Performance Risk Scoring of Risk-free Renewable Generation Bids

Aparna Gupta*
Lally School of Management
Rensselaer Polytechnic Institute
Troy, NY

Denis Osipov†
System Planning and Analysis
New York Power Authority
Albany, NY

January 4, 2024

Abstract

Renewable generators can bid a conservative segment of the forecasted power generated into the day-ahead market with the aim of generating greater risk-adjusted profitability. However, these bids and their underlying risks may vary across renewable generators, which must be calibrated for a reliable clearing of the day-ahead power market. A rigorous third-party performance risk scoring of these risk-free bids is a valuable input and validation in support of a robust day-ahead market clearing framework. In this paper, we develop an ensemble machine learning approach for performance risk scoring of risk-free renewable generation bids and evaluate the performance of these bids made by different renewable generators and in different geographies. The methodology predicts contractual shortfall with high accuracy, where we utilize it to study the significant features of risk-free bidding of renewable generators.

Keywords: renewable energy; energy markets; risk scoring; energy pricing; risk management; neural networks.

JEL Codes: Q42; Q41; C45; G32

1 Introduction and Motivation

Renewable generation participating as equal peers in power markets is both desirable and essential for meeting future energy demand in all geographies. This is desirable for enterprises investing in renewable generation with a goal of being profitable without relying on government subsidies. It is also desirable for the benefits of lower cost of renewable generation to translate to lower cost of

*Corresponding author. Email address: guptaa@rpi.edu
The information, data, or work presented herein was funded in part by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0001276. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. We acknowledge support from Dr. Joe Chow (RPI), Dr. Koushik Kar (RPI), Dr. Daniel Kirk-Davidoff (EPRI) and Ph.D students, Sai Palepu & Anweshita Das (Lally School, RPI), as well as valuable advisory direction from our ADVISORS, NYISO and ERCOT.
†Email address: denis.osipov@nypa.gov

energy for consumers. Therefore renewable generation investment and power market participation of renewable generators is globally essential for the future sustainability goals.

Large scale investment in renewable generation, such as, in solar or wind farms is inherently risky. The fundamental risk arising from the variation in weather conditions prevents an accurate prediction of generation throughput from these power generation resources, especially when compared to the traditional coal, natural gas or nuclear-based power generation. Accurate prediction is necessary for an advanced planning of demand and supply matching on power grids by the grid operators, often termed as the unit commitment problem. The uptake of generation assets participation for market clearing is done at a price point for each time period, termed as the economic dispatch.

Various risk management solutions have emerged for the management of inherent risk of intermittence in renewable power production using wind or solar assets. Energy storage technologies based on different principles pose as very promising mechanism for reducing the intermittence risk of renewable generation. While storage technologies can help mitigate the intermittence risk, cost constraints are likely to prevent entirely eliminating the intermittence. On the other hand, risk-responsive contracting mechanisms have been developed to enable renewable generators to offer power at matching reliability level as that of traditional generators (Gupta and Palepu, 2023). For a reliable and trusted incorporation of these risk-responsive contracts in the market clearing in a unit commitment/economic dispatch framework, a third-party validation and affirmation of these contracts is required.

In this paper, we develop a framework for performance risk scoring of risk-free renewable generation bids in the power markets. When renewable generation assets make risk-free bids comparable to traditional power generators, such as combined cycle natural gas generators, these bids arising from inherently intermittent generation must be assessed for their validity and reliability. Similar to risky debt markets, a third-party assessment and rating of the risk-free bids is a valuable service offered for a trusted incorporation of these bids in market clearing. We apply our framework for the performance risk scoring of risk-free bids made by wind and solar generation in different geographies for the evaluation of the accuracy of the performance scoring.

Deregulation of power markets in the past decades merited borrowing risk management principles from the financial domain (Cartea and Villaplana, 2008; Thompson, 2013; Wozabal and Rameseder, 2020; Lucheroni and Mari, 2021). Specifically, the extensive use of derivative instruments to hedge electricity price risk and strategies for power market participants (Vehviläinen and Keppo, 2003; Deng and Oren, 2006; Doege et al., 2009; Falbo et al., 2010; Coulon et al., 2013). Increasing trends of renewable energy penetration that have greater dependence on weather elements has supported the utilization of cross-hedging and weather derivatives based risk management strategies (Müller and Grandi, 2000; Bessembinder and Lemmon, 2002; Brockett et al., 2005; Pérez-González and Yun, 2013; Hain et al., 2018; Bhattacharya et al., 2020). In addition to various financial risk management principles being applied to renewable energy assets, prior studies have developed optimal bidding strategies for renewable resources integrated in microgrids by modeling uncertainties in renewable energy production (Ferruzzi et al., 2016; Wang et al., 2017; Das and Basu, 2020; Nikpour et al., 2021). In the larger scale power grids, existing renewable risk management solutions and bidding strategies assume renewable generators to be price takers, which limits their competitiveness, revenue generation capability, as well as subjects them to high degrees of curtailment (Prokhorov and Dreisbach, 2022; Prol et al., 2020; Bird et al., 2016). These traits are not supportive of sustainable growth and investment in renewable energy. To address these issues, recent literature has developed risk-responsive pricing strategies using securitization principles to support stochastic renewable generators to bid in the day-ahead power markets (Gupta and Palepu, 2023), specifically risk-free offering comparable to the energy bids made by traditional generation resources.

In the credit markets, securitization has been used for decades for risk pooling and carving out securities to match investors' risk-reward appetite (Gupta, 2014; Morkoetter et al., 2017). Securitization has also been proposed for renewable energy for the possible benefit from securitizing cashflow of renewable assets for risk mitigation, access to a large capital pool, improvement in financing, reduction in transaction costs, and other growth opportunities (Liu et al., 2007; Krupa and Harvey, 2017; Alafita and Pearce, 2014; Gabig et al., 2015; Jiang and Chen, 2005; Lowder and Mendelsohn, 2013; Hyde and Komor, 2014). Utilization of asset securitization principles for stochastic renewable generation resources to actively participate and integrate in the power markets requires innovations in third-party evaluation of these bids, similar to credit rating of securitized instruments.

Rating agencies and their service in terms of rating debt instruments, including securitized products, is invaluable (Fischer, 2015). This was most significantly highlighted in the role of rating agencies in the 2008 global financial crisis (Hossain and Kryzanowski, 2019; Wojtowicz, 2014; Stolper, 2009). Securitization may be done based on a bundle of individual debt instruments of retail loans or mortgages or commercial loans or bonds types, resulting in the structured product to be called mortgage-backed securities or collateralized loan obligations, respectively, or various other variants. Therefore, successful securitization must rely on reliable credit risk assessment of the constituent instruments (Jacobson and Roszbach, 2003). Historically, retail credit risk models have used a range of data analytics techniques from logistic regression to discriminant analysis (Steenackers and Goovaerts, 1989; Boyes et al., 1989; Rosenberg and Gleit, 1994). Similarly, the seminal work of Altman et al. (1977) paved the path for commercial credit risk assessment, followed by other approaches, such as, structural firm value models (Merton, 1974) and reduced-form models (Jarrow and Turnbull, 2000).

Beyond the initial work, much progress has been made in advancing credit risk assessment, both for retail and commercial credit. As data varieties and volumes have expanded, artificial intelligence and machine learning (AI/ML) approaches have also been investigated towards improving credit risk assessment. In the retail context, Khandani et al. (2010) utilized machine learning techniques for nonlinear non-parametric forecasting of consumer credit risk using customer transactions and credit bureau data to demonstrate significant improvements in classification rates of delinquencies. Similarly, convolution neural networks using consumer transaction data and deep learning techniques have been applied for predicting consumer default with very promising results, while also accounting for interpretability of the model (Kvamme et al., 2018; Albanesi and Vamossy, 2019). For the commercial context, Fraisse and Laporte (2022) utilize a suite of artificial intelligence approaches to compare against traditional models to evaluate corporate defaults toward improved assessment of bank capital requirements, and find that neural networks provide the strongest case for using AI techniques. The challenge of non-linearity in different quantiles of bank loan loss given default prediction is significantly improved by combining linear quantile regression with a neural network structure, which provides the benefit of not having to specify the exact nature of non-linearity beforehand (Kellner et al., 2022).

In this paper, we develop a scoring methodology for a third-party assessment of the performance of renewable generators' obligations in the day-ahead market. A renewable generator bids a certain conservative portion of its forecasted throughput in the day-ahead market, with the aim of matching the risk profile of bids from conventional generators, such as combined cycle natural gas generators. The highly dynamic nature and non-stationary conditions for the performance scoring, arising from the need for hourly scoring in the day-ahead setting in presence of changing weather and power grid conditions, requires devising an appropriate technique for the performance scoring. We develop an ensemble machine learning approach by combining neural networks with a random forest algorithm to predict whether a renewable generator will fall short of its obligations in the day-ahead market.

We apply the performance scoring methodology to a range of wind and solar generation farms located in the state of New York and Texas to evaluate its efficacy. The non-stationary conditions for the scoring requires assessing the methodology in different hours and seasons for the generation assets. Studying wind and solar generation allows us to compare the performance of these two renewable resources, as well as evaluate the geographical differences in renewable generation within and between the two states. We find the methodology developed in this article to provide high accuracy in all cases considered, with wind generation trailing in predictive accuracy relatively to the solar assets. New York wind and Texas solar based risk-free contracts are most reliably predicted contracts for their shortfall characteristics. Additionally, the ensemble learning method developed in this paper reveals the significant factors that allow predicting a reliable delivery of renewable energy risk-free contracts.

Our paper contributes to address the challenge of seamless renewable integration in power markets. Fundamental risk of intermittence of renewable generation is a formidable hindrance in achieving the aggressive investment targets for renewable generation. Performance scoring based third-party validation of renewable generators' risk-responsive bids into the day-ahead market can provide the necessary assurance to the power grid system operators in their market clearing decisions. The performance scores can be explicitly incorporated in the unit commitment and economic dispatch formulations for the market clearing function of the system operators.

We begin the next section with formally stating the performance scoring problem in terms of predictor and response variables. The logic and the description of the performance scoring algorithm are provided in Section 2.3. Details of the wind and solar assets, as well as the data used for each asset and geography, are provided in Section 3. In Section 4, accuracy characteristics of the scoring methodology are presented and compared for asset type, temporal and geographical variations. We provide our concluding remarks in Section 5.

2 Performance Risk Scoring Methodology

The time horizon for a day-ahead market clearing in power markets is 24 hours. The risk-free bids made by inherently risky renewable generators bear some risk of non-delivery, even though they are the most protected tranche in the hierarchy of risk-responsive bids made by renewable generators, as shown in Figure 1. Therefore, similar to the highest quality investment grade corporate bonds or senior secured tranche of a collateralized debt obligation, a third-party assessment and rating of the renewable assets' risk-free bids is essential and a valuable service for the reliable functioning of the power markets and the grid. This assessment can also serve as a crucial input to systems operators for their unit commitment and economic dispatch decisions.

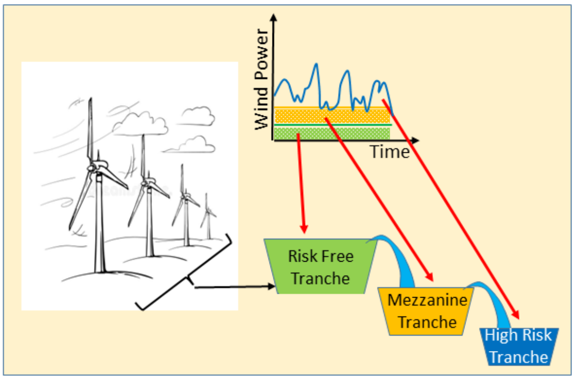

Figure 1: Schematic for the design of a risk-free contract offer of a renewable generator, following the tranche concept in securitization.

The specific context of performance scoring of renewable generators' risk-free bids bears specific challenges and need for consideration. The first of which is the time horizon and time duration for this performance scoring. The risk-free bids are placed at a specific time on the previous day for hourly delivery for each of the 24 hours of the next calendar day. Each of these 24 risk-free bids must be scored simultaneously, while accounting for the renewable asset-level and overall system-level characteristics and information available at the time. In this section, we define these relevant variables and develop a performance scoring methodology to assist in the risk assessment and utilization of the renewable risk-free bids in the power markets.

2.1 Variable Definition

A renewable asset's risk-free bid consists of a curve, referred to as a bid curve, that provides the price point per megawatt the generator is willing to accept for each megawatt of generation level for the hour. The market clearing day-ahead price, D_t, for each hour of the 24 hour calendar day is realized at time t in the timeline shown in Figure 2. This price level, D_t, determines the actual amount of risk-free power, Q_t, contracted by the renewable asset. The actual generation of the renewable asset, Y_{t+1}, determines the shortfall the generator faces in its ability to deliver the contracted power. We define the performance scoring response variable as the rate of shortfall, $S_t = \frac{\max(0, Q_t - Y_{t+1})}{Y_{t+1}}$. Prior to market clearing at time, t, in Figure 2, and after all the renewable assets have placed their risk-free bids at time, $t - \delta$, the rating agency must generate a shortfall forecast to provide a performance score of the renewable asset's risk-free bid.

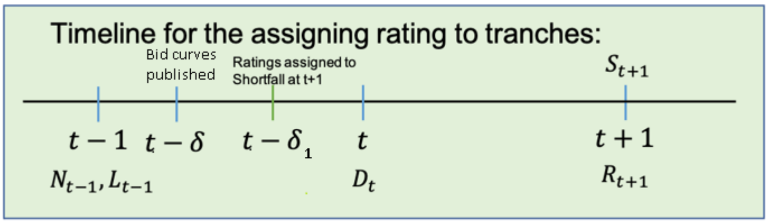

Figure 2: Timeline for the bid and performance scoring for the day-ahead power market.

The predictor variables available to a rating agency can be categorized as asset-level or system-level, where the former constitute information available regarding the asset's current and past attributes, while the system-level describe features of the grid as relevant to the asset in question. The key asset-level predictor variables available to the rating agency at time $t - \delta_1$ are the history of past shortfalls, $S_{t-u}; u = 1, \ldots, N$, generation forecast, F_{t+1}, the past generation forecast error, E_t, past generation, $Y_{t-u}; u = 1, \ldots, N$, and current and past risk-free bid curves of the asset.

Each asset accesses the grid at a node and the locational marginal price (LMP) for the node is the price point of relevance. Therefore, the system-level relevant variables for the performance scoring of the asset's risk-free bids are day-ahead electricity market price time series, $D_{t-u}; u = 1, \ldots, N$, real-time electricity market price time series, $R_{t-u}; u = 1, \ldots, N$, electricity demand (load) time series, $L_{t-u}; u = 1, \ldots, N$ at the asset's access node for the grid. Additionally, since in many grids natural gas plays and will continue to play a significant role in power generation, natural gas price time series, $N_{t-u}; u = 1, \ldots, N$, is also included.

As such day-ahead market price, real-time market price, natural gas price and electricity demand time series may not be as consequential to determine the shortfall, S_t. However, at time $t - \delta_1$ when performance scoring is done, the day-ahead market price, D_t is not known and must be estimated to obtain the power contracted under the risk-free bid of the asset, which is essential to determine the shortfall. These system level variables play a crucial role for predicting the day-ahead market price.

2.2 Distributional Characteristics of Shortfall

Before we design the performance scoring algorithm, it is important to characterize the distributional properties of the shortfall in the risk-free bid. A risk-free bid is a very conservative bid designed to match the reliability of a conventional asset's bid. Following Gupta and Palepu (2023), the conventional bids are taken to be dictated by a combined cycle natural gas generator with an average reliability of 96%, since these generators provide power generation with high reliability, some degree of flexibility and relatively lower carbon footprint. Moreover, combined cycle natural gas generators serve as a meaningful benchmark since they meet a high level of electricity demand now and are expected to do so in the future decades.

A renewable generator's risk-free bid made with a reliability of 96% has very low probability of non-delivery. Non-delivery will coincide with the realized generation level falling below the 4^{th} percentile of generation distribution for the hour. Therefore, the shortfall distribution is highly unbalanced, with very high fraction of realizations set at $S_t = 0$. When there is a non-zero shortfall, the realization of the shortfall level is highly affected by the seasonal and daily variations in almost all other variables of the domain. As a result, a single model for performance scoring cannot deliver the adequate accuracy for all windows of time by hours of the day and months of the year. The inherent non-stationarity requires a dynamic design of the performance scoring catering to different time windows. We develop the performance scoring algorithm next.

2.3 Algorithmic Logic and Design for Performance Scoring

The performance scoring response variable was defined earlier as the rate of shortfall, $S_t = \frac{\max(0, Q_t - Y_{t+1})}{Y_{t+1}}$. We further refined it as a categorical variable for the purpose of algorithm design. Noting the imbalanced nature of the response variable, we define the categories as: $\{0, (0, S_1], (S_1, S_2], \ldots, (S_N, 1.0]\}$. Each of these categories can be mapped to a symbol as done for bond debt ratings. For instance, when a risk-free contract is scored to meet its obligation, it earns a 'AAA' rating, and any drop from the target level is given a rating of 'AAA-', 'AA+', 'AA', 'AA-', etc. depending on the predicted degree of shortfall rate.

As described earlier, the rate of shortfall is non-stationary, therefore a single model cannot be used for its prediction at all time points. The diurnal pattern and seasonality in many variables results in the varying distributional characteristics of S_t. Therefore, the 24-hour window for which performance risk scoring needs to be done for the risk-free bids must be divided into time periods, $\{[0, t_1), \ldots, [t_M, 24]\}$, where in each time period S_t can be treated as independent identically distributed. We propose to build a model for each time period using the same algorithm described next to create a dynamic performance scoring methodology for renewable assets risk-free bidding.

The general principle for the algorithmic design is done in two stages as follows. In the first stage, the algorithm predicts if there will be any shortfall at all, while in the second stage, in the scenarios where shortfall occurrence is flagged, the algorithm seeks to find the size of the shortfall. In support of this two-stage design, the complete data set must be split into two training data sub-sets and a test set to validate the prediction. The fractions of data splits, $\{\alpha_1, \alpha_2, 1 - \alpha_1 - \alpha 2$, must be judiciously chosen to create the data sub-sets: 'Train 1' for Stage 1, 'Train 2' for Stage 2, and 'Test' for the validation. In each stage, an ensemble approach is chosen to assure high accuracy in prediction. The ensemble is designed with one supervised learning approach applied to predict the response variable, followed by the predicted response variable from the first method provided as an additional input to a second supervised learning approach. Random forests and neural networks are used as the two supervised learning approaches. Additionally, the interface between the stages, i.e. Stage 1 and Stage 2 and Stage 2 and Validation, is done with a conservative handling of the discrepancies in the confusion matrices of the two learning methods of the ensemble learning.

We enumerate the steps of the algorithm next.

- Define the three data sub-sets: Train 1, Train 2, Test using a split fractions of the complete data set: $\alpha_1, \alpha_2, 1 - \alpha_1 - \alpha_2$.
- **Stage 1**
  - **Step 1:** Using data in Train 1, build a random forest model to predict occurrence of shortfall.
  - **Step 2:** Add predicted outcome of shortfall from the random forest model as an additional input variable.
  - **Step 3:** Using data in Train 1 appended with additional variable, build a neural network model to predict the occurrence of shortfall.
- **Stage 2**
  - **Step 1:** Apply Stage 1 random forest followed by neural network model on Train 2 data set.
  - **Step 2:** Compute the Stage 1 neural network confusion matrix for Train 2 data.
  - **Step 3:** For any discrepancies seen in the prediction by the two ensemble methods of Stage 1 for the off-diagonal cells observations of the confusion matrix, label them as 'non-zero shortfall.'
  - **Step 4:** Using 'non-zero shortfall' labeled data in Train 2, build a Stage 2 random forest model to predict the level of shortfall.
  - **Step 5:** Add the predicted outcome of level of shortfall from the Stage 2 random forest model as an additional input variable.
  - **Step 6:** Using 'non-zero shortfall' Train 2 data appended with additional variable, build a Stage 2 neural network model to predict the level of shortfall.
- **Validation**

- **Step 1:** Apply the Stage 1 random forest, followed by Stage 1 neural network model on Test data set to identify occurrence of shortfall.
- **Step 2:** Label all instances with discrepancy between the two Stage 1 learning methods of the ensemble approach as 'non-zero shortfall'
- **Step 3:** Apply the Stage 2 random forest, followed by Stage 2 neural network model on 'non-zero shortfall' cases of Test data set.
- **Step 4:** Combine the Stage 1 neural network and Stage 2 neural network confusion matrices into a single confusion matrix for the entire algorithm to compute performance accuracy of the algorithm.

Based on the combined confusion matrix, various *Accuracy*, *Precision*, *Recall*, *F1 Score* metrics are computed, with a focus on 'no shortfall' versus 'non-zero shortfall' outcomes. We describe the renewable generation assets, their data and the corresponding system-level data to demonstrate the application of the performance risk scoring of risk-free bids.

3 Renewable Assets Data Description

Our study focuses on two states of the United States, namely, New York and Texas. Between these two states, which are reasonably big in their own right, there is ample diversity in the geographical characteristics to be able to broadly test the performance of the risk scoring methodology. The location of the solar and wind generation assets used in this study are shown in Figure 3. We provide a description of the data used for wind and solar generation assets and asset level variables located in New York and Texas. The power generation and forecast time series data are obtained for New York from the New York Independent System Operator (NYISO) and for Texas from the Electric Reliability Council of Texas (ERCOT). The generation capacity of the assets is provided in Table 1.

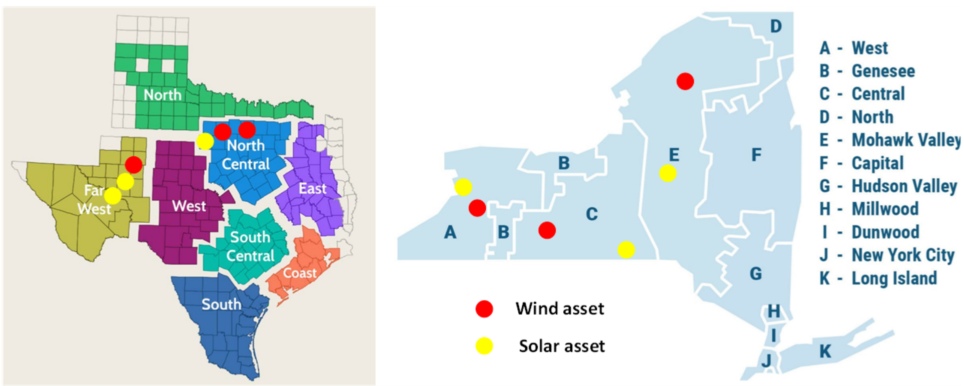

Figure 3: Location of renewable generation assets in the two states studied: New York (right) and Texas (left).

Geography	Zone	Asset Type	Generation Capacity (MW)
New York	A	Wind	135.5
New York	C	Wind	172.85
New York	E	Wind	395.75
Texas	North Central	Wind 1	120
Texas	North Central	Wind 2	110
Texas	Far West	Wind	65.8
New York	A	Solar	68.35
New York	C	Solar	83.2
New York	E	Solar	53.51
Texas	North Central	Solar	112
Texas	Far West	Solar 1	110.2
Texas	Far West	Solar 2	121.1

Table 1: Generation capacity of all the wind and solar assets located in New York and Texas included in the study.

3.1 Wind Resources Data

Texas wind generation and forecast data at hourly frequency are available for single wind generation farms for the time period Jan 2015- Dec 2018. As shown in Figure 3, the Texas wind assets belong to two different zones: North Central Zone and Far West Zone. New York wind generation and forecast data, aggregated for 3 wind farms in geographical proximity, are available for the time period Jan 2017- Dec 2020. This is the best granularity of data available for New York state since individual New York wind farms have privacy concerns for their data. The aggregated data from collocated units in a zone are treated as from a single renewable asset. As such these aggregations do not diminish the value of the implementation of the methodology, as the farms are in close geographically proximity, and hence experience very similar generation risks. As shown in Figure 3, the New York wind assets belong to three different zones: Zone A, C and E. Following the framework developed in Gupta and Palepu (2023), hourly risk-free bid curves are computed for all days in the generation data time span for all the wind assets.

3.2 Solar Resources Data

.

Texas solar generation and forecast data at hourly frequency are available for single solar generation farms for the time period Jan 2015- Dec 2018. As shown in Figure 3, the Texas solar assets belong to two different zones: Far West zone and North Central zone. New York solar generation and forecast data from behind-the-meter (BTM) installations aggregated at the zonal level are acquired from NYISO for the time period Jan 2018 - Dec 2020. This is the best granularity of solar generation data available for the New York state since large solar farms don't yet exist in New York state. Figure 3 shows that the New York solar generation data are for three different zones: Zone A, C and E. Once again following the framework developed in Gupta and Palepu (2023), hourly risk-free bid curves are computed for all days in the generation data time span for each solar asset.

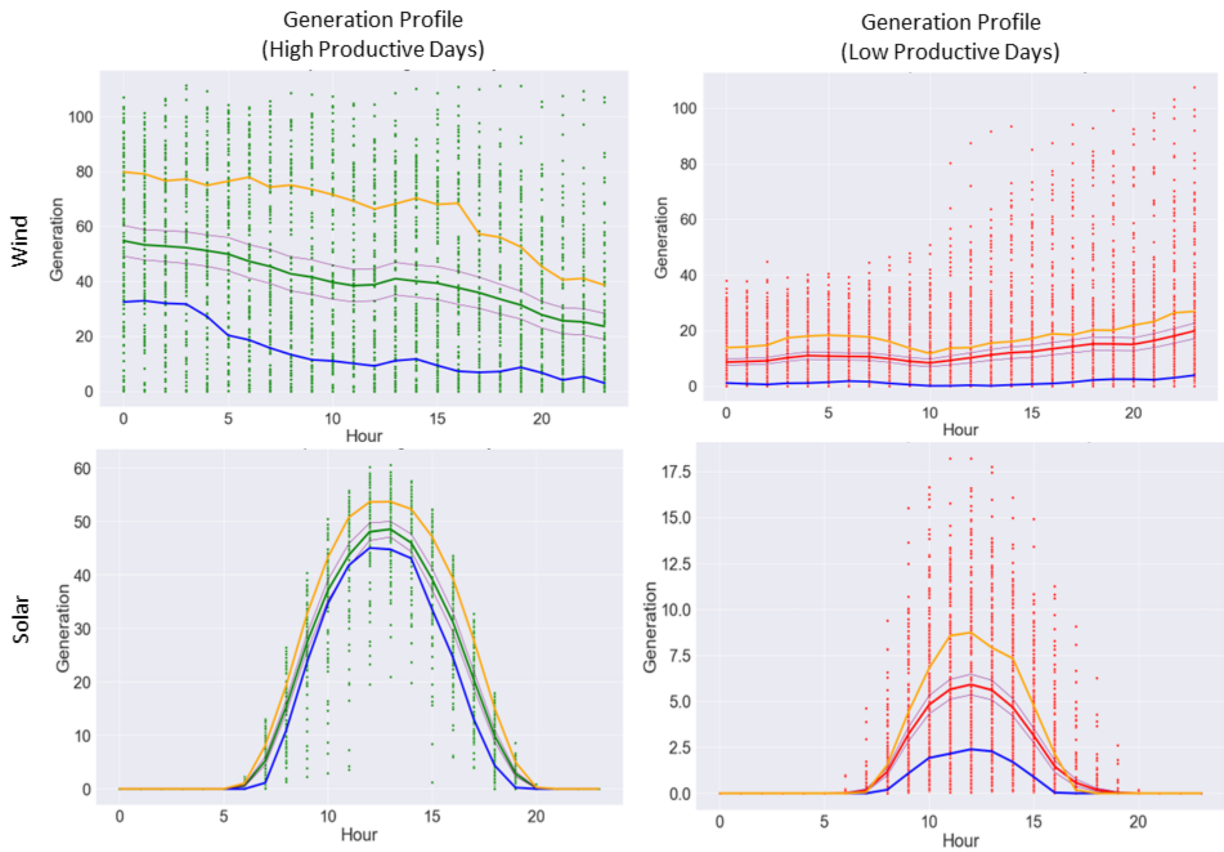

Figure 4: Sample generation profile of high productive (left) and low productive (right) 24 hour periods of wind (top) and solar (bottom) assets generation. The plot shows the mean, 95% confidence on the mean and 5th and 95th percentile of the type of day.

The hourly (diurnal) variation and seasonal differences are captured in the 24-hour generation profile of two different types of days in Figure 4. The differences in the patterns in the 24-hour window entail a different risk-free bidding strategy for different hours, and hence different shortfall rate characteristics. Notably, the solar generation is limited to daytime hours, with throughput gradually taking off and tapering down after sunrise and near sunset, respectively. The peak of generation is the most significant difference in the high and the low productive types of days.

3.3 Scoring Variables and Data

Based on the generation and forecast data for all the solar and wind generation assets, combined with the risk-free bid curves for each hour of all the assets, the actual amount of risk-free power contracted, Q_t, and actual shortfall rate, S_t, time series is generated. However, for this computation, system level price data are required. Day-ahead and real-time market electricity prices & real time load data for all New York Zones, namely A, C and E, for the time period Jan 2013-Dec 2020 are obtained from NYISO. The zonal electricity prices and load data are publicly available from NYISO website. The Henry-Hub daily natural gas spot price data are obtained from the U.S. Energy Information Administration, which is also publicly available. Day-ahead and real-time market electricity prices, real-time load data are similarly obtained for the same period of time for Texas zones from ERCOT. At the time of making the performance scoring, $t - \delta_1$, the predicted day-ahead price, $\bar{D}_t$, is determined in terms of the past values of relevant variables, from which expected amount of risk-free power contracted is determined.

Variable Name	Variable Symbol
Mean Generation	Mean Gen(t-1)
Standard deviation of Generation	Std of Gen(t-1)
Minimum Generation	Min Gen(t-1)
Generation Forecast	F(t)
Difference between 2PM and 10PM Forecast	2PM-10PM
Forecast Error	E(t-1)
Max Forecast Error in 24 hour window	Max forecast error(t-1)
Min Forecast Error in 24 hour window	Min forecast error(t-1)
Difference between Midnight and Noon Forecast Error	Midnight – Noon Err Diff
Predicted Power Contracted	Q_hat(t)
Predicted Day-Ahead Market Price	D_hat(t)
Predicted Real-Time Market Price	R_hat(t)
Natural Gas Futures Price	N(t)

Figure 5: Variables list and definition for the Risk-free Contract performance scoring.

Based on the variables and time series data described, we define and summarize the variables used in the performance risk scoring methodology in the table of Figure 5. The daily statistical summaries for renewable asset power generation for $t - 1$, namely, mean, minimum and standard deviation, are created. Additionally, we define the difference in peak versus off-peak generation forecast. Similarly, statistical summaries are created for the generation forecast error for $t - 1$, namely, maximum, minimum, difference in peak and off-peak, and the hourly forecast error. Using the data described in the section, we implement the performance scoring methodology of Section 2 to conduct a comparative study of performance of risk-free contracts of different assets located in different geographies.

4 Comparative Assessment of Performance Risk Scoring

The renewable generation assets, even within a state, can be in quite different territories that result in varying generation profile and risk-free contract characteristics. As summarized in Table 1 and shown in the maps of Figure 3, we have a sample of wind and solar generation farms in varied zones and regions of two reasonably large states of Texas and New York. In order to evaluate the performance of the risk scoring methodology, we conduct a detailed comparison by all the factors that may cause variation in the performance of risk scoring. We begin with displaying and discussing the variation in the overall shortfall distribution of the risk-free contracts by geography, asset type, zone and time of day in order to develop a basic view of the ground truth of the shortfall.

We implement the risk scoring methodology for predicting shortfall in the risk-free contracts issued by various assets and report the standard methodology performance indicators, such as, accuracy, precision, recall and F1 score. This provides an opportunity to investigate the variables that play an important role in scoring, serving both as an intuitive validation as well as offering insight on the factors that determine the shortfall. Finally, we conduct a comparison of risk scoring performance by time of day for risk-free contracts issued by geography, generation asset type, and season.

4.1 Risk-free Contract Shortfall Distribution

Figure 6 shows the complete risk-free contract shortfall distribution for a wind generation asset in the Far West zone of Texas during peak hour. Peak hour in the electricity market is defined by times in the day when the electricity demand, and as a result, the electricity price is at its peak in a 24 hour window. This typically occurs at around 2pm. The figure shows a high degree of imbalance in the shortfall outcomes with a very high percentage of outcomes at the zero shortfall level. This should be expected for a risk-free contract that must meet its obligation reliably. The non-zero shortfall outcomes are insignificant by scale, therefore an embedded plot is included in the figure that zooms into just the non-zero shortfall outcomes.

The conditional shortfall distribution shows a left skew and heavy right tail, implying that when the risk-free contract issued by this wind asset fails to meet its obligation, it tends to miss its target by a significant amount. This can occur from high forecast error or an aggressive bid when the day-ahead market price settles at a low level, thus picking up the aggressive risk-free bid by the wind asset. Investigation into the variable importance in the shortfall prediction will shed light on the role played by these variables in predicting the degree of shortfall.

For a more comprehensive baseline view of the shortfall distribution, we display the conditional shortfall distributions for non-zero outcomes of risk-free contracts of solar and wind assets in New York and Texas at peak versus non-peak hour in Figure 7. All the distributions display strong skewness, with right skew being typical of solar assets, while left skew appearing as a shared feature of the wind assets. Each asset maintains the type of skewness between peak and off-peak hour. This implies that when the risk-free contract offered by a solar asset fails to miss its obligation, it tends to do with a small margin, whereas a wind asset misses the target obligation quite drastically. This highlights the inherent heightened volatility of wind power generation, where the forecast error

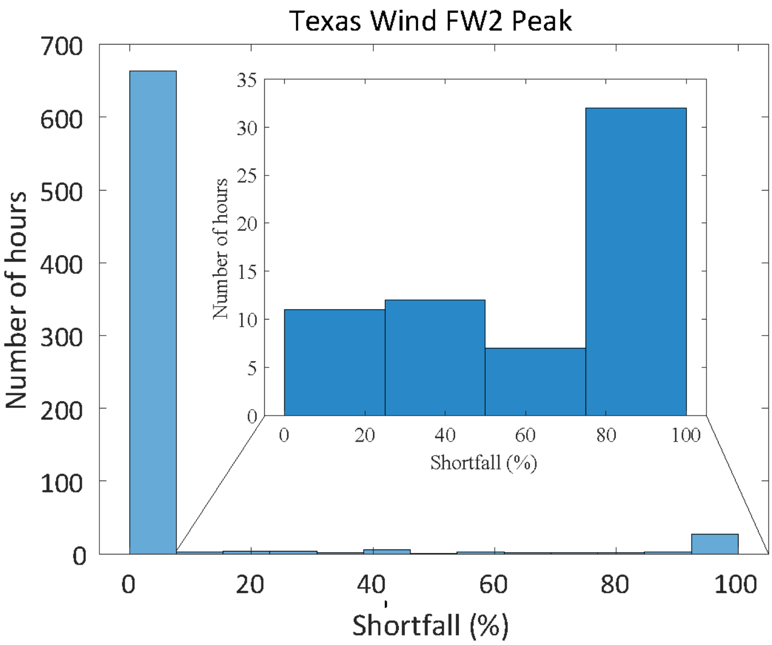

Figure 6: Shortfall distribution plot for the full data and the zoom into data points corresponding to non-zero shortfall. Shortfall is shown for a wind asset located in the Far West zone of Texas.

occasionally ends up being very high.

-

4.2 Performance Risk Scoring and Variable importance

We apply the performance risk scoring methodology to the risk-free contract offered by 12 renewable generation assets described in Table 3, which provides a variation by state, zone and asset type. The algorithm developed in Section 2.3 consists of two stages, where the first stage predicts whether there is a shortfall and the second one predicts the extent of shortfall in scenarios where a shortfall is predicted. The two stages were designed to address the imbalance in the shortfall outcomes being heavily dominated by the zero shortfall scenario, as displayed in Figures 6 and 7. In order to demonstrate the performance of the risk scoring methodology, we first demonstrate the performance of both stages individually, followed by the performance of the overall methodology for the chosen 12 generation assets' risk-free contracts.

Prediction performance is evaluated using the accuracy, precision, recall and F1 score metrics defined in terms of the prediction's confusion matrix. A confusion matrix consists of 'true positive' (TP), 'true negative' (TN), which are the diagonal entries of the confusion matrix, 'false positive' (FP), 'false negative' (FN), which are the off-diagonal entries of the confusion matrix. These performance metrics are defined as follows: Accuracy $= \frac{TN+TP}{TN+TP+FN+FP}$, Precision $= \frac{TP}{TP+FP}$, Recall $= \frac{TP}{TP+FN}$, and F1 Score $= 2\frac{Precision*Recall}{Precision+Recall}$. These metrics are best suited for a two outcome prediction task, which Stage 1 and Stage 2 predictions are. However, for the overall prediction task of more than 2 outcomes regarding the shortfall levels, we utilize an additional prediction performance metric

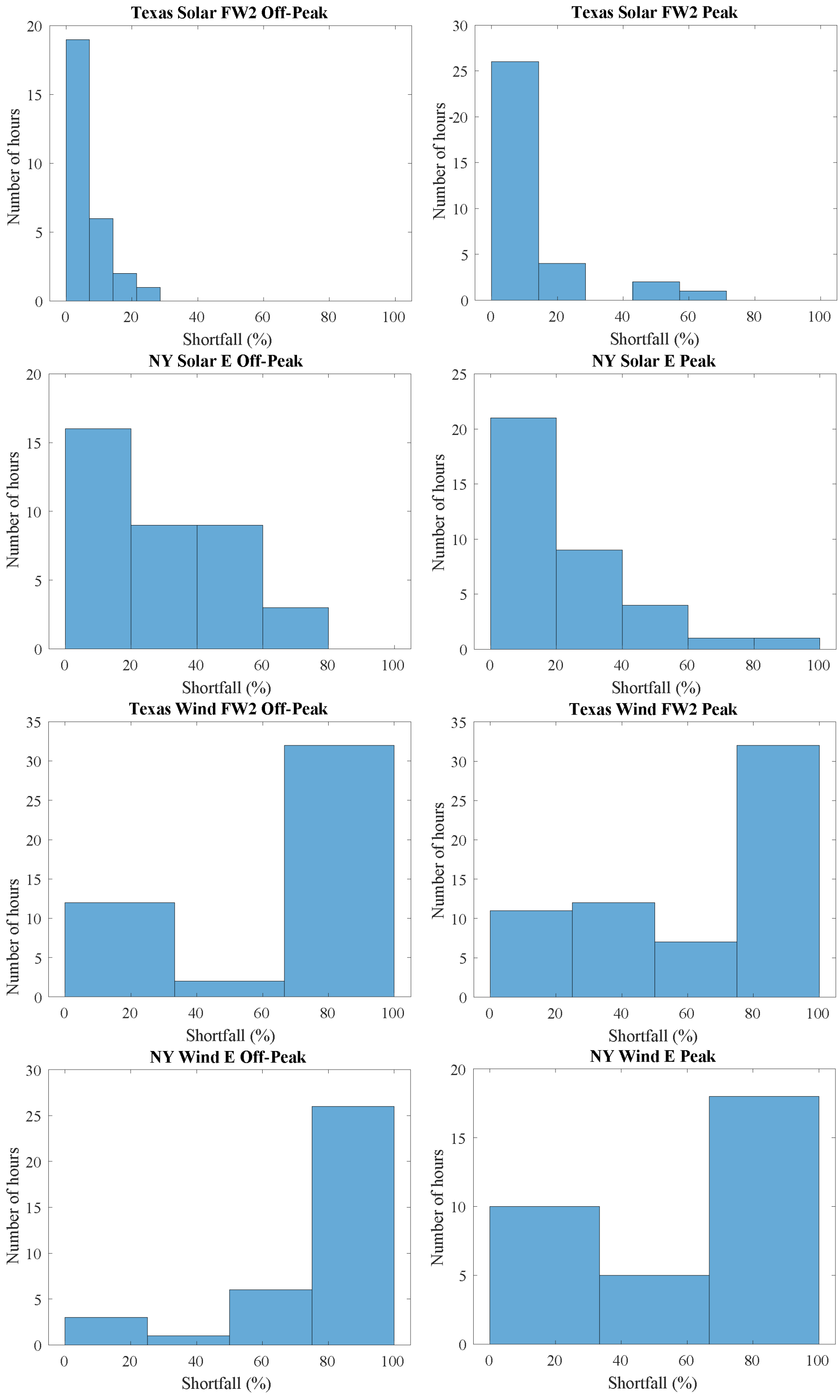

Figure 7: Conditional shortfall distributions for non-zero shortfall outcomes for a range of asset types, peak versus non-peak time in New York and Texas states.

beyond the above metrics, called the scoring performance metric (SPM), defined as: $\frac{||diagonal(C)||_{1,1}}{||C||_{1,1}}$, where C is the confusion matrix of the prediction task and $||.||_{1,1}$ is the entry-wise matrix 1-norm.

The table in Figure 8 displays the prediction performance metrics for the two outcomes prediction of Stage 1 for the risk-free contract issued by the chosen 12 generation assets. Overall the solar assets

Stage 1

Asset	Accuracy	Precision	Recall	F1 score
TX Solar NC1	1.00	1.00	1.00	1.00
TX Solar FW2	1.00	1.00	1.00	1.00
TX Solar FW3	1.00	1.00	1.00	1.00
TX Wind NC1	0.93	0.94	0.98	0.96
TX Wind NC2	0.94	0.94	0.99	0.97
TX Wind FW2	0.94	0.95	0.99	0.97
NY Solar A	0.99	1.00	0.99	0.99
NY Solar C	0.99	1.00	0.99	0.99
NY Solar E	0.99	1.00	0.99	1.00
NY Wind A	0.97	0.97	0.99	0.98
NY Wind C	0.96	0.96	0.99	0.98
NY Wind E	0.98	0.99	0.99	0.99

Figure 8: Accuracy, Precision, Recall and F1 Score of Stage 1 prediction performance of the performance risk scoring of risk-free contract.

show a better prediction performance than the wind assets, and Texas solar assets are more reliably predicted compared to their New York counterparts. These results are intuitively expected, thus supporting the quality of prediction obtained from our methodology. Among the wind assets, the prediction performance of New York wind assets is better than that of the Texas wind assets, which is somewhat counter-intuitive. The Accuracy metric is uniformly the lowest performance metric among the four considered in this table. F1 Score, which is considered a better indicator than Accuracy, shows a uniformly high level for Stage 1 prediction of all the 12 assets. The table in Figure 9 displays the Stage 2 two-outcome prediction performance metrics. The two outcomes for Stage 2 are a low non-zero shortfall level and a high shortfall level defined using a cut-off. The number of outcomes that get predicted at the low versus the high level is able to capture the left or the right skew in non-zero conditional shortfall distributions discussed in Section 4.1. The first 4 columns report the standard two-outcome confusion matrix performance metrics of Accuracy, Precision, Recall and F1 Score. Risk-free contracts issued by solar assets have a uniformly highly performance in Stage 2 as well. Wind assets have a worse performance in Stage 2 than in Stage 1, with worst indicated by the Recall metric, where the performance falls as low as 7% for New York wind asset in Zone E and 38% for Texas wind asset in North Central and Far West zones. This low Recall level reflects the high shortfall levels that are getting labeled as low shortfall outcomes in the Stage 2 prediction, which indicates an underestimation of a bad scenario. Therefore, appropriately adequate precaution should

Stage 2

Asset	Accuracy	Precision	Recall	F1 score	Combined
TX Solar NC1	1.00	1.00	1.00	1.00	1.0
TX Solar FW2	1.00	1.00	1.00	1.00	1.0
TX Solar FW3	1.00	1.00	1.00	1.00	1.0
TX Wind NC1	0.75	0.75	0.38	0.50	0.92
TX Wind NC2	0.82	0.60	0.86	0.71	0.93
TX Wind FW2	0.90	1.00	0.38	0.55	0.94
NY Solar A	1.00	1.00	1.00	1.00	0.99
NY Solar C	1.00	1.00	1.00	1.00	0.99
NY Solar E	1.00	1.00	1.00	1.00	0.99
NY Wind A	0.86	0.84	0.33	0.53	0.96
NY Wind C	0.87	0.83	0.56	0.67	0.95
NY Wind E	0.78	1.00	0.07	0.13	0.96

Figure 9: Accuracy, Precision, Recall and F1 Score of Stage 2 prediction performance of the performance risk scoring of risk-free contract. Last column of the table shows Scoring Performance Metric (SPM) using diagonal and off-diagonal confusion matrix norms for the overall performance risk scoring prediction of risk-free contract.

be taken regarding the level of shortfall prediction. The last column reports the combined Stage 1 and Stage 2 accuracy by the scoring performance metric, which is uniformly high, including in the cases where the Recall performance was low.

A crucial insight that can be sought from the performance risk scoring methodology is regarding the variables that feature as playing an important role in the prediction. Since the methodology consists of 2 stages and uses two supervised learning methods in an ensemble mode at each stage, variable importance can be inferred from each stage and supervised learning method. A summary of important variables in Stage 1 is provided in the table of Figure 10. These variables were described in the table of Figure 5 of Section 3.3. Stage 1 prediction using Random Forest method doesn't distinguish between asset type and geography for the variables that play an important role. The forecast of generation, F_t, emerges as the most important variable and the contracted amount of generation, $\hat{Q}_t$, under the risk-free contract is the second most important variable. These two variables intuitively appear to be important for judging whether the risk-free contract terms will be met. Neural network method chooses a different set of variables as most important, where a distinction in variable importance also emerges by asset type. Wind assets are best predicted by the forecast of generation, F_t, variable, while solar assets are predicted by a more customized combination of forecast indication, i.e. difference in the forecast for peak versus off peak times. The second order important variables maintain the asset distinction, with wind assets picking contracted amount and electricity market price variables and solar assets continue to focus on forecast and forecast error variables. Distinction in important variables by geography is most significant for New York versus Texas wind assets, where electricity market price variables appear to play a dominant role for Texas wind assets.

Asset	Random Forest 1st important	Random Forest 2nd Important	Neural Network 1st important	Neural Network 2nd Important
TX Solar NC1	E(t-1)	F(t)	2PM-10PM	Max forecast error
TX Solar FW2	F(t)	Q_hat(t)	2PM-10PM	Max forecast error
TX Solar FW3	F(t)	D_hat(t)	Max forecast error(t-1)	2PM-10PM
NY Solar A	F(t)	Q_hat(t)	Max forecast error(t-1)	D_hat(t)
NY Solar C	F(t)	Q_hat(t)	2PM-10PM	Max forecast error
NY Solar E	F(t)	E(t-1)	F(t)	Q_hat(t)
TX Wind NC1	F(t)	Q_hat(t)	D_hat(t)	F(t)
TX Wind NC2	F(t)	Q_hat(t)	D_hat(t)	R_hat(t)
TX Wind FW2	F(t)	Q_hat(t)	F(t)	D_hat(t)
NY Wind A	F(t)	Q_hat(t)	F(t)	Q_hat(t)
NY Wind C	F(t)	Q_hat(t)	F(t)	Q_hat(t)
NY Wind E	F(t)	Q_hat(t)	F(t)	Mean Gen(t-1)

Figure 10: Important variables in Stage 1 Prediction.

Asset	Random Forest 1st important	Random Forest 2nd Important	Neural Network 1st important	Neural Network 2nd Important
TX Solar NC1	F(t)	D_hat(t)	2PM-10PM	F(t)
TX Solar FW2	Q_hat(t)	F(t)	2PM-10PM	F(t)
TX Solar FW3	D_hat(t)	Q_hat(t)	2PM-10PM	F(t)
NY Solar A	F(t)	2PM-10PM	F(t)	Max forecast error
NY Solar C	Q_hat(t)	F(t)	2PM-10PM	F(t)
NY Solar E	Q_hat(t)	F(t)	2PM-10PM	F(t)
TX Wind NC1	Q_hat(t)	F(t)	F(t)	Std of Gen(t-1)
TX Wind NC2	F(t)	E(t-1)	F(t)	Std of Gen(t-1)
TX Wind FW2	Q_hat(t)	D_hat(t)	F(t)	Std of Gen(t-1)
NY Wind A	F(t)	Midnight – Noon Err Diff	F(t)	RF_pred
NY Wind C	Q_hat(t)	F(t)	Mean Gen(t-1)	E(t-1)
NY Wind E	F(t)	Q_hat(t)	Mean Gen(t-1)	Min Gen(t-1)

Figure 11: Important variables in Stage 2 Prediction.

A summary of important variables in Stage 2 is provided in the table of Figure 11. Other than the variables described in the table of Figure 5, 'RF_pred' refers to the prediction of outcome provided by the Random Forest method. The overall cast of characters of variables in Stage 2 is similar to that of Stage 1, however there are crucial differences in the important ones, including between the two ensemble methodologies of Stage 2. Random Forest quite uniformly elevates the amount contracted as the most important variable, irrespective of the asset type, which is a reassuring pick for the most important variable. Otherwise, the forecast of generation remains the most important variable for solar and wind assets. Most important variable in Neural Network for solar assets clearly shifts to the difference in peak versus off-peak forecast, while for wind assets the mean generation level assumes an important role. Random Forest's second most important variable appears to be a mixed bag, with most frequent occurrence of forecast of generation, and a new appearance of past forecast errors. In Neural Network method, for solar assets the forecast of generation is the dominant second most important variable, while for wind assets, especially in Texas, the standard deviation of generation is a regular feature. The New York wind assets each pick a different variable as a variable of second most importance, including using prediction of the Random Forest method. We additionally examined the third, fourth, etc. important variables of both the methods in both stages. These third and fourth important variables in Stage 1 and 2 are dominantly generation variables in the Neural Network, while for the Random Forest they are electricity market price and generation variables. All these variables were defined in the table for Figure 5. It is reassuring that prediction of Random Forest is not heavily relied upon in the Neural Network method, which implies that the two methods are drawing intelligence from different signals, and the ensemble method aggregates the learning for making the prediction.

4.3 Geographic Comparison of Performance Risk Scoring

There are noticeable differences in the properties and performance of risk scoring by assets and geographies. We examine these differences by each hour for a wind asset chosen in Texas and New York. There are significant hourly variations in all variables in the power sector, from load (or demand), prices, to wind and solar energy generation patterns. The electricity load and price diurnal pattern primarily arising from human daily activity and geographical proximity of power generation and consumption. Solar generation is governed by sunrise and sunset at different times of the day in different geographies.

Figure 12 demonstrates the performance risk scoring accuracy by the hour for a representative New York wind asset against a representative Texas wind asset. The plot is for all the 24 hours starting with the midnight 00:00hrs. On average the New York wind asset accuracy is higher than for the Texas wind asset, but the hour to hour variation in accuracy is significant for both assets. The overall performance for the New York wind asset using all hours combined is 0.95, which breaks down to an hourly performance that ranges between 0.87 and 0.96. The variation in the Texas wind asset is lower with an hourly minimum of 0.89 and maximum of 0.93. Difference in variation has a pattern, where Texas asset's performance is more accurate in the early morning hours before 05:00hrs, and New York asset's performance is better in the afternoon hours after 13:00hrs through early evening.

In contrast to the wind asset, the Texas solar asset has dominantly better scoring accuracy

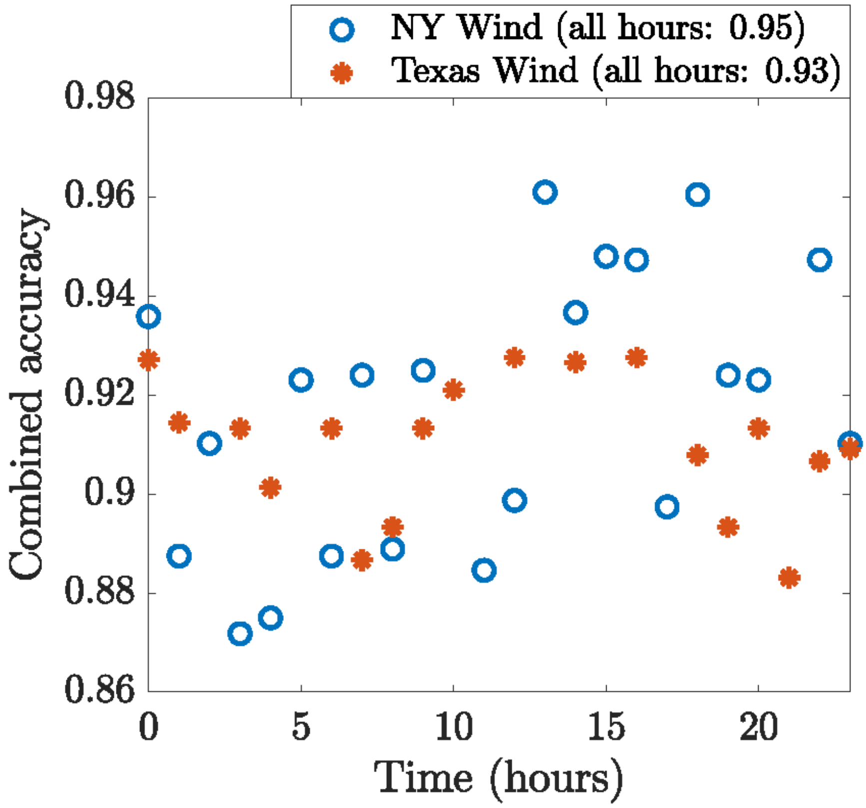

Figure 12: Hourly performance risk scoring of New York and Texas representative wind generation asset.

performance compared with New York solar asset, as shown in Figure 13. The figure shows hourly scoring accuracy for sunrise to sunset for a representative New York and Texas solar asset. As we have seen in all prior results, the risk-free contract drawn out of solar generation assets are more accurately assessed for their performance scoring. The lowest accuracy between the two geographies occurs in New York for 15:00hrs. At this same hour, the Texas solar asset shows perfect accuracy of risk scoring. This can be explained on the basis that New York solar asset contracts significantly for these hours due to high solar forecast of generation, but higher variation in the actual generation in New York compared to Texas results in the scoring accuracy to fall, even though it is at a high level of 0.92 in absolute terms.

4.3.1 Asset Type Comparison of Performance Risk Scoring

In Figure 12, a New York wind asset scoring performance was better than that of a Texas wind asset, while Texas solar outperformed New York solar in Figure 13. Figure 4.3 compares New York wind against solar, and shows that in the scale of performance by asset type, New York solar in fact has better scoring performance than New York wind asset. This is true in a more pronounced way for asset type comparison for Texas. For this combination of assets, the hourly asset type comparison

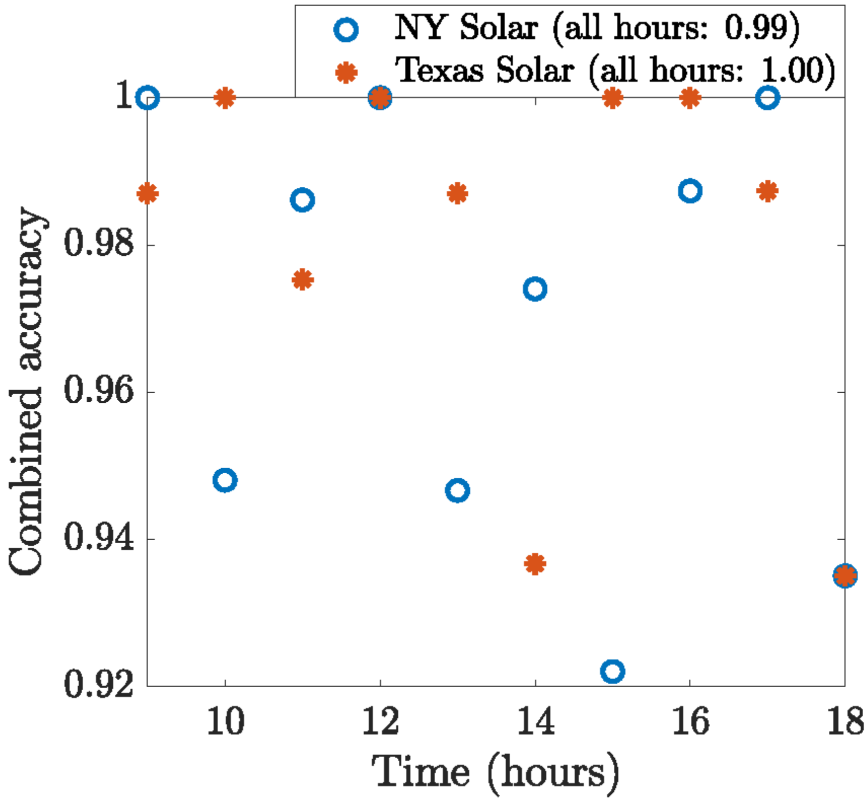

Figure 13: Hourly performance risk scoring of New York and Texas representative solar generation asset.

.

only makes sense in the sunrise to sunset period, as shown in the figure. The New York solar asset's performance is lowest in the peak hours of 13:00hrs to 15:00hrs, which is also the time when the New York wind asset shows the best scoring performance. This is a good complementary feature for grid level properties that when solar asset's scoring performance is relatively less reliable, the wind assets are scored more reliably.

4.4 Seasonal Comparison of Performance Risk Scoring

Electricity demand and renewable generation pattern both depend and vary by season. Therefore, we separate days by season, winter and summer, to compare the scoring performance by season. We specifically focus on a representative wind generation asset in Texas, for which Figure 15 shows the 24 hour scoring accuracy for the summer and winter months. Wind asset's risk-free contract is better predicted for its performance in the summer months compared to the winter months. The difference in performance of the scoring is the least during the peak hours of 12:00hrs and 15:00hrs between summer and winter months. The overall daily performance for the summer months for this Texas wind asset is 0.95, that breaks down to an hourly performance range of 0.9 to 0.96. The winter hourly scoring performance ranges from 0.88 to 0.93.

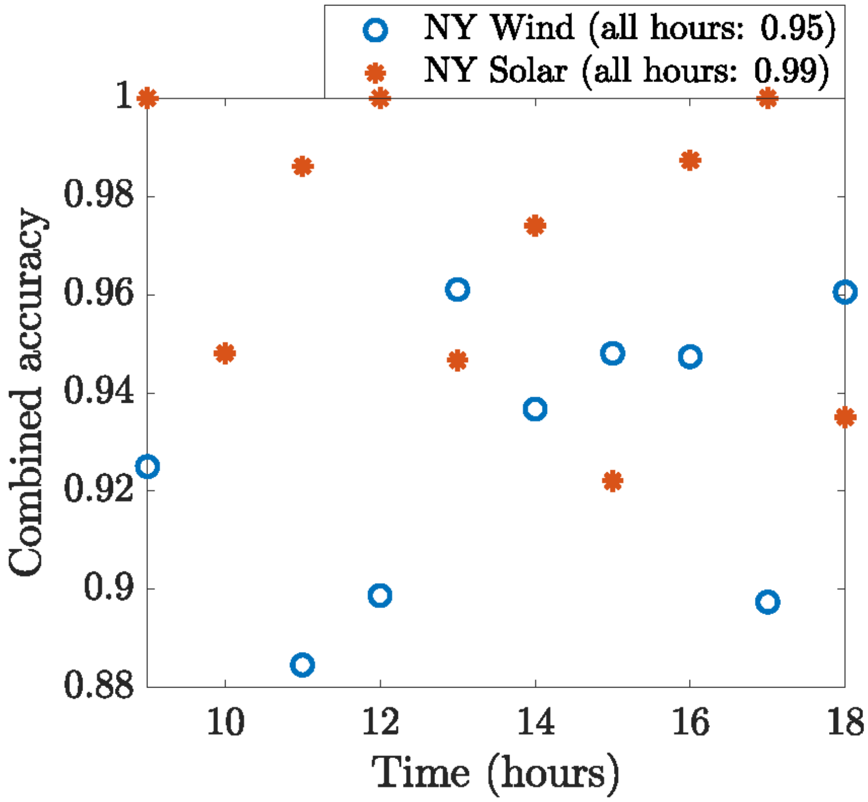

Figure 14: Hourly performance risk scoring of New York representative wind versus solar generation asset.

The Texas wind asset's difference in scoring performance between summer and winter months motivates to compare other asset types and geographies by season, as shown in the table of Figure 16. The New York solar and wind assets both demonstrate no difference in scoring performance in summer versus winter months. The Texas solar asset shows only a minor difference in scoring performance in summer versus winter months. In summary, the seasonal difference in scoring performance is most significant for the Texas wind asset. The uniformity in scoring performance across seasons is indicative of the strength of the methodology and its ability to recruit the features that deliver an accurate prediction of shortfall, but also points to a responsive definition of the risk-free contracts in terms of the important domain characteristics.

5 Conclusion and Future Work

The renewable generation targets in all geographies around the globe heavily rely on the ability to make the investment profitable. Profitable operation of renewable generation at the grid scale requires that these generators are treated at par with the traditional power generation resources in terms of their participation in the power markets. Mechanisms that allow the renewable generators to competitively sell their energy throughput into the market to command a fair revenue share are crucial

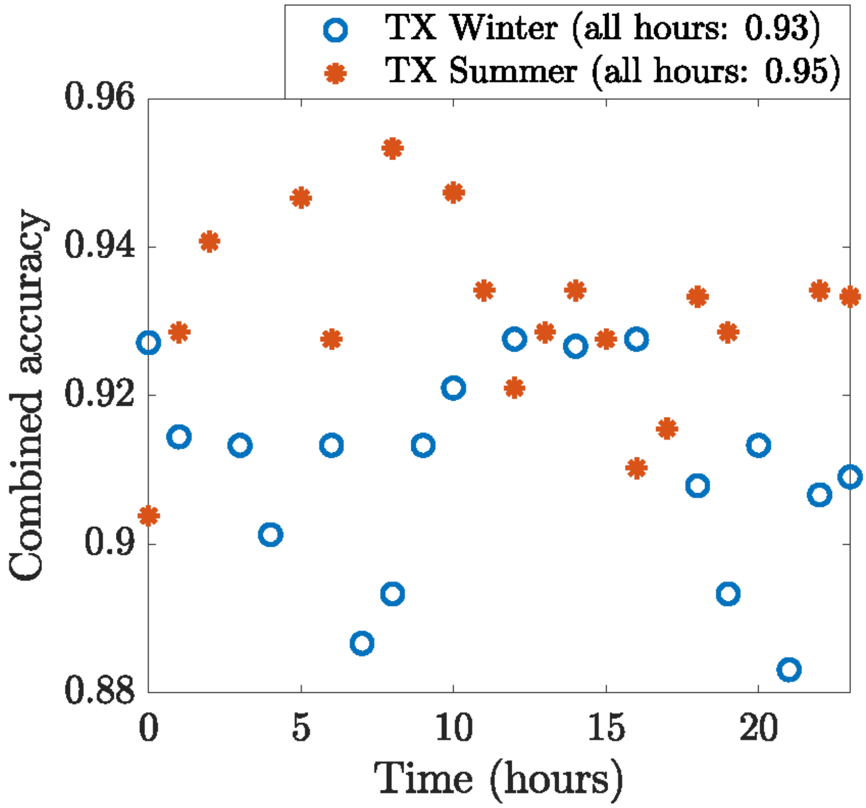

Figure 15: Hourly performance risk scoring of Texas representative wind generation asset in summer versus winter season.

Combined Accuracy for 2 Stages of the Scoring Algorithm

Asset	Season	
	Winter	Summer
NY Wind E	0.95	0.95
TX Wind FW2	0.93	0.95
NY Solar E	0.99	0.99
TX Solar FW2	1.00	0.99

Figure 16: Seasonal Scoring Performance Metric (SPM) for representative wind and solar risk-free contracts in New York and Texas.

for this purpose, rather than have the renewable generators depend on government subsidies for their viability. Innovations utilizing risk-responsive bidding strategies adapted to the inherent stochasticity and intermittence of renewable generation are emerging. This paper developed a performance risk scoring methodology for the validation of such risk-responsive bidding strategies, specifically aimed at risk-free contracting designed to compete with traditional and highly reliable natural gas based combined cycle generators.

The performance risk scoring of the renewable generators' risk-free contracting provides the necessary third-party input to the grid's system operator for a robust and reliable functioning of power markets. Similar to credit rating of debt issuance and securities issued based on debt securitization, our performance risk scoring provides indication of how likely a renewable generator will fulfill its risk-free contract's obligation. A key difference from debt rating is the power market's timeline and frequency requirement of risk-free contract performance risk scoring. In light of this requirement, we designed a two-stage ensemble machine learning methodology for the performance risk scoring of renewable generators' risk-free contracts.

We implemented and evaluated the risk-free contract's performance risk scoring for different renewable asset types, geographies, zones, time of the day, and seasons. The scoring algorithm performance, as measured using standard metrics, was overall high with some variations seen across wind versus solar assets, regions and hour of the day. The seasonal variation in algorithmic performance was minimal. The performance risk scoring algorithm additionally provided insights regarding variable importance for this prediction task. The predicted shortfall levels from the performance risk scoring can be mapped to a letter grading similar to debt rating, such as 'AAA,' 'AAA-,' 'AA,' 'AA-,' etc., where the accuracy of the prediction can be factored into the '+/-' adjustments. For instance, when a low shortfall is predicted for an asset's risk-free bid with very high accuracy, a letter grading of 'AAA+' is applied, but if a low shortfall level is predicted with lower accuracy, a letter grading of 'AAA-' may be assigned. A grid operator may take the grading explicitly into account in its unit commitment and economic dispatch considerations, as well as provide a system level guideline to renewable generators regarding permissible predicted shortfall and accuracy for bids to be incorporated for market clearing.

Beyond the risk-free bids, which is a very conservative bid, a renewable generator can also make riskier bids following the tranche principles of securitization, as shown in Figure 1. Similar to the performance risk scoring of the risk-free bids, the shortfall prediction algorithm can be extended to score the riskier bids made by a renewable generator. In a networked system, such as a power grid, each node and assets interfacing with the network at these nodes must function in tandem. Moreover, renewable assets in geographical proximity face similar weather conditions. These relations can be utilized for a network-based learning enhancement for contractual shortfall prediction, both for risk-free and riskier bids of renewable generators. These developments can greatly improve profitable integration of renewable generation in power grids, thus making cleaner energy more feasible across the globe.

References

Alafita, T. and Pearce, J. (2014). Securitization of residential solar photovoltaic assets: costs, risks and uncertainty. *Energy Policy*, 67:488–498.

Albanesi, S. and Vamossy, D. F. (2019). Predicting consumer default: A deep learning approach. Technical report, National Bureau of Economic Research.

Altman, E. I., Haldeman, R. G., and Narayanan, P. (1977). Zetatm analysis a new model to identify bankruptcy risk of corporations. *Journal of banking & finance*, 1(1):29–54.

Bessembinder, H. and Lemmon, M. L. (2002). Equilibrium pricing and optimal hedging in electricity forward markets. *the Journal of Finance*, 57(3):1347–1382.

Bhattacharya, S., Gupta, A., Kar, K., and Owusu, A. (2020). Risk management of renewable power producers from co-dependencies in cash flows. *European Journal of operational research*, 283(3):1081–1093.

Bird, L., Lew, D., Milligan, M., Carlini, E. M., Estanqueiro, A., Flynn, D., Gomez-Lazaro, E., Holttinen, H., Menemenlis, N., Orths, A., et al. (2016). Wind and solar energy curtailment: A review of international experience. *Renewable and Sustainable Energy Reviews*, 65:577–586.

Boyes, W. J., Hoffman, D. L., and Low, S. A. (1989). An econometric analysis of the bank credit scoring problem. *Journal of Econometrics*, 40(1):3–14.

Brockett, P. L., Wang, M., and Yang, C. (2005). Weather derivatives and weather risk management. *Risk Management and Insurance Review*, 8(1):127–140.

Cartea, Á. and Villaplana, P. (2008). Spot price modeling and the valuation of electricity forward contracts: The role of demand and capacity. *Journal of Banking & Finance*, 32(12):2502–2519.

Coulon, M., Powell, W. B., and Sircar, R. (2013). A model for hedging load and price risk in the texas electricity market. *Energy Economics*, 40:976–988.

Das, S. and Basu, M. (2020). Day-ahead optimal bidding strategy of microgrid with demand response program considering uncertainties and outages of renewable energy resources. *Energy*, 190:116441.

Deng, S.-J. and Oren, S. S. (2006). Electricity derivatives and risk management. *Energy*, 31(6-7):940–953.

Doege, J., Fehr, M., Hinz, J., Lüthi, H.-J., and Wilhelm, M. (2009). Risk management in power markets: the hedging value of production flexibility. *European Journal of Operational Research*, 199(3):936–943.

Falbo, P., Felletti, D., and Stefani, S. (2010). Integrated risk management for an electricity producer. *European Journal of Operational Research*, 207(3):1620–1627.

Ferruzzi, G., Cervone, G., Delle Monache, L., Graditi, G., and Jacobone, F. (2016). Optimal bidding in a day-ahead energy market for micro grid under uncertainty in renewable energy production. *Energy*, 106:194–202.

Fischer, T. (2015). Market structure and rating strategies in credit rating markets–a dynamic model with matching of heterogeneous bond issuers and rating agencies. *Journal of Banking & Finance*, 58:39–56.

Fraisse, H. and Laporte, M. (2022). Return on investment on artificial intelligence: The case of bank capital requirement. *Journal of Banking & Finance*, 138:106401.

Gabig, N., Cohen, B., and Kapoor, M. (2015). Solar abs: understanding market potential. *The Journal of Structured Finance*, 21(1):105–110.

Gupta, A. (2014). *Risk management and simulation*. Taylor & Francis, Boca Raton.

Gupta, A. and Palepu, S. (2023). Designing risk-free service for renewable wind and solar resources. *European Journal of Operational Research*.

Hain, M., Schermeyer, H., Uhrig-Homburg, M., and Fichtner, W. (2018). Managing renewable energy production risk. *Journal of banking & finance*, 97:1–19.

Hossain, A. T. and Kryzanowski, L. (2019). Global financial crisis after ten years: A review of the causes and regulatory reactions. *Managerial Finance*.

Hyde, D. and Komor, P. (2014). Distributed PV and securitization: made for each other? *The Electricity Journal*, 27(5):63–70.

Jacobson, T. and Roszbach, K. (2003). Bank lending policy, credit scoring and value-at-risk. *Journal of banking & finance*, 27(4):615–633.

Jarrow, R. A. and Turnbull, S. M. (2000). The intersection of market and credit risk. *Journal of Banking & Finance*, 24(1-2):271–299.

Jiang, J. N. and Chen, H. (2005). Integrating the power industry into the larger economy via electricity-backed asset securitization. *The Electricity Journal*, 18(6):46–54.

Kellner, R., Nagl, M., and Rösch, D. (2022). Opening the black box–quantile neural networks for loss given default prediction. *Journal of Banking & Finance*, 134:106334.

Khandani, A. E., Kim, A. J., and Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34(11):2767–2787.

Krupa, J. and Harvey, L. D. (2017). Renewable electricity finance in the United States: a state-of-the-art review. *Energy*, 135:913–929.

Kvamme, H., Sellereite, N., Aas, K., and Sjursen, S. (2018). Predicting mortgage default using convolutional neural networks. *Expert Systems with Applications*, 102:207–217.

Liu, W., Wang, J., Xie, J., and Song, C. (2007). Electricity securitization in China. *Energy*, 32(10):1886–1895.

Lowder, T. and Mendelsohn, M. (2013). Potential of securitization in solar pv finance. Technical report, National Renewable Energy Lab.(NREL), Golden, CO (United States). https://www.nrel.gov/docs/fy14osti/60230.pdf.

Lucheroni, C. and Mari, C. (2021). Internal hedging of intermittent renewable power generation and optimal portfolio selection. *Annals of Operations Research*, 299(1-2):873–893.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2):449–470.

Morkoetter, S., Stebler, R., and Westerfeld, S. (2017). Competition in the credit rating industry: Benefits for investors and issuers. *Journal of Banking & Finance*, 75:235–257.

Müller, A. and Grandi, M. (2000). Weather derivatives: a risk management tool for weather-sensitive industries. *The Geneva Papers on Risk and Insurance. Issues and Practice*, 25(2):273–287.

Nikpour, A., Nateghi, A., Shafie-khah, M., and Catalão, J. P. (2021). Day-ahead optimal bidding of microgrids considering uncertainties of price and renewable energy resources. *Energy*, 227:120476.

Pérez-González, F. and Yun, H. (2013). Risk management and firm value: Evidence from weather derivatives. *The Journal of Finance*, 68(5):2143–2176.

Prokhorov, O. and Dreisbach, D. (2022). The impact of renewables on the incidents of negative prices in the energy spot markets. *Energy Policy*, 167:113073.

Prol, J. L., Steininger, K. W., and Zilberman, D. (2020). The cannibalization effect of wind and solar in the california wholesale electricity market. *Energy Economics*, 85:104552.

Rosenberg, E. and Gleit, A. (1994). Quantitative methods in credit management: a survey. *Operations research*, 42(4):589–613.

Steenackers, A. and Goovaerts, M. (1989). A credit scoring model for personal loans. *Insurance: mathematics & economics*, 8(1):31–34.

Stolper, A. (2009). Regulation of credit rating agencies. *Journal of Banking & Finance*, 33(7):1266–1273.

Thompson, M. (2013). Optimal economic dispatch and risk management of thermal power plants in deregulated markets. *Operations Research*, 61(4):791–809.

Vehviläinen, I. and Keppo, J. (2003). Managing electricity market price risk. *European Journal of Operational Research*, 145(1):136–147.

Wang, J., Zhong, H., Tang, W., Rajagopal, R., Xia, Q., Kang, C., and Wang, Y. (2017). Optimal bidding strategy for microgrids in joint energy and ancillary service markets considering flexible ramping products. *Applied Energy*, 205:294–303.

Wojtowicz, M. (2014). Cdos and the financial crisis: Credit ratings and fair premia. *Journal of Banking & Finance*, 39:1–13.

Wozabal, D. and Rameseder, G. (2020). Optimal bidding of a virtual power plant on the spanish day-ahead and intraday market for electricity. *European Journal of Operational Research*, 280(2):639–655.