Exploring the Impact of Interannual Dynamics on Long-Duration Energy Storage in Energy System Models

Hawkins, Daniel; Van Den Broek, Machteld; Bruninx, Kenneth

DOI
10.1109/EEM64765.2025.11050355

Publication date
2025

Document Version
Final published version

Published in
2025 21st International Conference on the European Energy Market, EEM 2025

Exploring the impact of interannual dynamics on long-duration energy storage in energy system models.

Daniel Hawkins, Machteld van den Broek & Kenneth Bruninx
Faculty of Technology, Policy, and Management
Technische Universiteit Delft
Delft, The Netherlands
d.d.a.hawkins@tudelft.nl

2025 21st International Conference on the European Energy Market (EEM) | 979-8-3315-1278-1/25/$31.00 ©2025 IEEE | DOI: 10.1109/EEM64765.2025.11050355

***Abstract*—Due to computational limits, temporal details within Energy System Optimisation Models are often reduced, for example by reducing the time horizon or by resampling via Time Series Aggregation (TSA) techniques. In high RES energy systems, this may lead to undersizing of Long-Duration Energy Storage (LDES) capacities, necessary for system flexibility, due to the omission of long-term interannual weather effects. Via comparative analysis between the capacity expansion results for different subsets of weather years, this paper shows the extent to which single year models underpredict LDES. but also that a small cluster (n=2, 3) of weather years can adequately capture key system-defining weather patterns. Identifying these weather years ex-ante is non-trivial, as there is no obvious correlation with how well they describe the full set of weather years. As this assumed correlation underpins current time series aggregation techniques, new techniques are required. *Index Terms*-—Energy system optimisation model, long-duration energy storage, time series aggregation**

I. Introduction

In seeking carbon neutrality by 2050, some EU member states have set targets for 100% of electricity supply from renewable energy sources (RESs), which raises concerns over the weather-dependence of some of these technologies. Weather patterns exhibit both short- and long- term dynamics, ranging from predictable cycles (e.g., day-night cycles, seasons) to seemingly random emergent effects across longer time scales. Long-Duration Energy Storage (LDES, which refers here to technologies capable of storing energy in the TWh-range for days to years) is expected to play a critical role within a future climate-neutral power system, providing the flexibility to respond to seasonal and interannual fluctuations in renewable generation, and ensuring system adequacy [1].

Whereas several storage representations in energy system optimisation models (ESOMs) have been proposed [2], [3], [4], [5], centring on the value of mitigating intra-annual seasonal fluctuations, recent research [6], [7], [8] draws attention to the risks around extreme, non-cyclical weather patterns, and those beyond seasonal timescales.

Despite growing awareness of these longer-term dynamics, contemporary ESOMs typically do not consider multiple consecutive weather years due to concerns around computational tractability. Modelers make trade-offs between a variety of decision variables, interactions, and constraints. As modelers seek to include details pertinent to their research question, temporal details are often reduced. This is particularly detrimental for the modelling of LDES whose system purpose and economic value centre around arbitrage over time. Several researchers [7], [9], [10] have proposed to explore multi-year models but report computational challenges, necessitating model simplifications.

Modellers can reduce the temporal details of their model, and as such limit the computation burden, via ad-hoc modelling choices or formal time series aggregation techniques. In the first category, typical modelling decisions include the time horizon, the resolution of time series data (e.g., hourly or daily), and the number of weather years to include in a model with a single, representative weather-year being typical. In the second category, researchers have explored more advanced Time Series Aggregation (TSA) techniques that involve the selection of representative hours/days/weeks to capture the dynamics of the original dataset. Application of TSAs relies on the assumption that an ex-ante error metric can be defined for the timeseries that correlates well with the ex-post error of the planning model's (capacity expansion) results. Identifying an appropriate metric is an ongoing challenge [11]. Some alternative approaches, for example [9], have explored ex-post (i.e., non-TSA) techniques, seeking to infer the outcomes of long-term models from the results of single year models, but interannual dynamics remain excluded.

Although TSAs can be applied across multiple weather years ex-ante, neither these nor the ex-post approaches have been designed to capture interannual effects, i.e., the impact of one weather-year on the next, nor how sequential extremes can trigger cascading impacts. An important question remains outstanding of how to select the weather years that capture these long-term dynamics for the proper sizing of LDES whilst keeping a model computationally tractable. Novel ESOM formulations and/or TSA techniques are needed that include temporal details relevant to the short- and long- term

This work forms part of the HyTROS project, funded out of the Dutch Group Fund Program, GroenvermogenNL (groenvermogennl.org/), an R&D consortium of academic and industrial organisations. The contents of this publication reflect only the views of the authors and not necessarily those of wider HyTROS participants.

requirements of LDES. To better understand how ESOM formulations and TSAs come together, further research is needed into their co-interactions in this context. This research takes a first step in developing such formulations, exploring variations in weather-year input data. We optimise, evaluate, and compare system costs and capacity mixes for different subsets of weather-year data with the ambition of identifying the balance between ESOM detail and tractability. The contribution of this paper is twofold:

1. In a case study, we show that single weather year models underestimate LDES capacity and provide a sub-optimal capacity mix, regardless of the choice of weather year. Clustering of weather years allows the capture of system defining events that may span multiple weather years whilst limiting computational burden.
2. We show that the traditional resampling assumption based on the correlation between the ex-ante timeseries error and ex-post planning model results breaks down when considering long-term dynamics and LDES.

This paper is a starting point for new formulations and timeseries sampling techniques that consider interannual weather dynamics and the associated impact on LDES capacities within a climate-neutral energy system.

The remainder of this paper is organized as follows. In Section II we present our methodology, discussing the ESOM formulation, approach to weather year clustering, error metrics and study limitations. Section III details the case study, including economic and operational parameters. Section IV presents and discusses modelling results whilst Section V concludes with the key findings.

II. Methodology

To investigate the influence of weather-years on LDES capacity within an energy system, we construct an ESOM using the Calliope modelling framework [12] (Section II.A). We compare the results of ESOM runs for different subsets of 1-3 weather-years against a reference model (Section II.B). To enforce a fair comparison between runs, the model configuration is fixed across runs except for timeseries parameters (RES capacity factors, demand) which vary according to the input weather-year(s). Errors on the capacity mix and cost are used as performance indicators (Section II.C). All models, data and results including references to sources are available via GitHub: (Link to Repository).

A. ESOM formulation

The planning model assumes perfect foresight and minimizes total system cost. The decision variables are technology capacities and generation levels. We consider a single-node, electricity-only system for the sake of simplicity. Cost types considered vary between technologies but generally include annualized CAPEX, fixed OPEX, and variable OPEX, as detailed in the GitHub repository. Alongside capacity limits, several standard ESOM constraints are applied consistently across all models in this study. A minimum uptime condition is enforced for the Nuclear generation (only) to emulate typical operating conditions [13].

Energy balance is enforced for each hour across the time horizon such that final demand equals supply, with supply equalling the output of generation technologies, less curtailment, plus storage discharge for that hour. Technology capacities, power and energy, are decision variables that must be within lower and upper limits. Unmet demand is not permitted.

With respect to storage constraints. State of charge (SOC) at each timestep is equal to the sum of SOC, charge, and discharge of the previous timestep. Cyclical storage conditions are enforced across the time horizon such that SOC for the first and last hours of the timeseries are contiguous. Standby storage losses are omitted in this model but efficiencies for charging, discharging, and conversion are included.

Given this study's interest in the comparative findings across different subsets of weather-year inputs, temporal variables are the priority. Computational feasibility issues require trade-offs across other model details, broadly: candidate technologies, geographies (nodes and the networks connecting them), operational constraints and the treatment of uncertainty. The simplifications considered in this paper are deemed adequate for this study given the focus on comparative results, i.e. we focus on the error between test models and a reference model.

B. Weather year aggregation

Three types of instances of the ESOM are solved, differing only across the structure of the input timeseries, with all other modelling parameters and constraints fixed.

The first model represents a reference case against which other models are assessed. This full horizon case considers all available weather years (in this study, 2010-2019), a period which would typically be challenging to solve in more complex ESOMs and here requires simplification to resolve as described above.

The second category of models is concerned with single weather-years. In exploring the adequacy of single-year models, we compare the results for capacity mix and annualized costs of individual year models against the reference (full horizon) model.

The third category of models explores multiple weather-years, seeking to confirm whether clusters of weather-years can accurately reflect the results of the reference model. Two- and three-year clusters were evaluated.

Two-year clusters are formed via the concatenation of individual weather years. Both contiguous (e.g, [2012, 2013]) and non-contiguous clusters (e.g., [2012, 2020]) are evaluated. Weighting factors are applied to enforce the notion that one acts as a system defining 'bad' weather-year (weight: 1) whilst the other represents a typical operational year (weight: $N-1$), such that the sum of weights equals the number of weather years in the original dataset, N. All combinations of coupled weather-years in the original dataset are considered.

In the case of three-year clusters, clusters are formed by appending a representative operational weather-year (weight: $N-2$) onto a coupling of system-defining weather years (weights, for each: 1). Given the large number of possible permutations of three weather-years (e.g., assuming a data set of 10 weather years, $^{10}P3 = 720$), a selection was made to reduce the number of tested instances. System defining weather-year couples were selected as having been contiguous in the historical dataset (e.g. [2010, 2011] and [2015, 2016] but not [2012, 2018]). For each weather-year couple in the original data sets, models were run for all individual prepended weather-years not already included in that couple (e.g., [2014; 2017, 2018] was included, but not [2012; 2012, 2013]). This selection strategy retains the interannual dynamics that occurred historically across the coupled weather-years.

C. Error metrics

Several error statistics are used to evaluate the efficacy of single and clustered weather-year models at reproducing the results of the reference model.

The Cost Error (CE), E_C, is the relative difference between the annualized costs of the tested model and the reference:

$$E_C = \frac{C_{ref} - C_{model}}{C_{ref}} \quad (1)$$

Where C_m is the annualised total system cost in instance, m.

Capacity error for a given technology, E_K^x, e.g. LDES Capacity Error, is the difference between the capacity of the tested model and that of the reference:

$$E_K^x = \frac{K_{ref}^x - K_{model}^x}{K_{model}^x} \quad (2)$$

Where K_m^x is the installed capacity of technology, x, for model, m.

The Mean Absolute Capacity Mix Error (MACME), $\overline{E_K}$, is the mean, across all investable technologies, of the absolute difference between the capacities of the tested model and the reference model:

$$\overline{E_K} = \frac{1}{N^{techs}} \sum_{x}^{techs} |E_K^x| \quad (3)$$

Where N^{techs} is the number of technologies for averaging purposes. Note that installed capacity refers to power capacity for generation and conversion technologies, and storage capacity for storage technologies.

III. Case Study

Our case study is inspired by the Netherlands. The Netherlands presents an interesting case study where the adequacy of a 100% RES system presents a major challenge. The Netherlands is investing heavily into its offshore system [16], due to land space constraints for onshore wind and solar, subjecting the future energy system to highly intermittent sources with no obvious availability cycles. The lack of opportunity for pumped hydro forces a need for alternative large-scale energy storage solutions, with current interest centred around hydrogen storage in salt caverns [14]. Generation technologies available to the planning model for investment with relevant economic and performance parameters are described in Table I.

TABLE I. ECONOMIC AND PERFORMANCE PARAMETERS FOR INVESTABLE TECHNOLOGIES.

Economic Parameter	**Offshore Wind**	**Onshore Wind**	**Solar (Utility Scale)**	**Nuclear**	**Hydrogen CCGT**	**Li-Ion Battery (Grid Scale)**	**Electrolyser (Grid Scale)**	**Hydrogen Salt Cavern**
Lower Capacity Limit (GW)	2.5	4.2	-	1	-	-	-	-
Upper Capacity Limit (GW)	110	16	-	3.3	-	-	-	-
Fixed Capital Cost, € / kW	2,100	1,100	560	6,900	940	1,100 (€ / kWh)	1,200	3.2 (€ / kWh)
Annual O&M Cost, €M / kW / year	50	17	11	17	31	-	24	
Variable O&M Cost, € / kWh	5.0e-3	2.0e-3	-	2.7e-3	4.7e-3	2.1e-4	-	-
Variable Fuel Cost, € / kWh	-	-	-	4.4e-3	-	-	-	-
Lifetime, years	27	27	35	40	25	20	25	100
Efficiency	-	-	-	-	56%	Charge: 97% Discharge: 98%	65%	Injection: 99% Withdrawal: 99%

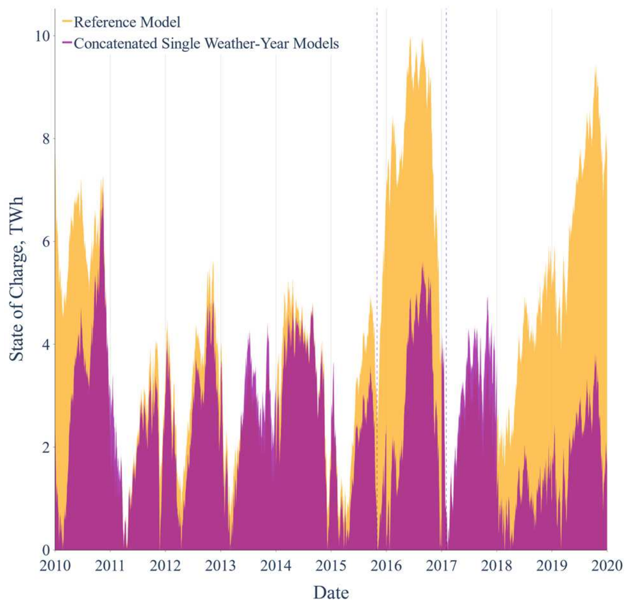

Figure 1. SOC for hydrogen storage over time for the reference model and concatenated single weather year models

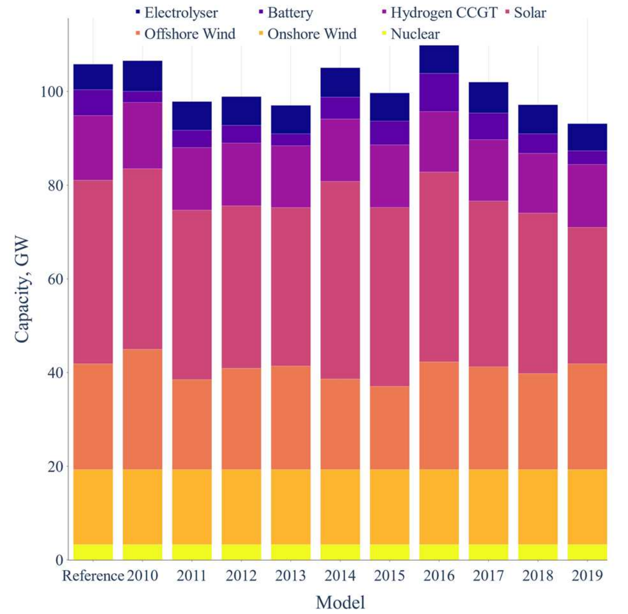

Figure 2. Capacity mix for the reference model and single weather year models

The country anticipates limited investment in nuclear power plant capacity (3.3 GW by 2032) [17], which we enforce in the generation mix.

Two storage technology systems are available for investment: batteries and hydrogen storage. Battery technology is intended to provide short-duration flexibility services on the order of hours. No capacity limits are enforced, however a power-to-energy ratio of 3 is applied [18]. The hydrogen storage system is comprised of separate investments into electrolysis capacity (power; charge), salt cavern capacity (energy; storage), and a hydrogen-fired CCGT (power; discharge). Disaggregation of these functions relieves the model of any assumptions with respect to power to energy ratios which do not apply for large-scale storage of hydrogen. No capacity limits are enforced.

Historic hourly timeseries data for solar, offshore wind and onshore wind for the period 2010-2019 was sourced from the MERRA-2 datasets made available by Renewables Ninja [19], whilst historic demand for the same period was sourced directly from the ENTSO-E Transparency platform [20].

IV. RESULTS

A. Single weather year models underestimate LDES capacity

Fig. 1 shows the SOC of the LDES over time for the reference model and single weather year models, concatenated chronologically. The hydrogen storage's SOC varies substantially across the time horizon as the cavern and hydrogen CCGT dispatch to meet insufficient renewable generation. Maximal peaks in the SOC chart correspond with installed storage capacities and are therefore considered the system-defining event in each case. The results demonstrate the scale of underinvestment across all single-year model instances by comparison with the reference, with the single year model for 2010 exhibiting the largest capacity investment at 70% (7 TWh) of the reference model's capacity (10 TWh).

Note that the reference model's system defining event extends from late-2015 to early-2017, spanning two winters as indicated in Fig. 1, but none of the single year models for this period alone lead to adequate investment into LDES capacity. In the single year models, we observe multiple periods of rapid withdrawal from the LDES system suggesting sequential periods of high demand and low renewable output. Of further consideration is that LDES is charged using energy from RESs, and therefore the extent of its energy capacity for responding to demand, depends on renewable output in the preceding periods. This supports the notion that interannual effects substantially impact LDES capacity requirements, which is not otherwise captured by single year models that have no relationship with adjacent weather years.

Designing a system to accommodate longer-term dynamics, extends beyond just LDES capacity and impacts the broader capacity mix, as highlighted in Fig. 2. The largest fluctuations occur across Offshore Wind and Solar, whilst Hydrogen CCGT and Electrolysis capacities are largely stable across models despite a lack of capacity limits. Whereas capacity factors for solar and wind technologies fluctuate with changing weather, influencing the economics of their capacity investments, Electrolysis and Hydrogen CCGT capacities correlate with demand during periods of mutually low wind and low solar generation. There is always some moment in a year where both the sun and wind are absent, and this defines the investment into the power capacities of the hydrogen pathway. Note that onshore wind and nuclear always hit their capacity limits.

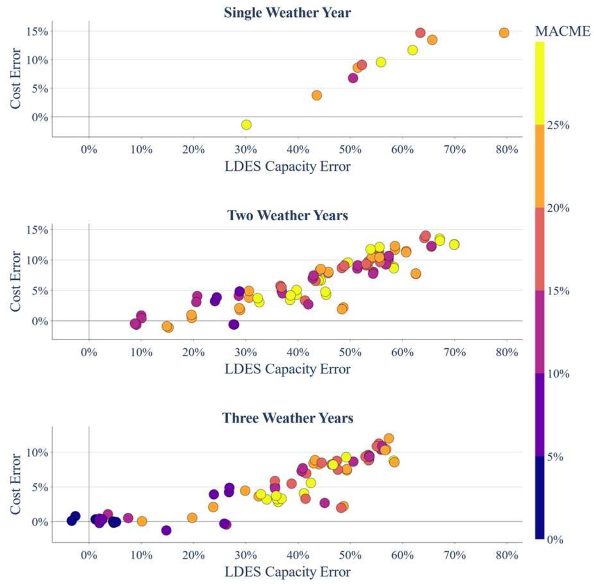

Figure 3. Distribution of LDES capacity error vs. Cost Error, with MACME (colour scale), for models of different cluster sizes.

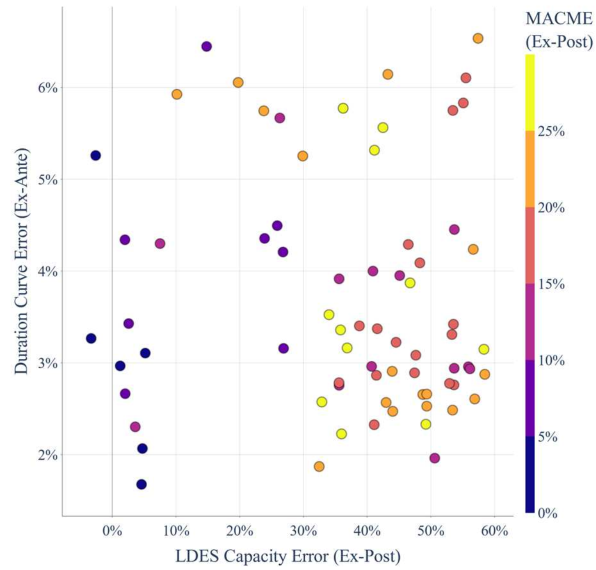

Figure 4. LDES capacity (ex-post) error vs. duration curve (ex-ante) error, with MACME (colour scale), for three-year instances.

B. *Weather year clustering*

In seeking ways to capture these interannual effects without modelling the full horizon, we investigate clusters of two and three weather year inputs, with results shown in Fig. 3. These results demonstrate the existence of weather year combinations that adequately reproduce the full horizon results. LDES capacity errors for three-year clusters range from -3% to 58%, demonstrating that increasing cluster size does not guarantee a reduction in error, and therefore the importance of selecting the right cluster of weather-years.

We find that 9 three-year clusters achieve a CE, MACME, and LDES capacity error all within 10%. In all cases, these clusters include 2016 as a system defining weather-year and one of 2015 or 2017, suggesting a strong correlation with the reference model's system defining event.

C. *Timeseries error versus planning model error*

We investigate the potential of using ex-ante (pre-optimisation) timeseries error to predict ex-post (post-optimisation) error for any given cluster, as is typical when applying TSA techniques. The assumption is that by minimising a "timeseries error" before solving the planning problem, we can minimise errors within the planning results. We use techniques described in [2] which assesses timeseries inputs by comparing the duration curves for time-varying parameters of the cluster (i.e., demand, and RES capacity factors) with the reference. We only show results for three-year clusters. As shown in Fig. 4, and in contrast with the single-year analyses of [2], [11], no obvious correlation exists between ex-ante timeseries error and the ex-post planning error (LDES capacity error and MACME). Whilst the [2013; 2016, 2017] model with the lowest ex-ante error, produces a reasonable LDES capacity error (5%), it does not identify the best performing model. Furthermore, the [2013; 2010, 2011] model with the second lowest ex-ante error, produces a high LDES capacity error (32%) and high MACME (25%). This seemingly erratic relationship between ex-ante and ex-post errors, suggests that the assumptions applied in the selection of representative samples by TSAs, do not hold up consistently for multi-year clusters. As with the capacity results, this may be attributable to the lack of consideration for long-term dynamics and the compounding impact of consecutive events. Existing single-year models and resampling techniques (i.e. TSAs) focus on absolute values, independent of chronology, and therefore do not consider long-term, interannual dynamics. This is relevant here where we focus on LDES, charged via RES, and where traditional dispatchable capacity is unavailable.

V. Conclusion

Our results demonstrate that single weather year models undersize LDES in a carbon-neutral energy system as they do not capture interannual dynamics. An appropriate two- or three-year cluster can reduce error but the ex-ante selection of weather years for inclusion remains an open challenge. Development of an intelligent weather year selection method, that considers the compounded impact of system defining events is key to improving the representation of LDES within ESOMs. Future research should explore how TSAs, applied to multi-year datasets, can be reformulated and/or constrained to ensure they retain the interannual dynamics necessary for the proper sizing of LDES capacity.

This study focuses on the comparative findings between different subsets of weather years with a full horizon reference model. Tractability of the reference model required simplification of the formulation and research should explore whether the errors for this simplified study are a good proxy for errors within a more complex ESOM instances.

REFERENCES

[1] T. Levin *et al.*, 'Energy storage solutions to decarbonize electricity through enhanced capacity expansion modelling', *Nat Energy*, vol. 8, no. 11, pp. 1199–1208, Sep. 2023, doi: 10.1038/s41560-023-01340-6.

[2] S. Gonzato, K. Bruninx, and E. Delarue, 'Long term storage in generation expansion planning models with a reduced temporal scope', *Applied Energy*, vol. 298, p. 117168, Sep. 2021, doi: 10.1016/j.apenergy.2021.117168.

[3] L. Kotzur, P. Markewitz, M. Robinius, and D. Stolten, 'Time series aggregation for energy system design: Modeling seasonal storage', *Applied Energy*, vol. 213, pp. 123–135, Mar. 2018, doi: 10.1016/j.apenergy.2018.01.023.

[4] T. Blanke, K. S. Schmidt, J. Göttsche, B. Döring, J. Frisch, and C. Van Treeck, 'Time series aggregation for energy system design: review and extension of modelling seasonal storages', *Energy Inform*, vol. 5, no. S1, p. 17, Sep. 2022, doi: 10.1186/s42162-022-00208-5.

[5] A. F. Castelli, L. Moretti, G. Manzolini, and E. Martelli, 'Robust optimization of seasonal, day-ahead and real time operation of aggregated energy systems', *International Journal of Electrical Power & Energy Systems*, vol. 152, p. 109190, Oct. 2023, doi: 10.1016/j.ijepes.2023.109190.

[6] A. Grochowicz, K. van Greevenbroek, and H. C. Bloomfield, 'Using power system modelling outputs to identify weather-induced extreme events in highly renewable systems', *Environ. Res. Lett.*, vol. 19, no. 5, p. 054038, Apr. 2024, doi: 10.1088/1748-9326/ad374a.

[7] O. Ruhnau and S. Qvist, 'Storage requirements in a 100% renewable electricity system: extreme events and inter-annual variability', *Environ. Res. Lett.*, vol. 17, no. 4, p. 044018, Mar. 2022, doi: 10.1088/1748-9326/ac4dc8.

[8] T. H. Ruggles *et al.*, 'Planning reliable wind- and solar-based electricity systems', *Advances in Applied Energy*, vol. 15, p. 100185, Sep. 2024, doi: 10.1016/j.adapen.2024.100185.

[9] A. Grochowicz, K. van Greevenbroek, F. E. Benth, and M. Zeyringer, 'Intersecting near-optimal spaces: European power systems with more resilience to weather variability', *Energy Economics*, vol. 118, p. 106496, Feb. 2023, doi: 10.1016/j.eneco.2022.106496.

[10] Z. Lei *et al.*, 'A compact time horizon compression method for planning community integrated energy systems with long-term energy storage', *Applied Energy*, vol. 361, p. 122912, May 2024, doi: 10.1016/j.apenergy.2024.122912.

[11] K. Poncelet, H. Höschle, E. Delarue, A. Virag, and W. D'haeseleer, 'Selecting Representative Days for Capturing the Implications of Integrating Intermittent Renewables in Generation Expansion Planning Problems', *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 1936–1948, May 2017, doi: 10.1109/TPWRS.2016.2596803.

[12] S. Pfenninger and B. Pickering, 'Calliope: a multi-scale energy systems modelling framework', *JOSS*, vol. 3, no. 29, p. 825, Sep. 2018, doi: 10.21105/joss.00825.

[13] IAEA, 'Nuclear Power Reactors in the World', IAEA, Vienna, Austria, 44, Jul. 2024. Accessed: Jan. 28, 2025. [Online]. Available: https://www-pub.iaea.org/MTCD/Publications/PDF/p15748-RDS-2-44_web.pdf

[14] TNO, 'LARGE-SCALE ENERGY STORAGE IN SALT CAVERNS AND DEPLETED FIELDS'. Accessed: Oct. 08, 2024. [Online]. Available: https://publications.tno.nl/publication/34637700/8sBxDu/TNO-2020-R12006.pdf

[15] J. Johnston, R. Henriquez-Auba, B. Maluenda, and M. Fripp, 'Switch 2.0: A modern platform for planning high-renewable power systems', *SoftwareX*, vol. 10, p. 100251, Jul. 2019, doi: 10.1016/j.softx.2019.100251.

[16] T. Macquart, D. Kucukbahar, and B. Prinsen, 'Dutch Offshore Wind Market Report', Netherlands Enterprise Agency (RVO), Apr. 2023. [Online]. Available: https://www.rvo.nl/sites/default/files/2023-04/Offshore-Wind-Market-Report.pdf

[17] Ministerie van Algemene Zaken, 'Kernenergie - Duurzame energie - Rijksoverheid.nl'. Accessed: Jan. 28, 2025. [Online]. Available: https://www.rijksoverheid.nl/onderwerpen/duurzame-energie/kernenergie

[18] 'Technology Data Catalogue for Energy Storage'. Danish Energy Agency, Oct. 2018. Accessed: Jan. 28, 2025. [Online]. Available: https://ens.dk/en/analyses-and-statistics/technology-data-energy-storage

[19] S. Pfenninger and I. Staffell, 'Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data', *Energy*, vol. 114, pp. 1251–1265, Nov. 2016, doi: 10.1016/j.energy.2016.08.060.

[20] 'ENTSO-E Transparency Platform', ENTSO-E Transparency Platform. Accessed: Dec. 19, 2025. [Online]. Available: https://transparency.entsoe.eu/