Identifying Representative Weeks for the Finnish Power Sector with Time-Series Clustering

Ilona Ylikoski

School of Science

Bachelor's thesis
Espoo October 1, 2025

Supervisor

Prof. Afzal Siddiqui

Advisor

Prof. Afzal Siddiqui

Aalto University
School of Science

Aalto University
School of Science

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi
Abstract of the bachelor's thesis

Author Ilona Ylikoski

Title Identifying Representative Weeks for the Finnish Power Sector with Time-Series Clustering

Degree programme Bachelor's Programme in Science and Technology

Major Mathematics and Systems Science **Code of major** SCI3029

Teacher in charge Prof. Afzal Siddiqui

Advisor Prof. Afzal Siddiqui

Date October 1, 2025 **Number of pages** 30+5 **Language** English

Abstract

The share of power generation from variable renewable energy sources (VRE) has increased in the Finnish power system. Intermittent production from VRE plants increases model complexity, creating demand for methods to reduce the temporal resolution of models. Time-series aggregation (TSA) methods, such as representative periods, achieve complexity reduction by representing long time periods with fewer time steps.

This thesis studies power demand as well as solar and wind availability data from the Finnish power system in 2023. Two different approaches to selecting four representative weeks from the data are studied with the goal of assessing their ability to retain key features of the original time series. The seasonal approach selects a week from four predetermined seasons based on their mean and standard error. By contrast, the partitional approach applies the clustering algortihm partitioning around medoids (PAM), which compares each pair of weeks to each other, and groups the data into four clusters. Various methods for measuring dissimilarity of weeks are compared with the partitional method. The representativity of the selected weeks is compared in terms of both statistical values and a benchmark optimisation model.

Both approaches were found to underrepresent extreme values, which led to significantly lower objective-function values than in the full time resolution (FTR) benchmark model. Including solar power availability as a factor in determining the representative weeks did not improve results.

The seasonal approach performed especially well in terms of the statistical value metrics. However, it had high errors in the objective-function value. The partitional approach struggled more in retaining the means of the data, but had smaller errors in the objective function value. Implementing dynamic time warping for the dissimilarity computation with the partitional approach was found to improve aggregation of wind availability.

Keywords Time-Series Aggregation , Power Systems , Clustering , Optimisation

Aalto University
School of Science

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi
Tekniikan kandidaatintyön tiivistelmä

Tekijä Ilona Ylikoski

Työn nimi Edustavien viikkojen määrittäminen Suomen sähköjärjestelmälle aikasarjojen ryvästysmenetelmillä

Koulutusohjelma Teknistieteellinen kandidaattiohjelma

Pääaine Matematiikka ja systeemitieteet **Pääaineen koodi** SCI3029

Vastuuopettaja Prof. Afzal Siddiqui

Työn ohjaaja Prof. Afzal Siddiqui

Päivämäärä October 1, 2025 **Sivumäärä** 30+5 **Kieli** Englanti

Tiivistelmä

Uusiutuvien energialähteiden osuus Suomen sähköntuotannosta on kasvanut merkittävästi viime vuosina. Tuuli-, vesi- ja aurinkovoiman tuotannon vaihtelevuus kasvattaa sähköverkkoa kuvaavien mallien laskennallista vaativuutta tehden tarpeen mallien ajallisen ulottuvuuden pienentämiseen. Aikasarjojen aika-aggregointimenetelmät (engl. time-series aggregation), kuten edustavien aikavälien (engl. representative periods) valitseminen, mahdollistavat kompleksisuuden vähentämisen, kun täyttä aikaresoluutiota (engl. full time resolution) voidaan kuvata pienemmällä määrällä aika-askelia.

Tässä kandidaatintyössä tutkitaan dataa sähkön kysynnästä sekä aurinko- ja tuulivoiman saatavuudesta Suomen sähköjärjestelmässä vuonna 2023. Datasta identifioidaan neljä edustavaa viikkoa kahden eri lähestymistavan varianttien avulla pyrkien samalla säilyttämään alkuperäisten aikasarjojen keskeisimmät piirteet. Vuodenaikamenetelmässä jokaisesta neljästä ulkoisesti määritellystä vuodenajasta valitaan edustava viikko perustuen eroihin keskiarvossa ja -hajonnassa. Ositusmenetelmässä hyödynnetään medoidien ympärille osittaminen -ryvästysalgoritmia (engl. Partitioning Around Medoids), joka ryhmittelee viikot perustuen niiden keskinäisiin eroavaisuuksiin. Viikkojen eroavaisuuden mittaamisen käytetään useita eri mittoja ja niiden valinnan vaikutuksia tuloksiin analysoidaan. Edustavien viikkojen laatua arvioidaan sekä tilastollisten tunnuslukujen että mittapuuoptimointimallin tulosten perusteella.

Kummallakin lähestymistavalla datan äärimmäiset arvot jäivät aliedustetuiksi edustavissa viikoissa, mikä johti merkittävästi matalampaan mittapuumallin kohdefunktion arvoon verrattuna täyden aikaresoluution malliin. Aurinkovoiman saatavuuden sisällyttäminen muuttujaksi edustavien viikkojen valinnassa ei parantanut tuloksia kummallakaan lähestymistavalla.

Vuodenaikamenetelmä suoriutui erityisen hyvin tilastollisten tunnuslukujen arvojen säilyttämisessä, mutta sai ison negatiivisen virheen kohdefunktiossa. Toisaalta ositusmenetelmällä oli enemmän vaikeuksia datan keskiarvon säilyttämisessä, mutta se saavutti vuodenaikamenetelmää selvästi pienempiä virheitä kohdefunktiossa. Dynaamisen ajan vääristämisen (engl. Dynamic Time Warping) hyödyntäminen ositusmenetelmässä paransi tuulivoiman saatavuuden aggregaatiota.

Avainsanat Ryvästys, Aika-aggregointi, Sähköjärjestelmät, Optimointi

Contents

Abbreviations and Acronyms

CHP	Combined Heat and Power
DTW	Dynamic Time Warping
Eucl.	Euclidean distance
Manh.	Manhattan distance
PAM	Partitioning Around Medoids
ROR	Run of the river
TSA	Time-series aggregation
VRE	Variable renewable energy

1 Introduction

Finland has the goal of carbon neutrality by the year 2035 (Ministry of Economic affairs and Employment, 2022). As a result, the power system is increasingly more reliant on variable renewable energy sources (VREs), such as wind and solar power. The irregular production of VREs and the storage dynamics of hydropower and possible future battery storage add complexity and pose challenges in modelling the power sector (Hoffmann et al., 2020).

The increasingly complex models can be computationally intractable in their full temporal resolution. To address this problem, time-series aggregation (TSA) methods have been applied to models in order to reduce their temporal dimension. TSA refers to representing a longer dataset with fewer time steps. This allows the formulation of simpler, lower time dimension problems, thereby reducing the computational cost when compared to the full time resolution.

TSA methods have been widely applied to modern energy-system models (Hoffmann et al., 2020). Research on the topic has shown that the performance of TSA methods depends strongly on the characteristics of the system they are applied to (Hoffmann et al., 2022). Hence, understanding the strengths and shortcomings of different TSA methods applied to such models is important for interpretations of their results. Modelling power systems adequately is crucial in analysing their functioning and economic impact.

Many of the legacy TSA methods are based on approximating the data, rather than utilising the original time series themselves. However, as VRE production increases in the system, the legacy TSA methods, such as load duration curve approximation, become incapable of retaining the intermittency of the data (Newbery, 2023). Hence, this thesis focuses on variants of the representative-period method, where representative weeks are chosen from the data. With the representative period method, the aggregation is formed using the data from the selected weeks, with weights corresponding to the share of the year they are representing.

Representative days are more commonly used than weeks (Hoffmann et al., 2020); however, when considering models with intertemporal constraints longer periods are required (Reichenberg et al., 2018). The Finnish power sector has a significant presence of hydropower that can store water in reservoirs, calling for the use of such intertemporal constraints. Thus, it is of interest to study how different TSA methods perform when applied to selecting longer time periods.

This thesis analyses two alternative approaches for aggregating a dataset from the Finnish power system. The seasonal and partitional approaches are used to identify four representative weeks for power demand as well as wind and solar power availabilities in the year 2023. The goal of both approaches is to represent the full year with only the selected four weeks. The seasonal approach chooses one week from each one of four predetermined seasons, corresponding to Winter, Spring, Summer, and Autumn (Hassanzadeh Moghimi et al., 2023), whereas the partitional approach applies a clustering algorithm to define four groups in the data.

The goal of this thesis is to analyse the strengths and weaknesses of each approach in capturing the characteristics of the original time series. Representativity is

measured by comparing the aggregated data to the original in terms of the values of descriptive statistics and the results of a benchmark optimisation model when compared to a full time resolution (FTR) model.

The structure of this thesis is as follows. First, Chapter 2 discusses the Finnish power sector and time series aggregation in more detail, before providing context for the objectives of the study. Chapter 3 introduces the dataset and discusses the main features of each time series. Then, Chapter 4 formulates the approaches assess in detail and describes the methodology used to analyse their performance. Chapter 5 shows the results of implementing the analysis, and Chapter 6 summarises the main findings.

2 Background

2.1 Finnish Power Sector

The Finnish power sector forms a zone within the larger Nordic power system, which also includes Sweden, Norway and Denmark (Nord Pool, 2025). The Nord Pool system is depicted in Figure 1. Transmission lines connect the Finnish system to its neighbouring countries in the Nordics and Baltics, allowing for import and export of electricity. Finnish electricity prices are set in the day-ahead market for each hour of the following day. Unlike the other Nord Pool member countries, Finland only has one price zone, which sets the same price throughout the country (Hassanzadeh Moghimi et al., 2023).

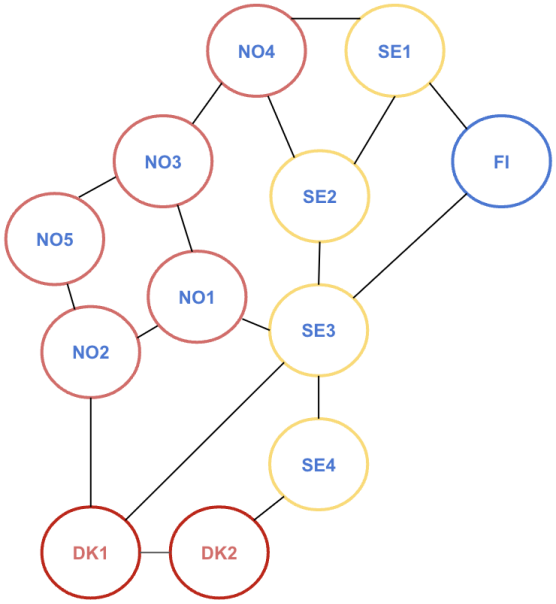


Figure 1: A visualisation of the Nord Pool system's price zones in Norway (NO1-NO5), Sweden (SE1-SE4), Denmark (DK1-DK2) and Finland (FI) (Nord Pool, 2025)

Events in the various nodes of the Nordic system affect each other. The dynamics of a multi-node system are more complex and computationally heavy to model.

To ensure computational tractability of the optimisation problem developed as a benchmark, this thesis focuses only on the Finnish system and price zone.

The power generation originates from a mix of different generation types: hydro, wind, solar, nuclear, and other thermal power plants. The net electricity generation by generation type is shown in Figure 2. Each generation method has its unique features and limiting factors that affect the system as a whole. Due to the cold climate, power demand and supply fluctuate greatly throughout the year (Figure 3). Therefore, the variety in production methods and their interactions in different seasons are of central interest when modelling the system.

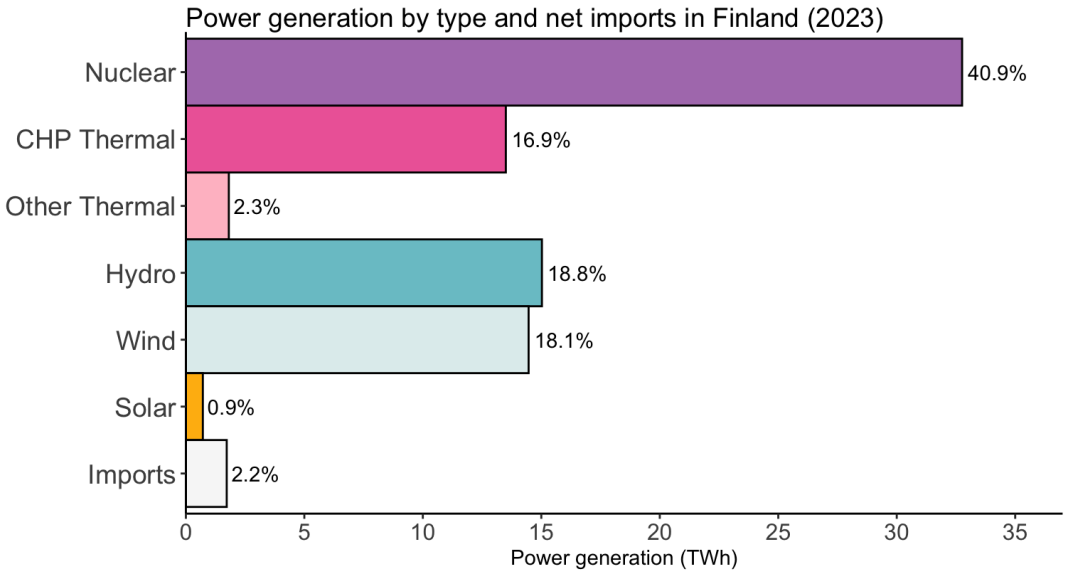


Figure 2: Electricity generation by type in Finland in 2023 (Statistics Finland, 2025)

The largest source of power generation in Finland is nuclear power, which produced 41% of total power in 2023 (Statistics Finland, 2025). Nuclear plants provide stable levels of CO_2-free power generation, but their output levels are slower to adjust than other types of thermal plants. It should be noted that the Olkiluoto 3 power plant began regular production in May 2023, which increased nuclear power generation capacity by 1.6 GW, leading to a 57% increase in total nuclear power capacity (Teollisuuden Voima Oy, 2023).

Additionally, there are non-nuclear thermal plants that can be utilised in the system. Most of these thermal plants are combined heat and power (CHP) plants, meaning they are capable of generating heat in addition to electricity. Finnish thermal plants use a variety of different fuels, including biomass, coal, peat coal, oil, waste, and gas (Virasjoki et al., 2018; Statistics Finland, 2025).

Hydropower is also a major source of power generation (Statistics Finland, 2025). According to Hassanzadeh Moghimi et al. (2023), hydropower plants in Finland can be further classified into ones solely utilising river waterflows and ones with a reservoir. Hydropower plants connected to a reservoir can regulate their production by adjusting the water level of the reservoir. This mechanism makes it possible to store power during low demand, and, to increase generation when demand is higher.

Generation from run-of-river (ROR) plants, which do not have reservoirs, on the other hand, is determined solely by the inflows of the watersheds, in which they are located.

Production capacity of wind power has grown in Finland in recent years (Fingrid, 2024d). During the year 2023 the production from wind power rose close to the level of hydropower (Statistics Finland, 2025). Solar power, on the other hand, has yet to gain a large foothold in the market, and it only has a very small share of total power generation (Fingrid, 2024c). Solar availability is also very seasonal, due to the high variance in the amount of daylight hours throughout the year.

2.2 Time-Series Aggregation for Energy-System Models

Time-series aggregation (TSA) refers to representing time-series data with a reduced number of time steps, while striving to maintain their general properties. In recent years, as the complexity of energy-system models have increased. Hence, numerous methods for TSA have been developed and discussed with the goal of applying them to such models (Hoffmann et al., 2020).

Research has shown that various factors must be considered when choosing a method for time-series aggregation. Many studies on TSA and representative periods on energy systems do not include models with intertemporal constraints (Teichgraeber and Brandt, 2022). The cross-time-step dynamics created by hydropower storage and the varying ramp-up and ramp-down rates of thermal power plants provide a restriction in modelling the Finnish power system (Hassanzadeh Moghimi et al., 2023). These dynamics are depicted with intertemporal constraints, meaning, that the values of decision variables at one time step constrain the values at other time steps. The interest in these dynamics increases as the share of renewables grows in power systems, and it becomes possible to introduce battery storage in the system.

Hoffmann et al. (2022) argue that the number, similarity, and cross-correlation of time series being aggregated simultaneously are important to consider when choosing an aggregation method. Hence, the features of the dataset must be analysed and taken into account before choosing a method. Teichgraeber and Brandt (2022) argues that aggregating wind data, requires "extra care". In practice, this means that the stochasticity of wind data must be specifically taken into account when choosing a method.

Finding representative weeks is a variant of the broader TSA method of representative periods, also known as typical periods. In the representative-days method, a preselected number of representative time periods are identified from a larger dataset. The original data is then depicted by assigning weights to each representative period, leading to fewer total time steps and, thus, decreased complexity (Reichenberg et al., 2018). In this thesis, the length of each period is a week; however, representative days have been studied more.

Although this thesis focuses on variants of the representative-weeks method, various other TSA methods have been developed. Alternative methods to representative periods include downsampling, merging and segmentation (Aghabozorgi et al., 2015; Hoffmann et al., 2020) among others.

In downsampling, the data is represented by a sampled subset of the original data. For example, only every other time step could be selected and thus the temporal dimension would be reduced to one half of the original (Hoffmann et al., 2020). Merging-based approaches combine multiple (adjacent) time steps into a longer time step, which thus represents a longer time period Aghabozorgi et al. (2015). The segmentation approach is similar to the representative-periods method, with the alteration that the representative segments used may vary in length (Hoffmann et al., 2022). More complex TSA methods combine multiple different types of approaches.

Load duration curve approximation based methods compare time periods by sorting the time steps of each period in descending order, leading to an approximation of the value distribution of time series. These curves can then be compared to find a representative one. However, this does not work for VRE, as it ignores the intermittency of the data due to the sorting that takes place (Newbery, 2023).

In Hassanzadeh Moghimi et al. (2023), a study of energy firms' strategic behaviour in the Nordic power sector, a season-based approach was used to find representative weeks for electricity price, consumption, and solar- and wind-power availabilities. In this approach, four seasons were defined in the data, from which representative weeks were selected based on minimising seasonal-weekly errors in weekly averages and standard errors (see Section 4.1.3).

Among the methods for TSA, time-series clustering is an increasingly common method to choose representative periods. In a review, Hoffmann et al. (2020) classified 130 different publications where TSA methods were implemented, and argued that "the future of TSA in energy-systems is in feature-based methods, including clustering," Hence, it is of interest to study how the simpler seasonal approach compares to a clustering-based approach.

The performance of some clustering-based TSA methods has been studied in the context of aggregation for power and energy-system models, for example by Reichenberg et al. (2018), Teichgraeber and Brandt (2019), and Hoffmann et al. (2022).

Reichenberg et al. (2018) applied k-means clustering to finding both representative days and system states (i.e., single representative time-steps) for demand, wind, and solar data. The results were compared to capacity expansion benchmark models from various European regions. The shares of VRE in production capacity and system cost were used as the metrics for representativity. The study found that both approaches struggled with underestimating VRE capacity if they were not taken into account in the aggregation process.

Hoffmann et al. (2022) compared aggregation results of k-means, k-medoids, and an algorithm based on preserving the value distribution of the full time resolution in finding typical days and time segments. One of the benchmark optimisation models analysed was a capacity-expansion model for 96 different regions in Europe. It was found that although objective-function values close to the original were achieved, the effect of aggregation VRE technologies varied by method and number of representative time periods used.

Teichgraeber and Brandt (2019) applied various partitional clustering methods and distance measures on energy price datasets. They found that centroid-based

clustering methods systematically undervalue the objective-function, whereas k-medoids approaches lead to higher variance in results. Similarly to Reichenberg et al. (2018) and Hoffmann et al. (2022), the authors also emphasised that it is important to analyse the performance of each method based on a benchmark optimisation problem, rather than the clustering error.

2.3 Objectives

As surveyed in Section 2.2, there are many possible approaches for the selection of representative weeks. Two approaches are selected for further analysis in this thesis; the seasonal approach used in Hassanzadeh Moghimi et al. (2023) and a partitional approach, where a clustering algorithm, similar to the approaches studied in Reichenberg et al. (2018) and Teichgraeber and Brandt (2022).

The two methods use a very different approaches for selecting which sections of the year each representative week represents. The exogenously defined four seasons of the seasonal approach, define the groups before selecting the weeks and thus, ensure the adjacency of the sections of the year each representative week stands in for in the aggregation. The clustering algorithm used in the partitional approach, on the other hand, analyses the dissimilarity of individual pairs of weeks and groups them without regarding their relative positions within the year. The selection of the grouping and the representative week is, thus, done at the same step. This fundamental difference in the two makes it of interest to study how well each approach performs relative to one another.

A common result of many TSA methods is the underrepresentation of extreme values (Hoffmann et al., 2020). Teichgraeber and Brandt (2022) argue that this outcome is due to the tendency of clustering to smooth time series. In the case of the power system, extreme values are, for example, peaks in demand during the winter or low wind production during a day with fair weather. As it is important for the power system to function in these situations, the effect of the lack of representation for these situations should be considered. Therefore, each approach is analysed specifically in terms of its ability to retain an objective-function value close to that of the full time resolution, even after the number of time steps is reduced. This is done by developing a benchmark optimisation model. However, it has also been shown by Reichenberg et al. (2018) and Hoffmann et al. (2022), that low error in the objective function does not necessarily indicate that VRE data is well aggregated. Hence, other metrics, such as descriptive statistics and VRE generation in the benchmark, are also studied to form a more comprehensive picture of the ability of each approach to retain key attributes of the time series.

Many different clustering algorithms have been used to cluster data for representative periods Hoffmann et al. (2020). This thesis focuses on one medoid-based algorithm (Partitioning Around Medoids) (see Section 4.2.3). To keep the comparison of the two approaches reasonable, the number of clusters is fixed at four, although the clustering algorithm would enable making more numerous clusters as well.

The algorithm is combined with six different metrics for measuring the dissimilarity of each pair of weeks, in order to assess how the challenging wind data can be best

compared (see Section 4.2.2). Specifically, dynamic time warping (DTW) alignments are considered as an alternative to comparing only the same time steps of each week. This is done to address the aperidicity of the wind data, where the peaks and valleys of wind may be misaligned.

Additionally, the effect of including solar availability in the selection criteria of each approach is analysed. This is due to the low installed solar-power-generation capacity in the Finnish power grid in relation to the other variables.

3 Dataset

The dataset used in this thesis consists of power demand as well as wind and solar power availabilities in the Finnish power sector during the year 2023. Each time series has 8760 hourly time steps, which correspond to the hours of the year. The profiles of the time series are shown in Figure 3.

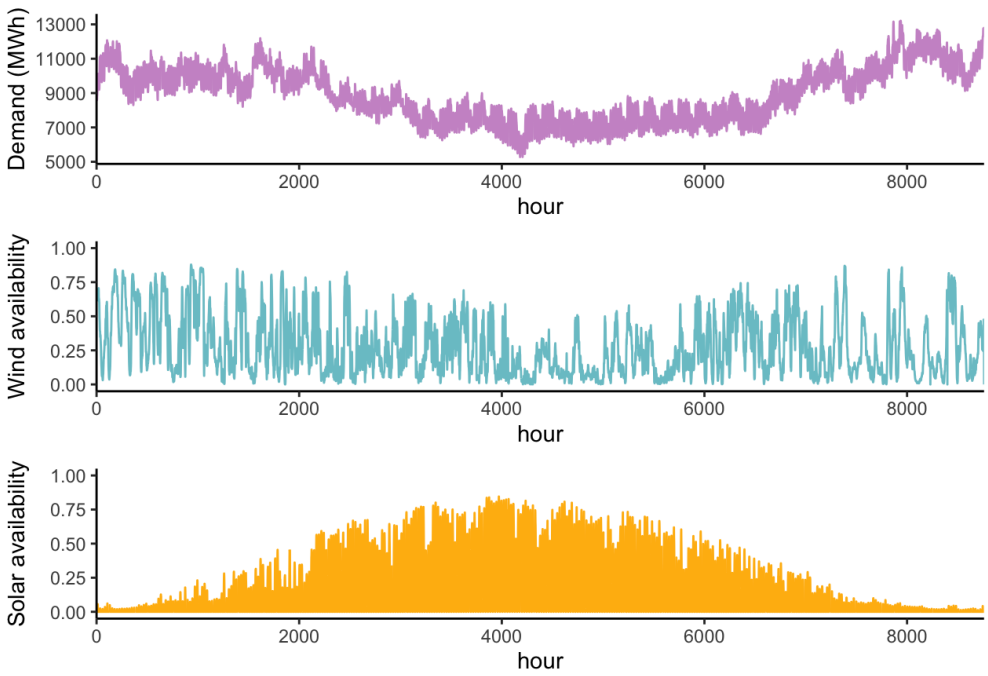


Figure 3: Time series for Finnish demand as well as wind and solar availabilities in 2023 (Nord Pool, 2024; Fingrid, 2024a,b,c,d)

52 full weeks are defined from the dataset representing the whole year. Each week consists of 168 time steps, beginning with the first hour of a Sunday. As $52 \times 168 = 8736$, the last 24 hours of the year are, hence, not included in the weeks examined. Weeks 1 and 27 are plotted next to each other in Figure 4. The weeks are examples of a week during the winter and the summer, respectively, and highlight the seasonal, daily, and hourly differences in the time series.

The data used for electricity demand is sourced from Nord Pool and is represented by the hourly data for consumption at the FI price zone of the market in megawatt

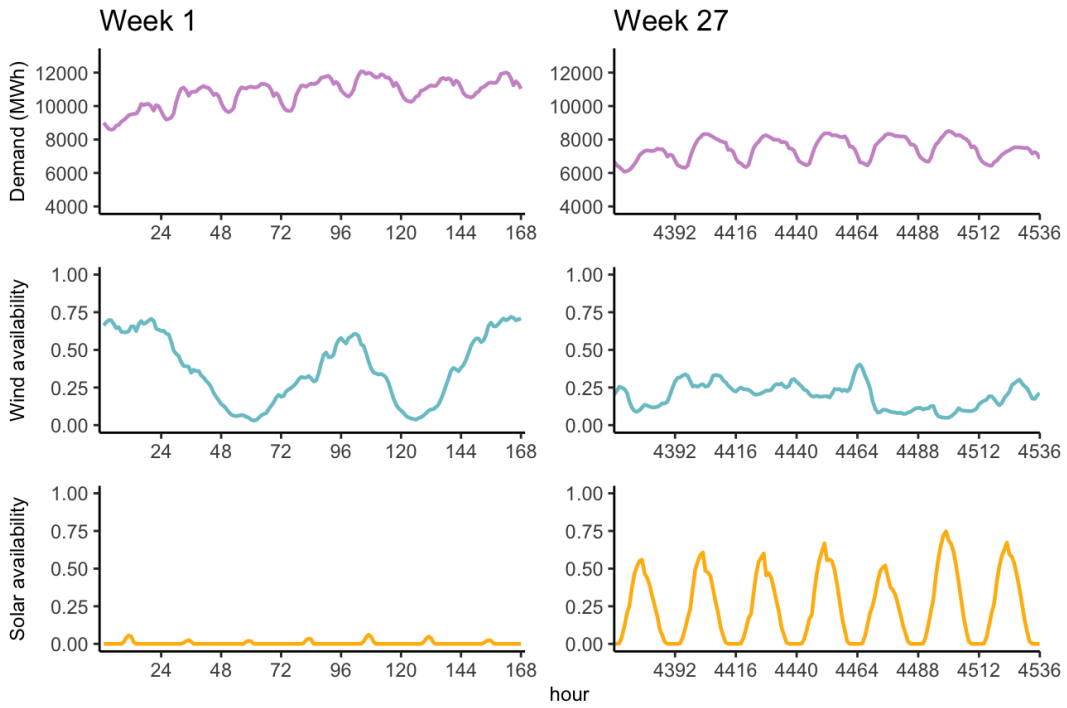


Figure 4: Data profiles of weeks 1 and 27, highlighting the differences in the time series in higher resolution

hours (MWh) (Nord Pool, 2024). Consequently, any demand for export is ignored. The demand data has periodicity on the daily and weekly levels. Demand is higher during the day and on weekdays and lower in the night and on weekends. There are also distinct differences between the demand of warmer and cooler months.

The data for wind and solar power availabilities were computed from the corresponding total production and generation capacity estimates by Fingrid (Fingrid, 2024a,b). The values for availability were obtained by dividing the hourly generation values by the estimated total generation capacity at said hour. This allows for normalisation in terms of the increase in generation capacity throughout the year. The values for wind and solar availabilities range from 0 to 1 and represent the share of current capacity utilised at each hour.

A 24-hour period is visible in the solar data, as well as a rising trend in the beginning of the year. The trend turns into a negative one after the summer solstice in June, after which the daylight hours start to decrease. On the other hand, the wind data is stochastic, and there are no visible trends in the data.

4 Methods

This chapter describes the two approaches applied to the dataset and the criteria used to evaluate their performance. First, Section 4.1 discusses the details of the seasonal approach and then Section 4.2 discusses the partitional approach. Section

4.3 presents the benchmark problem modelling the Finnish power sector as well as the statistical values for the aggregated time series that are also analysed.

4.1 Seasonal Approach

The seasonal approach is the first TSA method examined in this thesis. The approach starts by first exogenously defining four "seasons." One representative week is chosen for each season by choosing the week that has the closest mean and standard error to that of the whole season. The seasonal approach is adapted from Hassanzadeh Moghimi et al. (2023) and is implemented with an adjusted version of the code used in the project. The code, written for the optimisation software General Algebraic Modeling System (GAMS), was modified so that it could work with the dataset used in this thesis. Each step of the approach is described in the following subsections: 4.1.1 further specifies the definitions of the seasons; 4.1.2 describes the normalisation required in the approach, and 4.1.3 gives the precise formulation the statistic used to decide the representative weeks.

4.1.1 Definition of Seasons

In the seasonal approach, the hours of the year are first divided into four seasons. The same partition was used as by Hassanzadeh Moghimi et al. (2023), which was determined by analysing the annual hydropower inflow in the Nordic countries. The seasons are partitioned at hours 1896, 4104, 6312, and 8496.

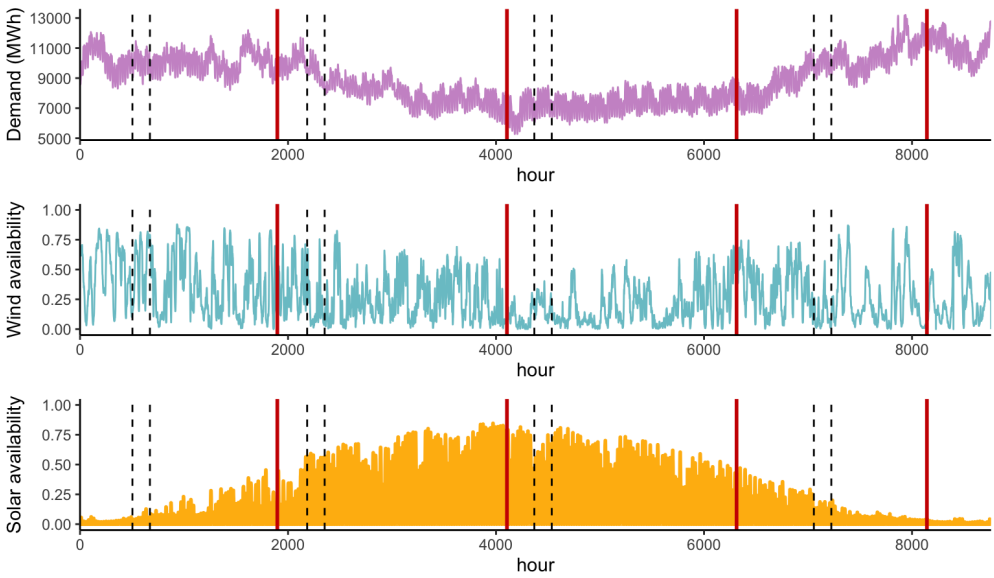


Figure 5: Divisions of the seasons are visualised with the demand data

The division into seasons is visualised in Figure 5, where season boundaries are indicated by the red lines. A representative week, depicted as the area between each pair of dotted lines, is chosen within each season. The data corresponding to the

same week represents, each of the three time series for that season. The seasons, referred to as Winter, Spring, Summer, and Autumn, are formed by an equal number of consecutive weeks. Notably, however, the Winter season is split into two periods: one in the beginning of the year and one in the end.

4.1.2 Normalisation

The second step of the seasonal approach is normalisation (Hassanzadeh Moghimi et al., 2023), which allows the time series to be scaled in such a way that their comparison is meaningful. Thus, the wind and solar availabilities, which represent percentages, come to have a similar weight to the demand data, which originally has a significantly wider range of values at approximately [5200, 13200].

Normalisation can be performed on the time series in several different ways (Teichgraeber and Brandt, 2019). In this case, the time series are normalised with min-max normalisations. In min-max normalisation, the normalised time series $\tilde{x}$ is obtained by subtracting the minimum value of the series $\min x$ from each time step and then dividing by the range of the series $\max x - \min x$:

$$\tilde{x} = \frac{x - \min x}{\max x - \min x} \tag{1}$$

The normalised values of the time series are, hence, set to the range $[0, 1]$.

4.1.3 Choosing Representative Weeks

As the third step of the approach, the test statistic is calculated for each week of the year. The value of the test statistic is then compared within each season, and the week with the smallest value is chosen as the representative week for that season. To calculate the seasonal-weekly error for a week, the values for the weekly mean μ^x_{week} and standard error s^x_{week} for each of the variables $x \in \{W, Q, S\}$ are computed. The seasonal-weekly squared error SW^2_x between the weekly means and the standard errors and those of the season as a whole $(s^x_{season}, \mu^x_{season})$ is calculated for each variable x:

$$SW^2_x = (\mu^x_{season} - \mu^x_{week})^2 + (s^x_{season} - s^x_{week})^2 \tag{2}$$

The final metric is then calculated by taking the square root of the weighted sum of the variables:

$$Score(week) = \sqrt{\sum_{x \in W,Q,S} w_x SW^2_x} \tag{3}$$

The week with the lowest dissimilarity score is chosen as the representative week for the whole season. The weights, w_x, can be adjusted to alter the influence of each variable, thereby allowing the method to be tuned.

4.2 Partitional Approach

The partitional approach is named after the fact that it applies a partitional clustering algorithm to group the weeks into clusters. The algorithm, Partitioning Around Medoids (PAM), compares each pair of weeks to each other at the time-step level, unlike the seasonal approach, which uses descriptive statistics as a tool for comparison. The approach is implemented with the R programming language (R Core Team, 2021).

The stages of the approach are described in the following sections. Section 4.2.1 discusses normalisation, Section 4.2.2 presents how the differences between weeks are compared, Section 4.2.3 describes the clustering algorithm in detail, and Section 4.2.4 shows how the order in which the clusters are arranged for the optimisation problem is chosen.

4.2.1 Normalisation

Similarly to the seasonal approach, normalisation is necessary in the partitional approach (Teichgraeber and Brandt, 2019). The method chosen is z-score normalisation on the full time series. Each of the three full time series is normalised by the formula:

$$\tilde{x} = \frac{x - \mu_x}{\sigma_x} \tag{4}$$

where $\tilde{x}$ is the normalised time series, and μ_x and σ are the mean and standard deviation of the original time series, respectively.

Min-max normalisation was also considered, as Hoffmann et al. (2020) found it to be the most common normalisation method. However, z-score normalisation was found to perform better, likely due to it being more robust to outliers as it also normalises in terms of standard deviation. After normalisation, weights can be applied to the chosen variables, similarly to the seasonal model. The weighted variable can be multiplied by the weight $w_x > 1$, which increases its influence in relation to the other variables.

4.2.2 Measuring Dissimilarity

Following normalisation, the next step of the partitional approach is to compute the difference between each pair of weeks. Numerous methods for computing distances between time series have been developed (Aghabozorgi et al., 2015). This thesis compares six different variations of measuring the dissimilarity of two weeks.

Computations are implemented with the R-library "proxy" (Meyer et al., 2015), which provides a variety of different proximity measures. Using the library, the distances can be arranged into a 52×52 matrix, known as a dissimilarity matrix, which streamlines the next stage of approach.

The first two methods for measuring dissimilarity, are the Euclidean and Manhattan distances. With these distance measures, the hours at the same relative position within the two compared weeks are compared to each other for each possible combination.

The Euclidean distance, also known as the L2 distance, is computed as the root of the squared distances between the values at the same position:

$$d_{l_2}(x, y) = \sqrt{\sum_{t \in T} (x_t - y_t)^2} \tag{5}$$

where x and y are the two time-series being compared and $t \in T$ is the set of time steps. The Manhattan distance (or L1 norm) is simply the sum of the absolute differences between each pair of time steps:

$$d_{l_1}(x, y) = \sum_{h \in w} |x_h - y_h| \tag{6}$$

The Euclidean and Manhattan distances are similar, but the Euclidean distance is more robust to outliers. In this thesis, both are used in order to analyse whether their performance is chosen clustering algorithm. Both the Euclidean and Manhattan distances are easily modified to work with multivariate time series by taking the sum of the distances computed for each individual series. This makes them both of them applicable to the dataset for this thesis, unlike some more complex alternatives.

When comparing two weeks, both of the measures compare the time step t of one time series to the same time step of the other, as seen in the leftmost panel of 6. This method, however, may have difficulties in adequately comparing time series that are similar but misaligned or without similar periodicity. In these cases, if the general profile of a pair of time series has similar peaks and valleys that do not occur at the same time steps, the similarity is not detected in the distance computed. In the case of the dataset used for this thesis, the wind availability time series has no clear periodicity, which could lead to the aforementioned problem. Hence, in addition to the two measures already introduced, the performance of alternative alignments created by applying dynamic time warping (DTW) to the distance measurement are also considered.

The DTW algorithm finds the increasing sequence a_i, b_i, mapping each time step of each time series to a time step of the other series, in the order that minimises the distance d between the connected time steps:

$$d_{DTW}(x, y) = \min_{a_i, b_i} \sum_{i \in w} d(x_{a_i}, x_{b_i}) \tag{7}$$

The distance d is most commonly the Euclidean distance (Giorgino, 2009). In this thesis, the Manhattan distance is also considered. DTW was originally designed for signal processing but has been applied to clustering for representative periods by Teichgraeber and Brandt (2019). The algorithm is implemented using the R-library "dtw" by Giorgino (2009).

The simplest variant of DTW does not limit or penalise how far each time step is warped (Giorgino, 2009). As illustrated in the middle plot in Figure 6 this can lead to a large share of all time steps of one series being warped to a single time step in the other. To study the effects of this phenomenon in the context of the energy data used, DTW with a boundary constraint is also considered in this thesis.

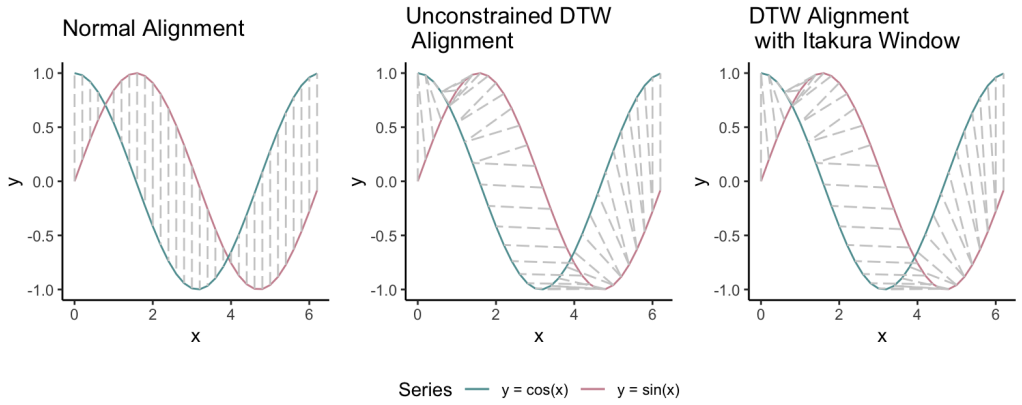


Figure 6: Normal, unconstrained DTW, and DTW with the Itakura window alignments for example time series.

Several methods for constraining DTW have been developed (Giorgino, 2009). Generally, the two main types of constraints are window functions and penalty terms. Window constraints set a limit on how far time can be warped, whereas penalty terms add a cost to the DTW objective function that increases the further the aligned time steps are from each other.

The Itakura window constraint was chosen for this thesis because it does not require determining additional parameters. The Itakura parallelogram constrains the DTW algorithm so that the values in each end of the time series can warp fewer time steps than the values in the middle. Figure 7 shows the effect of the window constraint when applied on the demand of two weeks. It can be seen that with DTW, the majority of time steps of week 1 are mapped into only two time steps of week 22. Including the window constraint decreases this effect.

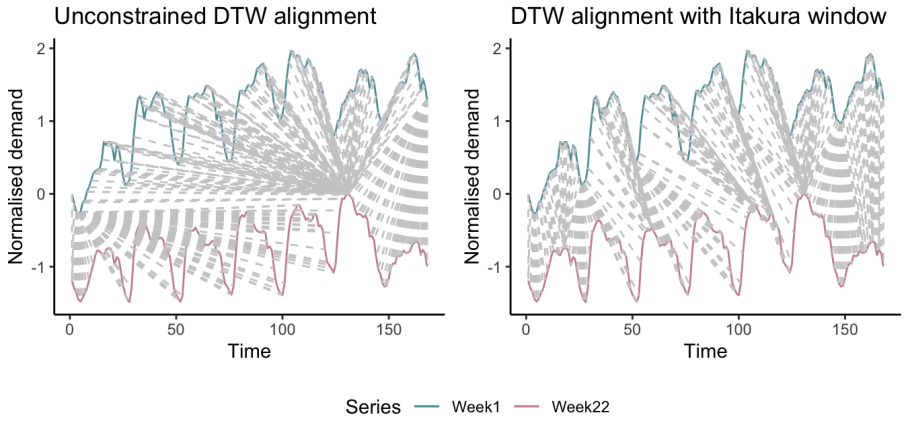


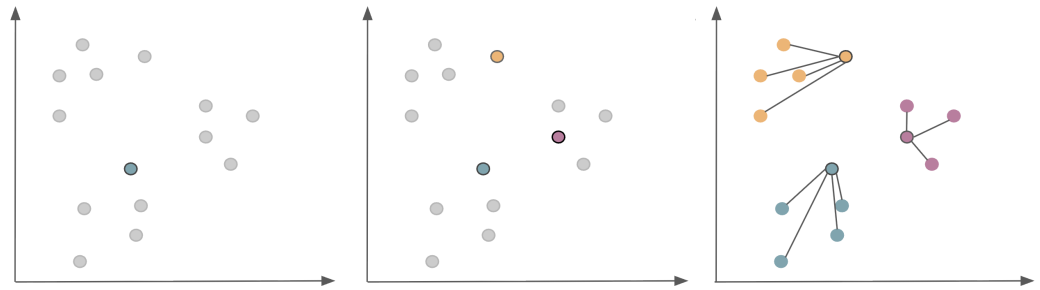
Figure 7: Unconstrained and Itakura window DTW alignments of two demand profiles

4.2.3 Clustering Algorithm

Once the distances between each pair of weeks were computed, the PAM clustering algorithm is applied to the data. This is implemented with the R-package "cluster" (Maechler et al., 2023). PAM is a partitional clustering algorithm, meaning that its objective is to partition n objects into k clusters. As a medoid-based algorithm, PAM chooses the partition so that it minimises the total distance of each object to the medoid object of its cluster. The medoid, also referred to as the representative object of a cluster by Kaufman and Rousseeuw (2005) is, thus, the centremost object of each cluster.

In the case of this analysis, the objects being considered are the 52 weeks being compared. The distance between objects is represented by the values in the dissimilarity matrix computed in the previous step. The advantage of using a medoid-based algorithm is that the medoids of the final clustering can be selected as the representative weeks. As described by Kaufman and Rousseeuw (2005), the algorithm has two phases referred to as the build stage and swap stage.

The build stage initialises the algorithm by successively selecting the k initial medoids. The first selected medoid is the one with the smallest total distance to all other objects (Figure 8a). The remaining $k-1$ are then chosen one-by-one as the objects that decrease the objective-function value the most by choosing them (Figure 8b). The remaining $n-k$ objects are then initially assigned to the medoids nearest to them (Figure 8c).



(a) The first initial medoid is selected

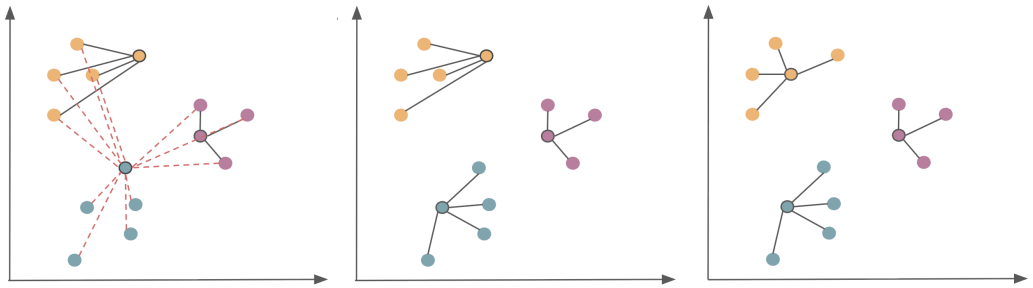
(b) The other initial medoids are selected one by one

(c) Each remaining object is assigned to its nearest initial medoid

Figure 8: A visualisation of the build stage of the PAM algorithm

After the initial medoids are chosen, the swap stage follows. All alternatives to each current medoid (i.e., all objects not assigned as medoids) are considered (Figure 9a). A medoid is swapped to an alternative if the objective function value can be further decreased by the swap (Figure 9b). The swap stage continues until the algorithm has found a local minimum, where no swap of any individual medoid will further decrease the objective function value (Figure 9c) (Kaufman and Rousseeuw, 2005).

PAM is a quite computationally heavy algorithm (Kaufman and Rousseeuw, 2005) However, due to the limited number of objects being clustered (i.e. 52 weeks) in the



(a) The alternatives to the blue medoid are considered

(b) The blue medoid is swapped, as the objective can be improved

(c) Swaps are repeated until the objective cannot be further decreased

Figure 9: A visualisation of the swap stage of the PAM algorithm

dataset and the ability to precompute distances between each pair of objects the computation time does not pose an issue. For larger datasets, a less computationally heavy algorithm may need to be considered.

Due to the greedy selection of the initial medoids in the build phase, PAM has deterministic end results unless there are equal distances between objects in some stage of the process, which would lead to randomisation (Kaufman and Rousseeuw, 2005). This feature is useful in the case of this thesis, as later stages of analysis are performed in different software and running multiple trials, hence, may be difficult.

A medoid-based algorithm is also more intuitively suitable for comparison to the seasonal approach than the more common centroid-based algorithms. Medoids are real objects representing real data, unlike centroids, which are computed averages. The more common centroid-based algorithms have also been found to be more prone to heavier suppression of extreme values by Teichgraeber and Brandt (2019) and Hoffmann et al. (2022).

4.2.4 Defining the Order of Representative Weeks

As stated earlier, the medoid of each cluster is chosen as a representative week for the members of said cluster. The size of (i.e., number of the weeks in) the cluster serves as a weight for that week. However, as the system modelled in this thesis has the element of hydropower storage, an optimisation model describing it will need to include intertemporal constraints. As a result, in addition to the data of representative weeks, the order in which they are presented in the aggregation must also be considered. PAM defines the clusters without taking into account how the weeks in each cluster are situated temporally in the year. As a result, there is no innate ordering defined by the algorithm, and one must be determined exogenously after clustering.

The method chosen to determine the order of the clusters is ranking them in ascending order in terms of the median index of the weeks in each cluster. By ordering the clusters in terms of the median, the representative weeks are usually run temporally in the position that most of the members of that cluster are situated

in relation to the other clusters.

Other alternatives considered for ranking the clusters were the numbers of the medoid weeks and the average of the week numbers of the cluster members. However, when using medoids as the criterion, it is possible that clustering results occur where the medoid is far from other members of the cluster. In this case, the representative week would be placed differently than the weeks it is representing would indicate. The average, on the other hand, would not work in a situation where the weeks of the cluster are at the beginning and at the end of the year. In this case, the rank of the representative week may end up in the middle of the year, even though it included only winter weeks.

4.3 Assessment of Representativity

As discussed in Section 2.2, various metrics have been used to measure the representativity of representative periods, which measure different features of the time series (Hoffmann et al., 2020). Therefore, multiple different types of metrics are used in this thesis to analyse the efficacy of the different approaches. Two types of metrics are used to determine how well the compared approaches perform in representing the original data: descriptive statistics and optimisation results of a benchmark problem.

The first type of metric is selected descriptive statistics for the aggregated demand and wind availability. The selected statistics are the mean and the standard deviation of the time series, which were chosen to consider the error in both the average and the variability of the aggregated data. As the share of solar power in the total generation is low, it is not used as a metric for performance. Table 1 shows the descriptive statistics and their reference values for the original time series. The relative error in these statistics is computed by comparing the value of the statistics for the original time series to the ones of the aggregated series.

Table 1: Descriptive statistics used to evaluate the aggregated time series and the reference values for the original series

Statistic	Definition	Reference value
μ_Q	Mean of demand	9010 (MWh)
σ_Q	Standard deviation of demand	1560
μ_W	Mean of wind availability	0.277
σ_W	Standard deviation of wind availability	0.215

Reichenberg et al. (2018) argue that similar statistical values for time series does not necessarily equal similar results in optimisation. Therefore, a benchmark optimisation problem is also implemented. With the benchmark problem, the results of the full time resolution (FTR) problem, which utilises the full original dataset, can be compared with a reduced time resolution problem using the aggregated data.

The benchmark problem formulated for this purpose is a linear program (LP), modelling capacity expansion for the Finnish power system, which is a simplified version of Tassi (2021) and Hassanzadeh Moghimi et al. (2023). The objective of the

benchmark problem is to minimise total cost from generation, maintenance, demand response and generation-capacity expansion for the Finnish power system. The full optimisation problem is described in detail in Appendix A.

Figure 10 shows the generation distribution by fuel type in the FTR model next to the real values also shown in Figure 10. The main difference is the nuclear generation corresponding to 7% more of the total in the FTR than in the real data. However, this discrepancy is affected by the fact that Olkiluoto 3 is considered to be functional for the whole year in the FTR model, although general production only began in May (Teollisuuden Voima Oy, 2023). The overgeneration of nuclear is reflected in undergeneration of both other thermal power types. There is also slight overgeneration by solar and wind power, which can mostly be explained by the use of end-of-year generation capacities.

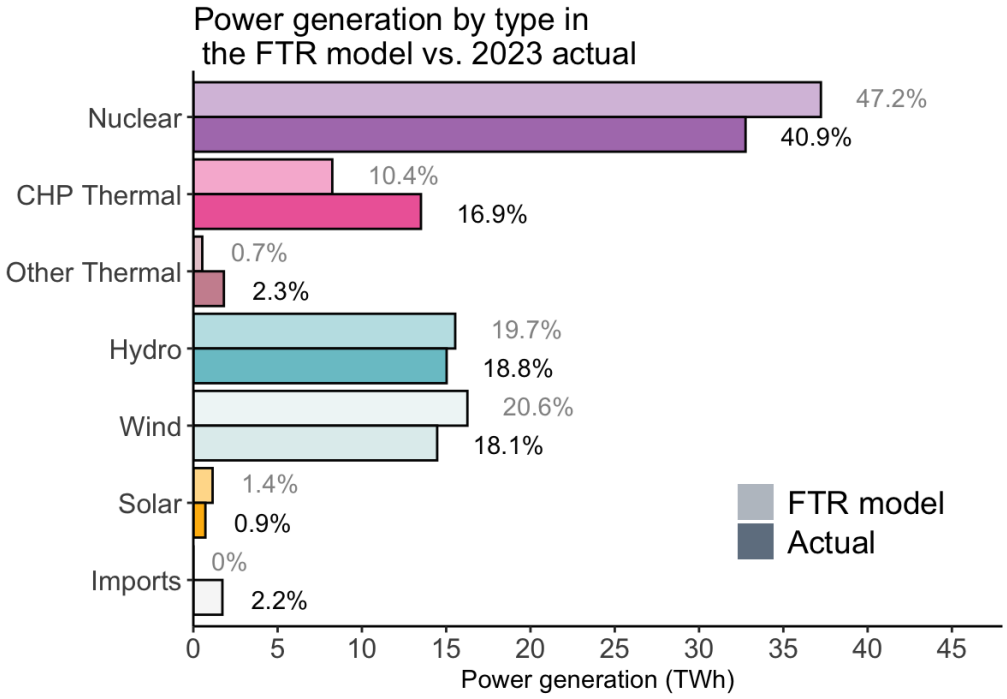


Figure 10: A comparison of power generation in the model and real data

To analyse the performance of the aggregated data in action, a reduced time dimension problem is also constructed. The model is only solved for the four representative weeks and applies the parameters of the sizes of the seasons or clusters as weights for each week. Hence, the reduced problem solves in 672 time steps rather than the 8760 of the FTR model. The changes made to the FTR model are described in Appendix B.

Table 2 presents the metrics used to compare results of the reduced time dimension problem with the FTR problem. The first metric is the error in the objective-function value, which indicates the overall performance of the aggregation in maintaining the features of the original data.

The findings of both Reichenberg et al. (2018) and Hoffmann et al. (2022),

however, show that analysing the objective-function value alone, may be misleading and not necessarily indicate performance of VRE aggregation. Consequently, the differences in wind and hydro generation are additionally considered. These two variables largely account for the amount of power generated with zero fuel costs. A large negative error in these two variables would lead to an increase in the value of the objective function. Although it is not considered in the aggregation due to the low-quality data, hydropower generation is also considered, as the weights and order of the representative weeks can affect how hydropower functions in the model.

Table 2: Values from the benchmark problem used to evaluate the aggregated time series and the reference values from the FTR model

Error	Definition	FTR value
Obj%	Relative error in the objective-function value	2716 (M€)
W%	Relative error in wind-power generation	16.11 (TWh)
H%	Relative error in hydropower generation	15.54 (TWh)

5 Results

5.1 Seasonal Approach

Table 3 presents the relative errors of the performance metrics when using the seasonal approach, where both wind availability and demand have equal weights of 1 (WQ-11). The same representative weeks were also selected when solar power was added as a variable with an equal weight (WQS-11). This shows that the seasonal approach is quite robust in relation to the solar variable.

Table 3: Results of the seasonal approach when using equal weights for each variable

Seasonal	μ_Q	σ_Q	μ_W	σ_W	Obj%	W%	H%
WQ-11 WQS-111	-0.05 %	-30.01 %	0.72 %	-2.33 %	-10.53 %	1.66 %	1.25 %

The results show low errors in μ_Q and μ_W as well as $Q\%$ and $W\%$. However, the objective-function value is 10.53% lower than the FTR. Additionally, the largest error is in σ_Q, which indicates that although the average demand is close to the original value, there is a lack of extreme values in the aggregated data. This is likely also the cause of the low objective function-value.

To address the problems with demand, the objective function of the seasonal approach was adjusted to have a higher weight on standard error (WQ-11, 2xSE) and demand (WQ-12 and WQ-1[1,5]). The results of these two alternate parameter choices are shown in Table 4.

From the results, it can be seen that increasing the weight of the standard error worsened the error in the objective-function value, and did not manage to decrease

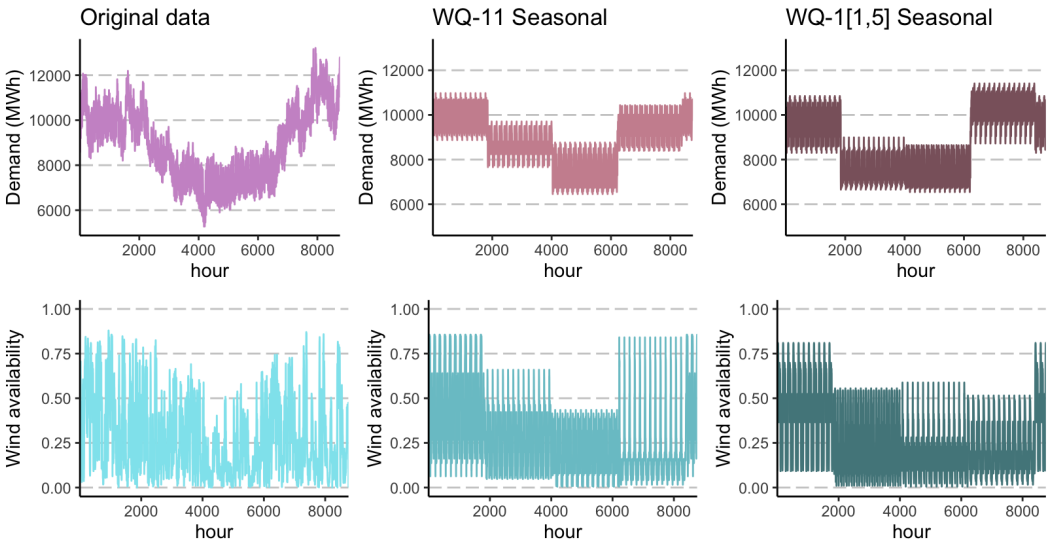


Figure 11: The original data and two alternative seasonal aggregations plotted next to each other

Table 4: Resulting aggregated time series of the seasonal approach after adjusting the method parameters

Seasonal	μ_Q	σ_Q	μ_W	σ_W	Obj%	W%	H%
WQ-11. 2xSE	-6.70 %	-30.26 %	3.61 %	-4.19 %	-23.48 %	4.78 %	0.77 %
WQ -1[1.5] WQ-12	-1.20 %	-12.79 %	0.36 %	-4.65 %	-6.32 %	1.53 %	-1.17 %

the error in the σ_Q. This result is unexpected, however, it is possible that the data are structured so that as the σ_Q improves the σ_W decreases so significantly that it counteracts the improvement in the objective function. Increasing the weight of demand, however, did manage to lower the error in the objective function, with the σ_Q error decreasing to -12.79%.

5.2 Partitional Approach

Table 5 shows the results of partitional method when using both wind availability and demand data with equal weights (WQ-11). Euclidean and Manhattan norms perform very similarly in these metrics when using the same type of alignment. From the table, it can be seen that the error in the objective function was the largest with the normal Euclidean and Manhattan norms, which do not apply DTW.

Table 5: Results of the partitional approach with wind availability and demand as equally weighted clustering variables (WQ-11).

Partitional WQ-11	μ_Q	σ_Q	μ_W	σ_W	Obj%	W%	H%
Manhattan	-5.08 %	-1.41 %	-8.30 %	-22.33 %	-7.64 %	-7.30 %	-1.39 %
Euclidean	-3.30 %	-10.24 %	-7.94 %	-27.91 %	-9.01 %	-6.81 %	0.64 %
DTW Manh.	-0.65 %	2.50 %	-6.14 %	-0.47 %	-2.70 %	-4.97 %	-2.14 %
DTW Eucl.	-0.63 %	4.93 %	-7.22 %	-1.40 %	-2.06 %	-6.18 %	-2.05 %
DTW Manh. + Itakura window	2.87 %	-14.92 %	-4.69 %	-2.79 %	-0.39 %	-3.67 %	-0.09 %
DTW Eucl. + Itakura window	3.04 %	-14.19 %	-6.50 %	-4.65 %	0.50 %	-5.66 %	-0.09 %

Additionally, it can observed that the lowest error in the objective function was reached with the dissimilarity measures applying DTW with the Itakura window constraint. These variants, however, have large negative errors in σ_Q, and μ_Q above the FTR. This indicates that the improved objective-function value is due to the above-average demand inflating the value, rather than a particularly good aggregation result.

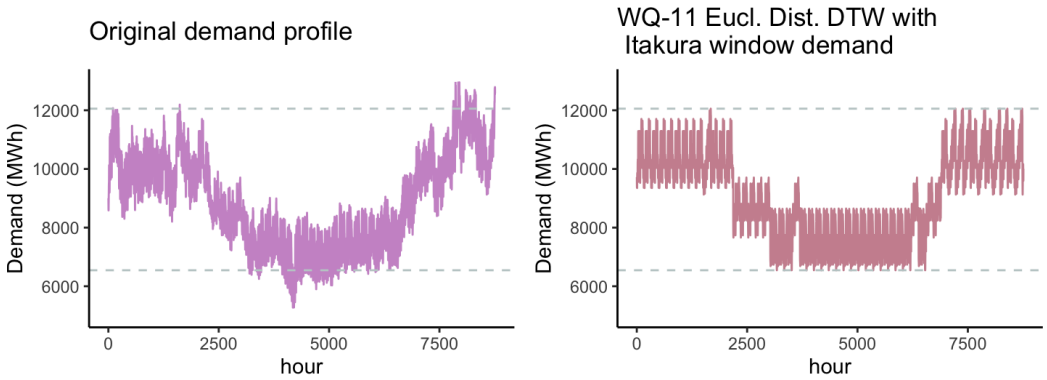


Figure 12: A comparison of the aggregated wind profile to the original shows, that the original TS has values both below the minimum and above the maximum of the aggregated TS

While all variants seem to struggle with lower than intended wind generation, variants using DTW have better performance in terms of σ_W. This can also be observed visually from the time series formed by the aggregation presented in Figure 13. Especially in the summer months, a very flat series is created when only the L2 standard is used (σ_W = -27.9%). DTW L1 (σ_W=-0.5%), on the other hand, also contains more spikes of higher production.

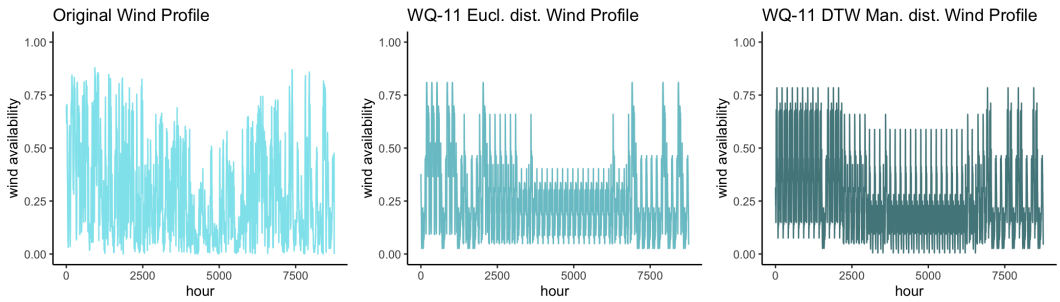


Figure 13: The original and aggregated wind profiles created by different variants of the partitional method

Table 6: Results of partitional method after including solar power as a variable

Partitional WQS-111	μ_Q	σ_Q	μ_W	σ_W	Obj%	W%	H%
Manhattan	-5.2 %	-2.8 %	-5.1 %	-20.0 %	-8.8 %	-4.2 %	-1.1 %
Euclidean	-5.6 %	-3.2 %	-5.8 %	-21.4 %	-9.2 %	-4.9 %	-0.9 %
DTW Manh.	1.7 %	-13.2 %	-4.7 %	-0.9 %	-2.4 %	-3.7 %	-0.4 %
DTW Eucl.	-0.9 %	1.8 %	-7.9 %	-5.6 %	-2.5 %	-6.8 %	-3.8 %
DTW Manh. + Itakura Window	1.1 %	-12.1 %	-10.8 %	-5.6 %	-1.0 %	-9.9 %	-2.4 %
DTW Eucl. + Itakura Window	-0.1 %	-0.8 %	-11.2 %	-7.4 %	-0.8 %	-10.2 %	-3.7 %

Table 6 presents the results when solar availability is included in the dissimilarity measurement along with wind availability and demand. It can be noted that the error of the target function remains similar for all variants as in the results without accounting for solar availability.

6 Conclusions

Taking into account power systems' growing reliance on VREs has increased complexity of power system models. This thesis compared and analysed variants of the representative-weeks method, developed to address this issue, thus addressing a gap in research.

Two alternative approaches to finding four representative weeks were analysed in this thesis. The seasonal approach exogenously defines seasons, from which the weeks were chosen and the partitional approach applied a k-medoids clustering algorithm to form clusters, whose medoids were selected as representative weeks. Both descriptive statistics and a benchmark optimisation problem were used to analyse the quality of the selected weeks.

Based on the results of the analysis, neither approach or variant of one was a clear winner. Both approaches have their strengths and weaknesses. The seasonal approach exceeded in more consistently retaining the mean behaviour of the time series but struggled with the objective-function value. The partitional approach, on the other hand, achieved closer objective-function values to that of the FTR model. However, this was not always due to better aggregation. It was also found that the introduction of dynamic time warping (DTW) to dissimilarity measurement improved wind aggregation in the partitional approach.

A common outcome of both methods was the underrepresentation of extreme values. This was reflected in the objective-function values in the lower time dimension benchmark problem, which were generally below the one of the FTR model.

The effect of introducing solar availability as a variable did not have a major impact in performance in either approach. The seasonal method was quite robust in terms of the solar output as an input to the clustering and including it did not affect the representative weeks chosen. The performance of the partitional method decreased in most metrics after adding solar, which is likely due to low solar generation capacity in the system.

This study focused solely on identifying four representative weeks. Further analysis could focus on analysing how the number of lengths of representative periods would affect the performance of the methods. For example, a larger number of representative days could be analysed to see the difference in behaviour. Additionally, analysing a longer multi-year time series could also be of interest.

References

Saeed Aghabozorgi, Ali Seyed Shirkhorshidi, and Teh Ying Wah. Time-series clustering–a decade review. *Information systems*, 53:16–38, 2015.

Fingrid, 2024a. Online dataset downloaded from the Fingrid open data portal, Cited: 24.6.2024, Accessible at: https://data.fingrid.fi/en/datasets/181.

Fingrid, 2024b. Online dataset downloaded from the Fingrid open data portal, Cited: 24.6.2024, Accessible at: https://data.fingrid.fi/en/datasets/181.

Fingrid. Datasets: Total production capacity used in the solar power forecast, 2024c. Online dataset downloaded from the Fingrid open data portal, Cited: 24.6.2024, Accessible at: https://data.fingrid.fi/en/datasets/267.

Fingrid. Datasets: Total production capacity used in the wind power forecast, 2024d. Online dataset downloaded from the Fingrid open data portal, Cited: 24.6.2024, Accessible at: https://data.fingrid.fi/en/datasets/268.

Toni Giorgino. Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, 31(7):1–24, 2009. doi: 10.18637/jss.v031.i07.

Farzad Hassanzadeh Moghimi, Hanna Ek Fälth, Lina Reichenberg, and Afzal S Siddiqui. Climate policy and strategic operations in a hydro-thermal power system. *The Energy Journal*, 44(5):67–94, 2023.

Maximilian Hoffmann, Leander Kotzur, Detlef Stolten, and Martin Robinius. A review on time series aggregation methods for energy system models. *Energies*, 13 (3):641, 2020.

Maximilian Hoffmann, Leander Kotzur, and Detlef Stolten. The pareto-optimal temporal aggregation of energy system models. *Applied Energy*, 315:119029, 2022.

Leonard Kaufman and Peter J Rousseeuw. *Finding groups in data: an Introduction to Cluster Analysis*, pages 37–50, 68–70, 102–104. John Wiley & Sons, 2005.

Martin Maechler, Peter Rousseeuw, Anja Struyf, Mia Hubert, and Kurt Hornik. *cluster: Cluster Analysis Basics and Extensions*, 2023. URL https://CRAN.R-project.org/package=cluster. R package version 2.1.6.

David Meyer, Christian Buchta, and Maintainer David Meyer. Package 'proxy'. 2015.

Ministry of Economic affairs and Employment. Carbon neutral finland 2035 – national climate and energy strategy, 2022. Accessible at: http://urn.fi/URN:ISBN:978-952-327-843-1.

David M Newbery. High renewable electricity penetration: Marginal curtailment and market failure under "subsidy-free" entry. *Energy Economics*, 126:107011, 2023.

Nord Pool. Power system data: Consumption, 2024. Online dataset accessible at the Nord Pool data portal https://data.nordpoolgroup.com/power-system/production.

Nord Pool. The power market: Bidding areas, 2025. Cited: 7.9.2025, Accessible at: https://www.nordpoolgroup.com/en/the-power-market/Bidding-areas/.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL https://www.R-project.org/.

Lina Reichenberg, Afzal S Siddiqui, and Sonja Wogrin. Policy implications of downscaling the time dimension in power system planning models to represent variability in renewable output. *Energy*, 159:870–877, 2018.

Kirstine Roberts, Dina Khastieva, Afzal S Siddiqui, Eeva Vilkkumaa, and Ilona Ylikoski. Industrial electrification and market power in a hydro-thermal power system. 2025. Working Paper.

Statistics Finland. Energian hankinta ja kulutus: 12sv sähkön hankita ja kokonaiskulutus 1960-2024*, 2025. Last updated: 5.6.2025, Cited: 23.8.2025, Accessible at: https://stat.fi/en/statistics/ehk,.

Tatiana Tassi. Multi-criteria decision analysis for portfolio planning in power generation assessing economic and environmental tradeoffs in the finnish power sector, 2021. MSc Thesis.

Holger Teichgraeber and Adam R Brandt. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. *Applied energy*, 239:1283–1293, 2019.

Holger Teichgraeber and Adam R Brandt. Time-series aggregation for the optimization of energy systems: Goals, challenges, approaches, and opportunities. *Renewable and Sustainable Energy Reviews*, 157:111984, 2022.

TVO Teollisuuden Voima Oy. Tiedotteet: Olkiluoto 3 epr:n säännöllinen sähköntuotanto on alkanut, 2023. Last updated: 3.6.2023, Cited: 16.4.2023, Accessible at: https://www.tvo.fi/ajankohtaista/tiedotteetporssitiedotteet/2023/olkiluoto3eprnsaannollinensahkontuotantoonalkanut.html.

Vilma Virasjoki, Afzal S Siddiqui, Behnam Zakeri, and Ahti Salo. Market power with combined heat and power production in the nordic energy system. *IEEE Transactions on Power Systems*, 33(5):5263–5275, 2018.

A Full Time Resolution Benchmark Model

The full time resolution benchmark model indices are presented in table A1, variables in table A2 and parameters in table A3. The parameter values for capacities, costs, hyrdopower inflow, ramp-up and down rates and emissions rates are the same as in Roberts et al. (2025).

Table A1: Indices and corresponding sets of the benchmark model

Indices	Description	Set
$t \in T$	Time-steps, hours of the year	$T = \{\mathbb{Z} \in [1, 8760]\}$
$e \in E$	VRE unit types	$E = \{wind, offshore, solar\}$
$u \in U$	Thermal unit types	$U = \{coal, oil, nuclear, chp_{coal}, chp_{waste}, chp_{gas}, chp_{peat}, chp_{biomass}\}$
$w \in W$	Hydro unit types	$W = \{ror, res\}$

Table A2: Variables of the benchmark model

Variable	Description	Unit
$g_t^{\{e,u,w\}}$	Energy generated by unit $\{e, u, w\}$ at time t	MWh
$x^{\{e,u,w\}}$	Investment in new capacity of type $\{e, u, w\}$	MW
$y^{\{e,u,w\}}$	Available production capacity of type $\{e, u, w\}$	MW
r_t^{sto}	Energy stored in reservoir at time t	MWh
z_t	Demand response at time t	MWh

The objective function of the problem is to minimise the sum of annual costs from investment, operations, generation, and demand response:

$$\min \sum_{e,u,w} \left[C^{inv}_{\{e,u,w\}} x^{\{e,u,w\}} + C^{ava}_{\{e,u,w\}} y^{\{e,u,w\}} \right] + \sum_{t \in T} \sum_{u \in U} \left[C^{gen}_{u} + C^{CO_2} P_u \right] g^u_t + \sum_{t \in T} C^{dem} \, z_t$$

, where $\sum_{e,u,w} = (\sum_{e \in E} + \sum_{u \in U} + \sum_{w \in W})$.

The problem is subject to the following constraints:

$$Q_t = \sum_{u \in U} g^u_t + \sum_{e \in E} g^e_t + \sum_{w \in W} g^w_t + z_t, \forall t \tag{A1}$$

(A1) guarantees that the total generation from all production types and possible demand response must meet the demand at every time step.

$$z_t \leq A^{dem}_t Q_t, \forall t \tag{A2}$$

(A2) ensures that the demand response at any given time-step cannot be higher than the specified percentage of demand A_t.

Table A3: Parameters of the benchmark model

Parameter	Description	Unit
A_t^e	Availability factor of VRE of type e at time t	-
A_t^{ror}	Availability factor of inflow to ROR hydro plants	-
A_t^{dem}	Maximum share of demand that can be curtailed at time t	-
$C_{\{e,u,w\}}^{ava}$	Operations and maintenance ($O\&M$) cost for type $\{e,u,w\}$	€/MW
C^{CO_2}	Cost of CO_2 emissions from generation	€/tCO_2
C^{dem}	Cost of demand response	€/MWh
C_u^{gen}	Cost of generation for thermal unit u	€/MWh
$C_{\{e,u,w\}}^{inv}$	Cost of investment in new capacity of type e,u,w	€/MW
I_t	Inflow of energy to hydro reservoirs during time-step t	MWh
$K^{\{e,u,w\}}$	Initial production capacity of type $\{e,u,w\}$	MW
$\underline{R}, \overline{R}$	Floor and ceiling levels of hydro reservoir	MWh
R^{start}, R^{end}	Reservoir levels at the start and end of the year	MWh
R_u^{up}, R_u^{down}	Ramp-up and ramp-down rates for thermal units of type u	-
P_u	Emissions rate of thermla generation of type u	tCO_2/MWh
T_{len}	Length of a single time-step	1h
Q_t	Demand at time t	MWh

Table A4: Initial capacities $K^{\{e,u,w\}}$, investment costs C^{inv}, O&M costs C^{ava} and generation costs C^{gen} of each generation type

Unit type	Set	K	C_{inv}	C_{ava}	C_{gen}
wind	e	6649,7	1340100	29754	0
offshore	e	44,3	3320100	138600	0
solar	e	1018	1351800	18207	0
coal	u	565	3692700	55440	40.67
oil	u	1290	3692700	55440	67
nuclear	u	4394	7074900	140580	21
chp$_{coal}$	u	438	3692700	55440	40.67
chp$_{waste}$	u	1652,6	11367900	235062	22
chp$_{gas}$	u	1631,6	752400	6183	61.5
chp$_{peat}$	u	1208,6	11367900	235062	15.5
chp$_{biomass}$	u	1619,5	11367900	235062	30.75
ror	w	930	6365700	30186	0
res	w	2275	6365700	30186	0

$$y^{\{e,u,w\}} \leq K^{\{e,u,w\}} + x^{\{e,u,w\}}, \forall \{e,u,w\} \tag{A3}$$

(A3) ensures that thet the available capacity of type e,u,w is bounded by the sum of initial capacity $K^{\{e,u,w\}}$ and the investment made into the generation type $x^{\{e,u,w\}}$.

$$T^{len} y^u - g_{i,t}^u \geq 0, \forall t,u \tag{A4}$$

(A4) asserts that the amount of energy generated by thermal of type u at time t, can be at maximum the available capacity:

$$g_t^u - g_{t-1}^u \leq T^{len} R_u^{up} y^u, \forall t,u \tag{A5}$$

$$T^{len} R_u^{down} y^u \leq g_t^u - g_{t-1}^u, \forall t,u \tag{A6}$$

(A5) and (A6) enforce that the rate of change in thermal generation between two time-steps cannot be faster than the defined up-ramp and down-ramp rates R_{up}, R_{down}.

$$T^{len} A_t^e y^e - g_t^e \geq 0, \forall t, e \tag{A7}$$

(A7) ensures that the VRE generation at every time step does not exceed available capacity.

$$r_t^{sto} = r_{t-1}^{sto} + I_t - g_t^{res}, \forall t \tag{A8}$$

(A8) asserts that the energy stored in the reservoir at time step t is equal to the reservoir level at the previous time step $t-1$ plus inflow I_t minus generation g_t^{res}

$$T^{len} y^{res} - g_t^{res} \geq 0, \forall t \tag{A9}$$

(A9) ensures that the reservoir can be depleted at maximum at the rate determined by the initial hydropower capacity.

$$T^{len} A_t^{ror} y^{ror} - g_t^{ror} \geq 0, \forall t \tag{A10}$$

(A10) ensures that the run-of-river hydro plants is constrained by the available inflow at that hour.

$$\underline{R} \leq r_t^{sto}, \forall t \tag{A11}$$

$$r_t^{sto} \leq \overline{R}, \forall t \tag{A12}$$

(A11) and (A12) guarantee that the reservoir level cannot go below the determined minimum and maximum levels of capacity, $\underline{R}$ and $\overline{R}$.

$$r_{t=1}^{res} = R^{start} \tag{A13}$$

$$r_{t=8760}^{res} = R^{end} \tag{A14}$$

(A13) and (A14) set the reservoir level at the start and the end of the year to be the same as the actual values in 2023, to ensure that not all of the reservoirs are depleted throughout the year.

B Reduced Time Dimension Benchmark Model

Reduced time dimension = only run problem for the four representative weeks and then have objective be the weighted sum of their contributions to the costs.

To analyse the aggregated data in relation to the benchmark, the model fomulation is adjusted so that it solves in fewer time steps. The representative weeks method lowers the number of time steps from 8760 to s $4 \times 168 = 672$ time steps. The program runtime is also reduced from 44 seconds for the FTR to less than one second, when solving the problem instance in GAMS using the CPLEX solver).

The reduced benchmark version problem replaces the time step index $t \in T = \{\mathbb{Z} \in [1, 8760]$, with the new sets $rh \in RH = [1, 168]$, representing representative hours i.e. each hour of a representative week and $c \in C = \{1, 2, 3, 4\}$, representing the cluster (partitional approach) or season (seasonal approach) of the week.

The parameters and variables withouta reference to the temporal dimension (t) remain the same as in the FTR. Parameters and variables with a time step reference are altered to reference representative hours: $t \rightarrow (rh, c)$.

Additionally, hydropower reservoir infow $I_{rh,c}$ and ROR availability $A^{ror}_{rh,c}$, are recalculated from the inflow data. The average weekly inflow during a season is divided equally to each hour of that season's representative week. Without this adjustment, differences in the performance of different representative weeks were heavily dependent on the inflow profiles of the representative weeks selected. This was done, because the low quality of the inflow data did not allow for it to be included in the selection process for the representative weeks with either approach analysed. Additionally a variable for the end-of-cluster reservoir level is added to the model $\hat{r}^{sto}_c$.

In the reduced time dimension problem, the objective function is the weighted sum of the costs associated with each week.

$$\min \sum_{e,u,w} \left[C^{inv}_{\{e,u,w\}} x^{\{e,u,w\}} + C^{ava}_{\{e,u,w\}} y^{\{e,u,w\}} \right]$$
$$+ \sum_{c \in C} N_c \left[\sum_{rh \in RH} \left[\sum_{u \in U} \left[C^{gen}_u + C^{CO_2} P_u \right] g^u_{rh,c} + C^{dem} z_{rh,c} \right] \right]$$

, where $\sum_{u,e,w} = (\sum_{u \in U} + \sum_{e \in E} + \sum_{w \in W})$ and N_c is the number of weeks in c.

Most constraints remain unchanged, however adjustments are made to the hydropower constraints.

Reservoir level at the end of season $\hat{r}^{sto}_c$ is the level at the end of the previous cluster $\hat{r}^{sto}_{c-1}$ plus the change in the reservoir level during the representative week of the season $\sum_{rh \in\ RH} I_{rh,c} - g_{rh,c} - s_{rh,c}$ multiplied by the number of weeks in the season N_c

$$\hat{r}^{sto}_c = \hat{r}^{sto}_{c-1} + N_c \sum_{rh \in RH} I_{rh,c} - g_{rh,c} - s_{rh,c}, \forall c > 1 \tag{B1}$$

(B2) sets the reservoir level at the end of the first cluster to refer to R^{start} instead

of the reservoir level of the previous cluster:

$$\hat{r}_c^{sto} = R^{\text{start}} + N_c \sum_{rh \in RH} I_{rh,c} - g_{rh,c} - s_{rh,c}, c = 1 \tag{B2}$$

The hydro generation during the first hour of each season is bound individually with constraint (B3)

$$r_{rh,c}^{sto} = \hat{r}_{c-1}^{sto} + I_{rh,c} - g_{rh,c}^{res} - s_{rh,c}, \forall\, (rh = 1, c > 1) \tag{B3}$$

The end-of-cluster reservoir level is set to also be bounded by the same minimum and maximum levels as the hourly ones by (B4) and (B5).

$$\hat{r}_c^{sto} \geq \underline{R}, \forall c \tag{B4}$$

$$\hat{r}_c^{sto} \leq \overline{R}, \forall c \tag{B5}$$