SOC-informed feedback enhancement of time-series aggregation

for storage-embedded power system expansion planning

MSc EE-EPE Thesis Project

Ziling Zhou

Delft University of Technology

TU Delft

SOC-informed feedback enhancement of time-series aggregation

for storage-embedded power system expansion planning

by

Ziling Zhou

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday July 2nd, 2026 at 9:00 AM.

Thesis supervisors:	Prof. Milos Cvetkovic,	IEPG, TU Delft, Thesis Supervisor
	Ruiqi Zhang MSc,	IEPG, TU Delft, Daily Supervisor
Thesis committee:	Prof. Milos Cvetkovic,	IEPG, TU Delft, Chair
	Dr.ir. K. (Kenneth) Bruninx,	Energy and Industry, TU Delft
Student number:	6165877	
Project duration:	November 19th, 2025 – July 2nd, 2026	

Abstract

The increasing penetration of renewable energy sources makes power system expansion planning with storage strongly dependent on chronological operational conditions. Time-series aggregation (TSA) can reduce the computational burden of full-year planning models, but it distorts the temporal information that determines storage operation and investment decisions.

This thesis investigates how SOC-based diagnostic information can improve representative-day-based expansion planning for storage-embedded power systems. The planning model jointly considers transmission expansion, wind investment, and energy storage sizing. Representative days are selected through hierarchical clustering with preserved extreme days, while sequentially linked days are used to reconstruct inter-day chronology. The framework aims to find investment decisions that achieve lower total cost under full chronological evaluation.

Reduced models with different numbers of representative days are evaluated, and their reconstructed SOC trajectories are analyzed. SOC-based diagnostic metrics are developed to characterize the components of the trajectory gap. These metrics are used as diagnostic signals for identifying where the current temporal representation is insufficient for storage-related operation.

The results show that daily-cycling storage is mainly affected by intra-day shape mismatch, while long-duration storage is more sensitive to accumulated drift and inventory-level bias. Natural days with large intra-day shape mismatch are useful feedback candidates because they reveal inadequacy in the current representative-day set for describing storage charging and discharging patterns. Preserving these day-shape-critical days improves storage-relevant temporal representation and can lead to better investment decisions.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my thesis supervisor, Prof. Milos Cvetkovic, for his guidance, support, and valuable feedback throughout this thesis project. His insights helped me set the research direction, refine the scope of the study, and improve the quality of this work.

I would also like to thank my daily supervisor, Ruiqi Zhang MSc, for his patient guidance, continuous support, and detailed discussions throughout the project. Our regular meetings greatly helped me understand the relevant literature and theoretical foundations of this research. His guidance also encouraged me to interpret experimental results more critically and to develop a more rigorous way of thinking about research questions, methodologies, and conclusions.

I am also grateful to the members of my thesis committee, Prof. Milos Cvetkovic and Dr.ir. K. (Kenneth) Bruninx, for their time, comments, and constructive suggestions.

Furthermore, I would like to thank my dear friends overseas and back home for their constant care and emotional support. During the most anxious and uncertain moments of this journey, your conversations, encouragement, and shared reflections helped me regain confidence and look forward to the future with more strength. Your warm companionship has been an important source of comfort and motivation during my study abroad experience.

I would also like to thank my family for their understanding, patience, and unconditional support. Their trust has always been a quiet but powerful source of strength, especially during the challenging stages of my study and thesis work.

Ziling Zhou

Contents

Chapter 1

Introduction

1.1 Overview

1.1.1 Energy transition and planning challenges

The transition towards low-carbon power systems is increasing the penetration of renewable energy sources (RES), such as wind and solar generation. These technologies are essential to reduce carbon emissions and support long-term decarbonization targets. However, their variability and uncertainty introduce significant challenges for both system operation and long-term planning. Power systems with high RES penetration require sufficient flexibility to balance supply and demand across multiple temporal scales, ranging from hourly ramping to multi-day or seasonal energy shifting.

Power system expansion planning models are widely used to determine long-term investment decisions in generation capacity, transmission reinforcement, and flexibility resources, such as energy storage and reserve-providing generation [7, 9]. In renewable-based systems, hourly chronological time-series data are often needed to capture the inherent variability of load demand and renewable generation, as well as their effects on ramping, storage operation, and other inter-temporal constraints [3, 5]. In this thesis, the expansion planning problem is formulated as a storage-embedded co-planning problem involving transmission expansion, wind capacity expansion, and ESS investment. The model aims to support cost-oriented investment decisions that achieve low full-time total cost while maintaining operational feasibility. These investment decisions are evaluated together with chronological operational constraints, including generation dispatch, transmission flows,

renewable curtailment, reserve requirements, and storage SOC dynamics.

However, representing these chronological constraints with full-year hourly data can make the planning model computationally demanding, because each time step adds operational variables and constraints to the investment problem. This creates a trade-off between temporal detail and model tractability, and motivates the use of time-series aggregation methods in expansion planning models with high renewable penetration [3, 19]. Time-series aggregation (TSA) has therefore been widely adopted to reduce the temporal dimension of energy system and power system planning models. Existing reviews show that TSA methods can substantially reduce computational requirements by replacing the original high-resolution time series with a smaller set of representative periods, while attempting to preserve the relevant characteristics which can let the reduced model generates a closer solution as the original model[1, 2]. Within TSA, representative-period approaches are commonly used, and representative days selected by clustering-based methods provide a practical compromise between model tractability and temporal representativeness [28, 3].

Based on these planning challenges, this thesis considers a storage-embedded co-planning problem that includes wind capacity expansion, transmission expansion, and ESS investment. Wind expansion represents the integration of low-carbon but variable generation, transmission expansion enables renewable power to be delivered across the network and reduces congestion-related restrictions, and ESS investment provides temporal flexibility by shifting surplus renewable energy to periods of higher demand or lower renewable availability. Considering these three investment options together provides a more comprehensive planning framework, because renewable integration depends not only on additional generation capacity, but also on the network capability to transfer power and the storage capability to balance temporal mismatches [8, 9].

1.1.2 Time-series aggregation and chronology representation

Conventional clustering-based TSA methods usually group time periods according to statistical similarity. For example, natural days with similar load, wind, and solar profiles may be assigned to the same representative day. This approach is effective in preserving major statistical patterns of the input data

and reducing model size [28, 3]. However, ordinary clustering does not necessarily preserve the chronological order of the original time series. Two days may be similar in shape but located in different seasons or under different preceding and following system conditions. And consecutive natural days may be assigned to different clusters when their profiles are dissimilar. This mismatch between profile similarity and chronological continuity leads to a loss of chronology in the reduced time series.

This loss of chronology is particularly relevant in systems with strong inter-temporal constraints. Renewable generation and load demand are not only characterized by their individual hourly values, but also by their temporal sequence. Prolonged low-renewable periods, multi-day high-load events, and transitions between surplus and deficit conditions are all sequence-dependent phenomena. If the reduced model only preserves representative profiles but not their chronological relationships, it may misrepresent the operational conditions that drive flexibility requirements.

To address this limitation, chronology-preserving TSA methods have been proposed. Pineda and Morales developed chronological time-period clustering for capacity expansion planning with storage, in which only adjacent time periods are merged so that the reduced sequence preserves the original chronological order [3]. Other studies have also proposed linked or interconnected representative-period frameworks to improve the representation of both intra-period and inter-period storage behavior [4]. These studies indicate that temporal reduction should not only approximate the distribution of input profiles, but also retain or reconstruct the chronological structure that determines system flexibility.

At the same time, chronology is not the only feature that matters in representative-period selection. A reduced time series may preserve average profiles but still miss a small number of planning-critical periods. Representative-day selection studies therefore emphasize that extreme operating conditions should also be considered, because they can strongly influence investment decisions and system operating costs, enabling more reliable planning results[13, 10, 31]. In renewable-based systems, such extreme conditions are not limited to peak-load days; they may also include high net-load days, low-renewable periods, and other scarcity-related periods[32, 17]. This thesis builds on this idea, but focuses specifically on whether storage-related diagnostic information can provide useful feedback for extreme-day selection.

1.1.3 Storage-embedded systems and SOC representation

The chronology-related limitations of representative-day and clustering-based TSA become more critical when energy storage systems (ESS) are included in power system planning. Even chronology-preserving or linked-period methods can only approximate the original SOC evolution, because storage operation depends on the accumulated charging and discharging history across time. Unlike conventional generation units, storage systems do not only provide instantaneous power output. Their operation is governed by the state of charge (SOC), which links charging and discharging decisions across time. Charging in one period increases the energy available for future discharge, while discharging reduces the remaining inventory. Therefore, storage operation is inherently path-dependent.

This feature makes storage systems highly sensitive to temporal aggregation. If representative days are treated independently, or if their chronological sequence is not properly reconstructed, the reduced model may fail to represent the actual evolution of storage inventory. Kotzur et al. highlight this issue in the context of seasonal storage and propose a storage formulation that separates intra-period and inter-period states, showing that storage inventories require explicit links between representative periods [5]. Similar concerns are also discussed in studies on long-duration storage and enhanced representative-period formulations, where storage feasibility and inter-period energy transfer are shown to depend strongly on the chosen temporal representation [12, 11]. This issue is also consistent with recent high-temporal-resolution expansion studies, which show that systems with wind, solar, and storage can be more sensitive to temporal representation choices than thermal-dominated systems [39].

This insight is important for expansion planning because the value of storage depends not only on within-day cycling, but also on the ability to shift energy across multiple days or longer horizons[5, 17]. Representative-day methods can capture daily operating profiles, but they may distort the inter-day sequence that determines long-duration storage operation. Hybrid representative-day and sequential-linking approaches have therefore been proposed to preserve selected extreme periods during clustering and to reconstruct part of the chronological relationship between representative days. In such approaches, predefined extreme days are kept as representative profiles when their clusters are merged, while sequentially linked days use the original

day-to-representative-day mapping to connect repeated representative-day blocks for inter-day SOC tracking [17].

In storage-embedded systems, SOC trajectories provide a direct way to observe how temporal aggregation represents storage operation. Even when a reduced model produces an economically competitive investment decision, replacing heterogeneous natural days with clustered representative days can still lead to considerable reconstructed SOC gaps. These gaps are useful because they indicate where the reduced temporal representation poorly captures storage behavior, especially in terms of intra-day profile mismatch and inter-day chronology reconstruction. Therefore, SOC trajectory analysis is not used as a separate planning objective, but as a diagnostic signal to identify poorly represented periods and to support feedback-enhanced representative-day selection, which can help the reduced model derive investment decisions with better full-time cost performance.

1.1.4 Need for storage-aware evaluation and feedback

Existing evaluation frameworks for TSA-based reduced models are largely cost-oriented. Cost-based metrics are commonly used because they directly measure whether the reduced model provides economically reliable planning decisions. Bahl et al. emphasized the importance of evaluating time-series aggregation in the domain of the objective function rather than only through statistical similarity of the input data [6]. This cost-oriented perspective is essential because the final purpose of expansion planning is to support reliable and low-cost investment decisions.

However, in storage-embedded planning models, cost-based evaluation alone may not fully reveal the mechanism behind reduced-model errors. Since SOC does not usually appear as an explicit cost term in the objective function, storage-state errors may remain hidden behind similar total costs. The SOC trajectory affects the feasible charging and discharging capability of storage, and therefore determines how much flexibility is actually available at each time step. If the reconstructed SOC is biased, the reduced model recognizes different system stress and storage value, thus making less-optimal investment decisions.

Recent representative-day selection studies also show that reduced temporal representations should be evaluated by their impact on optimization outcomes, not only by input similarity. Li et al. distinguish between input-based and cost-based representative-day selection and show that representa-

tive days can be selected with explicit consideration of their effect on capacity expansion outcomes [31]. Hilbers et al. further develop a posteriori TSA schemes in which aggregation is informed by model operational variables and chronology preservation, which supports the broader idea that diagnostic information from model behavior can be used to refine temporal aggregation [38]. This model-aware direction motivates the use of diagnostic information from the optimization model as feedback for improving representative-period selection.

This thesis therefore treats SOC-based analysis as a complementary diagnostic layer for cost-oriented reduced planning. The purpose is not to replace economic indicators or to directly minimize the SOC gap. Instead, SOC diagnostics are used to explain how temporal aggregation distorts storage-state evolution and to identify storage-relevant periods that are useful for improving representative-period selection. By explicitly analyzing SOC reconstruction errors, worst periods in current temporal representation can be identified for improvements, thus deriving more optimal investment results.

1.2 Research questions and contribution

1.2.1 Motivation

The previous sections show that temporal representation is a central issue in storage-embedded expansion planning. Since the value of ESS depends on shifting energy across time, especially across daily and multi-day horizons, planning models with storage require a careful representation of chronological relationships. Existing TSA approaches have improved this representation by describing intra-day operation through representative periods and reconstructing part of the inter-day chronology through linked-period formulations. However, it is still necessary to evaluate how well this reconstructed chronology represents storage behavior. In this context, SOC reconstruction gaps provide a direct indication of temporal-representation errors, because they show where the reduced model fails to reproduce the storage-state evolution implied by the full chronological model. This motivates the central research objective of this thesis:

> How can SOC-based diagnostic information be used to explain storage-related errors and guide feedback-enhanced extreme-day

selection in cost-oriented representative-day-based power system expansion planning models with storage?

To address this objective, this thesis develops a storage-aware evaluation and feedback framework for TSA-based reduced planning models. The framework aims to find planning results of lowest total cost, while using using SOC diagnostic information to guide later extreme-day selection.

1.2.2 Research questions

The thesis is organized around the following research questions, which is explored and answered respectively in chapter 4, 5, and 6.

RQ1: How reconstructed SOC trajectories connect between time-series aggregation and investment decisions in reduced planning models with storage?

RQ2: What additional storage-related error mechanisms can SOC-based diagnostics reveal in economically competitive reduced planning models?

RQ3: Can SOC-based diagnostic signals guide extreme-day selection and improve the total-cost performance of representative-day-based planning models?

1.2.3 Contribution of the thesis

The main contributions of this thesis are summarized as follows.

- **Cost-oriented evaluation of reduced planning models with storage.** The thesis evaluates representative-day-based planning models using operational estimation error and optimality gap, and identifies reduced models that remain economically competitive when evaluated under the full chronological time series.
- **Identification of the limitation of cost-only interpretation in storage-embedded planning models.** The thesis shows that cost-based metrics remain necessary for assessing planning performance, but they do not fully explain storage-state reconstruction errors caused by temporal aggregation.

- **Development of SOC-based diagnostic metrics.** The thesis proposes SOC-based indicators to characterize the discrepancy between reduced-model and full-model storage trajectories. These metrics captures different components of SOC reconstruction error, which are formally defined in Chapter 4.

- **Mechanism-based interpretation across storage types.** The thesis shows that daily-cycling storage and long-duration storage may be affected by different dominant SOC error mechanisms. This distinction helps explain why the same temporal aggregation method can distort different storage technologies in different ways.

- **Integration of chronology reconstruction into storage-aware evaluation.** The thesis analyzes how representative days and consecutive day blocks reconstruct inter-day SOC evolution, and how this reconstruction can introduce structured storage-state errors.

- **Design and testing of SOC-informed extreme-day feedback.** The thesis explores how SOC-based diagnostic signals can be used to guide extreme-day selection in TSA. The feedback analysis tests whether storage-relevant extreme days can improve the total-cost performance of representative-day-based planning models.

- **Reflection on the role of SOC diagnostics in cost-oriented planning.** The thesis clarifies that SOC diagnostics should not be interpreted as a method for globally minimizing SOC reconstruction error. Instead, they provide solution-dependent diagnostic signals that can help identify storage-relevant extreme-day candidates and improve the interpretation of planning results.

- **Demonstration through case studies.** The thesis evaluates reduced models under different numbers of representative days and different extreme-day settings. It analyzes how total cost, investment decisions, and SOC error mechanisms vary across storage types, time periods, and SLD blocks.

1.3 Structure of the thesis

The remainder of this thesis is organized as follows.

Chapter 2 presents a literature review of power system expansion planning and time-series aggregation methods. It first reviews representative-period approaches and clustering-based TSA methods, then discusses chronology-preserving aggregation, extreme-period selection, and storage-aware temporal reduction. The chapter also reviews existing evaluation metrics for reduced planning models and identifies the research gap addressed by this thesis: the need to connect storage-aware diagnostics with cost-oriented representative-period selection.

Chapter 3 formulates the power system co-planning model used in this study. The model includes investment decisions for transmission lines, wind generation, and energy storage systems, as well as operational constraints based on DC optimal power flow. Particular attention is given to the modeling of ESS operation and the RD-SLD chronology reconstruction, where sequentially linked days are introduced in Chapter 2 and formulated in Chapter 3, used to propagate SOC across representative and sequential periods.

Chapter 4 evaluates reduced planning models using both cost-based and SOC-based metrics. It first presents conventional economic indicators, including operational estimation error and optimality gap, to identify economically competitive reduced models. It then introduces SOC-based metrics to diagnose discrepancies between reduced-model and full-model storage trajectories.

Chapter 5 provides a detailed diagnostic analysis of SOC reconstruction errors. It examines how SOC errors differ across storage units, storage types, natural days, and SLD blocks. The chapter identifies whether the dominant error mechanism is intra-day shape mismatch, cross-day drift, or inventory-level bias.

Chapter 6 develops feedback-enhanced TSA strategies based on SOC diagnostic results. It investigates how storage-relevant critical days can be incorporated as extreme days in the representative-day selection process, and compares different SOC-informed extreme-day settings in terms of total-cost performance and investment decisions.

Chapter 7 concludes the thesis by summarizing the main findings, discussing the limitations of the proposed framework, and outlining future research directions. In particular, it reflects on why SOC-informed feedback can reduce total cost even when aggregate SOC gaps do not necessarily decrease, and discusses how iterative or sensitivity-based feedback could be explored in future work.

Chapter 2

Literature Review

2.1 Power system expansion planning and temporal reduction

2.1.1 Expansion planning under renewable and storage integration

Power system expansion planning determines long-term investment decisions for generation, transmission, and flexibility resources in order to ensure reliable and cost-effective system operation. Under increasing penetration of renewable energy sources (RES), this problem becomes more complex because renewable generation is variable, weather-dependent, and often geographically separated from major load centers [7, 8]. As a result, generation investment, transmission reinforcement, and operational flexibility must be considered jointly rather than independently.

Energy storage systems further extend the scope of expansion planning. Storage can charge during periods of low demand or high renewable generation and discharge during high-demand, low-renewable, or congested periods. It can therefore reduce curtailment, relieve congestion, and support renewable integration. Coordinated investment in transmission and storage is important because the two provide complementary forms of flexibility [9]. Long-duration storage further increases the importance of temporal representation, since its value depends on shifting energy across several days, weeks, or seasons [10, 11].

In this thesis, the planning problem is positioned as a co-planning prob-

lem involving transmission expansion, wind capacity expansion, and energy storage investment. The value of each investment option depends on the operation of the others: wind investment may be limited by network congestion, transmission expansion can reduce curtailment, and storage can shift energy across time and locations. This interdependence motivates a planning model that combines investment decisions with operational constraints.

2.1.2 Need for time-series aggregation

A central difficulty in renewable-based expansion planning is the need to represent short-term operational variability within a long-term investment model. Although investment decisions are made over long horizons, their feasibility and value depend on hourly or daily operating conditions, including load variation, renewable fluctuations, transmission congestion, curtailment, load shedding, and storage operation. This is particularly important for storage, because its state of charge (SOC) depends on previous charging and discharging decisions and therefore requires chronological information [3, 5].

Using the full chronological time series creates a severe computational burden. A single year of hourly data contains 8760 operating points, and each introduces dispatch variables and operational constraints. When these are combined with investment variables, especially binary variables for transmission expansion, the resulting optimization problem becomes computationally intractable [3, 10]. This creates a trade-off between temporal accuracy and computational tractability.

Time-series aggregation (TSA) addresses this trade-off by replacing the original chronological data set with a smaller set of representative periods. These periods are expected to preserve the most relevant characteristics of load, renewable generation, and system stress while reducing model size [1, 2]. The key issue is therefore not only whether the reduced time series approximates the original input data, but also whether it preserves the temporal features that influence planning decisions.

2.2 Representative-period selection methods

2.2.1 Representative days, weeks, and system states

Several temporal representations have been used in power system planning. Duration curves and system states provide compact representations of time-dependent parameters, but they usually remove part or all of the chronological order of the original time series. This limits their ability to represent inter-temporal constraints such as ramping and storage dynamics [3, 10]. System-state approaches can preserve correlations among multiple variables and may include transition information, but they do not naturally guarantee feasible storage energy trajectories [12].

Representative days and representative weeks are more common in expansion planning because they retain chronological information within each selected period. Representative days are suitable for intra-day load patterns, renewable generation profiles, and short-duration storage cycling. Representative weeks provide a longer temporal window and capture multi-day renewable patterns, but under a fixed time-step budget fewer weeks than days can be selected, which may reduce temporal diversity and still miss rare operating conditions. In practice, representative periods are often selected using clustering-based TSA methods such as k-means, k-medoids, or hierarchical clustering [13, 3, 14].

2.2.2 K-means, k-medoids, and hierarchical clustering

Clustering-based TSA treats each temporal unit as a feature vector. For a representative-day formulation, an original day i can be written as

$$\mathbf{x}_i = \left[\mathbf{p}_i^L, \mathbf{p}_i^W, \mathbf{p}_i^S, \cdots\right], \tag{2.1}$$

where $\mathbf{p}_i^L$, $\mathbf{p}_i^W$, and $\mathbf{p}_i^S$ denote the load, wind, and solar profiles of day i. The goal is to group similar days and select representative profiles that approximate the original data with fewer time periods.

In k-means clustering, each cluster is represented by its centroid μ_k, which is the mean feature vector of all temporal units assigned to that cluster, and the algorithm minimizes the within-cluster squared distance:

$$\min_{\mathcal{C}_1,\ldots,\mathcal{C}_K} \sum_{k=1}^{K} \sum_{i\in\mathcal{C}_k} \|\mathbf{x}_i - \boldsymbol{\mu}_k\|_2^2 . \tag{2.2}$$

This method is simple and computationally efficient, but the centroid smooth peak values and does not necessarily correspond to an actual historical period [15, 13]. K-medoids instead selects an actual observed period as the representative:

$$m_k = \arg\min_{j \in \mathcal{C}_k} \sum_{i \in \mathcal{C}_k} d(\mathbf{x}_i, \mathbf{x}_j). \tag{2.3}$$

Because the medoid is a real period, it can better preserve realistic profiles. However, k-medoids and k-means are both sensitive to the initial medoid selection and may converge to local solutions [12, 10].

Hierarchical clustering starts with each period as an individual cluster and iteratively merges clusters until the desired number of representative periods is reached. Compared with k-means and k-medoids, it does not require random initialization and provides a transparent merging structure. In the framework adopted in this thesis, this merging structure is convenient for explicitly preserving predefined extreme days during representative-day selection [17].

2.2.3 Ward's method and feature construction

A common linkage rule in agglomerative hierarchical clustering is Ward's method, which merges the pair of clusters that causes the smallest increase in within-cluster variance [18]. For two clusters A and B, the Ward dissimilarity can be written as

$$D(A, B) = \frac{2|A||B|}{|A| + |B|} \left\| \bar{\mathbf{x}}_A - \bar{\mathbf{x}}_B \right\|_2^2, \tag{2.4}$$

where $\bar{\mathbf{x}}_A$ and $\bar{\mathbf{x}}_B$ are the cluster centroids defined as below:

$$\bar{\mathbf{x}}_A = \frac{1}{|A|} \sum_{i \in A} \mathbf{x}_i. \tag{2.5}$$

Ward's method is therefore consistent with minimizing information loss during hierarchical grouping.

The performance of clustering-based TSA depends strongly on the selected features. In renewable-based planning, common features include load demand, wind generation, solar generation, spatial correlations, and net load. Net load represents the residual demand that must be served by dispatchable

generation, storage, or imports:

$$p_{i,t}^{NL} = p_{i,t}^{L} - p_{i,t}^{RES}. \tag{2.6}$$

Here, net load should be interpreted as a physical power quantity, not as a direct subtraction of normalized load factors and renewable generation factors. In principle, peak net-load periods can be more planning-relevant than peak load alone because they reflect both demand and renewable availability. Net-load-related features and extreme net-load conditions are therefore often considered in representative-day selection for renewable and storage systems [17, 10].

In the present thesis, however, the available time-series inputs for representative-day selection are load factors and wind factors rather than complete pre-investment MW-level net-load profiles. Moreover, wind capacity is itself an investment decision in the co-planning model, so the final renewable injection level is not fully known before solving the planning problem. Therefore, the conventional baseline extreme day is selected as the peak-load day. Later chapters then test whether SOC-based diagnostic information can identify alternative storage-relevant extreme days that improve total-cost-oriented planning performance.

2.3 Chronology and extreme-period preservation

2.3.1 Limitations of ordinary clustering

Ordinary clustering-based TSA is effective for reducing temporal complexity, but it has two important limitations. First, it groups periods according to similarity in feature space rather than chronological adjacency. As a result, it can preserve intra-period profiles but does not directly preserve the order of the original time series. This is problematic for modeling storage SOC evolution [3, 5].

Second, ordinary clustering may underrepresent extreme conditions. Centroid-based methods can average peak load, low renewable availability, or high net-load periods with normal periods. Medoid-based methods select real periods, but they still do not guarantee that planning-critical stress periods are included in the representative set. A representative set may therefore

Table 2.1: Summary of representative-period selection methods discussed in the literature.

Method	Main idea	Main limitation	References
K-means	Groups periods around centroid profiles	Sensitive to initial mean selection; may smooth peaks	[15, 13]
K-medoids	Selects actual historical periods as representatives	Sensitive to initial medoid selection	[12, 10]
Hierarchical clustering	Iteratively merges similar periods	Standard form does not preserve chronology or extremes	[16, 18]
Representative weeks	Preserve multi-day chronology	Lower temporal diversity under a fixed time-step budget; rare conditions may still be missed	[3, 11]
System states	Compactly represent typical operating conditions	Storage feasibility and inter-period SOC continuity can be difficult	[12, 10]

approximate average input patterns while missing critical chronological or extreme-value information [13, 10, 17].

These limitations motivate two main directions in the literature: chronology-preserving clustering and extreme-period preservation.

2.3.2 Chronological and priority chronological clustering

Chronological time-period clustering (CTPC) preserves chronology by allowing only adjacent periods to be merged. In ordinary clustering, the pair of clusters selected for merging can be written conceptually as

$$(A^*, B^*) = \arg\min_{A,B\in\mathcal{C}} D(A, B), \tag{2.7}$$

where $\mathcal{C}$ denotes all possible cluster pairs. In CTPC, the search is restricted to neighbouring cluster pairs:

$$(A^*, B^*) = \arg\min_{(A,B)\in\mathcal{N}} D(A, B), \tag{2.8}$$

where $\mathcal{N}$ denotes adjacent pairs in the original chronological order. This restriction allows the reduced sequence to retain chronological structure and makes it possible to model inter-temporal constraints over the reduced horizon [3].

CTPC is particularly relevant for storage because SOC can be propagated along the reduced chronological sequence. However, the adjacency restriction may reduce profile-similarity accuracy and may smooth local extreme values when only a limited number of clusters is used [3, 11]. Priority chronological time-period clustering (PCTPC) addresses part of this issue by assigning higher priority to critical periods, recognizing that chronology and extreme values should be considered together in expansion planning with long-term pattern [10].

2.3.3 Hybrid representative-day selection with extreme periods

A second route is to preserve the representative-day structure while explicitly protecting planning-relevant extreme days during clustering. Unlike an append-after-clustering approach, which appends extreme periods after clustering, preserved extreme days remain inside the hierarchical clustering procedure. If a cluster containing an extreme day is merged with another cluster, the representative profile of the merged cluster is set equal to the extreme-day profile. Therefore, the extreme day is not averaged out by centroid updating, but continues to represent the merged cluster in subsequent clustering steps.

Moradi-Sepahvand and Tindemans propose a hybrid representative-day approach that combines hierarchical clustering, extreme-value preservation, and sequentially linked days for planning transmission lines and long-term storage systems [17]. Let each natural day i be represented by a feature vector $\mathbf{x}_i$. The algorithm initializes each day as a separate cluster:

$$\mathcal{C}^{(0)} = \{\{1\}, \{2\}, \ldots, \{N\}\}\,. \tag{2.9}$$

For a cluster A, the centroid is

$$\bar{\mathbf{x}}_A = \frac{1}{|A|} \sum_{i \in A} \mathbf{x}_i. \tag{2.10}$$

The dissimilarity between two clusters is measured by a Ward-type criterion, and at each iteration, the pair of clusters with the smallest dissimilarity is

merged:

$$(A^*, B^*) = \arg \min_{A,B \in \mathcal{C},\, A \neq B} D(A, B). \tag{2.11}$$

This procedure is repeated until the desired number of representative days is reached.

The hybrid feature lies in the treatment of preserved extreme days. Let $\mathcal{E} \subset \mathcal{D}$ denote the set of preserved extreme days, where $\mathcal{D}$ is the set of original natural days. After two clusters A^* and B^* are merged into $C = A^* \cup B^*$, the representative profile of C is updated as

$$\mathbf{r}_C = \begin{cases} \mathbf{x}_e, & \text{if } C \cap \mathcal{E} \neq \emptyset,\ e \in C \cap \mathcal{E}, \\ \bar{\mathbf{x}}_C, & \text{if } C \cap \mathcal{E} = \emptyset. \end{cases} \tag{2.12}$$

where $\mathbf{x}_e$ is the feature profile of the preserved extreme day contained in cluster C, and $\bar{\mathbf{x}}_C$ is the centroid profile of the merged cluster. This rule means that a merged cluster containing an extreme day is represented by the extreme-day profile rather than by the averaged centroid. The resulting representative-day set is

$$\mathcal{R} = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{NRD}\}, \tag{2.13}$$

where each $\mathbf{r}_d$ is either a centroid profile of a non-extreme cluster or the preserved profile of an extreme day.

The hybrid representative-day route can be summarized as

$$\begin{aligned} &\text{natural days} \rightarrow \text{feature construction} \\ &\rightarrow \text{hierarchical clustering with preserved extremes} \\ &\rightarrow \text{representative days.} \end{aligned} \tag{2.14}$$

This route does not treat hierarchical clustering as universally superior to CTPC. Rather, CTPC preserves chronology by restricting the merging process to adjacent time periods, whereas the hybrid RD route keeps representative-day clustering and adds explicit protection for selected extreme days [3, 10, 17]. This distinction is important because extreme-period preservation is not limited to the maximum-load day. In renewable-based systems, planning stress may also be caused by high net load, low renewable availability, steep ramps, or consecutive scarcity periods [13, 10, 32, 17].

In this thesis, hierarchical clustering is adopted as the base representative-day selection method following the hybrid RD framework in [17]. The purpose is not to prove that this approach is universally better than CTPC, but

to establish a methodological route that combines representative-day selection, extreme-day preservation, and later chronology reconstruction through sequentially linked days. This structure provides the basis for the storage-aware evaluation in later chapters.

2.4 Storage representation under reduced temporal scope

2.4.1 Short-duration and long-duration storage

Energy storage systems introduce temporal coupling into power system planning models because they shift energy across time. Short-duration storage, such as battery energy storage systems, is mainly used for intra-day balancing, renewable smoothing, and short-term flexibility. In contrast, long-duration storage, such as pumped hydro or power-to-gas, can provide inter-day, weekly, or seasonal energy shifting [3, 5, 11].

This distinction is important for TSA. Representative days can often capture daily charging and discharging patterns, which makes them suitable for storage technologies with mainly intra-day cycles. Long-duration storage, however, depends on the chronological sequence of renewable surplus and deficit periods across multiple days or longer horizons. If a reduced temporal representation does not preserve such inter-period relationships, it may misrepresent the value and operation of long-duration storage [11, 17].

2.4.2 Inter-period SOC representation

The key modeling feature of storage is the state of charge (SOC). The SOC at a given time depends on previous charging and discharging decisions, making storage operation path-dependent. Therefore, storage is constrained not only by instantaneous power limits, but also by accumulated energy states.

Under representative-period aggregation, this inter-temporal structure becomes difficult to preserve. If each representative period is treated independently, the model may represent intra-period storage cycling but ignore energy transfer between representative periods. For the candidate ESS, which combines short-term and long-term ESS, retaining inter-period energy shift is central to storage value [5, 11].

Several formulations have been proposed to address this issue. Kotzur et al. separate storage states into intra-period and inter-period components, allowing storage inventory to evolve across typical periods while retaining a reduced temporal structure [5]. Tejada-Arango et al. develop enhanced representative days and system states to improve storage investment analysis under reduced temporal representations [12]. Chronological time-period clustering provides another route by preserving the order of the reduced sequence, allowing storage states to evolve through the aggregated horizon [3]. These studies show that storage-aware TSA requires both representative profiles and a mechanism for SOC continuity.

2.4.3 Sequentially linked days as chronology reconstruction

Sequentially linked days (SLDs) provide a representative-day-based way to reconstruct part of the lost inter-day chronology. Instead of treating representative days as weighted independent profiles, SLDs use the mapping from original days to representative days to identify consecutive blocks of original days represented by the same RD [17]. It provides a pathway to retain intra-day and inter-day energy shift for representative days derived from multiple clustering methods, including k-means, k-medoids, and hierarchical clustering as mentioned before.

After the representative days are obtained, each natural day is mapped to the closest representative day. Let n denote a natural day and let $\mathcal{R}$ denote the set of representative days. The mapping is

$$m(n) = \arg\min_{r \in \mathcal{R}} \|\mathbf{x}_n - \mathbf{r}\|_2 . \tag{2.15}$$

Here, $m(n)$ identifies which representative day best represents natural day n. Based on this mapping, an SLD is defined as a consecutive sequence of natural days with the same assigned representative day. When the mapping changes to another representative day, a new SLD begins. For an SLD s covering natural days from a_s to b_s, this means

$$m(a_s) = m(a_s + 1) = \cdots = m(b_s) = d_s, \tag{2.16}$$

where d_s is the representative day associated with SLD s. The length of the SLD is

$$L_s = b_s - a_s + 1. \tag{2.17}$$

Thus, SLDs are not new representative profiles; they are chronological blocks constructed from the original day-to-RD mapping.

The SLD construction can be summarized as

$$\begin{aligned} &\text{representative days} \rightarrow \text{day-to-RD mapping} \\ &\rightarrow \text{consecutive same-RD day blocks} \\ &\rightarrow \text{SLD chronology reconstruction.} \end{aligned} \tag{2.18}$$

Representative days describe intra-day operating profiles, while SLDs describe the inter-day repetition of these profiles. This distinction is important because the SOC trajectory depends not only on the daily charging and discharging pattern, but also on how many consecutive days this pattern is repeated and what SOC level is carried into the next chronological block.

SLD-based reconstruction remains an approximation. Original days mapped to the same representative day may still differ in load, renewable availability, congestion, and storage dispatch. Repeating the same representative-day SOC pattern across an SLD block can therefore introduce storage-state reconstruction errors, such as intra-day shape mismatch, cross-day accumulated drift, or persistent inventory-level bias. For this reason, SLDs improve the representation of chronology compared with independent RD weighting, but they do not eliminate the need for storage-aware evaluation.

2.5 Evaluation of reduced planning models

2.5.1 Input-based and cost-based evaluation

Evaluation metrics for reduced planning models can be divided into input-based and cost-based metrics. Input-based metrics assess how closely the reduced time series approximates the original data. Common criteria include time-series distance, duration-curve mismatch, correlation preservation, peak-value deviation, and statistical moments of load and renewable generation [1, 2, 13]. These metrics are useful because TSA is first applied to input data before the reduced optimization model is solved.

However, input similarity does not necessarily imply planning accuracy. A small number of high-stress periods may have limited influence on average input error but strongly affect investment decisions in generation, transmission, or storage. Therefore, several studies emphasize that TSA should also be evaluated using optimization outcomes, such as objective-function error,

operational cost error, model resolution sensitivity, or investment-decision quality [6, 19, 20].

Cost-based metrics are closer to the purpose of expansion planning– finding the investment decision of the smallest total cost. They assess whether the reduced model produces investment decisions that remain economically competitive when evaluated under full chronological data. In this thesis, operational estimation error and optimality gap are used as cost-based measures of reduced-model quality, because they evaluate the reduced model's cost-estimation accuracy and the full-time performance of its investment decision, respectively. Their detailed definitions are provided in Chapter 4.

2.5.2 Limitations of cost-based evaluation for storage

Although cost-based metrics are essential, they may not fully explain reduced-model behavior in storage-embedded systems. Total cost is mainly determined by investment cost, generation cost, curtailment cost, and load-shedding penalties. SOC usually does not appear as an explicit cost term; instead, it affects operation through feasibility constraints and charging/discharging availability [3, 5, 17].

As a result, two solutions can have similar total costs while relying on different storage trajectories. If the reduced model reconstructs SOC differently from the full model, the system may partly compensate through thermal redispatch, transmission flow adjustment, renewable curtailment, or load shedding. This is especially relevant for long-duration storage, where small errors in daily energy balance can accumulate over multiple days or weeks [11].

This does not weaken the importance of cost-based metrics. Rather, it means that they should be complemented by operational diagnostics when storage is a central planning resource. Cost-based evaluation answers whether the reduced model is economically competitive, while storage-aware diagnostics help explain how the reduced temporal structure represents storage operation.

2.5.3 Need for SOC-aware diagnostic evaluation and feedback

SOC-aware diagnostic evaluation is needed to identify how temporal aggregation distorts storage-state evolution. In this thesis, SOC-based metrics

are used as complementary diagnostic tools rather than replacements for cost-based metrics. They distinguish different forms of storage-state error, including total SOC deviation, intra-day shape mismatch, cross-day accumulated drift, and persistent inventory-level bias. Their formal definitions are provided in Chapter 4.

SOC-based diagnostics can also be interpreted as feedback information for TSA. Unlike input-based clustering indicators, SOC diagnostics are obtained from the behavior of the reduced and full planning models, and therefore reflect how temporal aggregation affects storage operation. Such diagnostic information can help identify storage-relevant days or SLD blocks that are poorly represented by the current aggregation. This supports a model-aware view of TSA, where representative-period selection is refined using optimization outcomes or post-solution information [6, 31, 20, 38].

Therefore, SOC-aware evaluation extends TSA assessment from economic accuracy to storage-aware diagnostic interpretation and feedback enhancement. A reduced model may perform well in terms of operational cost error or optimality gap while still producing distorted SOC trajectories; conversely, SOC diagnostics can indicate where the reduced temporal representation fails and guide feedback-enhanced TSA strategies [11, 17, 20].

2.6 Research gap and positioning of this thesis

The literature reviewed above shows that TSA has developed along several important directions. Classical clustering-based methods, such as k-means, k-medoids, and hierarchical clustering, reduce temporal complexity by grouping similar load and renewable profiles into representative periods [13, 14]. Chronology-preserving methods, such as CTPC, maintain the order of the input time series and are particularly relevant for inter-temporal constraints and storage operation [3]. Priority-based and hybrid methods further recognize that extreme periods should be preserved because they may strongly influence investment decisions and total-cost-oriented planning outcomes [10, 17]. Extreme-event studies also show that rare operating conditions may be lost in ordinary TSA, while their preservation can affect system design and cost results [41]. In addition, iterative extreme-period inclusion has been proposed as a way to refine TSA after identifying insufficiently

represented critical periods [42].

Existing studies have also improved storage representation under reduced temporal scope. Typical-period storage formulations, enhanced representative days, system-state approaches, chronological clustering, and SLD-based reconstruction address different aspects of the inter-period storage problem [5, 12, 11, 17]. In parallel, model-based evaluation metrics have been developed to assess whether reduced temporal representations produce reliable planning results, moving beyond purely input-based similarity [6, 19, 20]. More recent model-aware TSA studies further suggest that information from optimization outcomes can refine representative-period selection rather than relying only on input-profile similarity [31, 38].

Despite these advances, an important gap remains in storage-embedded planning models. Existing TSA methods can reduce temporal complexity, preserve selected extreme periods, and reconstruct part of the chronology. In particular, SLD-based reconstruction provides a practical way to link representative days across the original chronological sequence. However, the selection of extreme days is still often based on input-side or system-stress indicators, such as peak load, high net load, or high operating cost, rather than on storage-state representation itself [10, 31, 32, 17]. Therefore, it remains unclear how storage-related diagnostic information can be used to identify extreme days or periods that are more relevant to storage-embedded planning decisions.

Representative days capture intra-day profiles, and SLDs reconstruct part of the inter-day sequence, but the reduced model still replaces heterogeneous natural-day behavior with repeated representative-day patterns. As a result, SOC reconstruction errors may arise from intra-day shape mismatch, cross-day accumulated drift, and inventory-level bias. These mechanisms cannot be fully explained by cost-based metrics alone. They are also not equivalent to ordinary input-profile errors, because they are observed from the storage operation produced by the planning model.

This thesis addresses the gap by treating SOC-based metrics as diagnostic feedback signals rather than as a direct optimization objective. Cost-based metrics remain the primary criteria for judging whether a reduced planning model is economically competitive. SOC-based metrics are then used to explain why SOC trajectory gaps remain under economically competitive solutions, what these gaps are composed of, and which natural days or SLD blocks may indicate storage-relevant representative-day distortion. In this sense, the purpose of SOC diagnostics is not to guarantee a monotonic re-

duction of aggregate SOC error, but to provide storage-aware information that can guide extreme-day selection and be evaluated through total-cost performance.

This thesis therefore combines four elements. First, it adopts hybrid hierarchical clustering with explicit extreme-day preservation for representative-day selection, following the hybrid RD framework in [17]. Second, it uses sequentially linked days to reconstruct inter-day chronology for storage operation. Third, it evaluates reduced planning models through cost-based metrics to identify economically competitive solutions, and then uses SOC-based diagnostic metrics to explain storage-related error mechanisms. Fourth, it tests whether SOC-based diagnostic signals can guide extreme-day selection and improve total-cost performance. In this way, the thesis positions itself at the intersection of representative-period selection, storage-aware chronology reconstruction, model-based evaluation, and SOC-informed feedback enhancement.

Chapter 3

Formulation of the power system co-planning model

Nomenclature

Sets and indices

i, j	Bus indices
g, w, l, e	Indices for generators, wind farms, loads, and ESS
d, t	Representative day and time indices
s, k	SLD index and repetition index
G, W, W^{new}, E	Sets of generators, wind farms, candidate wind farms, and ESS
NL, EL	Sets of candidate and existing transmission lines
D, T_d^P, T_d^D	Representative days, time points, and time intervals
S	Set of sequence-length days (SLDs)

Decision variables

I_{ij}^{NL}	Binary decision for candidate transmission line
$P_{W,w}^{cap}$	Installed capacity of candidate wind farm
$E_e^{cap},\ P_{E,e}^{cap}$	Installed ESS energy and power capacity
$P_G[g,d,t],\ P_G^{res}[g,d,t]$	Generator output and reserve
$P_W[w,d,t],\ P_W^{curt}[w,d,t]$	Wind generation and curtailment
$P_L^{shed}[l,d,t]$	Load shedding
$F_{ij}[d,t],\ \theta_i[d,t]$	Line flow and voltage angle
$P_E^{ch}[e,d,t],\ P_E^{dis}[e,d,t]$	ESS charging and discharging power
$SOC[e,d,t]$	SOC trajectory within representative day
$\Delta SOC[e,d]$	Net SOC change over representative day
$SOC^{start}[e,s]$	SOC at the beginning of SLD

Parameters

$TC,\ IC,\ OC$	Total, investment, and operational cost
$P_{W,w}^{max}$	Maximum candidate wind capacity
α	Minimum wind build-out ratio
$\bar{E}_e,\ S_e^{max}$	Upper bounds of ESS energy and power capacity
$P_G^{max}[g]$	Generator capacity limit
$P_L[i,d,t]$	Load demand
$P_{line,ij}^{max}$	Line capacity limit
B_{ij}	Line susceptance
P^{base}	System base power
M_{ij}	Big-M parameter
$R_g,\ \tau$	Ramp rate and reserve response time
$\eta_e^c,\ \eta_e^d$	Charging and discharging efficiencies
$\Delta t_{d,t}$	Time interval duration
$L_s,\ d_s$	Length and mapped representative day of SLD

The formulation in this chapter adopts an existing deterministic mixed-integer linear co-planning structure and modifies it for storage-aware SOC evaluation. The transmission and wind expansion components follow the three-layer planning framework in [20], consisting of an investment layer, an operational cost layer, and a DC-OPF layer. The ESS investment and chronology-linking formulation follows the representative-day and sequentially linked day structure proposed in [17]. Therefore, the main contribution of this thesis is not a new co-planning formulation, but the use of this established model as the reference framework for evaluating SOC reconstruction errors under time-series aggregation.

The resulting model is a deterministic MILP for transmission expansion, wind generation investment, and ESS investment. Candidate transmission lines are represented by binary investment variables, while wind and storage capacities are represented by continuous investment variables. This structure is consistent with expansion-planning models that combine discrete network reinforcement decisions with continuous capacity decisions and linear operational constraints [21, 22].

3.1 Investment layer

This layer follows the transmission and wind investment structure used in [20]. The ESS energy and power capacity investment variables are incorporated following the storage co-planning formulation in [17].

The objective of the power system co-planning model is to determine the optimal investment decision set ID^* among all feasible investment options that minimizes the total system cost TC. Candidate transmission lines are modeled by binary variables, while wind and ESS investments are modeled as continuous capacity variables.

The total investment cost includes the cost of candidate transmission lines, planned wind capacity, ESS energy capacity, and ESS power capacity:

$$IC = \sum_{(i,j)\in NL} I_{ij}^{NL} \cdot IC_{ij}^{NL} + \sum_{w\in W^{new}} P_{W,w}^{cap} \cdot IC_{w}^{W} + \sum_{e\in E^{new}} E_{e}^{cap} \cdot IC_{e}^{E} + \sum_{e\in E^{new}} P_{E,e}^{cap} \cdot IC_{e}^{EP} \tag{3.1}$$

For each candidate transmission line $(i, j) \in NL$, the investment variable is:

$$I_{ij}^{NL} \in \{0, 1\} \tag{3.2}$$

For planned wind farms, the installed capacity is bounded by:

$$0 \leq P_{W,w}^{cap} \leq P_{W,w}^{max} \tag{3.3}$$

A minimum wind build-out requirement is imposed as:

$$\sum_{w\in W^{new}} P_{W,w}^{cap} \geq \alpha \sum_{w\in W^{new}} P_{W,w}^{max} \tag{3.4}$$

For candidate ESS, both installed energy and power capacities are optimized:

$$0 \leq E_{e}^{cap} \leq \bar{E}_{e} \tag{3.5}$$

$$0 \leq P_{E,e}^{cap} \leq S_e^{max} \tag{3.6}$$

Together, these constraints define the feasible investment region for transmission expansion, wind development, and ESS sizing.

3.2 Operational cost layer

The operational cost layer follows the piecewise-linear operating-cost structure adopted in [20]. The same structure is retained here and applied to the storage-embedded co-planning model.

The total operational cost OC is defined as the weighted sum of operating cost rates over all representative days and time intervals:

$$\begin{aligned} OC = \sum_{d \in D} \omega_d \sum_{t \in T_d^D} \Delta t_{d,t} \Bigg[& \sum_{g \in G} \left(C_G^1[g,d,t] + C_G^2[g,d,t] \right) \\ & + \sum_{l \in L} C_L^{shed}[l,d,t] + \sum_{w \in W} C_W^{curt}[w,d,t] \Bigg] \end{aligned} \tag{3.7}$$

where ω_d denotes the weight of representative day d.

Thermal generation cost The following piecewise-linear generation-cost constraints are adapted from [20]. The quadratic generation cost is represented by a linear lower convex envelope to keep the model within a MILP structure [23, 36].

The thermal generation cost is decomposed into two linear components:

$$C_G[g,d,t] = C_G^1[g,d,t] + C_G^2[g,d,t], \qquad \forall g \in G,\ d \in D,\ t \in T_d^D \tag{3.8}$$

For each breakpoint $k \in K$, the first cost component is constrained by:

$$\frac{C_G^1[g,d,t]}{\Delta t_{d,t}} \geq (c_{g,2}\pi_{g,k} + c_{g,1}) \cdot \frac{P_G[g,d,t] + P_G[g,d,t+1]}{2} - \frac{1}{2} c_{g,2} \pi_{g,k}^2 \tag{3.9}$$

The second component is represented by:

$$\frac{C_G^2[g,d,t]}{\Delta t_{d,t}} \geq \frac{c_{g,2}\pi_{g,k}}{12} \left(P_G[g,d,t] - P_G[g,d,t+1] \right) - \frac{1}{24} c_{g,2} \pi_{g,k}^2 \tag{3.10}$$

$$\frac{C_G^2[g,d,t]}{\Delta t_{d,t}} \geq -\frac{c_{g,2}\pi_{g,k}}{12}\left(P_G[g,d,t] - P_G[g,d,t+1]\right) - \frac{1}{24}c_{g,2}\pi_{g,k}^2 \qquad (3.11)$$

where the breakpoint is defined as:

$$\pi_{g,k} = \frac{k}{|K|-1}P_G^{max}[g], \qquad \forall g \in G,\ k \in K \qquad (3.12)$$

Load shedding cost Load shedding is penalized using the value of lost load:

$$C_L^{shed}[l,d,t] = VOLL[l]\cdot\frac{P_L^{shed}[l,d,t] + P_L^{shed}[l,d,t+1]}{2}\cdot\Delta t_{d,t}, \quad \forall l \in L,\ d \in D,\ t \in T_d^D \qquad (3.13)$$

Wind curtailment cost Wind curtailment cost is expressed as:

$$C_W^{curt}[w,d,t] = C^{WC}[w]\cdot\frac{P_W^{curt}[w,d,t] + P_W^{curt}[w,d,t+1]}{2}\cdot\Delta t_{d,t}, \quad \forall w \in W,\ d \in D,\ t \in T_d^D \qquad (3.14)$$

Total objective function The overall objective is:

$$TC = IC + OC \qquad (3.15)$$

3.3 DC-OPF layer

The DC-OPF layer is adapted from the planning formulation in [20] and extended by adding ESS charging and discharging terms to the nodal power balance. It determines thermal generation dispatch, wind dispatch, transmission flows, load shedding, and ESS operation for each representative day and time point.

A DC-OPF approximation is adopted to keep the operational model linear and computationally tractable. Compared with AC-OPF, DC-OPF neglects voltage-magnitude variations, reactive power, and network losses, and represents active-power flows through voltage-angle differences and line susceptances. This approximation is widely used in planning-level optimization models because it provides a tractable representation of network constraints [35, 36].

Power balance constraints At each bus i, nodal power balance is enforced as:

$$\begin{aligned}\sum_{g\in G_i} P_G[g,d,t] + \sum_{w\in W_i} P_W[w,d,t] + \sum_{e\in E_i} \left(P_E^{dis}[e,d,t] - P_E^{ch}[e,d,t]\right) \\ + \sum_{\ell\in L_i^{in}} F_\ell[d,t] - \sum_{\ell\in L_i^{out}} F_\ell[d,t] = P_L[i,d,t] - P_L^{shed}[i,d,t]\end{aligned} \tag{3.16}$$

DC power flow constraints For each existing transmission line $(i,j) \in EL$, the DC angle-flow relation is:

$$\frac{F_{ij}[d,t]}{P^{base}} = B_{ij}\left(\theta_i[d,t] - \theta_j[d,t]\right), \qquad \forall (i,j)\in EL,\ d\in D,\ t\in T_d^P \tag{3.17}$$

Reference bus constraint One bus is selected as the reference bus. In this case study, the reference bus is bus 12:

$$\theta_{ref}[d,t] = 0, \quad ref = 12,\ \forall d\in D,\ t\in T_d^P \tag{3.18}$$

Transmission line constraints For existing transmission lines, the flow limit is:

$$-P_{line,ij}^{max} \le F_{ij}[d,t] \le P_{line,ij}^{max}, \qquad \forall (i,j)\in EL,\ d\in D,\ t\in T_d^P \tag{3.19}$$

For candidate transmission lines, the flow is coupled with the investment variable:

$$-P_{line,ij}^{max} I_{ij}^{NL} \le F_{ij}[d,t] \le P_{line,ij}^{max} I_{ij}^{NL}, \qquad \forall (i,j)\in NL,\ d\in D,\ t\in T_d^P \tag{3.20}$$

The candidate-line angle-flow relation is enforced through a Big-M formulation [22]:

$$\frac{F_{ij}[d,t]}{P^{base}} - B_{ij}\left(\theta_i[d,t] - \theta_j[d,t]\right) \le \left(1 - I_{ij}^{NL}\right) M_{ij}, \qquad \forall (i,j)\in NL,\ d\in D,\ t\in T_d^P \tag{3.21}$$

$$\frac{F_{ij}[d,t]}{P^{base}} - B_{ij}\left(\theta_i[d,t] - \theta_j[d,t]\right) \ge -\left(1 - I_{ij}^{NL}\right) M_{ij}, \qquad \forall (i,j)\in NL,\ d\in D,\ t\in T_d^P \tag{3.22}$$

Wind generation constraints Wind generation is limited by the wind availability factor. For existing wind farms, the dispatched wind power is bounded by the existing installed capacity:

$$0 \leq P_W[w,d,t] \leq \overline{P}_W[w]\phi_W[d,t], \quad \forall w \in W^{old},\ d \in D,\ t \in T_d^P \tag{3.23}$$

For candidate wind farms, the dispatched wind power is bounded by the optimized installed wind capacity:

$$0 \leq P_W[w,d,t] \leq P_W^{cap}[w]\phi_W[d,t], \quad \forall w \in W^{new},\ d \in D,\ t \in T_d^P \tag{3.24}$$

Wind curtailment is defined as the difference between available wind power and dispatched wind power:

$$\begin{aligned} P_W^{curt}[w,d,t] = P_W^{avail}[w,d,t] - P_W[w,d,t], \\ \forall w \in W,\ d \in D,\ t \in T_d^P \end{aligned} \tag{3.25}$$

For compact notation, $P_W^{avail}[w,d,t]$ denotes the available wind power of wind farm w. For existing wind farms, it is given by $\bar{P}_{W,w}\phi_W[d,t]$, while for candidate wind farms, it is given by $P_{W,w}^{cap}\phi_W[d,t]$.

Thermal generator reserve requirement Thermal generation output and reserve are constrained by:

$$P_G[g,d,t] \leq P_G^{\max}[g] - P_G^{res}[g,d,t], \qquad \forall g \in G,\ d \in D,\ t \in T_d^P \tag{3.26}$$

$$P_G^{res}[g,d,t] \leq P_G[g,d,t], \qquad \forall g \in G,\ d \in D,\ t \in T_d^P \tag{3.27}$$

Following the "3+5 rule", the system reserve requirement is represented as 3% of load demand plus 5% of wind generation [27]:

$$\sum_{g \in G} P_G^{res}[g,d,t] = 0.03 \sum_{l \in L} P_L[l,d,t] + 0.05 \sum_{w \in W} P_W[w,d,t], \qquad \forall d \in D,\ t \in T_d^P \tag{3.28}$$

Ramping constraints with reserve delivery are:

$$\frac{P_G[g,d,t+1] - P_G[g,d,t]}{\Delta t_{d,t}} + \frac{P_G^{res}[g,d,t]}{\tau} \leq R_g, \qquad \forall g \in G,\ d \in D,\ t \in T_d^D \tag{3.29}$$

$$\frac{P_G[g,d,t] - P_G[g,d,t+1]}{\Delta t_{d,t}} + \frac{P_G^{res}[g,d,t]}{\tau} \leq R_g, \qquad \forall g \in G,\ d \in D,\ t \in T_d^D \tag{3.30}$$

$$\frac{P_G[g,d,t+1] - P_G[g,d,t]}{\Delta t_{d,t}} + \frac{P_G^{res}[g,d,t+1]}{\tau} \leq R_g, \qquad \forall g \in G,\ d \in D,\ t \in T_d^D \tag{3.31}$$

$$\frac{P_G[g,d,t] - P_G[g,d,t+1]}{\Delta t_{d,t}} + \frac{P_G^{res}[g,d,t+1]}{\tau} \leq R_g, \qquad \forall g \in G,\ d \in D,\ t \in T_d^D \tag{3.32}$$

3.4 Energy storage and chronology layer

The ESS and chronology layer is based on the representative-day and sequentially linked day formulation proposed in [17]. Representative days describe intra-day ESS operation, while SLDs are used to propagate SOC across consecutive natural days mapped to the same representative day. This thesis retains this formulation and later evaluates the SOC reconstruction errors introduced by the RD-SLD approximation.

In the present case study, no existing ESS is included in the input data, and all storage units are candidate investments.

ESS charging and discharging limits For each ESS unit e, charging and discharging power are constrained by:

$$0 \leq P_E^{ch}[e,d,t] \leq P_{E,e}^{cap}, \qquad \forall e \in E,\ d \in D,\ t \in T_d^P \tag{3.33}$$

$$0 \leq P_E^{dis}[e,d,t] \leq P_{E,e}^{cap}, \qquad \forall e \in E,\ d \in D,\ t \in T_d^P \tag{3.34}$$

The installed ESS power and energy capacities are bounded by:

$$0 \leq P_{E,e}^{cap} \leq S_e^{\max}, \qquad \forall e \in E \tag{3.35}$$

$$0 \leq E_e^{cap} \leq \bar{E}_e, \qquad \forall e \in E \tag{3.36}$$

SOC dynamics within a representative day The SOC transition within a representative day is:

$$\begin{aligned} SOC[e,d,t+1] = SOC[e,d,t] \\ + \frac{\Delta t_{d,t}}{2}\Bigg(\eta_e^c\big(P_E^{ch}[e,d,t] + P_E^{ch}[e,d,t+1]\big) \\ - \frac{1}{\eta_e^d}\big(P_E^{dis}[e,d,t] + P_E^{dis}[e,d,t+1]\big)\Bigg) \end{aligned} \tag{3.37}$$

The representative-day SOC trajectory is defined relative to its starting point:

$$SOC[e,d,0] = 0, \qquad \forall e \in E,\ d \in D \tag{3.38}$$

The net SOC change over one representative day is:

$$\Delta SOC[e,d] = SOC[e,d,T_d^{end}], \qquad \forall e \in E,\ d \in D \tag{3.39}$$

Relative SOC excursion within a representative day The intra-day SOC excursion is bounded by:

$$SOC^{rel,min}[e,d] \geq -SOC[e,d,t], \qquad \forall e \in E,\ d \in D,\ t \in T_d^P \tag{3.40}$$

$$SOC^{rel,max}[e,d] \geq SOC[e,d,t], \qquad \forall e \in E,\ d \in D,\ t \in T_d^P \tag{3.41}$$

Here, $SOC^{rel,min}[e,d]$ and $SOC^{rel,max}[e,d]$ describe the maximum downward and upward SOC deviations from the start-of-day SOC within representative day d. They are used to ensure that, after the representative-day SOC profile is repeated in an SLD block, the absolute SOC remains within feasible energy bounds.

SLD chronology linking Let $SOC^{start}[e,s]$ denote the absolute SOC at the beginning of SLD s. Each SLD has length L_s and is mapped to representative day d_s. The chronological link between adjacent SLDs is:

$$SOC^{start}[e,s{+}1] = SOC^{start}[e,s] + L_s \cdot \Delta SOC[e,d_s], \qquad \forall e \in E,\ s \in S \backslash \{s^{last}\} \tag{3.42}$$

The annual cyclic condition is:

$$SOC^{start}[e,s_{\text{first}}] = SOC^{start}[e,s_{\text{last}}] + L_{s_{\text{last}}} \cdot \Delta SOC[e,d_{s_{\text{last}}}], \qquad \forall e \in E \tag{3.43}$$

SOC feasibility over repeated days For the k-th repeated day inside SLD s, the lower and upper SOC bounds are:

$$SOC^{start}[e,s]+k\cdot\Delta SOC[e,d_s]-SOC^{rel,min}[e,d_s]\geq 0, \qquad \forall e\in E,\ (s,k)\in S^{rep} \tag{3.44}$$

$$SOC^{start}[e,s]+k\cdot\Delta SOC[e,d_s]+SOC^{rel,max}[e,d_s]\leq E_e^{cap}, \qquad \forall e\in E,\ (s,k)\in S^{rep} \tag{3.45}$$

This layer combines standard ESS charging/discharging constraints with SLD chronology reconstruction, allowing the reduced model to represent both intra-day SOC movement and inter-day SOC continuity.

3.5 Test system and input data

The case study is based on a modified RTS-24-bus test system. The network data, including buses, generators, loads, existing transmission lines, candidate transmission lines, and wind investment candidates, are taken from the modified RTS-24-bus dataset provided by Moradi-Sepahvand [34]. This network is used as the base system for the co-planning model formulated in this chapter.

Candidate ESS data are added based on the storage data used by Moradi-Sepahvand and Tindemans [17]. The ESS set contains both daily-cycling and long-duration storage candidates, which are distinguished later in the SOC diagnostic analysis.

The chronological input data use one-year hourly load and wind profiles for the Netherlands in 2019. The load data are obtained from the annual peak load and load curve published by TenneT [33], while the wind profile follows the Netherlands wind-factor dataset used in [20]. The hourly profiles are normalized into load and wind factors and then applied to the bus-level load and wind capacity data. The full year is divided into 365 natural days, with each day represented by 24 hourly intervals and 25 time points in the piecewise-linear formulation.

3.6 Full-time model without time-series aggregation

Before applying TSA, the planning model is solved using the original one-year hourly time series. This case is referred to as the full chronological planning model. It has the same investment variables, operational constraints, and objective-function structure as the reduced representative-day model, including transmission expansion, wind investment, ESS sizing, DC-OPF constraints, reserve and ramping constraints, and SOC dynamics. The difference is in the temporal representation: the full model uses all 365 natural days and hourly time steps directly, while the reduced model replaces the original year with a smaller set of representative days and reconstructs inter-day SOC chronology through SLD constraints. Therefore, the full model avoids temporal aggregation error but introduces a much larger number of operational variables and constraints, making it computationally difficult to solve to global optimality.

The solver log is summarized in Table 3.1. After 26143.02 seconds, the solver returned a feasible incumbent solution with an objective value of 2.4508×10^9. The corresponding best bound was 2.3608×10^9, resulting in a remaining optimality gap of 3.6999%. Therefore, the reported full-time solution is used as a computational reference point rather than as a proven global optimum.

Table 3.1: Solver result of the full-time model without TSA.

Item	Value
Solution time	26143.02 s
Explored nodes	4
Simplex iterations	1,874,601
Incumbent objective	$2.450757901945 \times 10^9$
Best bound	$2.36081628833 \times 10^9$
MIP gap	3.6999%
Threads	32

This result provides an anchor for the following evaluation. It confirms the computational burden of the full chronological formulation and provides a reference incumbent and best bound for comparing reduced models. In later chapters, representative-day models with different NRD values and extreme-

day strategies are evaluated against this reference.

Chapter 4

Evaluation of Simplified Planning Models with Storage Dynamics

4.1 Overview

The co-planning model developed in Chapter 3 incorporates energy storage systems (ESS) and inter-day chronological constraints through the sequentially-linked-day (SLD) formulation. This enables reduced planning models to retain part of the intertemporal structure of storage operation while significantly lowering computational complexity through representative days (RDs)[needs moradi citation]. However, this simplification inevitably alters the original chronology and introduces approximation in the reconstructed storage-state trajectories.

The analysis framework follows the model-based perspective in the TSA literature. Representative-day approaches are widely used in long-term power system modeling because full hourly resolution can become computationally restrictive [28]. Additionally, aggregation choices, such as clustering methodology, chronology representation, and extreme days preservation in this thesis, can affect planning outcomes, especially when operational variability, storage, and chronological constraints are important [29, 30]. Therefore, reduced models should not be evaluated only by input similarity, but also by their effect on optimization outcomes, such as objective-function error [6, 19]. Based on this outcome-oriented evaluation perspective, this thesis further

introduces reconstructed SOC trajectory metrics as a storage-specific diagnostic dimension for identifying SOC representation errors and supporting extreme-day feedback.

This chapter starts with evaluating the economic performance of current reduced planning models: since the objective of the co-planning model is to obtain low full-time total cost under a model-derived investment decision, cost-based metrics, especially the optimality gap, continue to serve as the primary criteria for finding economically competitive planning results. Then a competitive planning result within 20-50 range of representative days numbers is set as reference and diagnostic object.

Existing studies on storage under reduced temporal scope show that the representation of inter-period storage states is sensitive to chronology loss and temporal aggregation [5, 11, 12]. It can be inferred that storage states representation error, after certain data processing, can provide useful information to improve time-series-aggregation(TSA), and thus planning results. For this reason, this chapter introduces a complementary SOC-based metrics as a diagnostic layer for examining storage-state reconstruction gaps under reference planning results.

The chapter is organized as follows. Section 4.2 first presents the conventional cost-based evaluation results, including operational estimation error and optimality gap. Then the competitive planning results of NRD=21 is confirmed to be the reference diagnostic object. Section 4.3 then shows that noticeable SOC trajectory discrepancies may still remain within these economically competitive results, and positions SOC gap as an informative diagnostic tool instead of improvement target. Section 4.4 discusses the implications of SOC misrepresentation from a model-based and TSA perturbing perspective: infers that the reduced model SOC representation pattern has influence on investment decision planning. Finally, Section 4.5 defines a set of SOC-based diagnostic metrics designed to characterize different components of the SOC reconstruction gap. Together, these analysis establish a cost-oriented but storage-aware evaluation framework: cost-based metrics identify economically competitive planning result and set it as reference, while SOC-based metrics explain the storage-related errors inside those results and provide the basis for the diagnostic and feedback analysis in the following chapters.

4.2 Cost-based evaluation of planning results

Cost-based evaluation is used because it measures the impact of temporal aggregation in the same domain as the planning objective. Bahl et al. argue that TSA errors should be evaluated in the objective-function domain rather than only through statistical similarity of the aggregated time series [6]. Studies on electricity-sector model resolution also show that temporal and operational resolution can affect investment outcomes and scenario conclusions [19, 29]. Therefore, operational estimation error and optimality gap are used in this chapter to identify economically competitive planning results of the reduced model. These metric results are computed based on the co-planning model introduced in Chapter 3.

4.2.1 Definition of cost-based metrics

The two cost-based metrics used in this chapter distinguish between cost-estimation accuracy and investment-decision quality. Operational estimation error evaluates whether the reduced temporal representation estimates the operating cost of a fixed investment decision accurately. Optimality gap, by contrast, evaluates whether the investment decision obtained from the reduced model remains close to the best available full-space benchmark. This distinction is consistent with model-based TSA evaluation, where a reduced model may estimate costs inaccurately, select a suboptimal investment plan, or both [6, 20]. The definitions of operational estimation error and optimality gap used in this thesis follow the cost-based evaluation framework adopted in [20].

Let $\hat{ID}$ denote the investment decision obtained from the RD-based planning model, and ID^* denote the optimal investment decision obtained from the full-space model. Following [20], the following cost functions are defined:

- $TC_{RD}[\cdot]$: total system cost evaluated using representative days
- $TC_{Full}[\cdot]$: total system cost evaluated using the full time series

Operational estimation error The operational estimation error measures the discrepancy between the reduced model and the full-time model under a fixed investment decision [20]:

$$\Delta OC[\hat{ID}] = TC_{RD}[\hat{ID}] - TC_{Full}[\hat{ID}] \tag{4.1}$$

This metric reflects how accurately the reduced model estimates system operation when compared to the full chronological model. With hierarchical clustering, the operational estimation error in this study is always negative.

Optimality gap This metric directly measures the loss in economic optimality caused by selecting a less optimal investment decision [20].

$$\Delta TC_{Full}[\hat{ID}] = TC_{Full}[\hat{ID}] - TC_{Full}[ID^*] \tag{4.2}$$

Specifically, $TC_{Full}[ID^*]$ is substituted by the best bound obtained from the solver, namely 2360M, rather than a proven global optimum, since the global optimum is intractable even in the simplified test system. The optimal total full cost is expected to be the smallest, so the optimality gap is always positive. Therefore, the optimality gap is used as the main metric for evaluating TSA and investment-decision performance.

4.2.2 Results under varying number of representative days

To include the effect of temporal resolution, the number of representative days (NRD) is varied from 20 to 50. For each NRD configuration, the RD-based planning model described in Chapter 3 is solved to obtain the corresponding investment decision $\hat{ID}$. The resulting solutions are then evaluated using both the reduced model and the full time series model.

For each NRD, the following metrics are computed:

- Optimality gap $\Delta TC_{Full}[\hat{ID}]$
- Operational estimation error $\Delta OC[\hat{ID}]$

The optimality gap and operational estimation error capture the economic performance of the reduced model. The results are summarized in Fig. 4.1, where the variation of these metrics with respect to NRD is illustrated.

Figure 4.1 presents the variation of two cost-based performance metrics—*operational estimation error* and *optimality gap*—with respect to the number of representative days (NRD). The TSA setting is presented below: 1) Extreme day selection: The 21st natural day, the day with peak load. 2) Clustering method: Hybrid hierarchical clustering which emphasize extreme days incorporation. 3) Chronology preservation method: Sequentially linked days

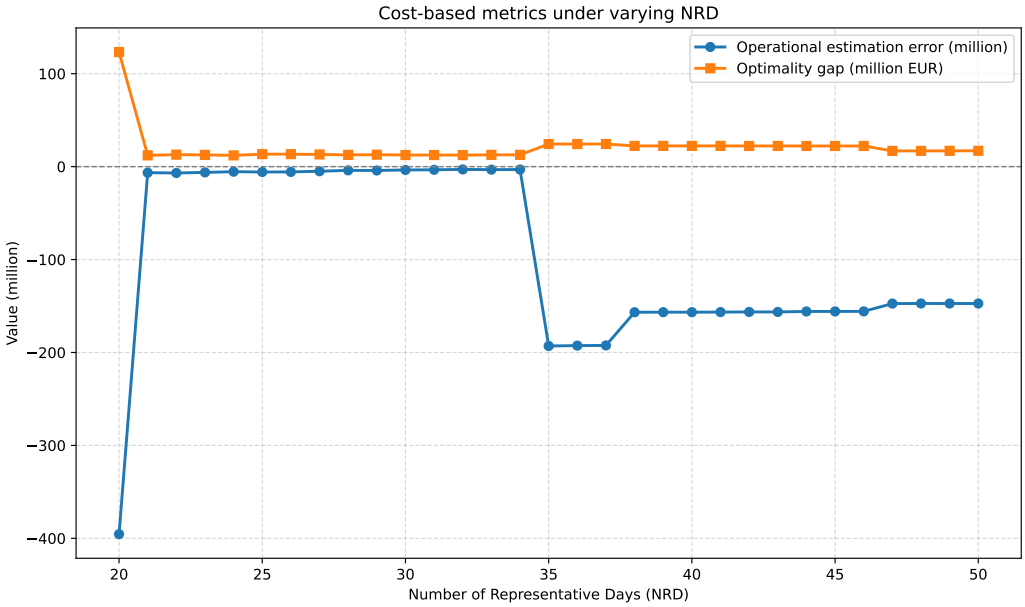

Figure 4.1: Cost-based metrics under varying NRD.

monitoring energy shift. As shown in the figure, both metrics exhibit similar trends and display a clear stepwise pattern across different NRD values.

A notable observation is that in the range of NRD = 21 to 34, both metrics remain relatively low and stable. The small operational estimation error indicates that within this interval, the reduced model is able to approximate the full time-series model with relatively high accuracy. The small optimality gap also indicates that the derived total cost is closer to the best bound.

The non-monotonic behavior observed across NRD values is consistent with the fact that representative-period selection does not only depend on the number of periods, but also on which operating conditions are preserved. Previous studies have shown that representative periods must capture not only average profiles, but also periods that are important for investment decisions, such as high-demand, low-renewable, or high net-load conditions [13, 10, 17]. Therefore, performing one more or less clustering may alter critical cluster characteristics and lead the model to derive different investment decisions.

From a cost-based metrics perspective, a pronounced difference can be observed between the solutions at NRD = 20 and NRD = 21. This large

deviation contributes to the stepwise pattern observed in Fig. 4.1, where the optimality gap does not decrease smoothly with increasing NRD, as commonly expected in traditional grid planning studies. Instead, the results exhibit non-monotonic behavior, which is illustrated in Chapter 6 to result from the effect of adding extreme days in storage-inclusive planning problems.

At NRD = 20, the model yields a relatively low investment cost, approximately EUR 676 million, but this results in a significantly higher operational cost, around EUR 1.81 billion. In contrast, increasing to NRD = 21 leads to a substantial rise in investment cost, approximately EUR 873 million, accompanied by a significant reduction in operational cost, approximately EUR 1.49 billion.

This difference is reflected in the underlying investment structure. At NRD = 20, only a single transmission line is constructed, four wind farms are built, and ESS deployment is limited to four buses, namely buses 3, 4, 5, and 7. In contrast, at NRD = 21, an additional transmission line is introduced, wind capacity becomes more diversified, and ESS is expanded to five buses, namely buses 1, 3, 4, 5, and 7, indicating improved spatial flexibility.

Overall, the comparison between NRD = 20 and NRD = 21 highlights a strong sensitivity of the reduced model to the clustering process. Due to the agglomerative nature of hierarchical clustering, one additional merging step is conducted when moving from NRD = 21 to NRD = 20. As a result, at NRD = 20, the model invests less in transmission, wind generation, and ESS, leading to an underestimation of system capacity requirements.

In summary, the results suggest that:

- With the selected TSA methodology, the reduced model achieves its best performance when NRD is between 21 and 34 within the 20–50 NRD range. In this interval, both cost-based metrics remain low and stable, indicating a good approximation of the full time-series model and a few competitive model planning results.

- The transition from NRD = 21 to NRD = 20 reveals a structural shift rather than a smooth improvement, highlighting the sensitivity of the reduced model to the clustering process. Thus editing on the clustering process/TSA also has dominant influence on planning results, including total cost and investment decisions.

- The scenario of NRD = 21 derives a small total cost while saving variables compared with the range NRD = 21–34, so it is selected as the reference planning results in this thesis and support further SOC-related diagnosis.

4.3 SOC trajectory discrepancies under economically competitive planning results

In storage-embedded co-planning problem, storage state representation requires separate attention because storage operation is governed by inter-temporal state dynamics, which is monitored by SLD chronology preservation in this thesis. In representative-period models, storage behavior may be approximated within each representative period, but the continuity of stored energy across periods is more difficult to preserve. Kotzur et al. address this issue by separating intra-period and inter-period storage states, while Gonzato et al. show that reduced temporal representations may lead to over- or underinvestment in long-duration storage if inter-period arbitrage is not properly captured [5, 11]. Therefore, the inter-day energy shifting pattern needs evaluation, as it manifests continuity of stored energy across periods and relate to model investment decisions.

In order to characterize storage-state reconstruction gaps under the RD-SLD approximation, it is necessary to compare the state-of-charge (SOC) trajectories obtained from the reduced model with those from the full model under the same fixed investment decision.

In the full model, SOC is directly obtained by sequentially solving the dispatch problem over the complete chronological time series. The SOC trajectory is therefore generated through exact inter-temporal coupling:

$$SOC_t^{full} = SOC_{t-1}^{full} + \Delta SOC_t^{full} \tag{4.3}$$

which reflects the chronological evolution of storage inventory over time.

Differently, the reduced model reconstructs the SOC trajectory using representative days (RDs) and sequence linking days (SLDs). The absolute SOC at time t, corresponding to the k-th repetition of representative day d within SLD block s, is reconstructed as:

$$SOC_{s,k,t}^{red} = SOC_s^{start} + k \cdot \Delta SOC_d + SOC_{d,t}^{RD} \tag{4.4}$$

where:

- SOC_s^{start} is the starting SOC of SLD block s,
- ΔSOC_d is the net daily SOC change associated with representative day d,
- $SOC_{d,t}^{RD}$ is the intra-day SOC profile of the representative day, defined relative to zero:

$$SOC_{d,0}^{RD} = 0 \tag{4.5}$$

This reconstruction replaces the heterogeneous day-by-day SOC evolution of the full model with repeated application of a single representative-day pattern and a fixed daily net change within each SLD block. In the full model, natural days may exhibit different intra-day charging and discharging shapes as well as different net daily SOC changes. Under the RD-SLD approximation, these heterogeneous patterns are compressed into a limited set of day templates and then propagated chronologically through block linking. As a consequence, considerable SOC reconstruction error is structurally inevitable under fixed computation budget.

With SOC trajectories created for both the reduced model and the full model, the overall SOC MAE and RMSE data across NRD range of 20-50, which provide an overall indication of deviations in storage behavior between reduced model and full model, can be easily derived for visualization. Here, overall SOC MAE denotes the average absolute pointwise difference between reconstructed and full SOC trajectories.

As shown in Fig. 4.2, both the SOC MAE and RMSE exhibit persistent fluctuations across the investigated NRD interval, without showing a clear convergence trend. This behavior differs from that of the cost-based metrics in Fig. 4.1 while sharing the same NRD range, as the operational estimation error and optimality gap remain relatively low and stable in the range of NRD = 21 to 34.

Within the NRD range of 21-34, it is observed that considerable and fluctuating overall SOC gap remains under economically competitive planning results. Combining with finding that considerable SOC reconstruction error is structurally inevitable under limited computation budget, the overall SOC gap is not an appropriate objective to improve or minimize.

Although the SOC gap at the system level lacks significance, the SOC trajectories with respect to time for the single planning result provide more information related to clustering process and chronology preservation. Under a fixed investment decision, reduced model and full time series model develop

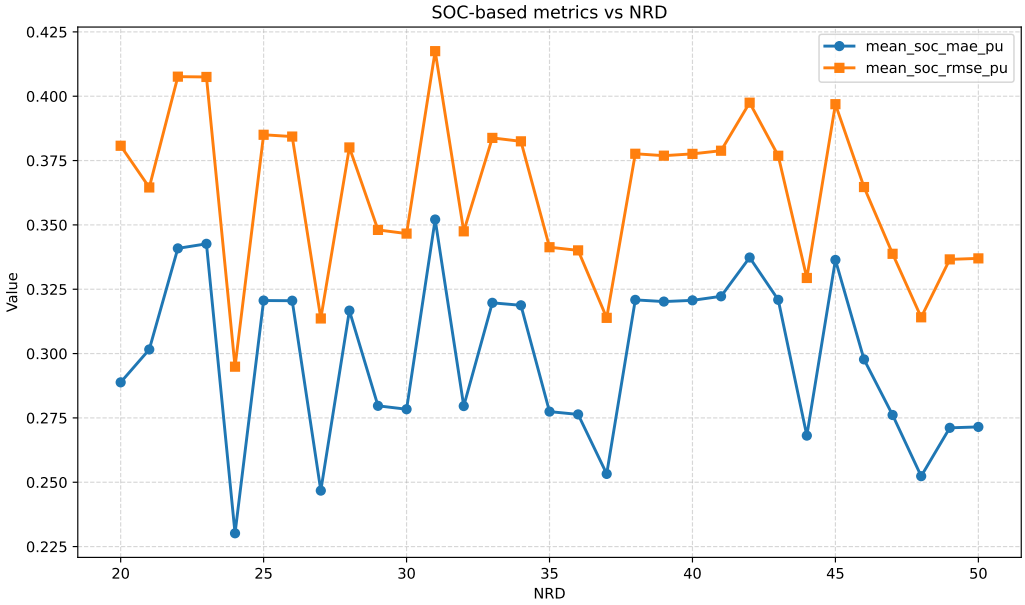

Figure 4.2: SOC-based metrics under varying NRD.

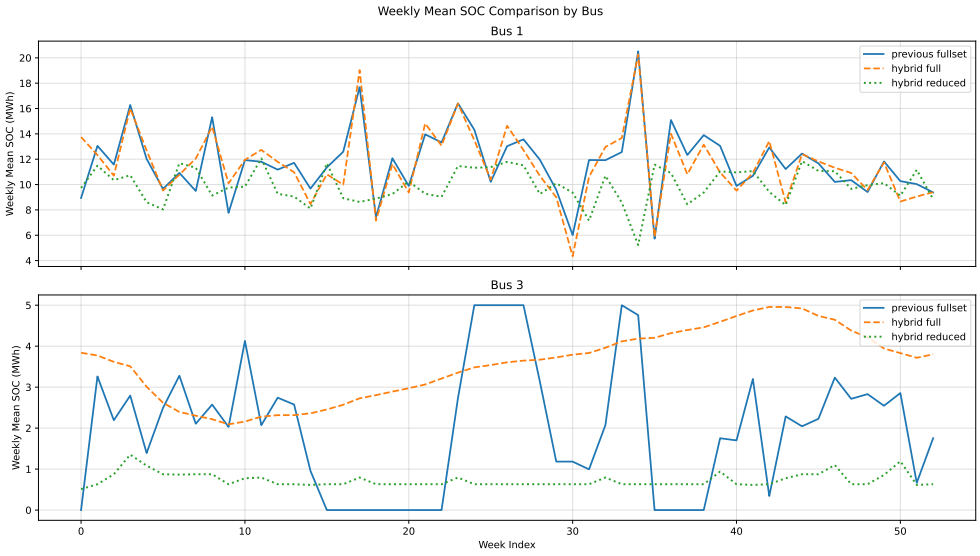

Figure 4.3: Weekly mean SOC comparison at Bus 1 and Bus 3.

distinct trajectories across the whole year, revealing energy-distorted days and periods.

For the reference model NRD = 21, where cost-based metrics indicate good economic performance, a comparison of the yearly reconstructed SOC trajectories is presented in Fig. 4.3. Bus1 and Bus3 respectively represent short-term storage and long-term storage pattern.

Key observations:

1) Different clustering process and investment decisions both lead to distinct SOC trajectories.

2) Comparing the reduced model and the full model, some periods or days correspond to huge point-wise SOC gap, meaning storage state is badly represented over these critical periods.

3) Short-term storage deviates from the full-model trajectory mainly through differences in local variation and weekly shape, and long-term storage has a sustained offset in storage level over time, indicating that TSA has different error mechanisms on different storage types.

The purpose of comparing SOC trajectories here is therefore not to evaluate the reduced model by requiring pointwise agreement with the full-model SOC, but to disclose what kinds of critical storage-state distortion remain under an economically competitive planning results.

For this reason, SOC-based analysis is introduced in this thesis not as an alternative optimization objective, but as a diagnostic perspective. It helps explain how temporal reduction and sequentially linked days distort storage-state evolution, why reduced-model and full-model SOC curves differ, and which periods are most informative for identifying storage-relevant extreme days or chronology-sensitive blocks. This diagnostic role becomes especially important in a system containing both daily-cycling storage and long-duration hydro storage, where TSA may affect different storage types through different dominant error mechanisms.

4.4 Implications of SOC misrepresentation in reduced models

Section 4.2 mentioned a strong sensitivity of the investment decision to the clustering process by comparing the planning results of NRD=20 and NRD=21. Additionally, Figure 4.3 demonstrates that different clustering

process results in distinct SOC trajectories by comparing SOC trajectories of the reduced time series and full time series.

From this perspective, it can be inferred that editing on clustering process, including adding extreme days, modifying cluster characteristics, and changing SLD blocks sequence, can lead to changing investment decisions and different SOC trajectories.

This section discusses the relationship between SOC gap, clustering process, and investment result from a model-based and TSA perturbing perspective.

In the co-planning model introduced in Chapter 3, the total cost is defined as

$$TC = IC + OC, \tag{4.6}$$

where IC is the investment cost and OC is the operational cost. The operational cost is mainly determined by thermal generation cost and load-shedding penalty. In simplified form, it can be written as

$$OC = \sum_{i,t} \left(c_g P_{g,i,t} + c_{LS} LS_{i,t} \right), \tag{4.7}$$

where $P_{g,i,t}$ denotes thermal generation and $LS_{i,t}$ denotes load shedding. SOC itself does not appear as an explicit cost term in the objective function. Instead, it affects the feasible operation of the system through storage charging and discharging constraints.

The role of SOC is therefore embedded in the operational feasibility constraints. The nodal power balance can be expressed as

$$\sum_{g \in G_i} P_{g,i,t} + P^{dis}_{E,i,t} - P^{ch}_{E,i,t} + \sum_{j:(j,i)\in\mathcal{L}} F_{j,i,t} - \sum_{j:(i,j)\in\mathcal{L}} F_{i,j,t} + LS_{i,t} = D_{i,t}, \quad \forall i, t, \tag{4.8}$$

while the storage dynamics are governed by

$$SOC_{i,t} = SOC_{i,t-1} + \eta_c P^{ch}_{E,i,t} - \frac{1}{\eta_d} P^{dis}_{E,i,t}. \tag{4.9}$$

When the SOC value at time t is underestimated as observed in Figure 4.3, flexibility from line power, thermal generator is required to support the load demand. Thus, the system stress increases.

$$\Delta P_{g,i,t} + \sum_{j:(j,i)\in\mathcal{L}} \Delta F_{j,i,t} - \sum_{j:(i,j)\in\mathcal{L}} \Delta F_{i,j,t} + \Delta LS_{i,t} \approx -\Delta P^{dis}_{E,i,t} + \Delta P^{ch}_{E,i,t}, \tag{4.10}$$

where the symbol ≈ emphasizes that this is a mechanism-based interpretation rather than a strict accounting identity.

More than that, underestimated SOC trajectories directly refers to underestimated ESS economic value, as the planned ESS support is weaker in the reduced model than it is in full time series. In this way, the model tend to build more energy storage device. So, from the model-based perspective: SOC trajectories have influence on the investment decision, and most severe periods should support extreme periods selection.

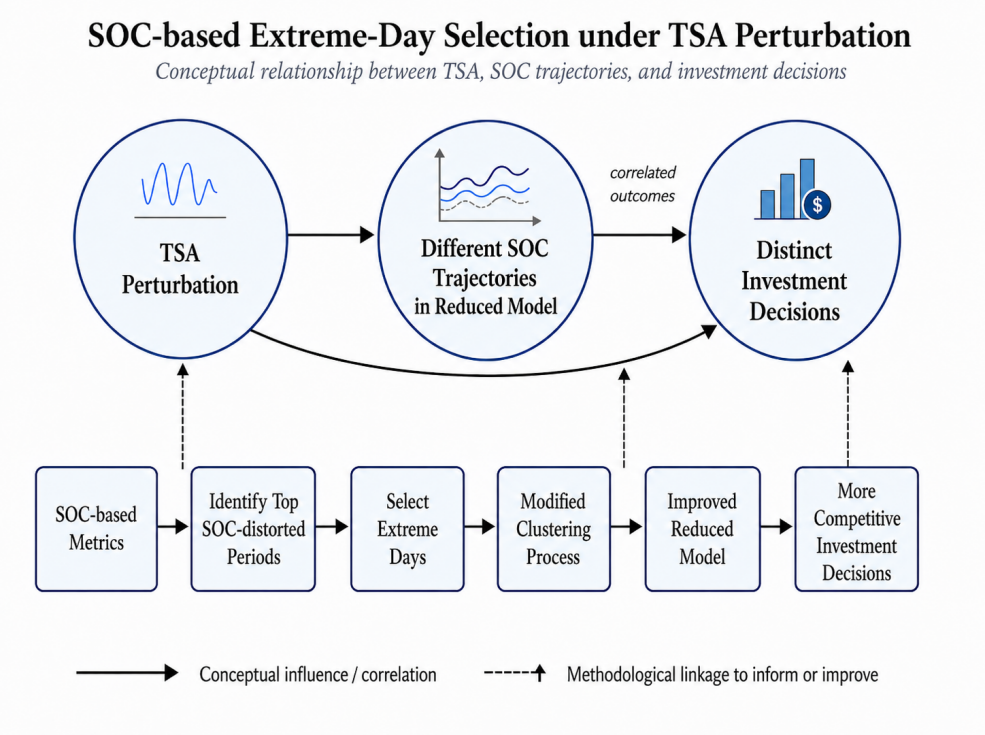

Figure 4.4: SOC-based extreme-day selection workflow under TSA perturbation.

Another perspective is Time-Series-Aggregation(TSA) perturbing. Changing the TSA leads to different SOC trajectories in the reduced model and also the distinct investment decisions that correlates with each other. So the methodology in the following parts is: extreme days are selected from SOC-based metrics, and the modified clustering process make the reduced model derive better investment decisions. In the reference case, the top few

SOC-level-distorted periods forms the extreme days, and more competitive investment decisions are expected.

4.5 SOC-based evaluation metrics

According to the formula 4.3 and 4.4, the SOC gap between the reduced model and full-time model propagates with time, which means estimation error in time point, day, and period would pass across to the whole time span. For diagnostic purpose, the metrics below are designed to decompose the SOC reconstruction gap into interpretable components, so that the dominant error mechanism can be identified. In this study, three structurally distinct error components are distinguished:

- **Intra-day shape error:** deviation in the relative SOC trajectory within each day due to imperfect RD representation;
- **Cross-day accumulated drift:** systematic accumulation of daily net SOC errors due to approximating varying daily changes by a fixed ΔSOC_d;
- **Inventory level bias:** persistent offset in absolute SOC levels, arising when accumulated drift within an SLD is carried over as the starting point of the next SLD.

Among these components, intra-day shape error mainly originates from the replacement of heterogeneous natural-day SOC profiles by a limited number of representative-day templates. Even if the mapped representative day captures the general operating regime, differences in the detailed charging and discharging pattern within the day will remain. Cross-day accumulated drift arises because the reduced model propagates storage states using a fixed daily net SOC change for repeated occurrences of the same representative day, whereas in the full model the daily net change may vary from day to day. Inventory level bias is then formed when such drift accumulates within an SLD block and is passed forward as the starting SOC of subsequent blocks. Hence, the SOC reconstruction error in the reduced model is not merely a numerical residual, but a direct consequence of the RD-SLD approximation structure itself.

For this reason, noticeable SOC discrepancies may remain even under near-optimal cost solutions. Unlike aggregate cost indicators, SOC-based errors remain sensitive to how natural days are approximated and how chronology is propagated across SLD blocks.

To quantify these effects, define the pointwise SOC error:

$$e_t = SOC_t^{red} - SOC_t^{full} \tag{4.11}$$

The following six metrics are introduced to characterize the different components of the SOC reconstruction error.

(1) Absolute MAE

$$MAE_{abs} = \frac{1}{T} \sum_{t=1}^{T} |e_t| \tag{4.12}$$

Mathematical logic: Measures the average absolute deviation of the reconstructed SOC trajectory.

Interpretation: Captures the total error, combining intra-day shape error, accumulated drift, and inventory bias.

Function: Summary indicator of the aggregate SOC reconstruction gap.

(2) Increment MAE

$$MAE_{inc} = \frac{1}{T-1} \sum_{t=2}^{T} \left| \Delta SOC_t^{red} - \Delta SOC_t^{full} \right| \tag{4.13}$$

with

$$\Delta SOC_t = SOC_t - SOC_{t-1} \tag{4.14}$$

Equivalently,

$$\Delta SOC_t^{red} - \Delta SOC_t^{full} = e_t - e_{t-1} \tag{4.15}$$

Mathematical logic: Measures discrepancies in temporal SOC transitions.

Interpretation: Captures dynamic mismatch in charging/discharging behavior.

Function: Diagnostic indicator of intra-day transition mismatch; insensitive to constant inventory offsets.

(3) Day-shape MAE

For each day d, define the relative SOC trajectory:

$$\widetilde{SOC}_{d,t} = SOC_{d,t} - SOC_{d,0} \tag{4.16}$$

$$MAE_{shape} = \frac{1}{T} \sum_{d} \sum_{t} \left| \widetilde{SOC}_{d,t}^{red} - \widetilde{SOC}_{d,t}^{full} \right| \tag{4.17}$$

Equivalently, using the pointwise SOC error e_t defined in (4.6),

$$MAE_{shape} = \frac{1}{T} \sum_{d} \sum_{t} |e_t - e_0| \tag{4.18}$$

Mathematical logic: Removes daily starting level and compares intra-day trajectories.

Interpretation: Pure measure of intra-day shape mismatch.

Function: Direct indicator of representative-day shape distortion and a useful basis for identifying shape-critical days.

(4) Day-start bias

$$Bias_{day} = \frac{1}{D} \sum_{d=1}^{D} \left| SOC_{d,0}^{red} - SOC_{d,0}^{full} \right| \tag{4.19}$$

Equivalently, using the pointwise SOC error e_t defined in (4.6),

$$Bias_{day} = \frac{1}{D} \sum_{d=1}^{D} |e_0| \tag{4.20}$$

Mathematical logic: Measures deviation at daily starting points.

Interpretation: Captures inventory level bias at day boundaries.

Function: Indicator of cross-day inconsistency and inventory misalignment.

(5) Block bias

Let B_s denote the set of time steps within SLD block s:

$$\bar{e}_s = \frac{1}{|B_s|} \sum_{t \in B_s} e_t \tag{4.21}$$

$$Bias_{block} = \frac{1}{S} \sum_{s=1}^{S} |\bar{e}_s| \tag{4.22}$$

Mathematical logic: Measures the mean error over each SLD block.

Interpretation: Reflects persistent inventory offset resulting from accumulated drift.

Function: Captures the propagation of drift into inventory bias across SLD blocks.

(6) Drift ratio

$$Drift\ Ratio = \frac{MAE_{abs}}{MAE_{inc}} \tag{4.23}$$

Mathematical logic: Compares total error magnitude to dynamic error magnitude.

Interpretation:

- High value: error dominated by accumulated drift and inventory bias
- Low value: error dominated by intra-day dynamic mismatch

Function: Compact diagnostic indicator of the dominant error mechanism.

Overall relationship

A single aggregate SOC error is insufficient because different reconstruction errors may have different operational meanings. For example, a constant inventory offset may lead to a large absolute SOC error but small transition error, while an intra-day charging-shape mismatch may lead to a large increment error without persistent inventory bias. Therefore, the metrics are intentionally separated into total-error, transition-error, shape-error, and bias-related indicators. This decomposition supports the diagnostic objective of the chapter: identifying how the RD-SLD approximation distorts storage-state evolution rather than only measuring the size of the deviation.

- **Intra-day shape error:** MAE_{shape}, MAE_{inc}
- **Cross-day accumulated drift:** reflected by high Drift Ratio
- **Inventory level bias:** $Bias_{day}$, $Bias_{block}$
- **Total effect:** MAE_{abs}

4.6 Summary and insights

This chapter establishes the evaluation logic for reduced planning models with storage by connecting cost-based performance, reconstructed SOC gaps, TSA representation, and investment-decision quality. Cost-based metrics remain the primary criteria for judging whether a reduced model is economically competitive. In this chapter, operational estimation error and optimality gap are used to identify a competitive reduced planning result, with NRD = 21 selected as the reference case for subsequent SOC-based diagnosis.

However, the analysis also shows that a reduced model with acceptable cost performance may still exhibit substantial reconstructed SOC gaps. For storage-embedded systems, this is important because reduced temporal scope and chronology approximation can distort inter-period storage behavior [11, 17]. Under the RD-SLD approximation, heterogeneous natural-day storage behavior is replaced by repeated representative-day profiles and reconstructed SLD sequences. Therefore, the SOC gap is not only a numerical trajectory error, but also a inevitable consequence of TSA representation.

The main insight of this chapter is that the reconstructed SOC gap with respect to time provides a diagnostic link between TSA and investment decisions. SOC does not directly appear as a cost term in the objective function, but it constrains storage charging, discharging, and available flexibility. As a result, SOC misrepresentation may be partly compensated by dispatch adjustment under a fixed investment decision, but it can still influence how the reduced model evaluates storage value, system stress, and investment needs.

For this reason, SOC-based metrics are introduced as a complementary diagnostic layer rather than as replacements for cost-based metrics. Their role is not to force the reduced-model SOC trajectory to exactly reproduce the full-time trajectory. Instead, they decompose SOC reconstruction gaps into interpretable mechanisms, including intra-day shape mismatch, cross-day accumulated drift, and inventory-level bias. This makes the SOC gap useful for identifying storage-relevant days, distorted representative-day patterns, and chronology-sensitive SLD blocks.

Overall, this chapter shows that cost-based metrics answer whether a reduced planning result is economically competitive, while SOC-based diagnostics explain how TSA affects storage-state representation and why this may matter for investment decisions. This provides the basis for the mechanism-oriented SOC analysis in the next chapter and for the later use of SOC-based diagnostic signals in feedback-enhanced extreme-day selection.

Chapter 5

Diagnostic analysis of SOC reconstruction errors

5.1 Overview of the diagnostic objective

Chapter 4 established that, under the RD-SLD approximation, noticeable SOC reconstruction gaps may remain even when the reduced model is economically competitive. These gaps are structurally unavoidable, but they are not uniform in origin. A large SOC discrepancy may be associated with intra-day shape mismatch, cross-day accumulated drift, or persistent inventory bias. Accordingly, the purpose of this chapter is not to treat SOC error as a single deviation measure, but to examine what it reveals about the mechanisms by which temporal reduction distorts storage-state reconstruction.

The analysis is carried out for the reference reduced model with NRD = 21, where the extreme-day setting is the original maximum-load day. This case is economically competitive and therefore provides a suitable reference for diagnostic analysis. The deep evaluation results are examined from several complementary perspectives: storage-level comparison, grouped comparison by storage type, temporal concentration of daily errors, identification of diagnostically important days, and block-level persistent bias.

In the present case study, Buses 1, 4, and 5 correspond to daily-cycling storage units, whereas Buses 3 and 7 correspond to long-duration storage units. Since long-duration storage relies more strongly on inter-day energy transfer, it is expected to be more sensitive to chronology reconstruction and inventory propagation. By contrast, daily-cycling storage is expected to

be more sensitive to representative-day selection and intra-day SOC shape reconstruction. As shown below, the diagnostic results reveal different dominant error mechanisms for these two storage categories.

5.2 Storage-level comparison of SOC reconstruction errors

Figure 5.1 compares the three primary SOC-based indicators at the storage level: absolute MAE, increment MAE, and day-shape MAE. Together, these metrics distinguish between total SOC discrepancy, mismatch in local SOC transitions, and intra-day shape distortion after removing daily starting levels.

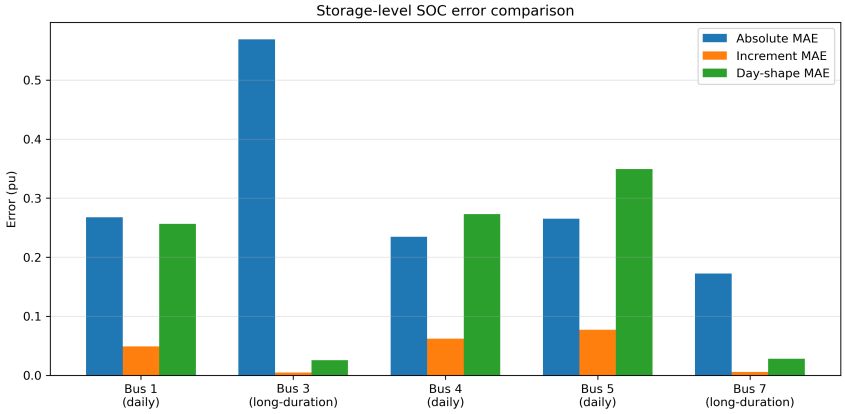

Figure 5.1: Storage-level comparison of absolute MAE, increment MAE, and day-shape MAE.

The most prominent result in Fig. 5.1 is the behavior of Bus 3. It exhibits by far the largest absolute MAE, while its increment MAE and day-shape MAE remain very small. This indicates that the dominant discrepancy at Bus 3 is not local intra-day mismatch, but persistent deviation in the reconstructed inventory level. In other words, the reduced model reproduces the local SOC movement of Bus 3 reasonably well, but does not preserve its absolute storage level over time.

A different pattern is observed for the daily-cycling units, especially Bus 5. Its absolute MAE is much smaller than that of Bus 3, but its increment

MAE and day-shape MAE are the largest among the daily storage units. This identifies Bus 5 as the clearest shape-driven case, in which the main discrepancy originates from inaccurate reconstruction of the intra-day SOC pattern. Bus 4 shows the same qualitative behavior at lower magnitude, while Bus 1 lies between the two extremes and can be interpreted as a mixed case.

Bus 7 follows the same qualitative mechanism as Bus 3, but with a much smaller overall error magnitude. Its absolute MAE is non-negligible, while both increment MAE and day-shape MAE remain small. This suggests that Bus 7 is also mainly affected by chronology-related drift rather than by intra-day shape distortion, although much less severely than Bus 3.

To further examine inventory-level misalignment, Fig. 5.2 compares the day-start bias and block bias of all storage units.

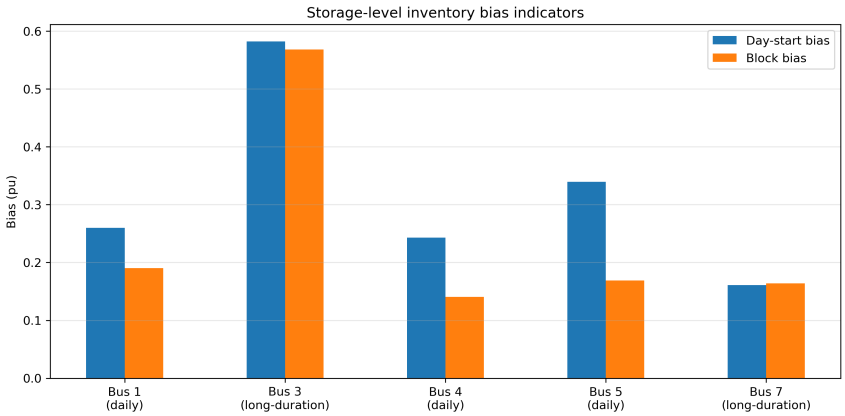

Figure 5.2: Storage-level comparison of day-start bias and block bias.

The bias indicators reinforce the interpretation above. Bus 3 again stands out with the highest day-start bias and the highest block bias, confirming that its discrepancy is not limited to isolated time points but develops as a persistent inventory offset. For Buses 1, 4, and 5, day-start bias is still visible, but block bias remains clearly lower than for Bus 3, indicating that cross-day inconsistency exists but does not propagate into the same degree of persistent drift. Bus 7 remains only mildly affected.

A compact summary is provided by the drift ratio in Fig. 5.3, where a high ratio indicates that total error is much larger than local transition error and is therefore dominated by accumulated drift and inventory bias.

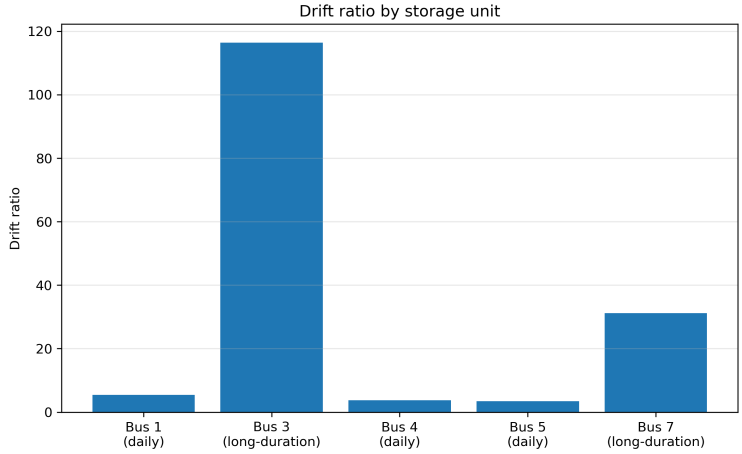

Figure 5.3: Drift ratio by storage unit.

Figure 5.3 shows a clear separation between storage units. Bus 3 is strongly drift-dominated, Bus 7 is a weaker version of the same mechanism, and Buses 4 and 5 are more shape-driven. Bus 1 again lies in between. Taken together, Figs. 5.1–5.3 show that SOC reconstruction error is heterogeneous across storage units and that the dominant mechanism differs systematically across them.

5.3 Error mechanisms across storage technologies

The storage-level results can also be interpreted from the perspective of storage type. In the present system, Buses 1, 4, and 5 represent daily-cycling storage, whereas Buses 3 and 7 represent long-duration storage. Figure 5.4 compares the storage units in terms of increment MAE and absolute MAE, while Figs. 5.5 and 5.6 compare the mean daily error patterns of the two storage groups.

In Fig. 5.4, Bus 3 is clearly separated from all other units: it combines a very small increment MAE with the largest absolute MAE, confirming that its dominant mechanism is accumulated drift rather than local dynamic mismatch. Bus 7 shows the same qualitative pattern at much lower severity. This difference between the two long-duration storage units suggests that

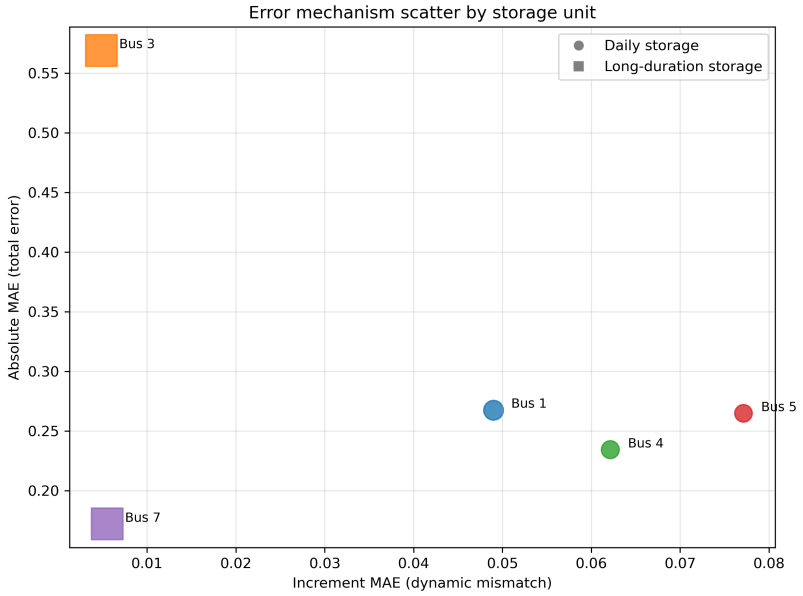

Figure 5.4: Error mechanism scatter by storage unit.

the same chronology-related error can be amplified or mitigated by local network conditions. Bus 7 is located in a more strongly connected area, linked to Buses 6, 8, and 9, with thermal generation at Bus 6 and a candidate reinforcement on the 6–7 corridor. Such surrounding flexibility may absorb part of the local energy-balance mismatch introduced by representative-day reconstruction. By contrast, Bus 3 has fewer nearby compensation options, so the mismatch is more likely to remain in the storage trajectory and appear as persistent inventory-level deviation. By contrast, Buses 4 and 5 are positioned farther to the right, indicating stronger local mismatch in charging and discharging dynamics. Bus 5 is again the clearest shape-driven case, while Bus 1 remains a mixed case.

The same distinction appears in the temporal aggregation of daily errors. Figure 5.5 compares the mean daily absolute error of daily-cycling and long-duration storage.

The long-duration storage group exhibits a higher and smoother absolute SOC error profile over substantial parts of the year. This persistence is consistent with chronology-related accumulation: once inventory deviation begins, it is propagated over consecutive days. By contrast, the daily-cycling group shows a more fragmented pattern with stronger day-to-day variation,

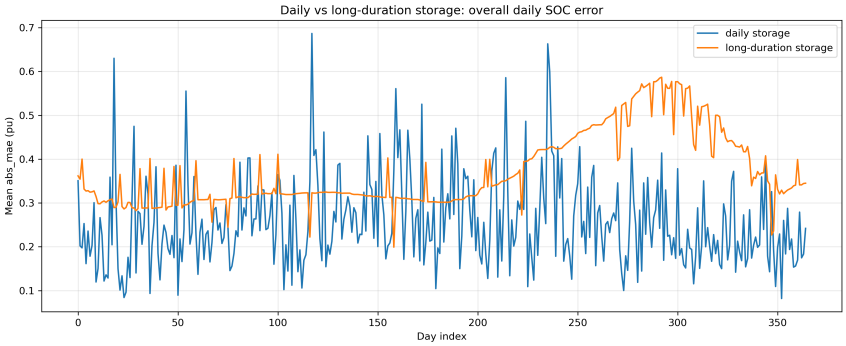

Figure 5.5: Comparison of mean daily absolute SOC error between daily-cycling and long-duration storage.

indicating higher sensitivity to specific operating days rather than to persistent block-level propagation.

A complementary picture is provided by Fig. 5.6, which compares the mean day-shape MAE of the two storage groups.

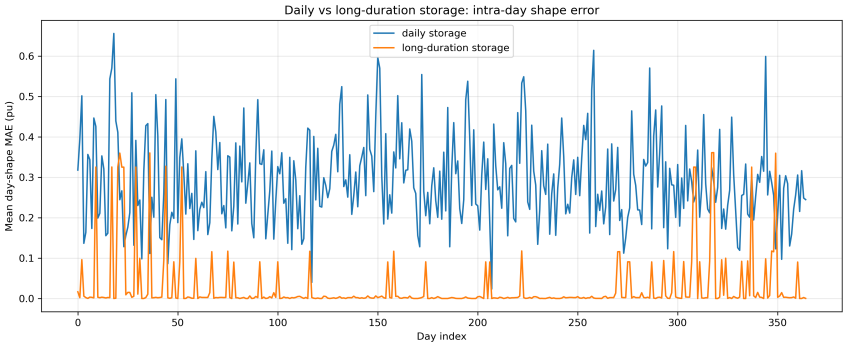

Figure 5.6: Comparison of mean daily day-shape error between daily-cycling and long-duration storage.

Figure 5.6 shows the opposite contrast: daily-cycling storage exhibits substantially larger day-shape error across most of the year, whereas long-duration storage remains close to zero for most days. This indicates that, for long-duration storage, the main issue is not the reconstruction of intra-day shape, but the preservation of chronology and inventory level across days. For daily-cycling storage, by contrast, the main issue lies in the representation of intra-day charging and discharging patterns.

Taken together, Figs. 5.4–5.6 show that TSA affects different storage types in systematically different ways. Long-duration storage mainly exposes chronology-related drift and inventory bias, whereas daily-cycling storage mainly exposes intra-day shape mismatch.

5.4 Temporal distribution of daily SOC errors

The storage-level and technology-level comparisons indicate that SOC reconstruction errors are not homogeneous across time. To examine their temporal concentration, Fig. 5.7 shows the daily absolute SOC error for each storage unit, while Fig. 5.8 shows the corresponding day-shape error.

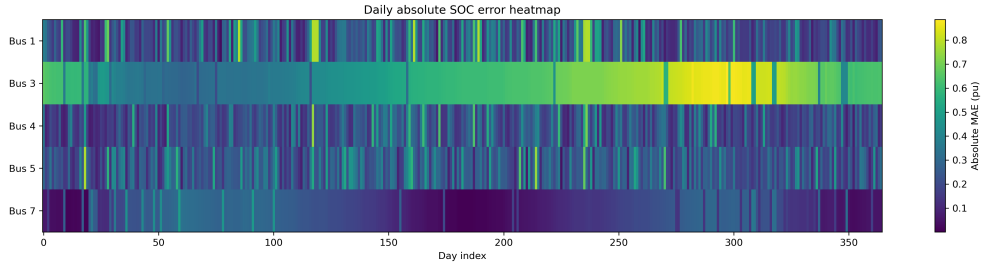

Figure 5.7: Daily absolute SOC error heatmap across storage units.

Figure 5.7 shows that absolute SOC error is concentrated in specific periods rather than distributed uniformly over the year. Bus 3 exhibits the clearest persistent high-error regime, especially in the later part of the year, which is consistent with the drift-dominated interpretation established above. Bus 7 shows a similar but much weaker structure. In contrast, Buses 1, 4, and 5 display more fragmented patterns with shorter and more localized peaks, indicating that the discrepancy of daily-cycling storage is more closely tied to particular operating days than to continuous inventory accumulation.

Figure 5.8 reveals a different structure. The most striking contrast is again Bus 3: although it exhibits the largest absolute error, its day-shape error remains small for most days. This confirms that its main discrepancy is not failure to reproduce the relative intra-day SOC profile, but drift in the absolute inventory level. Bus 7 shows the same pattern at lower magnitude. By contrast, Bus 5 exhibits frequent and pronounced day-shape errors

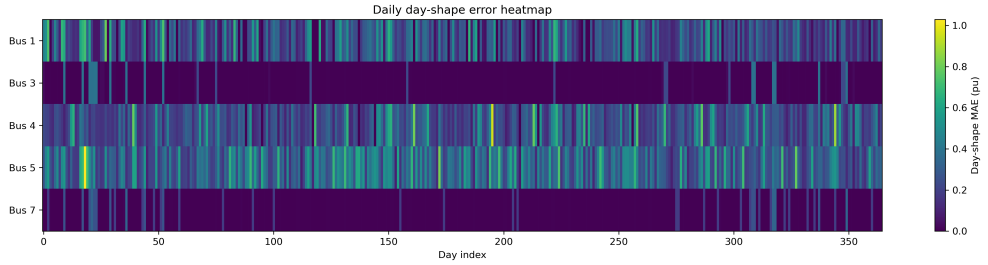

Figure 5.8: Daily day-shape SOC error heatmap across storage units.

throughout the year, while Buses 1 and 4 show the same qualitative behavior more moderately. Comparing the two heatmaps therefore makes the diagnostic distinction clear: some periods are dominated by shape-related error, while others are dominated by drift and inventory-level bias.

5.5 Identification of diagnostically important days

The heatmaps above show that SOC reconstruction error is temporally concentrated. To identify the most informative periods more explicitly, the daily results are ranked from two complementary perspectives: overall SOC-critical days, identified by aggregated absolute MAE across all storage units, and shape-critical days, identified by aggregated day-shape MAE.

Figure 5.9 shows that Days 17 and 44 are the most severe shape-critical days, followed by Days 344, 18, and 9. These days are particularly informative for diagnosing deviations in intra-day SOC pattern reconstruction and are therefore especially relevant to daily-cycling storage.

A different ranking emerges in Fig. 5.10, where Days 235, 117, and 236 are the most critical from the perspective of total SOC discrepancy, followed by Days 18, 214, and 292. Compared with the shape-based ranking, the composition changes substantially. This confirms that the days that dominate total SOC mismatch are not identical to those that dominate intra-day shape mismatch.

An important exception is Day 18, which appears in both rankings. This makes it a particularly informative diagnostic day, since it is simultaneously challenging from the perspectives of both total SOC fidelity and intra-day

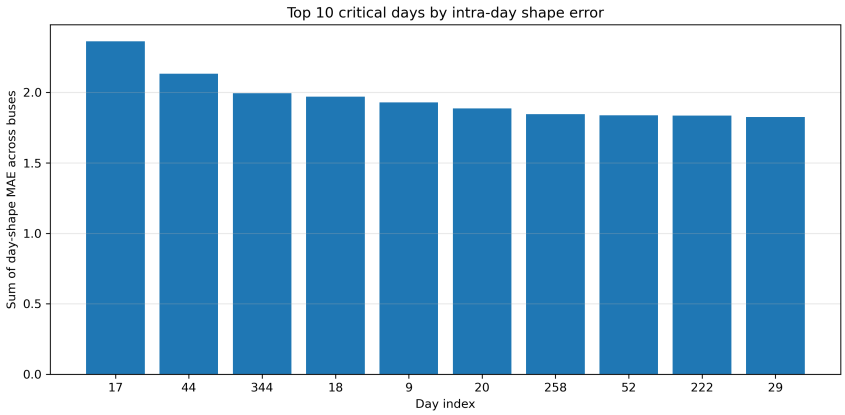

Figure 5.9: Top 10 critical days ranked by aggregated day-shape error.

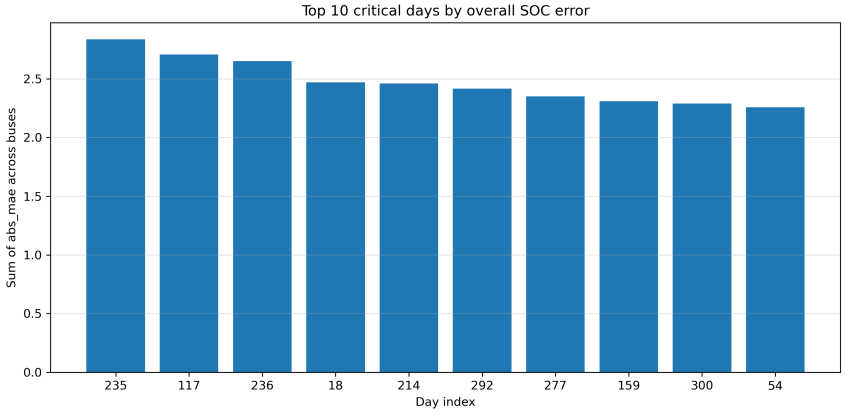

Figure 5.10: Top 10 critical days ranked by aggregated absolute SOC error.

shape reconstruction.

The comparison between Figs. 5.9 and 5.10 therefore leads to an important diagnostic observation: important days should not be identified using a single criterion only. Shape-critical days are more closely related to representative-day pattern mismatch, whereas overall SOC-critical days are more closely related to sustained drift and inventory deviation. For this reason, these two categories are kept distinct in the subsequent interpretation. They should be understood as candidate signals for storage-aware sensitivity analysis and refinement, rather than as a one-step correction rule.

5.6 Identification of chronology-sensitive SLD blocks

The daily perspective is still incomplete for storage units dominated by accumulated drift. If chronology reconstruction deviates over consecutive days, the resulting discrepancy should also appear at the SLD block level. To examine this mechanism more directly, Fig. 5.11 ranks the ten SLD blocks with the largest aggregated block bias across all storage units.

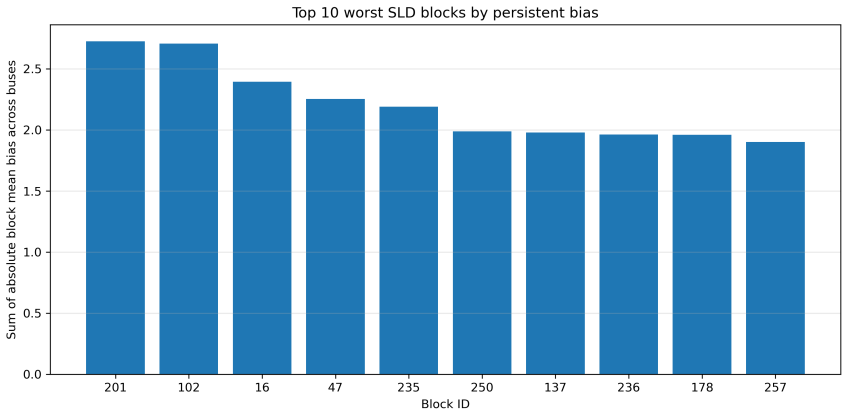

Figure 5.11: Top 10 worst SLD blocks ranked by aggregated persistent SOC bias.

Figure 5.11 shows that persistent block-level bias is also highly concentrated. A limited number of SLD blocks account for the strongest inventory deviation, with Blocks 201 and 102 showing the largest aggregated bias, fol-

lowed by Blocks 16, 47, and 235. This indicates that chronology-related SOC error is not uniformly distributed across all SLDs, but is concentrated in specific chronology segments.

This block-level ranking is consistent with the earlier observations. In particular, the prominence of high-bias blocks supports the conclusion that part of the discrepancy, especially for long-duration storage, is driven by persistent drift rather than by isolated day-shape mismatch. It also helps explain why some of the overall SOC-critical days identified earlier appear near top-ranked blocks, such as Days 235 and 236 near Block 201: these critical days can in some cases be interpreted as visible manifestations of deeper block-level propagation problems.

The block-level results should therefore be read as a chronology-sensitive diagnostic layer that complements the daily rankings. They are not intended to replace critical-day identification, but to reveal propagation structures that cannot be inferred from day-wise rankings alone.

5.7 Discussion on diagnostic findings

This chapter has examined the SOC reconstruction error of the reference reduced model at NRD = 21 from multiple diagnostic perspectives. The results show that the observed SOC gap is not a uniform discrepancy, but a combination of distinct mechanisms distributed across storage units, days, and chronology blocks.

At the storage level, Bus 3 is the clearest drift-dominated case: it exhibits the largest absolute SOC error while maintaining very small increment and day-shape MAE, together with the highest day-start bias, block bias, and drift ratio. Bus 7 shows the same qualitative mechanism at much lower magnitude. This combination of low local-transition error and high inventory-level error reflects the slow-moving nature of long-duration storage. Increment MAE measures the change in the SOC error between adjacent time steps, rather than the accumulated SOC offset itself. For daily-cycling storage, frequent intra-day charging and discharging may generate errors with alternating signs, so local overestimation and underestimation can partly cancel within or across daily cycles. Long-duration storage, however, is governed by slower inter-day inventory trends. Once the reduced model assigns an incorrect long-term SOC level, the error $SOC^{red} - SOC^{full}$ may remain sign-consistent over extended periods, producing persistent inventory bias

even when local transition errors remain small. In contrast, Bus 5 is the clearest shape-driven case, with the largest increment MAE and day-shape MAE among the daily-cycling units. Bus 1 remains a mixed case.

When aggregated by storage type, a systematic distinction emerges. Long-duration storage mainly exposes chronology-related distortion, including cross-day drift and inventory bias, whereas daily-cycling storage mainly exposes intra-day representative-day mismatch. This confirms that TSA affects different storage technologies through different dominant mechanisms.

The temporal analysis further shows that SOC errors are concentrated rather than uniformly distributed over the year. Based on this concentration, two complementary categories of diagnostically important days are identified: shape-critical days, such as Days 17, 44, and 344, and overall SOC-critical days, such as Days 235, 117, and 236. Day 18 is particularly informative because it appears in both rankings. In addition, the block-level diagnosis shows that persistent inventory bias is also concentrated in a limited number of SLD blocks, especially Blocks 201 and 102, indicating that part of the discrepancy is propagated through chronology segments rather than generated only as isolated daily events.

The main insight of this chapter is therefore that SOC-based diagnostics are valuable not because they produce another aggregate performance score, but because they reveal how temporal reduction affects different storage types through different mechanisms. More specifically, the results support distinguishing between shape-critical days and chronology-sensitive periods, rather than expecting a single undifferentiated list of extreme days to capture the full structure of SOC reconstruction error. In this way, the chapter provides a diagnostic basis for the feedback discussion in the next chapter.

Chapter 6

Feedback discussion based on SOC diagnostics

6.1 Motivation for feedback

Chapter 5 showed that the SOC reconstruction gap in the reduced model is not a uniform error, but a structured combination of different mechanisms. In particular, the diagnostics distinguished between periods associated with intra-day shape mismatch and periods associated with chronology-related drift and inventory bias. This suggests that the conventional extreme-day setting based only on the maximum-load day may overlook storage-relevant periods that are important for the representation of SOC trajectories.

The idea of using diagnostic information as feedback is consistent with the model-aware direction in TSA research. Traditional representative-period selection is often based on input-profile similarity, but this does not necessarily identify the periods that are most important for planning outcomes. Recent representative-day selection studies therefore distinguish between input-based and cost-based perspectives, where the quality of selected periods is evaluated by their effect on the optimization model rather than only by clustering distance [31]. Similarly, feedback-enhanced TSA methods use model-output errors to identify poorly represented periods and update the aggregation procedure [20]. These studies motivate the use of post-solution diagnostic information as a feedback signal for improving representative-period selection.

Based on this observation, a first feedback attempt is considered in this

chapter. The purpose is not to directly optimize SOC error within the planning model, since the reduced model still minimizes total cost rather than SOC discrepancy. Instead, the SOC diagnostics are used to modify the extreme-day setting in the TSA procedure, in order to test whether diagnostically important days can improve the reduced-model planning outcome.

6.2 Positioning of extreme-day selection methods

Extreme-day selection methods in representative-period-based planning can be broadly understood from three perspectives: input-defined stress conditions, cost- or model-output-defined critical periods, and storage- or chronology-aware temporal aggregation. These perspectives provide different ways to identify days that should be preserved in the reduced temporal representation.

The first perspective is input-defined extreme-day selection. In this type of method, extreme days are selected directly from the original time-series inputs before the planning model is solved. The most conventional example is the peak-load day, which represents the maximum demand condition and is often used as an adequacy-related stress day. In systems with high renewable penetration, however, peak load alone may not represent the most stressful operating condition, because system stress also depends on renewable availability. Therefore, high net-load days and low net-load days have been proposed to capture the combined effect of demand and renewable generation [32]. Low-renewable days can also be used to represent renewable scarcity, while high-ramping or high-variability days are relevant when the main concern is short-term flexibility and reserve requirement. These input-defined methods are intuitive and easy to implement, but they are selected independently of the optimization results. As a result, they may identify physically extreme days without necessarily identifying the days that are most important for the final investment decision.

The second perspective is cost- or model-output-defined extreme-day selection. In this type of method, critical days are identified according to the behavior of the planning model or the error observed in model evaluation. For example, Li et al. distinguish between input-based and cost-based representative-day selection and use indicators such as load-shedding cost

and highest operating cost to identify extreme days under difficult operating conditions [31]. Compared with input-defined rules, cost-based rules are more directly connected to the expansion-planning objective, because they identify periods that have a large economic impact on the planning result. More generally, a posteriori time-series aggregation methods use model outputs after solving the optimization problem to refine the temporal representation. Hilbers et al. propose an a posteriori aggregation approach that incorporates operational variables and chronology information into the aggregation process, including information related to generation, transmission, and storage operation [38]. Teichgraeber et al. further show that extreme periods can be included in time-series aggregation and that iterative inclusion of such periods can improve the reliability of energy system designs under rare but important operating conditions [41], [42]. These studies indicate that extreme-day selection can be guided not only by input profiles, but also by the consequences of temporal aggregation in the optimization model.

The third perspective is storage- and chronology-aware aggregation. Storage systems introduce inter-temporal coupling because charging and discharging decisions are linked through the state of charge. Therefore, storage-aware temporal aggregation must consider not only individual representative days, but also the chronological relationship between periods. Chronological time-period clustering preserves temporal order and is therefore relevant for capacity expansion planning with storage [3]. Hybrid representative-day approaches further combine representative-day clustering, explicit extreme-value preservation, and sequentially linked days to represent both extreme conditions and part of the inter-day chronology [17]. These methods are especially important for long-duration storage, because storage stress may arise from consecutive low-renewable periods, repeated net-load stress, or accumulated multi-day energy imbalance rather than from a single peak-load day.

However, although the literature has considered input-defined extremes, cost-based critical days, and storage-aware chronology reconstruction, storage-related reconstruction error is still less frequently used as the direct criterion for selecting extreme days. Existing storage-aware aggregation methods usually focus on preserving chronology, maintaining storage feasibility, or improving temporal representation for storage operation. They do not commonly select extreme days according to explicit SOC reconstruction errors, such as absolute SOC deviation, intra-day SOC shape mismatch, cross-day accumulated drift, or inventory-level bias. This leaves a gap for storage-embedded planning models: a day may be important not because it has the

highest load, lowest renewable output, largest net load, or highest operating cost, but because it reveals a severe distortion in the reconstructed SOC trajectory.

This chapter therefore positions SOC-informed extreme-day selection as a complementary extension of RD-SLD methodology. It does not replace peak-load, net-load, low-renewable, ramping-based, or cost-based extreme-day rules. Instead, it adds a solution-dependent storage-state diagnostic perspective. After economically competitive reduced models are identified, SOC-based metrics are used to locate natural days or SLD blocks where the representative-day reconstruction poorly captures storage behavior. These SOC-critical days are then preserved as extreme days and evaluated through their impact on full-time total cost and investment decisions. In this sense, the proposed feedback does not aim to minimize SOC error directly. Rather, it examines whether storage-related diagnostic information can provide useful guidance for extreme-day selection in cost-oriented planning models with storage.

6.3 Modified extreme-day setting

Based on the positioning in Section 6.2, this section constructs the modified extreme-day settings used for the feedback experiment. The purpose is not to search for a globally optimal extreme-day list, but to test whether storage-related diagnostic signals from Chapter 5 can provide useful candidates for representative-day preservation.

The baseline setting follows the conventional maximum-load-day rule. In the present case study, the preserved peak-load extreme day is natural day 21, so the baseline extreme-day set is

$$E_{\text{peak}} = \{21\}.$$

This setting provides the reference case for comparing whether SOC-informed days can improve the planning outcome beyond a conventional load-based extreme-day rule.

The SOC-informed settings are constructed from the diagnostic rankings obtained in Chapter 5. Two types of SOC signals are used. The first is the intra-day day-shape error, which reflects whether the storage pattern within a natural day is poorly represented by its mapped representative day. The second is the aggregated absolute SOC discrepancy, which identifies days

where the reconstructed SOC trajectory differs most strongly from the full-time reference trajectory.

A first modified setting is designed as a hybrid compromise between these two diagnostic signals. It contains:

- Days 17 and 44, which are the two most severe days ranked by intra-day shape error;
- Day 235, which is the day with the largest aggregated absolute SOC discrepancy;
- Day 18, which appears in both the shape-critical and overall SOC-critical rankings.

The resulting hybrid SOC-informed extreme-day set is

$$E_{\text{hybrid}} = \{17, 44, 235, 18\}.$$

The rationale is that Days 17 and 44 represent the strongest signals of representative-day shape distortion, Day 235 represents the strongest signal of overall SOC discrepancy, and Day 18 is retained because it is jointly identified by both rankings. Therefore, the hybrid set is interpreted as a compromise between shape-oriented and overall-SOC-oriented diagnostics.

To avoid evaluating only one ad hoc SOC-informed list, two additional settings are constructed. The first uses the four most critical days ranked by intra-day day-shape error:

$$E_{\text{shape}} = \{17, 44, 344, 18\}.$$

The second uses the four most critical days ranked by aggregated absolute SOC discrepancy:

$$E_{\text{ABS}} = \{235, 117, 236, 18\}.$$

In summary, the following four extreme-day settings are compared:

$$E_{\text{peak}} = \{21\},$$

$$E_{\text{shape}} = \{17, 44, 344, 18\},$$
$$E_{\text{ABS}} = \{235, 117, 236, 18\},$$
$$E_{\text{hybrid}} = \{17, 44, 235, 18\}.$$

All three SOC-informed settings are tested over the NRD range from 20 to 50 and compared with the peak-load baseline. The evaluation focuses on planned total cost after full-time re-evaluation, rather than on whether the SOC error itself is monotonically reduced. This is consistent with the cost-oriented role of the planning model: SOC diagnostics are used as feedback signals for selecting candidate extreme days, while the effectiveness of each setting is judged by the resulting full-time cost and investment decision.

Changing the extreme-day set explicitly modifies the TSA configuration, because preserved days are imposed during the representative-day selection process rather than added only after clustering. In the hierarchical clustering procedure, a preserved extreme day is not averaged out when its cluster is merged; instead, it remains as the representative profile of the merged cluster. As a result, different extreme-day settings can change the final representative-day profiles, the assignment of natural days to representative days, and the resulting SLD structure. Since the SLD formulation uses this assignment to reconstruct inter-day SOC propagation, the choice of preserved days affects both the intra-day profiles and the inter-day storage chronology represented by the reduced model. This forms the direct link between extreme-day selection, system-stress representation, storage-value estimation, and the resulting investment decisions.

6.4 Comparison of planned total cost under different SOC-informed extreme-day lists

Although the feedback signal is derived from SOC diagnostics, the effectiveness of each modified extreme-day setting is evaluated primarily through the planning objective. This follows the model-based evaluation principle that representative-period selection should ultimately be assessed by its impact on optimization outcomes [6, 31, 20]. Therefore, the SOC-informed lists are not compared only by whether they reduce SOC error, but by whether the resulting investment decisions perform well when evaluated under the full chronological model.

To move beyond a single modified setting, this section compares the planned total cost obtained under three SOC-informed extreme-day lists over the NRD range from 20 to 50. According to Section 6.2, the three lists are defined as follows:

- **ABS:**

$$\{235,\ 117,\ 236,\ 18\},$$

corresponding to the top four overall SOC-critical days;

- **Dayshape:**

$$\{17,\ 44,\ 344,\ 18\},$$

corresponding to the top four shape-critical days;

- **Hybrid extreme-day set:**

$$\{17,\ 44,\ 235,\ 18\},$$

combining the two strongest shape-critical days, the strongest overall SOC-critical day, and the overlapping diagnostic day.

Figure 6.1 compares these three SOC-informed settings directly.

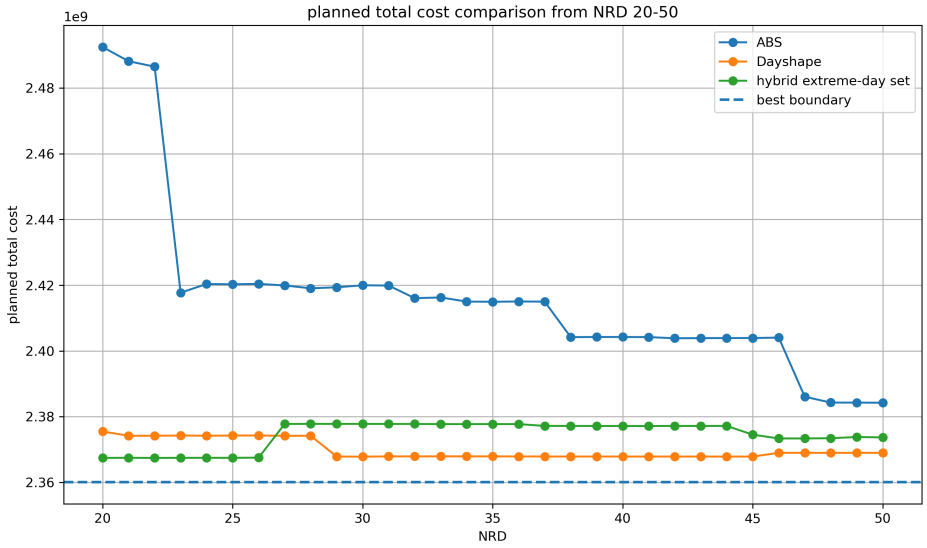

Figure 6.1: Comparison of planned total cost under the ABS, Dayshape, and hybrid extreme-day settings from NRD 20 to 50.

As shown in Fig. 6.1, the Dayshape and hybrid extreme-day settings produce very similar cost levels over most of the investigated NRD range, while the ABS-based setting remains clearly higher. This indicates that the

planning impact of the hybrid set is, to a large extent, dominated by its day-shape component. In other words, preserving the most severe shape-critical days appears to be more effective than preserving the top-ranked overall SOC-critical days when the objective is to improve the economic outcome of the reduced model.

At the same time, the two best-performing settings are not identical. The best result obtained with the Dayshape setting is a total cost of

$$2{,}367{,}832{,}153$$

at NRD = 30, whereas the hybrid extreme-day set reaches a lower best value of

$$2{,}367{,}455{,}161$$

at NRD = 20. Therefore, although the overall behavior of Dayshape and Hybrid is similar, the hybrid extreme-day set still provides the best total cost among the tested SOC-informed settings.

For a broader perspective, Fig. 6.2 further compares all five cases considered in this chapter: no extreme day, the conventional peak-load extreme day, and the three SOC-informed lists.

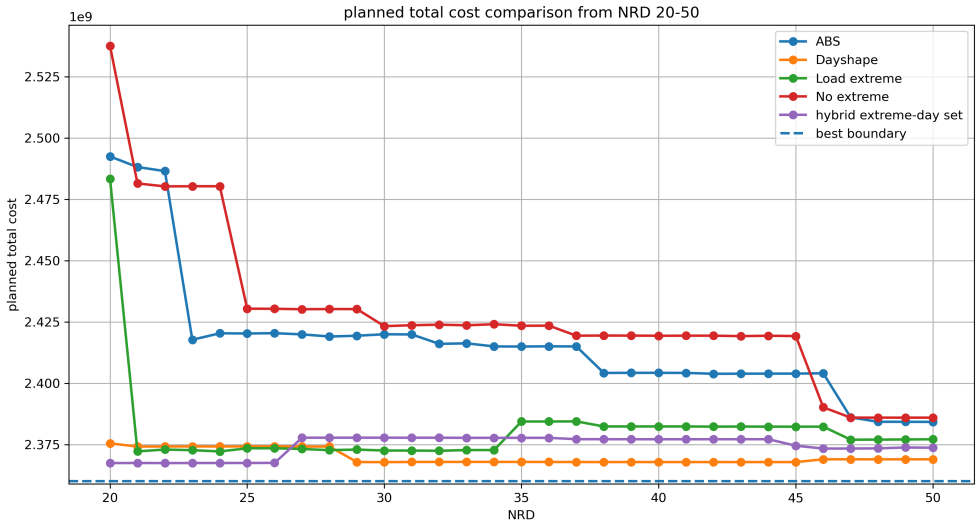

Figure 6.2: Comparison of planned total cost under five extreme-day settings from NRD 20 to 50.

Figure 6.2 provides a more intuitive comparison of the role of extreme-day selection. The no-extreme case gives the worst overall performance, while

preserving only the conventional peak-load day improves the result substantially. However, the best cost levels are obtained by the Dayshape and hybrid extreme-day settings, which outperform both the no-extreme case and the conventional load-extreme setting over most of the relevant range. By contrast, the ABS-based setting does not achieve the same level of improvement.

These results suggest that, for the present test system, shape-oriented SOC diagnostics are more effective than absolute-SOC rankings in identifying economically useful extreme days. At the same time, the slight advantage of the hybrid set indicates that including one strong overall-SOC signal can still be beneficial when combined with the dominant day-shape information. The implication is not that one universally optimal diagnostic rule has been found, but that different SOC-based rankings perturb the TSA process in different ways, and that the day-shape signal appears to be the most influential among them in the current case.

For this reason, the next section focuses on the two best-performing settings, namely the Dayshape and hybrid extreme-day sets. The purpose is to examine whether their similar total costs arise merely from close objective values, or whether they also correspond to similar investment structures. In addition, the comparison with the conventional peak-load extreme-day setting provides a basis for explaining how different preserved days lead to different investment decisions and what mechanisms inside the TSA procedure may be responsible for these differences.

6.5 Interpretation of the investment decisions under different extreme-day settings

6.5.1 Limitations of absolute-SOC critical days as extreme-day candidates

The comparison between the ABS, Dayshape, and Hybrid settings shows that not every SOC diagnostic indicator is equally suitable as an extreme-day selection rule. In model-aware TSA, feedback should identify periods whose preservation can improve the representative-period structure and the resulting optimization outcome, rather than only periods where the reduced trajectory visibly deviates from the benchmark [31, 20]. For storage systems, this distinction is important because storage representation is sensitive to

inter-period chronology and long-duration energy shifting [5, 11]. Consistent with this storage-aware interpretation, the diagnostic results in this thesis show that a large absolute SOC error may reflect accumulated inter-period drift or inherited inventory bias rather than a poor intra-day representative-day match.

In the present case, the ABS-based extreme-day setting does not improve the planning outcome and performs worse than the conventional peak-load extreme-day rule. This suggests that aggregated absolute SOC error is not well suited as a direct criterion for preserved extreme days. The main reason is that absolute SOC error combines multiple components of reconstruction discrepancy, including intra-day shape mismatch, cross-day accumulated drift, and inventory-level bias. Therefore, a day with large absolute SOC error is not necessarily a day whose daily storage pattern is poorly represented by its mapped representative day. It may simply be the point where accumulated drift or inherited inventory bias becomes most visible.

This matters because extreme-day preservation in the present framework acts mainly through representative-day selection and natural-day mapping. Preserving a day is most useful when its intra-day pattern would otherwise be distorted or averaged out by clustering. If a day is selected only because its absolute SOC level differs strongly from the full-model trajectory, preserving it may not correct the underlying representative-day mismatch.

By contrast, the day-shape error is more closely aligned with the level at which clustering acts. Since it removes the daily starting SOC level, it filters out much of the inventory-level bias and focuses on the intra-day SOC pattern. Days ranked by day-shape error therefore provide a cleaner signal of representative-day distortion. This explains why the Dayshape and Hybrid settings perform better than the ABS setting in Section 6.3.

Therefore, absolute SOC error is treated in this thesis mainly as a diagnostic indicator of where the reconstructed trajectory differs most strongly from the full model. For extreme-day design, metrics that are more directly linked to representative-day distortion, especially the day-shape error, provide more effective guidance.

6.5.2 Comparison of investment decisions under the Dayshape and Hybrid settings

To further interpret the cost results in Section 6.4, the two best-performing SOC-informed settings are compared at the level of investment decisions. The purpose of this comparison is to determine whether the similar total costs obtained under the Dayshape and Hybrid settings arise merely from close objective values, or whether they also correspond to similar expansion structures.

Table 6.1: Comparison of the best investment decisions obtained from the Dayshape and Hybrid SOC-extreme settings.

Category	**Dayshape best solution**	**Hybrid SOC-extreme best solution**
Best NRD	30	20
Investment cost	766,897,250	779,530,882
Full-model operation cost	1,600,934,903	1,587,924,278
Full-evaluated total cost	2,367,832,153	2,367,455,161
Transmission expansion	Line 36 (Bus 6–7) Line 37 (Bus 13–15)	Line 36 (Bus 6–7) Line 37 (Bus 13–15)
Wind investment	Bus 4: 400 Bus 5: 266.46 Bus 13: 78.06 Bus 18: 400 Bus 22: 396.36	Bus 4: 400 Bus 5: 276.71 Bus 13: 95.49 Bus 18: 400 Bus 22: 400
ESS investment	Bus 1: 1.03 Bus 3: 5 Bus 4: 80 Bus 5: 80 Bus 7: 8	Bus 1: 3.19 Bus 3: 5 Bus 4: 80 Bus 5: 80 Bus 7: 8

Table 6.1 shows that the two solutions are highly similar in their overall expansion structure. Both settings produce the same transmission-expansion backbone, namely Lines 36 and 37, and both adopt the same main ESS

configuration at Buses 3, 4, 5, and 7. The only storage difference appears at Bus 1, where the Hybrid solution installs a slightly larger ESS capacity than the Dayshape solution.

The main distinction between the two solutions lies in wind investment. Compared with the Dayshape solution, the Hybrid solution allocates slightly more capacity to Buses 5 and 13 and fully utilizes the candidate wind capacity at Bus 22. These differences are modest in magnitude and do not alter the basic expansion pattern. Instead, they indicate that the Hybrid setting acts mainly as a refinement of the wind-allocation structure already identified by the Dayshape setting.

This interpretation is also consistent with the cost decomposition. The Hybrid solution has a somewhat higher investment cost than the Dayshape solution, but it achieves a lower full-model operation cost. As a result, the Hybrid setting yields the best full-evaluated total cost among the tested SOC-informed settings. This means that the advantage of the Hybrid setting does not come from a fundamentally different expansion strategy, but from a limited adjustment of investment details that produces a more favorable operational outcome under full chronological evaluation.

From the perspective of feedback design, this comparison suggests that the Dayshape signal already captures the dominant planning-relevant information contained in the SOC diagnostics. The Hybrid setting, which combines the strongest day-shape signals with one strong overall-SOC signal and one overlapping day, does not overturn the planning logic of the Dayshape-based solution. Rather, it provides a small additional refinement on top of a Dayshape-dominated expansion structure.

Therefore, the similarity between the two solutions is itself an important result. It indicates that the low total cost obtained by the Dayshape and Hybrid settings is not caused by two entirely different decision patterns reaching similar objective values by coincidence. Instead, the two settings converge to closely related investment structures, with the Hybrid solution providing a slightly more effective fine-tuning of the final wind and storage allocation.

6.5.3 Comparison with the conventional peak-load extreme-day setting

The comparison with the peak-load rule should be interpreted in light of the role of input-defined extremes. Preserving a peak-load day can improve

adequacy representation, but it may also emphasize one type of system stress more strongly than other storage-relevant conditions. Prior work on extreme-period preservation shows that different extreme definitions, such as load peaks, renewable scarcity, and net-load peaks, can affect planning outcomes in different ways [10, 32]. Therefore, the peak-load setting is not necessarily inferior in general, but in the present case it produces a more conservative expansion pattern than the SOC-informed settings.

The comparison above shows that the Dayshape and Hybrid settings converge to highly similar investment structures. A further question is how these SOC-informed solutions differ from the conventional peak-load extreme-day rule. To address this point, Table 6.2 compares the best Dayshape and Hybrid solutions with a representative best solution obtained under the peak-load extreme-day setting.

Table 6.2: Comparison of investment decisions under the Dayshape, Hybrid, and peak-load extreme-day settings.

Category	**Dayshape best**	**Hybrid SOC-extreme best**	**Peak-load extreme best**
Investment cost	766,897,250	779,530,882	873,125,346
Full-model operation cost	1,600,934,903	1,587,924,278	1,499,101,185
Full-evaluated total cost	2,367,832,153	2,367,455,160	2,372,226,531
Transmission expansion	Line 36 (Bus 6–7) Line 37 (Bus 13–15)	Line 36 (Bus 6–7) Line 37 (Bus 13–15)	Line 36 (Bus 6–7) Line 37 (Bus 13–15)
Wind investment	Bus 4: 400 Bus 5: 266.46 Bus 13: 78.06 Bus 18: 400 Bus 22: 396.36	Bus 4: 400 Bus 5: 276.71 Bus 13: 95.49 Bus 18: 400 Bus 22: 400	Bus 4: 400 Bus 5: 375.44 Bus 13: 228.02 Bus 18: 400 Bus 22: 400
ESS investment	Bus 1: 1.03 Bus 3: 5 Bus 4: 80 Bus 5: 80 Bus 7: 8	Bus 1: 3.19 Bus 3: 5 Bus 4: 80 Bus 5: 80 Bus 7: 8	Bus 1: 25 Bus 3: 5 Bus 4: 80 Bus 5: 80 Bus 7: 8

Table 6.2 shows that the three extreme-day settings lead to the same transmission-expansion backbone, namely Lines 36 and 37. This suggests that, in this case study, the main network reinforcement requirement is relatively robust to the choice of extreme-day setting. The difference between the settings therefore does not mainly lie in the selected transmission topology, but in how the reduced model evaluates the need for wind and storage flexibility.

The peak-load extreme-day setting results in substantially larger wind capacities at Buses 5 and 13 and a much larger ESS investment at Bus 1. This can be interpreted as the effect of a stress-oriented temporal representation. Since the preserved extreme day is the maximum-load day, the reduced model gives explicit priority to a high-demand operating condition. This does not mean that the peak-load day is irrelevant; rather, it means that this setting emphasizes one particular type of system stress. Under this representation, additional wind and storage capacity become valuable as a hedge against the preserved high-load condition. As a result, the peak-load case produces a more conservative investment structure, with higher investment cost but lower full-time operational cost.

The SOC-informed settings modify this investment signal. Their preserved days are selected from storage-related diagnostic results instead of only from the maximum-load condition. In particular, the Dayshape setting preserves natural days whose intra-day SOC trajectories are poorly represented by the current representative-day set. These days are not necessarily conventional extreme days in terms of load or renewable availability. Their importance is that they reveal where the reduced model fails to capture storage-relevant daily charging and discharging patterns. Since representative days directly determine intra-day operating profiles, preserving such days can improve the reduced model's representation of storage operation in a way that is more directly connected to the representative-day mechanism.

This explains why the Dayshape and Hybrid settings lead to more restrained wind and ESS investment. They do not simply reduce investment arbitrarily. Instead, they reduce the dominance of the peak-load stress signal and replace it with a more storage-relevant representation of temporal mismatch. The reduced model therefore no longer interprets flexibility value mainly through the maximum-load day, but through days that expose weaknesses in the storage-operation representation. This produces lower wind expansion at Buses 5 and 13 and much smaller ESS capacity at Bus 1, while still maintaining good full-time operational performance.

The Hybrid solution can be understood as a compromise between the storage-shape signal and the broader SOC-error signal. It does not change the transmission backbone and remains close to the Dayshape solution in its overall investment structure. However, it introduces a limited adjustment in wind allocation compared with the pure Dayshape setting. This indicates that, in this case, most of the improvement already comes from correcting the intra-day storage-shape representation, while the additional SOC-error information in the Hybrid setting provides only a secondary refinement.

Therefore, the investment differences in Table 6.2 should not be interpreted only as numerical differences between three tested cases. They reflect how different extreme-day rules change the temporal signal seen by the reduced planning model. The peak-load setting emphasizes a demand-stress condition and therefore encourages a more conservative expansion strategy. The Dayshape and Hybrid settings emphasize storage-relevant representation errors and therefore provide a more targeted flexibility signal. This mechanism helps explain why the SOC-informed settings can achieve lower total cost without requiring a fundamentally different transmission-expansion pattern.

6.6 Summary and discussion

This chapter examined how SOC-based diagnostic information can be incorporated into the TSA process through modified extreme-day settings. The purpose was not to treat SOC-informed days as a guaranteed correction rule, but to test whether storage-related diagnostic signals can provide useful feedback for representative-day selection and improve the planning outcome of the reduced model.

The results are consistent with the broader movement from purely input-based TSA toward model-aware and feedback-based aggregation. Previous studies have shown that representative periods should be evaluated according to their effect on planning outcomes, and that extreme-period selection can materially change investment and operational decisions [32, 20]. This chapter extends this idea to storage-aware diagnostics by using SOC reconstruction errors to identify periods where the reduced representation poorly captures storage operation.

The main interpretation from the comparison is that different extreme-day settings change how the reduced model represents two planning-relevant

signals: system stress and storage value. System stress is mainly related to high demand or scarcity conditions, and is therefore naturally emphasized by the conventional peak-load extreme-day rule. Storage value, however, depends on whether the model correctly represents when and where storage can charge and discharge to shift energy from surplus periods or locations to scarcity periods or locations. Therefore, changing the preserved extreme days can change not only the represented stress condition, but also the model's interpretation of how useful storage is for balancing the system.

This mechanism helps explain the difference between the peak-load and SOC-informed settings. The peak-load setting emphasizes a demand-stress condition and therefore leads to a more conservative expansion pattern, with larger wind and storage investment in the tested case. By contrast, the Dayshape setting identifies natural days whose intra-day SOC trajectories are poorly represented by the current representative-day structure. These days are not necessarily extreme in terms of load demand, but they reveal that the reduced model does not reproduce important daily charging and discharging patterns. Since representative days directly determine intra-day operating profiles, preserving these days provides a more direct correction to the storage-value signal seen by the reduced model.

The performance difference between the SOC-informed settings also shows that not all SOC-based rankings are equally useful for extreme-day selection. The ABS-based setting is less effective because a large absolute SOC discrepancy may combine intra-day shape mismatch, accumulated drift, and inventory-level bias. Therefore, the day with the largest absolute SOC error is not always the day whose own profile should be preserved as a representative day. In contrast, the day-shape metric is more closely aligned with the representative-day mechanism because it focuses on whether the daily storage-operation pattern itself is poorly represented.

The Dayshape and Hybrid results suggest that, in this case study, correcting intra-day storage-shape representation is the main useful feedback mechanism. The Hybrid setting provides a small additional improvement by combining the day-shape signal with broader SOC-error information, but its investment structure remains close to the Dayshape solution. This indicates that most of the planning-relevant SOC feedback in this case is already captured by the intra-day shape mismatch signal.

These findings should be interpreted as conditional rather than universal. If a system is mainly affected by short-duration or daily-cycling storage behavior, intra-day shape mismatch is likely to be an informative feedback

signal because it directly reflects charging and discharging patterns within representative days. If long-duration storage or multi-day energy shifting becomes more dominant, then feedback signals based on consecutive-day blocks, accumulated drift, or inventory-level bias may become more relevant. Therefore, the general lesson is not that one specific SOC-informed rule is always superior, but that extreme-day selection should be matched with the storage mechanism that most strongly affects planning outcomes.

Taken together, the results show that SOC-based diagnostics are useful for feedback enhancement not because they directly minimize aggregate SOC error, but because they help identify storage-relevant temporal representation errors. In the present case, day-shape-oriented diagnostics improve the reduced model by giving a more targeted representation of storage value, while avoiding the over-conservative stress signal produced by the peak-load rule. More generally, the proposed approach supports a mechanism-aware form of extreme-day selection, where preserved days are chosen according to the temporal errors that most affect storage operation and investment decisions.

Chapter 7

Conclusion, Reflection, and Future Work

This thesis investigates how SOC-based diagnostic information can be used to interpret storage-state representation errors and guide feedback-enhanced extreme-day selection in representative-day-based power system expansion planning. Existing time-series aggregation methods can reduce computational complexity, preserve selected extreme periods, and reconstruct part of the chronology through representative days and sequentially linked days. However, in storage-embedded planning models, extreme-day selection is still usually driven by input-side stress indicators, such as peak load, high net load, or renewable scarcity, or by general model-output indicators. Storage-state representation itself is less frequently used as a feedback signal for selecting extreme days. This thesis therefore focuses on whether SOC diagnostics can reveal storage-related temporal representation errors and provide useful guidance for modifying the TSA setting.

The first research question concerns how reduced planning results can be positioned as the basis for storage-aware SOC reconstruction analysis. Chapter 4 establishes the reduced-model evaluation setting and identifies economically relevant reduced cases for further diagnostic study. These reduced cases provide the planning-result context in which SOC trajectories can be reconstructed and compared with the full-time-series model. The analysis shows that even when a reduced model produces a competitive planning result, the RD-SLD representation may still reconstruct storage trajectories differently from the full-time operation. This motivates the use of SOC trajectory analysis as a storage-aware diagnostic layer for understanding how the reduced

temporal structure represents storage behavior.

The second research question asks what storage-related error mechanisms can be revealed by SOC-based diagnostics. Chapter 5 shows that SOC reconstruction errors are not uniform. They can arise from intra-day shape mismatch, cross-day accumulated drift, and persistent inventory-level bias. Daily-cycling storage is more strongly affected by intra-day shape mismatch, while long-duration storage is more sensitive to accumulated drift and inventory-level bias. This distinction is important because a large absolute SOC error does not necessarily indicate that the daily representative profile itself is poorly selected. It may instead show where inter-period drift or inherited inventory bias becomes visible. Therefore, SOC diagnostics provide a mechanism-based interpretation of storage-state distortion under the RD-SLD representation.

The third research question asks whether SOC-based diagnostic signals can guide extreme-day selection and improve representative-day-based planning outcomes. Chapter 6 shows that SOC-informed extreme-day selection can be useful, but its effectiveness depends on which SOC signal is used. The day-shape-based and hybrid SOC-informed settings perform better than the conventional peak-load extreme-day setting in the tested case, while the absolute-SOC-error-based setting is less effective. This indicates that SOC diagnostics are most useful when they identify temporal representation errors that are directly related to the clustering and mapping structure. In particular, day-shape error is more suitable as feedback for representative-day selection because it focuses on intra-day SOC pattern mismatch after removing the daily starting SOC level.

The results also clarify why different extreme-day settings can lead to different investment decisions. Changing the preserved extreme days explicitly changes the TSA configuration. The preserved days affect the representative-day profiles generated by hierarchical clustering, the assignment of natural days to representative days, and the resulting SLD structure used for inter-day SOC propagation. As a result, different extreme-day settings change how the reduced model represents both system stress and storage value. The peak-load setting emphasizes a demand-stress condition, while the SOC-informed settings emphasize storage-relevant representation errors. In the case of intra-day shape mismatch, the diagnostic signal reflects whether the reduced model reproduces the daily net storage-operation pattern implied by charging and discharging. This changes the model's interpretation of when and where storage is valuable for shifting energy across time and, under

network constraints, across locations.

Overall, the thesis concludes that SOC diagnostics extending from the RD-SLD structure provide a useful storage-aware feedback layer for representative-day-based planning models with energy storage. Their role is not to directly minimize SOC reconstruction error, but to reveal storage-relevant temporal representation errors and identify candidate extreme days that can modify the reduced temporal structure in a planning-relevant way. In the present case, day-shape-oriented diagnostics provide the most effective feedback because they are closely aligned with the representative-day mechanism and the intra-day cycling behavior of storage. More generally, the proposed methodology suggests a potential way to improve power system planning models with energy storage by connecting storage-state diagnostics with mechanism-aware extreme-day selection.

7.1 Reflection

Several points should be interpreted carefully. First, the full-time benchmark used in this thesis is a computational reference solution rather than a proven global optimum. The full chronological model remains difficult to solve because it combines a large hourly operational space with investment variables. Therefore, the comparison should be understood as an evaluation against the best available full-time incumbent and bound, rather than as a proof of global optimality.

Second, SOC-informed feedback should not be interpreted as a rule that guarantees lower SOC reconstruction error after adding SOC-critical extreme days. Preserving new extreme days changes the representative-day profiles, the natural-day mapping, and the SLD reconstruction. This may improve the planning signal even if some aggregate SOC-error indicators do not decrease monotonically. Therefore, the value of SOC-informed feedback lies in how it modifies the reduced temporal representation and the resulting planning outcome, rather than in simple SOC-error minimization.

Third, the interpretation of investment changes should be treated as mechanism-based evidence rather than strict causal proof. When day-shape-error extreme days are preserved, the reduced model changes its representative-day structure and its reconstruction of intra-day cycling and inter-period energy shifting. The resulting investment decision may include less storage or flexibility investment because the model estimates system stress and storage

value differently. However, investment decisions and SOC trajectories are jointly determined by the optimization model. Therefore, the results support the interpretation that SOC-informed temporal representation changes the investment signal, but they do not isolate a single causal pathway from SOC error to investment cost.

Fourth, the performance difference between absolute-SOC and day-shape-based feedback highlights the importance of metric design. Aggregated absolute SOC error is useful for locating where the reconstructed trajectory differs most strongly from the full-time trajectory, but it mixes intra-day mismatch, accumulated drift, and inventory bias. It is therefore more suitable as a diagnostic summary than as a direct extreme-day selection rule. Day-shape error is more closely aligned with the representative-day selection problem because it focuses on the intra-day storage pattern that clustering directly affects.

Finally, the proposed SOC-informed feedback framework should be understood as mechanism-dependent. If a system is mainly driven by short-duration or daily-cycling storage, intra-day shape mismatch is likely to be an informative feedback signal because it reflects daily charging and discharging patterns. If long-duration storage or multi-day energy shifting becomes more dominant, feedback signals based on consecutive-day blocks, accumulated drift, or inventory-level bias may become more relevant. Therefore, the broader contribution is not one fixed extreme-day rule, but a framework for selecting feedback signals according to the storage mechanism that most strongly affects the reduced planning model.

7.2 Future Work

Future research can extend this work in several directions. First, the proposed SOC-informed feedback can be developed into an iterative procedure. In the present thesis, SOC-critical days are selected from one diagnostic stage and then tested in modified extreme-day settings. A natural extension is to repeat this process: solve the reduced model, reconstruct and evaluate SOC trajectories, identify new storage-critical days or SLD blocks, update the representative-day set, and re-evaluate the resulting investment decision. Such an iterative framework could test whether SOC-informed feedback converges toward more stable cost and investment outcomes.

Second, future work can develop more rigorous sensitivity-based feed-

back indicators. The present SOC metrics are diagnostic and descriptive: they identify where SOC reconstruction differs and what type of error mechanism is dominant. A stronger method would estimate how much preserving a candidate day is expected to change investment cost, operational cost, storage value, or feasibility margins. This could connect SOC diagnostics more directly with marginal planning value.

Third, the feedback framework can be extended from individual natural days to multi-day periods. Long-duration storage is often affected by consecutive low-renewable periods, accumulated energy imbalance, and seasonal inventory evolution. Therefore, selecting only single extreme days may not fully capture the stress conditions that drive long-duration storage investment. Future work could select SOC-critical SLD blocks or multi-day extreme periods instead of isolated days.

Fourth, the method should be tested on larger systems, multiple weather years, different renewable penetration levels, and different storage technologies. The present case study provides evidence that day-shape-based SOC feedback can improve total-cost performance, but the robustness of this conclusion should be evaluated across broader system conditions. In particular, systems with stronger solar penetration, more seasonal storage, different transmission bottlenecks, or higher renewable curtailment may exhibit different dominant SOC-error mechanisms.

Finally, future work can compare SOC-informed feedback with other model-aware extreme-day selection rules, such as operational-cost-error days, congestion-error days, load-shedding days, or renewable-curtailment days. Such comparison would clarify whether SOC diagnostics provide unique information beyond existing cost-based or operational-output-based feedback. This would further strengthen the positioning of SOC-informed extreme-day selection as a storage-aware extension of representative-period aggregation for expansion planning.

Bibliography

[1] M. Hoffmann, L. Kotzur, D. Stolten, and M. Robinius, "A review on time series aggregation methods for energy system models," *Energies*, vol. 13, no. 3, Art. no. 641, Feb. 2020. DOI: 10.3390/en13030641.

[2] H. Teichgraeber and A. R. Brandt, "Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities," *Renewable and Sustainable Energy Reviews*, vol. 157, Art. no. 111984, Apr. 2022. DOI: 10.1016/j.rser.2021.111984.

[3] S. Pineda and J. M. Morales, "Chronological time-period clustering for optimal capacity expansion planning with storage," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7162–7170, Nov. 2018. DOI: 10.1109/TPWRS.2018.2842093.

[4] R. Novo, P. Marocco, G. Giorgi, A. Lanzini, M. Santarelli, and G. Mattiazzo, "Planning the decarbonisation of energy systems: The importance of applying time series clustering to long-term models," *Energy Conversion and Management: X*, vol. 15, Art. no. 100274, Sep. 2022. DOI: 10.1016/j.ecmx.2022.100274.

[5] 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.

[6] B. Bahl, A. Kümpel, H. Seele, M. Lampe, and A. Bardow, "Time-series aggregation for synthesis problems by bounding error in the objective function," *Energy*, vol. 135, pp. 900–912, Sep. 2017. DOI: 10.1016/j.energy.2017.06.082.

[7] R. Hemmati, R.-A. Hooshmand, and A. Khodabakhshian, "Comprehensive review of generation and transmission expansion planning," *IET Generation, Transmission & Distribution*, vol. 7, no. 9, pp. 955–964, Sep. 2013. DOI: 10.1049/iet-gtd.2013.0031.

[8] E. Spyrou, J. L. Ho, B. F. Hobbs, R. M. Johnson, and J. D. McCalley, "What are the benefits of co-optimizing transmission and generation investment? Eastern interconnection case study," *IEEE Transactions on Power Systems*, vol. 32, no. 6, pp. 4265–4277, Nov. 2017. DOI: 10.1109/TPWRS.2017.2660249.

[9] X. Zhang and A. J. Conejo, "Coordinated investment in transmission and storage systems representing long- and short-term uncertainty," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7143–7151, Nov. 2018. DOI: 10.1109/TPWRS.2018.2842045.

[10] A. García-Cerezo, R. García-Bertrand, and L. Baringo, "Priority chronological time-period clustering for generation and transmission expansion planning problems with long-term dynamics," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4325–4339, Nov. 2022. DOI: 10.1109/TPWRS.2022.3151062.

[11] S. Gonzato, K. Bruninx, and E. Delarue, "Long term storage in generation expansion planning models with a reduced temporal scope," *Applied Energy*, vol. 298, Art. no. 117168, Sep. 2021. DOI: 10.1016/j.apenergy.2021.117168.

[12] D. A. Tejada-Arango, M. Domeshek, S. Wogrin, and E. Centeno, "Enhanced representative days and system states modeling for energy storage investment analysis," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6534–6544, Nov. 2018. DOI: 10.1109/TPWRS.2018.2819578.

[13] 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.

[14] H. Teichgraeber and A. R. Brandt, "Clustering methods to find representative periods for the optimization of energy systems: An initial

framework and comparison," *Applied Energy*, vol. 239, pp. 1283–1293, Apr. 2019. DOI: 10.1016/j.apenergy.2019.02.012.

[15] R. Green, I. Staffell, and N. Vasilakos, "Divide and conquer? k-means clustering of demand data allows rapid and accurate simulations of the British electricity system," *IEEE Transactions on Engineering Management*, vol. 61, no. 2, pp. 251–260, May 2014. DOI: 10.1109/TEM.2013.2284386.

[16] Y. Liu, R. Sioshansi, and A. J. Conejo, "Hierarchical clustering to find representative operating periods for capacity-expansion modeling," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 3029–3040, May 2018. DOI: 10.1109/TPWRS.2017.2746379.

[17] M. Moradi-Sepahvand and S. H. Tindemans, "Capturing chronology and extreme values of representative days for planning of transmission lines and long-term energy storage systems," in *Proc. 2023 IEEE Belgrade PowerTech*, Belgrade, Serbia, Jun. 2023, pp. 1–6. DOI: 10.1109/PowerTech55446.2023.10202993.

[18] J. H. Ward Jr., "Hierarchical grouping to optimize an objective function," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 236–244, Mar. 1963. DOI: 10.1080/01621459.1963.10500845.

[19] D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, C. L. Lara, and I. E. Grossmann, "Impact of model resolution on scenario outcomes for electricity sector system expansion," *Energy*, vol. 163, pp. 1231–1244, Nov. 2018. DOI: 10.1016/j.energy.2018.08.015.

[20] R. Zhang, "Feedback enhancement of time series aggregation for power system expansion planning," M.Sc. thesis, Delft University of Technology, Delft, The Netherlands, 2024.

[21] L. L. Garver, "Transmission network estimation using linear programming," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-89, no. 7, pp. 1688–1697, Sep. 1970. DOI: 10.1109/TPAS.1970.292825.

[22] C. Li, A. J. Conejo, P. Liu, B. P. Omell, J. D. Siirola, and I. E. Grossmann, "Mixed-integer linear programming models and algorithms for

generation and transmission expansion planning of power systems," *European Journal of Operational Research*, vol. 297, no. 3, pp. 1071–1082, Mar. 2022. DOI: 10.1016/j.ejor.2021.06.024.

[23] A. Viana and J. P. Pedroso, "A new MILP-based approach for unit commitment in power production planning," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 997–1005, Jan. 2013. DOI: 10.1016/j.ijepes.2012.08.046.

[24] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011. DOI: 10.1109/TPWRS.2010.2051168.

[25] B. Stott, J. Jardim, and O. Alsac, "DC power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009. DOI: 10.1109/TPWRS.2009.2021235.

[26] S. Frank and S. Rebennack, "An introduction to optimal power flow: Theory, formulation, and examples," *IIE Transactions*, vol. 48, no. 12, pp. 1172–1197, 2016. DOI: 10.1080/0740817X.2016.1189626.

[27] E. Ela, M. Milligan, and B. Kirby, "Operating reserves and variable generation," National Renewable Energy Laboratory, Golden, CO, USA, Tech. Rep. NREL/TP-5500-51978, Aug. 2011.

[28] P. Nahmmacher, E. Schmid, L. Hirth, and B. Knopf, "Carpe diem: A novel approach to select representative days for long-term power system modeling," *Energy*, vol. 112, pp. 430–442, Oct. 2016. DOI: 10.1016/j.energy.2016.06.081.

[29] B. A. Frew and M. Z. Jacobson, "Temporal and spatial tradeoffs in power system modeling with assumptions about storage: An application of the POWER model," *Energy*, vol. 117, pp. 198–213, Dec. 2016. DOI: 10.1016/j.energy.2016.10.074.

[30] M. Hoffmann, L. Kotzur, D. Stolten, and M. Robinius, "Typical periods or typical time steps? A multi-model analysis to determine the optimal temporal aggregation for energy system models," *Applied Energy*, vol. 304, Art. no. 117825, Dec. 2021. DOI: 10.1016/j.apenergy.2021.117825.

[31] C. Li, A. J. Conejo, P. Liu, B. P. Omell, and I. E. Grossmann, "On representative day selection for capacity expansion planning of power systems under extreme operating conditions," *International Journal of Electrical Power & Energy Systems*, vol. 137, Art. no. 107697, May 2022. DOI: 10.1016/j.ijepes.2021.107697.

[32] A. Yeganefar, A. G. Givisiez, P. A. Ahmadi, and M. P. Moghaddam, "Improvement of representative days selection in power system planning by incorporating the extreme days of the net load demand," *Applied Energy*, vol. 272, Art. no. 115224, Aug. 2020. DOI: 10.1016/j.apenergy.2020.115224.

[33] "Annual peak load and load curve - TenneT." [Online]. Available: https://netztransparenz.tennet.eu/electricity-market/transparency-pages/transparency-germany/network-figures/annual-peak-load-and-load-curve/.

[34] M. Moradi-Sepahvand, *Mojtabamoradii/Modified-RTS-24-bus: The data for Modified-RTS-24-bus test system used as study system for paper entitled "Representative Days and Piecewise Linear Hours for Power System Planning with Long-Term Storage"*. [Online]. Available: https://github.com/Mojtabamoradii/Modified-RTS-24-bus/tree/main.

[35] B. Stott, J. Jardim, and O. Alsaç, "DC power flow revisited," *IEEE Transactions on Power Systems*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009. DOI: 10.1109/TPWRS.2009.2021235.

[36] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 12–19, Feb. 2011. DOI: 10.1109/TPWRS.2010.2051168.

[37] S. Frank and S. Rebennack, "An introduction to optimal power flow: Theory, formulation, and examples," *IIE Transactions*, vol. 48, no. 12, pp. 1172–1197, 2016. DOI: 10.1080/0740817X.2016.1189626.

[38] A. P. Hilbers, D. J. Brayshaw, and A. Gandy, "Reducing climate risk in energy system planning: A posteriori time series aggregation for models

with storage," *Applied Energy*, vol. 334, Art. no. 120624, Mar. 2023. DOI: 10.1016/j.apenergy.2022.120624.

[39] T. Levin, P. L. Blaisdell-Pijuan, J. Kwon, and W. N. Mann, "High temporal resolution generation expansion planning for the clean energy transition," *Renewable and Sustainable Energy Transition*, vol. 5, Art. no. 100072, Aug. 2024. DOI: 10.1016/j.rset.2023.100072.

[40] L. Reichenberg and F. Hedenus, "The error induced by using representative periods in capacity expansion models: System cost, total capacity mix and regional capacity mix," *Energy Systems*, vol. 15, pp. 215–232, 2024. DOI: 10.1007/s12667-022-00533-4.

[41] H. Teichgraeber, C. P. Lindenmeyer, N. Baumgärtner, L. Kotzur, D. Stolten, M. Robinius, A. Bardow, and A. R. Brandt, "Extreme events in time series aggregation: A case study for optimal residential energy supply systems," *Applied Energy*, vol. 275, Art. no. 115223, Oct. 2020. DOI: 10.1016/j.apenergy.2020.115223.

[42] H. Teichgraeber and A. R. Brandt, "Designing reliable future energy systems by iteratively including extreme periods in time-series aggregation," *Applied Energy*, vol. 304, Art. no. 117696, Dec. 2021. DOI: 10.1016/j.apenergy.2021.117696.