Investigation of endogenous renewables and demand resolutions for energy market simulations

1st Florian Maurer
University of Applied Sciences Aachen
Aachen, Germany
0000-0001-8345-3889

2nd Jonathan Sejdija
University of Applied Sciences Aachen
Aachen, Germany
0009-0009-3615-4160

3rd Volker Sander
University of Applied Sciences Aachen
Aachen, Germany
0009-0006-8850-8520

***Abstract*—This study presents a comprehensive analysis of energy market simulation results across varying spatial resolutions, focusing on the impact of different NUTS (Nomenclature of Territorial Units for Statistics) aggregations for Germany, corresponding to the simulation of renewables from weather data. The research aims to provide insights into the trade-offs between computational complexity and result accuracy in energy system modeling. Further on, it explains and discovers various limitations in input data. The exploration employs the agent-based energy market simulation model ASSUME to calculate results for scenarios with different data resolutions. Spatially, the model considers different NUTS levels, ranging from NUTS1 (states) to NUTS3 (districts) aggregations, allowing for a detailed examination of how geographical granularity affects market outcomes. One complexity is the generation of localized time series data for renewables which is done by calculating the average weather data in each NUTS area. All agents of a given type in a NUTS area are modeled together, which increases the number of simulation agents as well as increasing the spatial resolution. Results demonstrate that finer spatial resolutions do not always provide more accurate representations of localized energy dynamics, particularly in regions with significant solar generation with unknown self-consumption. The study also shows that the change in energy usage and unknown consumptions make it hard to create a fine forecast of energy demand from standard profiles. This research demonstrates that an improved spatial resolution does not necessarily result in more accurate simulation results. Hence, there is the need for an awareness of side effects caused by deficits in data accuracy. This highlights the need for further improvements to available input data, although significant progress has been made in recent years.**

***Index Terms*—agent-based modeling, energy market, energy dispatch, open source, comparative simulation**

I. INTRODUCTION

Latest developments in data availability made it possible to simulate energy markets with higher spatial resolution of input data sets from mostly endogenous time series. While most market simulations are highly adjusted towards the input data sets, which is selected and calibrated in a time-consuming manner, open data makes it possible to simulate markets with just a handful of additional assumptions.

In this study, the open source market simulation ASSUME is used in combination with public data of the German energy system to simulate the day-ahead market. This is solely based on the available generation from the German power plant register MaStR, historic load from ENTSO-E, historic weather from ECMWF and historic fuel prices from Instrat PL. All of which is publicly available for the years 2018-2024.

Here, the question arises if a higher spatial resolution of the weather data, demand data, and agents also provides a better fit to historic market results. Additionally, a general overview of simulation results from such an endogenous simulation is given for different years, providing insights of the simulation impact.

II. RELATED WORK

Energy market simulations have been extensively studied using various modeling techniques. Among them, agent-based modeling (ABM) has proven particularly effective in capturing market dynamics, offering valuable insights into energy systems [1]. High-resolution models, which integrate detailed spatial and temporal factors, have been systematically compared in [2], revealing comparable dispatch results between agent-based and fundamental market models.

The importance of real-world data for model validation and calibration has been emphasized in [3], enhancing the accuracy and applicability of simulations. Efficiency comparisons of ABM approaches are explored in [4], while [5] evaluates two agent-based market models using a calibrated input data set. Furthermore, the role of spatially and temporally resolved models in improving simulation precision is discussed in [6]. However, existing literature does not provide a methodology for developing market simulations solely from publicly available data. Instead, price adjustments and additional modifications are typically made to align results with expectations. When evaluating scientific models it can be beneficial to reproduce the results.

Furthermore, individual investigation of simulation results on different spatial resolutions are a core contribution to the energy market research community.

III. METHODOLOGY

Accurate market modeling depends on careful calibration and evaluation of input data. Often, only partial datasets are available, extracted from various public and semi-public energy system studies. This study introduces a broader approach

979-8-3315-1278-1/ 25/$31.00

that enables modeling any year with available historical data while also serving as a benchmark tool for future simulations.

Since all input data is publicly available, the methodology prioritizes a balanced simulation approach, avoiding overfitting to historical results. Most data is therefore used in its raw form, which may reduce the potential fit to past observations but ensures generalizability.

Market simulations rely on several key datasets, including power plant inventories, renewable generation time series, price data, and demand series. Each dataset is analyzed for historical fit to assess data quality and identify necessary improvements for market modeling. Finally, multiple simulations with varying spatial aggregations are conducted to evaluate key drivers of simulation performance and explore strategies for accurately modeling real market scenarios.

Demand time series

One approach to creating a demand time series for the future relies on estimations of Germany's total energy demand for a period of one year. In the second step, an actual load time series is created the BDEW standard load profiles[1]. This approach requires total energy demand data at the NUTS regional level, as energy consumption varies across regions. This can be done using the demand calculation from openEgo[2], which includes aggregated demand for each NUTS area. The energy demand varies from region to region, as can be seen in Figure 1, which shows the synthetic data per NUTS3 area of Germany.

To create a demand time series from such totals, the BDEW data is used by adding up standard load series for specific sectors, such as commercial, household, and agricultural sectors. In this study, the open-source tool demandlib is used to generate demand profiles [3].

While BDEW standard load profiles help generate aggregated demand profiles, they are insufficient for reconstructing exact yearly time series for total national demand. To address this, a linear regression model is used to optimize coefficients a_1 to a_{10}, adjusting the weight of the standard load profiles to best match historical demand data.

$$\begin{aligned} y = a_1 g_1 + a_2 g_2 + a_3 g_3 + a_4 g_4 + a_5 g_5 + a_6 g_6 \\ + a_7 g_7 + a_8 h_0 + a_9 \cdot l_1 + a_{10} \cdot l_2 + \epsilon \end{aligned} \tag{1}$$

The model optimizes the coefficients for the BDEW standard load profiles across commercial consumers (g_1 to g_7, households (h_0), and agriculture (l_1 and l_2). The resulting linear combination is then used to generate a synthetic hourly load profile, which is scaled according to the local annual demand of each NUTS3 region.

This approach yields an overall acceptable fit, as illustrated in Figure 2. However, it is important to note that demand forecasting for future scenarios may not fully replicate historical observations. One key limitation is the absence of industrial load profiles for consumers that fall outside standard load categories but have continuous load measurement (RLM) data. These consumers, typically with an annual consumption exceeding 100,000 kWh, are not captured in publicly available load series.

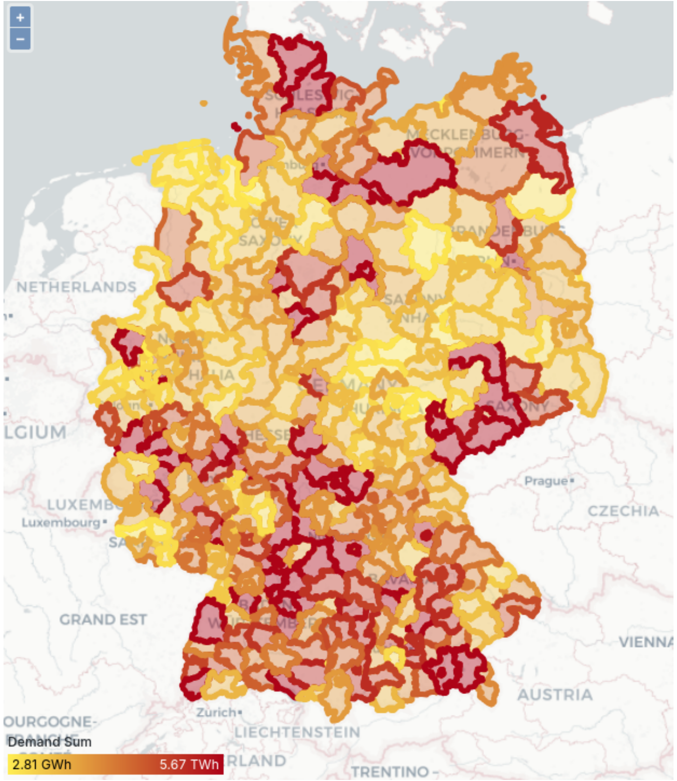

Fig. 1. Heatmap of yearly energy demand per NUTS3 area for Germany, publicly available from OpenEgo project, based on synthetic data generated with Openstreetmap

Despite these limitations, the use of standard load profiles based on public data provides a reasonable approximation of demand patterns. The mean absolute error (MAE) for the fitted time series in 2015 is 3 GW, representing 8.1 % of the average load of 57 GW. When comparing the fitted demand model to actual demand in 2023, the MAE increases to 8 GW, or 15.8 % of the observed load.

Of course, there are various other options to generate synthetic demand data for future scenarios, such as simply adjusting the weekly pattern of a different year, which provides average errors of about 10%, yet a standardized method does not exist. Various other methods include usage of ARIMA or Machine Learning based models [7]–[9], which generally require retraining throughout time and can not easily be adjusted to model assumptions [10]. As illustrated in Figure 3, the proposed method ensures continuity and incorporates seasonality while disregarding peak loads. The decline in load during off-peak hours is estimated with a degree of exaggeration, while the impact of solar generation has the effect of reducing the adaptability of the system to future scenarios.

---

[1] https://www.bdew.de/energie/standardlastprofile-strom/

[2] https://github.com/openego/data_processing

[3] https://github.com/oemof/demandlib

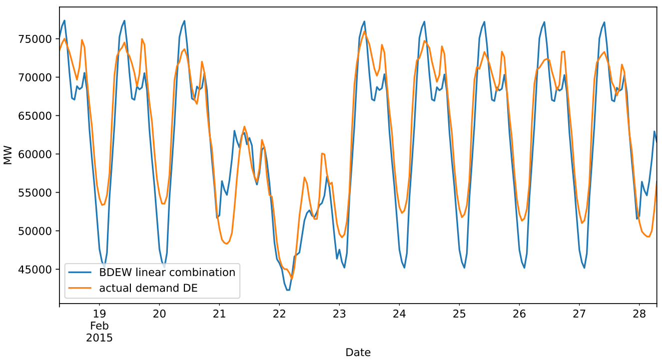

Fig. 2. Comparison of best linear combination from BDEW profiles to the actual power demand of Germany in 2015

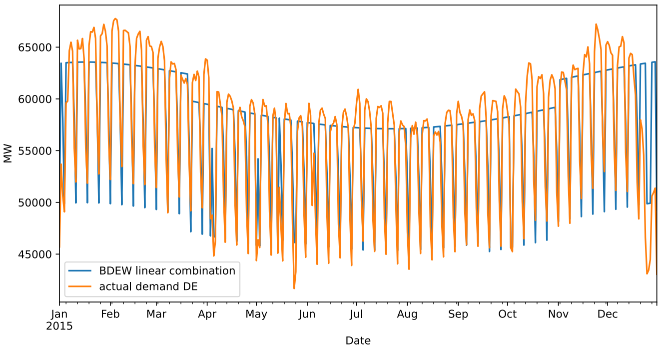

Fig. 3. Comparison of best linear combination from BDEW profiles to the actual power demand of Germany in 2015

Power plants

The existing power plants are obtained from the German register of existing power plants (Marktstammdatenregister)[4]. This data set encompasses the commissioning date and the decommissioning dates of power plants along with technical power parameters making it a valuable resource for modeling individual power plants. By using the Open-Energy-Data-Server[5], one can efficiently query the power plant database according to the different scenarios and filter by generation technologies [11]. The Marktstammdatenregister contains locational information as postal code, which can be translated into NUTS areas and subsequently employed in spatial simulations.

This capability enables the creation of customized simulations for specific NUTS areas or regions. Moreover, the continuous availability of the data sources in the database facilitates adaptation to future scenarios with minimal effort.

While some of the fuel costs are taken from a recent study [12], accounting for time-dependent fuel costs of conventional power plants requires publicly available, machine-readable data. The Energy Instrat Platform[6] provides fuel costs for hard coal, gas, oil and EU ETS prices under a permissive license as shown in Table I. This source is used to accommodate the fluctuating prices when simulating different years without larger adjustments to each year.

Technology	Prices
lignite	2.3 €/MWh
biomass	20 €/MWh
nuclear	8 €/MWh
hard coal	dynamic coal price
oil	dynamic oil price
gas	dynamic gas price
co2	EU ETS dynamic price

TABLE I

FUEL COSTS OF THE SIMULATIONS. TO ADHERE TO PRICE CHANGES THROUGHOUT THE DIFFERENT YEARS, KNOWN HISTORIC PRICES ARE USED WHERE AVAILABLE FROM INSTRAT PL, WHILE STATIC VALUES ARE TAKEN FROM EXISITING RESEARCH [12]

Renewables time series

Similar to demand series, typically only the national generation series exists for renewables power production of wind and solar. One can disaggregate this into individual NUTS areas using a bottom-up approach. The European Center for Medium-Range Weather Forecasts (ECMWF) provides individual weather data on a 10-kilometer per 10-kilometer grid, enabling the calculation of solar and wind generation from irradiation and respective wind speed. The installed power of solar and wind can as well be used from the register of power plants, as this even includes small-scale household solar installations.

The conversion of meteorological data and documented power plants to a renewable generation time series is accomplished through the utilization of two prominent Python libraries: windpowerlib and pvlib. Pvlib utilizes disparate diffusion models, thereby enabling the calculation of direct normal irradiation (DNI) and diffuse horizontal irradiation (DHI) from the general horizontal irradiation (GHI) provided by ECMWF [13]. The study adopts the Erbs diffusion model [14] as a baseline to generate realistic solar power estimates from the GHI weather information provided in the ECMWF data set [15]. The deviations from the actual solar generation are illustrated in Fig. 4.

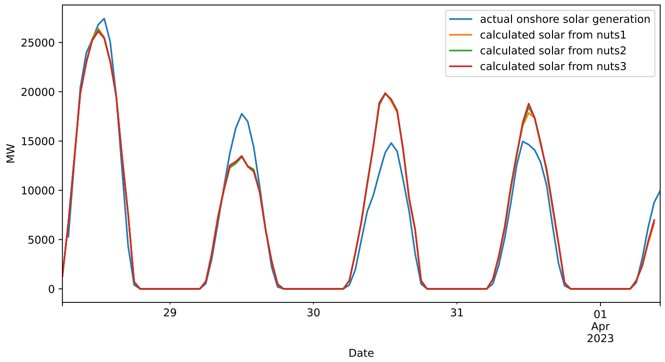

Fig. 4. Visualization of calculated solar generation using average DE weather of the year 2023 for different spatial NUTS aggregations

---

[4] https://www.marktstammdatenregister.de/MaStR

[5] https://github.com/open-energy-data-server/open-energy-data-server

[6] https://energy.instrat.pl/en/about-energy-instrat-pl/

Similar, the wind speed at 10 meters height is calculated from the meridional wind speed and zonal wind speed from ECMWF at 10 meters height through Equation 2.

$$V_{wind} = \sqrt{V_{zonal}^2 + V_{meridional}^2} \tag{2}$$

which is then used with windpowerlib to calculate the wind speed at a known height using the hellman equation. The interpolation of the wind power curve is finally derived from the wind power equation in Eq. 3 or empirical model data derived from the turbine type where available [16].

$$P_{wind} = \frac{1}{2}\rho A v^3 \tag{3}$$

This methodology is analogous to the solar power calculated for different aggregations of NUTS areas of Germany. The deviations from the actual wind generation are shown in Fig.5. One has to note, that the overall quality of the wind time series improves with greater spatial resolution, while the overall performance slightly decreases for the solar forecast in the spatial resolution. This is connected to the generally high data quality of wind power generation, while solar generation takes place in a much more distributed way, suggesting a lower data quality. The publicly available historic solar generation data from ENTSO-E does not include self-consumption and home battery storages which leads to additional inaccuracies, which are piling up with increasing market penetration of distributed renewables.

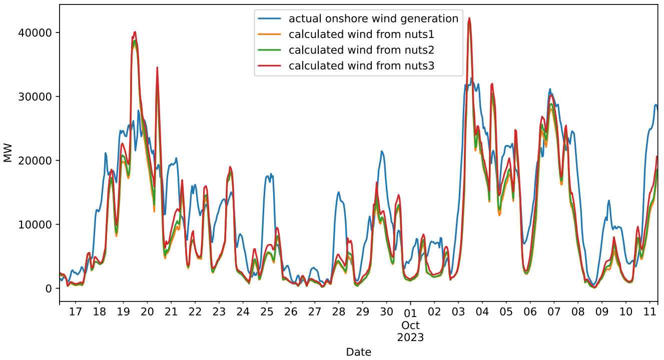

Fig. 5. Visualization of calculated solar generation using average DE weather of the year 2023 with different spatial NUTS aggregations

Concluding the introspection of input data series, the demand for better data quality in the power plant register in solar exists. This issue is further complicated by the limited availability and irregular updates of such data, with only broad information on azimuth and tilt generally accessible. Moreover, standardized load profiles for demand data do not account for potential solar generation, self-consumption, or storage integration, as these elements remain invisible to the grid. While such factors are crucial for analyzing local energy markets, they are not considered in this study. A potential solution could involve leveraging publicly available consumption data from smart meters, though their deployment in Germany remains limited [17].

In general, widely accepted load forecasts that reflect the evolving energy system are essential to minimize the need for subjective assumptions by energy system modelers. When simulating historical energy systems, using actual historical demand data - when available - rather than standardized load profiles allows for a more accurate representation of peak loads and demand fluctuations.

Market Model

The comprehensive market model ASSUME (Agent-based Simulation for Studying and Understanding Market Evolution) is used for this study to model the market behavior. It makes it possible to model different market configurations [18], including complex orders [19]. ASSUME has been compared to the established market modeling tool AMIRIS in [5] and provides detailed visualization capabilities of dispatch and market results.

ASSUME is configured using an object oriented approach, while also including interoperability scripts to load examples from CSV files, generic databases, AMIRIS [20] or PyPSA [21] input data sets. It includes an interface to the Open-Energy-Data-Server which holds the simulation data in different geospatial resolution. A fundamental component is the ad-hoc investigation of market scenarios and outcomes, which facilitates a comparison with the actual results of the EPEX spot auction, as illustrated in Figure 6 for the case of Germany in 2023 for NUTS1. This figure reveals a cost structure analogous to that of the EPEX spot market results, although the volume that has been cleared differs. This disparity arises because the market simulation clears the entirety of the volume in its spot market, instead of multiple different market options.

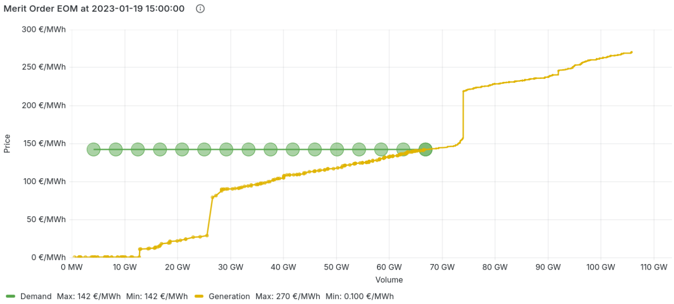

Fig. 6. Merit order visualization of the market result of a given hours from the simulation dashboard of Germany 2023 NUTS1. Results for other NUTS aggregation vary only slightly.

Consequently, this comparison exclusively encompasses the most salient market stage, while the potential for additional markets, such as intra-day or balancing markets, is considered for future scenarios.

Each market participant submits bids for the maximum capacity of each unit covered by the market participant, with the respective marginal cost of the unit. Due to the simplicity of the bidding strategy employed in the simulation, it is statically defined and does not adapt to market results.

IV. SIMULATION RESULTS

Qualitative and quantitative criteria are used to compare the diverse scenarios under consideration. The results are compared with historically realised data in order to establish a benchmark for the different energy market simulations.

To this end, a set of scenarios concerning the German electricity market from 2019 to 2024 is employed. The scenario input data is published under a permissive licence and can be reproduced with the Open-Energy-Data-Server. The applied workflow facilitates the comparison of disparate metrics (e.g. electricity prices, power plant dispatch and system costs) and the subsequent comparison with actual historical prices and dispatch.

To account for the spatial resolution of distinct NUTS areas, multiple agents are used. Each NUTS region utilizes the average weather time series for the area, as demonstrated for renewable energy generation.

Fig. 7 presents an example of a market simulation dispatch time series, highlighting the impact of omitted heat demand modeling. Since additional heat demand is not publicly available, it was excluded from the market input datasets. This omission leads to stronger ramping of conventional generation compared to historical data. This underscores the importance of accurate input data and how trade-offs in model assumptions can significantly impact the simulated dispatch.

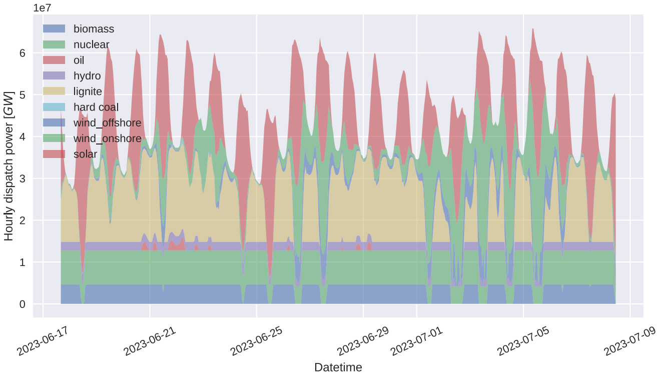

Fig. 7. Exemplary visualization of energy market dispatch from the hourly market simulation of 2023 with NUTS1. Simulated conventional power plants do not provide base load and have higher gradients than usual, due to missing heat modeling

Finally in Fig. 8 the correlation coefficient of all the modeled simulations with the respective NUTS areas are shown. It is evident that the impact of employing different resolutions for modeling renewable energy sources, as implemented in this study, is negligible, as the variations observed are minimal. A fixed pattern of which aggregation provides best results can not be fully estimated, though, due to the reduced complexity, simulations with a higher aggregation or lower resolution can be handled more easily. On the other hand, simulations employing higher resolutions do offer enhanced individual bids within the market, thereby enabling a more comprehensive examination of bidding behavior. In conclusion, it has been demonstrated that the market model is capable of producing satisfactory results in the modeling of the German energy market across all aggregation levels. However, it should be noted that years including the gas crises exhibit a lower correlation factor. Performance metrics of the simulated market to the historical market dispatch for all years and spatial resolutions is given in the Appendix.

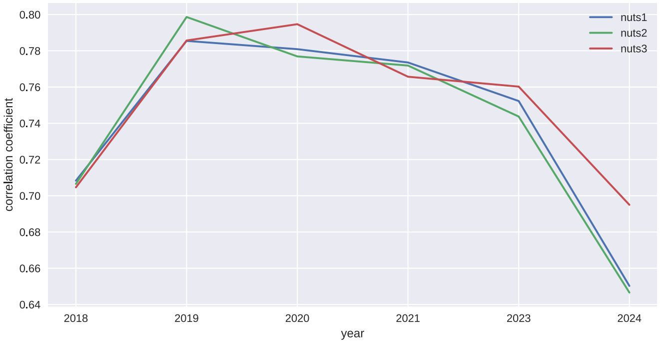

Fig. 8. Summary of correlation coefficient with historical benchmark data set of ENTSO-E between market simulation with different NUTS aggregations. One can see that the focus on correct price modeling and bidding behavior have a higher influence on the results than the aggregation, while all simulations provide good out of the box results

V. CONCLUSION AND FUTURE WORK

This paper showcases a market simulation based on open-access data, making it easy to reproduce results in a scientific context and adapt it to different years using existing data. While the general price structure of the simulation aligns with historical trends, exact price spikes cannot be accurately reproduced. This indicates that the actual dispatch volume differs from the simulated market dispatch. Additionally, negative prices, typically caused by ramping requirements and district heat demand obligations, cannot be accurately simulated. Further improvements regarding storage operation are necessary, as these have a significant effect on storage prices.

It is shown that the spatial resolution between NUTS areas has a minor effect on renewable time series, which only slightly influences the market dispatch. The impact of wind generation improves with higher NUTS resolution, thanks to better data quality, while the resolution of solar generation slightly degrades with higher resolution. This highlights the need for further improvements to available input data, although significant progress has been made in recent years.

In this paper, market agents use simplified bidding strategies. While this reduces model complexity, it limits the ability to integrate price forecasts and reinforcement learning, as shown by Harder et al. [22]. This is also why changes in conventional power plants are not observed when aggregating agents for different NUTS areas.

It is recommended to run simulations with a lower resolution, as local weather fluctuations tend to average out. While this approach provides generally good results, further fine-tuning with historical data is still required to improve accuracy. In conclusion, this methodology offers a straightforward approach for applying simulations to different years, providing reliable results while highlighting areas for further improvement.

REFERENCES

[1] S. Mishra, T. L. Silva, L. Hellemo, S. Jaehnert, L. E. Egner, S. A. Petersen, T. Signer, F. Zimmermann, and C. Bordin, "Agent-based modeling: Insights into consumer behavior, urban dynamics, grid management, and market interactions," *Energy Strategy Reviews*, vol. 57, p. 101613, Jan. 2025.

[2] S. Misconel, R. Leisen, J. Mikurda, F. Zimmermann, C. Fraunholz, W. Fichtner, D. Möst, and C. Weber, "Systematic comparison of high-resolution electricity system modeling approaches focusing on investment, dispatch and generation adequacy," *Renewable and Sustainable Energy Reviews*, vol. 153, p. 111785, Jan. 2022.

[3] A. Bublitz, D. Keles, F. Zimmermann, C. Fraunholz, and W. Fichtner, "A survey on electricity market design: Insights from theory and real-world implementations of capacity remuneration mechanisms," *Energy Economics*, vol. 80, pp. 1059–1078, May 2019.

[4] L. Torralba-Díaz, C. Schimeczek, M. Reeg, G. Savvidis, M. Deissenroth-Uhrig, F. Guthoff, B. Fleischer, and K. Hufendiek, "Identification of the Efficiency Gap by Coupling a Fundamental Electricity Market Model and an Agent-Based Simulation Model," *Energies*, vol. 13, no. 15, p. 3920, Jan. 2020.

[5] F. Maurer, F. Nitsch, J. Kochems, C. Schimeczek, V. Sander, and S. Lehnhoff, "Know Your Tools - A Comparison of Two Open Agent-Based Energy Market Models," in *2024 20th International Conference on the European Energy Market (EEM)*, Jun. 2024, pp. 1–8.

[6] H. C. Gils, T. Pregger, F. Flachsbarth, M. Jentsch, and C. Dierstein, "Comparison of spatially and temporally resolved energy system models with a focus on Germany's future power supply," *Applied Energy*, vol. 255, p. 113889, Dec. 2019.

[7] L. Suganthi and A. A. Samuel, "Energy models for demand forecasting—A review," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 2, pp. 1223–1240, Feb. 2012.

[8] I. Ghalehkhondabi, E. Ardjmand, G. R. Weckman, and W. A. Young, "An overview of energy demand forecasting methods published in 2005–2015," *Energy Syst*, vol. 8, no. 2, pp. 411–447, May 2017.

[9] M. M. Forootan, I. Larki, R. Zahedi, and A. Ahmadi, "Machine Learning and Deep Learning in Energy Systems: A Review," *Sustainability*, vol. 14, no. 8, p. 4832, Jan. 2022.

[10] E. Pla and M. Jiménez Martínez, "Dealing with change: Retraining strategies to improve load forecasting in individual households under Covid-19 restrictions," *Energy Reports*, vol. 9, pp. 82–89, Oct. 2023.

[11] F. Maurer, J. Sejdija, and V. Sander, "Decentralized energy data storages through an Open Energy Database Server," *NFDI4Energy*, Feb. 2024.

[12] C. Kost, P. Müller, and J. S. Schweiger, "Studie: Stromgestehungskosten erneuerbare Energien," *Fraunhofer ISE*, Jul. 2024.

[13] K. S. Anderson, C. W. Hansen, W. F. Holmgren, A. R. Jensen, M. A. Mikofski, and A. Driesse, "Pvlib python: 2023 project update," *Journal of Open Source Software*, vol. 8, no. 92, p. 5994, Dec. 2023.

[14] D. G. Erbs, S. A. Klein, and J. A. Duffie, "Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation," *Solar Energy*, vol. 28, no. 4, pp. 293–302, Jan. 1982.

[15] copernicus, "Climate reanalysis | ECMWF," https://climate.copernicus.eu/climate-reanalysis, 2023.

[16] S. Haas, U. Krien, B. Schachler, S. Bot, V. Zeli, F. Maurer, K. Shivam, F. Witte, S. J. Rasti, Seth, and S. Bosch, "Wind-python/windpowerlib: Update release," Zenodo, Feb. 2024.

[17] Bundesnetzagentur, "Monitoring Report 2022 - Germany," p. 341, Nov. 2022.

[18] F. Maurer, K. K. Miskiw, R. R. Acosta, N. Harder, V. Sander, and S. Lehnhoff, "Market Abstraction of Energy Markets and Policies - Application in an Agent-Based Modeling Toolbox," in *Energy Informatics*, B. N. Jørgensen, L. C. P. Da Silva, and Z. Ma, Eds. Cham: Springer Nature Switzerland, Dec. 2023, vol. 14468, pp. 139–157.

[19] J. Adams, N. Harder, and A. Weidlich, "Do Block Orders Matter? Impact of Regular Block and Linked Orders on Electricity Market Simulation Outcomes," in *2024 20th International Conference on the European Energy Market (EEM)*, Jun. 2024, pp. 1–7.

[20] C. Schimeczek, K. Nienhaus, U. Frey, E. Sperber, S. Sarfarazi, F. Nitsch, J. Kochems, and A. A. E. Ghazi, "AMIRIS: Agent-based Market model for the Investigationof Renewable and Integrated energy Systems," *JOSS*, vol. 8, no. 84, p. 5041, Apr. 2023.

[21] T. Brown, J. Hörsch, and D. Schlachtberger, "PyPSA: Python for Power System Analysis," *JORS*, vol. 6, no. 1, p. 4, Jan. 2018.

[22] N. Harder, R. Qussous, and A. Weidlich, "Fit for purpose: Modeling wholesale electricity markets realistically with multi-agent deep reinforcement learning," *Energy and AI*, vol. 14, p. 100295, Oct. 2023.

APPENDIX

	mean	MAE	RMSE	correlation
Historical 2018	44.66			
2018 nuts1	41.88	10.08	12.69	0.71
2018 nuts2	40.24	10.61	13.20	0.71
2018 nuts3	40.19	10.61	13.18	0.70
Historical 2019	37.83			
2019 nuts1	42.04	7.57	11.01	0.79
2019 nuts2	40.47	6.82	10.34	0.80
2019 nuts3	40.76	7.05	10.52	0.79
Historical 2020	30.37			
2020 nuts1	36.85	9.64	13.28	0.78
2020 nuts2	37.52	10.11	13.71	0.78
2020 nuts3	35.76	8.98	12.67	0.79
Historical 2021	97.28			
2021 nuts1	78.38	30.50	50.72	0.77
2021 nuts2	78.63	30.65	50.87	0.77
2021 nuts3	75.42	31.55	52.69	0.77
Historical 2023	95.78			
2023 nuts1	107.11	24.72	33.49	0.75
2023 nuts2	107.78	25.53	34.25	0.74
2023 nuts3	104.03	23.39	32.30	0.76
Historical 2024	79.74			
2024 nuts1	86.22	23.73	41.39	0.65
2024 nuts2	85.30	23.57	41.37	0.65
2024 nuts3	77.59	21.74	39.78	0.69

TABLE II
RESULTING SIMULATION CHARACTERISTICS IN COMPARISON TO HISTORIC PRICE SERIES FOR DIFFERENT AGGREGATIONS OF GERMAN ENERGY MARKET