The Journal of Open Source Software

NEMSEER: A Python package for downloading and handling historical National Electricity Market forecast data produced by the Australian Energy Market Operator

Abhijith Prakash [1,3¶], **Anna Bruce** [2,3], **and Iain MacGill** [1,3]

1 School of Electrical Engineering and Telecommunications, University of New South Wales, Australia **2** School of Photovoltaics and Renewable Energy Engineering, University of New South Wales, Australia **3** Collaboration on Energy and Environmental Markets (CEEM), University of New South Wales, Australia
¶ Corresponding author

DOI: 10.21105/joss.05883

Software

- Review
- Repository
- Archive

Editor: Arfon Smith

Reviewers:

- @mfleschutz
- @amandadsmith

Submitted: 18 May 2023
Published: 09 December 2023

Summary

In electrical power systems and their associated market frameworks, actions close to or during real-time (i.e. the time of power delivery) may be critical to ensuring that:

1. Supply and demand are balanced, and that the power system is operated within its technical envelope.
2. Market participants can maximise revenues from their generating and/or demand-side resources.

However, *real-time* actions and their outcomes are, to some degree, dependent on decisions made *ahead-of-time* — e.g. starting a gas turbine, charging a battery energy storage system and maintaining a greater level of spare capacity in reserve during periods of system stress. Given physical and financial uncertainties, no ahead-of-time decision is perfect; instead, they are made in light of the best available information, which includes demand, generation and market price forecasts (Mays et al., 2022).

Though the Australian National Electricity Market (NEM) lacks the ahead-of-time market platforms that are present in many restructured electricity industries in Europe and North America (Cramton, 2017; Roques, 2021), market participants provide resource and market offer information to the Australian Energy Market Operator (AEMO). AEMO then uses these submitted data alongside demand and renewable energy generation forecasts to run several centralised ahead-of-time processes (Australian Energy Market Operator, 2020, 2021b, 2022). These processes produce ahead-of-time information, or "forecasts", that market participants can use to inform their operational decision-making[1], and that trigger AEMO to prepare for (or, in the worst case, undertake emergency actions before or during) periods of potential system risk (Australian Energy Market Commission, 2021).

`NEMSEER` is a Python 3 package that facilitates access to and analysis of *historical* ahead-of-time operational information from AEMO-run processes. Specifically, it enables users to download data from these processes, manipulate it using pandas (The pandas development team, 2020; Wes McKinney, 2010) or `xarray` (Hoyer & Hamman, 2017) data structures, and cache the data in Parquet (*Parquet*, 2023) or netCDF (*NetCDF 4*, 2022) formats. Other major dependencies used by `NEMSEER` include `attrs` (Schlawack, 2022), `requests` (*Psf/Requests*, n.d.) and `tqdm` (Costa-Luis et al., 2023).

[1]That is, how they participate in the central dispatch process that is used to clear the gross-pool markets in each region of the NEM.

Statement of need

AEMO publicly releases data from five of its operational ahead-of-time processes:

- 5-minute pre-dispatch
- Pre-dispatch
- Pre-dispatch Projected Assessment of System Adequacy
- Short Term Projected Assessment of System Adequacy
- Medium Term Projected Assessment of System Adequacy

However, significant effort and prerequisite knowledge is required to obtain and process this data for analysis. Firstly, a user must be familiar with how AEMO's data repositories are organised. Secondly, a user must have knowledge of what type of data each ahead-of-time process generates (i.e. the range of tables and columns available), and of each process' lookahead horizon (i.e. for a given time at which the process is *run*, how many periods into the future are *forecasted*?). Finally, a user must download, unzip and clean CSV files before being able to load and handle tables of interest using data analysis tools.

`NEMSEER` solves these issues by:

1. Providing learning resources and references (via the README and a glossary in the documentation) that unpack what each ahead-of-time process does and what data they offer.
2. Making it easier to download and handle this data. `NEMSEER` can inform the user of the date range of available data, which data tables are available and even generate the appropriate range of *run* times for a set of *forecasted* times that a user is interested in. Once a user queries a subset of data, `NEMSEER` will download, unzip and process the CSV files into `pandas` or `xarray` data structures.

Though existing software solutions (e.g. NemSight, ez2view and NEOpoint) can provide access to some of the same data, most lack a programmatic interface useful for deeper analysis and all are proprietary commercial software. Furthermore, `NEMSEER` adds significant value to users interested in deeper analysis through its documentation. It contains examples showing how users can analyse demand forecast errors and energy price convergence using pre-dispatch demand and price forecast data (obtained using `NEMSEER`) and historical *actual* NEM system and market data (obtained using `NEMOSIS`) (Gorman et al., 2018). Figure 1 is an output of one such example.

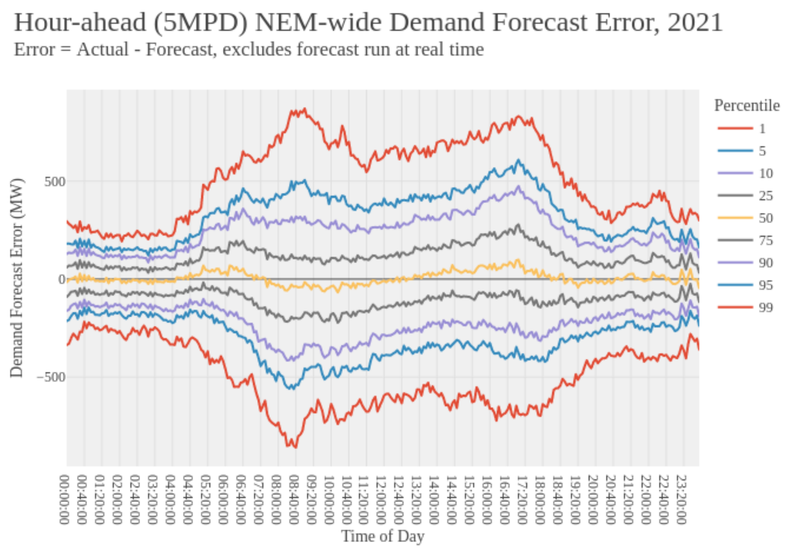

Figure 1: NEM-wide time of day demand error percentiles for 2021 for hour-ahead demand forecasts (i.e. those run during 5-minute pre-dispatch, or 5MPD).

`NEMSEER` use cases include:

- **Modelling system operator or market participant decision-making under uncertainty**. The latter could involve using pre-dispatch market price forecast data to understand the implications of using imperfect information to schedule energy storage systems (ongoing work by Prakash (2022/2022)), or calculating price forecast errors that are used as inputs for stochastic modelling frameworks (e.g. Yurdakul & Billimoria (2022/2023)).
- **Identifying periods of interest for market bidding behaviour analysis**. Significant divergence of the *actual* market price from *forecast* market prices (as explored in the energy price convergence example) might be due to participants changing their market offers as conditions change.
- **Obtaining specific NEM data that is only published in ahead-of-time datasets**. This includes some dynamic risk measures (e.g. capacity that would be lost in a credible contingency) and "sensitivities" that explore changes to market prices and interconnector flows across a range of demand change scenarios in each market region of the NEM (Australian Energy Market Operator, 2021a).

Acknowledgements

We acknowledge contributions from Nicholas Gorman, Declan Heim and Dylan McConnell.

Development of `NEMSEER` was supported by the UNSW Digital Grid Future Institute, and constitutes broader research work supported by an Australian Government Research Training Program Scholarship.

References

Australian Energy Market Commission. (2021). *Reserve services in the National Electricity Market, Directions Paper*. https://www.aemc.gov.au/sites/default/files/2021-01/Reserve%20services%20directions%20paper%20-%205.01.2021%20-%20FINAL.pdf

Australian Energy Market Operator. (2020). *Reliability Standard Implementation Guidelines*. https://aemo.com.au/-/media/files/electricity/nem/planning_and_forecasting/rsig/reliability-standard-implementation-guidelines.pdf?la=en

Australian Energy Market Operator. (2021a). *Pre-Dispatch Sensitivities*. https://www.aemo.com.au/-/media/files/electricity/nem/security_and_reliability/dispatch/policy_and_process/pre-dispatch-sensitivities.pdf

Australian Energy Market Operator. (2021b). *Pre-dispatch operating procedure*. https://www.aemo.com.au/-/media/files/electricity/nem/security_and_reliability/power_system_ops/procedures/so_op_3704-predispatch.pdf?la=en

Australian Energy Market Operator. (2022). *Pre dispatch*. https://aemo.com.au/energy-systems/electricity/national-electricity-market-nem/data-nem/market-management-system-m pre-dispatch

Costa-Luis, C. da, Larroque, S. K., Altendorf, K., Mary, H., richardsheridan, Korobov, M., Raphael, N., Ivanov, I., Bargull, M., Rodrigues, N., Chen, G., Lee, A., Newey, C., CrazyPython, JC, Zugnoni, M., Pagel, M. D., mjstevens777, Dektyarev, M., … Nordlund, M. (2023). *Tqdm: A fast, Extensible Progress Bar for Python and CLI*. Zenodo. https://doi.org/10.5281/zenodo.7697295

Cramton, P. (2017). Electricity market design. *Oxford Review of Economic Policy*, *33*(4), 589–612. https://doi.org/10.1093/oxrep/grx041

Gorman, N., Haghdadi, N., Bruce, A., & MacGill, I. (2018). NEMOSIS – NEM Open Source Information Service; open-source access to Australian National Electricity Market Data. *Asia-Pacific Solar Research Conference*. https://www.researchgate.net/publication/329798805

Hoyer, S., & Hamman, J. (2017). Xarray: N-D labeled arrays and datasets in python. *Journal of Open Research Software*, *5*(1). https://doi.org/10.5334/jors.148

Mays, J., Craig, M. T., Kiesling, L., Macey, J. C., Shaffer, B., & Shu, H. (2022). Private risk and social resilience in liberalized electricity markets. *Joule*, *6*(2), 369–380. https://doi.org/10.1016/j.joule.2022.01.004

NetCDF 4. (2022). Unidata. https://doi.org/10.5065/D6H70CW6

Parquet. (2023). The Apache Software Foundation. https://github.com/apache/parquet-format

Prakash, A. (2022). NEMStorageUnderUncertainty. In *GitHub repository*. GitHub. https://github.com/prakaa/NEMStorageUnderUncertainty (Original work published 2022)

Psf/requests: A simple, yet elegant, HTTP library. (n.d.). GitHub. Retrieved March 14, 2023, from https://github.com/psf/requests

Roques, F. (2021). The evolution of the European model for electricity markets. In *Handbook on Electricity Markets* (pp. 308–330). Edward Elgar Publishing. https://doi.org/10.4337/9781788979955.00017

Schlawack, H. (2022). *Attrs*. Zenodo. https://doi.org/10.5281/zenodo.7467074

The pandas development team. (2020). *Pandas-dev/pandas: Pandas*. Zenodo. https://doi.org/10.5281/zenodo.3509134

Wes McKinney. (2010). Data Structures for Statistical Computing in Python. In S. van der Walt & Jarrod Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 56–61). https://doi.org/10.25080/Majora-92bf1922-00a

Yurdakul, O., & Billimoria, F. (2023). Online Companion to "Risk-Averse Self-Scheduling of Storage in Decentralized Markets". In *GitHub repository*. GitHub. https://github.com/oyurdakul/pesgm23 (Original work published 2022)