

STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's Thesis in Economics
Academic Year 2022–2023

Analysis of 2022 Demand Response Events in California: Baseline Modelling using the SARIMAX-GARCH Model

Norbert Oros (42022)

Abstract

This study investigates the use of a combination of SARIMAX and GARCH models to create baselines for electricity demand, generation, and interchange in order to analyze deviations during demand response events. Demand response plays and will continue to play a crucial role in the management of the electric grid during periods of extreme imbalances. By examining demand response events in California during 2022, the study reveals that deviations in demand leading up to and during these events are not random, suggesting the presence of load shifting and effective peak management strategies. As for net generation, the hourly deviations during the pre-demand response period of 8 hours were on average positive and statistically significant, suggesting that the increased demand was largely satisfied by in-state generation capacities. The hourly deviations of the total interchange with neighboring balancing authorities during peak demand response periods were positive, which suggests that imports were systematically lower than predicted during the demand response event window. In summary, this study provides an insight into the characteristics of demand response events and highlights areas for further investigation and refinement in the context of California's electricity system.

Key words: demand response, energy modeling, electricity demand
JEL: Q41, Q43, Q48

Supervisor: Elena Paltseva
Date submitted: 15 May 2023
Date examined: 29 May 2023
Discussant: Gabriel Grundin
Examiner: Karl Wärneryd

Acknowledgements

I would like to thank my advisor, Elena Paltseva, for her guidance provided during the writing process of this thesis. I would also like to thank professor Rickard Sandberg for sharing his expertise in time series analysis that was instrumental for jump-starting this piece of work. Any errors remaining in this work are solely my own.

Table of Contents

1	Introduction	1
2	Background	3
2.1	Electricity market in California	3
2.1.1	Overview	3
2.1.2	Electricity Pricing	5
2.2	Demand Response	6
2.2.1	CAISO Grid Emergencies: Flex Alerts EEA Watch	9
2.2.2	Emergency Load Reduction Program (ELRP)	9
3	Literature Review	11
4	Methodology and Hypotheses	16
5	Data	20
6	Model Selection and Estimation	23
7	Results	27
7.1	Demand	27
7.2	Net Generation	30
7.3	Total Interchange	33
8	Discussion	36
9	Conclusion	41
	References	42
A	Appendix I	45
A.1	Demand Response Events in 2022	45
A.2	Demand for Electricity	45
A.3	Net Generation	50
A.4	Total Interchange	52
B	Appendix II	55

List of Figures

1	Western Energy Imbalance Market (Source: S&P Global Commodities Insights, S&P Global Market Intelligence)	4
2	Electricity Demand in the CAISO grid	21
3	Close-up: Weekly variation in the variables of interest.	22
4	Close-up: Daily variation in variables of interest.	23
5	17 August 2022 Observed vs. Predicted Demand	28
6	Observed vs. Predicted net generation, 17 August 2022	31
7	Observed vs. Predicted total interchange, 17 August 2022	34
8	Emergency and Demand Response Days in 2022	45
9	Residuals Diagnostics - Demand Time Series	49
10	Residuals Diagnostics - Net Generation Time Series	50
11	Residuals Diagnostics - Total Interchange Time Series	52
12	Decomposed Training Set Times Series	61
13	ACF and PACF Plots of Demand, Net Generation and Interchange	62
14	Net Generation and Interchange with Neighboring Balancing Authorities	63

List of Tables

1	Sectoral breakdown of consumers enrolled in DR programs in the CAISO grid	8
2	Descriptive Statistics	20
3	Tests for the presence of stationarity in the studied time series	24
4	Residual Diagnostics	24
5	Model Comparison - Demand for Electricity	25
6	Model Comparison - Net Generation	25
7	Model Comparison - Interchange	26
8	OLS Estimates - Demand Residuals	29
9	OLS Estimates - Net Generation Residuals	32
10	OLS Estimates - Total Interchange Residuals	35
11	Demand for Electricity (MWh) - Coefficient estimates of the best models by Mean Absolute Percentage Error	55
11	Demand for Electricity (MWh) - Coefficient estimates of the best models by Mean Absolute Percentage Error	56
12	Net Generation -Model Estimates of the three best models by AIC	56
12	Net Generation -Model Estimates of the three best models by AIC	57
12	Net Generation -Model Estimates of the three best models by AIC	58
13	Total Interchange (MWh) - Coefficient estimates of the best models by AIC	58
13	Total Interchange (MWh) - Coefficient estimates of the best models by AIC	59
13	Total Interchange (MWh) - Coefficient estimates of the best models by AIC	60

1 Introduction

In late summer 2020, California experienced an extreme heatwave that pushed the temperature in parts of the state above 40 °C. As people turned on fans and air conditioning to escape the heat, the demand for electricity exceeded the available supply and rolling blackouts ensued. This was the first time in 20 years that California was unable to meet its electricity needs (California Public Utilities Commission, 2021b). Events of that summer have been widely perceived as a harbinger of what may become a more common occurrence and authorities took note. The California Public Utilities Commission opened an Emergency Reliability Rulemaking R.20-11-003 that set out a road map for securing the reliability of the grid and preventing future blackouts (California Public Utilities Commission, 2021b). This piece of legislature instructed utilities to, among other objectives, expand existing and design new demand response programs and variable pricing plans that would increase the grid's flexibility and provide a potent tool for grid operators to manage peak load demand. Many of these programs became operational in 2022.

The economic value of demand response stems primarily from improving the efficiency of matching supply and demand on the grid and the overall market performance. Consumers may take advantage of lower prices during off-peak times, producers benefit from increased reliability of the grid and being able to deliver electricity to consumers, and grid operators become better equipped to keep the grid operational at all times and handle potential contingencies. Demand response also paves way for greater deployment of intermittent renewable resources as it helps shift the demand to times when renewable energy is plentiful (California Public Utilities Commission, 2022a).

It is important to study the impact of demand response to inform a global conversation about effective and not-so effective initiatives. California and its experience with grid contingencies and the deployment of demand response programs has the potential to provide valuable insight for designing and operating effective demand response programs worldwide. This study's focus on year 2022 is intentional as the state experienced a significant expansion of demand response programs as well as relatively high occurrence of demand response inducing events during this time period.

In terms of demand response assessment, demand response programs are traditionally evaluated on the basis of how much peak load was reduced and the total savings delivered. The evaluation entails the construction of a baseline, attributing and quantification of any deviations from the baseline. Baseline modelling in itself is an involved procedure as the target audience must be kept in mind at all times. On one hand, baselines need to be accurate and unbiased, on the other people and businesses enrolled need to be able to interpret and understand them to actively participate in DR programs (Demand Side Analytics, LLC, 2023).

What studies about demand response have in common is the universal focus on demand for electricity in perceived isolation from the other side of the equation: supply. We shall keep in mind that electricity producers and traders are also affected by demand response initiatives as demand and supply in electricity markets must match at all times. This study therefore extends the traditional approach to demand response evaluation of focusing only on demand by exploring how in-state net generation and interchange with neighboring grids changed on a grid level prior to and during demand response events in California in 2022. More specifically, I asked the following questions:

1. Did the observed time series of interest (demand for electricity, net generation and interchange) deviate from the predicted baseline in any systematic way on days when grid emergencies were declared and during periods when demand response programs were in effect?
2. Were these deviations, if present at all, meaningful statistically and practically?

The study employs the OLS regression method in combination with a SARIMAX and a GARCH model proposed by (Bollerslev, 1987), which has been considered suitable for analyzing seasonal data exhibiting non-linearity. Methodologically, the study presented herewith draws closely on Wang et al. (2012) and Garcia et al. (2005). I estimated a SARIMAX-GARCH model and generated counterfactual baselines, captured any deviations of observed time series from the predicted values and looked for any systematic patterns in these deviations prior to and during demand response events.

The results of the study show that the demand for electricity on average exceeded predicted baselines in the period of 8 hours prior to a demand response event in a non-random fashion, which indicates the presence of load shifting. The deviation from the baseline during peak demand response event hours was however negative, which would imply that the demand response resulted in peak loads below the predicted baselines.

Findings of this study bear relevance to a wide audience of policy makers, grid operators, electricity suppliers, traders and potentially residential consumers. Demand response as a form of load management is growing in popularity as grid conditions worldwide evolve in response to changing conditions.

This study opens with background information on the electricity market in California, followed by a literature review covering demand response research and energy modelling techniques. Then there are sections on methodology and model selection specifics. The study finally concludes with the presentation and discussion of results.

2 Background

2.1 Electricity market in California

2.1.1 Overview

California’s electricity market is an attractive case to study due to its size, complexity and the variety of challenges it faces. In 2021, California ranked 2nd in electricity consumption (247,249.87 GWh) and 4th in net generation (197,165.10 GWh) among all U.S. states. California is also the most populous state and there were a combined 17.7 million customers connected to the grid (U.S. Energy Information Administration, 2021a).

Electric grids in the United States are by and large operated by balancing authorities. Roughly 80% of the grid in California is operated by the California Independent System Operator (CAISO). The CAISO facilitates trading in wholesale and retail markets, assesses and allocates transmission capacities across the grid, evaluates generation capacity reserves and handles potential contingencies. In addition to the CAISO, there are a few smaller balancing authorities servicing the remaining 20% of the state’s grid (Balancing Authority of Northern California (BANC), Turlock Irrigation District (TIDC), Imperial Irrigation District (IID), Los Angeles Department of Water and Power (LADWP), Nevada Power Company (NEVP), PacifiCorp - West (PACW), Western Area Power Administration - Desert Southwest Region (WALC)) (U.S. Energy Information Administration, 2021a).

The electricity market in California is considered competitive and partially restructured, meaning that generation, transmission and distribution capacities are owned by different entities. Some residents and businesses in specific locations do not have a choice of utility providers, though, and must enter into agreement with the one provider covering their service territory. (United States Environmental Protection Agency, 2020).

1. **Retail market.** Most consumers can choose their utility provider freely, but their choice is limited by their location, hence the label “partially restructured market”. There were 235 full service ¹ utility companies operating in the state that are either investor-owned (IOUs), publicly-owned (POUs) or they are companies, whose ownership resembles a cooperative. The top five by sales in 2021 were *Southern California Edison Co.*, *Pacific Gas & Electric Co.*, *Los Angeles Department of Water & Power*, *San Diego Gas & Electric Co.*, and the *Clean Power Alliance of Southern California* (U.S. Energy Information Administration, 2021a). These five utilities together were responsible for delivering 54.0% of all electricity

¹Full service providers sell bundled electricity services (e.g., both energy and delivery) to end users (EIA, 2022)

consumed in California in 2021. The retail market as a whole is regulated by the California Public Utilities Commission (California Public Utilities Commission, 2022a).

2. **Wholesale market.** The wholesale market consists of the spot and day-ahead markets(California Independent System Operator, CAISO, n.d.-b). The state’s grid is also part of the Western Energy Imbalance Market (1), which covers large parts of the western United States and enables its participants to trade electricity across state and balancing authority borders in real time. This means that electricity generators from outside of the CAISO grid can trade in the day-ahead as well as the spot markets.

As for the state’s generation mix, California produced 49.0% of its electricity from natural gas, 17.4% from solar, 8.4% from nuclear, 7.9% from wind farms and the rest from a combination of hydro power, geothermal, biomass and coal in 2021(U.S. Energy Information Administration, 2021a). California is not self-sufficient in electricity generation and on average the state covers about 25.0% of its annual electricity consumption by importing from other states and Mexico (U.S. Energy Information Administration, 2020).

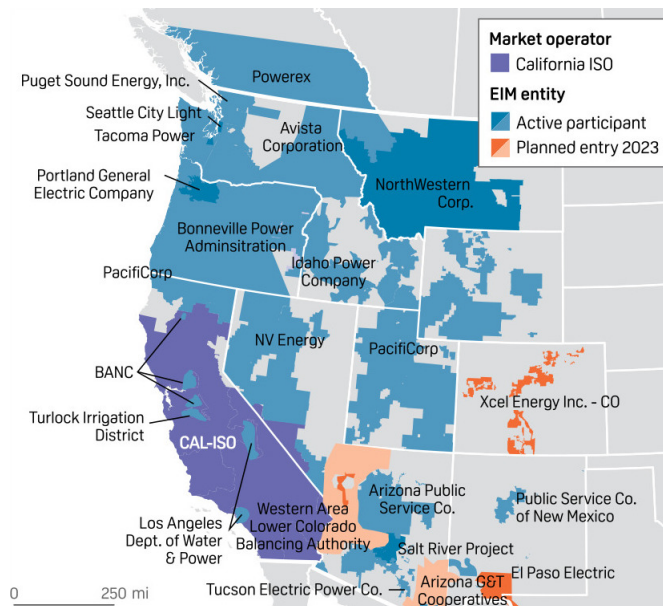


Figure 1: Western Energy Imbalance Market (Source: S&P Global Commodities Insights, S&P Global Market Intelligence)

Regulators, planners, grid operators and electricity market participants in California have been and continue to face numerous challenges that impact the grid’s reliability. Some of these challenges are linked primarily to climate change, aging infrastructure and the increasing share of intermittent renewable resources and a gradual phase-out of fossil fuels.

2.1.2 Electricity Pricing

Electricity prices that residential (retail) consumers in California pay for electricity are regulated by the California Public Utilities Commission (CPUC) (California Public Utilities Commission, 2021a). While wholesale prices on the spot and day-ahead markets fluctuate frequently, prices for residential consumers change only once a month. The CPUC receives periodic pricing proposals from public and investor-owned utilities, which it has the authority to accept or reject. These proposals state how much revenue the utility companies intend to collect from all their customers. The requested revenue amounts are then allocated across a variety of consumer classes (residential, commercial, industrial, agricultural and municipal) and prices set accordingly to generate the targeted revenue (California Public Utilities Commission, 2021a). Each utility company charges their residential customers different rates. In January 2023, average prices per kWh for selected sectors were as follows: residential (26.45 cents/kWh), commercial (21.99 cents/kWh), and industrial (16.82 cents/kWh) (U.S. Energy Information Administration, 2021b).

The default pricing plan for residential consumers in 2022 was the *Time-of-Use (T-o-U)* plan. Each 24-hour period is split into several price categories with early morning hours usually being the cheapest and late afternoon & early evening most expensive. As of March 2019, all segments of consumers (residential, commercial, industrial, municipal and agricultural) are subject to T-o-U tariffs by default (California Public Utilities Commission, 2021a).

Commercial and industrial consumers can choose from a wider range of pricing plans that are less regulated by the CPUC. In addition to the default T-o-U plan, they can choose from *critical peak pricing* or *real-time pricing* plans, too. Their descriptions are available in the section on *Demand Response* programs. In addition, customers whose peak demand exceeds a critical threshold and who draw significant amounts of electricity at once attract extra charges. This is often the case for large industrial, commercial and agricultural consumers (California Public Utilities Commission, 2021a).

For purposes of this study, it is important to keep in mind that a demand response event does not lead to explicit increases in electricity prices for residential consumers the way it does for commercial and industrial consumers on flexible pricing plans. If residential consumers respond to a call for voluntary reduction in consumption, they are not being motivated or discouraged by an increase in prices.

2.2 Demand Response

The term Demand response (DR) covers a broad range of intervention mechanisms and tools used by grid operators to mitigate extreme peak load demand and maintain balanced conditions in an electric grid (Albadi & El-Saadany, 2008; U.S. Department of Energy, 2022). A demand response commonly entails a reduction in consumption below a consumer’s normal level in times when the grid experiences an amplified strain on its capacities (California Public Utilities Commission, 2022a). The strain could be caused by natural disasters, weather extremes, or even scheduled and unscheduled shutdowns of generating capacity without adequate replacements. DR programs tend to fall within two categories: (1) Incentive-based and (2) Price-based programs. To illustrate some of the most common types of DR programs in both categories, I draw on descriptions provided by Paterakis et al. (2017). The primary trait that sets the two categories apart is the way that consumers are induced to alter their electricity consumption.

Types of Demand Response

1. **Incentive-based DR programs.** Consumers opt to receive regular payments for providing flexibility to the grid and reduce their consumption when called upon by their utility provider.
 - (a) **Direct Load Control.** This type of DR implies that consumers allow utilities to control some of their electric appliances remotely and adjust their usage based on the conditions in the grid. They do so in exchange for a regular payment. Direct Load Control programs are designed specifically for small, residential consumers.
 - (b) **Curtable Load Programs.** Such programs target medium and large consumers in the commercial and industrial sectors. Enrolled participants receive periodic payments for providing stand-by capacity. These participants must respond by reducing consumption when called upon, otherwise they may face penalties for failing to meet their contractual obligations.
 - (c) **Demand-side Bidding.** Medium and large consumers engaged in demand-side bidding may offer to reduce their consumption on an exchange in order to get compensated for the flexibility they are willing to provide to the grid operators. Small consumers can participate, too, by entering into agreements with third-party aggregators and Load Serving Entities (LSEs).
2. **Price-based DR programs.** Consumers participating in this type of DR programs are

incentivised to reduce electricity consumption by receiving signals in the form of higher prices.

- (a) **Time-of-Use (T-o-U) Tariffs.** The price of electricity for enrolled participants depends on the time of day with different rates prevailing during peak and off-peak times. Prices do not, however, reflect short-term fluctuations in the wholesale markets and are generally known to consumers in advance.
- (b) **Critical Peak Pricing.** Unlike with T-o-U tariffs, critical peak pricing allows utilities to increase the price of electricity in addition to the contracted rates during times of extreme peak load. Consumers on critical peak pricing plans can be notified of critical peak prices being charged on short notice and these elevated prices can remain in effect from a few hours to a whole day. In return, participating consumers may be offered lower prices when the conditions on the grid are more stable.
- (c) **Real-time Pricing.** Consumers pay real-time hourly prices that prevail in the wholesale markets. These prices can fluctuate widely and real-time pricing is most commonly available to commercial and industrial consumers, though utilities in some countries make such pricing plans available to residential consumers. As of spring 2023, there are no real-time pricing plans available to residential consumers in California. However, T-o-U plans are often referred to as real-time pricing plans without a direct link to the current wholesale electricity rates.

Demand response programs offer valuable benefits to consumers as well as grid operators and electricity suppliers. They allow consumers to manage their electricity consumption more efficiently and save money by taking advantage of lower off-peak prices (California Public Utilities Commission, 2022a; Lawrence Berkeley National Laboratory, 2020). They also increase the overall stability of the grid by helping reduce the strain during peak times and better matching of demand to available supply that is increasingly more intermittent due to the deployment of renewable energy resources (Albadi & El-Saadany, 2008; California Public Utilities Commission, 2022a; Gerke et al., 2020). Environmental benefits of demand response also stem from the improved ability to shift demand to times when renewable energy is plentiful, thus avoiding the deployment of more carbon-intensive power plants during peak load periods.

Another benefit of demand response is the ability to optimize investment in energy infrastructure. Dyson, Mark, James Mandel, et al. (2015) estimated that “increasing the flexibility of residential demand in the U.S. could help avoid \$9 billion per year in grid investment costs

Demand Response Statistics, California CAISO Grid 2021				
Sector	Number	Share of Total	Enrolled Capacity (MW)	Share of Total Enrolled Capacity
Residential	422 782	84.9%	256.3	24.9%
Commercial	72 984	14.7%	536.1	52.0%
Industrial	2 166	0.4%	238.1	23.1%
Total	497 932	100.0%	1,030.5	100.0%

Table 1: Sectoral breakdown of consumers enrolled in DR programs in the CAISO grid

(10% of the national forecast needs) and additional \$4 billion per year in building new generation capacity and ancillary services”. These findings underscore the economic importance of developing and implementing effective demand response programs that would cover all sectors, including residential.

Several demand response programs of each type mentioned earlier are in operation in California in 2023. These programs are “administered by California’s three regulated investor-owned utilities (IOUs), CPUC jurisdictional entities, such as Community Choice Aggregators (CCAs), or third-party entities known as Aggregators or Demand Response Providers (DR Providers)” (California Public Utilities Commission, 2022a). Participation in DR programs in general, not just in California, is by default conditional on customers meeting hardware specification of their electric meters and appliances (in case of direct load control programs) (Gagne et al., 2018). In 2021 as the last year, for which data are available, the U.S. Energy Information Administration reported that 497,932 consumers with a total capacity of 1,030.5 MW were enrolled in DR programs in California (U.S. Energy Information Administration, 2021a).

Demand response program administrators have discretion over when to declare a demand response event and not all programs and participants are called on to reduce consumption simultaneously. The administrators rely primarily on information and forecasts provided by CAISO. The grid operator may choose to send out an emergency notification to utilities if they expect a shortage of electricity. CAISO emergency notifications are tiered by severity of the anticipated or prevailing conditions on the grid. *Flex Alert* and *Energy Emergency Watch* warnings are generally issued a day in advance, while higher degree emergencies can be declared anytime (California Independent System Operator, CAISO, n.d.-a). Consumers are urged to conserve energy and suppliers are requested to provide additional capacity if available. This study focuses specifically on Flex alerts and EEA Watch declarations as they are the primary trigger for demand response programs in California. However, DR programs are deployed on an as-needed basis and not all of them come into effect simultaneously. Utilities, together with the grid operator, use discretion in deciding how much reduction to induce and by what consumer categories.

2.2.1 CAISO Grid Emergencies: Flex Alerts EEA Watch

The California Independent Grid Operator constantly monitors the situation on the grid to match the demand for electricity with available supply. It also produces short term and day-ahead forecasts that serve as a basis for load planning and preparing for any potential contingencies such as shortages. Those can be caused by weather extremes, power plant shutdowns, damages to the transmission lines, wildfires or other natural and man-made events that may threaten the functionality of the state's grid (California Independent System Operator, CAISO, n.d.-a).

When the CAISO forecast indicates that the buffer capacity falls below a critical threshold, it issues a *Flex Alert*, which is the lowest tier notification that calls on consumers to voluntarily reduce their electricity consumption in a specified time window, generally between 16:00 and 21:00 local time. Flex alerts have most commonly been issued by 15:00 a day ahead in the summer and early fall months when high air temperatures tend to drive up the demand for electricity. The alert is delivered to consumers through channels managed by their respective utility providers, demand response aggregators, social media, TV, the CAISO website, e-mail, and even social media. The alert may target selected regions of California or the state as a whole.

When a Flex alert is in effect, consumers are advised to avoid using appliances consuming significant amounts of energy (California Independent System Operator, CAISO, n.d.-a). A Flex alert can be in effect concurrently with different types of grid emergency notifications. In summary, a Flex alert is a form of preventative measure to avoid other emergency procedures such as rolling blackouts from being enforced. The first Flex alert was issued in 2000. In 2022, the alert was in effect 11 times.

In times when preventative measures turn out to be insufficient and the CAISO anticipates one or more hours to be in fact supply-deficient, it issues an Energy Emergency Alert Watch (EEA Watch) notification. This time, the grid operator instructs utilities to be on stand-by and dispatch demand response measures in amounts requested. In addition, the operator encourages power generators to make supplemental offers to cover potential shortfalls (California Independent System Operator, CAISO, n.d.-a). In 2022, the EEA Watch was declared 9 times.

2.2.2 Emergency Load Reduction Program (ELRP)

In 2021, the California Public Utilities Commission rolled out a novel demand response program named the *Emergency Load reduction Program* (ELRP), or the Power Saver Rewards (PSR) program as advertised to residential consumers. This program receives particular attention in this study primarily due to its sheer size (4,053,348 enrolled residential and commercial customers rel-

ative to the total of 17.7 million customers (all sectors, including industrial) in the whole state), relatively recent implementation (2021 for commercial consumers and 2022 for residential consumers) (California Public Utilities Commission, 2022b), and notable budget allocation size (\$ 185 million set aside for the program in 2022-2023, of which \$150 million are earmarked for incentives alone) (Southern California Edison, 2022). It is still a pivotal project to be in operation for the next 5 years and its target segments are residential and commercial consumers (California Public Utilities Commission, 2022b). The program is triggered in response to CAISO’s emergency notifications, including Flex alerts. Of the 11 days that a Flex alert was in effect in 2022, the ELRP was activated on 10 of them (Demand Side Analytics, LLC, 2023). During an ELRP demand response event, participants are urged to avoid using large appliances such as ovens or washing machines, to adjust their air conditioning to a higher temperature and postpone charging their electric vehicles to either before or after the Flex alert comes into effect (California Public Utilities Commission, 2022b).

Participants enrolled in the ELRP receive a compensation for voluntarily reducing their electricity consumption below their individually estimated baseline when a grid emergency is declared. In 2022, the rate of compensation was 2.00\$/kWh compared to \$1/kWh in 2021 when only commercial consumers were eligible (California Public Utilities Commission, 2022b). The ELRP can be activated for a maximum of 60 hours in a given year and remain for several consecutive days. The program is administered by California’s investor-owned utilities (IOUs) Pacific Gas & Electric, San Diego Gas & Electric, and Southern California Edison. These companies also facilitate communication with customers and their enrollment. It is up to the them to decide the order, in which participants should be prompted to respond to an ongoing demand response event. In addition, the IOUs decide on baseline methodologies, according to which they model customers’ normal demand baselines, estimate electricity savings, and calculate the compensation earned.

A particular point to note about the design of the program is that an estimated “half of the CA consumers were to be enrolled automatically” (California Public Utilities Commission, 2022b) and special outreach has been made to low-income and public assistance receiving communities. The reductions in electricity are voluntary and program participants do not get penalized for not delivering any energy savings. The same participants can be enrolled in other demand response programs operating in California, but there are additional terms and conditions in place for them to avoid fraud and double counting. (California Public Utilities Commission, 2022b) In its first full year of operation the ELRP delivered observably positive results as stated in the Statewide Residential Emergency Load Reduction Program Baseline Evaluation released in January 2023.

The estimated average per-event peak load demand savings were positive for two of the three utility companies operating in the CAISO grid (PG&E: 203.6MW, SCE: 496.4 MW, SDG&E: -30.3MW) (Demand Side Analytics, LLC, 2023).

3 Literature Review

The academic discourse on demand response branches out in several directions and touches upon equally many related fields such as energy modelling, engineering, economics, public policy and even sociology. Some researches have pondered over why some consumers may respond to demand response calls for reduced consumption while others would not (Ito et al., 2018). Suasion and appealing to individual responsibility of consumers to reduce consumption has been found to have an effect, but this effect diminished in the context of demand response programs as calls for reduction became more frequent. More promising results were achieved when economic and monetary incentives were introduced (Ito et al., 2018). There is evidence that residential consumers generally respond to price incentives when it comes to lowering electricity= consumption (Spees & Lave, 2007). Their responsiveness, however, may depend on how much electricity they use to begin with and for what purposes (Gyamfi et al., 2013). Some households do are not responsive to prices unless they use relatively more electricity, for instance for space heating. The ability to participate in demand response programs might be further limited by lacking the required hardware, which can be out of reach for some, especially low income households.

In terms of enrollment and active participation, media and public campaigns have been found helpful in achieving demand response goals (Moezzi et al., 2002). Campaigns appear to have positively influenced people’s decision making, though the context of a looming crisis heightened people’s responsiveness to campaign messaging. At the time the impact of campaigns was studied, calls for a voluntary demand reduction were deemed decisive for avoiding blackouts in 2001 in California. (Goldman et al., 2002). Interruptible load programs and more traditional approaches to DR failed partly due to consumers opting out or not complying when called upon.

Other than price-based DR initiatives, Direct Load Control (DLC) as a form of incentive-based DR have been identified as a viable and cost-effective tool for peak load demand management (Gagne et al., 2018). While many households are needed to participate for a DLC to be deliver meaningful reduction in electricity consumption, targeting industrial consumers may deliver greater committed capacity with fewer participants. It was noted that compensation to customers must outweigh the cost of procuring necessary hardware and software to assure sufficient enrollment

rates. (Gagne et al., 2018). Deciding on an appropriate assessment of demand reductions and the corresponding reimbursements to consumers is the topic of an ongoing conversation. There is a consensus among utility providers and regulatory authorities in the United States that baselines should be easy to understand and derived from the individual’s or the entity’s historical consumption. (Demand Side Analytics, LLC, 2023). Some researchers argued for the importance of accounting for appliances that residential consumers in DR programs use when modelling individual baselines (Gabaldón et al., 2021; Heshmati, 2014). This would potentially lower estimation and billing errors. In addition it would allow for identification of “licit gaming - intentional altering of the baseline prior to a DR event to maximize the potential reimbursement” (Gabaldón et al., 2021). In addition, same-day morning of an DR event adjustments to customer baselines have been deemed beneficial in improving the overall baseline accuracy (Coughlin et al., 2008). Customers with highly variable load profiles represent an exception as their baselines are notoriously challenging to construct accurately. The issue of heterogeneity of DR participants’ characteristic needs in all sectors (residential, commercial and industrial) is prominent in the literature and warrants a specialized approach to each sector when it comes to designing effective DR programs. Smart meters can help enable such specialization (Yan et al., 2018). The deployment of smart meters have been found to have helped extend DR programs to residential consumers by enabling access to variable electricity pricing plans. They also have the ability to manage household electricity consumption based on real time prices (Siano & Sarno, 2016; Widergren et al., 2012). A smart metering device or another piece of intelligent hardware could adjust consumption based on current prices and turn appliances on or off based on a set of thresholds defined by consumer. When it comes to the selection of industrial consumers to participate in DR programs, a careful consideration must be given to the ability of the consumer to respond to a demand response event (McKane et al., 2008). Some industrial entities are better candidates for enrollment and the time sensitivity of their production processes makes the difference. For instance, meat and dairy processing plants are generally unable to alter their electricity demand without risking spoilage or facing regulatory fines. Other reasons for why industrial consumers may not enroll in or respond to DR events include “technical risks of disrupting the production processes, loss of revenues and insignificant cost savings” (Leinauer et al., 2022). A potential remedy could be the formation of a market for electricity savings (Chao, 2011; Wu et al., 2015). All participating consumers would be obliged to buy electricity in excess of a contractual demand baseline and be also able to sell any quantity not consumed below this baseline in an open market.

Energy modeling is another category of academic literature that this study is closely related

to. Modeling electricity supply and demand can be notoriously challenging as one must contend with multiple types of seasonality, non-linear relationships, volatility and extreme outliers when building a forecasting model. The data may also often exhibit abnormalities and require additional processing before any work can be done using them. It is crucial to verify whether applicable assumptions are upheld, too. Modeling and forecasting are crucial for the efficient operation of an electric grid and a successful management of demand response programs. It involves the construction of aggregate and individual baselines, assessment of their accuracy and regular revision as conditions on the grid change. Modeling is a popular academic exercise and there is a large variety of different methods in the literature. Here is a small sample of some of the most popular approaches.

Machine learning and neural networks have been gaining in popularity among researchers as their computational power has unlocked new avenues to explore and model complex data. Random forest models have been tried to predicted hourly prices in California and showed promising results (Sadeghi-Mobarakeh et al., 2017). Gradient boosting, which is another method from the machine learning family based on combining decision trees proved superior to random forest or regression based methods at forecasting electric loads when judged by forecasting error (Zhang et al., 2019). This model appeared particularly well suited to forecast the load during night time and early morning. Other neural networks that have shown significant modelling potential include a Long short-term memory (LSTM) recurrent network. This is a deep learning method that can engage in both supervised and unsupervised learning. It has been found to deliver superior forecasts when modelling individual demand baselines (Kong et al., 2017). These baselines are generally more challenging to produce due to high load variability and increased random noise, which other modeling techniques are less equipped to handle (Jazaeri et al., 2016; Kong et al., 2017). Deep Convolution Neural Networks (DCNN) from the same model family have also been tried out to forecast electric loads in Victoria, Australia (Khan et al., 2019). The DCNN outperformed other models in the study, including ARIMA, in terms of forecast accuracy.

When perusing more recent literature on energy modeling, one might gain the impression that machine learning has dominated the field, but other statistical modeling techniques remain just as commonplace and popular. There are certain trade-offs between the complexity, accuracy and interpretability of models and a principle of parsimony could apply to general model selection. (Valentini et al., 2022). This summary suggests that more complex baseline methodologies tend to lead to greater forecasting accuracy, but also greater cost and initial investment in building the needed skillset. Methodologies with highest accuracy include unsupervised learning (UL) tech-

niques, other probabilistic estimation techniques and neural networks (NN). Time series analysis techniques such as autoregressive (integrated) moving average models (AR(I)MA) rank in the upper-middle pack by accuracy.

In another comprehensive overview of the most popular groups of modelling techniques, the authors summarize strengths, weaknesses and applications that each modeling group is best suited to (Verwiebe et al., 2021). ARMA and ARIMA models are considered generally reliable methodologies used for energy modeling. One of their main advantages is that they are more easily interpretable compared to machine learning methods (Verwiebe et al., 2021). They are also well equipped for uncovering patterns of auto-correlation and seasonality in the underlying data. To their detriment, though, ARIMA models are considered less suitable for long-term forecasting as they perform worse at forecasting extremes, handling outliers and working with large numbers (Verwiebe et al., 2021). One notable observation to be made about this summary is that Day-matching, which is the preferred baseline modelling method applied by utilities in the CAISO grid in California, is considered to show low to medium accuracy.

Other methods that appear in the literature about demand and price forecasting but are somewhat less common include the Fourier series expansion (Yukseltan et al., 2020), functional non-parametric and semi-functional partial linear models (Vilar et al., 2012). All three have shown competitive accuracy measurements when compared to ARIMA and its variations.

Variations of the ARIMA model such as seasonal ARIMA (SARIMA), or ARIMA with external regressors (ARIMAX, SARIMAX) are often combined with additional methods to increase the model’s accuracy. In Wang et al. (2012), the authors fit an additional model to the residuals produced by a seasonal ARIMA model predicting the electric load in northwestern China to increase the model’s accuracy. Residual errors were “modelled by three methods, namely Particle Swarm Optimization (PSO) optimal Fourier approach, the seasonal ARIMA model and the combined model, to enhance the prediction precision” (Wang et al., 2012). The combined model delivered the greatest improvement in accuracy compared to the original SARIMA model.

One of the basic assumptions of ARIMA models is that residuals produced by such models are not autocorrelated, have constant variance, and need to resemble a white noise process. This assumption is oftentimes violated when the modelled data is characterized by periods of low and high volatility. This variability, also known as non-constant variance, appears as residuals expanding and shrinking in magnitude in a non-linear fashion. A class of models called the *Autoregressive Conditional Heteroscedasticity (ARCH)* and by extension the *Generalized Autoregressive Conditional Heteroscedasticity (GARCH)* models has been developed for time series exhibiting this trait

by Engle (1982) and built on by Bollerslev (1987). ARCH models are designed to control for the time-varying volatility by including lagged squared error terms in the model equation (Enders, 2013; Engle, 1982). The volatility at time t is related to the previous squared residuals, allowing the model to capture the clustering of volatility in highly volatile data such as electric loads and prices. The GARCH model is an extension to the ARCH model and it is characterized by the incorporation of lagged squared residuals as well as lagged conditional variances. By including both the past squared errors and past conditional variances, GARCH models are able to model volatility dynamics more accurately (Enders, 2013). In Garcia et al. (2005), the authors made use of the GARCH framework and estimated an ARMA-GARCH model to predict day-ahead electricity prices in Spain and California. The choice of model was intentional as it relaxed some of the assumptions imposed on ARIMA models about its residuals being uncorrelated, having constant variance and a zero mean (Garcia et al., 2005). The residuals were “assumed to be serially correlated and can [therefore] be modeled by an Auto Regressive (AR) process” (Garcia et al., 2005). The forecast accuracy of the GARCH model in this study was found to exceed that achieved by a general ARIMA model.

While the available literature on demand response and energy modelling is varied and increasing in the number of publications, macro-level analyses of grid conditions during demand response events are sparse. Potential insights that such studies might provide could be valuable to policy makers and market participants who engage in electricity generation and cross-border and cross-balancing authority trading. The study I conduct in this thesis contributes to the existing literature in two ways: (1) The study extends the common evaluation approach to demand response by looking at the supply side of the equation and provides a grid-level analysis of the interaction between aggregate demand for electricity, net generation and interchange during recent demand response events in 2022 and (2) It combines several classes of statistical models (Autoregressive Integrated Moving Average (ARIMA), OLS, and the Generalized Autoregressive Conditional Heteroskedasticity (GARCH)) in an novel way to achieve improved accuracy and identify potential effects of California’s demand response initiatives on the variables of interest.

4 Methodology and Hypotheses

The methodology that I implement to explore the effects of demand response on grid-level demand and supply is most strongly inspired by Garcia et al. (2005) who used a GARCH model to predict volatile electricity prices, and Wang et al. (2012), whose approach to residual analysis and extraction of additional information from the error term provided for improved accuracy and the ability to address potential assumption violations of ARIMA-generated residuals. Now, let us define the referenced models mathematically. The definitions of a SARIMA and SARIMAX models are quoted *verbatim* from Shumway et al. (2000) and complemented by information from Enders (2013).

Definition 1. *The multiplicative seasonal auto-regressive integrated moving average model, or SARIMA model is given by*

$$\Phi_P(B^s)\phi(B)\nabla_s^D\nabla^d y_t = \delta + \Theta_Q(B^s)\theta(B)w_t$$

where w_t is the usual Gaussian white noise process $w_t \sim N(0, \sigma_{w_t}^2)$. The general model is denoted as $ARIMA(p, d, q) \times (P, D, Q)_s$. The ordinary auto-regressive and moving average components are represented by polynomials $\phi(B)$ and $\theta(B)$ of orders p and q , respectively, the seasonal auto-regressive and moving average components by $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ of orders P and Q and ordinary and seasonal difference components by $\nabla^d = (1 - B)^d$ and $\nabla_s^D = (1 - B^s)^D$.

The seasonal component of the SARIMA model shall enable us to capture the many types of seasonality that are present in data on electricity demand and supply. In addition, we can improve the accuracy of such a model by including exogenous regressors.

Definition 2. *SARIMA with Exogenous Regressors - SARIMAX*

$$\Phi_P(B^s)\phi(B)\nabla_s^D\nabla^d y_t = \delta + \Theta_Q(B^s)\theta(B)w_t + \sum_{i=1}^n \beta_i x_t^i$$

where $\sum_{i=1}^n \beta_i x_t^i$ consists of n exogenous regressors and their coefficients β_i for $i \leq n$ at time t .

The model residuals in the previous formulation above are omitted, but for purposes of this analysis it is beneficial to state them explicitly. Below is a highly simplified version of the SARIMAX model :

$$\Delta Y_t = ARIMA(p, d, q,) \times (P, D, Q)_s + \sum X_t + \epsilon_t$$

It would be reasonable to expect the error term ϵ_t to violate one or more of the standard assumptions about ARIMA residuals (particularly the constant variance of the error term assumption) due to the volatile nature of electricity data as frequently noted in the existing literature. The error term may still contain valuable information that the SARIMAX model is not well equipped to unlock. One of the goals of this study is to uncover this information, if present, and find out if demand response has any effect on the variables in question (demand, net generation, and interchange). To this end, I estimate a GARCH model that is trained on the SARIMAX-generated residuals. For better illustration of the GARCH model specifics I draw on Shumway et al. (2000) and course materials published by Pennsylvania State University (2021).

Please note that our target time series is a collection of residuals generated by a SARIMAX, model, therefore

$$\epsilon_t = y_t$$

ARCH and GARCH models were designed to model variance of a volatile time series (Engle, 1982). An ARCH(1) model to predict the variance at time t would look as follows

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2$$

where σ_t^2 is the conditional variance of a time series. The variance depends on the squared values of the preceding observation. For modelling residuals using a simple ARCH(1), we may assume the mean of the series to be zero and therefore write our model as

$$y_t = \sigma_t z_t$$

where σ_t is the conditional standard deviation and z_t is the residual of the ARCH model

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 \epsilon_{t-1}^2}$$

A generalized conditional heteroskedasticity model proposed by Bollerslev (1987) extends the ARCH framework and considers the values of past variances in addition to the previous squared values. A simple GARCH (1,1) would therefore look this:

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

Our time series (Residuals obtained from the SARIMAX model) can be modelled as follows

$$y_t = \sigma_t z_t$$

where the conditional variance at time t is obtained as follows

$$\sigma_t = \sqrt{\alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2}$$

Piecing the SARIMAX-GARCH Together

The equation below is an expanded SARIMAX model:

$$(1 - \nabla^d)(1 - \nabla_s^D)(1 - \nabla^s)Y_t = \beta_0 + \beta_1 X_t + \dots + \beta_k X_t^k + \\ \phi_1(1 - \nabla)^p Y_t + \dots + \phi_p(1 - \nabla)^p Y_{t-p} + \theta_1(1 - \nabla)^q \varepsilon_{t-1} + \dots + \theta_q(1 - \nabla)^q \varepsilon_{t-q} + \phi_1^*(1 - \nabla^s)^P Y_t + \\ \dots + \phi_P^*(1 - \nabla^s)^P Y_{t-Ps} + \theta_1^*(1 - \nabla^s)^Q \varepsilon_{t-s} + \dots + \theta_Q^*(1 - \nabla^s)^Q \varepsilon_{t-Qs} + \varepsilon_t$$

$(1 - \nabla^d)$ is the differencing term of the d order of non-seasonal differencing;

$(1 - \nabla_s^D)$ is the differencing term of the D order of seasonal differencing;

$(1 - \nabla^s)$ is the within-season differencing of the s seasonal period;

$\beta_0, \beta_1, \dots, \beta_k$, are the coefficients of the exogenous variables t ;

$\phi_1, \phi_2, \dots, \phi_p$, are the autoregressive (AR) coefficients for the lagged values of Y_t ;

$\theta_1, \theta_2, \dots, \theta_q$, are the moving average (MA) coefficients for the lagged residuals *var* ε_{t-s}

$\phi_1, \phi_2, \dots, \phi_P$, are the seasonal autoregressive (SAR) coefficients for the lagged values of Y_t at the seasonal period s

$\theta_1, \theta_2, \dots, \theta_Q$, are the seasonal moving average (SMA) coefficients for the lagged residuals ε_t at the seasonal period s

and ε_t represents the error term at time t

GARCH Model

The error term from the SARIMAX model can modeled by a GARCH model.

$$\varepsilon_t = \sigma_t z_t + a_t$$

σ_t represents the conditional standard deviation or volatility at time t estimated by the GARCH model, and z_t represents a standardized residual term, often assumed to be drawn from a standard normal distribution.

The GARCH model estimates the conditional volatility of the error term, which captures the time-varying variance in the residual series. By multiplying the conditional standard deviation σ_t with the standardized residuals z_t , we obtain the predicted error term ε_t at each time point. The GARCH model residuals are represented by the a term at the end. These are the residuals that I then use in an OLS analysis.

The procedure implemented in this study can be summarized in the following way.

1. First, an appropriate SARIMAX model is selected to model the time series of interest (demand, net generation and interchange). Residuals are tested for whether they conform to standard ARIMA assumptions.
2. Residuals are tested for the presence of conditional variance (heteroscedasticity). An ARCH-LM test is to be carried out. The presence of heteroscedasticity would enable us to estimate a GARCH model to capture the error variance that the SARIMAX model was unable to and unlock additional information from the residuals.
3. After fitting a suitable GARCH model, the residuals produced by this model are captured for purposes of an OLS regression estimation with dummy variables representing time periods when a grid emergency was in effect and demand response programs were dispatched.
4. Obtained coefficients are eventually evaluated for their statistical and economic significance. The coefficients that can be considered different from zero would indicate that observed values differ from predicted baselines in a non-random fashion, hinting at potential causality.

Hypothesis

To reiterate the objective of the study, I aim to identify any non-random patterns in the deviation of demand, generation, and interchange from their predicted baselines in order to assess whether these deviations, if present, can be attributed to the dispatch of demand response initiatives in California. The OLS regressions shall take the following form:

$$y_t = \alpha_0 + \beta_1 DR_event_period + \beta_2 pre_DR_period + \beta_3 EmergencyDay + \sum X + \epsilon$$

where $\sum X$ is a vector of control variables. The null hypothesis I work with is that coefficients $\beta_1 = \beta_2 = \beta_3 = 0$ and demand response does not have an impact on how the time series variables in question (demand, generation and interchange) deviate from their predicted baselines. If one or more of these coefficients turn out to differ from zero and bear statistical significance, then we may conclude to have found evidence of an impact of demand response on the time series in question.

5 Data

The data used in the analysis were obtained from a variety of sources. Descriptive statistics of continuous variables are presented in the table below.

Table 2: Descriptive Statistics

	Count	Mean	Median	Max	Min
Demand (MWh)	2952	30841.40	29870.00	51104.00	20992.00
Net Generation (MWh)	2952	23929.85	23229.50	43767.00	10874.00
Total Interchange (MWh)	2952	-6308.52	-6654.00	464.00	-10764.00
Generation from Solar (MWh)	2952	4576.93	1462.00	13126.00	-89.00
Temperature °C	2952	21.30	21.01	37.48	-1.51

1. **Demand for electricity.** The data were obtained from the U.S. Energy Information Administration (U.S. Energy Information Administration, n.d.). Demand for electricity is measured in MWh on an hourly on a balancing authority level (in this case the California Independent System Operator). This measure covers all electric load demand in the grid (residential, commercial and industrial) combined. As shown in Figures 2, 3 and 4, the time series exhibits multiple types of seasonality. Winters are characterized by lower demand while summers and first few weeks of fall tend to record the highest demand. Peaks appear to increase in magnitude in recent years. There is also an observable weekly pattern with Saturdays representing a weekly trough and Wednesdays . appearing as a weekly peak. Daily troughs tend to be observed at around 5:00 and peaks at 19:00. These are typical for summer months. Winter, spring and fall daily load profiles exhibit different patterns. From the three figures it is also apparent that California consumes more electricity than it generates.
2. **Total Interchange.** The data were obtained from the U.S. Energy Information Administration (U.S. Energy Information Administration, n.d.). Similar to the demand, total interchange is measured in MWhs on the CAISO grid level. Total interchange is the sum of all hourly inflows and outflows of electricity to and from the CAISO grid. California

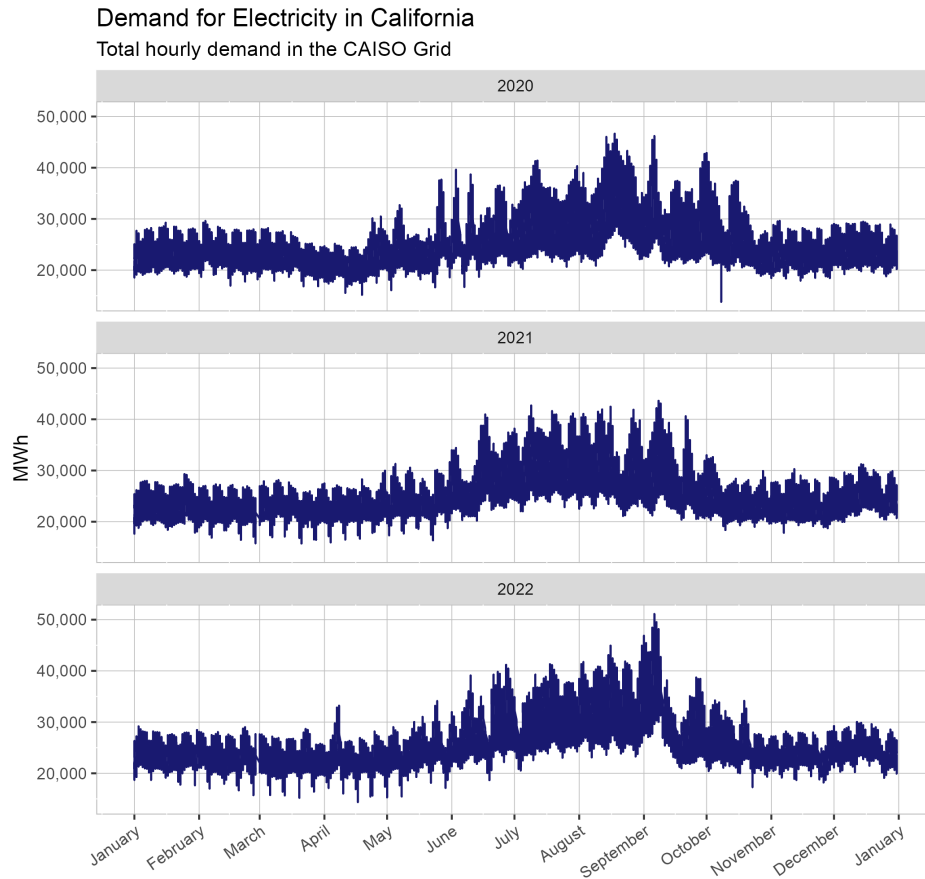


Figure 2: Electricity Demand in the CAISO grid

has been traditionally a net importer of electricity and the total interchange is largely negative. Imports follow a common daily pattern, while weekly and annual patterns are less pronounced.

3. **Net Generation, Solar Generation.** Obtained from the U.S. Energy Information Administration (U.S. Energy Information Administration, n.d.). Net generation is the sum of all generated electricity, minus any that has been stored or used up in the process of generation. Solar generation is a subset of the net generation variable. Both are measured in MWhs on a balancing authority level. In general, California is not self-sufficient to cover its electricity needs, but net generation covers most of the domestic consumption. In terms of seasonality, net generation follows the pattern of demand relatively closely (See Figure 3). Peaks and troughs appear to form usually an hour before the demand's daily minima and maxima.
4. **Temperature.** Hourly temperature data were obtained from the University of Utah MesoWest website run by the Department of Atmospheric Sciences (Department of Atmospheric

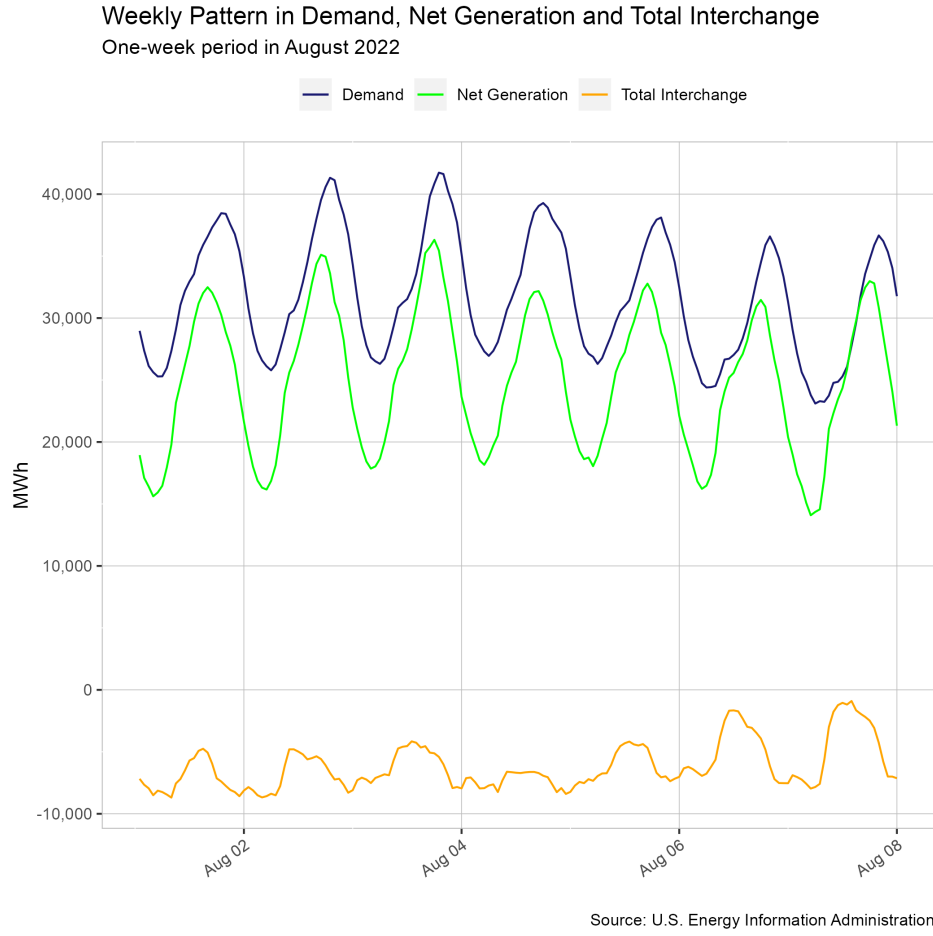


Figure 3: Close-up: Weekly variation in the variables of interest.

Sciences, University of Utah, n.d.). This variable is measured in degrees Celsius in hourly intervals at the Los Angeles International Airport (LAX). The temperature variable has been included in the analysis as a proxy measure as outdoor temperature influences energy consumption. The location of the data source has been chosen intentionally due to the fact that the Los Angeles census area is one of the most populous parts of California. The local temperature and weather therefore affect a sizeable portion of the CA population and influences energy consumption in a manner that may show in grid-level data.

5. **Emergency and Demand Response-related dummy variables.** These variables were produced using information from the California Independent System Operator’s website (California Independent System Operator, CAISO, n.d.-b). In 2022, there were 10 Energy Emergency Watch days, Flex alert was in effect 11 times (for 5 to 6 hours each time). A pre-DR event dummy variable covers the 8 hours immediately before the start of a demand response event that starts at 16:00 local time. Demand response events are most commonly declared

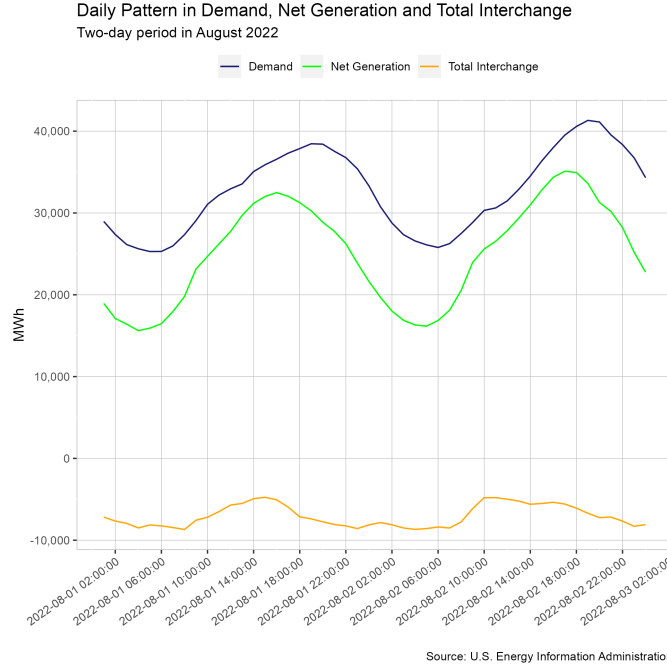


Figure 4: Close-up: Daily variation in variables of interest.

a day in advance and be in effect between 16:00 and 21:00 local time. A *DR-event period* dummy takes a value of 1 if a Flex alert was in effect. An *Emergency Day* dummy is equal to 1 for all hours in a day, for which there was a Flex Alert issued or the Energy Emergency Watch was declared.

6 Model Selection and Estimation

In this study, I explore the use two types of statistical models: the seasonal Autoregressive Moving Average with exogenous regressors (SARIMAX) model and the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model. The model selection process involved several steps as outlined below.

First, I performed two types of unit roots tests on the time series data, and the results are presented in Table 3. All three time series appeared to be stationary, contrary to my expectations. Based on trends shown in Figure 12, I would have expected non-stationarity. This finding suggested that initial non-stationarity and further pre-processing might not be necessary after all. Obligatory ACF and PACF plot were produced and can be seen in Figure 13.

To identify the most suitable model for each time series, I experimented with a variety of model specifications and exogenous regressors. Tables 5, 6, and 7 provide an overview of explored model specifications for the three analyzed time series (demand, net generation, and total interchange).

To fit these models, I utilized the `auto.arima()` function from the `forecast` package in R. Differencing was applied to induce stationarity in the majority of the models, which contradicted the findings of stationarity in the preceding step.

Time Series	Test	Test Statistic	Lag	p-value	H ₀
Electricity Demand	Augmented Dickey-Fuller (ADF)	-9.2058	11	<0.01	Unit root
	Phillips-Perron	-228.63	7	<0.01	Unit root
Net Generation	Augmented Dickey-Fuller (ADF)	-7.0925	11	<0.01	Unit root
	Phillips-Perron	-224.06	7	<0.01	Unit root
Total Interchange	Augmented Dickey-Fuller (ADF)	-7.8474	11	<0.01	Unit root
	Phillips-Perron	-174.95	7	<0.01	Unit root

Table 3: Tests for the presence of stationarity in the studied time series

Next, I trained the models on data from July and August 2021 and tested them on September 2021 data. The root mean square error (RMSE) and mean absolute percentage error (MAPE) values for both the training and testing datasets are presented in the respective tables. The models obtained from 2021 data were then employed to predict unaffected baselines for the analysis of demand response events in the summer and early fall of 2022. Coefficient estimates for the three most promising models for each time series are shown in Tables 11, 12 and 13 in Appendix II.

Model	Data	Test	Test statistic	Lags	H ₀	p-value
Model 10: ARIMA(2,1,2)(0,0,1)[24]	Residuals (Demand)	Ljung-Box	$Q^* = 131.43$	48	White noise	1.087e-09
Model 10: ARIMA(2,1,2)(0,0,1)[24]	Residuals (Demand)	ARCH-LM	$\chi^2 = 406.51$	NA	No ARCH effects	2.2e-16
Model 12: ARIMA(2,1,3)(0,0,2)[24]	Residuals (Net Generation)	Ljung-Box	$Q^* = 68.76$	48	White noise	0.02626
Model 12: ARIMA(2,1,3)(0,0,2)[24]	Residuals (Net Generation)	ARCH-LM	$\chi^2 = 40.074$	NA	No ARCH effects	6.989e-05
Model 11: ARIMA(2,1,1)(0,0,2)[24]	Residuals (Interchange)	Ljung-Box	$Q^* = 75.227$	48	White noise	0.001707
Model 11: ARIMA(2,1,1)(0,0,2)[24]	Residuals (Interchange)	ARCH-LM	$\chi^2 = 84.63$	NA	No ARCH effects	5.357e-13

Table 4: Residual Diagnostics

In evaluating the predictive potential of the models, I assigned the greatest weight to two metrics: the Akaike Information Criterion (AIC) and the Mean Absolute Percentage Error (MAPE) calculated from the test data set. I also conducted a Ljung-Box test on the residuals of each model before final selection. The test results indicated that these models did not produce residuals resembling white noise processes. Based on the presence of autocorrelation and heteroskedasticity in the residuals, as shown in Table 4 (And also in Residual Diagnostics charts in Appendix I), I estimated appropriate GARCH models for each of the three time series (See Appendix A). The tests confirmed the presence of autocorrelation and heteroskedasticity, further justifying the decision to fit GARCH models to the residuals. The selection of the GARCH models was based on their corresponding AIC values. Selected GARCH model specifications can be found in the appendix.

Table 5: Model Comparison - Demand for Electricity

Model	AIC	BIC	RMSE (Train)	RMSE (Test)	MAPE (Train)	MAPE (Test)	Hours	Days	°C	Fourier	Variance	Ljung-Box p-value
1	ARIMA(2,0,3)	22614.62	22651.75	479.65	5588.95	1.21	18.38	No	No	No	230995.21	0.00
2	ARIMA(4,0,1)(2,1,0)[24]	21376.64	21418.95	352.19	3608.48	0.84	10.23	No	No	No	126675.66	0.00
3	ARIMA(2,1,2)(0,0,1)[24]	21395.83	21549.67	315.72	3296.65	0.77	8.92	Yes	No	No	101657.58	0.00
4	ARIMA(2,1,2)(0,0,1)[24]	21400.38	21586.03	314.92	3272.31	0.76	8.87	Yes	No	No	101565.18	0.00
5	ARIMA(2,1,2)(0,0,1)[24]	21399.04	21590.00	314.57	3275.07	0.76	8.85	Yes	Yes	No	101406.71	0.00
6	ARIMA(4,0,2)(2,1,0)[24]	21376.76	21429.65	351.72	3576.85	0.84	10.09	No	Yes	No	126510.23	0.00
7	ARIMA(2,1,2)	21458.91	21607.43	322.81	3331.46	0.80	9.12	Yes	No	Yes	106207.47	0.00
8	ARIMA(3,0,0)(2,1,0)[24]	21385.75	21454.50	352.09	3666.38	0.84	10.43	No	No	No	127037.84	0.00
9	ARIMA(2,1,5)	21442.31	21606.75	320.33	3297.99	0.79	8.97	Yes	No	No	104796.36	0.00
10	ARIMA(2,1,2)(0,0,1)[24]	21394.48	21553.61	315.36	3294.59	0.77	8.92	Yes	No	No	101497.61	0.00

Model	AIC	BIC	RMSE (Train)	RMSE (Test)	MAPE (Train)	MAPE (Test)	Hours	Days	°C	Fourier	Solar	Variance	Ljung-Box p-value
1	ARIMA(4,1,1)	24121.89	24153.72	801.41	5436.42	2.77	22.41	No	No	No	No	644861.49	0.00
2	ARIMA(3,0,0)(2,1,0)[24]	22917.42	22949.16	597.40	3294.22	2.04	11.44	No	No	No	No	363982.72	0.00
3	ARIMA(2,1,2)(0,0,2)[24]	22959.17	23118.31	533.62	4103.53	1.83	16.74	Yes	No	No	No	290607.37	0.03
4	ARIMA(2,1,2)(0,0,2)[24]	22984.42	23175.39	536.07	3676.68	1.84	14.55	Yes	No	No	No	294492.14	0.02
5	ARIMA(2,1,2)(0,0,2)[24]	22969.52	23165.79	532.96	4065.27	1.82	16.45	Yes	Yes	No	No	291285.76	0.03
6	ARIMA(2,0,2)(2,1,0)[24]	22917.62	22959.93	596.63	3278.18	2.04	11.41	No	Yes	No	No	363538.04	0.00
7	ARIMA(2,1,2)	23027.84	23176.37	547.05	3667.43	1.89	14.32	No	No	Yes	No	305005.68	0.00
8	ARIMA(3,0,0)(2,1,0)[24]	22927.30	22990.77	596.96	3317.19	2.04	11.53	No	No	No	No	364940.98	0.00
9	ARIMA(2,1,2)	23027.86	23176.39	547.05	3674.70	1.89	14.39	Yes	No	No	No	305008.50	0.00
10	ARIMA(2,1,2)(0,0,2)[24]	22959.69	23124.13	533.35	4128.37	1.82	16.93	Yes	No	No	No	290517.97	0.10
11	ARIMA(4,0,1)(2,1,0)[24]	22620.65	22673.54	551.90	3192.72	1.91	11.39	No	Yes	No	Yes	310455.28	0.00
12	ARIMA(2,1,3)(0,0,2)[24]	22607.18	22776.93	484.96	3735.04	1.67	14.67	Yes	No	No	Yes	239551.21	0.00
13	ARIMA(2,1,3)(0,0,2)[24]	22686.69	22755.45	484.96	3197.54	1.67	11.45	No	Yes	No	Yes	324638.83	0.00

Table 6: Model Comparison - Net Generation

	Model	AIC	BIC	RMSE (Train)	RMSE (Test)	MAPE (Train)	MAPE (Test)	Hours	Days	°C	Fourier	Solar	Variance	Ljung-Box p-value
1	ARIMA(4,1,1)	22937.72	22969.54	538.53	2123.46	9.93	32.86	No	No	No	No	No	291190.00	0.00
2	ARIMA(1,0,1)(2,1,0)[24]	22201.38	22227.82	468.13	1916.92	8.95	30.15	No	No	No	No	No	223350.57	0.00
3	ARIMA(2,1,1)(0,0,2)[24]	22269.02	22422.85	423.44	1961.35	8.30	29.21	Yes	No	No	No	No	182867.89	0.00
4	ARIMA(1,1,4)(0,0,2)[24]	22244.28	22440.54	417.69	1949.68	8.16	28.74	Yes	Yes	No	No	No	178913.25	0.00
5	ARIMA(1,1,4)(0,0,2)[24]	22245.64	22447.21	417.60	1905.35	8.15	28.11	Yes	Yes	Yes	No	No	178960.10	0.00
6	ARIMA(1,0,1)(2,1,0)[24]	22202.89	22234.63	468.06	1916.84	8.96	30.15	No	No	Yes	No	No	223430.34	0.00
7	ARIMA(0,1,1)	22412.73	22545.34	445.91	1541.66	8.33	22.33	No	No	No	Yes	No	202234.47	0.00
8	ARIMA(1,0,1)(2,1,0)[24]	22170.55	22228.73	461.31	1801.81	8.89	28.19	No	Yes	No	No	No	217778.69	0.00
9	ARIMA(0,1,1)	22412.73	22545.34	445.91	1541.47	8.33	22.32	Yes	No	No	No	No	202234.48	0.00
10	ARIMA(2,1,4)(0,0,2)[24]	22269.86	22444.91	422.44	2090.29	8.10	31.33	Yes	No	Yes	No	No	182501.59	0.06
11	ARIMA(2,1,1)(0,0,2)[24]	22161.77	22320.90	417.94	1898.71	8.15	28.05	Yes	No	No	No	Yes	177664.66	0.00

Table 7: Model Comparison - Interchange

7 Results

7.1 Demand

The goal of the study was to identify and estimate the impact of demand response programs on demand, net generation, and interchange during demand response events in California in 2022. Figure 5, provides a visual a representation of the exercise.

Figure 5 displays a 30-step ahead prediction (red line) using a combination of SARIMAX and GARCH models ($\text{ARIMA}(2,1,2)(0,0,1)[24]$, $\text{GARCH}(1,2)$). The first demand response event day of 2022 with a Flex alert in effect was observed on 17 August 2022. It is important to note that the model does not incorporate information about individual active demand response programs or grid emergencies as there may be overlaps and the activation of the programs is dependent on the decision of the grid operator and DR program administrators. The shaded area represents the time window during which a Flex alert was in effect, dispatching a range of demand response programs.

Throughout the day, the observed demand consistently exceeds the predicted demand, continuing until 22:00. However, it is worth noting that the observed demand starts converging to the predicted baseline at 16:00, coinciding with the demand response event period. Additionally, the daily peak occurs one hour earlier than predicted. The observed demand exhibits random deviations from the baseline. The figure also includes a chart indicating a temperature proxy, demonstrating that demand lags behind temperature by approximately one hour.

As expected, the 85% and 95% prediction intervals expand over time. Now, the focus is on assessing whether deviations from the baseline during demand response periods, pre-demand response periods, and demand response event days differ from deviations observed during all hours.

The OLS estimates for this exercise are presented in Table 8, with particular interest in columns 4 through 6. Columns 1 through 3 contain estimates of ARIMA residuals regressed on control variables and the relevant dummies. However, these residuals violate the assumption of non-constant variance.

The coefficient for the *Pre-DR Period (8hrs)* is 180.13, indicating that, on average, the deviation from the baseline during the pre-demand response period was 180.13 MWh larger compared to other hours on non-demand response event days. This coefficient is statistically significant at the 0.1% level, suggesting that load shifting occurred, and consumers smoothed their electricity demand during this period.

Regarding the *Demand Response Period* coefficient, it is -120.91 and statistically significant at the 5% level. This implies that the hourly deviation from the baseline during the demand response

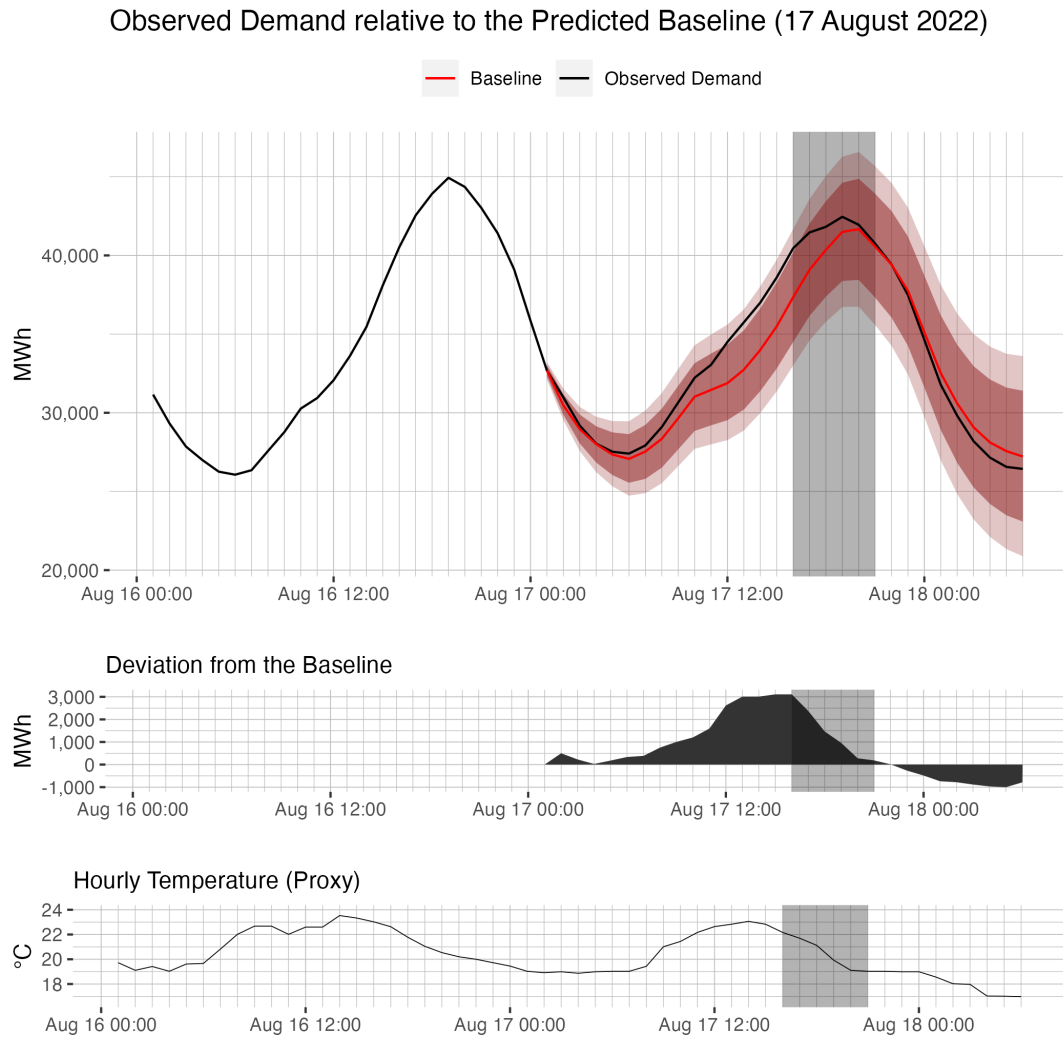


Figure 5: 17 August 2022 Observed vs. Predicted Demand

period was, on average, 120.91 MWh lower relative to other off-peak hours on non-demand response event days. When compared to peak hours on non-demand response event days, the magnitude decreases to -57.32MWh. The direction and magnitude of this coefficient indicate that peak load demand was effectively reduced during demand response event days in 2022. As for R^2 , it is extremely low, which is to be expected. If the models producing these residuals are sufficiently specified, there should be minimal to no systematic or meaningful variation there left.

Dependent Variable:	Residuals ARIMA(2,1,2)(0,0,1) ₂₄			Residuals GARCH(1,2)		
	(1)	(2)	(3)	(4)	(5)	(6)
(Constant)	−39.42 (83.48)	100.25 (83.50)	70.92 (86.40)	−79.88 (78.91)	33.12 (80.00)	−7.11 (82.72)
Electricity Demand (GWh)	−0.04 (1.89)	0.94 (1.80)	1.59 (1.87)	2.88 (1.78)	2.97 (1.72)	3.86* (1.79)
EEA Watch (Flex Alert) Day	−14.79 (28.53)		−44.42 (33.72)	−39.34 (26.96)		−60.93 (32.29)
Temperature (°C)	2.30 (3.22)	−5.88 (3.35)	−5.24 (3.39)	0.43 (3.05)	−5.26 (3.21)	−4.39 (3.24)
Peak Load Time	−63.51* (28.00)	−12.89 (27.47)	−21.42 (28.22)	−88.97*** (26.46)	−51.89* (26.32)	−63.59* (27.02)
Demand Response Period		−243.53*** (52.02)	−209.18*** (58.18)		−168.02*** (49.84)	−120.91* (55.71)
Pre-DR Period (8hrs)		248.30*** (44.16)	279.87*** (50.23)		136.84** (42.31)	180.13*** (48.10)
R ²	0.01	0.05	0.05	0.01	0.02	0.03
Adj. R ²	0.00	0.04	0.05	0.01	0.02	0.02
Num. obs.	1464	1464	1464	1464	1464	1464

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 8: OLS Estimates - Demand Residuals

7.2 Net Generation

In Figure 6, another 30-step ahead prediction is presented, focusing on net generation. The prediction is generated using a combination of SARIMAX and GARCH models ($\text{ARIMA}(2,1,3)(0,0,2)[24]$, $\text{GARCH}(1,2)$) for the same date as mentioned earlier.

Throughout the day, the observed net generation consistently surpasses the predicted baseline, persisting until 17:00. However, starting at 16:00, which aligns with the demand response event period, the observed net generation begins converging towards the predicted baseline. Similar to the demand analysis, the daily peak in net generation occurs one hour earlier than predicted. The deviations from the baseline in net generation also appear to be random. The accompanying temperature proxy chart indicates a lag of approximately one hour between temperature and net generation.

In contrast to the demand analysis, the 85% and 95% prediction intervals for net generation exhibit both shrinkage and expansion over time. This suggests that net generation is relatively more stable and predictable compared to demand. The subsequent analysis focuses on assessing whether deviations from the baseline during demand response periods, pre-demand response periods, and demand response event days differ from deviations observed throughout all hours.

The OLS estimates in Table 8 are of interest, particularly columns 4 through 6. Columns 1 through 3 present estimates of ARIMA residuals regressed on control variables and relevant dummies, although these residuals violate the assumption of non-constant variance.

The coefficient for the *Pre-DR Period (8hrs)* is 271.57, indicating that, on average, the deviation from the baseline during the pre-demand response period was 271.57 MWh larger compared to other off-peak hours on non-demand response event days. This coefficient is statistically significant at the 0.1% level, suggesting that domestic generation significantly increased to meet a substantial portion of the electricity demand on this particular day.

On the other hand, the *Demand Response Period* coefficient is 16.82 and statistically insignificant at all meaningful levels. This implies that the average hourly deviation from the baseline during the demand response period did not significantly differ from the predicted baseline. Therefore, there is no evidence to suggest that demand response had a noticeable impact on net generation during the demand response event period (16:00 - 21:00) on emergency days in 2022.

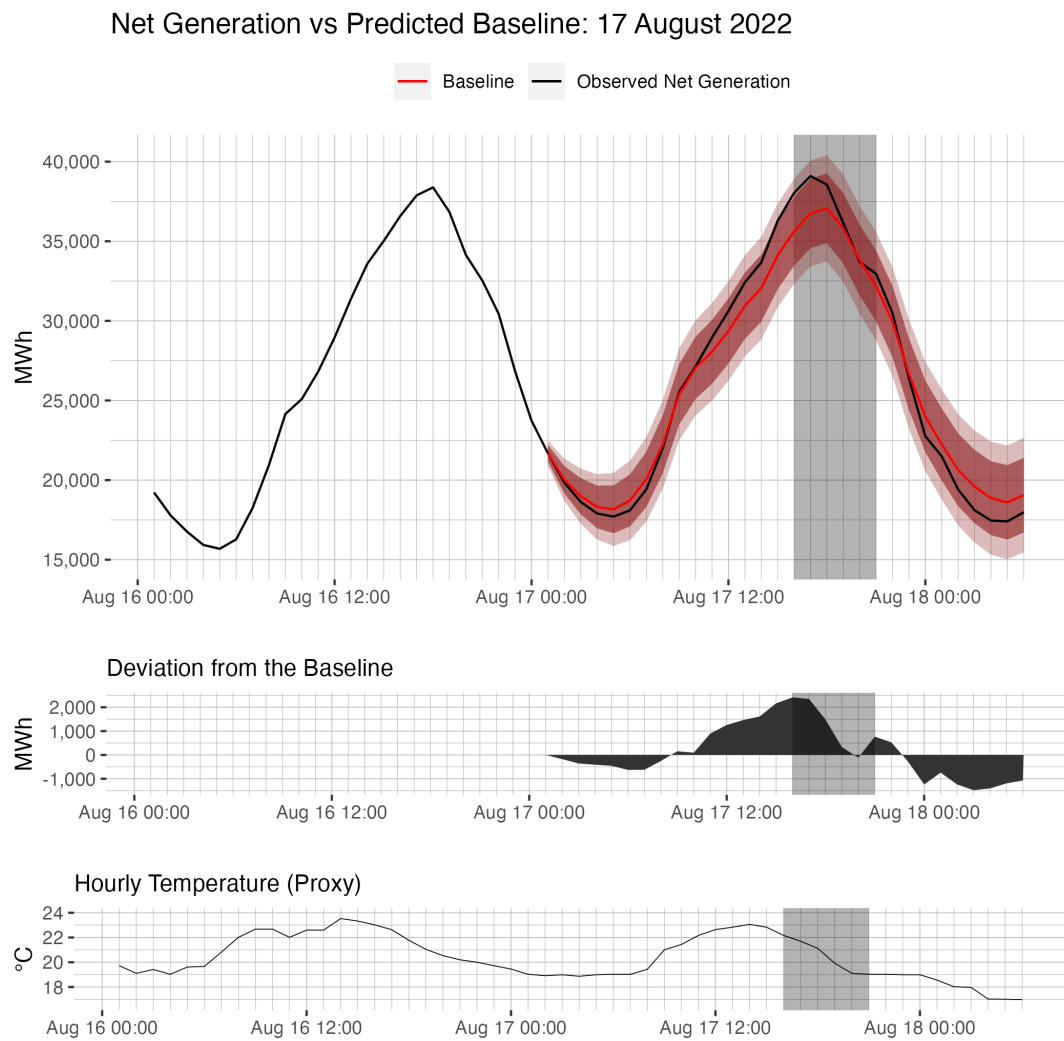


Figure 6: Observed vs. Predicted net generation, 17 August 2022

Dependent Variable:	Residuals ARIMA(2,1,3)(0,0,2) ₂₄			Residuals GARCH(1,2)		
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	-9.45 (144.68)	102.25 (137.01)	93.86 (148.06)	48.22 (143.81)	161.11 (136.18)	151.92 (147.16)
Electricity demand (GWh)	4.65 (2.85)	4.95 (2.80)	5.05 (2.87)	4.71 (2.83)	5.00 (2.78)	5.11 (2.85)
EEA Watch (Flex Alert) Day	67.85 (46.19)		-8.34 (55.73)	67.44 (45.91)		-9.15 (55.39)
Temperature (°C)	-7.13 (6.37)	-12.49* (6.13)	-12.17 (6.49)	-9.81 (6.33)	-15.21* (6.09)	-14.86* (6.45)
Peak Load Time	-88.45* (42.65)	-71.25 (43.12)	-72.31 (43.72)	-89.32* (42.39)	-72.10 (42.86)	-73.27 (43.45)
Solar Net Generation (GWh)	5.91 (3.25)	4.66 (3.04)	4.48 (3.27)	5.88 (3.23)	4.65 (3.02)	4.45 (3.25)
Demand Response Period		9.81 (80.35)	15.73 (89.58)		10.32 (79.86)	16.82 (89.03)
Pre-DR Period (8hrs)		265.18*** (67.90)	270.87*** (77.85)		265.32*** (67.48)	271.57*** (77.38)
R ²	0.01	0.02	0.02	0.01	0.02	0.02
Adj. R ²	0.00	0.01	0.01	0.00	0.01	0.01
Num. obs.	1464	1464	1464	1464	1464	1464

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 9: OLS Estimates - Net Generation Residuals

7.3 Total Interchange

Figure 7 presents the final 30-step ahead prediction for the total interchange, employing SARI-MAX and GARCH models (ARIMA(2,1,1)(0,0,2)[24], GARCH(2,2)), focusing on the same date as mentioned earlier.

Observations indicate that, up until 15:00, the total interchange surpasses the predicted baseline. This suggests that California imported less electricity than initially forecasted in the first 15 hours. However, at the same hour, there was a shift, and imports exceeded the predicted baseline, remaining elevated until the following day without signs of convergence. Notably, the total interchange appears to move inversely to temperature, implying that solar generation, which follows a similar daily pattern, may significantly influence the magnitude of imports.

Unlike the previous analyses, the 85% and 95% prediction intervals for total interchange demonstrate significant expansion over time. This indicates that total interchange exhibits greater fluctuations and is less predictable relative to the other two analyzed time series. Let us now shift our focus towards assessing whether deviations from the baseline during demand response periods, pre-demand response periods, and demand response event days differed from deviations observed at all hours in the studied time periods of 2022.

Table 10 presents respective OLS estimates, with specific attention given to columns 4 through 6. Columns 1 through 3 provide estimates produced from ARIMA. As stated before, these residuals violate the assumption of non-constant variance and may therefore be too problematic for drawing conclusions.

The coefficient for the *Pre-DR Period (8hrs)* is -65.33, indicating that, on average, the deviation from the baseline during the pre-demand response period was 65.33 MWh larger compared to other hours on non-demand response event days. However, this coefficient is not statistically significant at any meaningful level. Consequently, there is no evidence to suggest that total interchange consistently deviated from the predicted baseline during the pre-demand response period.

In contrast, the *Demand Response Period* coefficient is 200.37 and statistically significant at the 1.0% level. This implies that, on average, the hourly deviation from the baseline during the peak demand response period was lower by 200.37 MWh. Thus, hourly imports were systematically on average lower during the peak demand response hours compared to off-peak hours on non-emergency days. These results suggest that demand reductions induced during the peak demand response hours translated into lower imports, while net generation within the CAISO grid remained largely unchanged.

Total Interchange relative to the Predicted Baseline: 17 August 2022

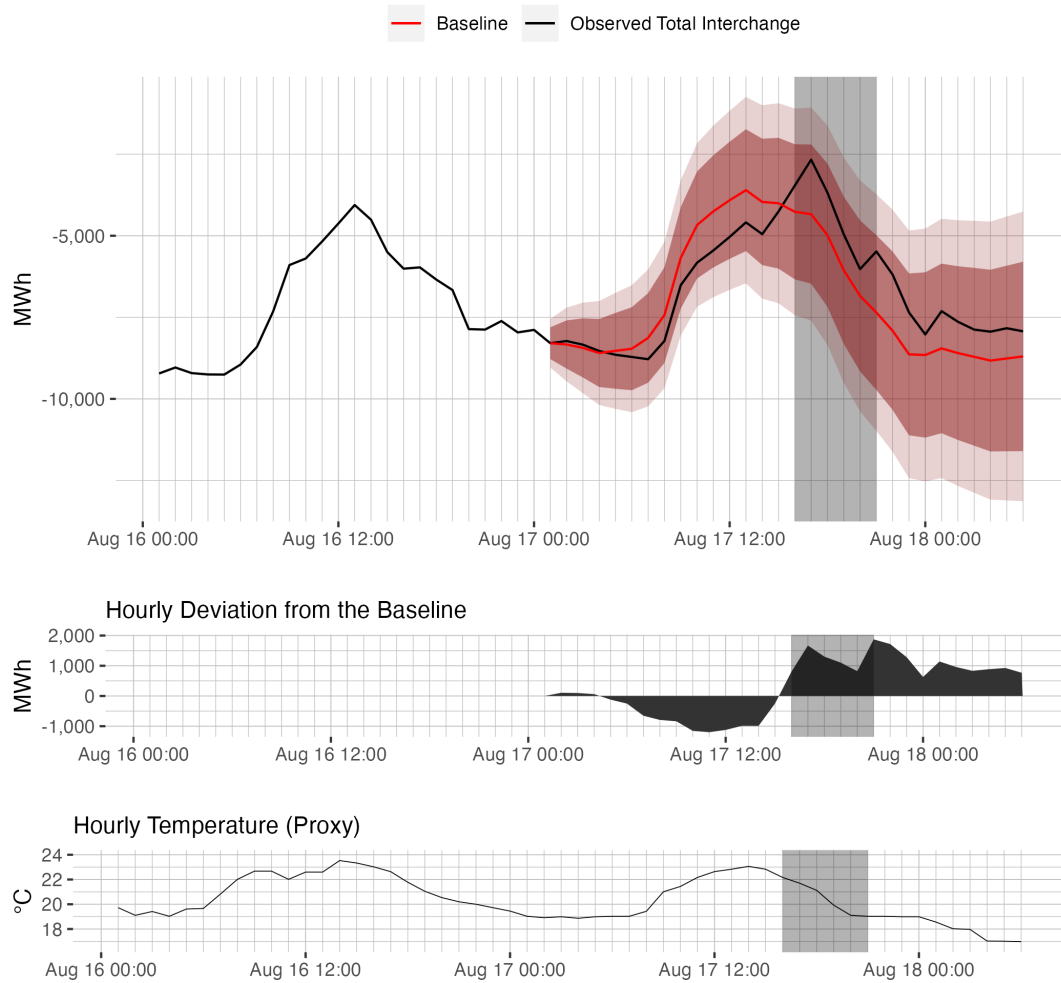


Figure 7: Observed vs. Predicted total interchange, 17 August 2022

Dependent Variable:	Residuals ARIMA(2,1,1)(0,0,2) ₂₄			Residuals GARCH(2,2)		
	(1)	(2)	(3)	(4)	(5)	(6)
(Intercept)	216.36 (127.01)	246.33* (120.44)	215.96 (130.15)	88.88 (126.50)	114.35 (119.92)	86.06 (129.58)
Electricity Demand (GWh)	0.50 (2.50)	-0.95 (2.46)	-0.60 (2.52)	0.44 (2.49)	-1.02 (2.45)	-0.70 (2.51)
EEA Watch (Flex Alert) Day	-8.82 (40.55)		-30.21 (48.98)	-8.03 (40.39)		-28.14 (48.77)
Temperature (°C)	-11.22* (5.59)	-10.67* (5.39)	-9.52 (5.71)	-5.68 (5.57)	-4.89 (5.37)	-3.82 (5.68)
Peak Load Time	-18.18 (37.44)	-42.31 (37.91)	-46.16 (38.43)	-17.44 (37.29)	-42.88 (37.74)	-46.47 (38.26)
Solar Net Generation (GWh)	6.33* (2.85)	7.11** (2.67)	6.46* (2.87)	6.56* (2.84)	7.33** (2.66)	6.72* (2.86)
Demand Response Period		174.68* (70.64)	196.11* (78.74)		180.41* (70.33)	200.37* (78.40)
Pre-DR Period (8hrs)		-78.25 (59.69)	-57.63 (68.43)		-84.54 (59.43)	-65.33 (68.13)
R ²	0.01	0.01	0.01	0.00	0.01	0.01
Adj. R ²	0.00	0.01	0.01	0.00	0.01	0.01
Num. obs.	1464	1464	1464	1464	1464	1464

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 10: OLS Estimates - Total Interchange Residuals

8 Discussion

The goal of this study was to estimate the impact of California’s demand response initiatives on demand, net generation, and interchange by constructing unaffected baselines using SARIMAX and GARCH models from aggregate grid-level data, and then analyzing the resulting deviations. The study aimed to determine whether the observed time series deviated from the baselines in an anticipated, meaning random way, or systematic deviations from the baselines during demand response events, pre-demand response periods, and emergency days occurred. The motivation behind this analysis stemmed from the apparent lack of research focusing on how the supply adjusts during demand response events.

The findings of the study indicate that the demand for electricity exhibited systematic deviations that could not be attributed to randomness. Specifically, hourly deviations were on average positive during the pre-demand response period (+180.13 MWh) relative to off-peak non-demand response days, and became on average negative (-120.91 MWh) during the peak demand response period between 16:00 and 21:00.

As for net generation, the hourly deviations during the pre-demand response periods were on average positive (271.57 MWh), exceeding the baseline. This suggests that the increased demand was satisfied by higher net generation within California. There appeared to be no meaningful hourly deviations of net generation from its baseline during the peak demand response.

Looking at the dummy coefficients when regressing residuals generated by the SARIMAX model predicting total interchange, there was some evidence of on average positive deviation from the baseline during demand response peak hours (Table 9). The coefficient (200.37 MWh) was positive and significant at the 5% level, indicating that imports decreased below the predicted baseline during demand response event hours. It is also useful to note that all sets of OLS regressions for the three variables in question (demand, net generation and interchange) shared a common characteristic: the adjusted R-squared values were very low. This implies that there might be very little remaining variation that is not random in the residuals.

On a broader scale, the findings suggest that demand response initiatives were used effectively during the demand response events of 2022 in California and this information can be observed even through the analysis of aggregate, grid-level data. However, the GARCH model estimation did not add significant insights for net generation relative to the coefficients obtained by regressing ARIMA generated residuals, implying that the model might not be improperly specified or that an alternative approach is required. In other papers utilizing ARCH and GARCH models (Bollerslev, 1987; Engle, 1982; Garcia et al., 2005), the time series analyzed were of much lower frequency

(Daily or monthly), which may provide a key insight into how to refine the methodology for any future study. In terms of total interchange, the study found that imports were on average smaller than predicted during peak demand response hours, highlighting the impact of demand response on the regional energy exchange.

When compared to existing research, this study contributes to the limited body of knowledge on the supply side effects of demand response. The findings partially with previous literature regarding effects on demand (Demand Side Analytics, LLC, 2023; Goldman et al., 2002), but there is a scarcity of research on how the supply side adjusts during demand response events.

Limitations, Internal Validity and Future Research

The results of the presented study shall be interpreted with caution as justifiable concerns for internal validity persisted and were not satisfactorily addressed, primarily due to data limitations. The following issues have been identified as the most potent sources of influence that may obscure the true effect of demand response on electricity demand, generation and interchange in California in 2022.

1. When a grid emergency is declared, the notification is likely received also by consumers who are not enrolled in any demand response programs. These consumers may still choose to respond and alter their consumption patterns for moral or ethical reasons. This could reduce the consumption of electricity beyond what could be attributable to DR programs on variables in question and falsely inflate the effect of DR, unless the notification of grid emergency itself is considered an integral part of DR, in which case this issue would be of no concern.
2. Small service-oriented businesses or manufacturing operations with insufficient air conditioning may close during extreme heatwaves as no customers are expected or working conditions become dangerous. This would reduce the demand for electricity that cannot be attributed to demand response initiatives.
3. The increased volatility of demand for electricity during grid emergencies may induce large price swings in both directions in the spot electricity markets. Utility scale battery operators and businesses on real time pricing plans may take advantage of temporary, though significant price drops and buy up the electricity saved by other consumers. The net effect of DR would be minimal, if detectable at all. To address this concern, customer-specific data would be needed for additional analysis. The analysis should also differentiate along sectoral lines

and estimate how the demand for electricity in each sector (residential, commercial and industrial) behaved during demand response events. Existing literature suggests the need for specific baselines at a sectoral level (Yan et al., 2018), as consumers in each sector may respond differently to demand response initiatives, and the intensity of response may vary. By taking a high-level view, valuable information may be diluted, making it hard to capture underlying nuances of sector-specific adjustments.

4. The electricity generation from in-state power plants may increase above the usual levels during grid emergencies. This additional electricity could to some extent crowd out electricity imports due to grid congestion charges and reduce electricity imports beyond what demand response initiatives could be associated with. To test this hypothesis and verify whether congestion charges play a significant role during grid emergencies, locational marginal prices and rates of congestion charges would need to be incorporated in addition to narrowing the scope to smaller geographic units (counties or districts surrounding grid nodes).
5. The heatwaves that triggered grid emergencies in California in 2022 may have affected other regional grids in the Western Energy Imbalance Market and the demand for electricity could have exceeded normal levels in the entire region of western United States. This would result in less electricity available for export to California, hence lower imports. A reduction in imports when DR was in effect would therefore not be fully attributable to demand response initiatives.
6. The analysis does not consider the influence of wholesale electricity prices, which are a crucial factor in understanding the response of generation and interchange to demand response events. Electricity generation and interchange are largely governed by prevailing prices in the day-ahead and spot markets. Pricing elements are fully absent in the models utilized in the study, though. The complexity of incorporating price into the analysis goes beyond the scope of this study. Determining which price to look at poses a challenge, as different generators in and out of the state may be receiving day-ahead prices or participate in spot markets with yet again different sets of prices. Moreover, locational marginal pricing adds a layer of complexity, as the physical distance between where the electricity is generated and where it is consumed affects prices (e.g. due to congestion charges). With hundreds of nodes in the CAISO grid quoting different wholesale prices to suppliers and traders, it would require a careful design to enable the incorporation of wholesale pricing into the analysis. The absence of price data limits our ability to fully comprehend the adjustments in total

interchange and net generation in the current study as these adjustments are driven by price dynamics. A potential remedy to the absence of pricing would be narrowing the focus to the district (county) or nodal level and incorporating locational marginal pricing into the working models.

The selected methodology of SARIMAX-GARCH modeling also has its limitations that became apparent in the study. ARIMA models, which are a component of the SARIMAX model, have been found to handle extremes less successfully in certain cases (Verwiebe et al., 2021). Demand response days can certainly be considered extreme in some ways. Machine learning approaches, on the other hand, have the potential to handle complexity and non-linearity better. Exploring machine learning techniques could be a promising avenue for future research, but it does require considerable investment in becoming acquainted with the methods. Additionally, it should be noted that authors using GARCH models have not always been transparent about their residual analyses, and it remains unclear whether the GARCH model in this study adequately captures the additional variation in the residuals. Also, the use of hourly data with large magnitudes as in this study can be computationally taxing on hardware, and this methodology may not be suitable for analysis performed on older hardware just as in the case of this thesis.

This study serves as a building bloc for further research. In terms of methodology, a shift towards machine learning, such as deep learning or neural networks, should be explored to better handle non-linear relationships and complex patterns in the data, especially on the hourly level (Valentini et al., 2022). These more advanced approaches may provide improved insights into the behavior of electricity demand, generation and interchange during demand response events. Additionally, building sector-specific baselines and examining the effects of demand response at a more granular level could enhance our understanding of the dynamics within different sectors and their responses to demand management strategies.

Another important limitation considering the granularity of data is the potential issue of availability. Demand response programs in California are typically administered by private entities, and gaining access to their data is not always guaranteed. The CPUC does make provisions for research that is in the public interest, but obtaining the data can be a lengthy process with uncertain outcomes. Demand response programs generate confidential and proprietary information, making it challenging to obtain comprehensive and granular data for analysis. It is worth noting that obtaining such data from jurisdictions outside of the European Union for research purposes requires the ability to navigate stringent laws and patience to complete lengthy application procedures. The poor accessibility may restrict the ability of researchers to conduct a more detailed

and thorough investigation of the effects of demand response on demand, net generation, and interchange. Future research should aim to establish collaborations and partnerships with relevant stakeholders to overcome these limitations and gain access to more comprehensive, and especially granular data sets, enabling a more robust analysis of demand response initiatives.

Acknowledging and addressing these limitations in future research can contribute to a better understanding of the effects of demand response on the electric grid, including the interactions between price dynamics, sector-specific responses, and the suitability of different modeling approaches.

9 Conclusion

This study provides insights into the impact of demand response initiatives on grid dynamics in California during the 2022 heatwaves. The deployment of the SARIMAX GARCH model to estimate counterfactual baselines and the analysis of deviations through OLS regression allowed for uncovering several interesting findings. The analysis revealed that demand during the periods preceding demand response events exceeded the predicted baselines, indicating increased demand. However, during the peak demand response period, the hourly deviations exhibited a negative trend when compared to off-peak hours on non-event days. This negative deviation suggests that the implementation of peak load management was successful in reducing demand during the designated response period.

In terms of net generation, the study found that it on average exceeded the predicted baseline during the pre-demand response periods, relative to other hours on non-event days. This observation indicates that the increased demand during this period was met by an increase in in-state generation, highlighting the capacity of the grid to meet the shifted demand.

Regarding total electricity interchange with neighboring balancing authorities, the study revealed that the hours preceding peak demand response periods did not show any meaningful deviations from the baseline. However, during the peak demand response period, the total interchange on average exceeded the predicted baseline from above. This suggests that the reduction in peak demand that was observed was primarily reflected in lower imports of electricity. The findings of this study have implications for a variety of market participants, including generators, wholesale traders, grid operators and policy makers. By understanding the grid-level dynamics during demand response events, market participants can make better informed decisions regarding generation capacity, trading strategies, and grid management. This knowledge can contribute to the improved efficiency and stability of the electricity market in California.

References

- Albadi, M. H., & El-Saadany, E. F. (2008). A summary of demand response in electricity markets. *Electric Power Systems Research*, 78, 1989–1996. <https://doi.org/10.1016/J.EPSR.2008.04.002>
- Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 542–547.
- California Independent System Operator, CAISO. (n.d.-a). CAISO Emergency Warning Notifications. Retrieved February 28, 2023, from <https://www.caiso.com/informed/Pages/Notifications/NoticeLog.aspx>
- California Independent System Operator, CAISO. (n.d.-b). Understanding the ISO. Retrieved March 8, 2023, from <https://www.caiso.com/about/Pages/OurBusiness/Default.aspx>
- California Public Utilities Commission. (2021a). Electric Rates. Retrieved April 14, 2023, from <https://www.cpuc.ca.gov/industries-and-topics/electrical-energy/electric-rates>
- California Public Utilities Commission. (2021b). Summer Reliability. Retrieved March 19, 2023, from <https://www.cpuc.ca.gov/news-and-updates/newsroom/summer-2021-reliability>
- California Public Utilities Commission. (2022a). Demand Response (DR). Retrieved March 8, 2023, from <https://www.cpuc.ca.gov/industries-and-topics/electrical-energy/electric-costs/demand-response-dr>
- California Public Utilities Commission. (2022b). Emergency Load Reduction Program. Retrieved March 8, 2023, from <https://www.cpuc.ca.gov/industries-and-topics/electrical-energy/electric-costs/demand-response-dr/emergency-load-reduction-program>
- Chao, H.-p. (2011). Demand response in wholesale electricity markets: The choice of customer baseline. *Journal of Regulatory Economics*, 39(1), 68–88.
- Coughlin, K., Piette, M. A., Goldman, C., & Kiliccote, S. (2008). *Estimating demand response load impacts: Evaluation of baselineload models for non-residential buildings in california* (tech. rep.). Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
- Demand Side Analytics, LLC. (2023). *Statewide Residential Emergency Load Reduction Program Baseline Evaluation* (tech. rep.). Retrieved March 8, 2023, from <https://rmi.org/insight/the-economics-of-demand-flexibility/>
- Department of Atmospheric Sciences, University of Utah. (n.d.). MesoWest Data. Retrieved February 28, 2023, from <https://mesowest.utah.edu/>
- Dyson, Mark, James Mandel, et al. (2015). *The Economics of Demand Flexibility: How “flexiwatts” create quantifiable value for customers and the grid* (tech. rep.). Retrieved March 8, 2023, from <https://rmi.org/insight/the-economics-of-demand-flexibility/>
- Enders, W. (2013). *Applied Econometric Time Series*. John Wiley & Sons.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.
- Gabaldón, A., Garcia-Garre, A., Ruiz-Abellón, M., Guillamón, A., Álvarez-Bel, C., & Fernandez-Jimenez, L. (2021). Improvement of customer baselines for the evaluation of demand response through the use of physically-based load models. *Utilities Policy*, 70, 101213.
- Gagne, D. A., Settle, D. E., Aznar, A. Y., & Bracho, R. (2018). *Demand response compensation methodologies: Case studies for Mexico* (tech. rep.). National Renewable Energy Lab.(NREL), Golden, CO (United States).
- Garcia, R. C., Contreras, J., Van Akkeren, M., & Garcia, J. B. C. (2005). A GARCH forecasting model to predict day-ahead electricity prices. *IEEE transactions on power systems*, 20(2), 867–874.
- Gerke, B., Gallo, G., Smith, S., Liu, J., Alstone, P., Raghavan, S., Schwartz, P., Piette, M. A., Yin, R., & Stensson, S. (2020). The California demand response potential study, phase 3: Final report on the shift resource through 2030.
- Goldman, C. A., Barbose, G. L., & Eto, J. H. (2002). California customer load reductions during the electricity crisis: Did they help to keep the lights on? *Journal of Industry, Competition and trade*, 2, 113–142.
- Gyamfi, S., Krumdieck, S., & Urmee, T. (2013). Residential peak electricity demand response - Highlights of some behavioural issues. *Renewable and Sustainable Energy Reviews*, 25, 71–77.

- Heshmati, A. (2014). Demand, customer base-line and demand response in the electricity market: A survey. *Journal of Economic Surveys*, 28(5), 862–888.
- Ito, K., Ida, T., & Tanaka, M. (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, 10(1), 240–267.
- Jazaeri, J., Alpcan, T., Gordon, R., Brandao, M., Hoban, T., & Seeling, C. (2016). Baseline methodologies for small scale residential demand response. *2016 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia)*, 747–752.
- Khan, S., Javaid, N., Chand, A., Khan, A. B. M., Rashid, F., & Afridi, I. U. (2019). Electricity load forecasting for each day of week using deep cnn. *Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019) 33*, 1107–1119.
- Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., & Zhang, Y. (2017). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE transactions on smart grid*, 10(1), 841–851.
- Lawrence Berkeley National Laboratory. (2020). Demand Flexibility, Energy Technologies Area. Retrieved March 8, 2023, from <https://buildings.lbl.gov/demand-flexibility>
- Leinauer, C., Schott, P., Fridgen, G., Keller, R., Ollig, P., & Weibelzahl, M. (2022). Obstacles to demand response: Why industrial companies do not adapt their power consumption to volatile power generation. *Energy Policy*, 165, 112876.
- McKane, A. T., Piette, M. A., Faulkner, D., Ghatikar, G., Radspieler Jr, A., Adesola, B., Mur-tishaw, S., & Kiliccote, S. (2008). *Opportunities, barriers and actions for industrial demand response in California* (tech. rep.). Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).
- Moezzi, M., Gossard, M. H., & Lutzenhiser, L. (2002). Using mass media to influence energy consumption behavior: California’s 2001 Flex Your Power campaign as a case study. *Teaming for Efficiency: Human and social dimensions of energy use: understanding markets and demand*, 8, 15.
- Paterakis, N. G., Erdinç, O., & Catalão, J. P. (2017). An overview of demand response: Key-elements and international experience. *Renewable and Sustainable Energy Reviews*, 69, 871–891. <https://doi.org/10.1016/J.RSER.2016.11.167>
- Pennsylvania State University. (2021). 11.1 ARCH/GARCH Models — STAT 510. Retrieved February 27, 2023, from <https://online.stat.psu.edu/stat510/lesson/11/11.1>
- Sadeghi-Mobarakeh, A., Kohansal, M., Papalexakis, E. E., & Mohsenian-Rad, H. (2017). Data mining based on random forest model to predict the California ISO day-ahead market prices. *2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 1–5.
- Shumway, R. H., Stoffer, D. S., & Stoffer, D. S. (2000). *Time Series Analysis and Its Applications* (Vol. 3). Springer.
- Siano, P., & Sarno, D. (2016). Assessing the benefits of residential demand response in a real time distribution energy market. *Applied Energy*, 161, 533–551.
- Southern California Edison. (2022). Emergency Load Reduction Program (ELRP): California IOUs Taking Actions to Prepare for Potential Extreme Weather in Summers 2022 and 2023. Retrieved March 8, 2023, from <https://2022.utilityforum.org/Data/Sites/7/media/presentations/martinez--elrp-overview.pdf>
- Spees, K., & Lave, L. B. (2007). Demand response and electricity market efficiency. *The Electricity Journal*, 20(3), 69–85.
- United States Environmental Protection Agency. (2020). U.S. Electricity Grid Markets. Retrieved April 8, 2023, from <https://www.epa.gov/green-power-markets/us-electricity-grid-markets>
- U.S. Department of Energy. (2022). Demand Response. Retrieved March 6, 2023, from <https://www.energy.gov/oe/demand-response>
- U.S. Energy Information Administration. (n.d.). Real-time Operating Hourly Electric Grid Monitor. Retrieved February 10, 2023, from https://www.eia.gov/electricity/gridmonitor/dashboard/electric_overview/US48/US48

- U.S. Energy Information Administration. (2020). California was the largest net electricity importer of any state in 2019. Retrieved April 1, 2023, from <https://www.eia.gov/todayinenergy/detail.php?id=46156>
- U.S. Energy Information Administration. (2021a). Annual Electric Power Industry Report, Form EIA-861. Retrieved March 20, 2023, from <https://www.eia.gov/electricity/data/eia861/>
- U.S. Energy Information Administration. (2021b). Electricity Data Browser [Demand Response Research Center at the Lawrence Berkeley National Laboratory]. Retrieved March 20, 2023, from <https://tinyurl.com/2ftwh7y5>
- Valentini, O., Andreadou, N., Bertoldi, P., Lucas, A., Saviuc, I., & Kotsakis, E. (2022). Demand response impact evaluation: A review of methods for estimating the customer baseline load. *Energies*, 15(14), 5259.
- Verwiebe, P. A., Seim, S., Burges, S., Schulz, L., & Müller-Kirchenbauer, J. (2021). Modeling energy demand—a systematic literature review. *Energies*, 14(23), 7859.
- Vilar, J. M., Cao, R., & Aneiros, G. (2012). Forecasting next-day electricity demand and price using nonparametric functional methods. *International Journal of Electrical Power & Energy Systems*, 39(1), 48–55.
- Wang, Y., Wang, J., Zhao, G., & Dong, Y. (2012). Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China. *Energy Policy*, 48, 284–294.
- Widergren, S., Marinovici, C., Berliner, T., & Graves, A. (2012). Real-time pricing demand response in operations. *2012 IEEE Power and Energy Society General Meeting*, 1–5.
- Wu, H., Shahidehpour, M., Alabdulwahab, A., & Abusorrah, A. (2015). Demand response exchange in the stochastic day-ahead scheduling with variable renewable generation. *IEEE Transactions on Sustainable Energy*, 6(2), 516–525.
- Yan, X., Ozturk, Y., Hu, Z., & Song, Y. (2018). A review on price-driven residential demand response. *Renewable and Sustainable Energy Reviews*, 96, 411–419.
- Yukseltan, E., Yucekaya, A., & Bilge, A. H. (2020). Hourly electricity demand forecasting using Fourier analysis with feedback. *Energy Strategy Reviews*, 31, 100524.
- Zhang, N., Li, Z., Zou, X., & Quiring, S. M. (2019). Comparison of three short-term load forecast models in Southern California. *Energy*, 189, 116358.

A Appendix I

A.1 Demand Response Events in 2022

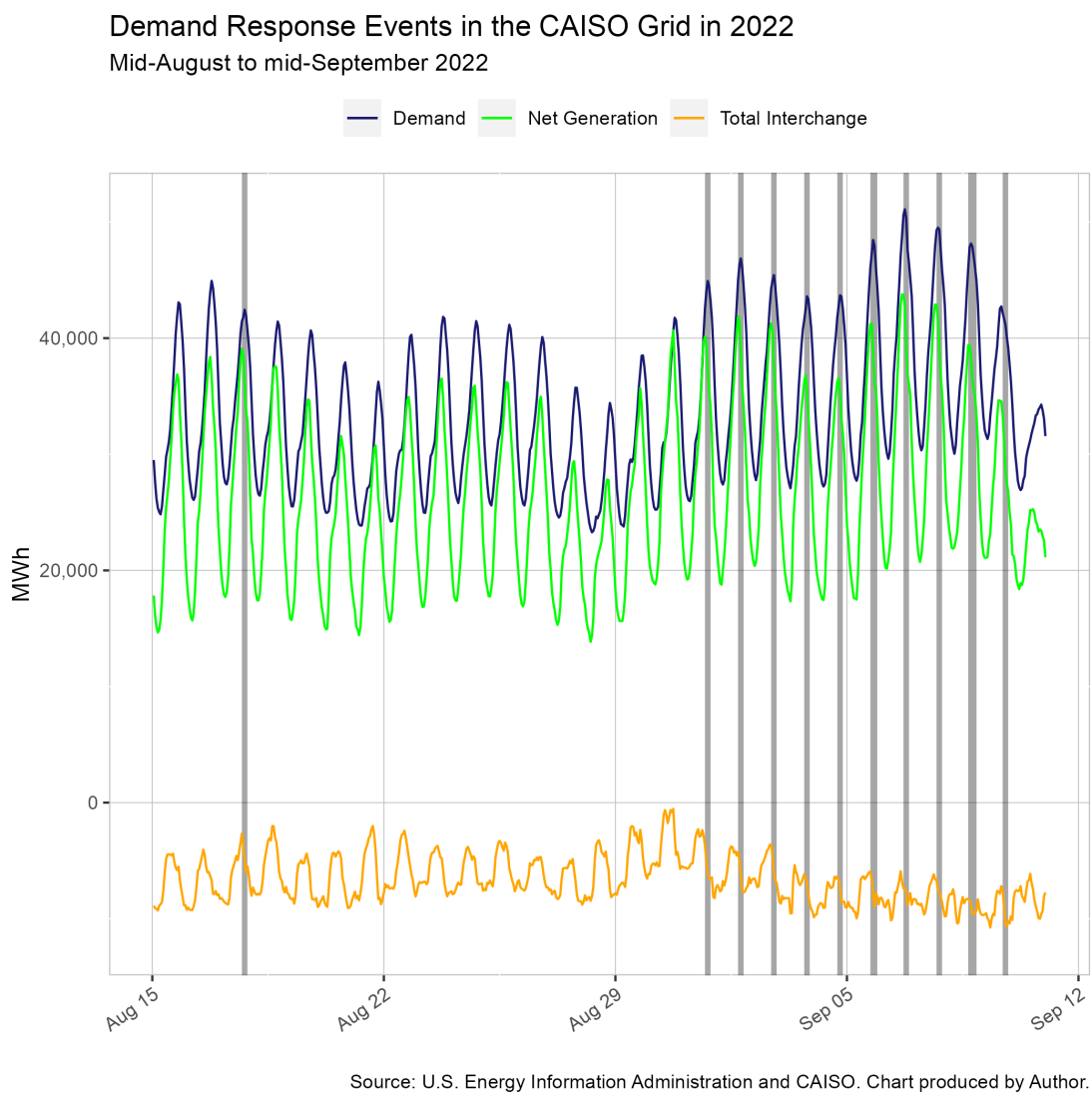


Figure 8: Emergency and Demand Response Days in 2022

A.2 Demand for Electricity

Appendix A Table 1: GARCH Model fitted to Demand Residuals	
*	*
GARCH Model Fit	*
*	*
Conditional Variance Dynamics	
GARCH Model	: sGARCH(1,2)
Mean Model	: ARFIMA(4,0,1)
Distribution	: norm

Optimal Parameters

	Estimate	Std. Error	t value	Pr(> t)
mu	12.989987	38.664176	0.335970	0.736894
ar1	0.894602	0.038968	22.957383	0.000000
ar2	0.006317	0.035597	0.177448	0.859156
ar3	0.057620	0.035157	1.638918	0.101230
ar4	-0.119530	0.028420	-4.205777	0.000026
ma1	-0.929802	0.029608	-31.404235	0.000000
mxreg1	-16.840918	54.597575	-0.308455	0.757736
mxreg2	-12.162553	54.749247	-0.222150	0.824197
mxreg3	-25.862622	52.667219	-0.491057	0.623386
mxreg4	-19.403941	56.669056	-0.342408	0.732044
mxreg5	-6.837892	55.937313	-0.122242	0.902707
mxreg6	-11.117219	55.693813	-0.199613	0.841783
mxreg7	-22.990261	55.041359	-0.417691	0.676173
mxreg8	-21.552734	56.575600	-0.380955	0.703237
mxreg9	-18.744247	55.156062	-0.339840	0.733977
mxreg10	-20.685863	55.236432	-0.374497	0.708035
mxreg11	-24.532882	55.269420	-0.443878	0.657131
mxreg12	-9.910037	56.046380	-0.176819	0.859651
mxreg13	-6.015882	55.009730	-0.109360	0.912917
mxreg14	-0.892417	55.543705	-0.016067	0.987181
mxreg15	15.544697	55.005942	0.282600	0.777483
mxreg16	4.468969	55.559346	0.080436	0.935891
mxreg17	3.055180	55.615845	0.054934	0.956191
mxreg18	-22.159514	55.819433	-0.396986	0.691378
mxreg19	-17.328743	55.808036	-0.310506	0.756176
mxreg20	-38.058864	56.220615	-0.676956	0.498434
mxreg21	-20.619006	53.082050	-0.388437	0.697693
mxreg22	-27.181300	54.778392	-0.496205	0.619750
mxreg23	-13.258068	54.498198	-0.243275	0.807792
omega	556.030888	562.220358	0.988991	0.322668
alpha1	0.057075	0.016257	3.510684	0.000447
beta1	0.161692	0.024305	6.652659	0.000000
beta2	0.776997	0.024512	31.698204	0.000000

Robust Standard Errors:

	Estimate	Std. Error	t value	Pr(> t)
mu	12.989987	22.023589	0.589822	0.555310
ar1	0.894602	0.045148	19.814899	0.000000
ar2	0.006317	0.042684	0.147982	0.882357
ar3	0.057620	0.030778	1.872095	0.061193
ar4	-0.119530	0.027872	-4.288521	0.000018
ma1	-0.929802	0.031766	-29.270231	0.000000
mxreg1	-16.840918	33.443467	-0.503564	0.614568
mxreg2	-12.162553	29.446699	-0.413036	0.679580
mxreg3	-25.862622	32.987020	-0.784024	0.433026
mxreg4	-19.403941	32.550370	-0.596120	0.551095
mxreg5	-6.837892	33.190123	-0.206022	0.836774
mxreg6	-11.117219	47.156635	-0.235751	0.813626
mxreg7	-22.990261	58.477128	-0.393150	0.694209
mxreg8	-21.552734	44.653898	-0.482662	0.629336
mxreg9	-18.744247	54.483566	-0.344035	0.730820
mxreg10	-20.685863	54.695626	-0.378200	0.705282

mxreg11	-24.532882	65.997500	-0.371724	0.710098
mxreg12	-9.910037	53.231064	-0.186170	0.852311
mxreg13	-6.015882	60.292024	-0.099779	0.920520
mxreg14	-0.892417	57.197111	-0.015602	0.987552
mxreg15	15.544697	47.241755	0.329046	0.742121
mxreg16	4.468969	46.750252	0.095592	0.923844
mxreg17	3.055180	45.611136	0.066983	0.946595
mxreg18	-22.159514	47.985383	-0.461797	0.644227
mxreg19	-17.328743	44.415251	-0.390153	0.696423
mxreg20	-38.058864	41.119369	-0.925570	0.354669
mxreg21	-20.619006	35.296020	-0.584174	0.559103
mxreg22	-27.181300	34.344802	-0.791424	0.428697
mxreg23	-13.258068	42.395712	-0.312722	0.754492
omega	556.030888	681.546988	0.815836	0.414594
alpha1	0.057075	0.035750	1.596487	0.110380
beta1	0.161692	0.019283	8.385069	0.000000
beta2	0.776997	0.014015	55.438768	0.000000

LogLikelihood : -10435.21

Information Criteria

Akaike	14.301
Bayes	14.420
Shibata	14.300
Hannan-Quinn	14.345

Weighted Ljung-Box Test on Standardized Residuals

	statistic	p-value
Lag [1]	1.387e-04	9.906e-01
Lag [2*(p+q)+(p+q)-1][14]	1.039e+01	6.594e-06
Lag [4*(p+q)+(p+q)-1][24]	2.654e+01	1.827e-05

d.o.f=5
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals

	statistic	p-value
Lag [1]	32.72	1.065e-08
Lag [2*(p+q)+(p+q)-1][8]	51.43	1.913e-13
Lag [4*(p+q)+(p+q)-1][14]	62.42	1.199e-14

d.o.f=3

Weighted ARCH LM Tests

	Statistic	Shape	Scale	P-Value
ARCH Lag [4]	0.315	0.500	2.000	0.5747
ARCH Lag [6]	2.918	1.461	1.711	0.3201
ARCH Lag [8]	4.263	2.368	1.583	0.3381

Nyblom stability test

Joint Statistic: no.parameters>20 (not available)
Individual Statistics:

mu	0.02931
ar1	0.12594
ar2	0.11946
ar3	0.08867
ar4	0.03395
ma1	0.13778
mxreg1	0.14148
mxreg2	0.04862
mxreg3	0.25682
mxreg4	0.07196
mxreg5	0.12251
mxreg6	0.01845
mxreg7	0.07164
mxreg8	0.69621
mxreg9	0.15044
mxreg10	0.26670
mxreg11	0.34905
mxreg12	0.10094
mxreg13	0.27423
mxreg14	0.06590
mxreg15	0.04029
mxreg16	0.17130
mxreg17	0.07650
mxreg18	0.05085
mxreg19	0.10873
mxreg20	1.01389
mxreg21	1.76671
mxreg22	0.82768
mxreg23	0.22302
omega	0.07850
alpha1	0.04178
beta1	0.07058
beta2	0.07319

Asymptotic Critical Values (10% 5% 1%)
Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

Adjusted Pearson Goodness-of-Fit Test:

	group	statistic	p-value(g-1)
1	20	51.66	7.448e-05
2	30	69.85	3.176e-05
3	40	75.95	3.617e-04
4	50	88.19	5.055e-04

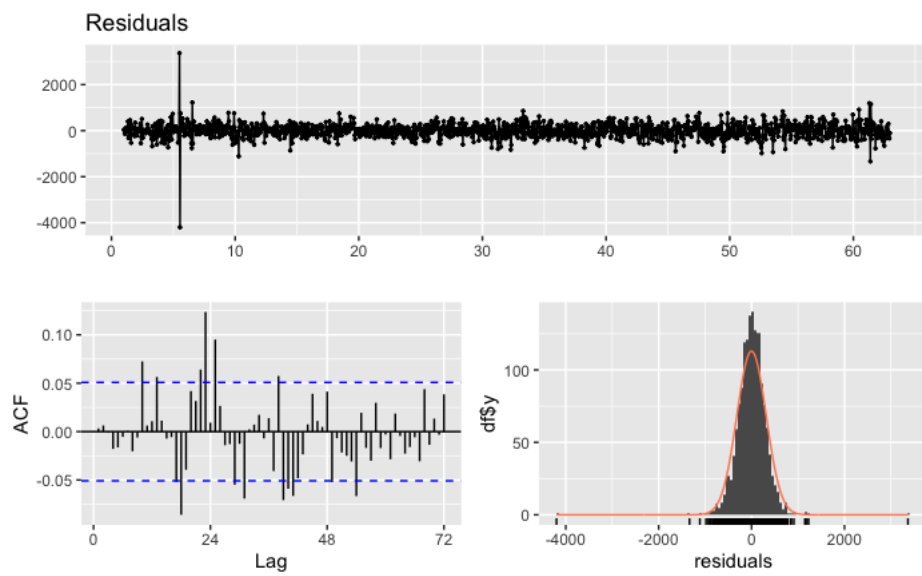


Figure 9: Residuals Diagnostics - Demand Time Series

A.3 Net Generation

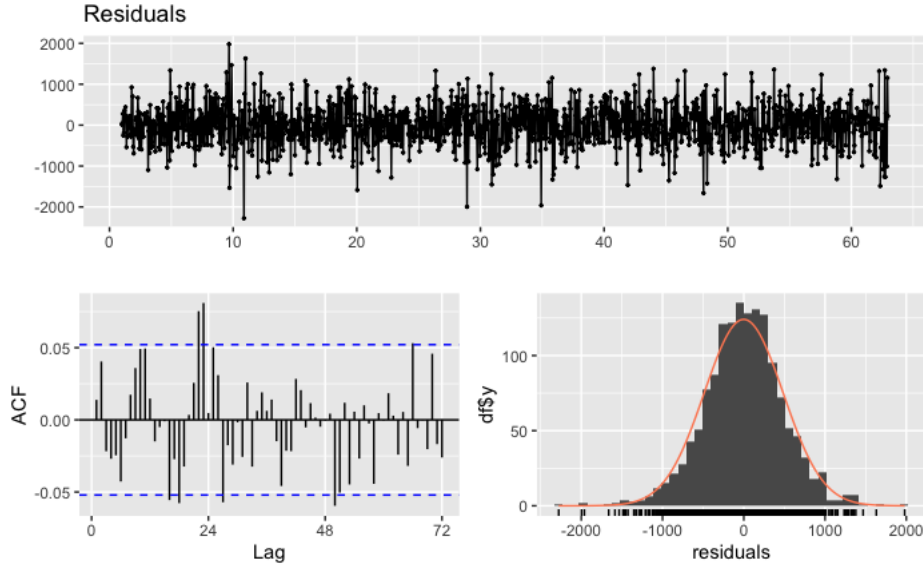


Figure 10: Residuals Diagnostics - Net Generation Time Series

Appendix A Table 2: GARCH Model fitted to Net Generation Residuals

* GARCH Model Fit *				

Conditional Variance Dynamics				

GARCH Model	:	sGARCH(1,2)		
Mean Model	:	ARFIMA(3,0,3)		
Distribution	:	norm		
Optimal Parameters				

	Estimate	Std. Error	t value	Pr(> t)
mu	-6.2030e+01	9.6183e+01	-0.64491	0.518984
ar1	-2.2712e+00	1.0350e-03	-2193.39431	0.000000
ar2	-1.6507e+00	9.3700e-04	-1762.34323	0.000000
ar3	-3.4606e-01	7.8300e-04	-441.89912	0.000000
ma1	2.2637e+00	3.4200e-04	6612.51515	0.000000
ma2	1.6310e+00	3.7400e-04	4361.67187	0.000000
ma3	3.2889e-01	6.3000e-05	5238.62683	0.000000
mxreg1	2.7975e+00	4.2466e+00	0.65876	0.510049
omega	1.3246e+04	2.6877e+03	4.92851	0.000001
alpha1	4.4981e-02	1.5354e-02	2.92957	0.003394
beta1	3.9772e-01	1.9859e-01	2.00270	0.045210
beta2	5.0182e-01	1.8628e-01	2.69384	0.007063
Robust Standard Errors:				
	Estimate	Std. Error	t value	Pr(> t)
mu	-6.2030e+01	1.0314e+02	-0.60140	0.547573
ar1	-2.2712e+00	4.2740e-03	-531.35718	0.000000
ar2	-1.6507e+00	3.4390e-03	-479.94594	0.000000

ar3	−3.4606e−01	1.2620e−03	−274.11190	0.000000
ma1	2.2637e+00	7.2800e−04	3109.45296	0.000000
ma2	1.6310e+00	7.8800e−04	2068.92359	0.000000
ma3	3.2889e−01	1.3600e−04	2419.84654	0.000000
mxreg1	2.7975e+00	4.7258e+00	0.59196	0.553876
omega	1.3246e+04	6.8632e+03	1.93004	0.053601
alpha1	4.4981e−02	3.0862e−02	1.45749	0.144981
beta1	3.9772e−01	1.3136e−01	3.02773	0.002464
beta2	5.0182e−01	1.2695e−01	3.95296	0.000077

LogLikelihood : −11125.06

Information Criteria

Akaike	15.215
Bayes	15.258
Shibata	15.214
Hannan–Quinn	15.231

Weighted Ljung–Box Test on Standardized Residuals

	statistic	p-value
Lag [1]	0.02088	8.851e−01
Lag [2*(p+q)+(p+q)−1][17]	10.86720	1.488e−03
Lag [4*(p+q)+(p+q)−1][29]	35.27425	8.320e−08
d.o.f=6		

H0 : No serial correlation

Weighted Ljung–Box Test on Standardized Squared Residuals

	statistic	p-value
Lag [1]	0.8314	0.3619
Lag [2*(p+q)+(p+q)−1][8]	2.7234	0.7360
Lag [4*(p+q)+(p+q)−1][14]	5.8378	0.6606
d.o.f=3		

Weighted ARCH LM Tests

	Statistic	Shape	Scale	P-Value
ARCH Lag [4]	0.3739	0.500	2.000	0.5409
ARCH Lag [6]	0.7854	1.461	1.711	0.8099
ARCH Lag [8]	1.6517	2.368	1.583	0.8108

Nyblom stability test

Joint Statistic: 2.6935

Individual Statistics:

mu	0.53907
ar1	0.04540
ar2	0.05167
ar3	0.08381
ma1	0.06722
ma2	0.04283
ma3	0.04208
mxreg1	0.71453

omega	0.18216
alpha1	0.18793
beta1	0.19210
beta2	0.19234
Asymptotic Critical Values (10% 5% 1%)	
Joint Statistic:	2.69 2.96 3.51
Individual Statistic:	0.35 0.47 0.75
Sign Bias Test	

Adjusted Pearson Goodness-of-Fit Test:

	group	statistic	p-value (g-1)
1	20	23.08	0.2340
2	30	33.66	0.2518
3	40	39.33	0.4550
4	50	55.40	0.2461

A.4 Total Interchange

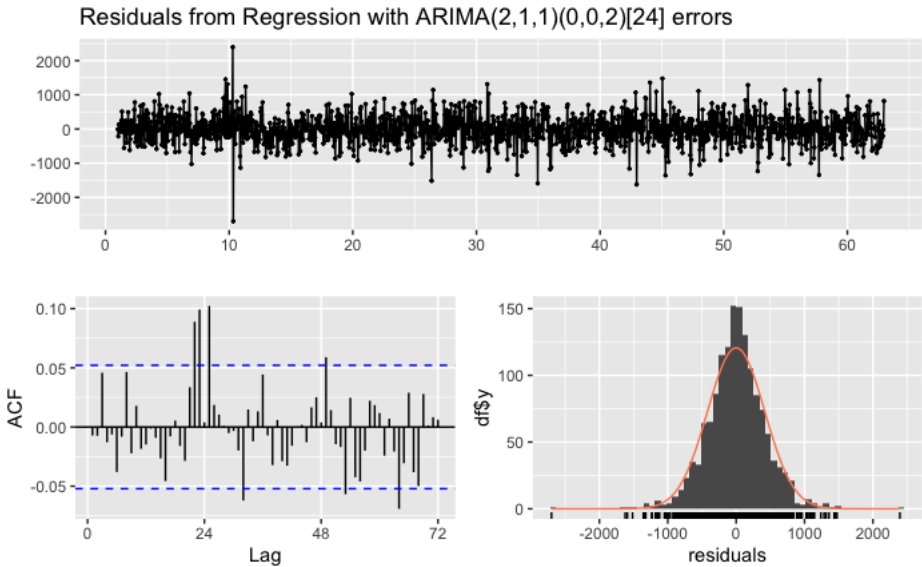


Figure 11: Residuals Diagnostics - Total Interchange Time Series

Appendix A Table 3: GARCH Model fitted to Total Interchange

*		*
*	GARCH Model Fit	*
*		*
Conditional Variance Dynamics		
GARCH Model	:	sGARCH(2,2)
Mean Model	:	ARFIMA(4,0,2)
Distribution	:	norm

Optimal Parameters

	Estimate	Std. Error	t value	Pr(> t)
mu	134.758815	81.073042	1.6622e+00	0.096475
ar1	-1.089543	0.026384	-4.1296e+01	0.000000
ar2	-0.145588	0.038081	-3.8231e+00	0.000132
ar3	-0.048799	0.039093	-1.2483e+00	0.211937
ar4	-0.020728	0.026397	-7.8524e-01	0.432315
ma1	1.106073	0.000001	8.4401e+05	0.000000
ma2	0.115995	0.002753	4.2127e+01	0.000000
mxreg1	-5.834937	3.536890	-1.6497e+00	0.098997
omega	195.410443	485.601196	4.0241e-01	0.687383
alpha1	0.000000	0.008404	6.0000e-06	0.999995
alpha2	0.000000	0.008228	1.8000e-05	0.999986
beta1	0.116009	0.000171	6.7676e+02	0.000000
beta2	0.882991	0.000005	1.7234e+05	0.000000

Robust Standard Errors:

	Estimate	Std. Error	t value	Pr(> t)
mu	134.758815	65.757703	2.0493e+00	0.040430
ar1	-1.089543	0.030900	-3.5260e+01	0.000000
ar2	-0.145588	0.039988	-3.6408e+00	0.000272
ar3	-0.048799	0.036695	-1.3298e+00	0.183575
ar4	-0.020728	0.022938	-9.0365e-01	0.366184
ma1	1.106073	0.000002	5.6152e+05	0.000000
ma2	0.115995	0.003161	3.6695e+01	0.000000
mxreg1	-5.834937	2.920389	-1.9980e+00	0.045717
omega	195.410443	711.725751	2.7456e-01	0.783655
alpha1	0.000000	0.011283	4.0000e-06	0.999997
alpha2	0.000000	0.010654	1.4000e-05	0.999989
beta1	0.116009	0.000171	6.7872e+02	0.000000
beta2	0.882991	0.000002	5.5256e+05	0.000000

LogLikelihood : -10946.84

Information Criteria

Akaike	14.972
Bayes	15.019
Shibata	14.972
Hannan-Quinn	14.990

Weighted Ljung-Box Test on Standardized Residuals

	statistic	p-value
Lag[1]	2.583e-08	0.9999
Lag[2*(p+q)+(p+q)-1][17]	3.982e+00	1.0000
Lag[4*(p+q)+(p+q)-1][29]	1.862e+01	0.1098
d.o.f=6		
H0 : No serial correlation		

Weighted Ljung-Box Test on Standardized Squared Residuals

statistic	p-value
-----------	---------

Lag [1]	1.301	0.2541
Lag [2*(p+q)+(p+q)-1][11]	8.541	0.1749
Lag [4*(p+q)+(p+q)-1][19]	12.060	0.2596

d.o.f=4

Weighted ARCH LM Tests

	Statistic	Shape	Scale	P-Value
ARCH Lag [5]	0.00697	0.500	2.000	0.9335
ARCH Lag [7]	0.95061	1.473	1.746	0.7712
ARCH Lag [9]	2.57017	2.402	1.619	0.6470

Nyblom stability test

Joint Statistic: 2.1552

Individual Statistics:

mu	0.04700
ar1	0.07602
ar2	0.10654
ar3	0.11690
ar4	0.14125
ma1	0.09205
ma2	0.11744
mxreg1	0.04177
omega	0.07304
alpha1	0.06736
alpha2	0.06973
beta1	0.07296
beta2	0.07296

Asymptotic Critical Values (10% 5% 1%)

Joint Statistic: 2.89 3.15 3.69

Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

Adjusted Pearson Goodness-of-Fit Test:

group	statistic	p-value(g-1)
1	20	40.34
2	30	60.96
3	40	64.20
4	50	71.31

B Appendix II

Table 11: Demand for Electricity (MWh) - Coefficient estimates of the best models by Mean Absolute Percentage Error

	Model 10	Model 5	Model 4
ar1	1.78*** (0.02)	1.78*** (0.02)	1.78*** (0.02)
ar2	-0.85*** (0.02)	-0.85*** (0.02)	-0.85*** (0.02)
ma1	-1.42*** (0.04)	-1.42*** (0.04)	-1.42*** (0.04)
ma2	0.49*** (0.04)	0.49*** (0.04)	0.49*** (0.04)
sma1	0.21*** (0.03)	0.21*** (0.03)	0.21*** (0.03)
Hour 1	-1413.85*** (62.56)	-1410.96*** (62.56)	-1411.00*** (62.53)
Hour 2	-2408.24*** (110.53)	-2409.98*** (110.53)	-2410.48*** (110.40)
Hour 3	-2956.92*** (154.88)	-2969.32*** (155.06)	-2970.35*** (154.83)
Hour 4	-3086.15*** (195.42)	-3103.39*** (195.71)	-3105.03*** (195.40)
Hour 5	-2624.54*** (231.83)	-2642.42*** (232.11)	-2644.57*** (231.74)
Hour 6	-1865.67*** (263.82)	-1886.87*** (264.16)	-1889.31*** (263.76)
Hour 7	-853.34** (291.18)	-876.55** (291.53)	-878.95** (291.12)
Hour 8	122.43 (313.74)	98.05 (314.09)	95.89 (313.67)
Hour 9	651.67* (331.40)	626.71 (331.73)	624.66 (331.32)
Hour 10	1241.33*** (344.06)	1217.34*** (344.36)	1215.33*** (343.97)
Hour 11	2223.72*** (351.69)	2201.89*** (351.94)	2199.83*** (351.57)
Hour 12	3294.61*** (354.25)	3276.07*** (354.44)	3274.00*** (354.09)
Hour 13	4721.10*** (351.73)	4707.77*** (351.85)	4705.55*** (351.51)
Hour 14	6339.43*** (344.12)	6331.27*** (344.20)	6329.04*** (343.86)
Hour 15	7950.41*** (331.46)	7945.91*** (331.51)	7943.61*** (331.18)
Hour 16	9190.05*** (313.80)	9187.31*** (313.84)	9184.90*** (313.50)
Hour 17	10110.26*** (291.20)	10109.10*** (291.23)	10106.41*** (290.90)
Hour 18	10263.04*** (263.80)	10262.76*** (263.83)	10260.03*** (263.49)
Hour 19	9416.16*** (231.77)	9415.26*** (231.79)	9412.96*** (231.48)
Hour 20	8413.84***	8412.14***	8410.43***

Table 11: Demand for Electricity (MWh) - Coefficient estimates of the best models by Mean Absolute Percentage Error

	Model 10	Model 5	Model 4
	(195.35)	(195.37)	(195.09)
Hour 21	6892.22***	6889.96***	6888.90***
	(154.85)	(154.86)	(154.64)
Hour 22	4424.75***	4421.43***	4420.79***
	(110.64)	(110.65)	(110.52)
Hour 23	1929.13***	1928.98***	1928.14***
	(62.56)	(62.54)	(62.50)
Tuesday	129.64	124.46	
	(88.98)	(88.93)	
Wednesday	102.53	89.83	
	(127.93)	(128.06)	
Thursday	34.21	14.07	
	(142.12)	(142.50)	
Friday	111.91	99.49	
	(141.43)	(141.50)	
Saturday	-5.22	-5.67	
	(126.41)	(126.24)	
Sunday	-126.14	-126.65	
	(90.32)	(90.19)	
Temperature °C		5.07	5.06
		(2.77)	(2.76)
AIC	21400.38	21399.04	21394.48
AICc	21402.11	21400.87	21395.76
BIC	21586.03	21590.00	21553.61
Log Likelihood	-10665.19	-10663.52	-10667.24
Num. obs.	1487	1487	1487

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 12: Net Generation -Model Estimates of the three best models by AIC

	Model 11	Model 12	Model 13
ar1	0.28***	1.88***	1.14***
	(0.02)	(0.02)	(0.03)
ar2	0.98***	-0.94***	-0.03
	(0.02)	(0.02)	(0.04)
ar3	-0.17***		-0.10*
	(0.03)		(0.04)
ar4	-0.17***		-0.05*
	(0.03)		(0.03)
ma1	0.89***	-1.73***	
	(0.01)	(0.03)	
sar1	-0.45***		-0.37***
	(0.02)		(0.03)
sar2	-0.21***		
	(0.02)		
Temperature °C	-4.11		
	(4.58)		
Solar (MWh)	0.44***	0.44***	0.45***
	(0.04)	(0.03)	(0.04)
ma2		0.73***	
		(0.05)	

Table 12: Net Generation -Model Estimates of the three best models by AIC

	Model 11	Model 12	Model 13
ma3		0.05 (0.03)	
sma1		0.21*** (0.03)	
sma2		0.06* (0.03)	
Hour 1		−1210.44*** (101.54)	
Hour 2		−1977.70*** (173.83)	
Hour 3		−2283.38*** (240.81)	
Hour 4		−2168.61*** (303.02)	
Hour 5		−1652.08*** (359.96)	
Hour 6		−1081.51** (419.22)	
Hour 7		−178.49 (506.54)	
Hour 8		536.89 (576.21)	
Hour 9		1144.00 (621.69)	
Hour 10		1972.47** (650.53)	
Hour 11		3119.27*** (664.54)	
Hour 12		4492.33*** (667.82)	
Hour 13		5989.29*** (660.45)	
Hour 14		7724.69*** (643.49)	
Hour 15		9126.70*** (611.32)	
Hour 16		10288.38*** (566.03)	
Hour 17		11082.97*** (491.21)	
Hour 18		11173.14*** (413.39)	
Hour 19		10031.21*** (359.66)	
Hour 20		8546.40*** (302.88)	
Hour 21		5906.76*** (240.85)	
Hour 22		3368.72*** (174.16)	
Hour 23		1443.09*** (101.29)	
Tuesday			−82.91

Table 12: Net Generation -Model Estimates of the three best models by AIC

	Model 11	Model 12	Model 13
			(137.93)
Wednesday			-154.18
			(192.53)
Thursday			-172.77
			(219.18)
Friday			-14.32
			(216.59)
Saturday			77.57
			(187.16)
Sunday			61.16
			(136.55)
AIC	22620.65	22607.18	22686.69
AICc	22620.80	22608.64	22686.94
BIC	22673.54	22776.93	22755.45
Log Likelihood	-11300.33	-11271.59	-11330.34
Num. obs.	1459	1459	1459

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$

Table 13: Total Interchange (MWh) - Coefficient estimates of the best models by AIC

	Model 11	Model 5	Model 8
ar1	1.09*** (0.03)	0.82*** (0.04)	0.92*** (0.01)
ar2	-0.20*** (0.03)		
ma1	-0.97*** (0.01)	-0.68*** (0.04)	0.19*** (0.03)
sma1	0.22*** (0.03)	0.21*** (0.03)	
sma2	0.13*** (0.03)	0.13*** (0.03)	
Hour 1	50.67 (73.18)	49.95 (72.97)	
Hour 2	-38.22 (110.03)	-41.70 (110.97)	
Hour 3	-89.04 (135.55)	-87.83 (138.26)	
Hour 4	-142.55 (154.17)	-138.94 (159.88)	
Hour 5	-358.06* (168.24)	-340.56 (175.19)	
Hour 6	-779.05*** (195.11)	-280.59 (186.74)	
Hour 7	-666.58* (278.90)	628.83** (195.31)	
Hour 8	177.31 (339.13)	1926.75*** (201.60)	
Hour 9	777.31* (371.83)	2746.67*** (206.11)	
Hour 10	1143.57** (389.60)	3227.06*** (209.06)	
Hour 11	1339.34***	3456.65***	

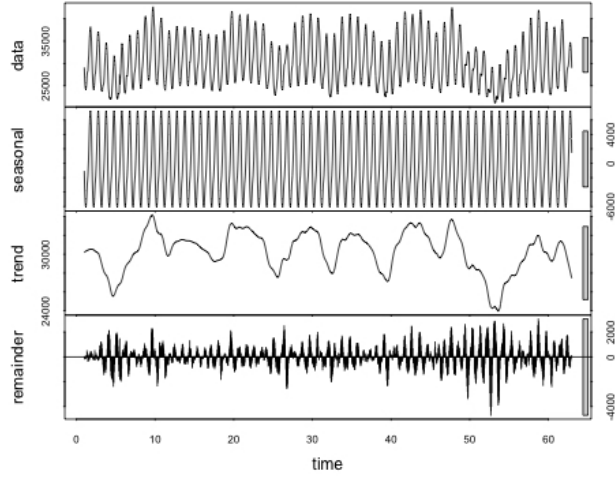
Table 13: Total Interchange (MWh) - Coefficient estimates of the best models by AIC

	Model 11	Model 5	Model 8
	(395.50)	(210.67)	
Hour 12	1527.41***	3640.74***	
	(395.46)	(211.07)	
Hour 13	1532.34***	3604.05***	
	(389.70)	(210.26)	
Hour 14	1472.40***	3475.90***	
	(379.78)	(208.33)	
Hour 15	1552.19***	3399.15***	
	(357.23)	(205.21)	
Hour 16	1536.57***	3155.47***	
	(324.71)	(200.70)	
Hour 17	1679.25***	2753.34***	
	(255.10)	(194.44)	
Hour 18	1671.21***	1956.45***	
	(185.06)	(185.99)	
Hour 19	984.78***	990.11***	
	(168.47)	(174.69)	
Hour 20	386.87*	385.95*	
	(154.43)	(159.41)	
Hour 21	-56.09	-58.93	
	(135.84)	(138.17)	
Hour 22	-407.29***	-407.42***	
	(110.45)	(111.37)	
Hour 23	-300.83***	-305.37***	
	(72.94)	(72.97)	
Solar(MWh)	0.19***		
	(0.03)		
ma2		-0.17***	
		(0.03)	
ma3		0.00	
		(0.03)	
ma4		-0.09**	
		(0.03)	
Tuesday		-469.12***	-470.45***
		(119.69)	(113.83)
Wednesday		-447.20**	-413.85**
		(166.41)	(148.38)
Thursday		-671.59***	-643.28***
		(188.51)	(158.49)
Friday		-403.72*	-385.87*
		(187.57)	(157.36)
Saturday		-145.74	-123.86
		(163.81)	(145.13)
Sunday		260.99*	265.49*
		(119.38)	(112.94)
Temperature °C		-3.29	
		(4.14)	
sar1			-0.53***
			(0.03)
sar2			-0.22***
			(0.03)
AIC	22161.77	22245.64	22170.55
AICc	22163.04	22247.69	22170.74

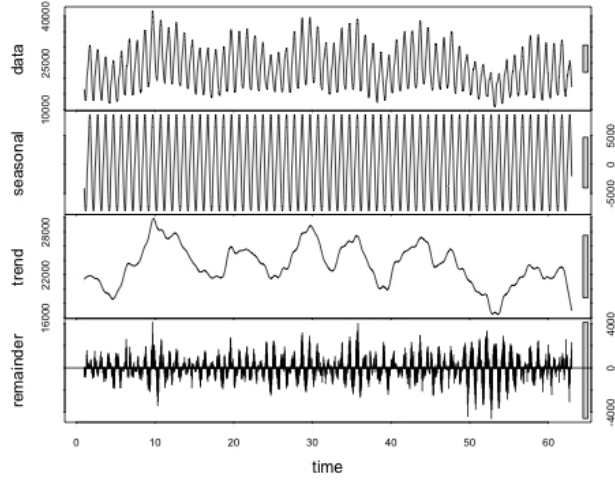
Table 13: Total Interchange (MWh) - Coefficient estimates of the best models by AIC

	Model 11	Model 5	Model 8
BIC	22320.90	22447.21	22228.73
Log Likelihood	-11050.88	-11084.82	-11074.28
Num. obs.	1459	1459	1459

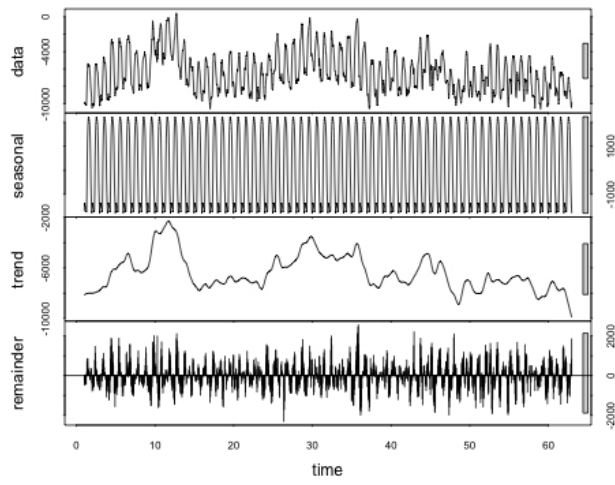
*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$



(a) Demand

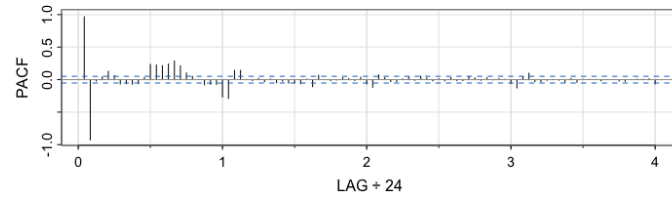
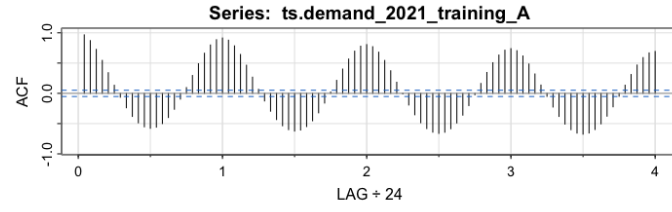


(b) Net Generation

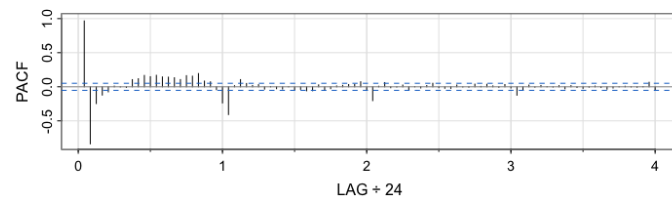
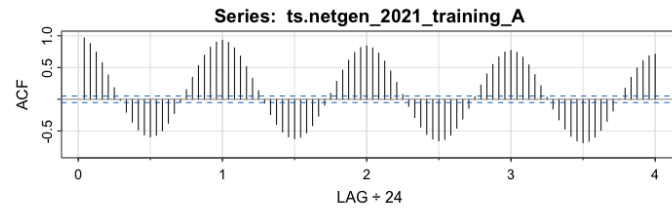


(c) Total Interchange

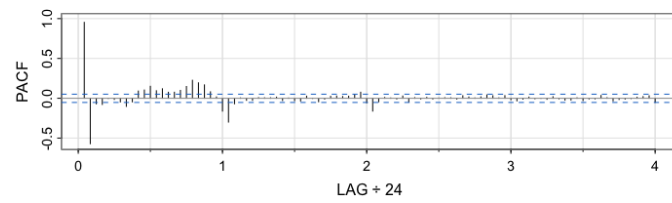
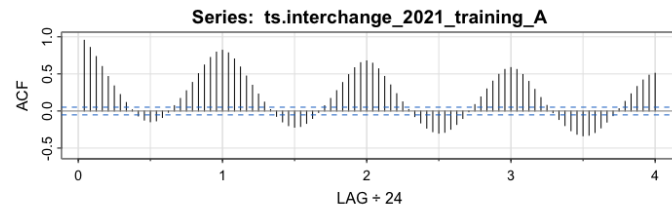
Figure 12: Decomposed Training Set Times Series



(a) Demand

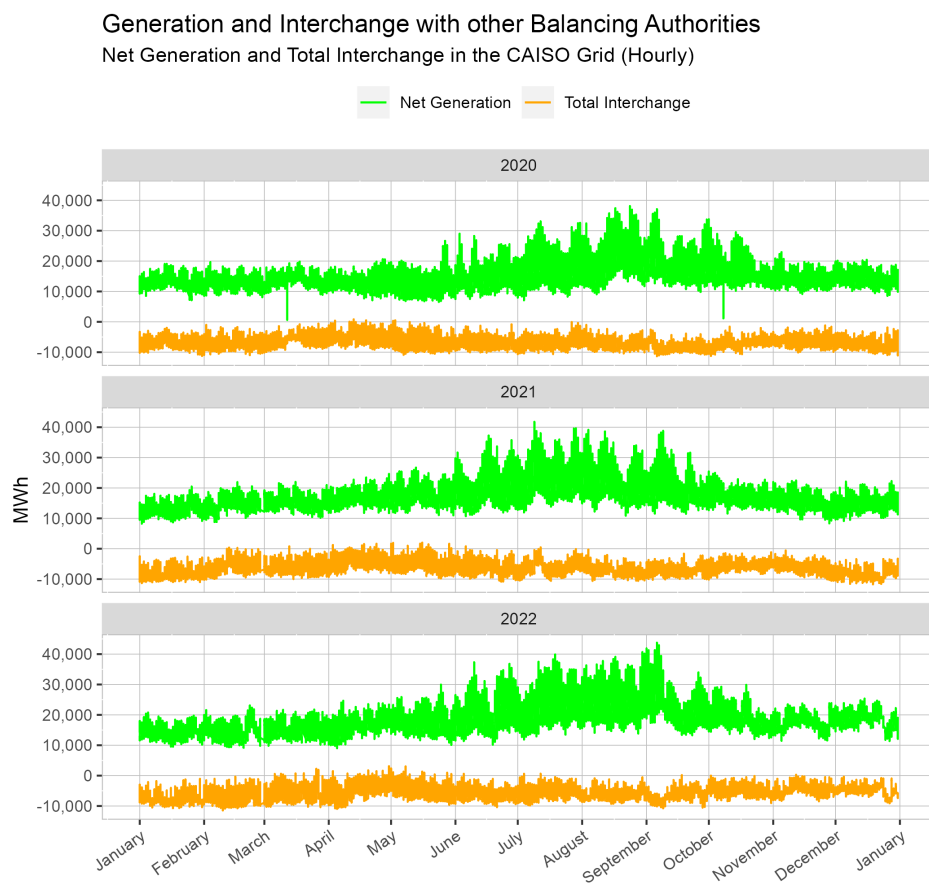


(b) Net Generation



(c) Total Interchange

Figure 13: ACF and PACF Plots of Demand, Net Generation and Interchange



Source: U.S. Energy Information Administration

Figure 14: Net Generation and Interchange with Neighboring Balancing Authorities