Stochastic Price Generation for Evaluating Wholesale Electricity Market Bidding Strategies

Nicholas D. Laws,[1,2] Dylan Cutler,[3] Hallie Dunham,[1] Sam Johnson,[4] Daniel Inman,[1] Nikita Granger,[4] Wayne Jones,[4] and David Chalenski[4]

1 National Renewable Energy Laboratory
2 The University of Texas at Austin
3 Camus Energy
2 Royal Dutch Shell

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-7A40-82005
May 2023

Stochastic Price Generation for Evaluating Wholesale Electricity Market Bidding Strategies

Nicholas D. Laws,[1,2] Dylan Cutler,[3] Hallie Dunham,[1] Sam Johnson,[4] Daniel Inman,[1] Nikita Granger,[4] Wayne Jones,[4] and David Chalenski[4]

1 National Renewable Energy Laboratory
2 The University of Texas at Austin
3 Camus Energy
2 Royal Dutch Shell

Suggested Citation
Laws, Nicholas D., Dylan Cutler, Hallie Dunham, Sam Johnson, Daniel Inman, Nikita Granger, Wayne Jones, and David Chalenski. 2023. *Stochastic Price Generation for Evaluating Wholesale Electricity Market Bidding Strategies*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-7A40-82005. https://www.nrel.gov/docs/fy23osti/82005.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Technical Report
NREL/TP-7A40-82005
May 2023

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Support for the work was also provided by Shell International and Exploration, Inc. under Agreement CW179873. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

Table of Contents

List of Figures

List of Tables

Abstract

This work presents a novel method for generating electricity price scenarios from statistical properties of past electricity prices using a hybrid statistical and reduced-form stochastic model. Previous work in applying stochastic differential equations (SDE) to model electricity prices has focused on daily average prices. To extend stochastic price generation methods to hourly or sub-hourly pricing, we address several weaknesses in the state-of-the-art: (1) we replace the mean-reversion component of the SDE with an ARIMA process that is better able to characterize the daily and weekly trends; (2) we extend the price-spike, or jump process to account for conditional probabilities of price spikes occurring in consecutive time steps by replacing the traditional Poisson process for modeling jumps with a generalized point process model inspired by brain neuron models; and (3) we replace the traditional method of estimating spike intensity with empirical variance with a Markov process based on observed price spike intensity transitions. The method is demonstrated with electricity prices from the US ERCOT market and a use-case example is provided for bidding an energy storage unit into the day-ahead and real-time energy markets of ERCOT using stochastic optimization methods. Results show that the synthetic price model out performs a (naive) persistence forecast model by resulting in 24% to 47% more in profits over 168 simulated days.

1 Introduction

1.1 Motivation

Advances in regulatory policies that expand market access to renewable energy (RE) generation and energy storage technologies are bringing a broader and more diverse set of market participants to wholesale electricity markets in the United States. Regulatory directives such as Federal Energy Regulatory Commission (FERC) Order 841 focus on creating avenues for the participation of energy storage resources in wholesale markets [1]. Similarly, FERC Order 2222 is opening up market participation to aggregations of distributed energy resources (DER), ensuring market access to a diverse set of new market actors [2]. The expanded market access that these regulatory directives afford is happening in conjunction with rapid cost declines for RE and storage technologies [3]. These factors—as well as other considerations such as decarbonization and energy resilience—are combining to drive more installations of RE and storage assets, bringing more resources online at a rapid pace [4].

The increase in number and diversity of energy resources in wholesale electricity markets is happening in tandem with an increase in the uncertainty in wholesale electricity prices. Increased price volatility is driven in large part by growing penetrations of variable renewable energy generation (primarily wind and solar) [5], but also by extreme weather events and overall net demand variability [6] [7].

Participants need effective bidding models to enable profitable participation in these increasingly diverse and volatile markets. While many approaches to strategic bidding exist—including optimization, equilibrium, and simulation approaches—there is a significant amount of on-going research in establishing the most effective methods. One attribute that is common to most of the bidding approaches is that they rely on the ability to form an accurate expectation of market prices [8]. Thus, to assess the performance of the various approaches being proposed, we need methods to model expected market prices and associated uncertainty. This will enable researchers to understand (and improve) the performance of different bidding approaches. Accurate characterization of price forecast uncertainty can help market participants evaluate the efficacy of potential bidding strategies under many different possible pricing regimes and futures.

A key challenge in modeling day-ahead and real-time electricity prices is that the underlying probability distribution often has heavy or fat tails [9]. Fat-tailed distributions are highly skewed and have an undefined variance, which makes them inconvenient to model. Additionally, occurrences of extreme price events are typically characterized as a stochastic process. Stochastic processes that arise from fat-tailed distributions are not mathematically well-behaved; using a probability function with a defined variance will under-represent the probability of extreme observations [10]. In electricity markets, extreme observations can be disruptive and/or result in high profits or loss, which strongly impacts the profitability of a given bidding strategy.

This paper presents an approach for generating statistically representative time series that exhibit high-periodicity and heavy tails, and capture the uncertainty observed in electricity markets. To achieve an accurate characterization of electricity prices, we divide the price generation process down into two components: a deterministic com- ponent and a stochastic component, and then

synthesize the two components into hourly price profiles. We present the methodology to stochastically generate these synthetic prices and a discussion of their goodness of fit and moments as compared to realized prices. To evaluate the performance of the methods, we generate synthetic, day-ahead price data from the Electricity Reliability Council of Texas's (ERCOT) Houston Hub location for the years 2018 and 2019 [11]. These data are used to demonstrate our method, but the techniques are applicable to any hourly (or 15-minute) electricity pricing data from deregulated wholesale markets across the United States, Europe, and throughout the world.

In addition to presenting the results on the quality of the synthetic price time series, we demonstrate a use-case of the synthetic price generation method for testing market bidding approaches in a stochastic optimization model. For each bidding period we generate 1,000 price scenarios and associated probabilities using our synthetic pricing methodology. We then apply k-means clustering on those 1,000 scenarios to generate 10 scenarios that are used as input to the stochastic optimization model and compare the profit performance against a persistence forecast model.

Our synthetic price generation methodology builds on considerable research in two distinct areas: stochastic differential equations (SDE) for forecasting electricity prices (specifically mean-reversion and jump-diffusion modeling) and auto-regressive integrated moving average (ARIMA) modeling for price forecasting. Previous work in applying SDEs to electricity prices has generally focused on daily average prices [12]. To extend these methods to hourly or 15-minute pricing, we replace the mean-reversion component of the SDE with an ARIMA process that is better able (than standard linear regression approaches) to characterize the daily and weekly patterns. We also extended how the jump process is modeled to account for conditional probabilities of price spikes occurring in consecutive hours. To capture the conditional probabilities, we re- placed the traditional Poisson process for modeling jumps with a generalized point process model, including a conditional intensity function. Finally, we replace traditional methods of estimating spike intensity (e.g. sampling into estimated distribution parameters) with a Markov chain approach where the transition matrices are derived from observed price spike transitions.

1.2 Literature Review

In his comprehensive review of electricity price forecasting [12], Weron identifies five main categories of price forecasting: multi-agent (including game theoretic and pro- duction cost models), fundamental/structural models, reduced- form quantitative/stochastic models, statistical or econo- metric models, and artificial intelligence-based models. Additionally, Weron notes that many of the approaches fall into a 'hybrid' category where two or more categories are combined. We focus our literature review on the relevant category for the research presented herein: a hybrid of reduced-form stochastic and statistical models. Along that vein, we focused on work for characterizing the stochastic nature, magnitude, and occurrence of price spikes.

The primary form that the reduced-form stochastic models takes is the jump-diffusion model. Jump-diffusion models are special cases of a general stochastic differential equation, which draws heavily from the finance literature for risk management in price dynamics and derivatives [12]. Mean-reversion, spot price spikes, and non-normality characteristics observed in stock markets are even more pronounced in electricity markets [13]. Electricity markets are unique in

that the commodity being traded cannot be easily stored for long periods of time. This constraint creates a system that is highly reactive to small changes in supply and demand. As with other commodities, electricity is traded on a spot market, which is based on day-ahead prices [14]. Trading based on day-ahead prices allows the system operator to verify that supply will meet demand in the next market period.

Weron et al. uses two separate approaches to model the "jumpy" character of spot prices in electricity markets [14]. One approach adds a jump term to a mean-reverting diffusion-type SDE, with the assumption that a negative jump always follows a positive one. This approach reflects the authors' observation that spikes usually only last one day. The second approach, which can handle spikes that last multiple time steps, uses a Markov chain with two states: a base state and a spike state. Each state is modeled by an independent price process, one mean-reverting and one log-normal. Both spike models are fit to data from which an annual sinusoidal term and a week-long moving average term have been subtracted. A key contribution to the mean-reverting jump diffusion approach comes from Cartea and Figueroa [13]. In contrast to Weron et al. [14], only weekly periodicity is removed from the price data (by subtracting the mean of data corresponding to the same day of the week). Deterministic annual seasonality is modeled with a time-varying mean reversion equilibrium. Cartea and Figueroa demonstrate that stripping spikes from electricity price data improves the result of a normality test dramatically, which makes a jump-diffusion approach applicable [13].

Geman and Roncoroni [15] propose a similar model to Cartea and Figueroa [13] with two main differences. First, instead of assuming that every positive jump is followed by a downward jump, they introduce a jump sign that is dependent on the current price (+1 if below a threshold, -1 if above). Second, they introduce a separate threshold above which the probability of a spike increases. Benth et al. compares the mean-reverting jump-diffusion model in Cartea, and Figueroa [13] to the threshold variation in [15], and a third approach previously proposed by Benth et al. [16]. The latter is a multi-factor model made up of a super- position of Ornstein-Uhlenbeck processes. Hayfavi and Talasli [17] propose an alternative multifactor model combining deterministic seasonality terms, Brownian motion, and three jump processes. One is an Ornstein-Uhlenbeck type process that reverts in the next time step after a jump, another is an Ornstein-Uhlenbeck process with slower mean reversion that takes a few time steps, and the third is a pure jump process that captures long term jump effects.

Weron proposes the idea of non-homogenous Poisson processes in [18], but is not able to fully apply this method due to lack of data to fit the model. The non-homogeneous approach considers arrival rates indexed on time of year and does not consider how spike probability also depends on prices in preceding hours, which is an important dependency in day-ahead pricing.

Borovkova and Schmeck [19] present another jump-diffusion variation that employs stochastic time change, which transforms the price distribution to be closer to Gaussian using the "activity rate". They build from prior research demonstrating a positive correlation between demand and the occurrence of price spikes and use temperature (as a proxy for demand) as the driving factor in the activity rate for the stochastic time change. The resulting model captures similar key characteristics as the jump-diffusion models described above (e.g., seasonality, mean reversion, and spikes) and uses the stochastic time change to incorporate exogenous variables such as temperature as a proxy for electricity demand. We incorporate the temperature variable as an

exogenous variable in our ARIMA component, while utilizing generalized, non-homogeneous point processes to characterize seasonal dependency in jump arrival rates.

Most of the literature described above has focused on forecasting of daily average spot prices rather than hourly day-ahead or sub-hourly real-time pricing. A recent review article by Lago et al. summarizing state-of-the-art algorithms in day-ahead price forecasting [20] identified that most of the recent research has primarily been in statis- tical models and deep learning models. Lago et al. note that the leading statistical models are generally autoregressive yet utilize least absolute shrinkage and selection operator (LASSO) or elastic nets for model estimation and feature selection tools (e.g., the Lasso Estimated AutoRegressive (LEAR) model) and employ variance stabilizing tools. The leading deep learning models generally fall into the deep neural net (DNN), recurrent neural net (RNN), and convolution networks (CNN) with multiple levels and long short-term memory (LSTM) models [20]. While many of these models have shown promising improvements in point forecast accuracy, they are not focused on generating stochastic price scenarios and capturing the total uncertainty in the market. Thus, our focus on combining some of the recent learnings from the statistical model space and combining that in a hybrid model with the reduced-form stochastic approach was selected for addressing our research questions.

Finally, while we have focused our literature review here on the price forecasting, we note that our primary focus has not been to create the perfect forecast of hourly pricing. Rather we are interested in creating a methodology for capturing the uncertainty inherent in the market, being able to generate thousands of potential realizations of what a next day pricing could look like, and applying the synthetic pricing to evaluating bidding models and associated risk. Repositories and synthetic generation approaches for time series data exist, but there is a dearth of adequate data for testing bidding approaches where the underlying time series exhibit high periodicity and heavy tails. Even with the increased collection and curation of large time series datasets, any available datasets that have these qualities are of limited use for testing, calibrating, and vetting time series models [21]. It is imperative that time series modeling and forecasting approaches use diverse training/testing data to be generalizable to future data realizations, especially as market conditions shift [21] [22].

1.3 Our Contributions

We developed a synthetic price generation method- ology that brings together jump diffusion modeling and ARIMA modeling to generate hourly electricity prices that capture both the seasonal (including weekly/hourly and annual seasonality) mean-reverting trend and the stochastic nature of price spikes (including inter-hour spike dependencies). To our knowledge this is the first application of jump-diffusion modeling to electricity markets, specifically enabled by our integration of ARIMA models and generalized point processes for jump arrivals. Specific contributions include:

- The integration of ARIMA models as the deterministic component of an SDE
- Replacement of the standard Poisson process in jump modeling by a generalized point process allowing for non-homogeneous and history-dependent spike modeling
- Integration of Markov chain transition matrices to model spike intensities and enable improved estimation of price transitions during multi-hour spike events

2 Method

Our methodology is laid out as follows: first, we pro- vide an overview of the SDE used in our price generation approach (contrasting it to previous work in the field); second, we discuss each of the three terms in the SDE and outline the derivation of relevant parameters.

2.1 Summary of the Stochastic Differential Equation

Our approach builds from a general SDE that captures the mean-reversion and jump-diffusion characteristics of electricity spot pricing (St). The baseline approach for the mean-reverting, jump-diffusion model (following notation from [13]) is as follows. The log-price process is defined as:

$$\ln S_t = g(t) + Y_t \tag{1}$$

where g(t) is a deterministic function and Yt is a stochastic process. The stochastic process, Yt, is defined as:

$$dY_t = -\alpha Y_t dt + \sigma(t) dZ_t + \ln J dq_t \tag{2}$$

The first term, −αYtdt, is the mean-reversion term. The speed of mean-regression α is typically derived by linear regression and often combined with g(t) to capture the deterministic component of the log-price process. The second term, σ(t)dZt, is the diffusion term, which combines the volatility observed in the de-spiked prices (σ(t)) and the increment of a standard Brownian motion (dZt). The last term, ln Jdqt, is the jump process that combines a random variable (J), which characterizes the jump size, with a Poisson process (dqt), which determines the arrival rate of the jump. The arrival rate is based on some intensity λ, which is often empirically derived from the data.

In our approach, we modeled actual prices (rather than log-prices) for two reasons. First, the log-transform can affect the skewness of the data that the jump component is aiming to capture, and level prices have been empirically shown to have better performance in capturing spikes [20] [23]. Second, hourly electricity pricing in many markets can take on negative values, which are inherently undefined in log-pricing.

We replace the deterministic component (the combination of g(t) and the mean reversion term) with an ARIMA process, Xt, that is able to capture the inter-day seasonality of hourly day-ahead and real-time pricing. The modeling of price-levels rather than log-prices and the use of an ARIMA model to capture the deterministic component of the SDE allows us to rewrite equation (1) and combine with an integrated equation (2) to derive our implementation of the SDE:

$$S_t = X_t + \int_t^T \sigma(t) dZ_t + \int_t^T J dq_t \tag{3}$$

In the following sections of the methodology, we describe our approach to modeling the individual terms in Equation 3 and deriving the associated parameters.

2.2 Mean Reversion Term - ARIMA Modeling

Data Preparation: To capture the mean-reversion com- ponent of the realized price time series, we fit an ARIMA model to the data. Prior to fitting the ARIMA model, we use the interquartile method to remove prices that are considered extreme (i.e., spikes). Equations (4) and (5) describe the process for spike identification.

$$IQR \triangleq Q_3 - Q_1 \tag{4}$$

$$K_t \notin [(Q_1 - c * IQR), (Q_3 + c * IQR)] \tag{5}$$

IQR is the inter-quartile range, Q1 is the 25th empirical quartile, Q3 is the 75th empirical quartile, c is a multiplier on the IQR, and Kt is a spike at time t.[1]

It is important to note that prior to fitting the ARIMA model we tested the de-spiked time series for a unit root using the Augmented Dickey-Fuller (ADF) test [10]. The ERCOT day-ahead pricing data used in this study do not have a unit root at the 0.01-level, which provides further rationale for using price levels instead of log-pricing.

In addition to the price levels, we included hourly temperature data as a covariate in the ARIMA models. The ADF test was used to verify that these hourly temperature data also do not have a unit root. Based on this, we performed first-order differencing on the hourly temperatures and used the hourly temperature differences in the ARIMA models. Figure 1 shows the data preparation process described here for a 9-month period in 2019.

ARIMA Fitting: To better capture short-term seasonality (diurnal and weekly), we used a rolling window approach to fit an ARIMA model for each consecutive one- week (168 hours) subset of the time series (3). We used the Python package pmdarima to fit a unique ARIMA model to each 168-hour window [24]. These models were used to generate the price forecasts discussed below.

$$\text{ARIMA}(\phi, \theta) : X_t = \phi y_{t-1} + e_t - \theta e_{t-1} + \beta(T_t - \phi T_{t-1}) \tag{6}$$

In equation (6), the auto-regressive portion is the product of φ and the time-lagged price; the moving average component is a white noise coefficient at time t (et) less the product of θ and the time-lagged white noise, and the covariate component is the product of β and the hourly temperature difference at time t (Tt) less the product of φ and the time-lagged temperature difference.

ARIMA Forecasts: Using each trained ARIMA model, the next 24-hour period is forecasted. The forecasted results are integrated using the last one-hour value from the training set, then scaled using a bootstrapping approach with a resample size of 5000.

[1] In the results and use-case example we use a c-value of 3 to capture a broader non-spike band than the commonly used 1.5 multiplier.

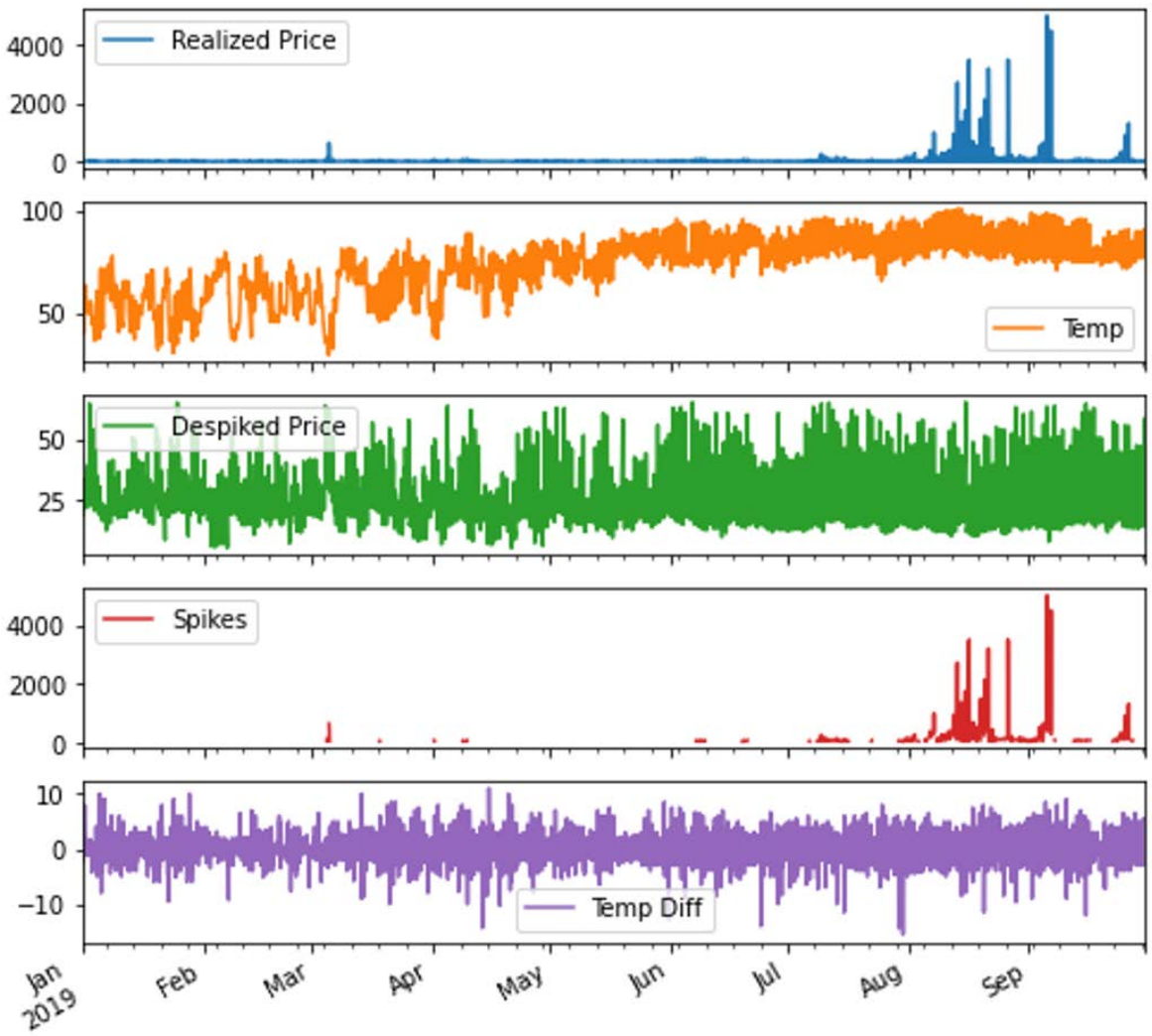


Figure 1: time series plots of (i) raw electricity prices used in this study, (ii) raw hourly temperatures, (iii) de-spiked electricity prices, (iv) electricity price spikes, and (v) first-order difference of the hourly temperatures.

2.3 Diffusion Term

The diffusion term in our model aligns closely with traditional diffusion modeling where volatility ($\sigma(t)$) is the standard deviation of the residuals from the ARIMA process and the stochastic portion (dZt) is standard Brownian motion, using the normal distribution with a zero mean. We allowed for the volatility to be defined on an hourly basis or as a single value. Either way, the volatility is defined on a rolling basis with the number of look-back days, or samples, acting as a hyperparameter in the model.

2.4 Jump Term

The jump component of the model extends and modifies the standard approaches used in jump diffusion modeling in two significant ways: (1) we utilize a generalized point process that allows for history dependent arrival rates (as opposed to a standard Poisson process), and (2) we use a

Markov chain with transition matrices to characterize jump intensities instead of modeling them as independent and identically distributed random variables.

First, we replace the Poisson process commonly used in jump diffusion models with a more generalized point process that includes a conditional intensity function, allowing us to characterize arrival rates that are non-homogeneous and history dependent. This extension was based on work in neural spike trains where the firing of different neurons in the brain are dependent both on the timing of the events and on preceding spike events [25] [26]. This switch was driven primarily by our transition to hourly price generation instead of the daily average price estimation seen in the literature. In hourly pricing, price spikes are more likely to occur during certain times of the day (e.g. more likely in the mid-afternoon than the middle of the night). It is also common to observe spikes that are longer than 1-hour in duration. We implemented the generalized point process to model both conditions.

To model time-dependent, multi-hour spike events, we utilize two sets of expected probabilities: (1) a set of initial arrival probabilities indexed on hour of the day (λo(h)) and (2) a conditional intensity function (λc(p)) that is de- pendent on the number of prior consecutive hours in which spikes occurred. These two sets of probabilities combine into the general point process (dqt), which uses the initial intensity and conditional intensity functions to determine whether or not a spike occurs at time t:

$$dq_t = \begin{cases} 1 & \text{with prob. } \lambda_0(h) \text{ if no spike in prior hour or prob. } \lambda_c(p) \text{ if } p \text{ spikes have occurred in succession.} \\ 0 & \text{with prob. } 1-\lambda_0(h) \text{ if no spike in prior hour or prob. } 1-\lambda_c(p) \text{ if } p \text{ spikes have occurred in succession.} \end{cases} \quad (7)$$

In equation (7) λ0(h) is the probability of a spike occurring in hour h of the day if no spike occurred in hour h − 1; and λc(p) is the probability of a spike occurring given that a spike has occurred in p consecutive hours prior. Both sets of probabilities are determined empirically based on spikes observed over the chosen look-back window (e.g., 30 days prior).

An example of calculated probabilities for two randomly selected days in 2019 (August 30th and November 2nd) is shown in Figure 2. Note that the summer day (left column) shows much higher initial arrival rate probabilities, with the noon hour having a probability over 40%. The summer day also has a higher likelihood of additional spikes after the initial spike, and those probabilities stay high longer than the on late fall day shown in the right column.

The preceding discussion on jumps has centered around the arrival of the jumps (dqt), but the spike intensity (J) should be considered as well. Traditionally, J has been defined as a normal, log-normal, or exponential random variable, but this approach presents challenges due to the heavy tails of the underlying distribution [12]. Using a random variable to characterize the spike intensity especially presents problems in the context of multi-hour spike events. Most multi-hour spike events follow a pattern (e.g. spike in the first hour is lower than the second, and then the intensity fades in the final hours), but using a random variable to calculate intensity leads to spikes in adjacent hours that are uncorrelated. To address these issues our method implements a Markov process to determine the spike intensity.

The Markov process for spike intensity is built from conditional transition probabilities from one price-bin to another price bin. Once the Poisson process indicates that an initial spike is

occurring, the Markov process begins for determining the spike intensity, or the market price. The zeroth bin is defined from the lowest observed price up to the highest de-spiked price. Subsequent bins can be determined in many ways. It is important to capture relevant price ranges to reflect realistic spike chains. Thus, determining the price bins should be done based on familiarity with the market data. Transition probabilities are determined empirically from the data. The length of any chain of price spikes is determined by the Bernoulli sampling that determines dqt.

Figure 3 shows a Hinton diagram of the transition prob- abilities, where the increasing size and darkness of the squares signify a higher likelihood of transitioning from the state on the vertical axis to the state on the horizontal axis. For example, upon the first spike in a chain one samples from the zeroth row in Figure 3. The most likely outcome is that the first spike intensity lands in the first price bin so let us assume that occurs. Say that the first price bin extends from $100/MWh to $300/MWh. Then, the spike intensity is determined by uniformly sampling within the bin, resulting in a spike intensity between $100/MWh to $300/MWh. Now, when a second consecutive spike occurs, one would sample from the distribution in the first row. Let us say that the second most likely out- come happens, which is a transition to the second bin or row of the transition matrix shown in Figure 3. Now, the second price intensity is determined by uniformly sampling from the second bin, which might extend from $300/MWh to $600/MWh for example. The process continues for as many time steps as indicated by the arrival process described above.

In summary, the combination of the non-homogeneous, history-dependent point process for spike arrival and the Markov chain transition matrix for spike intensity allows our approach to more fully capture the characteristics of hourly electricity pricing. First, our approach can model clusters of spikes of different durations and their varying likelihoods throughout the year. Next, the approach developed here can handle spike intensity relationships both within a multi-hour spike event and conditioned on time of the year.

2.5 Parameter Weighting

The primary hyperparameter to the model in both the diffusion term (Section 2.3) and the jump term (Section 2.4) is the number of look-back days, or number of samples used to derive the empirical values. The model considers this hyperparameter when building the parameters that define the diffusion and jump terms from the historical realized pricing. Sometimes it is better to use a larger window (e.g., 3-6 months of data) to obtain large sample sizes for estimating the arrival probabilities or transition matrix elements and one might not want to treat all of those days equally. This is especially true when estimating parameters for periods with higher spikes (in both arrival and intensity), which may persist for a few weeks to a month or two.

To address this issue, we add weighting hyper-parameters to the model that allows one to select among a few features for ordering the importance of days in the look-back window. For example, we suggest "days prior", "aver- age pricing", and "weekend/weekday" as features of the data along with linear, quadratic, or exponential weighting of sample days based on those features. Using weighting hyper-parameters allows the model to preferentially weight recent days, days with similar average (de-spiked) pricing, or weekends/weekdays (relative to the day being modeled). Additionally, it allows for three methods for calculating those weights, with linear smoothly applying the weights, quadratic extending the weights further from the 'similar' days (e.g. similar average price or very re- cent days), and exponential putting the highest weight

on 'similar' days. In the following results section, we show the impacts of switching between these weighting approaches for two different look-back lengths.

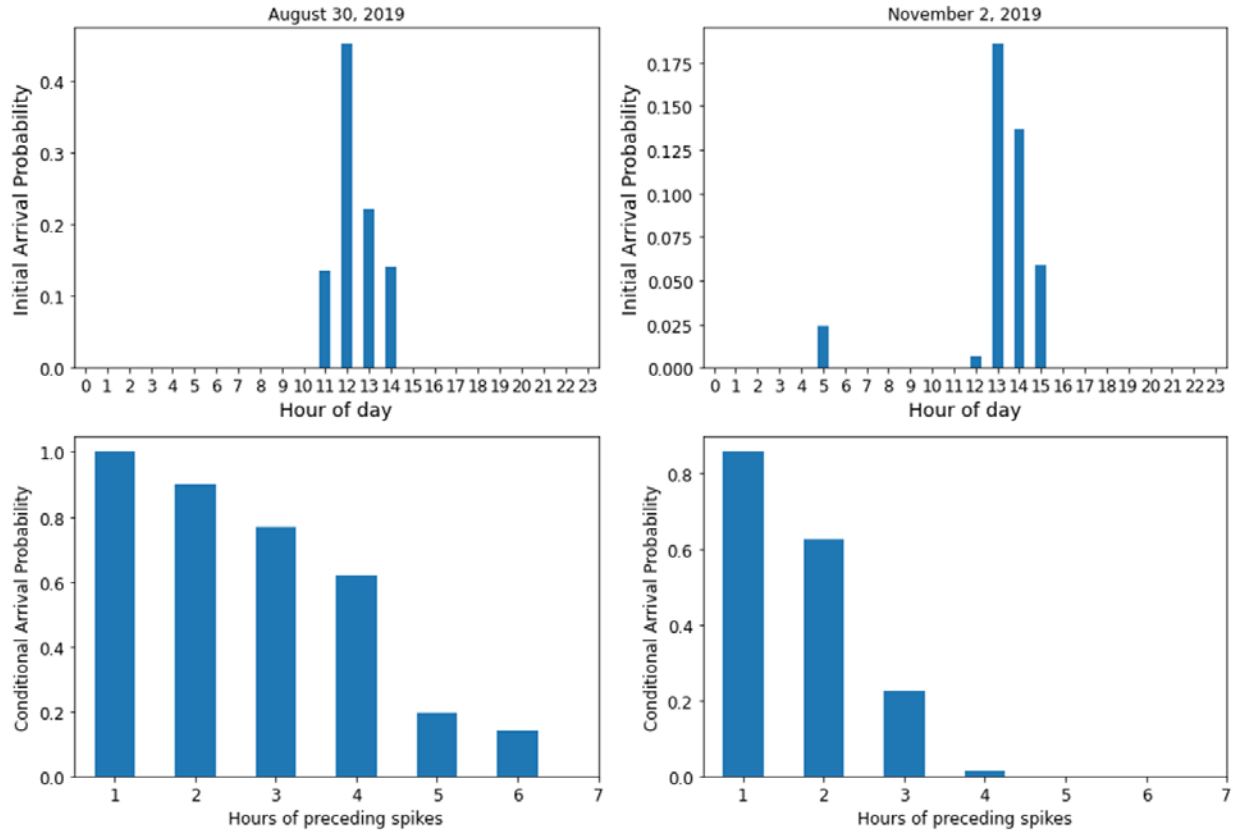


Figure 2: Initial arrival rate probabilities (top row) and conditional intensity functions (bottom row), shown for two days in 2019 (August 30th in left column and November 2nd in right column). Probabilities are derived empirically from the previous 30 days of price samples.

3 Results

We evaluated the efficacy of the approach outlined in this paper by evaluating the model-generated synthetic price data for the year of 2019, with look-back windows stretching back into 2018. We used two different look-back window lengths (30- and 180-day windows) and two different weighting approaches ('prior days' and 'average price' weighting). We used a linear function when applying each weighting setting. For example, in a 30-day look-back, the 15th day prior receives 0.5 weighting, while the day immediately preceding the modeled day receives a weighting of 1.0.

We generated one-thousand synthetic prices traces for each day of 2019 and each window/weighting combination. Each of these traces contains the same deterministic signal generated with the forecast from the ARIMA model, combined with an instance of the jump-diffusion components from the SDE. Figures 4, 5, and 6 display synthetic pricing traces for three distinct days using the 30-day look- back window and linear weighting on prior days. Figures 4 and 5 show two different days in the summer period, with the August day exhibiting significant price spikes (over $2,000/MWh for two consecutive hours) and the July day exhibiting more moderate price spikes ($150/MWh for three consecutive hours). The realized price from the settled day-ahead market is shown in blue. The realized price is not exposed to the model during generation of synthetic pricing.

A range of spike intensities and durations are captured on these days, and they generally cover the range of prices that are realized in the day-ahead market. While many spikes are included in the generated pricing, many of the simulated traces have no spike, or a much lower spike (demonstrated by the dark gray clustering under the realized price spike).

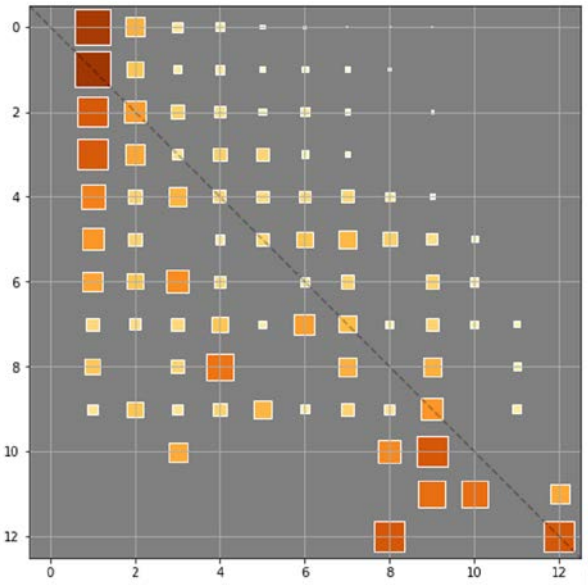


Figure 3: Hinton diagram of an example Markov chain transition matrix for spike intensity. The axes are integer price bins. Individual square sizes and darkness represent the probability of transition from one price bin to another. Rows of probabilities sum to one and are sampled uniformly. The zeroth row is sampled first for every set of price spikes to determine the bin for the first price spike.

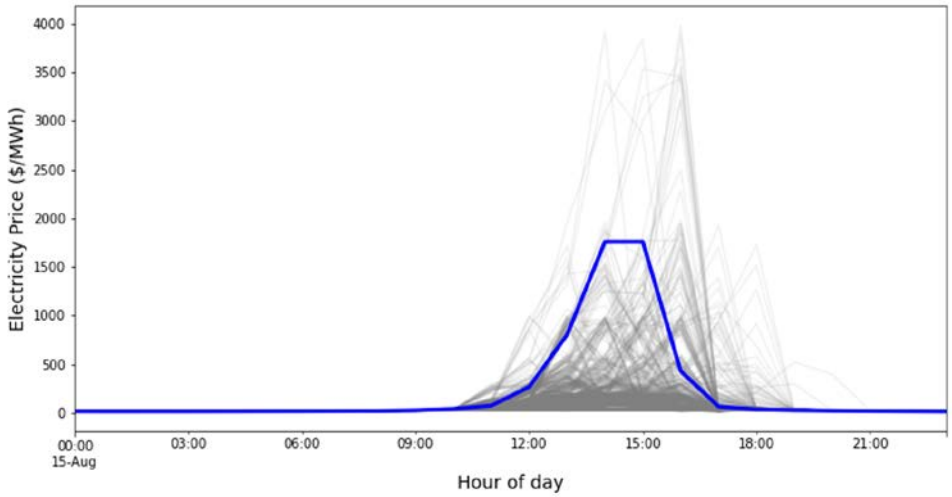


Figure 4: Synthetic pricing traces and realized pricing for August 15th of 2019.

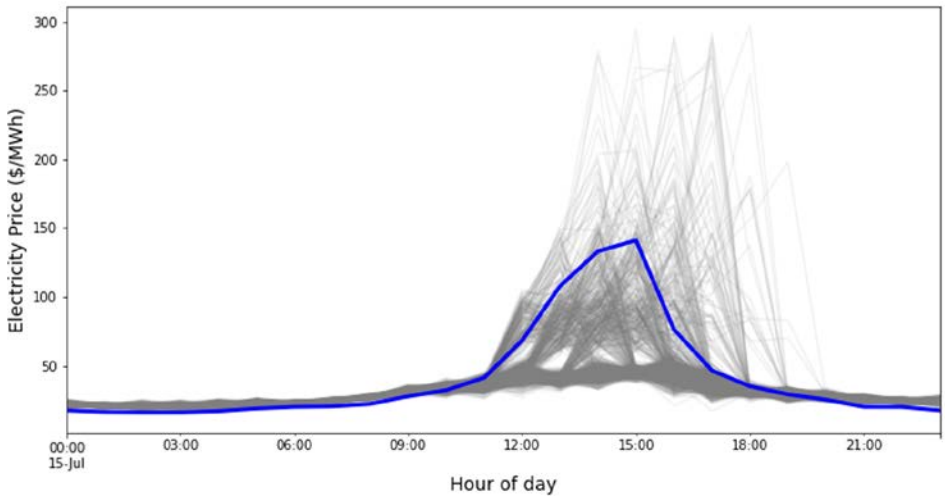


Figure 5: Synthetic pricing traces and realized pricing for July 15th of 2019.

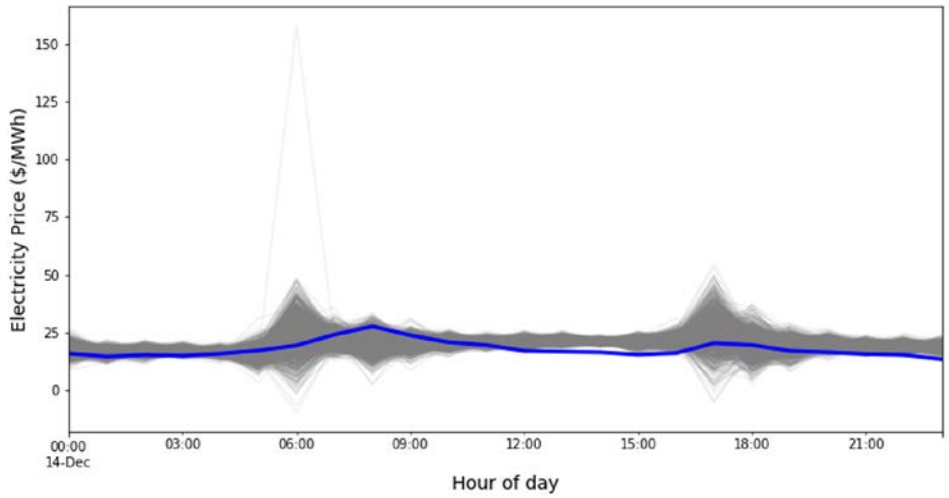


Figure 6: Synthetic pricing traces and realized pricing for December 14th of 2019.

Figure 6 shows a winter day in December that has no realized price spikes. The synthetic pricing for this day primarily shows the impacts of the ARIMA and diffusion portions of the model. The ARIMA portion sets the deterministic signal and the diffusion term simulates the expected variability around that mean signal. Since the diffusion term is a function of the hour of the day, we see more volatility in the generated prices during the morning and afternoon, which ends up aligning with the two small peaks in the realized pricing for the day. One trace out of the thousand simulated pricing sets includes a price spike. This result indicates that the initial arrival rates for this day had a non-zero value at 6am (the value is in fact 0.02652, suggesting that we would expect more predicted morning spikes if we generated another 1,000 traces).

To characterize the performance of the model, we also calculated the moments of the data for each of the window/weighting combinations and compared these results to the moments of the realized pricing. Comparing model- based moments to empirical moments is a common approach to evaluate the performance of the model [17]. Figure 7 and Figure 8 show the mean and standard deviation of the generated prices for each look-back length/weighting approach combination compared with the moments of the realized pricing. The figures show the moments by season of the year, as the mean price shape varies significantly throughout the year. The shorter, 30-day look-back window performs better overall than the 180-day look-back in terms of matching the realized price moments. In contrast, the weighting approach does not appear to have a significant impact on the accuracy of the moments. The model captures the mean signal of the empirical data well, but it sometimes fails to capture the full magnitude of price spikes observed in the data (for example in summer and fall months). This shortcoming could be a result of how the bins were defined in the Markov chain transition matrices. If the model assumes a low likelihood of entering high price bins or the sampling within bins occurred at their lower end, it could skew the mean in those times of the year.

4 Use-case Example

To demonstrate the value of the stochastic price generation method we present a use-case example for bidding a battery energy storage system (BESS) in an energy market. The energy market is modeled after the Electric Reliability Council of Texas (ERCOT) system, with day-ahead and real-time energy markets. Optimal price-quantity bids are determined using a stochastic optimization program modeled after the work in [27], which we extend by considering the Conditional Value at Risk (CVaR) as well as adding a two-stage program for the real-time bidding model to address the issue of infeasible BESS dispatch strategies that might result from the bid clearing process.

The goal of the BESS bidding model is to maximize the expected profit from executing energy arbitrage. Stochastic price scenarios are used to represent a discrete distribution of market prices in the stochastic programs to form a sample average approximation of the expected profit [28]. The probabilities of price scenarios are determined by first generating 1,000 price scenarios using the method de- scribed above with a rolling look-back window of 30 days. Second, the 1,000 price scenarios are clustered using the k- means algorithm to create 10 clusters and 10 price scenarios from the centroids of the clusters [29]. The probability of each of the 10 scenarios is determined as the ratio of the number of members in each cluster to the total number of scenarios.

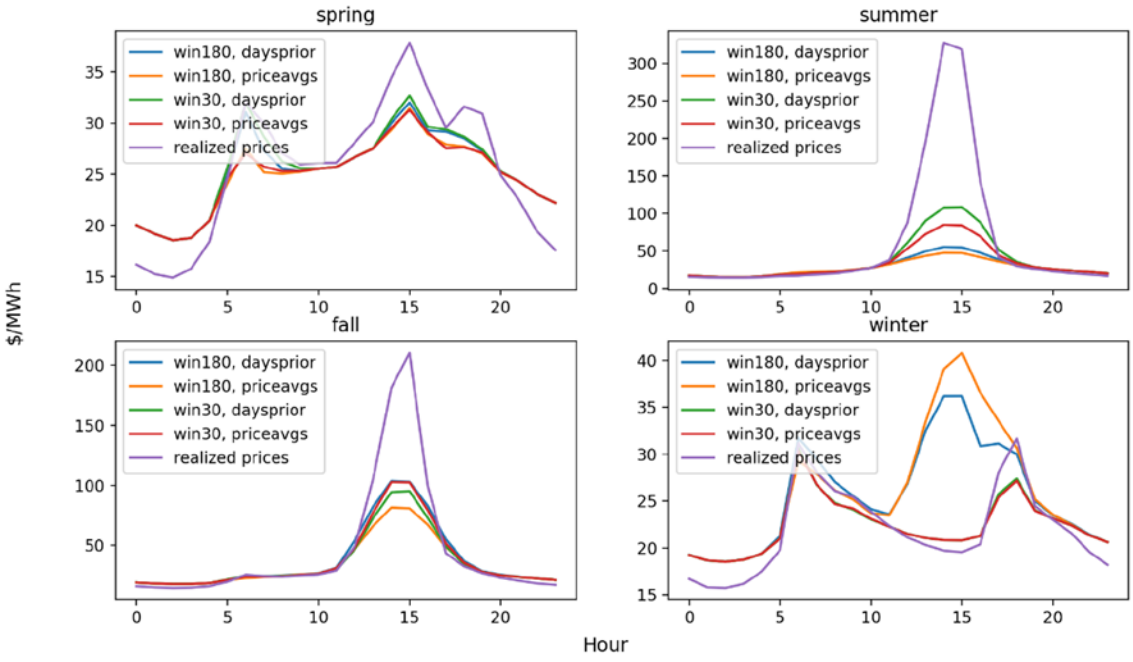


Figure 7: First moment (mean) of the synthetic pricing compared to realized prices. The synthetic pricing mean is shown for four cases: two look-back windows (30 and 180 day) and two weighting functions (prior days and average price).

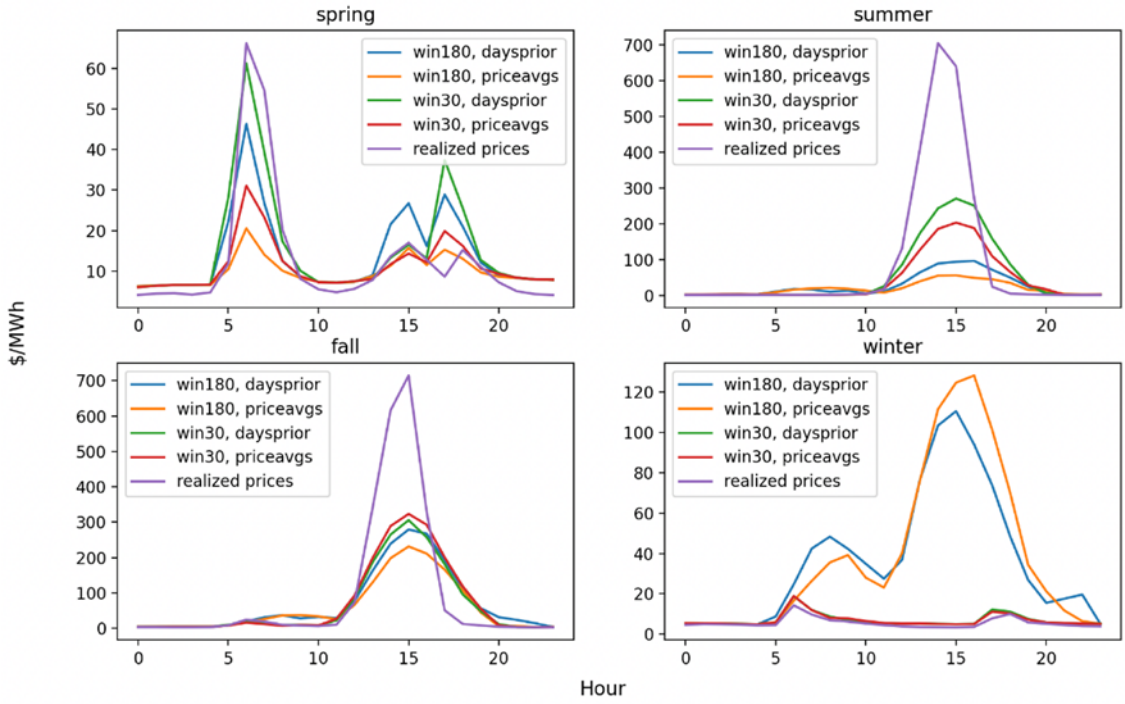


Figure 8: Second moment (standard deviation) of synthetic pricing compared to realized prices. The synthetic pricing mean is shown for four cases: two look-back windows (30 and 180 day) and two weighting functions (prior days and average price).

Once the ten price scenarios and their probabilities are determined, the first step in the stochastic bidding model is to form price-quantity bids in the day-ahead market. Equation A.1 in Appendix A shows the day ahead bidding model. The day-ahead market model uses price scenarios to determine the optimal price-quantity pairs to bid in each hour for tomorrow.

Next, the real-time market model A.2 takes the cleared day-ahead market energy quantities to maximize the weighted sum of the real-time market profit and the CVaR. The real- time market model is a two-stage stochastic program, with here-and-now decisions for the real-time price-quantity bids in the next clearing period and wait-and-see decisions for the price-quantity decisions in all subsequent time-steps. For this use-case we execute the real-time market model in a rolling fashion, starting with optimizing the price- quantity bid in the first hour of the operating day. As the simulation time advances the real-time market model horizon shrinks because we do not know the cleared day-ahead market quantities until the 14th hour in the operating day. In each operating day the day-ahead price-quantity bids for the next day are optimized in the 9th hour to reflect the 10 AM deadline in ERCOT. The day-ahead market is assumed to clear by the 14th hour, at which point the real- time model horizon is extended to 24 hours again. Figure 9 shows a graphical representation of the simulation horizons in the day-ahead and real-time models. Table 1 lists the BESS technical assumptions.

Table 1: Technical assumptions

Battery assumptions	Value
Storage Capacity	2 MWh
Inverter Capacity	1 MW
Total Round-Trip Efficiency	90.25%
Inverter Efficiency	95%
Rectifier Efficiency	95%
Initial State of Charge	0%

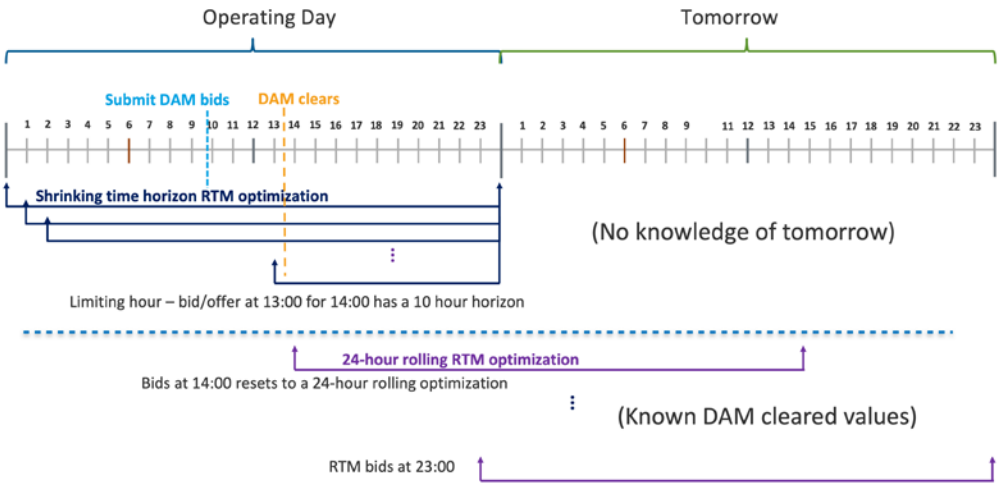


Figure 9: Simulation horizon for the real-time market model simulation, which is designed to reflect the ERCOT market. Day-ahead bids are due by 10 AM in the operating day and our cleared by the 14th hour. The real-time market model horizon shrinks and grows to reflect the knowledge of cleared day-ahead market quantities.

Table 2 shows the total profits from the simulation of 168 days in 2019 using ERCOT day-ahead and average hourly real-time market prices. For comparison we also present the results from using a persistence forecast for the expected energy market prices. We also show two different CVaR weightings: zero and 50%. The scenarios with a 50% CVaR weighting show the highest profit, with a 47% improvement over the persistence forecast profit. Considering no CVaR results in a 24% improvement over the persistence forecast profit. The benefit of considering the CVaR is likely due to the fact that being conservative in the day-ahead market allows for more opportunities in the real-time market. Indeed, as shown in Table 2, the stochastic model with a 50% CVaR weight resulted in lower day-ahead market profits than the stochastic model with no CVaR considered; yet when considering CVaR the stochastic model is able to get more net profits by performing better in the real-time market than the model with zero CVaR weight.

Figures 10 and 11 show a side-by-side comparison of the realized price, persistence forecast, and stochastic price scenarios for a certain day in the day-ahead and real-time markets respectively. It interesting to note that the poorer performance in terms of profits in the real-time market is likely due to the more volatile nature of the real-time market. From Figure 11 we can see that the stochastic scenarios also suffer from some persistence assumptions: most of the price scenarios show a price peak much later than actually occurs - similar to the persistence forecast.

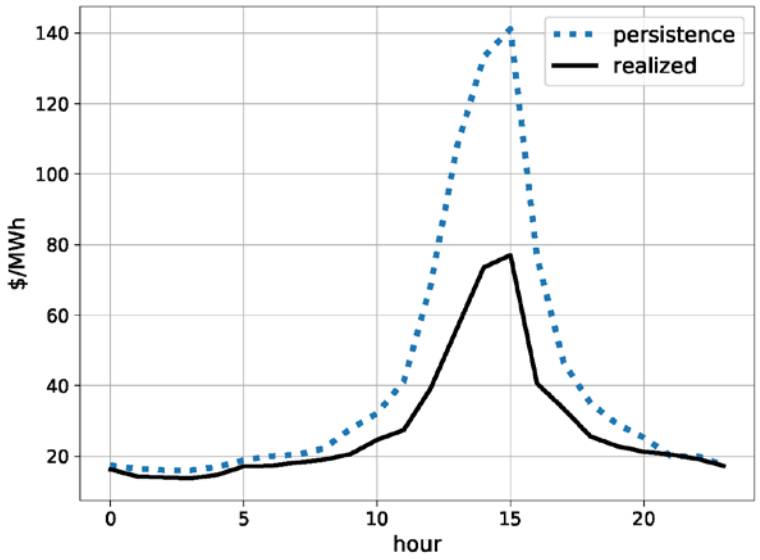
5 Conclusions

This work extends the state-of-the-art for modeling daily average electricity market prices with stochastic differential equations to sub-daily time intervals. Our goal is not provide better price forecasts, though the methods may be applied to price forecasting; rather, our method is intended to generate realistic price time series for evaluating different market bidding strategies. Like any model built upon statistics of past data, the method cannot predict new features of price time series (such as a consecutive spike series twice as long as any series observed in the past). However, our method can generate an unlimited amount of price series that our statistically and qualitatively similar to historical prices. With this capability re- searchers can compare different market bidding strategies over long (synthetic) time periods that may be unachievable due to data availability or recent changes in market rules that make all but the most recent prices irrelevant.

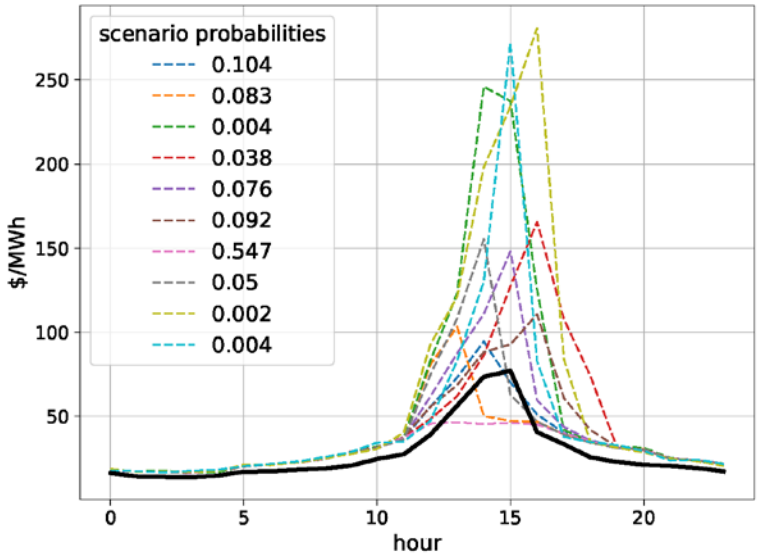
Despite the limitations of generating price scenarios from past prices, the model parameters can be modified by practitioners to represent market behavior that has not been witnessed but might be anticipated. For example, adjustments to the conditional intensity function, the initial arrival rates, the variance of the diffusion term or the state switching probabilities of the Markov chain all can be used to simulate hitherto unseen market conditions. Futhermore, quantifying the range of these market price attributes in different wholesale electricity markets is an area of future work and would enable more informed re- search into how market evolution will impact stochasticity and volatility. This research could be valuable in assessing the costs (and risk) associated with procuring 24/7 renewable power or in providing firm renewable generation from hybrid renewable plus energy storage plants. While the application of the synthetic pricing model to stochastic bidding model is limited in nature, it demonstrates that these data can be utilized to programmatically test bidding models and provide a mechanism to quantify the value in acknowledging and incorporating the increasing uncertainty that the day-ahead and real-time electricity markets are experiencing.

Table 2: Total profits from stochastic optimization models and persistence forecast simulations. The total profits are from simulating 168 days from 2019 with ERCOT day-ahead and average hourly real-time market prices.

Battery assumptions	CVaR weight	Total profit	Day-ahead	Real-time
Persistence forecast	n/a	$35,040	$48,147	-$13,107
Stochastic model	0	$43,318	$57,134	-$13,815
Stochastic model	0.5	$51,514	$53,729	-$2,215

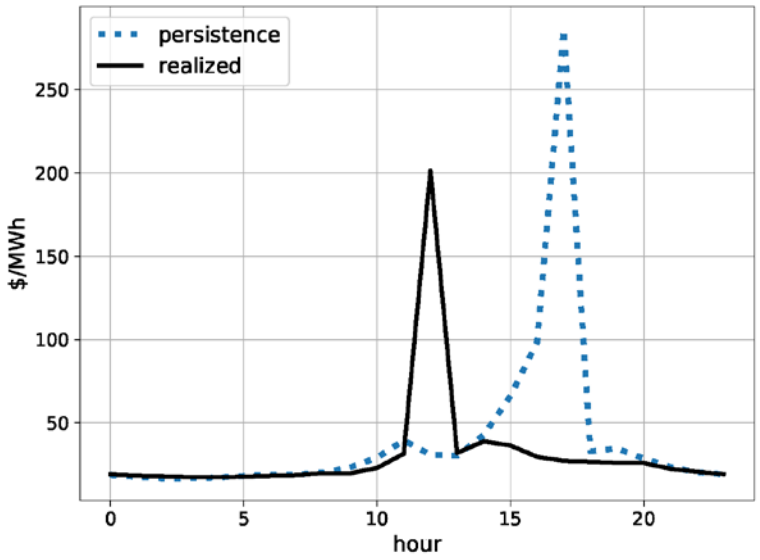


(a) Realized price and persistence forecast in day-ahead market on July 7th, 2019.

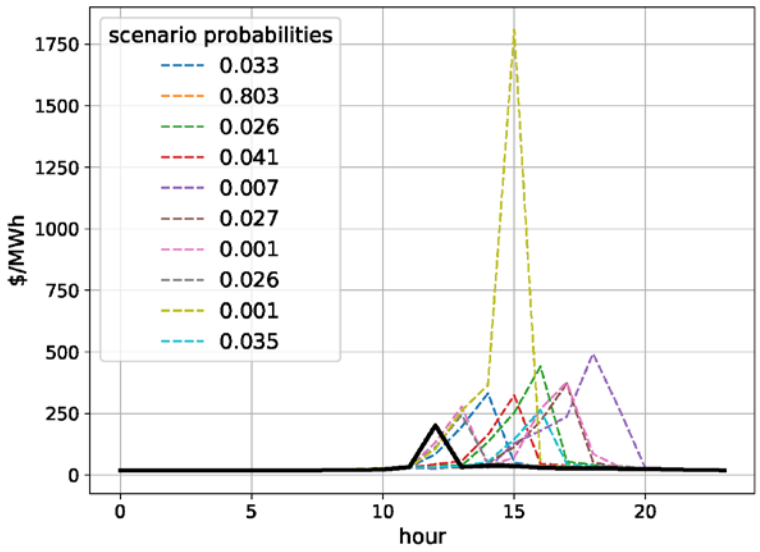


(b) Realized price and stochastic price scenarios in day-ahead market on July 7th, 2019. The legend shows the probability of each scenario, which is determined via the k-means clustering of 1,000 synthetic price scenarios. The thicker black line is the realized price.

Figure 10: Comparison of realized price, persistence forecast, and stochastic price scenarios in the day-ahead market on July 7th, 2019.



(a) Realized price and persistence forecast in real-time market on July 7th, 2019.



(b) Realized price and stochastic price scenarios in real-time market on July 7th, 2019. The legend shows the probability of each scenario, which is determined via the k-means clustering of 1,000 synthetic price scenarios. The thicker black line is the realized price.

Figure 11: Comparison of realized price, persistence forecast, and stochastic price scenarios in the real-time market on July 7th, 2019.

Appendix A. Energy Market Stochastic Optimization Model

Here we present the stochastic optimization models used to demonstrate a use-case for the stochastic price generation method. The models are based on the work in [27]. The day-ahead stochastic optimization model first determines the optimal price-quantity pairs to bid (to buy) or offer (to sell) in the day-ahead market. The real-time stochastic optimization model then takes the cleared day- ahead market energy obligations as inputs to determine the optimal price-quantity pair to bid or offer in the next real-time clearing period. For simplicity both models assume hourly time-steps.

A.1 . Day-ahead market price-quantity stochastic optimization model

The day-ahead model maximizes the weighted sum of the expected profit and the conditional value at risk. Table A.3 presents the sets, parameters, and variables of the day- ahead model, which is shown in A.1.

Constraints A.1b through A.1i are operational constraints for the BESS. The constraints A.1j through A.1s require that one price-quantity strategy be chosen over the horizon of T time-steps from all scenarios. Constraints A.1t and A.1u define the CVaR value [28]. And constraints A.1v through A.1x define non-negative variables.

$$\max_{x\in\mathcal{R},z\in\mathcal{Z}} \beta\sum_{s\in\mathcal{S}} p_s \sum_{t\in\mathcal{T}} \pi_{s,t}(x_{s,t,\text{Esell}} - x_{s,t,\text{Ebuy}}) - (1-\beta)x_{\text{CVaR}} \tag{A.1a}$$

$$\text{s.t. } x_{s,t,\text{charge}} \le B_{\text{MW}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1b}$$

$$x_{s,t,\text{discharge}} \le B_{\text{MW}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1c}$$

$$x_{s,t,\text{MWh}} \le B_{\text{MWh}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1d}$$

$$x_{s,t,\text{MWh}} \le B_{\text{MWh}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1e}$$

$$x_{s,t,\text{charge}} = x_{s,t,\text{Ebuy}}\eta_{\text{charge}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1f}$$

$$x_{s,t,\text{discharge}} = x_{s,t,\text{Esell}}/\eta_{\text{discharge}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1g}$$

$$x_{s,1,\text{MWh}} = f_{t=1}B_{\text{MWh}} - x_{s,1,\text{discharge}} + x_{s,1,\text{charge}} \qquad \forall s\in\mathcal{S} \tag{A.1h}$$

$$x_{s,t,\text{MWh}} = x_{s,t-1,\text{MWh}} - x_{s,t,\text{discharge}} + x_{s,t,\text{charge}} \qquad \forall s\in\mathcal{S}, \forall t\in\{2,\ldots,T\} \tag{A.1i}$$

$$x_{s,t,\text{Ebuy}} \le z_{s,t,\text{Ebuy}}B_{\text{MW}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1j}$$

$$x_{s,t,\text{Esell}} \le z_{s,t,\text{Esell}}B_{\text{MW}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1k}$$

$$x_{t,\text{Ebuy}} - x_{s,t,\text{Ebuy}} \le (1 - z_{s,t,\text{Ebuy}})B_{\text{MW}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1l}$$

$$x_{t,\text{Esell}} - x_{s,t,\text{Esell}} \le (1 - z_{s,t,\text{Esell}})B_{\text{MW}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1m}$$

$$x_{t,\text{Ebuy}} - x_{s,t,\text{Ebuy}} \ge 0 \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1n}$$

$$x_{t,\text{Esell}} - x_{s,t,\text{Esell}} \ge 0 \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1o}$$

$$x_{t,\$\text{buy}} \le \pi_{s,t} + Mz_{s,t,\text{Ebuy}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1p}$$

$$x_{t,\$\text{buy}} \ge \pi_{s,t} - M(1 - z_{s,t,\text{Ebuy}}) \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1q}$$

$$x_{t,\$\text{sell}} \ge \pi_{s,t} - Mz_{s,t,\text{Esell}} \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1r}$$

$$x_{t,\$\text{sell}} \le \pi_{s,t} + M(1 - z_{s,t,\text{Esell}}) \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1s}$$

$$x_{\text{CVaR}} = x_{\text{VaR}} + \frac{1}{\alpha}\sum_{s\in\mathcal{S}} p_s x_{s,M} \tag{A.1t}$$

$$x_{s,M} \ge -\sum_{t\in\mathcal{T}} \pi_{s,t}(x_{s,t,\text{Esell}} - x_{s,t,\text{Ebuy}}) - x_{\text{VaR}} \qquad \forall s\in\mathcal{S} \tag{A.1u}$$

$$x_{t,\$\text{sell}} \ge 0,\ x_{t,\$\text{buy}} \ge 0,\ x_{t,\text{Ebuy}} \ge 0,\ x_{t,\text{Esell}} \ge 0 \qquad \forall t\in\mathcal{T} \tag{A.1v}$$

$$x_{s,t,\text{charge}} \ge 0,\ x_{s,t,\text{discharge}} \ge 0,\ x_{s,t,\text{MWh}} \ge 0,\ x_{s,t,\text{Ebuy}} \ge 0,\ x_{s,t,\text{Esell}} \ge 0 \qquad \forall s\in\mathcal{S}, \forall t\in\mathcal{T} \tag{A.1w}$$

$$x_{s,M} \ge 0 \qquad \forall s\in\mathcal{S} \tag{A.1x}$$

Sets and Indices

$s \in \mathcal{S}$	integer scenarios for market prices and bidding strategies, $\mathcal{S} \triangleq \{1,\ldots,S\}$
$t \in \mathcal{T}$	integer time-steps, $\mathcal{T} \triangleq \{1,\ldots,T\}$

Parameters		**Units**
B_{MW}	battery inverter capacity	MW
B_{MWh}	battery energy capacity	MWh
$f_{t=1}$	fraction of battery energy capacity stored in first time-step	unitless
$\pi_{s,t}$	market price in scenario s, time-step t	\$/MWh
$p_{s,t}$	probability of market price scenario s,	unitless
β	weighting factor for profit vs. CVaR in objective function	unitless
α	quantile for CVaR (for example $\alpha = 0.05$ considers the 5% tail of the loss)	unitless

Decision Variables		**Units**
$x_{s,t,\text{charge}}$	power sent to the battery in scenario s, time-step t	MW
$x_{s,t,\text{discharge}}$	power discharged from the battery in scenario s, time-step t	MW
$x_{s,t,\text{MWh}}$	battery state of charge in scenario s, time-step t	MWh
$x_{s,t,\text{Ebuy}}$	energy purchased in scenario s, time-step t	MWh
$x_{s,t,\text{Esell}}$	energy sold in scenario s, time-step t	MWh
$z_{s,t,\text{Ebuy}}$	binary decision to purchase energy in scenario s, time-step t	MWh
$z_{s,t,\text{Esell}}$	binary decision to sell energy in scenario s, time-step t	MWh
$x_{t,\text{Ebuy}}$	energy purchased in time-step t, used for day-ahead market bids	MWh
$x_{t,\text{Esell}}$	energy sold in time-step t, used for day-ahead market offers	MWh
$x_{t,\$\text{buy}}$	market bid in time-step t	\$/MWh
$x_{t,\$\text{sell}}$	market offer in time-step t	\$/MWh
x_{CVaR}	conditional value at risk	\$
x_{VaR}	value at risk	\$
$x_{s,M}$	maximum of loss minus the value at risk and zero, where loss is the profit multiplied with negative one; used to define the conditional value at risk	\$

Table A.3: Sets, indices, parameters, and decision variables for Equation A.1.

A.2 Real-time market price-quantity stochastic optimization model

The two-stage real-time market price-quantity stochastic optimization model maximizes the weighted sum of the expected profit and the conditional value at risk. The real- time model takes the cleared day-ahead energy obligations as inputs to determine the real-time market profit using the method from the ERCOT market [30]. The primary difference from the day-ahead model is that only the price- quantity bid in the first time-step is constrained across the price scenarios. In other words, the price-quantity bids for all time steps beyond the first are "wait-and-see" decisions that vary over the price scenarios, whereas the "here-and- now" decision for the first time step is the same in all price scenarios.

Table A.4 presents the sets, parameters, and variables of the real-time model, which is shown in A.2.The here- and-now decisions are reflected in constraints A.2c through A.2h. Constraints A.2i through A.2m define whether or not the price bids and offers clear the market (in a similar fashion to [27]). Constraint A.2n is used to define the CVaR value. And constraints A.2o and A.2p define non- negative variables.

$$\max_{\boldsymbol{x}\in\mathcal{R},\boldsymbol{z}\in\mathcal{Z}} \beta \sum_{s\in\mathcal{S}} p_s \sum_{t\in\mathcal{T}} \pi_{s,t}\left[(x_{s,t,\text{Esell}} - E_{t,\text{DAsell}}) - (x_{s,t,\text{Ebuy}} - E_{t,\text{DAbuy}})\right] - (1-\beta)x_{\text{CVaR}} \tag{A.2a}$$

$$\text{s.t. } A.1b \ldots A.1k, A.1t \tag{A.2b}$$

$$x_{s,1,\text{charge}} = x_{\text{charge}} \qquad \forall s \in \mathcal{S} \tag{A.2c}$$

$$x_{s,1,\text{discharge}} = x_{\text{discharge}} \qquad \forall s \in \mathcal{S} \tag{A.2d}$$

$$x_{s,1,\$\text{buy}} = x_{\$\text{buy}} \qquad \forall s \in \mathcal{S} \tag{A.2e}$$

$$x_{s,1,\$\text{sell}} = x_{\$\text{sell}} \qquad \forall s \in \mathcal{S} \tag{A.2f}$$

$$x_{s,1,\text{Ebuy}} = x_{\text{Ebuy}} \qquad \forall s \in \mathcal{S} \tag{A.2g}$$

$$x_{s,1,\text{Esell}} = x_{\text{Esell}} \qquad \forall s \in \mathcal{S} \tag{A.2h}$$

$$x_{s,t,\$\text{buy}} \leq \pi_{s,t} + M z_{s,t,\text{Ebuy}} \qquad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \tag{A.2i}$$

$$x_{s,t,\$\text{buy}} \geq \pi_{s,t} - M(1 - z_{s,t,\text{Ebuy}}) \qquad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \tag{A.2j}$$

$$x_{s,t,\$\text{sell}} \geq \pi_{s,t} - M z_{s,t,\text{Esell}} \qquad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \tag{A.2k}$$

$$x_{s,t,\$\text{sell}} \leq \pi_{s,t} + M(1 - z_{s,t,\text{Esell}}) \qquad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \tag{A.2l}$$

$$x_{s,t,\$\text{sell}} \geq 0, \; x_{s,t,\$\text{buy}} \geq 0 \qquad \forall s \in \mathcal{S}, \forall t \in \mathcal{T} \tag{A.2m}$$

$$x_{s,M} \geq -\sum \pi_{s,t}\left[(x_{s,t,\text{Esell}} - E_{t,\text{DAsell}}) - (x_{s,t,\text{Ebuy}} - E_{t,\text{DAbuy}})\right] - x_{\text{VaR}} \qquad \forall s \in \mathcal{S} \tag{A.2n}$$

Sets and Indices

$s \in \mathcal{S}$	integer scenarios for market prices and bidding strategies, $\mathcal{S} \triangleq \{1, \ldots, S\}$
$t \in \mathcal{T}$	integer time-steps, $\mathcal{T} \triangleq \{1, \ldots, T\}$

Parameters		**Units**
B_{MW}	battery inverter capacity	MW
B_{MWh}	battery energy capacity	MWh
$f_{t=1}$	fraction of battery energy capacity stored in first time-step	unitless
$\pi_{s,t}$	market price in scenario s, time-step t	\$/MWh
$p_{s,t}$	probability of market price scenario s,	unitless
β	weighting factor for profit vs. CVaR in objective function	unitless
α	quantile for CVaR (for example $\alpha = 0.05$ considers the 5% tail of the loss)	unitless
$E_{t,\text{DAsell}}$	cleared energy sold obligations in day-ahead market	MWh
$E_{t,\text{DAbuy}}$	cleared energy purchased obligations in day-ahead market	MWh

Decision Variables		**Units**
$x_{s,t,\text{charge}}$	power sent to the battery in scenario s, time-step t	MW
$x_{s,t,\text{discharge}}$	power discharged from the battery in scenario s, time-step t	MW
$x_{s,t,\text{MWh}}$	battery state of charge in scenario s, time-step t	MWh
$x_{s,t,\text{Ebuy}}$	energy purchased in scenario s, time-step t	MWh
$x_{s,t,\text{Esell}}$	energy sold in scenario s, time-step t	MWh
$z_{s,t,\text{Ebuy}}$	binary decision to purchase energy in scenario s, time-step t	MWh
$z_{s,t,\text{Esell}}$	binary decision to sell energy in scenario s, time-step t	MWh
x_{charge}	power sent to the battery in first time-step	MW
$x_{\text{discharge}}$	power discharged from the battery in first time-step	MW
x_{Ebuy}	energy purchased in first time-step, used for real-time market bid	MWh
x_{Esell}	energy sold in first time-step, used for real-time market offer	MWh
$x_{\$\text{buy}}$	market bid in first time-step	\$/MWh
$x_{\$\text{sell}}$	market offer in first time-step	\$/MWh
x_{CVaR}	conditional value at risk	\$
x_{VaR}	value at risk	\$
$x_{s,M}$	maximum of loss minus the value at risk and zero, where loss is the profit multiplied with negative one; used to define the conditional value at risk	\$

Table A.4: Sets, indices, parameters, and decision variables for Equation A.2.

Bibliography

References

[1] Federal Energy Regulatory Commission, Electric storage participation in markets operated by regional transmission organizations and independent system operators, Federal Register 61 (2018). URL: https://ferc.gov/sites/default/files/ 2020-06/Order-841.pdf.

[2] Federal Energy Regulatory Commission, Participation of distributed energy resource aggregations in markets operated by regional transmission organizations and independent system operators, Federal Register 85 (2020). URL: https://www. federalregister.gov/d/2020-20973.

[3] P. Donohoo-Vallett, Revolution... now the future arrives for five clean energy technologies-2016 update, Technical Report, DOE, EERE, 2016.

[4] N. E. F. Bloomberg, New energy outlook 2019, Bloomberg New Energy Finance: New York, NY, USA (2019).

[5] M. Milligan, B. A. Frew, A. Bloom, E. Ela, A. Botterud, A. Townsend, T. Levin, Wholesale electricity market design with increasing levels of renewable generation: Revenue sufficiency and long-term reliability, The Electricity Journal 29 (2016) 26-38.

[6] M. Auffhammer, P. Baylis, C. H. Hausman, Climate change is projected to have severe impacts on the frequency and intensity of peak electricity demand across the United States, Proceedings of the National Academy of Sciences 114 (2017) 1886-1891.

[7] J. Seel, A. D. Mills, R. H. Wiser, Impacts of high variable renewable energy futures on wholesale electricity prices, and on electric-sector decision making (2018).

[8] M. Prabavathi, R. Gnanadass, Energy bidding strategies for re- structured electricity market, International Journal of Electrical Power & Energy Systems 64 (2015) 956-966.

[9] R. Weron, Heavy-tails and regime-switching in electricity prices, Mathematical Methods of Operations Research 69 (2009) 457- 473.

[10] D. A. Dickey, W. A. Fuller, Distribution of the estimators for autoregressive time series with a unit root, Journal of the American statistical association 74 (1979) 427-431.

[11] ERCOT, ????, Historical dam load zone and hub prices, URL: http://www.ercot.com/mktinfo/prices.

[12] R. Weron, Electricity price forecasting: A review of the state- of-the-art with a look into the future, International journal of forecasting 30 (2014) 1030-1081.

[13] A. Cartea, M. G. Figueroa, Pricing in electricity markets: a mean reverting jump diffusion model with seasonality, Applied Mathematical Finance 12 (2005) 313-335.

[14] R. Weron, M. Bierbrauer, S. Truck, Modeling electricity prices: jump diffusion and regime switching, Physica A: Statistical Mechanics and its Applications 336 (2004) 39-48.

[15] H. Geman, A. Roncoroni, Understanding the fine structure of electricity prices, The Journal of Business 79 (2006) 1225-1261.

[16] F. E. Benth, R. Kiesel, A. Nazarova, A critical empirical study of three electricity spot price models, Energy Economics 34 (2012) 1589-1616.

[17] A. Hayfavi, I. Talasli, Stochastic multifactor modeling of spot electricity prices, Journal of Computational and Applied Mathematics 259 (2014) 434-442.

[18] R. Weron, Market price of risk implied by asian-style electricity options and futures, Energy Economics 30 (2008) 1098-1115.

[19] S. Borovkova, M. D. Schmeck, Electricity price modeling with stochastic time change, Energy Economics 63 (2017) 51-65.

[20] P. Villaplana, Pricing power derivatives: A two-factor jump- diffusion approach, Available at SSRN 493943 (2003).

[21] Y. Kang, R. J. Hyndman, F. Li, Gratis: Generating time series with diverse and controllable characteristics, Statistical Analy- sis and Data Mining: The ASA Data Science Journal 13 (2020) 354-376.

[22] K. Smith-Miles, S. Bowly, Generating new test instances by evolving in instance space, Computers & Operations Research 63 (2015) 102-113.

[23] J. J. Lucia, E. S. Schwartz, Electricity prices and power derivatives: Evidence from the nordic power exchange, Review of derivatives research 5 (2002) 5-50.

[24] T. G. Smith, pmdarima: Arima estimators for python, ???? URL: http://www.alkaline-ml.com/pmdarima.

[25] D. H. Johnson, Point process models of single-neuron dis- charges, Journal of computational neuroscience 3 (1996) 275- 299.

[26] D. J. Daley, D. Vere-Jones, An introduction to the theory of point processes: volume I: elementary theory and methods, Springer, 2003.

[27] D. Krishnamurthy, C. Uckun, Z. Zhou, P. R. Thimmapuram, A. Botterud, Energy storage arbitrage under day-ahead and real-time price uncertainty, IEEE Transactions on Power Sys-tems 33 (2017) 84-93.

[28] A. Shapiro, D. Dentcheva, A. Ruszczynski, Lectures on stochastic programming: modeling and theory, SIAM, 2014.

[29] J. A. Hartigan, M. A. Wong, Algorithm as 136: A k-means clustering algorithm, Journal of the royal statistical society. series c (applied statistics) 28 (1979) 100-108.

[30] ERCOT Real Time Market, 2021. URL: https://www.ercot. com/mktinfo/rtm.