26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

Towards the generalised compression of weather data for technoeconomic optimisations

David L. Green [a]

[a]*Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia*
Email: d.green@csiro.au

Abstract: Capacity planning models for future energy systems are frequently bottlenecked by the temporal resolution of variable renewable energy (VRE) traces. Prior work has mostly pursued *a priori* time-series aggregation with mixed success. We explore an alternative *a posteriori* pathway: solve the full model and use primal/dual information to rank hours by their contribution to capacity decisions, then retain only the most informative hours while preserving inter-hour dynamics via non-uniform time steps. We compare *a priori* schemes (time step coarsening and horizon truncation) against *a posteriori* hour-selection based on constraint duals and temporal block sensitivity. Across six test cases (solar-dominant, wind-dominant, balanced; years 2017 and 2019) we quantify compression factor versus accuracy trade-offs. We find that the particular a posteriori methods applied here lacked sufficient robustness for reliable compression, likely due to capacity vector degeneracy and discretisation error neglect. However, simple time step coarsening consistently outperforms horizon truncation, with dt=2 (two-hourly time resolution) providing reliable compression with less than 1% error.

Keywords: *Energy system modelling, Time-series compression, Capacity expansion, Variable renewable energy, A posteriori selection*

1 INTRODUCTION

Capacity expansion and planning models must resolve long horizons with high spatial and technological detail. Hourly VRE traces drive storage state transitions and peaking requirements, but a full-year, full-resolution run can be prohibitive once many regions and technologies are included. Common practice aggregates or selects time periods *a priori* to construct representative days or weeks that reduce the model horizon (the number of hours simulated) while attempting to preserve accuracy. While effective on some instances, such methods can distort capacity decisions when they miss rare-but-binding conditions that vary by location. Here we investigate whether an *a posteriori* approach—ranking hours by their influence on capacity decisions using dual or sensitivity information—can produce VRE trace compressions with useful accuracy properties. For this work we define 'compression factor' as the ratio of the number of native hours simulated (i.e., hourly time resolution for a year) to the number of informative hours retained - to which the runtime of the resulting linear program is strongly dependent on. We explore three *a posteriori* methods: ranking hours by constraint duals, temporal block sensitivity, and a hybrid approach combining both. For the particular methods we explore, our results reveal that while theoretically appealing, such methods require careful implementation in practice, leading us to recommend, for now, simple *a priori* baselines (Time-step coarsening over Horizon shortening) for reliable compression.

A posteriori compression by itself has limited operational utility because it requires a full-resolution solution to compute hour importance. The approach was initially conceived as a labelling mechanism: from each solved instance we would extract the subset of informative hours and pair it with the input traces to form training data. This would enable learning an *a priori* predictor of informative hours from trace features, potentially generalising across regions, years, and technology mixes. However, as our results demonstrate, there is considerable future work in first producing a robust *a posteriori* method that consistently performs well across different scenarios and datasets before such training can be attmpted.

A priori temporal reduction is commonly done by selecting representative periods (days or weeks) or by clustering hours into a small set of typical profiles and repeating them across the year. The "Carpe diem" approach selects representative days that jointly capture variability across multiple regions and VRE technologies (Nahmmacher et al., 2016). A broad comparison of clustering methods (e.g., k-means, k-medoids, hierarchical) shows that the chosen representation can materially bias designs, with storage and ramping particularly sensitive to chronological distortions (Teichgraeber and Brandt, 2019). To mitigate the loss of long-term chronology for storage, inter-/intra-period inventory decompositions have been proposed so that seasonal storage remains feasible when using typical periods (Kotzur et al., 2018). Still, a priori schemes struggle with rare-but-binding hours (e.g., extreme residual-load or low-VRE events); dedicated treatments that embed extremes into the aggregated set improve robustness but add design choices (Teichgraeber et al., 2020). Over multi-decade datasets, method performance can vary substantially with input traces and constraints; a comparison of downsampling, clustering, and heuristics across 25 years of wind/PV data found no one-size-fits-all approach and highlighted the planning importance of inter-annual variability (Pfenninger, 2017).

An alternative is a posteriori aggregation: run the full-resolution model (or a high-resolution proxy) and derive which hours matter most for decisions, then compress accordingly while preserving chronology. Recent work formalises such a posteriori schemes for models with storage and shows accuracy gains over a priori approaches at comparable compression (Hilbers et al., 2023). In brief, hours are scored by storage scarcity signals and selected while maintaining chronological connections via effective-step weights, improving fidelity of storage dynamics.

We evaluate whether dual-derived metrics (generation capacity linkages, storage bounds/rates, nodal scarcity, residual load and ramps) can reliably predict informative hours across solar-dominated, wind-dominated, and balanced scenarios for two weather years. We also demonstrate the simple application of coarser time step (`dt` > 1) and model horizon truncation. Our analysis reveals that while a posteriori methods occasionally show promise, our particular implementation lacks the robustness needed for reliable compression, whereas simple time-step coarsening provides consistent and predictable performance improvements.

2 MODEL SUMMARY

We use a standard single-region capacity expansion model following prior CSIRO work (Green et al., 2024). Decision variables include generation capacities, storage energy and power capacities, hourly generation, storage charge/discharge, and state-of-charge. The model minimises annualised system cost subject to generator limits, storage power/energy limits, and hourly demand balance. Non-uniform inter-hour spacing is supported via an effective step dt_e between consecutive retained hours. Other feaures of the model include: objective components cover capital, fixed/variable O&M, fuel, and emissions costs aggregated with dt_e; storage state-of-charge transitions propagate across retained hours using dt_e, with a cyclic wrap around end hours; hour subsets are defined by an index set of retained hours as determined by either *a priori* Time-step/Horizon-shortening or *a posteriori* rank-based selection using dual/neutralisation metrics.

3 HOUR-RANKING METRICS AND SELECTION

We (*a posteriori*) compute per-hour, normalised scores indicating the potential influence of each hour on capacity decisions, based on primal and dual outputs of the solved model. All compression methods in this paper operate by dropping hours from the full year and retaining the remaining hours unchanged; the retained set is chosen either a priori (regular Time-step coarsening or Horizon shortening) or a posteriori (top-K hours by a metric), and non-uniform dt_e preserves chronology and energy accounting. The dual-based ranking metrics we examined included (all per-hour): i) Generation-capacity linkages: sums of absolute duals on thermal generation limits and on VRE generation limits (weighted by capacity factor); ii) Storage energy (SOC) bounds and power-rate limits: sums of absolute duals on SOC upper bounds and charge/discharge rate limits; iii) Nodal scarcity and price $\times$ load: max and load-weighted sums of absolute duals on demand-supply balance constraints; iv) Residual load and ramps: residual load is net demand after non-dispatchable VRE, and ramps are the absolute hour-to-hour change in residual load; these proxy scarcity and flexibility needs. Of these, only two yielded objective errors below 10% consistently across cases: storage SOC scarcity and VRE capacity scarcity (capacity-factor weighted). These are the only dual metrics we present in figures.

The types of hourly subset selection methods we consider are time-step coarsening, horizon shortening, dual-based, and time-block neutralisation. The specific instances of each type are as follows: **Time-step coarsening:** keep every dt-th native hour with uniform weights; chronology is preserved at coarser resolution. **Horizon shortening:** truncate the native horizon to some fraction of a year < 8760, running the model on an initial segment only. **Storage SOC scarcity (dual):** the absolute Lagrange multiplier on the storage state-of-charge (SOC) upper bound at each hour (summed across storage assets, if applicable). High values indicate hours when the SOC ceiling binds and additional energy capacity would relax the system, flagging times that are influential for storage sizing. **VRE capacity scarcity (capacity-factor weighted dual):** the absolute Lagrange multiplier on the hourly VRE generation limit—multiplied by that hour's capacity factor, summed across VRE technologies. High values identify hours where permitting more VRE capacity would most relieve scarcity, emphasising hours when VRE can actually produce. **Time-block neutralisation:** for each fixed-length time block (1, 2, or 4 hours), we replace the block's portion of the VRE traces with a neutralised version - defined as a three-month rolling, per-clock-hour median; the degradation in capacity accuracy defines the block's importance and induces an hour ranking.

All metrics are normalised to [0,1] across the year. Hours are ranked by metric value (descending). For a desired compression factor c (where $c = 1$ represents no compression and $c = 2$ represents 50% reduction in temporal resolution), we keep the top $8760/c$ ranked hours plus endpoints to maintain SOC propagation.

4 EXPERIMENTAL DESIGN

We consider six single-region cases: solar-dominant, wind-dominant, and balanced resource profiles for weather years 2017 and 2019. "Solar-dominant" ("wind-dominant") indicates PV (wind) capacity factors that shape baseline designs toward PV (wind), while "balanced" has comparable contributions from both resources under the baseline. Technologies include utility PV, onshore wind, and battery storage; up to 10% annual energy from gas (open-cycle gas turbine) is permitted. We evaluate each base case on a grid of compression factors with baselines (Time-step and Horizon) and *a posteriori* selection using SOC-dual, VRE-capacity-dual, and Neutralisation. A compression factor c, for *a priori* methods is simply the 8760 divided by the number of time steps retained, while for *a posteriori* methods, it defines the top $8760/c$ most important hours by metric (plus endpoints) and compute effective steps dt_e from gaps between retained hours. We instantiate the six scenarios as defined above (balanced, solar-dominant, wind-dominant; years 2017 and 2019).

VRE traces are obtained using Renewables.ninja methodologies for PV (Pfenninger and Staffell, 2016) and wind (Staffell and Pfenninger, 2016).

After runs, we compute error-versus-compression curves in terms of two distances: (i) capacity distance, defined as the L2 norm of the vector of relative differences between the capacity and that of the baseline by

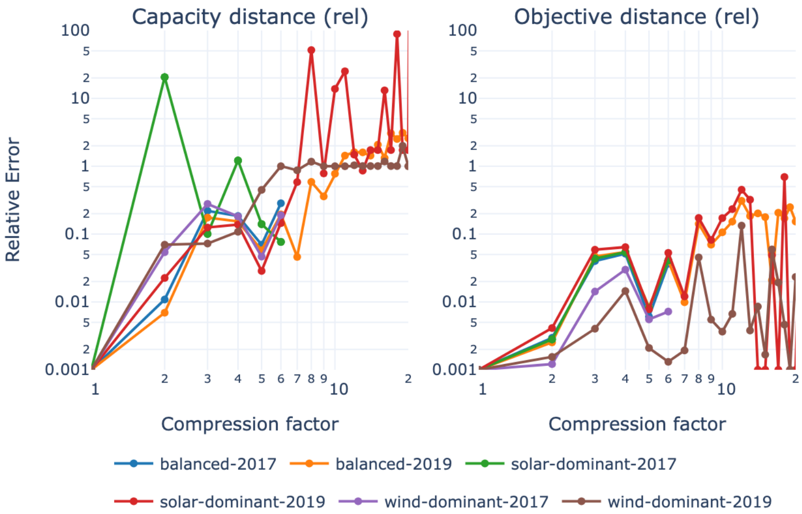

Figure 1: Baseline a priori coarsening via Time-step across six cases. Left: capacity distance; right: objective distance. Log scales on both axes.

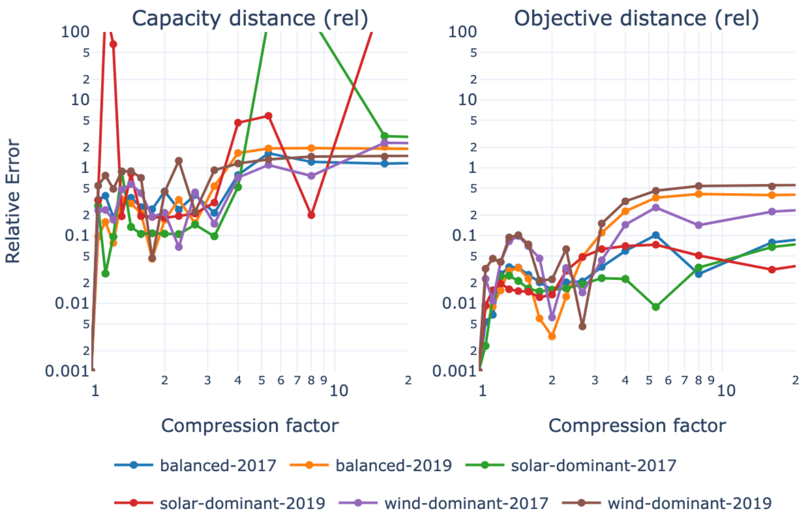

Figure 2: Baseline a priori truncation via Horizon shortening across six cases. Left: capacity distance; right: objective distance.

technology (with a small epsilon guard on denominators), and (ii) objective distance, the magnitude of the relative difference vs the baseline objective.

5 RESULTS

On the log–log charts below, an error value of 0.01 corresponds to a 1% deviation from the baseline. For most planning applications, uncertainties in wind data, cost forecasts, and discount rates suggest objective errors up to $\sim 10\%$ are often acceptable; we therefore seek curves that remain near zero as compression increases, and in particular below ~ 0.1 in objective distance for moderate compression.

An observation which persists across all methods is that the capacity errors can be large while objective errors remain small. This reflects near-degeneracy in the capacity design space: multiple capacity vectors achieve similar costs, so capacity distance may increase even when the objective is well preserved. Future work should consider this when designing hour-ranking methods that use capacity error as an importance signal. For the remainder of this work we focus on the objective error measure.

Dual-informed methods (Figures 3 and 4) sometimes reduce capacity-vector error more rapidly with compression than a priori Time-step/Horizon baselines (Figures 1–2), but exhibit significant variability across cases

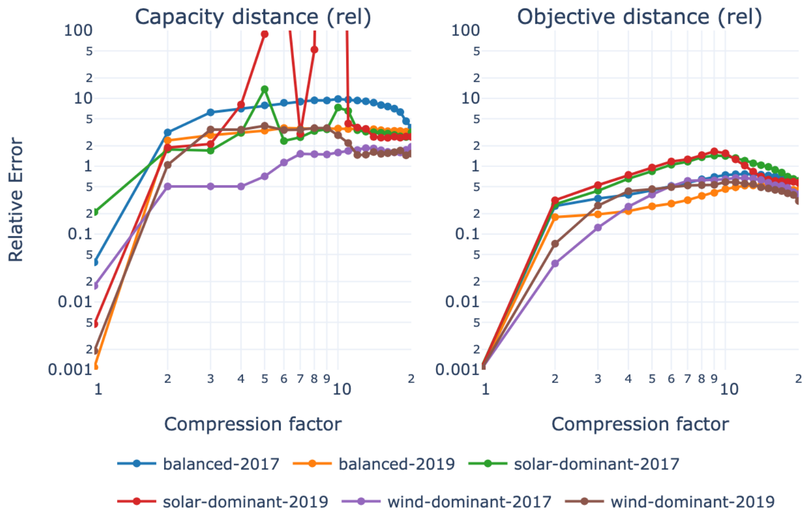

Figure 3: A posteriori selection using SOC-dual metric (sum of absolute duals on SOC upper bounds). Left: capacity distance; right: objective distance.

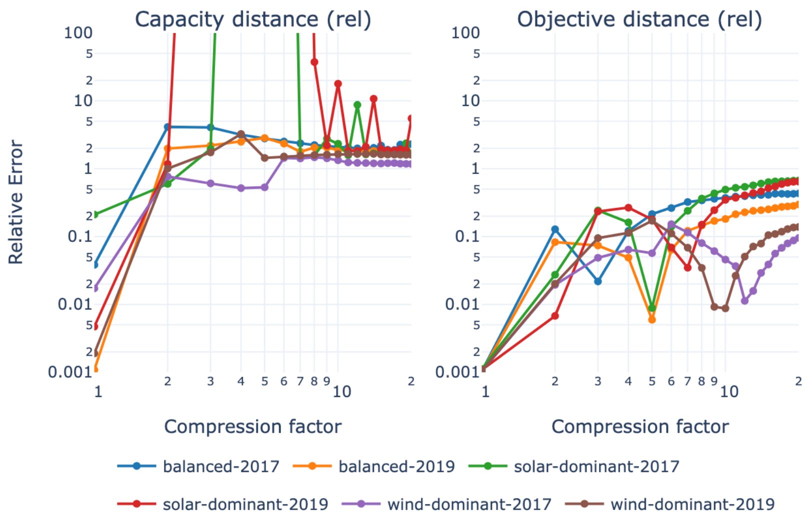

Figure 4: A posteriori selection using VRE-capacity-dual metric. Left: capacity distance; right: objective distance.

that undermines their reliability. While the dual series occasionally achieve lower capacity distance than the baselines for moderate compression, this performance is inconsistent and case-dependent.

Quantitatively, aggregating across cases, the median capacity distance at $c \approx 2$ is ~ 0.06 for Time-step and ~ 0.18 for Horizon, with Time-step consistently outperforming Horizon truncation. At $c \approx 4$, Time-step attains ~ 0.18 while Horizon remains higher (~ 0.75). The dual-based approaches show promising results in some instances but lack the robustness needed for reliable application across different scenarios.

Neutralisation curves (Figure 5) exhibit similar inconsistency, performing well in some cases but poorly in others, although since these are based on the the ranking of distance between the capacity vectors of the base case and the neutralised solutions, in future work it will be worth revisting this approach using the objective measure of distance.

6 DISCUSSION AND CONCLUSIONS

The comparison between a priori and a posteriori compression methods reveals mixed results that challenge the initial hypothesis that dual-informed hour rankings would provide consistently superior compression performance. While dual-based methods occasionally demonstrate lower error rates than simple baselines, their

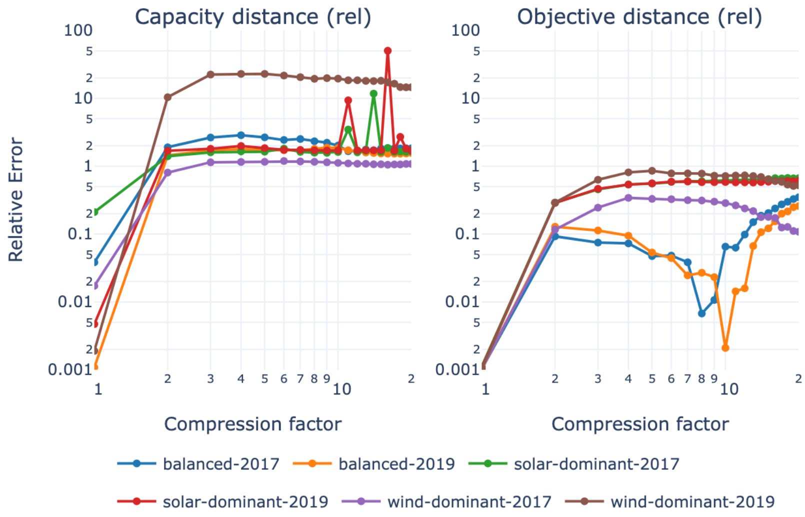

Figure 5: A posteriori selection using neutralisation-derived hour importance. Left: capacity distance; right: objective distance.

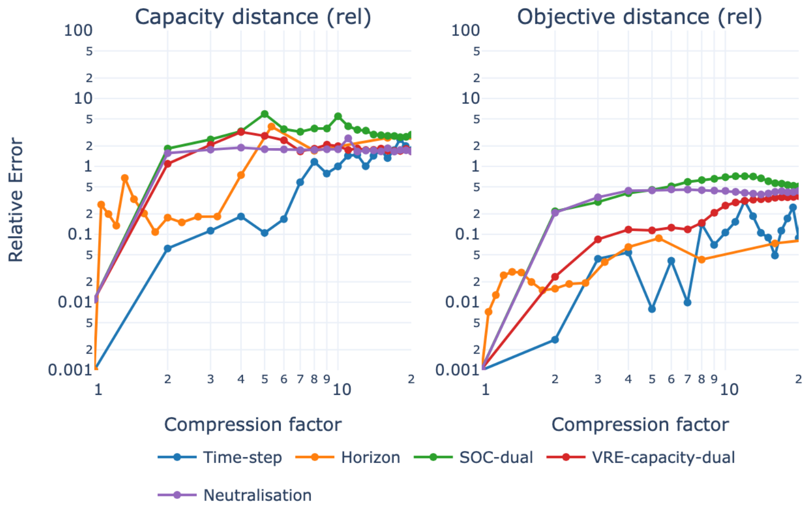

Figure 6: Method comparison summary across all cases: median capacity and objective distances versus compression for Time-step, Horizon, SOC-dual, VRE-capacity-dual, and Neutralisation.

inconsistent performance across test cases undermines their practical utility for reliable compression.

The fundamental limitation appears to stem from the nature of the capacity expansion problem itself. Many of these single-region, three-technology scenarios exhibit near-degeneracy in the capacity design space, where multiple capacity vectors achieve very similar objective values. This characteristic makes capacity-distance-based sensitivity metrics unreliable indicators of hour importance, as small perturbations can lead to seemingly large capacity changes with minimal objective impact.

Additionally, the a posteriori methods tested do not account for discretisation errors that would be introduced when transitioning to large, non-uniform time steps without correspondingly increasing the order of the temporal discretisation scheme used for storage state-of-charge linking. This oversight likely contributes to the observed performance degradation at higher compression factors.

Practical recommendations. Based on our summary findings (cf. 6), we recommend:

- **Prefer time-step coarsening over horizon truncation** in all cases. The consistent superiority of time-step methods across all scenarios makes this the clear choice for practical applications.

- **Use dt=2 as a reliable starting point.** Shifting from 1-hour to 2-hour resolution consistently introduces error of less than 1% across all test cases (as shown in Figure 1) while providing at least a 50% reduction in computational burden.
- **Exercise caution with higher compression factors.** While dt=3–6 may be acceptable in some cases, the error accumulation becomes less predictable and more scenario-dependent.
- **A posteriori methods require further investigation** given their inconsistent performance and computational overhead of requiring a full-resolution solve.

Limitations and future work. This study is limited to single-region scenarios with three technologies (solar, wind, 4-hour battery) and up to 10% gas generation. The conclusions may not generalise to multi-region systems with transmission constraints, longer-duration storage, or different technology mixes. Future work should investigate whether a posteriori methods perform more consistently in systems with clearer binding constraints and less degenerate capacity solutions.

Specific methodological improvements could include: (i) multi-objective sensitivity analysis that considers both capacity and objective perturbations simultaneously; (ii) robust optimization techniques that account for solution degeneracy when ranking hours; (iii) higher-order temporal discretisation schemes (e.g., variable-step integrators) that better handle large time steps while preserving storage dynamics; and (iv) hybrid approaches that combine the reliability of time-step coarsening with targeted selection of critical hours during extreme events.

The original goal of assembling training datasets for ML-based predictors remains valid but requires these more robust a posteriori selection methods to generate consistent labels across diverse scenarios.

REFERENCES

Green D. L., Graham P., Havas L., Browne G., Foster J., Hayward J., Reedman L., Khandoker T., Ross A. and Rogers J. (2024), Northern territory low emissions carbon capture storage and utilisation hub, power generation options analysis – task 7 report, Technical Report EP2024-6162, CSIRO, Australia. 72 pp.

Hilbers A. P., Brayshaw D. J. and Gandy A. (2023), Reducing climate risk in energy system planning: A posteriori time series aggregation for models with storage, *Applied Energy* **334**, 120624.

Kotzur L., Markewitz P., Robinius M. and Stolten D. (2018), Time series aggregation for energy system design: Modeling seasonal storage, *Applied Energy* **213**, 123–135.

Nahmmacher P., Schmid E., Hirth L. and Knopf B. (2016), Carpe diem: A novel approach to select representative days for long-term power system modeling, *Energy* **112**, 430–442.

Pfenninger S. (2017), Dealing with multiple decades of hourly wind and pv time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability, *Applied Energy* **197**, 1–13.

Pfenninger S. and Staffell I. (2016), Long-term patterns of european pv output using 30 years of validated hourly reanalysis and satellite data, *Energy* **114**, 1251–1265.

Staffell I. and Pfenninger S. (2016), Using bias-corrected reanalysis to simulate current and future wind power output, *Energy* **114**, 1224–1239.

Teichgraeber H. and Brandt A. R. (2019), Clustering methods to find representative periods for the optimization of energy systems, *Applied Energy* **239**, 1283–1293.

Teichgraeber H., Lindenmeyer C. P., Baumg"artner N., Kotzur L., Stolten D., Robinius M. and Brandt A. R. (2020), Extreme events in time series aggregation: A case study for optimal residential energy supply systems, *Applied Energy* **275**, 115223.