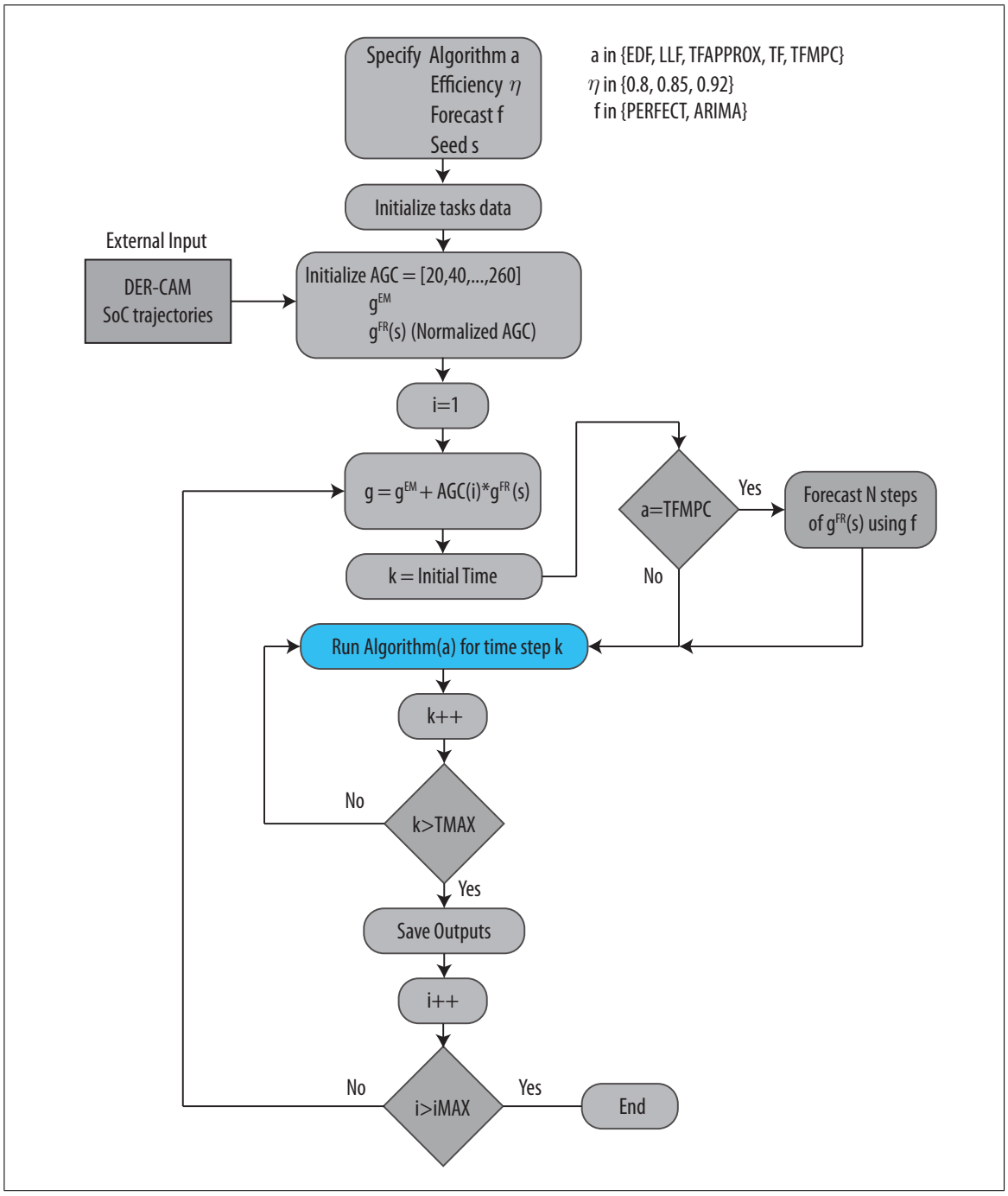


Figure B.1. Simulation flowchart

APPENDIX C. AGC SIGNAL GENERATION AND FORECASTS

ARIMA stands for Auto-Regressive Integrated Moving Average. These models were developed in the 1970s by statisticians G. Box and G. Jenkins, so they are sometimes referred as *Box-Jenkins Models*. ARIMA models are a general way of identifying a time series, under some assumptions, in terms of its own characteristics, with the objective of estimating future values of the time series.

An ARIMA model can be explained by three parameters:

- p: the amount of autoregressive terms.
- d: the amount of nonseasonal differences.
- q: the amount of moving average terms.

Thus, they are called **ARIMA(p,d,q)** models. A common simplification is that $d = 0$, so the model can be abbreviated and called an **ARMA(p,q)** model.

Some definitions are necessary to further understand ARIMA models, which were taken from Shumway & Stoffer (2006):

DEFINITION C.1. *We define the* ***backshift operator*** *by*

$$Bx_t = x_{t-1} \tag{C.1}$$

and extend it to powers $B^2x_t = B(Bx_t) = Bx_{t-1} = x_{t-2}$*, and so on. Thus,*

$$B^k x_t = x_{t-k} \tag{C.2}$$

Then, the first difference $\nabla x_t = x_t - x_{t-1}$ can be written as:

$$\nabla x_t = (1 - B)x_t \tag{C.3}$$

DEFINITION C.2. ***Differences of order*** d *are defined as*

$$\nabla^d = (1 - B)^d \tag{C.4}$$

DEFINITION C.3. *An* ***autoregressive model of order p****, abbreviated* ***AR(p)****, is of the form*

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t \tag{C.5}$$

where x_t is stationary, $\phi_1, \phi_2, ...\phi_p$ are constants ($\phi_p \neq 0$). Unless otherwise stated, we assume that w_t is a Gaussian white noise series with mean zero and variance σ_w^2. The mean of x_t in Eq. C.5 is zero. If the mean, μ, is not zero, replace x_t by $x_t - \mu$ in Eq. C.5, i.e.,

$$x_t - \mu = \phi_1(x_{t-1} - \mu) + \phi_2(x_{t-2} - \mu) + \cdots + \phi_p(x_{t-p} - \mu) + w_t \tag{C.6}$$

or write

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t \tag{C.7}$$

where $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$.

A useful form of writing this is by using the autoregressive operator,

DEFINITION C.4. *The* ***autoregressive operator*** *is defined as*

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \tag{C.8}$$

So the AR(p) model can be concisely written as:

$$\phi(B) x_t = w_t \tag{C.9}$$

DEFINITION C.5. *The* ***moving average model of order q****, or* ***MA(q)*** *model, is defined to be*

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}, \tag{C.10}$$

where there are q lags in the moving average and $\theta_1, \theta_2, ..., \theta_q$ ($\theta_q \neq 0$) are parameters. The noise w_t is assumed to be Gaussian white noise.

A useful of writing this is by using the moving average operator:

DEFINITION C.6. *The **moving average operator** is*

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots \theta_q B^q \tag{C.11}$$

So the MA(q) model can be concisely written as:

$$x_t = \theta(B) w_t \tag{C.12}$$

DEFINITION C.7. *A time series $\{x_t; t = 0, \pm 1, \pm 2, ...\}$ is **ARMA(p,q)** if it is stationary and*

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}, \tag{C.13}$$

with $\phi_p \neq 0, \theta_q \neq 0$, and $\sigma_w^2 > 0$. The parameters p and q are called the autoregressive and the moving average orders, respectively. If x_t has a nonzero mean μ, we set $\alpha = \mu(1 - \phi_1 - \cdot - \phi_p)$ and write the model as

$$x_t = \alpha + \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}. \tag{C.14}$$

Unless stated otherwise, $\{w_t, t = 0, \pm 1, \pm 2, ...\}$ is a Gaussian white noise sequence.

Finally, all the definitions needed to define an ARIMA model were stated.

DEFINITION C.8. *A process, x_t is said to be **ARIMA(p,d,q)** if*

$$\nabla^d x_t = (1 - B)^d x_t \tag{C.15}$$

is ARMA(p,q). In general, we will write the model as

$$\phi(B)(1 - B)^d x_t = \theta(B) w_t. \tag{C.16}$$

If $E(\nabla^d x_t) = \mu$, we write the model as

$$\phi(B)(1 - B)^d x_t = \alpha + \theta(B) w_t, \tag{C.17}$$

where $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$.

With the aforementioned theoretical background, the procedure used to both generate the AGC signals used in the simulations for all the controllers, and forecast these signals for the MPC controllers, can be described in Figure C.1. The resulting AGC signals that were used for the simulations are shown in Figure C.2.

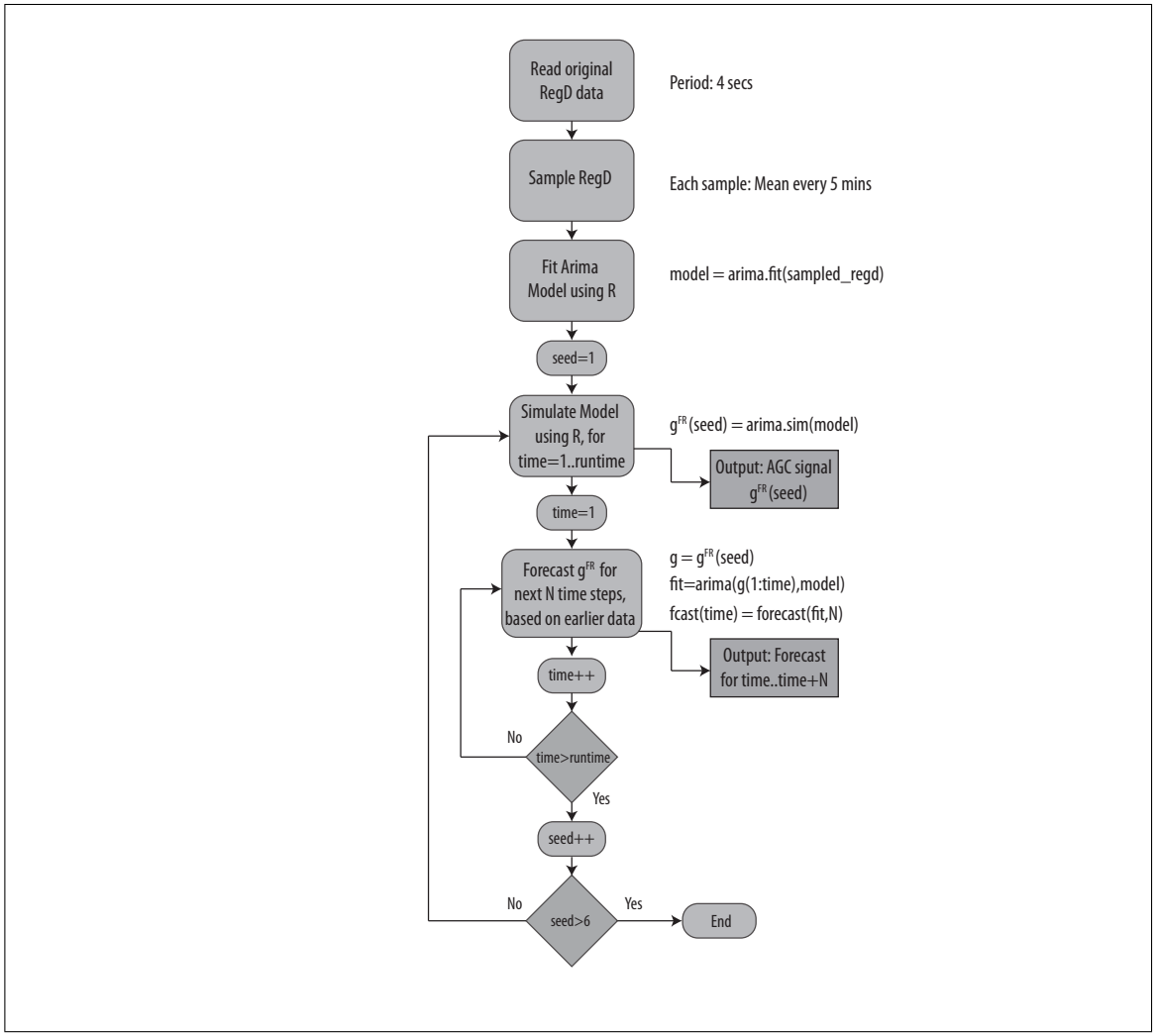


Figure C.1. AGC signal generation and forecast

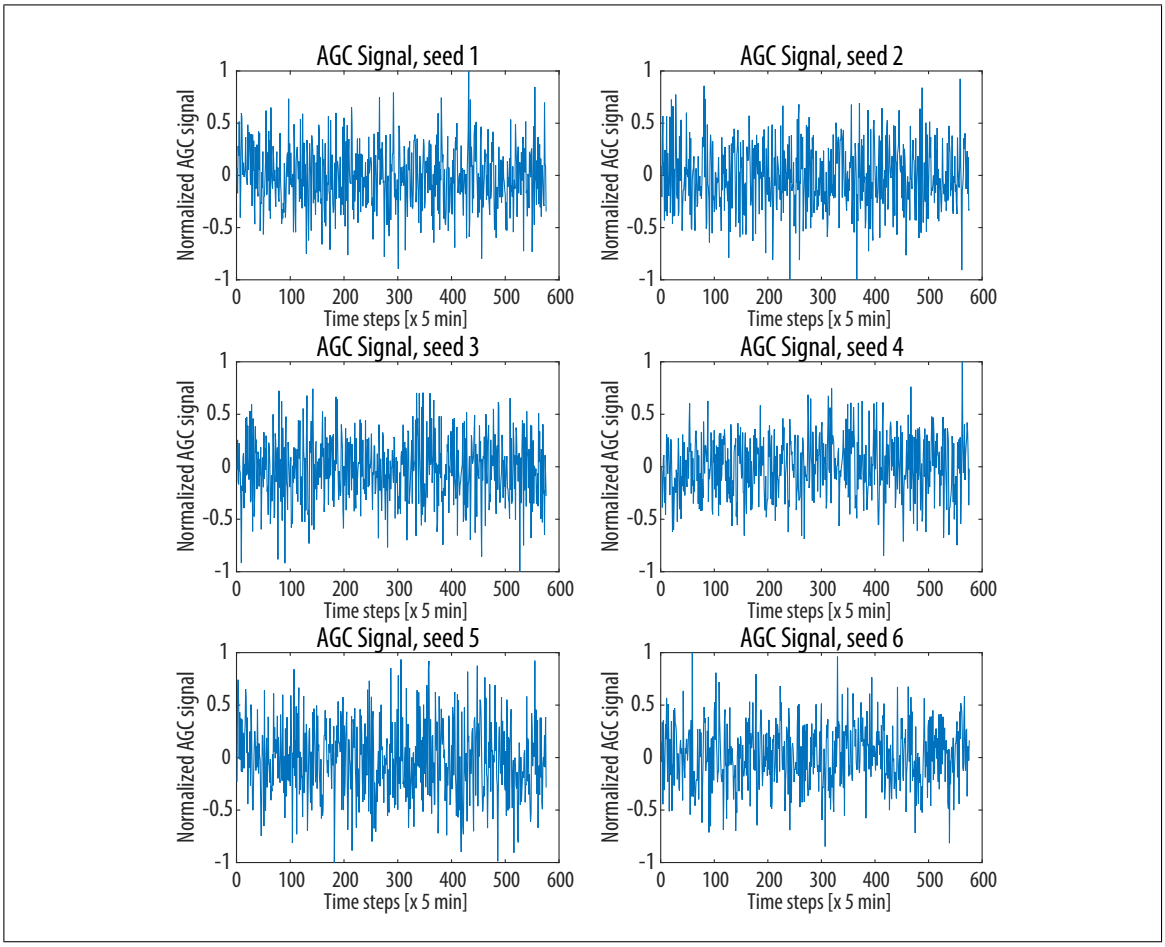


Figure C.2. Normalized AGC signals generated for each seed