Aalto-yliopisto
Aalto-universitetet
Aalto University

Master's Programme in Mathematics and Operations Research

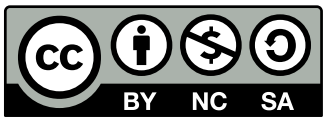
On Optimal Siting of Public Services in Helsinki

Einari Stenberg

Master's Thesis
2026

© 2026

This work is licensed under a Creative Commons "Attribution-NonCommercial-ShareAlike 4.0 International" license.



Aalto-yliopisto
Aalto-universitetet
Aalto University

Author Einari Stenberg

Title On Optimal Siting of Public Services in Helsinki

Degree programme Mathematics and Operations Research

Major Mathematics and Operations Research

Supervisor and advisor Prof. Pauliina Ilmonen

Collaborative partner City of Helsinki

Date 3.8.2026 **Number of pages** 74 **Language** English

Abstract

Population growth, aging, and cost pressures are forcing cities to evaluate how they provide and locate public services. Service networks are commonly built cumulatively with long time horizons, which can leave them suboptimal as a whole. To address this, we introduce a general framework for optimizing public service networks. We model service accessibility with a need-based utility function that combines travel-time accessibility, distance decay, and diminishing returns from multiple nearby services. Before that, we divide the population into age groups and weight each group based on its estimated need for the service. The utility function is used as the objective in an optimization model that selects service locations by maximizing total population utility, with an equity constraint that enforces a floor of accessibility to the whole area. The model is applied to public libraries in Helsinki, using public-transport travel time as the basis for the accessibility measure and a population prediction for 2036. We then calculate the optimal service network under several scenarios that include different numbers of libraries as well as adding or removing a single location from the network.

We find that 99.0% of adults reach their nearest library within 25 minutes. However, even with seemingly high accessibility, optimization shows clear inefficiency. The baseline utility of the 38 libraries can be achieved with 25 optimally placed locations and optimally reallocating these 38 libraries increases the total utility by 12.3%. In addition, optimal reallocation caused a slight shift from western toward eastern districts in library counts.

The results show that optimization methods can help decision makers organize public services in ways that increase the accessibility and efficiency of the service network without increasing total capacity. The inefficiency identified reflects the state of Helsinki's library network in 2036 if no changes are made, and shows that better outcomes do not depend on increased budgets but on more optimal service placement. The framework built in this thesis generalizes beyond libraries and gives cities a first step towards a tool for need-based evaluation of public service accessibility.

Keywords accessibility, distance decay, facility location problem, Helsinki, mixed-integer linear programming, public libraries, service network optimization

Aalto-yliopisto
Aalto-universitetet
Aalto University

Tekijä Einari Stenberg

Työn nimi Julkisten palvelujen optimaalinen sijoittelu Helsingissä

Koulutusohjelma Mathematics and Operations Research

Pääaine Matematiikan ja operaatiotutkimuksen maisteriohjelma

Työn valvoja ja ohjaaja Prof. Pauliina Ilmonen

Yhteistyötaho Helsingin Kaupunki

Päivämäärä 3.8.2026 **Sivumäärä** 74 **Kieli** englanti

Tiivistelmä

Väestönkasvu, ikääntyminen ja kustannuspaineet ohjaavat kaupunkisuunnittelijoita arvioimaan julkisten palveluiden sijoittelua. Palveluverkot ovat tyypillisesti suboptimaalisia, sillä verkot ovat muodostuneet kumulatiivisesti pitkien aikavälien kuluessa. Tässä työssä ehdotetaan yleistettävää mallia julkisten palveluverkkojen optimointiin sekä mallinnetaan palveluiden saavutettavuutta tarvepohjaisen hyötyfunktion avulla. Funktio huomioi matka-aikaan perustuvan saavutettavuuden, etäisyyden kasvaessa heikkenevän kysynnän sekä useiden lähistöllä sijaitsevien palveluiden tuottaman vähenevän rajahyödyn. Tätä ennen väestö jaetaan ikäryhmiin, joita painotetaan arvioidun palvelutarpeen perusteella.

Hyötyfunktiota käytetään optimointimallin kohdefunktiona. Malli valitsee palveluiden sijainnit maksimoimalla väestön kokonaishyödyn ja sisältää tasa-arvorajoitteen, joka asettaa alueellisen alarajan saavutettavuudelle. Mallin avulla arvioidaan Helsingin julkisten kirjastojen olemassa olevaa palveluverkkoa. Saavutettavuuden mittarina käytetään joukkoliikenteen matka-aikoja sekä vuoden 2036 väestöennustetta. Lisäksi työssä määritetään optimaalinen palveluverkko eri skenaarioille.

Valtaosa (99%) aikuisväestöstä saavuttaa lähimmän kirjastonsa 25:ssä minuutissa. Hyvältä vaikuttavasta saavutettavuuden tasosta huolimatta optimointi osoittaa, että nykyisessä verkossa on parannettavaa. Nykyisen 38:n kirjaston tuottama hyöty voidaan saavuttaa 25:llä optimaalisesti sijoitetulla kirjastolla ja 38:n kirjaston optimointi kasvattaa kokonaishyötyä 12,3%. Tuloksista on myös nähtävissä, että optimaalinen uudelleensijoittelu siirtää kirjastoja lievästi läntisistä kaupunginosista itäisiin.

Tulokset osoittavat, että optimointimenetelmien avulla päättäjät voivat organisoida julkisia palveluja siten, että palveluverkoston saavutettavuus ja tehokkuus parantuvat ilman, että kokonaiskapasiteettia kasvatetaan. Tulokset näyttävät palveluverkon tilan vuodelle 2036, jos muutoksia ei tehdä. Tässä työssä rakennettu malli on yleistettävissä myös muihin palveluihin ja se tarjoaa kaupungeille ensimmäisen version työkalusta palveluverkkojen tarvepohjaiseen saavutettavuuden arviointiin.

Avainsanat saavutettavuus, etäisyysvaimeneminen, Helsinki, lineaarinen sekakokonaislukuohjelmointi, yleiset kirjastot, palveluverkon optimointi

Preface

I want to thank my supervisor Pauliina Ilmonen for her supportive and attentive guidance throughout this work. Her commitment was very valuable.

I thank representatives at the City of Helsinki for the collaboration, in particular Tapani Heino and Mikko Niemi at the City Executive Office.

I also thank my dear friend Veikko Sompa for the fruitful late night discussions along the way. Finally, I thank my friends and family for their support and their genuine interest in my work.

Otaniemi, 3 August 2026

Einari Stenberg

Contents

Abstract 3

Abstract (in Finnish) 4

Preface 5

Contents 6

1 Introduction 8
- 1.1 Scope of the thesis . . . 9
- 1.2 The use of technical tools and artificial intelligence . . . 10
- 1.3 Structure . . . 11

2 Background 12
- 2.1 Urban livability . . . 12
  - 2.1.1 Measures of livability . . . 14
- 2.2 Heterogeneity in urban preferences and service proximity . . . 15
  - 2.2.1 Lifestyle vs. life-stage profiling . . . 16
  - 2.2.2 Discrepancies between stated preference and residence location 17
  - 2.2.3 Demographic weighting in livability assessment . . . 17
  - 2.2.4 Relationship between distance from facilities and residence location . . . 18
- 2.3 Facility location and accessibility measurement . . . 19
  - 2.3.1 Classical formulations . . . 19
  - 2.3.2 Distance decay in facility location . . . 21
  - 2.3.3 Diminishing returns and concave utility . . . 23
  - 2.3.4 Efficiency and equity in service location . . . 23

3 Methodology 25
- 3.1 Notation . . . 25
- 3.2 Accessibility measurement . . . 26
  - 3.2.1 Travel time weight function . . . 26
  - 3.2.2 Diminishing value of additional services . . . 27
  - 3.2.3 Necessity scores . . . 29
- 3.3 Livability utility function . . . 29
- 3.4 Optimization constraints . . . 31
  - 3.4.1 Equity constraint . . . 31
- 3.5 Optimization model . . . 31
  - 3.5.1 Linearization of the objective . . . 33

4 Data and implementation 35
- 4.1 Persona profiles . . . 35
  - 4.1.1 Finnish household segmentation . . . 35
  - 4.1.2 Acceptable travel time measures . . . 36

4.2 Travel times . . . 37
4.3 Population prediction and service locations . . . 40
4.4 Implementation . . . 40

5 Results **44**
5.1 Baseline . . . 44
5.1.1 Existing-network coverage . . . 44
5.1.2 Baseline utility . . . 45
5.2 Optimal service locations . . . 50
5.2.1 Adding one service . . . 50
5.2.2 Removing one service . . . 54
5.2.3 50% of baseline utility ($q = 5$) . . . 56
5.2.4 70% of baseline utility ($q = 10$) . . . 58
5.2.5 100% of baseline utility ($q = 25$) . . . 60
5.2.6 Equivalent network size ($q = 38$) . . . 62

6 Discussion and concluding remarks **66**
6.1 Limitations and assumptions . . . 67
6.2 Data considerations . . . 68
6.3 For city planners . . . 68
6.4 Concluding remarks . . . 69

References **70**

1 Introduction

People are moving to cities at an accelerating rate and most of the population growth is expected to happen in cities [48]. Due to this, focus on improving the quality of life in the urban area is becoming an increasingly important topic of research. Furthermore, as the population concentrates in cities, location selection and resource allocation for public services become important to ensure equitable access to all. In addition to good access, public services should take into account the needs of different subpopulations with different profiles.

The population in the City of Helsinki in Finland is aging, and the age distributions are projected to change significantly by the year 2036 [7]. For instance, the elderly population (aged at least 85) is projected to increase by 75% compared to the total population increase of 15%. Due to the shifting population dynamics it is ever more important to have the tools to determine how well the current service network can serve the future population, and what changes might be needed.

Urban design and service network optimization are often difficult tasks due to the nature of urban planning decisions. Most often, high dimensionality of the decision space makes the problems difficult to model, and visual inspection or brute force methods often cannot conclude how optimal the service network is. In addition, realistic justification for model assumptions can be difficult to achieve. In [50], Wang et al. categorize urban data types into seven areas in their analysis of urban renewal: environmental, social, economic, political, accessibility-related, cultural and physical. We select accessibility as the main area of evaluation, because it can be measured objectively, and the measure is directly connected to the location of the service unlike cultural or political measures.

Optimization models designed for urban areas often include distance-decay [15], the diminishing returns of additional nearby services [2, 9], and spatial equity [32] separately. In our review of the literature, we did not find an approach that combines them. In this thesis we develop a model that combines distance decay, diminishing returns of nearby services, need-based weighting of heterogeneous population and travel-time-based accessibility to create a model that aims to measure and optimize a service network. As a practical example, we optimize the library network in the City of Helsinki, and we use the population prediction for the year 2036 to support proactive decision-making.

The main research question of this thesis is how an accessibility-based optimization model can be used to evaluate and improve the placement of public services in the City of Helsinki. Furthermore, we answer the following questions:

- How can service accessibility be modeled so that it includes distance decay, need-based demographic differences, and the diminishing returns of multiple nearby services?
- How accessible is the current service network of libraries in Helsinki?
- How much room is there to improve the current service network of 38 libraries?

This thesis was written in collaboration with the City of Helsinki. The city provided domain expertise and access to population prediction data for the study area.

1.1 Scope of the thesis

Figure 1 shows population density by sub-district of Helsinki with library locations over each area. The visualization combines three datasets from two data sources from the City of Helsinki. Sub-district boundaries and library locations are from the Helsinki City geographic data service [6]. Population statistics by sub-district are from the Helsinki Region open statistics database [8].

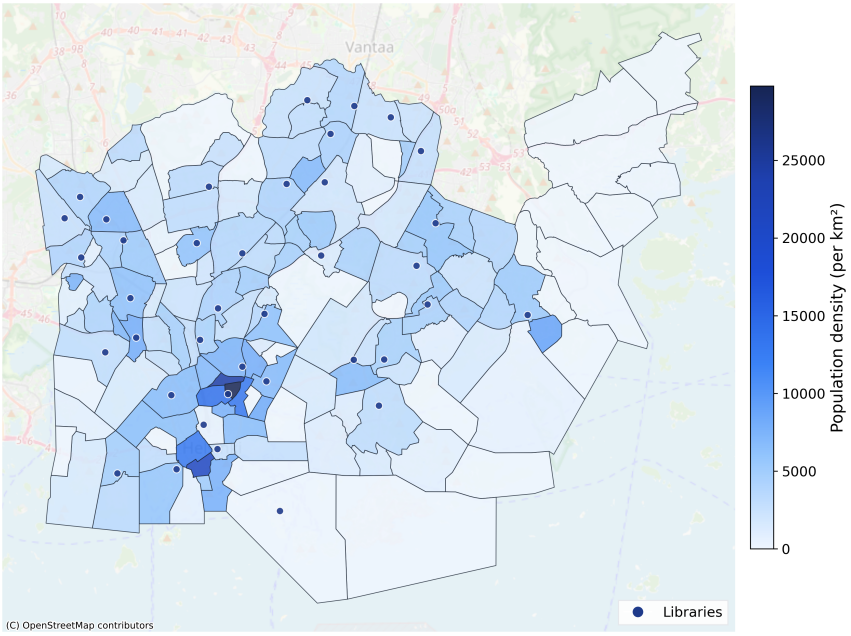


Figure 1: Population density by sub-district, Helsinki, with library locations. The color shows the population density of each sub-district and the markers show the library locations.

Figure 2 illustrates population density of residents aged 65 and over. We can see a clear difference in the population distribution compared to the population of all people living in Helsinki. This age group seems to be more evenly distributed across the city. For this age group the existing locations might serve residents better, as a larger portion of the population lives near a library with lower demand.

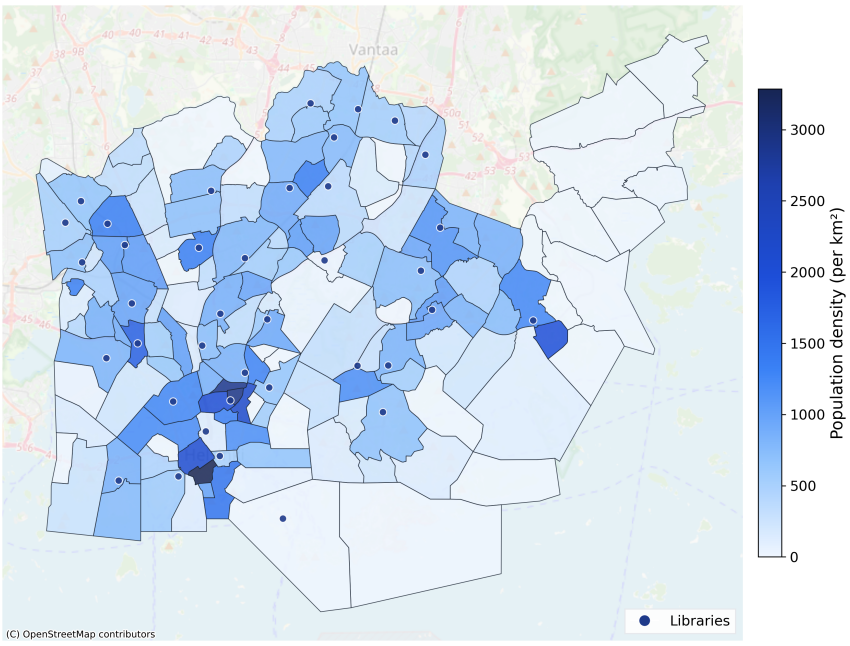


Figure 2: Population density of residents aged 65 and over, by sub-district. Library locations are marked as in Figure 1.

The geographic scope of this thesis is the municipality of Helsinki, and public libraries are selected as the service network of study. The network is managed by the City of Helsinki and the service was selected based on data availability, a manageable scope, and practical potential for improvement. The libraries serve as an example of a public service, but the framework developed in this thesis can be generalized to any service network. The model itself uses public-transport travel times to estimate accessibility, and population predictions for 2036 to support proactive decision-making. Moreover, we segment the population into profiles in order to capture the needs of different population groups. The optimization considers only service location and not capacity, construction costs or building constraints. These were identified to be out of scope for this thesis.

1.2 The use of technical tools and artificial intelligence

The technical tools used in this thesis include Visual Studio Code for software development and implementation and GitHub for version control. Furthermore, an AI-coding assistant was used to support data processing and visualizations. Artificial intelligence was also used to support grammar checks and to suggest improvements in the coherence of the written text as well as to help with LaTeX syntax. No interpretations, results or research findings were generated by AI.

The author of this thesis has verified all code, analysis, results and written content produced by any technical tool or method and takes full responsibility for all work displayed in this thesis.

1.3 Structure

The structure of this thesis is the following: Section 2 evaluates the current literature on livability, accessibility and facility location. Section 3 introduces the accessibility measure, the livability utility function and the service network optimization model. Section 4 describes the data used in the calculations and gives an outline on the implementation pipeline. Section 5 presents accessibility measurement for the current library network as well as results for the optimization scenarios and finally in Section 6 we discuss the findings, limitations and conclude how the model can be used by city planners.

2 Background

In this section we review existing literature on livability, accessibility and methods to optimize and measure service networks. We start by looking at livability: we review how it is defined, the separation of subjective and objective measures and how qualities of urban areas can be evaluated. Then, we focus on population heterogeneity and segmentation based on lifestyle and life-stage preferences and needs. Finally, we scope and review operations research literature on how facility location problems can be defined and modeled.

In a review article of 77 studies [17], Gonzalez Bohorquez et al. looked at how livability is defined across domains and topics. They included peer-reviewed empirical studies in which livability was a central concept. After excluding theoretical studies, they took a three-step approach for the review: first, how livability was defined and in what context. Second, what places and populations were used in the studies and why, and third, what measurements were taken to evaluate livability.

The authors conclude that livability is a contested concept that is characterized by ambiguity and inconsistency. For example, as many as 40 unique conceptual frameworks for livability were found. The authors also identified and ranked domains of livability from the reviewed research. Among many others, the three highest-ranked livability domain metrics were public space and infrastructure, transportation and neighborhood amenities.

2.1 Urban livability

Generally, measures of livability fall into two domains. First are the objective factors, which for instance include spatial patterns of the urban network [42] or traffic volume in the area [12]. Second are the subjective factors that include perceived safety, satisfaction with the neighborhood or sense of belonging in the community.

Higgs et al. conducted a study in Melbourne, Australia that highlights the objective approach by developing an urban livability index (ULI) [22]. To develop the ULI, they looked at measures across six domains: transport, social infrastructure, employment, walkability, housing, and green infrastructure. Then they tested the index against travel behavior data from an existing survey of over 12 000 adults to see if behavior and the calculated ULI have a relationship. Clear connections to several demographic differences were observed. For example, residents living in lower ULI areas were on average more likely to have children and less likely to live in a single-person household. Furthermore, residents in high ULI areas were more likely to choose active transportation (walking or cycling). These results suggest that differences in human behavior and well-being can be measured by looking at objective measures of the environment. However, these represent only one side of the livability concept.

Recent studies in the Finnish context suggest that evaluating objective physical measures in livability analysis may be insufficient alone. Pasanen et al., in [35], studied five Finnish cities including Helsinki and found that objective measures like urban zones have an impact, but subjective measures often play a stronger role for

livability. For example, residents' satisfaction with green spaces was more strongly associated with livability than the actual percentage of green space coverage.

Pasanen et al. argue further that community participation is often overlooked in livability studies despite being a key element of the WHO healthy cities initiative [10]. Pasanen et al. also found that possibilities to influence decisions regarding the neighborhood were one of the strongest predictors of perceived livability. A strong argument can be made for social livability measures: it seems that social factors have the highest impact on the perceived quality of life in urban areas, after a certain minimum level of living standard is achieved.

In addition to social factors, Pasanen et al. identified other important subjective measures. For example, these include perceived safety and maintenance of traffic routes. In general, it seems that for a livability framework to be sufficiently granular in the context of developed countries, it should account for social factors alongside the key infrastructural properties of an urban area.

In [40], Riepl and Grabow took a different approach in quantifying livability through human need theory. They developed a foundational livability indicator (FLI), a measure that defines foundational livability as the capacity of a place to provide services that satisfy a person's needs. The indicator has two key components of accessibility: availability and affordability. Availability was measured via 13 sectors across Vienna, and affordability was estimated through a measure which calculates households' residual income after covering costs for housing, utilities, transport, and food. The study revealed significant regional differences in availability and affordability that can arguably be extended to represent measures of livability.

The studies of Vienna and Melbourne show that the livability analysis at micro level can highlight significant patterns in the small scale urban environment. However, livability can also be reviewed at the macro level though the analysis can become much more complicated and debated. Paul and Sen, in their critical review [36], argue that three findings from the literature confirm the unreliability of global indexes. First, approaches vary regionally: Western approaches tend to prioritize physical and financial factors, whereas Eastern approaches favor socio-cultural measures. Second, livability is a complex problem comprising physical, environmental, and cultural factors rather than a simple measure. Third, global rankings often produce inconsistent outcomes because different indexes apply unequal weights for domains such as healthcare and infrastructure.

These findings show that livability has no single definition. What is measured, and how it is weighted, largely determines the result. Instead of building a complete livability index, in this thesis we focus on one measure, accessibility, and we choose our model weights based on the local context and input from City of Helsinki experts. However, to fully represent livability, it seems that both objective and subjective measures are needed. Objective measures provide the baseline for the livability of the urban environment and subjective measures add a layer of social factors that arguably have a significant impact on a person's well-being. In this thesis we focus on objective measures, and leave the evaluation of subjective measures to future research.

2.1.1 Measures of livability

As discussed, estimating livability is complicated and often involves a large variety of measures. As Table 1 illustrates, they range from socio-demographic and economic characteristics to environmental accessibility, subjective satisfaction, and governance. In this thesis we focus mainly on objective environmental measures, specifically the accessibility of public services, weighted by population need, because these are used by city planners and require no survey data.

Table 1: Livability measures used across reviewed studies. Each row lists a measure and the reviewed studies that use it.

Measure	Used in study
Socio-demographic factors	
Age	[35], [22], [12], [49]
Gender	[35], [12]
Employment status	[35], [22], [12]
Education level	[35], [12], [27], [42]
Household composition	[35], [40], [49]
Population density	[28], [23], [20]
Health status	[27], [13]
Economic factors	
Household income	[35], [28], [49], [22]
Area-level socioeconomic deprivation	[35], [42]
Affordability (residual income after essentials)	[40]
Housing costs	[40], [28], [49]
Employment opportunities	[42], [22]
Dwelling and housing characteristics	
Dwelling type and ownership	[35], [13], [20]
Floor space and dwelling quality	[40], [25]
Housing price	[28], [42]
Dwelling density	[22], [12]
Green view from home	[35]
Objective environmental & service accessibility	
Service accessibility (healthcare, education, food, culture)	[40], [36], [28], [22], [41], [17]
Public transport accessibility	[40], [28], [22], [27], [3], [41]
Green and blue space	[35], [40], [28], [22], [20], [23], [17]
Walkability and street connectivity	[22], [27], [3], [49], [50], [17]
Air quality and noise pollution	[22], [50], [41]

Measure	Used in study
Subjective satisfaction & sentiment	
Life satisfaction and well-being	[13], [27]
Perceived safety (traffic, crime)	[35], [13], [27], [41]
Satisfaction with green and blue spaces	[35]
Satisfaction with cleanliness and maintenance	[35], [13]
Satisfaction with public transport	[35]
Neighborhood appearance and attractiveness	[27], [13], [25]
Social media sentiment	[28]
Sense of belonging and community attachment	[41], [20]
Governance, participation, and social factors	
Participation in neighborhood decisions	[35], [10], [17]
Social cohesion and trust	[27], [41], [49]
Crime rate	[28], [27]
Community empowerment and inclusion	[10], [17]

2.2 Heterogeneity in urban preferences and service proximity

Treating the population as a single average hides important differences. For example, a young adult and an elderly resident living in the same area might use a public service in very different ways. It might be beneficial to divide the population into segments that capture how these groups behave differently.

In [25], Kauko used expert judgement to show that homebuyers in inner-city Budapest rank the same location attributes differently when choosing where to live. He argues that in a diverse market, demand is split into several preference profiles rather than following one average. Kauko concludes that in Budapest there are four apparent preference profiles. These profiles are illustrated in Table 2.

Table 2: Housing-preference profiles in Budapest [25]. Each row lists a profile identified in the study and its key preferences.

Profile name	Key preferences and characteristics
Traditional urban	Prefers historical architecture and urban aesthetics over practical factors like access to services or transport.

Profile name	Key preferences and characteristics
Segregation	Chooses neighborhoods based on socio-economic status. Wants to live near people of similar social class.
Tiebout (municipality)	Picks locations by district reputation rather than building quality. The district name matters more than the actual property.
CBD accessibility	Wants short commutes to the city center, schools, and public transport.

On the other hand, Palicki, in [34], found empirically that housing feature preferences are linked to life stage. The study was conducted by surveying 2537 people in Poznań, Poland. He showed that younger adults prefer location more than other attributes, compared to older adults.

Palicki's research considered flat features such as technical condition, heating, location, size, and brightness, and ranked these attributes within specific age groups. He draws three important conclusions. First, interest in flat features increases with age but decreases after the age of 70. Second, people older than 45 tend to have the most consistent preferences across age groups, which might imply that preferences converge after a certain age. Third, the most stable features were those ranked either lowest or highest among all age groups. Palicki's analysis links housing preferences to life stage but is limited to flat features and does not consider public service needs or how lifestyle might affect these preferences beyond life stage.

2.2.1 Lifestyle vs. life-stage profiling

In [24], Jansen et al. define two approaches to profile an urban resident: lifestyle preference and life-stage needs. The authors evaluate both approaches by analyzing the reasons for the selected residence location.

The life-stage need approach is based on the assumption that housing needs are mostly driven by aging or other life-stage changes. The approach takes into consideration characteristics such as age, household composition and labor market position and uses them to predict housing choice. This framework assumes that demand for housing space and services is driven by household transitions such as formation, expansion and contraction. For instance, the transition from a couple to a family with small children creates a demand for new services and could drive the need to change the living environment.

On the other hand, the lifestyle preference approach assumes that housing choices are driven by values, attitudes, and identity. Two individuals of the same age and household composition may make very different housing choices depending on the individual's values and identity. For example, urban-oriented or nature-oriented preferences can strongly affect a person's housing choice [20]. Naturally, the lifestyle

approach relies on assumptions that do not generalize well to the population level and the method often requires survey data that is resource intensive to gather.

In practice, service needs are often characterized by both life-stage factors and lifestyle preferences and the two approaches should not be assumed independent. In this thesis, the life-stage approach is more practical. Furthermore, Palicki showed that preferences change with life stage [34], meaning that life-stage segmentation also captures at least part of the preference variation. However, even well-motivated segmentation does not guarantee that modeled preferences reflect actual residential choices.

2.2.2 Discrepancies between stated preference and residence location

In [20], Hasanzadeh et al. studied the difference between the residents' stated preference and the actual housing choice in the city of Tampere, Finland. They categorized residents into urban tribes (urbanists vs. nature lovers) based on their behavioral preferences and found that these groups hold different criteria for a desirable living environment. For instance, urbanists prioritize service density and accessibility, while nature lovers favor green space even at the cost of lower accessibility. Their results showed a considerable difference between the stated preference of housing location and the actual residence location. One takeaway is that when modeling individuals' preferences, it is difficult to consider all the possible externalities that affect housing location choice.

Vasanen, in the study [49], analyzed stated and revealed preferences further. Vasanen found that the drivers of residential mobility differ significantly depending on the direction of the move. Moves outward from the city center of Turku are driven by a clear preference for a specific type of living environment. On the other hand, people moving to the city center were often driven by needs (age, life situation) rather than preference. This pattern might hold for Helsinki as well. People living in the suburban areas might be more preferentially inclined to do so compared to people living in the city center of Helsinki.

Based on these findings, stated preferences may not reliably predict residential choices. It also highlights a key issue of stated preference: the collected data is likely to be biased from numerous human factors. For service location modeling, observed demographic characteristics might provide a better foundation for segmentation compared to self-reported preferences.

2.2.3 Demographic weighting in livability assessment

As discussed in previous subsections, the needs and usage patterns of residents vary across the population profiles. Accessibility measures can model the heterogeneity in the population by weighting services based on usage frequency or by weighting population profiles based on their need for the service.

Abbiasov et al. conducted a large-scale evaluation of local living patterns across US residential areas in [1]. They found a strong relationship between income and the locality of trips to nearby services. The bottom income decile takes 41% of its

trips locally, compared with only 11% in the top decile. Furthermore, the authors introduced an accessibility measure called 15-minute access, which weights each service type based on the observed usage patterns of the service. As a result, the authors could calculate an index that measures the accessibility of an area based on whether services are within a 15-minute travel time. The outcome was a measurement that effectively compares US neighborhoods based on walking distance and service usage.

In comparison to weighting services by observed usage, one can weight the population based on the need of each service. With this approach, it is possible to assess the gap between usage and the modeled needs of different population groups. There are numerous examples of need-based population weighting in the literature. Eyles et al. [14] introduced a methodology for allocating health-care resources that weights populations by the standardized mortality ratio. Kephart and Asada [26] worked with a similar approach by comparing alternative need-based weighting approaches. They showed that the choice of weighting logic can shift the results significantly. The need-based principle has also been used directly in accessibility modeling. Ngui and Apparicio [33] weighted the population by age-specific utilization rates to measure accessibility to medical clinics in Montreal. Liu et al. [29] applied a stricter form of the same idea by measuring accessibility to elderly care facilities with demand restricted to elderly residents. The study most closely related to this thesis in terms of weighting method is the SDA-2SFCA method of Shao and Luo [43]. They weight the population at each location of their model by age- and gender-specific needs. In this thesis, we also use a weighting approach that is based on age groups.

2.2.4 Relationship between distance from facilities and residence location

The fundamental assumption in several studies is that proximity to services improves well-being. However, this assumption could be questioned. A study by Ettema and Schekkerman suggests that this might not be a straightforward analysis [13]. Their key objectives were to compare the impact of objective and subjective spatial variables. The authors found that life satisfaction and affective well-being are much more strongly influenced by subjective variables than objective variables. They also found that features of the nearby neighborhood have a substantially greater impact on well-being compared to the wider accessibility of services. However, a notable finding was on the relationship between distance to urban facilities and well-being. People living farther from services such as department stores, bars and restaurants reported higher satisfaction, suggesting that there is a positive correlation between distance to services and life satisfaction. They hypothesized that living close to these facilities introduces nuisances like crowding and noise pollution.

While this finding is intriguing, a few limitations need to be addressed. First, the study showed that desirable neighborhoods are associated with higher satisfaction, but the direction of causality was not clear. The study could not establish if good neighborhoods make people happy, or if happy people simply move to good neighborhoods. This might also explain the peculiar result mentioned above. Furthermore, the results could be context specific and they might not generalize to other urban areas.

In [27], Leyden et al. investigated if living in a walkable neighborhood in Dublin, Ireland makes people happier. The authors accounted for the age of studied people and thus could determine if the meaning of walkability changes across the lifespan. Furthermore, the authors found that living in a walkable neighborhood had a strong direct positive effect on happiness for adults aged 36–45. They also concluded that for the older population the effect was similar but indirect. Walkability created conditions for happiness from opportunities like easier social interaction. The study of Leyden et al. shares the same limitations regarding causality and context specificity as Ettema and Schekkerman's.

Even though the discussion in literature is conflicting, the majority of the research on this topic agrees on the direction of causality. Sahu and Agrawal, in [41], summarize results across 66 studies and show that spatial proximity is positively linked to residents' perceptions of quality of life. A possible explanation for the conflicting results is the externalities of local living conditions: in densely populated areas, there seems to be a proximity threshold for services after which the satisfaction of residence location starts to decline, with noise and pollution as significant contributors. The seemingly contradictory findings show that the relationship between service proximity and residential satisfaction is multidimensional and needs to be assessed case by case.

2.3 Facility location and accessibility measurement

The complexity of residential preferences discussed above can be addressed through a formal facility location optimization model. The facility location problem has been studied extensively in operations research and location science [38]. In this section, we look at three classical formulations, and we review modern models that measure and optimize facility locations.

2.3.1 Classical formulations

The so-called p-median problem was first studied by Hakimi [18, 19]. The problem aims to locate p facilities in a network to minimize the total distance from demand points to their nearest facility. In other words the goal is to minimize the sum of distances (for example total wire length or travel cost) between demand points and the center.

Hakimi proved that there always exists a p-median located entirely at the vertices of the network. This means that in order to find an optimal solution, we only have to examine subsets of vertices, reducing the optimization problem from continuous to discrete.

This problem is commonly expressed as a binary integer program [39]. Let I be the set of demand locations, J the set of candidate facility sites, d_{ij} the distance from demand i to facility j, and P_i the demand (e.g. population) at location i. We define $y \in \{0,1\}^{|J|}$ as a binary decision vector of length $|J|$, where each element y_j indicates whether a facility is placed at site j. The decision variable $x \in \{0,1\}^{|I|\times|J|}$ is a binary assignment matrix, where each element x_{ij} indicates whether demand point i

is assigned to facility j. The problem can be stated as:

$$\begin{aligned} \underset{x,\,y}{\text{minimize}} \quad & \sum_{i \in I} \sum_{j \in J} P_i d_{ij} x_{ij} \\ \text{subject to} \quad & \sum_{j \in J} x_{ij} = 1, \quad \forall i \in I, \\ & x_{ij} \leq y_j, \quad \forall i \in I, \forall j \in J, \\ & \sum_{j \in J} y_j = p. \end{aligned}$$

The objective is to minimize resource use, expressed as the total distance traveled by residents. This does not mean that all residents are within an acceptable service distance.

Toregas et al. formulated facility location as a set covering problem (SCP) in 1971 [47]. Compared to the p-median problem that optimizes for efficiency, the SCP is based on equity. The goal is to find the minimum number of facilities required to ensure that every demand point is within a distance d_0 of a facility. We define $C_i = \{j \in J : d_{ij} \leq d_0\}$ as the set of candidate sites that can cover demand i, and $y_j \in \{0, 1\}$ indicates whether a facility is placed at site j. The problem can be formulated as:

$$\begin{aligned} \underset{y}{\text{minimize}} \quad & \sum_{j \in J} y_j \\ \text{subject to} \quad & \sum_{j \in C_i} y_j \geq 1, \quad \forall i \in I. \end{aligned}$$

This approach guarantees full coverage of the network, but may need more resources than are available. The SCP also treats every demand point's coverage as a binary variable. In a more realistic case, we should also consider the gradual decay of accessibility with distance.

The maximal covering location problem (MCLP) was introduced by Church and ReVelle in 1974 [5]. The model maximizes the total demand covered within the threshold distance, using a fixed number of facilities p. This means that the model addresses both efficiency and coverage. We define $z_i \in \{0, 1\}$ to indicate whether demand point i is covered by at least one facility, and $y_j \in \{0, 1\}$ to indicate whether a facility is placed at site j. The problem can be defined as follows:

$$\begin{aligned} \underset{y,\,z}{\text{maximize}} \quad & \sum_{i \in I} P_i z_i \\ \text{subject to} \quad & \sum_{j \in J} y_j = p, \\ & \sum_{j \in C_i} y_j \geq z_i, \quad \forall i \in I. \end{aligned}$$

Table 3 compares the three classical formulations. The p-median problem focuses on efficiency by minimizing the total distance traveled. However, it does not guarantee

coverage, so some demand points can still be unreasonably far away. SCP, on the other hand, ensures that a minimum level of equity is achieved through coverage, but it does not consider resource availability. MCLP uses a similar approach to SCP, but the number of facilities is fixed. As a result, we can find a network that maximizes coverage under resource constraints, but it does not consider varying the number of facilities.

Table 3: Classical facility-location problem variants (p-median, SCP and MCLP) compared. Each column summarizes the objective, constraint, orientation and distance measure of one formulation.

	p-median	SCP	MCLP
Objective	Minimize total distance	Minimize facility count	Maximize covered demand
Constraint	Fixed facility count	Full coverage required	Fixed facility count
Orientation	Efficiency	Equity	Coverage under fixed resources
Distance	Linear	Binary	Binary

The classical facility location problems form the foundation for more complex and nuanced models. Before developing a more realistic model, the assumptions of the classical formulations need to be examined. For example, they all consider the distance measure to be either linear or a binary function. Farhan and Murray, in [15], identify issues that classical facility location models do not consider simultaneously: coverage range, distance decay, and partial regional service. As a solution, they developed a general linear optimization model that adjusts the demand P_i by a distance-decay function, using the exponential form $\hat{P}_{ij} = P_i e^{-\beta d_{ij}}$ in their application. Furthermore, Farhan and Murray argue that a continuous distance-decay representation captures demand more realistically than the binary, step-function coverage assumed in the classical models. They adopt the exponential form because it describes the short-distance interactions better for urban areas. However, their model still assumes homogeneous demand and equal attractiveness of different facilities, and does not consider marginal benefits, capacity, or real travel times.

2.3.2 Distance decay in facility location

Distance decay describes how the demand for a service declines as the travel distance or time to the service increases. The approach is widely used for modeling facility demand across different applications. Drezner et al., in [11], showed that the distance decay parameter should not be uniform across facilities. Services with higher perceived attractiveness tend to have a slower decline in demand with distance. The findings of Drezner et al. further support the viewpoint that the decay parameter should be selected case by case if differences in demand by characteristic can be shown.

One established method for implementing distance decay in accessibility evaluation and modeling is the floating catchment area method. Radke and Mu [37] introduced a

method for decomposing overlapping service regions and computing the probability of each household being served. Luo and Wang [31] then built on their work and formalized the two-step floating catchment area (2SFCA) method, which calculates a supply-to-demand ratio for each service within its catchment and then aggregates these ratios for each population location.

The 2SFCA method assumes that inside a certain catchment area, every point has equal access to the service. To improve the method, Luo and Qi [30] introduced the enhanced two-step floating catchment area (E2SFCA) method. The E2SFCA method uses the same two-step approach, but includes a distance decay function to account for the decrease in accessibility as travel time increases instead of assuming that the relationship is binary inside the catchment.

The decay function can take several forms, of which the Gaussian is a common choice [44, 23, 4]. Using the Gaussian is also intuitive. The function naturally takes into account that changes at very short and very long travel times do not have a large effect on the weight given to a particular service location.

The E2SFCA accessibility score is calculated in two steps, which are defined as follows. Let I denote the set of population locations and J the set of service locations. For each population location $i \in I$ and service location $j \in J$, let d_{ij} denote the travel time between them. The distance decay weight function w_{ij} assigns higher weights to shorter travel times. The weights are calculated based on a Gaussian function with an appropriate decay parameter.

In the first step, a supply-to-demand ratio is calculated for each service location. For each service location $j \in J$, define the catchment area $C_j = \{i \in I : d_{ij} \leq d_0\}$ containing all population locations within the threshold travel time. Let K_j denote the supply capacity at service location j (e.g., number of service units or a service quality measure). The weighted population demand within the catchment area is calculated as:

$$D_j = \sum_{i \in C_j} P_i \cdot w_{ij},$$

where P_i is the population at location i. The supply-to-demand ratio ρ_j for service location j is then:

$$\rho_j = \frac{K_j}{D_j} = \frac{K_j}{\sum_{i \in C_j} P_i \cdot w_{ij}}.$$

In the second step, an accessibility score is calculated for each population location. For each population location $i \in I$, define the catchment area $C_i = \{j \in J : d_{ij} \leq d_0\}$ containing all service locations within the threshold travel time. The accessibility score A_i for population location i is calculated by summing the supply-to-demand ratios of all accessible services, weighted by the distance decay function:

$$A_i = \sum_{j \in C_i} \rho_j \cdot w_{ij} = \sum_{j \in C_i} \frac{K_j}{\sum_{\ell \in C_j} P_\ell \cdot w_{\ell j}} \cdot w_{ij}.$$

This two-step process accounts for both the supply capacity of services and the competition for those services from surrounding populations. The accessibility score

A_i quantifies how well-served location i is, taking into account the availability of services and the travel time required to reach them.

2.3.3 Diminishing returns and concave utility

Distance decay models the relationship between demand and distance, but it does not address how demand for a facility changes when multiple services are near the same location. For example, if a library is added next to an existing one, total utility gained should not double. In reality, the added utility diminishes as the number of nearby facilities increases.

In [2], Ahmed and Atamtürk studied optimization problems where the utility function has the form

$$g(T) = f\left(\sum_{j \in T} v_j\right),$$

where $T \subseteq E$ and E is a finite ground set, v_j is the contribution of element j (in the facility location context this would be the accessibility contribution of each site), and f is strictly concave and increasing.

A set function g is submodular if

$$g(T_1) + g(T_2) \geq g(T_1 \cup T_2) + g(T_1 \cap T_2),$$

for all $T_1, T_2 \subseteq E$. Intuitively, this means that the utility function has diminishing marginal returns. The property of submodularity is directly applicable to facility location problems. If a concave utility function is defined, we can define a model where an additional service contributes less than the previous one. Ahmed and Atamtürk showed that the exponential utility function $f(a) = 1 - e^{-a/\lambda}$ satisfies the submodularity condition when the contributions are nonnegative making it a good candidate to model diminishing returns of utility for nearby facilities.

Coniglio et al. [9] extended these results to maximal covering location problems. They developed mixed-integer linear programming (MILP) reformulations for maximizing concave utility in covering problems under constraints such as fixed facility count. The fixed facility count case is close to the structure of the model in this thesis.

In this thesis, we use a concave function over the weighted accessibility score to prevent the model from clustering services in a single area. Without the marginal benefit property, the model would likely prefer to place new services next to existing ones in densely populated areas rather than spreading them across the network.

2.3.4 Efficiency and equity in service location

Even with diminishing returns, maximizing total utility alone can favor high-density areas, potentially leaving sparsely populated locations without reasonable access to services. One can argue that such access is an important consideration for public services, and that equity should be enforced even if it reduces efficiency.

Marsh and Schilling [32] reviewed how equity has been measured in facility location problems. They examined 20 different equity measures and proposed a

framework to organize them along three dimensions: choice of reference distribution, distance metric, and scaling. They found no consensus in the literature on how or which equity measures to use. Furthermore, they highlighted that there is often a conflict between equity and efficiency. As the equity of the facility network increases, its efficiency usually decreases.

Facility location models inevitably involve value judgments, and spatial equity is one example that can be built into the objective of the model. The balance between efficiency and equity should be selected based on the priorities of the decision maker, and this choice significantly affects the outcome.

3 Methodology

The background in Section 2 highlights that livability has no single definition, and that even within a restricted scope, several challenges remain. This section suggests an optimization approach that addresses many of them.

The section is divided into three parts. First, we define the accessibility measure, which combines the travel-time weights over the sites of each service type. Second, we construct the utility function, which applies diminishing returns to the accessibility score and weights it by profile-specific need. Finally, we formulate a discrete optimization problem that selects the best locations for new services and the least harmful removals of existing ones.

3.1 Notation

Note that the notation in this section is slightly different from the classical formulations in Section 2.3. Here I denotes the set of grid cells of demand, J the set of existing service locations, and the candidate sites for new services are separated into a distinct set N. The sites of service type s form the set $\bar{J}_s \subseteq J \cup N$.

Table 4 summarizes the sets and parameters used throughout this section.

Table 4: Notation used throughout this section.

Symbol	Description
I	Set of grid cells, indexed by i
J	Set of existing service locations
N	Set of candidate locations for new services
S	Set of service types, indexed by s
$\bar{J}_s$	Set of active and candidate sites for service type s, $\bar{J}_s \subseteq J \cup N$, indexed by j
R	Set of profiles, indexed by r
d_{ij}	Travel time from cell i to site j (minutes)
P_{ir}	Population of profile r in cell i
n_{rs}	Necessity score of service s for profile r
β_{rs}	Distance decay parameter for profile r and service type s
h_{rs}	Threshold travel time for calibrating β_{rs}
α	Saturation parameter for diminishing value function
a^*	Saturation threshold for calibrating α
$\tau_{\max}$	Maximum allowable travel time for coverage
C_i	Set of sites that can cover cell i within $\tau_{\max}$, $C_i = \{j \in \bar{J}_s : d_{ij} \leq \tau_{\max}\}$
q	Number of service sites to open
M_i	Penalty constant for uncovered population at cell i
w_{ijrs}	Travel-time decay weight between cell i and site j for profile r and service s
$y_j \in \{0, 1\}$	1 if site $j \in \bar{J}_s$ is open, 0 otherwise (decision variable)
$\delta_i \in \{0, 1\}$	1 if cell i is uncovered, 0 otherwise (coverage slack variable)
a_{irs}	Raw accessibility score at cell i for profile r, $\sum_{j \in \bar{J}_s} w_{ijrs}\, y_j$
u_{irs}	Individual utility of profile r at cell i for service s

3.2 Accessibility measurement

3.2.1 Travel time weight function

As discussed in Section 2.3.2, the relationship between demand and travel time to a service location is nonlinear, and the Gaussian function is a common choice for modeling it. We select the Gaussian function here as well. For each location $i \in I$, profile $r \in R$ and service site $j \in \bar{J}_s$, the distance-based weight is then defined as

$$w_{ijrs} = e^{-d_{ij}^2/\beta_{rs}},$$

where d_{ij} is the travel time between location i and site j, and β_{rs} is a scale parameter that controls how quickly accessibility declines with travel time for profile r and service type s.

To calibrate β_{rs} for each profile-service pair, we set a threshold travel time h_{rs} such that the weight becomes smaller than 0.1 after h_{rs}, i.e. $w_{ijrs} < 0.1$ for $d_{ij} > h_{rs}$.

We choose 0.1 as the calibration point because at the threshold travel time the service retains only 10% of its maximum attractiveness. Fixing $w_{ijrs} = 0.1$ at travel time h_{rs} gives us the following equation:

$$0.1 = e^{-h_{rs}^2/\beta_{rs}} \iff \ln(0.1) = -\frac{h_{rs}^2}{\beta_{rs}} \iff \beta_{rs} = -\frac{h_{rs}^2}{\ln(0.1)}.$$

Given a behaviorally meaningful threshold travel time h_{rs} (for example, an acceptable maximum travel time) for each profile-service pair, we can derive a corresponding β_{rs} that reflects that profile's travel behavior for that service type. This threshold h_{rs} replaces the single global catchment threshold d_0 of Section 2.3.2 with a profile- and service-specific calibration point. Unlike d_0, it is not a hard cutoff and sites beyond h_{rs} still contribute a small weight.

However, we set $w_{ijrs} = 0$ when it falls below 0.01 to increase computational efficiency. Each weight is below 1% of the weight of a single service at zero travel time, so dropping it has no meaningful effect on the score.

Figure 3 illustrates the decay weight as a function of travel time for different threshold values h.

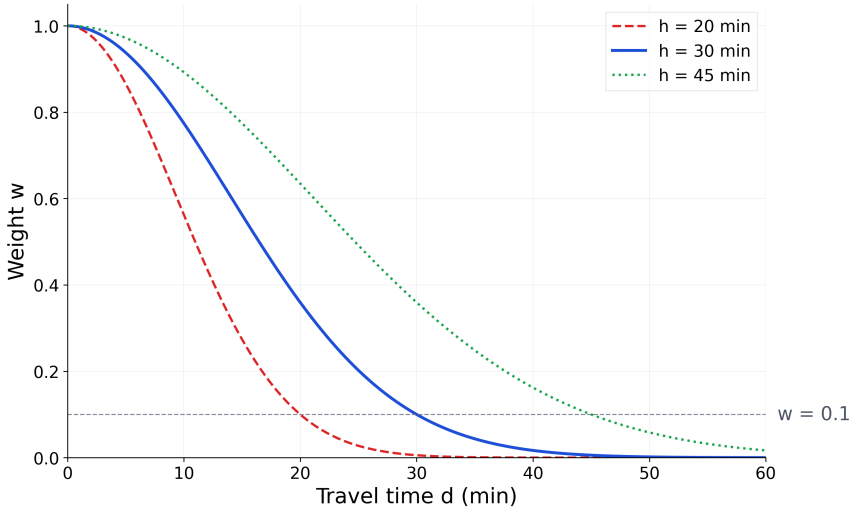


Figure 3: Travel-time decay weight w_{ijrs} for varying threshold h. Each curve shows the weight as a function of public transport travel time for one threshold value $h \in \{20, 30, 45\}$ minutes. The dashed line shows the calibration point $w = 0.1$, which is reached at the threshold travel time.

3.2.2 Diminishing value of additional services

Let us consider a cell that has a service. If we were to add a new service to that grid cell, or to a very nearby cell, it would be reasonable to assume that the total utility increase for a person in that cell would not double. Instead, there is a diminishing amount of utility that is added to the existing utility of the location. Due to this, we introduce a concave function f that controls the utility gained from additional nearby

services. We calibrate it with a fixed reference point, similarly to the travel time weight function.

Let us define an accessibility score at location i for profile r as the weighted sum of active services

$$a_{irs}(y) = \sum_{j \in \bar{J}_s} w_{ijrs}\, y_j.$$

To model diminishing returns, we apply the exponential concave function discussed in Section 2.3.3. We select this function because it satisfies the submodularity condition [2], it requires only a single parameter α for calibration, and it has a natural interpretation ($f(0) = 0$ and $f(a) \to 1$ as $a \to \infty$). The function is defined as

$$f(a) = 1 - e^{-\alpha\, a},$$

where $\alpha > 0$ controls how quickly the marginal benefit reaches a saturation point. This is the exponential utility function of Section 2.3.3 with $\alpha = 1/\lambda$. To calibrate α, we set a saturation threshold a^* such that $f(a^*) = 0.8$, i.e. the benefit has reached 80% of its maximum at accessibility level a^*. The 80% is selected as it fits the assumption that a single well-accessible service captures most of the utility and the remaining 20% represents the additional value from backup or redundancy provided by further services. Fixing $f(a^*) = 0.8$ gives

$$0.8 = 1 - e^{-\alpha\, a^*} \iff \ln(0.2) = -\alpha\, a^* \iff \alpha = -\frac{\ln(0.2)}{a^*}.$$

We set $a^* = 0.5$ based on discussions with experts at the City of Helsinki [21]. The assumption is that residents primarily benefit from their nearest service and a single service captures most of the accessibility benefit. The physical interpretation for the value, using the maximum travel times defined in Section 4.1.2, is that a young adult would reach $a^* = 0.5$ with a single library at about 11 minutes away, two libraries at about 15.5 minutes away or five libraries at the boundary of 20 minutes. Furthermore, a library at the same cell would receive $f(a) \approx 96\%$, meaning only very marginal improvement for stacking libraries can be achieved. The selected a^* gives $\alpha = \frac{\ln(5)}{0.5} \approx 3.22$. Figure 4 shows the concave function, where the effect of marginal return is seen.

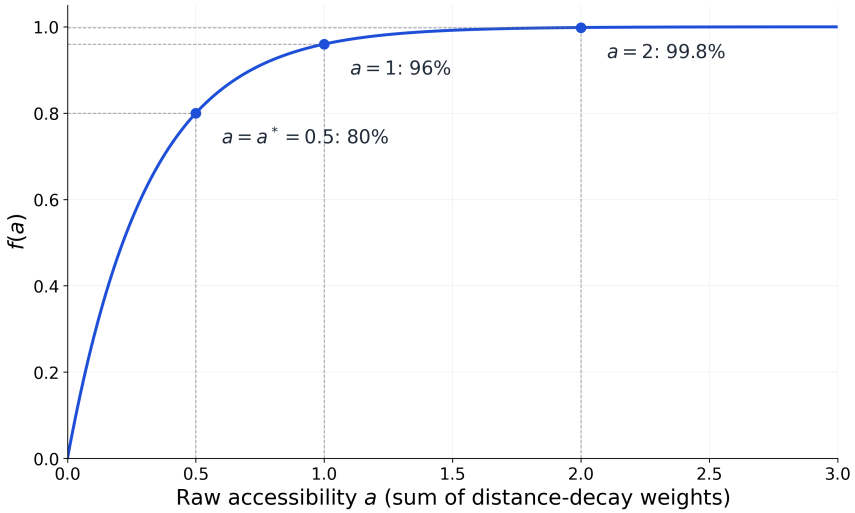


Figure 4: Diminishing-returns function $f(a)$ with $\alpha \approx 3.22$. The annotated points show the calibration $f(a^*) = 0.8$ at $a^* = 0.5$ and the flattening marginal returns at $a = 1$ and $a = 2$.

3.2.3 Necessity scores

Necessity scores are defined to differentiate population segments from each other by service need. As an example, an elderly person might depend on a library nearby more than a young student. We give each profile a necessity score $n_{rs} \in \{0, 1, 2, 3, 4, 5\}$ that weights how important service s is for profile r. In our model, the necessity score replaces the supply-to-demand ratio ρ_j of the E2SFCA method discussed in Section 2.3.2, shifting the weighting from a supply-side capacity measure to a demand-side need measure. The integer scale from 0 to 5 is chosen so that experts at the City of Helsinki can assign importance levels with consistency and interpretability in mind. It should be noted that the model treats these scores as numeric weights, meaning that a score of 4 has twice the effect of a score of 2. The specific profiles and their necessity score values are defined in Section 4.

Our model does not include a site-specific capacity parameter. We assume that all service locations of a given type provide a similar baseline level of service (in other words all services have infinite capacity). This assumption is reasonable for the Helsinki library network, where all branches offer similar core services and libraries rarely approach a hard capacity limit. For other service types with large capacity differences and well-defined limits, the model should be extended to include a capacity constraint.

3.3 Livability utility function

This thesis applies the model to a single service type (public libraries), but we define the framework over a general set S of service types to allow extension to multi-service analysis in future work. Before introducing the optimization model, we define the livability utility for a fixed service type and a fixed configuration of sites.

Figure 5 illustrates a catchment area for a selected location. At location X, multiple service sites are included in the calculation. In Figure 6, a new potential service location is placed inside the catchment area. In this scenario we would calculate the livability metric again for the location and see how much the new service location affects the livability calculation of the location. When the same logic is expanded to the whole analysis set, we can, for example, find a location for a new service that has the highest total increase in utility.

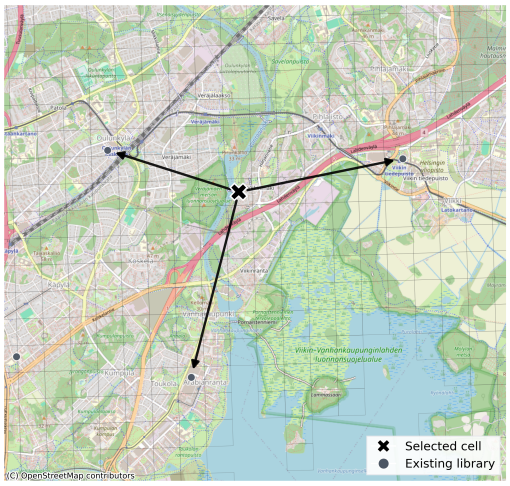


Figure 5: Accessibility of a selected grid cell, marked X. Arrows connect the cell to the existing libraries that contribute to its accessibility score.

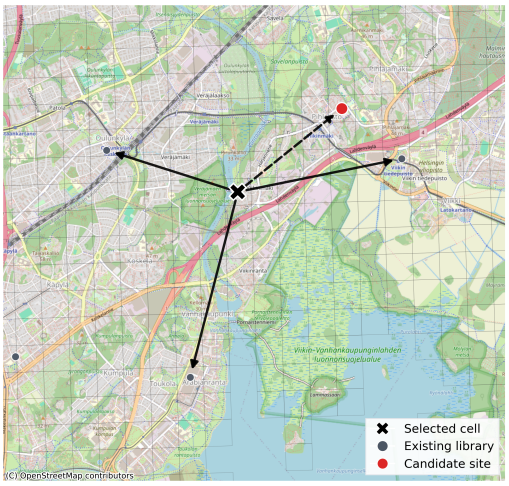


Figure 6: The same cell after adding a candidate site (red). The new site adds a further travel-time weight, raising the cell's accessibility score and its utility.

Let I denote the set of grid locations in Helsinki and R the set of profiles. We denote $\bar{J}_s \subseteq J \cup N$ as the set of active and candidate sites of the selected service type in a given scenario. For example, a scenario can be a set of the current network without any new additions of service s, or a set where we add a single service to the existing service network. Furthermore, let us introduce a binary decision vector $y \in \{0,1\}^{|\bar{J}_s|}$ that has elements y_j for each site $j \in \bar{J}_s$, where $y_j = 1$ if the service is active and 0 otherwise.

The individual utility of profile r at location i for service s is defined as a function of the decision vector y:

$$u_{irs}(y) = n_{rs}\, P_{ir}\, f\left(\sum_{j \in \bar{J}_s} w_{ijrs}\, y_j\right),$$

where n_{rs} is the necessity score of service s for profile r, P_{ir} is the population at location i belonging to profile r, and f is the diminishing value function defined above.

The total utility at location i is obtained by aggregating over all profiles,

$$U_{is}(y) = \sum_{r \in R} u_{irs}(y),$$

and the overall network's utility is the sum over all locations,

$$U_{\text{tots}}(y) = \sum_{i \in I} U_{is}(y).$$

The utility defined here should be interpreted as a need-weighted access value and utilities of different profiles are not directly comparable as well-being levels. Since the term $\sum_{i \in I} \sum_{r \in R} n_{rs} P_{ir}$ does not depend on the open sites, maximizing total utility is equivalent to minimizing the unmet need $\sum_{i \in I} \sum_{r \in R} n_{rs} P_{ir}(1 - f(a_{irs}))$.

Defining the utility as a function of the decision vector y allows us to first compute the utility of any given service network by setting $y_j = 1$ for all $j \in J$, and then to maximize the utility by adjusting the binary choices in y.

3.4 Optimization constraints

As we defined above, the total utility function $U_{\text{tots}}(y)$ aggregates all weighted travel time calculations for the whole grid of locations. The diminishing value function defined above naturally discourages service clustering. This means that the model does not gain much utility if it were to place a new service near another service. However, the utility function is still weighted by population as we want to take the population distribution into consideration. Even with diminishing returns, the model could heavily favor high density areas, leaving sparsely populated areas unserved. To solve this we introduce an equity constraint.

3.4.1 Equity constraint

Spatial equity is enforced by requiring a baseline level of accessibility: every populated grid cell i should be within a maximum allowable travel time $\tau_{\max}$ of an active service. We set $\tau_{\max} = 60$ minutes based on expert input from the City of Helsinki [21]. This threshold can be adjusted depending on the service type and priorities of the policymakers.

We formulate this as a soft constraint using a slack variable δ_i to prevent the model from becoming infeasible in scenarios where remote locations cannot physically be reached within $\tau_{\max}$. Let C_i denote the set of sites in $\bar{J}_s$ that can cover location i within the threshold:

$$C_i = \{j \in \bar{J}_s : d_{ij} \leq \tau_{\max}\}.$$

The constraint requires that at least one service in C_i is active, or the penalty variable δ_i is set to 1:

$$\sum_{j \in C_i} y_j + \delta_i \geq 1, \quad \forall i \in I \text{ where } \sum_{r \in R} P_{ir} > 0.$$

3.5 Optimization model

In this section we describe how the livability utility defined in Section 3.3 is formulated as an optimization model that chooses where services should be located.

We introduce a cell-specific penalty constant M_i to penalize uncovered population. The penalty makes leaving a populated cell uncovered cost more than any utility that the cell itself could contribute. This steers the model to cover populated areas of low accessibility before optimizing the utility score. We set

$$M_i = 1.1 \cdot \sum_{r \in R} n_{rs} P_{ir},$$

which is 10% higher than the maximum utility of cell i, reached when f saturates completely. We use a cell-specific constant instead of a single global big-M constant, because a global constant that dominates the total network utility makes the linear relaxation of the model loose and increases the solve time of the model significantly.

For the optimization model, there are two decision variables. These are $y_j \in \{0, 1\}$ for all $j \in \bar{J}_s$ and $\delta_i \in \{0, 1\}$ for all $i \in I$. The decision variable y_j indicates whether a service is active and δ_i activates the penalty for a location where the accessibility is too low. Furthermore, let q denote the number of service sites to be opened.

The complete optimization model maximizes total utility with diminishing returns and prioritizes locations where accessibility is low:

$$\begin{aligned}
\underset{y,\,\delta}{\text{maximize}} \quad & \sum_{i \in I} U_{is}(y) - \sum_{i \in I} M_i \delta_i \\
\text{subject to} \quad & \sum_{j \in N} y_j = q, \\
& y_j = 1, \quad \forall j \in J, \\
& \sum_{j \in C_i} y_j + \delta_i \geq 1, \quad \forall i \in I \text{ where } \textstyle\sum_{r \in R} P_{ir} > 0.
\end{aligned}$$

This formulation covers the service-addition scenarios: all existing services are kept active through the constraint $y_j = 1$ for $j \in J$, and the model selects q additional sites from the candidate set N. In the ground-up scenarios of Section 5, the network is built from scratch. The fixing constraint is dropped and the constraint on the number of selected sites is applied over all sites, so the model selects the full network of q sites freely:

$$\begin{aligned}
\underset{y,\,\delta}{\text{maximize}} \quad & \sum_{i \in I} U_{is}(y) - \sum_{i \in I} M_i \delta_i \\
\text{subject to} \quad & \sum_{j \in \bar{J}_s} y_j = q, \\
& \sum_{j \in C_i} y_j + \delta_i \geq 1, \quad \forall i \in I \text{ where } \textstyle\sum_{r \in R} P_{ir} > 0.
\end{aligned}$$

Both models are solved for a single service type s at a time. For a model that is generalized further to multiple service types, the objective should be summed over all service types, and the site-count and coverage constraints should be applied separately for each type.

3.5.1 Linearization of the objective

In order to solve the MILP problem, the objective function and the constraints have to be linear. All other equations fulfill this requirement except $f(a_{irs}) = 1 - e^{-\alpha\, a_{irs}}$. We linearize f using linear approximation.

For each cell-profile (i, r), we introduce auxiliary variables for the accessibility score $a_{irs} \in [0, \infty)$ and the linearized substitute for f as $\tilde{f}_{irs} \in [0, 1]$. The accessibility score is defined by the linear equality

$$a_{irs} = \sum_{j \in \bar{J}_s} w_{ijrs}\, y_j .$$

We define $b \in \{0, 0.25, 0.5, 1.0, 1.5\}$ as the breakpoints for a_{irs}. The maximum breakpoint is selected to be 1.5 as $f(1.5) \approx 0.99$, which we consider to be adequate accuracy for this thesis. For every breakpoint b, we calculate the tangent lines of f and add a constraint

$$\tilde{f}_{irs} \le f'(b) \cdot a_{irs} + \left(f(b) - f'(b) \cdot b\right),$$

where $f'(b) = \alpha\, e^{-\alpha\, b}$. Since f is concave and is being maximized, the solver naturally selects the tightest upper bound at each point. Figure 7 shows the shape that forms as a result. The theoretical maximum gap between f and $\tilde{f}$ is $\approx 5.4\%$. However, for the existing 38 libraries, this relaxation gives a gap of $\approx 2\%$. Furthermore, the reported utility values are recalculated with the exact f after optimization, so the gap does not affect the reported values. The approximation can in principle affect which sites are selected, but we did not quantify this effect.

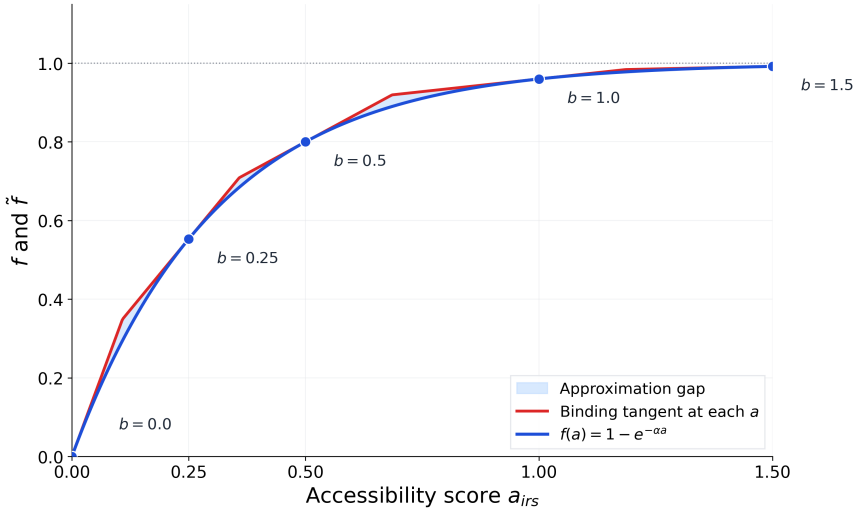


Figure 7: Tangent-line linearization of $f(a)$. The red line is the binding tangent at each a and forms the outer envelope of f. The shaded area is the approximation gap.

In the final model, we replace $f(a_{irs})$ with $\tilde{f}_{irs}$ in the objective. The total number of variables is $|\bar{J}_s|$ binary site variables, $|I|$ binary coverage slack variables, and

$2 \cdot |I| \cdot |R|$ continuous variables for the accessibility scores and linearized f-values. The constraint rows are generated only for cell-profile pairs with positive population, because pairs without population cannot contribute to the objective. The final model has five tangent inequalities and one accessibility constraint for each populated cell-profile pair, one count constraint, one coverage row for each populated cell, and, in the service-addition scenarios, $|J|$ fixing equalities.

4 Data and implementation

In this section we describe the data used in the thesis and the implementation pipeline that produces the results. The data sources are a mix of open and closed sources. In addition, we use expert opinion to weight the model according to the needs of the City of Helsinki.

For determining travel times in Helsinki, we use the Helsinki Region Travel Time Matrix 2018–2023 dataset (Version 2023.4) by Fink, Willberg, and Toivonen [16]. The dataset provides travel times between grid cells in the Helsinki region. The Helsinki region is divided into a grid of 250 m × 250 m cells with 13 132 cells in total. For each origin to destination cell pair, the grid reports travel times and distances with several different transport methods, including walking, cycling, public transport and by car. For the service locations we use the data from the City of Helsinki Map Service [6]. After filtering university libraries out from the dataset, we have 38 public libraries that form the baseline service network. Furthermore, we use a closed-source dataset from the City of Helsinki [7], which contains the population predictions of Helsinki for different age groups mapped to the same grid as [16]. Table 5 summarizes the data sources used in this thesis.

Table 5: Data sources used in this thesis. Each row lists a dataset, its provider, whether it is openly available, and the content used in this thesis.

Dataset	Provider	Access	Content
Travel Time Matrix 2023	Digital Geography Lab	Open	OD travel times, 13 132 cells
Helsinki Map Service	City of Helsinki	Open	Service locations, postal codes
Closed source population prediction dataset	City of Helsinki	Closed	Population predictions

4.1 Persona profiles

4.1.1 Finnish household segmentation

Statistics Finland has collected statistics that contain information on Finnish households [46]. In their dataset households are segmented by age group, household composition and income decile. The largest segments, in order of size, are retirees living alone, established singles aged 35 to 64, families with children, and young singles under 35. One key feature that separates these segments is life stage.

We use age as the basis of our segmentation. As discussed in Section 2.2.3, stated preferences are an unreliable basis for segmentation, and observed demographic characteristics can provide a more robust approach for the categorization. Age is a good candidate for the segmentation, and the population prediction dataset [7] restricts the analysis to age groups, as age is its lowest granularity with no further segmentation

available. Based on this we create the persona profiles defined in Table 6. The profiles focus on the adult population, ages 20 and older.

Table 6: Persona profiles. Each profile corresponds to one age group of the adult population.

Profile	Age category
Young adults	20–29
Adults	30–49
Middle-aged adults	50–64
Older adults	65+

In addition, we define the necessity score $n_{rs} \in \{0, 1, 2, 3, 4, 5\}$ for each profile r and service s as

$$n_{rs} = \begin{cases} 0 & \text{no need for the service} \\ 1 & \text{very low importance} \\ 2 & \text{low importance} \\ 3 & \text{moderately important} \\ 4 & \text{important} \\ 5 & \text{very important.} \end{cases}$$

Table 7 shows the selected necessity scores for a library across the four persona profiles. These scores are selected based on expert discussion with City of Helsinki representatives [21]. The main logic behind the necessity scores is to consider libraries as public investments in literacy, knowledge and culture, and to allocate the resource based on the unit of improvement per profile.

Table 7: Necessity scores n_{rs} for libraries, by profile. A higher score means that the library is more important for the profile.

Profile	$n_{r\text{library}}$
Young adults	4
Adults	5
Middle-aged adults	3
Older adults	3

The necessity scores become increasingly important for other generalizations of the model defined in Section 3. For example, school locations or hospital locations have inherently different necessity scores for the profiles defined here.

4.1.2 Acceptable travel time measures

Acceptable travel time thresholds h_{rs} are selected based on the expected behavior of the defined profile and the expert opinion of City of Helsinki representatives [21].

Each value represents a public transport travel time (midday, average walking speed to stops at the grid cells) where the Gaussian weight drops to 0.1. The full equation is defined in Section 3.2.1. Beyond this point, the library is considered to add only marginal value for that profile. Table 8 shows the values used as the maximum travel times per profile.

Table 8: Threshold travel times $h_{r,\text{library}}$ (minutes), by profile. The threshold is the public transport travel time at which the decay weight drops to 0.1. After this point the library adds only marginal value for the profile.

Profile	$h_{r,\textbf{library}}$ **(min)**
Young adults	20
Adults	15
Middle-aged adults	25
Older adults	15

A common reference point for service accessibility is the 15-minute city principle. The thresholds in Table 8 start from this reference and are refined based on expert feedback [21] to follow the behavioral preferences of the defined age groups.

4.2 Travel times

As mentioned, we use the Helsinki Region Travel Time Matrix to calculate the travel times. The dataset contains travel times from every cell to every other cell with multiple different transport methods in Helsinki metropolitan area (Espoo, Vantaa, Helsinki, Kauniainen). To make the calculations faster, we filter the dataset to include only the City of Helsinki. The original dataset has 172 million rows of data with 6 dimensions of travel times and 13 132 grid cells of 250 m × 250 m each. The filtering uses official postal code boundaries (00100–00990) obtained from Helsinki City's geographic data service [6]. After filtering, the analysis grid contains 3 868 cells and 15 million rows of data. The final dataset of the travel time matrix has travel times for walking, cycling, car and public transit. Figure 8 shows the resulting grid overlaid on a map of Helsinki.

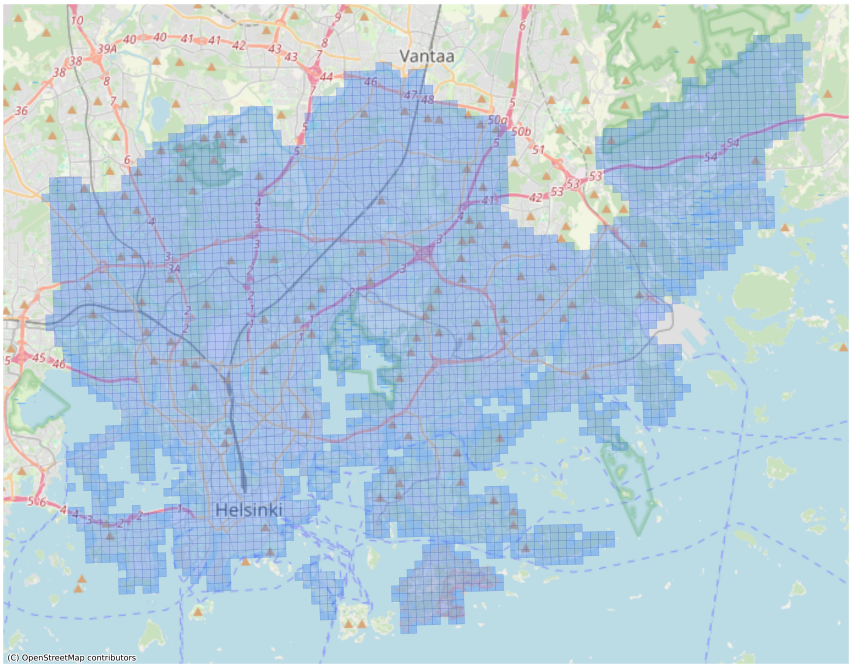


Figure 8: 250 m × 250 m analysis grid covering Helsinki (3 783 cells). The grid covers the municipality of Helsinki after removing cells with insufficient travel-time data.

The grid contains some cells that have incomplete travel-time data. A cell is dropped if there are fewer than 100 valid travel times. In total, 85 cells are removed, reducing the number of cells from 3 868 to 3 783. If these cells were kept the equity constraint would force the optimizer to select sites near these cells. The resident count of the dropped cells is 255, so no significant population loss is introduced.

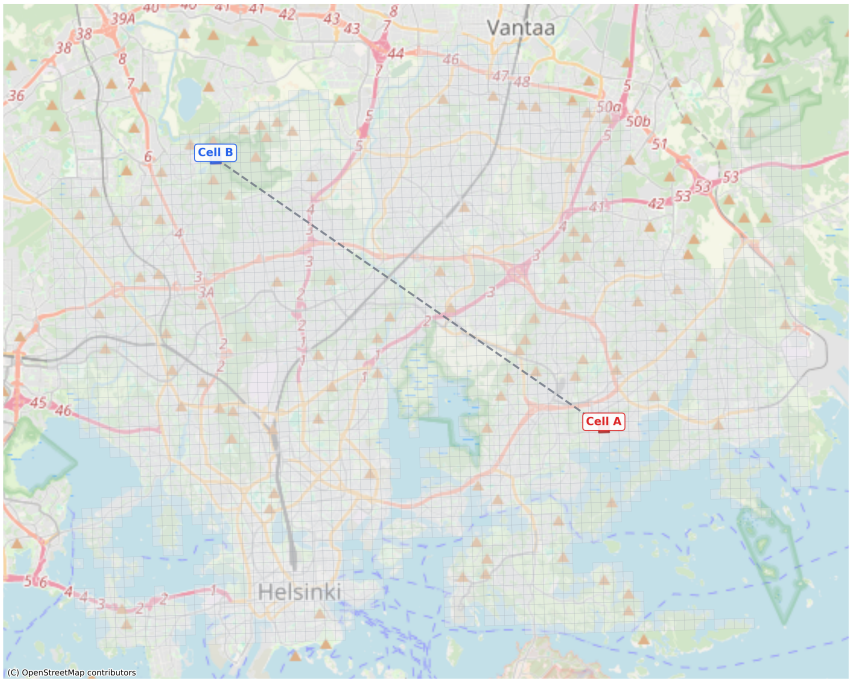


Figure 9: Two example grid cells for the travel-time calculation. The highlighted cells are the origin and destination used in the travel time comparison of Table 9.

Figure 9 shows an example of two grid cells used for travel time calculation. Table 9 shows the travel times between these cells for different transport modes. As the table illustrates, travel times vary significantly by transport mode. For the selected cell pair, travel by car is the fastest option at 34 minutes, followed by cycling at 58 minutes and public transit at 62 minutes. Walking takes nearly three hours due to the distance between the cells. In contrast, if we were to choose a cell pair where both of the cells are close to the city center, walking would be the fastest mode of transport.

Table 9: Travel times between two example cells, by transport mode. The travel times are from the Helsinki Region Travel Time Matrix for the cell pair shown in Figure 9.

Transport Mode	**Travel Time (min)**
Walking (average speed)	179
Cycling (average speed)	58
Car (midday traffic)	34
Public transit (midday)	62

All the calculations use the travel time estimate with public transport. The travel time from one cell to another represents the fastest time that a resident can achieve with average midday traffic and average walking distance for an adult. We select public transport, because it is the most accessible form of vehicle transport within the city.

4.3 Population prediction and service locations

The population dataset used for this thesis is an aggregated prediction of the population distribution on a 250 m grid [45] over the City of Helsinki. The dataset is acquired from the City of Helsinki and is closed source [7]. For all the results in this thesis, we use the population prediction for the year 2036. Each profile used in the calculations is assigned a value for every cell in the dataset. The population prediction data is closed source, so further description of the dataset is restricted.

4.4 Implementation

Figure 10 illustrates the end-to-end flow of the computations. Each phase feeds its output into the next, beginning from raw open data and concluding with the optimized service network and equity analysis.

The model is implemented in Python. The key libraries used are GeoPandas for spatial data handling, NumPy and SciPy for numerical computation, and Gurobi for solving the mixed-integer linear program. In addition, the solver is configured with a time limit of 1000 seconds per scenario and a relative gap of 1%.

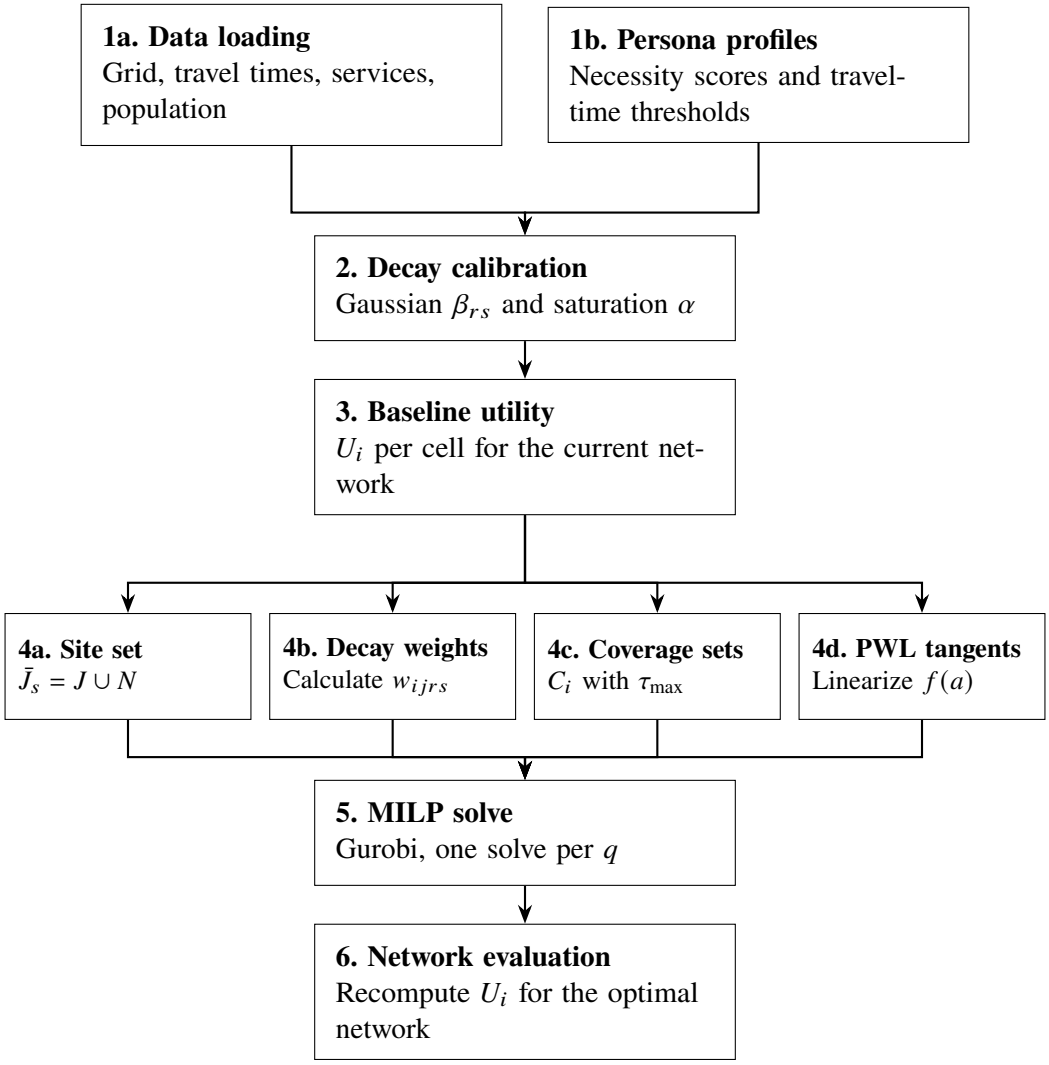


Figure 10: Implementation pipeline. Steps 1-3 load the data and compute the baseline utility, step 4 assembles the MILP inputs, and steps 5-6 solve the model and evaluate the resulting network.

The data processing starts with step 1 that loads the data into the model. We use the statistical grid from Statistics Finland [45], and all spatial values are mapped to this grid. The following data is processed and used in the latter steps: travel time from every cell to all other cells, current library network information, the persona profiles with their necessity scores and threshold travel times, and a population dictionary that contains the population information for every cell. In step 2, we precompute the decay parameters β_{rs} from the threshold travel times and the saturation parameter α from a^*. The formulas are derived in Sections 3.2.1 and 3.2.2. Step 3 is the baseline calculation that is used to benchmark the optimization and to calibrate the color gradient for the map visualizations. This step is computed only once as it is used as the reference point for the scenario computations.

Step 4 assembles the MILP inputs: the site set, the decay weights, the coverage sets, and the piecewise-linear tangents. The model is constructed over a set $\bar{J}_s = J \cup N$

where J is the 38 existing libraries and N is the set of candidate cells. Then we pair each non-zero population cell with all decay weights w_{ijrs} that serve as the coefficients in the accessibility constraints. In addition, we drop all coefficients below 0.01 to save compute time. Then, again for each populated cell, we build $C_i = \{j \in \bar{J}_s : d_{ij} \leq \tau_{\max}\}$, that is the set of sites that can serve i within the maximum travel time $\tau_{\max}$. Finally, as described in Section 3.5.1, we linearize the $f(a)$ in order to make the model linearly solvable.

After step 4 prepares the model, we can run the optimization for any number of libraries q. For a scenario where we build the library network from the ground up, the model has 37 830 variables and 57 918 constraints. With the Gurobi optimizer, the whole pipeline for a single q runs in approximately 300 seconds using a laptop (Lenovo Thinkpad E14). Step 6 evaluates the solved network and calculates key figures and creates graphs for analysis.

Figure 11 illustrates the matrix that is created for the optimizer. Each pixel's darkness corresponds to the non-zero values in that location of the matrix. The top block is the single count constraint that fixes the number of libraries at exactly q. It is not visible as it is only a single row, but is marked by the top line. Below is the block of accessibility constraints, where the number of non-zero coefficients is restricted by the decay weight minimum of 0.01. The lower blocks contain the linearized diminishing returns constraints, and the bottom block has the equity constraints, where the membership of each row is determined by the maximum travel time $\tau_{\max}$.

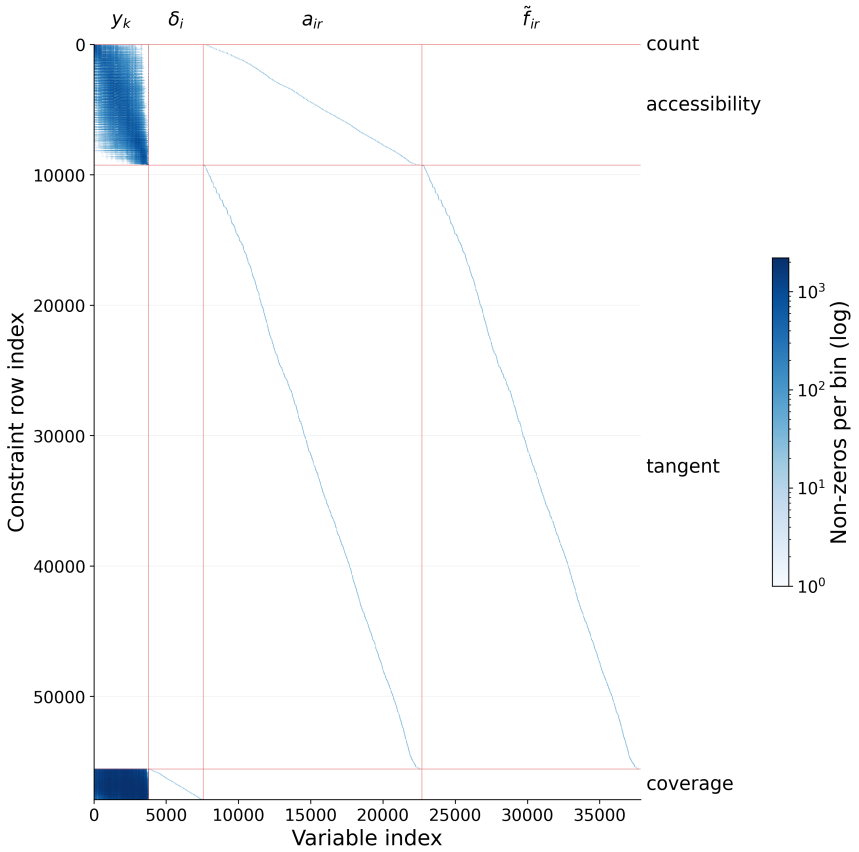


Figure 11: Structure of the ground-up MILP constraint matrix at $\tau_{\max} = 60$ min. Each dark pixel is a non-zero coefficient. Rows are constraints and columns are variables of the model. The column blocks are the site selection variables y_k, the coverage slack variables δ_i, the accessibility scores a_{ir}, and the linearized utility values $\tilde{f}_{ir}$. The row blocks are the count constraint that fixes the number of libraries, the accessibility constraints, the tangent inequalities that linearize the diminishing returns, and the coverage constraints.

5 Results

In this section we present the results of the utility calculations for a baseline scenario of 38 libraries, and different scenarios for optimizing the service network. We run the MILP optimization for 1 to 38 libraries to determine what level of utility is achieved with networks of different sizes. Furthermore, we determine the optimal location of a new library, and we select the library whose removal reduces the total utility the least. We use the population prediction for the year 2036 for all the calculations in this section. The main reason for using the predicted population density is to help the City of Helsinki assess the current service level of the library network for proactive decision-making. Since the analysis considers a single service type (libraries), we drop the service index s from the notation in this section.

We aggregate the variables of Section 3 by summing over cells or profiles as indicated. The variable P_r is the total population of profile r, P_i the population of cell i, and P the total adult population. Similarly, U_r is the total utility of profile r over the network and U_i the total utility of cell i. Per-capita utility is $u_{ir}(y)/P_{ir}$ for a single profile and U_i/P_i for all profiles combined. The mean and median per-capita utility of profile r are denoted by $\bar{u}_r^{\mathrm{pc}}$ and $\tilde{u}_r^{\mathrm{pc}}$.

A resident with profile r is defined as well served if the per-capita utility of their cell, $u_{ir}(y)/P_{ir}$, exceeds the comfortable-access threshold $u_r^{\mathrm{threshold}}$ defined in Section 5.1.2. Each resident is judged against the threshold of their own profile, so the combined well-served count is the sum of the per-profile counts. The population-weighted combined threshold $\bar{u}^{\mathrm{threshold}}$ is used as the reference level for the combined per-capita maps and for the share of cells above the threshold.

5.1 Baseline

5.1.1 Existing-network coverage

Before the optimization, it is useful to understand the current accessibility level and how much of the population is within reach of a library at different travel times. Figure 12 shows the cumulative share of Helsinki's adult population (ages 20+) that has a library within a given travel time.

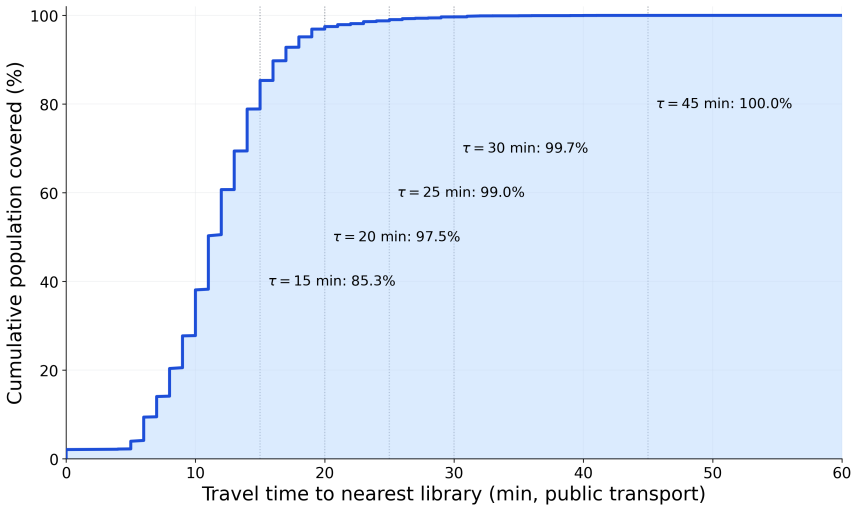


Figure 12: Cumulative share of adult population within travel time of the nearest library, 2036. The curve shows the share of the adult (20+) population in the 2036 prediction whose public transport travel time to the nearest of the 38 existing libraries.

The figure shows that with $\tau = 20$ min, only about 2.5% of the population falls outside the catchment. At $\tau = 25$ min, 99.0% of the adult population can reach their nearest library. At $\tau = 45$ min, the entire population can reach their nearest library. In Section 3.4.1 we defined an equity constraint for the optimization model. The constraint activates when a service network is sparse. The maximum travel time defined was $\tau_{\max} = 60$ min, and every location in Helsinki satisfies this threshold for the current 38 libraries.

5.1.2 Baseline utility

We first evaluate the current Helsinki library network by setting $y_j = 1$ for all $j \in J$ and computing the utility for all grid cells. The per-capita utility figures use a color gradient centered on the theoretical threshold for comfortable access to a service. The reference level of comfortable access that determines the color gradient, together with other example levels, is shown in Table 10. The reference level is highlighted as the bolded row. For other figures we use a linearly scaling gradient from zero to 95th-percentile per-cell utility. Note that the per-capita figures and the total utility figures have different interpretations and should not be compared directly.

Table 10: Reference accessibility levels a_{ir} and corresponding $f(a)$. Each row gives an accessibility score a, the corresponding value $f(a)$, and its physical interpretation for the library network. The bolded row is the reference level used as the center of the per-capita color gradient.

Accessibility a	$f(a)$	Physical interpretation
0.1	0.275	One library exactly at the profile's threshold travel time h_r; minimum-acceptable access.
0.5	**0.80**	**Comfortable access, used as the definition of well served and as the reference level for the color gradient.**
1.0	0.96	Defines the satisfied service level, i.e. a single library very close by.
$\to \infty$	$\to 1.0$	Saturation as the number of libraries approaches infinity.

The center of the color gradient (the dashed line in the colorbar) is the utility that a resident of profile r receives when the cell reaches $a = 0.5$:

$$u_r^{\text{threshold}} = n_r \cdot f(0.5) = n_r\left(1 - e^{-\alpha \cdot 0.5}\right) \approx 0.80\, n_r.$$

For the combined per-capita utility, we define the reference point as the population-weighted average:

$$\bar{u}^{\text{threshold}} = \sum_r \frac{P_r}{P} u_r^{\text{threshold}}.$$

The per-capita analysis shows network equity by measuring the utility the average resident receives in each cell. The per-capita utility distribution is shown in Figure 13. The figure can be interpreted as what a typical resident in that location might experience. As expected, the highest accessibility is around the library locations. However, there are many areas where the threshold is not met. For example, in the Östersundom district in eastern Helsinki, where the nearest library is in Vuosaari, the per-capita utility is near zero.

Table 11 summarizes the baseline network. For every profile we show the total utility U_r, the mean per-capita utility $\bar{u}_r^{\text{pc}}$, the median per-capita utility $\tilde{u}_r^{\text{pc}}$, and the comfortable-access threshold $u_r^{\text{threshold}}$. The remaining rows show the share of populated cells above the threshold, the population P_r, and the number of residents above and below the threshold.

There are three main findings from Table 11. First, the well-served shares differ widely across profiles. As noted in Section 3.3, utilities are scaled by the necessity score n_r, so mean utilities are not comparable across profiles. The well-served share does not have this effect, so these values can be compared directly. The well-served requirement is $a_{ir} > 0.5$ for every profile, so the share depends only on where the profile lives and on its threshold travel time h_r. Second, the Adults and Older adults

profiles have the lowest well-served shares of 42.0% and 38.8%, respectively. Both profiles have the strictest threshold travel time of 15 minutes, so their accessibility decays the fastest. Given the assumed travel-time tolerances, the current network serves these two profiles the worst. As the optimization model prioritizes profiles with higher population densities and necessity coefficients, these profiles can be expected to gain the most utility from the optimization. Third, around 34.3% of all populated cells and 60.8% of the population are well served. One reason for this could be that libraries tend to be located in areas where the population density is higher, which raises the well-served population share more than the well-served cell share.

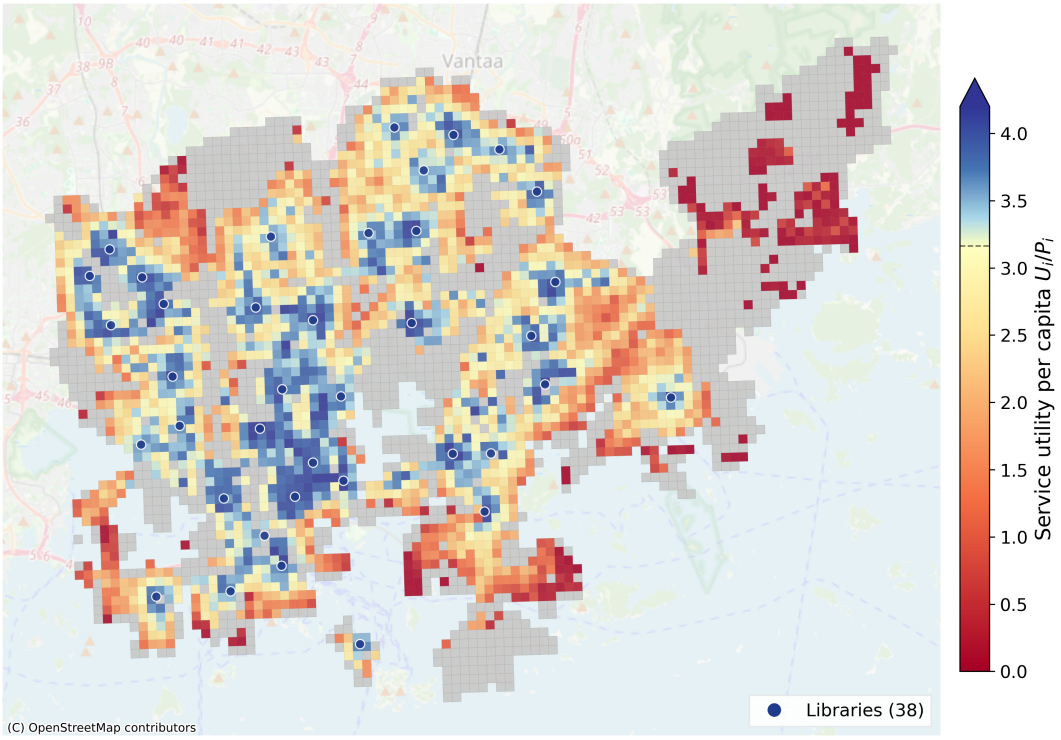


Figure 13: Baseline utility per capita U_i/P_i, all profiles. Each cell is colored by the average utility a resident of the cell receives from the 38 existing libraries, using the 2036 population prediction. The gradient is centered on the population-weighted comfortable-access threshold, shown as a dashed line in the colorbar. Warm colors fall below the threshold and cool colors above it.

	Young 20–29	Adults 30–49	Middle-aged 50–64	Older 65+	**Adults 20+ (combined)**
U_r	407,771	826,812	400,801	278,184	**1,913,568**
$\bar{u}_r^{\text{pc}}$	3.62	3.33	2.92	1.97	**2.99**
$\tilde{u}_r^{\text{pc}}$	3.75	2.96	2.99	1.79	**2.77**
$u_r^{\text{threshold}}$	3.20	4.00	2.40	2.40	**3.16**
cells above threshold	71.5 %	27.3 %	88.8 %	27.5 %	**34.3 %**
P_r	112,765	248,170	137,080	141,199	**639,213**
well served	96,445	104,235	132,917	54,842	**388,439**
(% of P_r)	(85.5 %)	(42.0 %)	(97.0 %)	(38.8 %)	**(60.8 %)**
not well served	16,320	143,935	4,162	86,357	**250,774**
(% of P_r)	(14.5 %)	(58.0 %)	(3.0 %)	(61.2 %)	**(39.2 %)**

Table 11: Baseline utility summary statistics (2036, 38 libraries). Rows: total utility, mean and median per-capita utility, the comfortable-access threshold, the share of populated cells above it, the population, and the residents above and below their own profile's threshold. The combined well-served count is the sum of the per-profile counts.

Figure 14 shows the combined baseline utility across all persona profiles. Some locations are clearly better served than others. The figure shows the absolute utility per cell, meaning that every cell is weighted with the population count per profile. The highest utility cells cluster around dense central areas (Kallio, Vallila, Kamppi, Pasila, Töölö) and sub-centers around libraries (Itäkeskus, Vuosaari, Malmi, Kannelmäki). For example, Viikki, even with a library, does not receive as much utility as Vuosaari in the southeast. The baseline utility figure complements the per-capita figure. It also shows that some libraries contribute much less to the total utility than others.

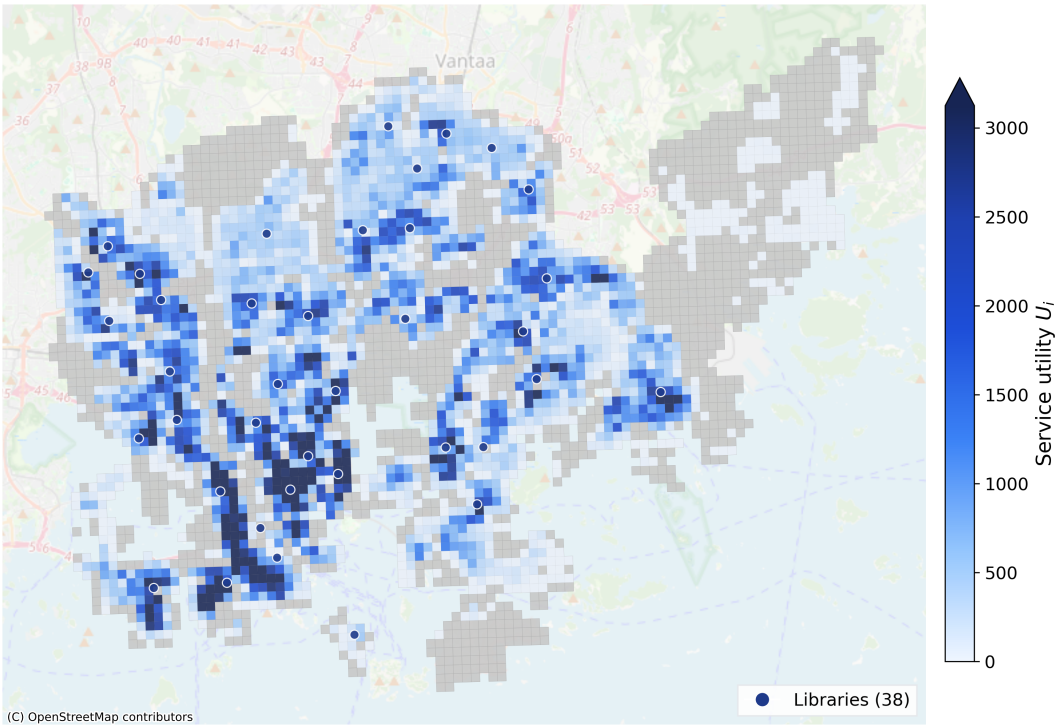


Figure 14: Baseline utility U_i for the current 38-library network, all profiles. Each populated cell is colored by its total utility U_i for the 2036 population, so the map weights accessibility by population count. The scale is linear from zero to the 95th percentile of per-cell utility and cells above the cap share the darkest shade.

Since there are clear differences among the profiles, we visualize each profile separately. Figure 15 shows that the per-capita utility distributions differ clearly between the profiles. The figure confirms the observed pattern from Table 11, where Middle-aged adults are much better served than the other profiles.

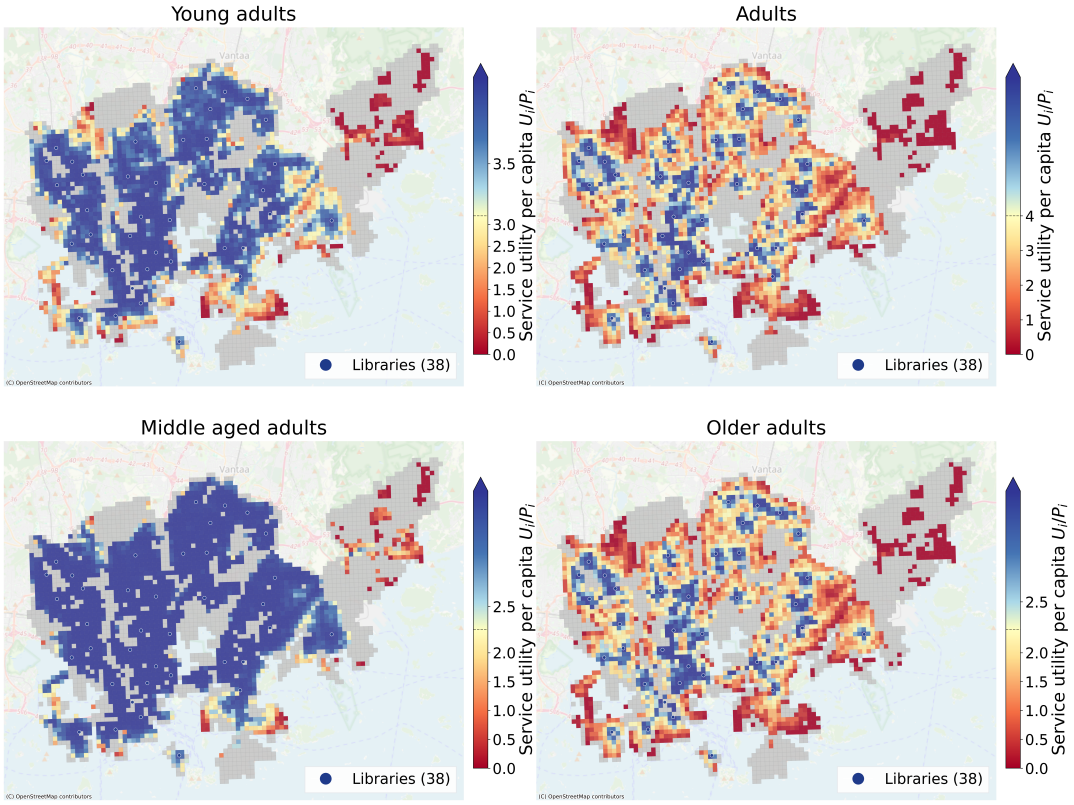


Figure 15: Baseline per-capita utility per persona profile. Each panel shows the per-capita utility of one profile, with the color gradient centered on that profile's own comfortable-access threshold. Blue cells are above the threshold and red cells below it.

5.2 Optimal service locations

We evaluate four scenarios to assess and improve the library network in Helsinki. First, we evaluate a scenario where we add a new library. The goal is to define a location where the total utility would increase the most. Second, we evaluate the removal of a single service location. Third, we build an optimal service network from the ground up. We answer how many libraries are needed to achieve 50%, 70% and 100% of the baseline utility of 38 libraries. To answer these, we solve the optimization model for increasing values of q and report the smallest number of libraries that reaches the desired share of the baseline utility. Finally, we compare the optimal service network of 38 libraries to the baseline.

5.2.1 Adding one service

In this scenario we fix all the existing libraries and let the model select one new service location. We do this by setting $y_j = 1$ for all $j \in J$ and $q = 1$, which means the optimization model selects exactly one site from the candidate set N. Furthermore, we drop the equity coverage constraint, because the baseline network already satisfies

the $\tau_{\max} = 60$ min limit for every populated cell.

Figure 16 illustrates the results. As the 38 existing libraries already cover Helsinki well, the utility gain from one additional service is small. The total utility increases by approximately 1.5%. The model places the new service close to the city center, where the population density and the utility gain are the highest. Figure 17 shows the location in more detail. The nearest address of the new service location is Merimiehenkatu 23, Punavuori (60.16099°N, 24.93920°E).

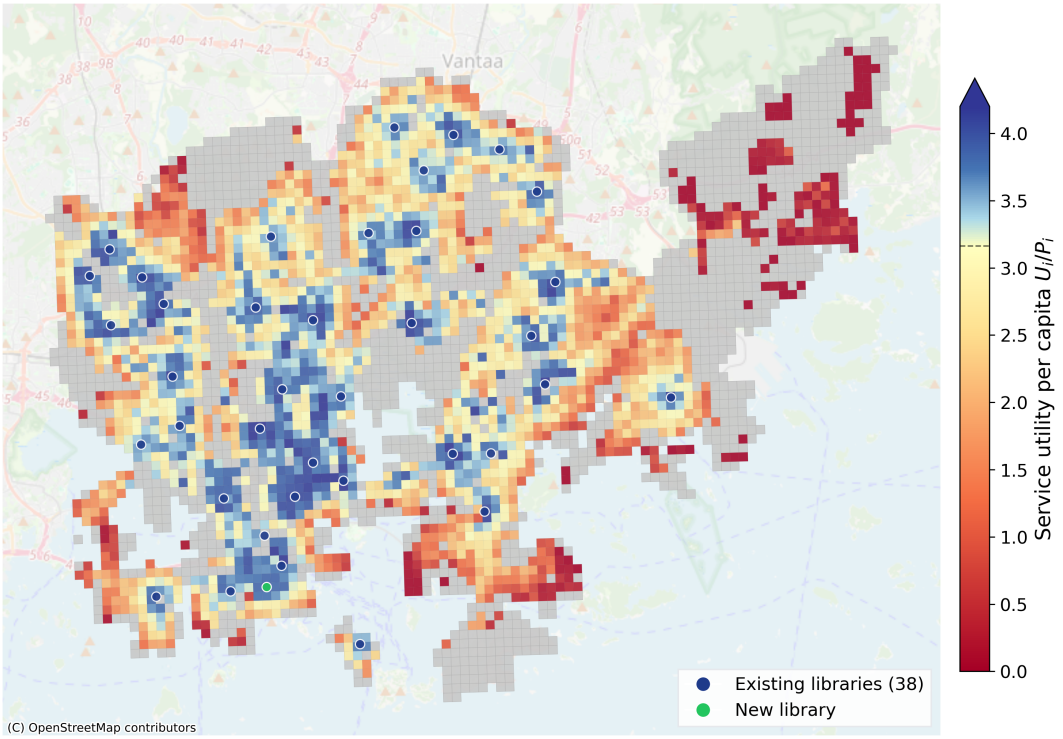


Figure 16: Baseline per-capita utility U_i/P_i with the proposed new Punavuori site added (all profiles). Same coloring as the baseline per-capita map, computed with the proposed site added to the 38 existing libraries (2036 population).



Figure 17: Zoomed view of the proposed new site in Punavuori. The map shows the immediate surroundings of the selected cell. The nearest address is Merimiehenkatu 23 (60.16099°N, 24.93920°E).

Table 12 summarizes the key figures and compares the new values to the baseline. With the single location addition, approximately 12 600 residents move above the threshold of comfortable access, mostly near the newly added location. However, the share of well-served cells increases by only 0.6 percentage points. This can also be explained by the high population density in central Helsinki compared to other districts.

	Young 20–29	Adults 30–49	Middle-aged 50–64	Older 65+	**Adults 20+ (combined)**	Δ vs. baseline
U_r	411,078	844,239	401,542	284,864	**1,941,723**	+28,156
(% of U_r^{base})	(100.8 %)	(102.1 %)	(100.2 %)	(102.4 %)	**(101.5 %)**	(+1.5 %)
$\bar{u}_r^{\text{pc}}$	3.65	3.40	2.93	2.02	**3.04**	+0.04
$\tilde{u}_r^{\text{pc}}$	3.76	3.00	2.99	1.82	**2.78**	+0.01
$u_r^{\text{threshold}}$	3.20	4.00	2.40	2.40	**3.16**	0.00
cells above threshold	72.0 %	28.2 %	89.0 %	28.4 %	**34.9 %**	+0.6 pp
P_r	112,765	248,170	137,080	141,199	**639,213**	0
well served	97,775	111,292	132,946	59,023	**401,037**	+12,598
(% of P_r)	(86.7 %)	(44.8 %)	(97.0 %)	(41.8 %)	**(62.7 %)**	(+2.0 pp)
not well served	14,990	136,878	4,133	82,175	**238,176**	-12,598
(% of P_r)	(13.3 %)	(55.2 %)	(3.0 %)	(58.2 %)	**(37.3 %)**	(-2.0 pp)

Table 12: Utility summary, scenario: add one library (Punavuori). Rows: total utility, mean and median per-capita utility, the comfortable-access threshold, the share of populated cells above it, the population, and the residents above and below their own profile's threshold. The combined well-served count is the sum of the per-profile counts. The last column shows the change compared to the 38-library baseline.

Figure 18 visualizes the new service network with the new location highlighted in red. No significant changes can be seen compared to the baseline figure.

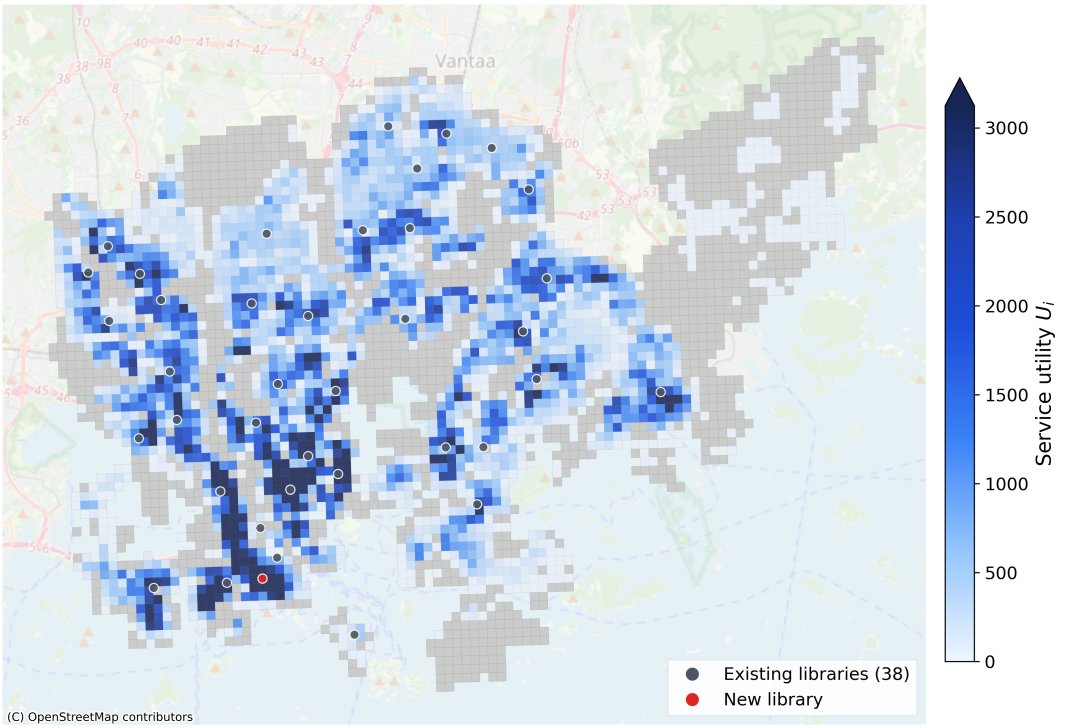


Figure 18: Optimal addition of one library ($q = 1$, fix existing). Total utility U_i per cell after the addition, on the baseline color scale. Gray circles are the 38 existing libraries and the red circle is the added site.

5.2.2 Removing one service

In this subsection we determine which library would cause the smallest loss in total utility over the grid. No MILP solve is required, as we can enumerate all libraries and recompute the utility with each one removed.

Figure 19 shows the result, with the selected library highlighted with a green circle. Table 13 summarizes the results of removing the service location. The least-critical library to remove is Suomenlinna. Removing the library would lower the total utility $U_{\text{tot}}^{\text{baseline}}$ from 1 913 568 to 1 911 895, a decrease of 0.1% compared to the baseline. The results are intuitive as the library meaningfully serves only the residents on the island. Furthermore, 418 people would move below the threshold of comfortable access. Figure 20 shows the total utility map with the remaining 37 libraries. As mentioned, the change in total utility is small, and the only visible change in the figure is the lower utility in the area of Suomenlinna. For comparison, the most critical library to remove would be the Vuosaari library. Removing it would decrease the total utility by 3.5%. The library serves a large area with no substitute libraries nearby, whereas in denser areas the surrounding libraries compensate for the removed one.

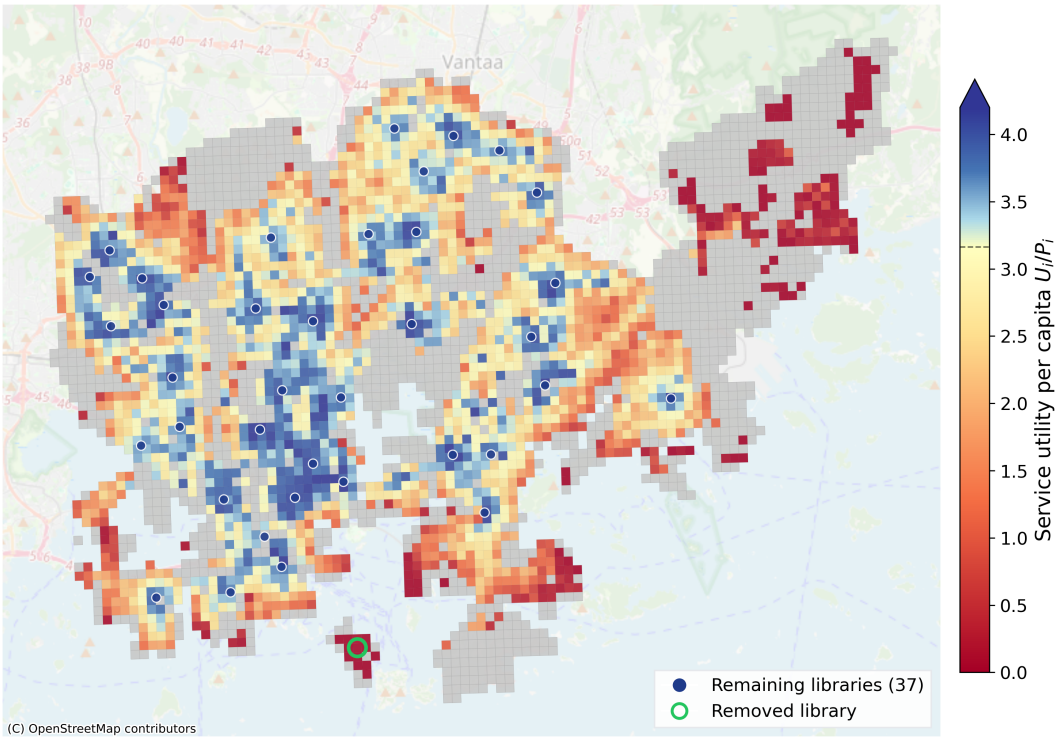


Figure 19: Baseline per-capita utility U_i/P_i with the Suomenlinna library removed (all profiles). Same coloring as the baseline per-capita map, computed with the Suomenlinna library removed (2036 population).

	Young 20–29	Adults 30–49	Middle-aged 50–64	Older 65+	**Adults 20+ (combined)**	Δ vs. baseline
U_r	407,511	825,954	400,434	277,996	**1,911,895**	-1,673
(% of U_r^{base})	(99.9 %)	(99.9 %)	(99.9 %)	(99.9 %)	**(99.9 %)**	(-0.1 %)
$\bar{u}_r^{\text{pc}}$	3.61	3.33	2.92	1.97	**2.99**	-0.00
$\tilde{u}_r^{\text{pc}}$	3.75	2.95	2.99	1.77	**2.77**	-0.00
$u_r^{\text{threshold}}$	3.20	4.00	2.40	2.40	**3.16**	0.00
cells above threshold	71.3 %	27.1 %	88.4 %	27.3 %	**34.1 %**	-0.2 pp
P_r	112,765	248,170	137,080	141,199	**639,213**	0
well served	96,391	104,056	132,794	54,780	**388,021**	-418
(% of P_r)	(85.5 %)	(41.9 %)	(96.9 %)	(38.8 %)	**(60.7 %)**	(-0.1 pp)
not well served	16,374	144,114	4,286	86,419	**251,192**	+418
(% of P_r)	(14.5 %)	(58.1 %)	(3.1 %)	(61.2 %)	**(39.3 %)**	(+0.1 pp)

Table 13: Utility summary, scenario: remove the Suomenlinna library. Rows: total utility, mean and median per-capita utility, the comfortable-access threshold, the share of populated cells above it, the population, and the residents above and below their own profile's threshold. The combined well-served count is the sum of the per-profile counts. The last column shows the change compared to the 38-library baseline.

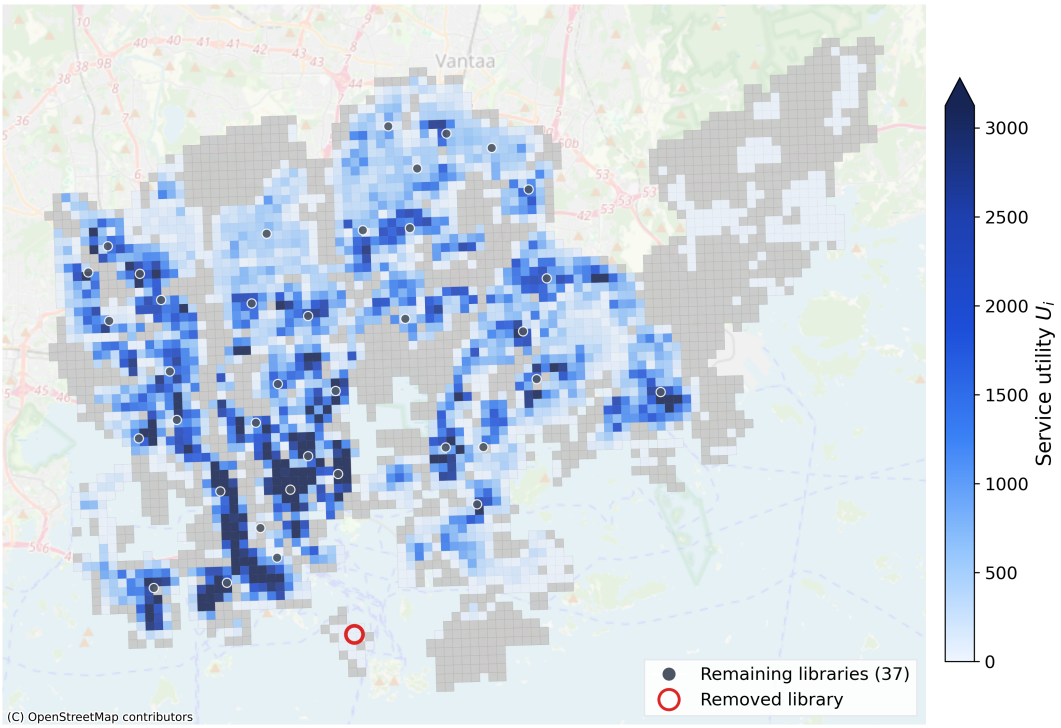


Figure 20: The least-critical library to remove is the Suomenlinna library. Total utility U_i with the remaining 37 libraries (gray circles), on the baseline color scale. The red ring marks the removed library.

5.2.3 50% of baseline utility ($q = 5$)

In this subsection we determine how many libraries are needed to reach at least 50% of the total utility of the baseline of 38 libraries. With optimal service locations, 5 libraries reach 51.7% of the baseline utility. Figure 21 shows the per-capita view of the 5 service locations. It is clear that far fewer cells reach the comfortable-access threshold. Table 14 shows that only 4.3% of the cells and 17.3% of the population are above the comfortable-access threshold. Moreover, Figure 22 shows that the network is spread evenly across the grid.

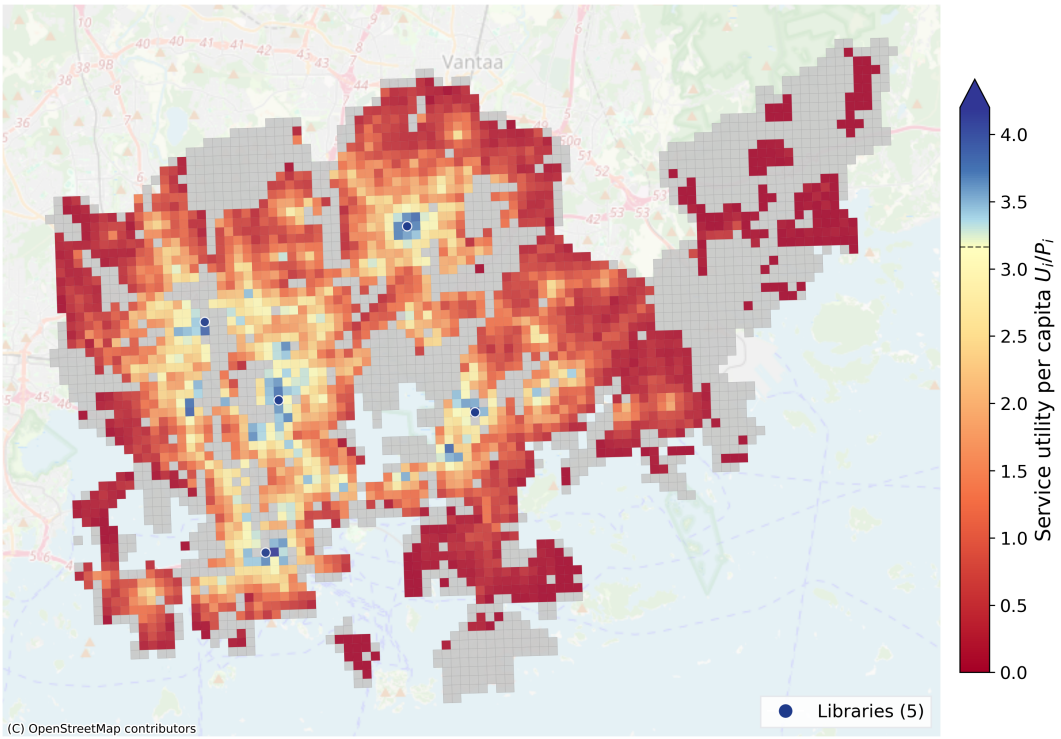


Figure 21: Per-capita utility U_i/P_i for the ground-up optimal network at $q = 5$ (all profiles). The network is optimized with $\tau_{\max} = 60$ min and the 2036 population prediction. The color gradient is centered on the comfortable-access threshold.

	Young 20–29	Adults 30–49	Middle-aged 50–64	Older 65+	**Adults 20+ (combined)**	Δ vs. baseline
U_r	253,771	325,157	307,295	102,270	**988,492**	-925,075
(% of U_r^{base})	(62.2 %)	(39.3 %)	(76.7 %)	(36.8 %)	**(51.7 %)**	(-48.3 %)
$\bar{u}_r^{\text{pc}}$	2.25	1.31	2.24	0.72	**1.55**	-1.45
$\tilde{u}_r^{\text{pc}}$	1.57	0.38	2.14	0.23	**0.97**	-1.80
$u_r^{\text{threshold}}$	3.20	4.00	2.40	2.40	**3.16**	0.00
cells above threshold	13.8 %	2.4 %	40.8 %	2.3 %	**4.3 %**	-30.0 pp
P_r	112,765	248,170	137,080	141,199	**639,213**	0
well served	26,354	5,943	75,018	3,127	**110,442**	-277,997
(% of P_r)	(23.4 %)	(2.4 %)	(54.7 %)	(2.2 %)	**(17.3 %)**	(-43.5 pp)
not well served	86,411	242,227	62,062	138,071	**528,771**	+277,997
(% of P_r)	(76.6 %)	(97.6 %)	(45.3 %)	(97.8 %)	**(82.7 %)**	(+43.5 pp)

Table 14: Utility summary, scenario: ground-up optimal network reaching ~ 50% of baseline utility ($q = 5$). Rows: total utility, mean and median per-capita utility, the comfortable-access threshold, the share of populated cells above it, the population, and the residents above and below their own profile's threshold. The combined well-served count is the sum of the per-profile counts. The last column shows the change compared to the 38-library baseline.

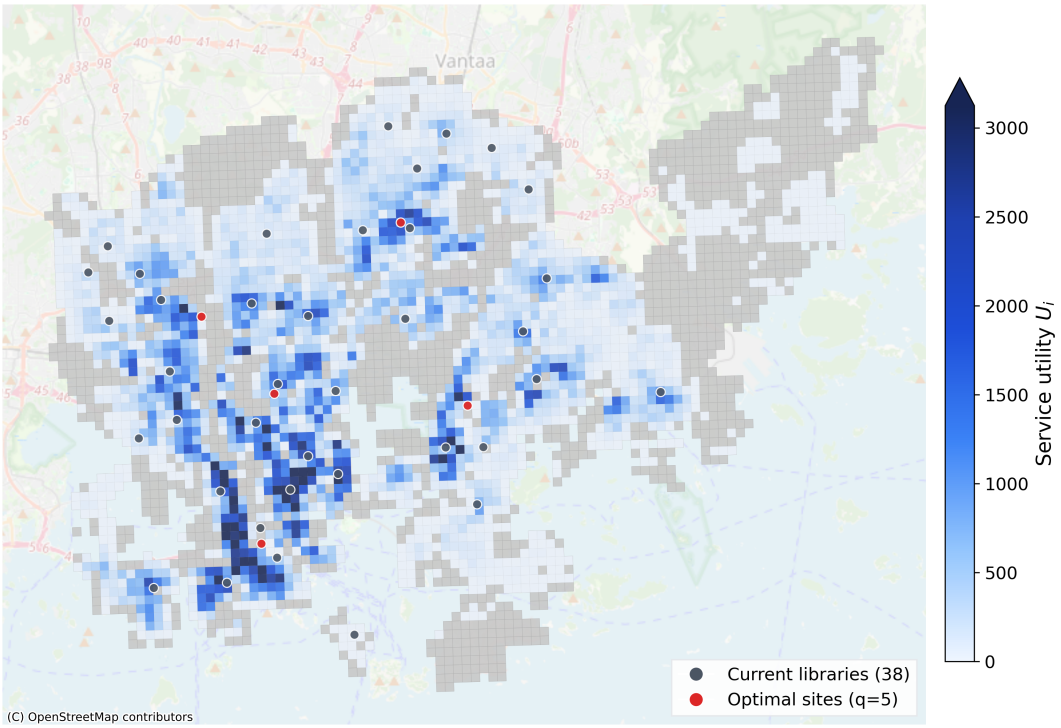


Figure 22: Optimal ground-up network reaching 50% of baseline utility. The selected sites are shown as red circles and the existing 38 libraries as gray circles for reference. Cells are colored by total utility on the baseline color scale.

5.2.4 70% of baseline utility ($q = 10$)

In this subsection we increase the library count until we reach 70% of the baseline utility. The 10-library network reaches 72.6% of the baseline utility. Most of the utility is captured in the densest areas, which highlights the uneven distribution of residents in Helsinki. Figure 23 shows the per-capita view of the optimized library network. The well-served areas cluster around the library locations. Table 15 shows the summary values for this scenario. With 10 libraries placed to maximize utility, only 12% of cells and 35% of the population are above the comfortable-access threshold. Figure 24 also allows comparing the optimized library locations to the baseline. Many of the optimized locations are near the existing libraries, which in part supports the quality of the current service locations.

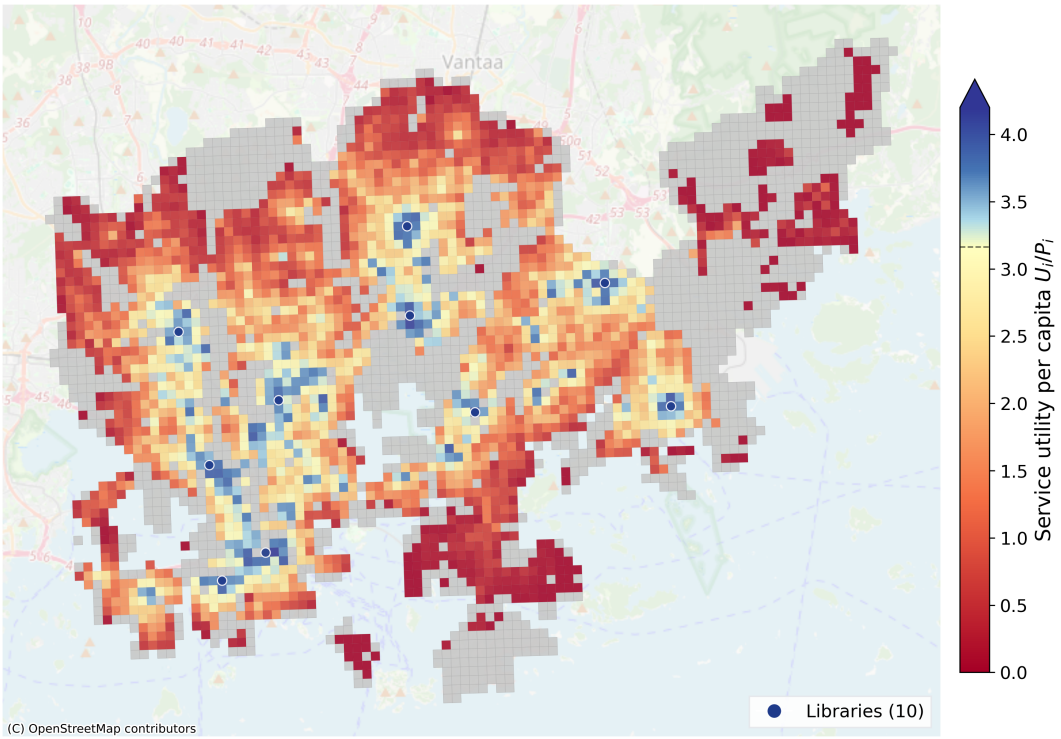


Figure 23: Per-capita utility U_i/P_i for the ground-up optimal network at $q = 10$ (all profiles). The network is optimized with $\tau_{\max} = 60$ min and the 2036 population prediction. The color gradient is centered on the comfortable-access threshold.

	Young 20–29	Adults 30–49	Middle-aged 50–64	Older 65+	**Adults 20+ (combined)**	Δ vs. baseline
U_r	332,895	525,793	360,281	171,010	**1,389,980**	-523,587
(% of U_r^{base})	(81.6 %)	(63.6 %)	(89.9 %)	(61.5 %)	**(72.6 %)**	(-27.4 %)
$\bar{u}_r^{\text{pc}}$	2.95	2.12	2.63	1.21	**2.17**	-0.82
$\tilde{u}_r^{\text{pc}}$	2.57	1.04	2.71	0.63	**1.60**	-1.18
$u_r^{\text{threshold}}$	3.20	4.00	2.40	2.40	**3.16**	0.00
cells above threshold	33.8 %	7.2 %	63.8 %	7.2 %	**12.0 %**	-22.3 pp
P_r	112,765	248,170	137,080	141,199	**639,213**	0
well served	65,144	30,549	110,722	16,354	**222,768**	-165,671
(% of P_r)	(57.8 %)	(12.3 %)	(80.8 %)	(11.6 %)	**(34.9 %)**	(-25.9 pp)
not well served	47,621	217,622	26,358	124,844	**416,445**	+165,671
(% of P_r)	(42.2 %)	(87.7 %)	(19.2 %)	(88.4 %)	**(65.1 %)**	(+25.9 pp)

Table 15: Utility summary, scenario: ground-up optimal network reaching ~ 70% of baseline utility ($q = 10$). Rows: total utility, mean and median per-capita utility, the comfortable-access threshold, the share of populated cells above it, the population, and the residents above and below their own profile's threshold. The combined well-served count is the sum of the per-profile counts. The last column shows the change compared to the 38-library baseline.

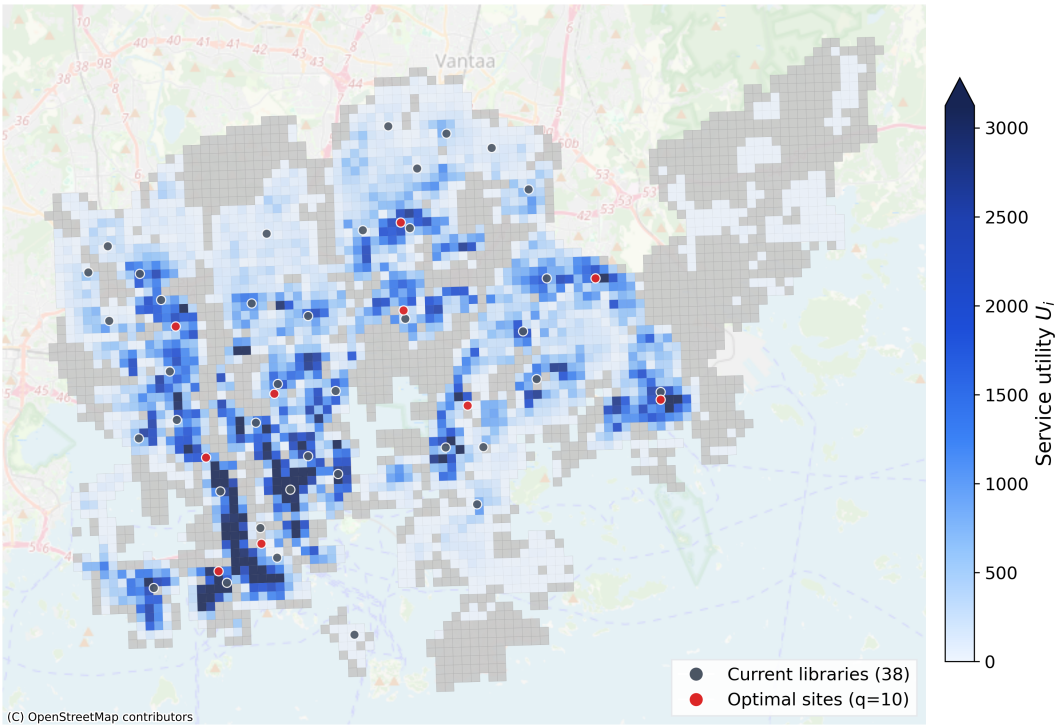


Figure 24: Optimal ground-up network reaching 70% of baseline utility. The selected sites are shown as red circles and the existing 38 libraries as gray circles for reference. Cells are colored by total utility on the baseline color scale.

5.2.5 100% of baseline utility ($q = 25$)

In this subsection we look at a scenario where we match the utility of the baseline network while optimizing the service locations. With the optimization, 25 libraries reach 101.2% of the baseline utility. Figure 25 shows the per-capita utility for the optimized service network, and Figure 26 shows the optimized service locations. The utility distribution is very similar to the baseline service network. Moreover, compared to the smaller networks, more of the optimized sites match the current locations, but notable differences are apparent. For example, the northwest part of Helsinki has notably fewer libraries than in the baseline and the libraries are in general more spread out.

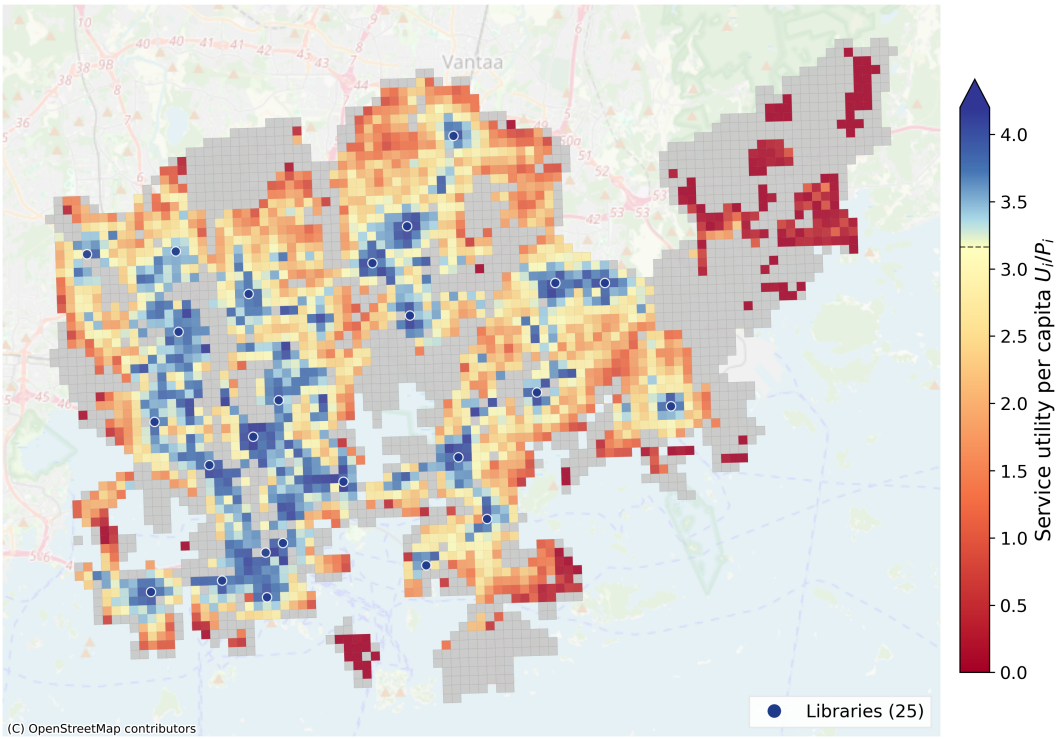


Figure 25: Per-capita utility U_i/P_i for the ground-up optimal network at $q = 25$ (all profiles). The network is optimized with $\tau_{\max} = 60$ min and the 2036 population prediction. The color gradient is centered on the comfortable-access threshold.

	Young 20–29	Adults 30–49	Middle-aged 50–64	Older 65+	**Adults 20+ (combined)**	Δ vs. baseline
U_r	414,401	841,864	402,689	278,517	**1,937,471**	+23,903
(% of U_r^{base})	(101.6 %)	(101.8 %)	(100.5 %)	(100.1 %)	**(101.2 %)**	(+1.2 %)
$\bar{u}_r^{\text{pc}}$	3.67	3.39	2.94	1.97	**3.03**	+0.04
$\tilde{u}_r^{\text{pc}}$	3.64	2.73	2.98	1.64	**2.64**	-0.13
$u_r^{\text{threshold}}$	3.20	4.00	2.40	2.40	**3.16**	0.00
cells above threshold	67.8 %	23.9 %	89.5 %	24.0 %	**31.1 %**	-3.2 pp
P_r	112,765	248,170	137,080	141,199	**639,213**	0
well served	98,527	105,619	134,381	53,809	**392,336**	+3,897
(% of P_r)	(87.4 %)	(42.6 %)	(98.0 %)	(38.1 %)	**(61.4 %)**	(+0.6 pp)
not well served	14,238	142,551	2,698	87,390	**246,878**	-3,897
(% of P_r)	(12.6 %)	(57.4 %)	(2.0 %)	(61.9 %)	**(38.6 %)**	(-0.6 pp)

Table 16: Utility summary, scenario: ground-up optimal network matching the baseline utility ($q = 25$). Rows: total utility, mean and median per-capita utility, the comfortable-access threshold, the share of populated cells above it, the population, and the residents above and below their own profile's threshold. The combined well-served count is the sum of the per-profile counts. The last column shows the change compared to the 38-library baseline.

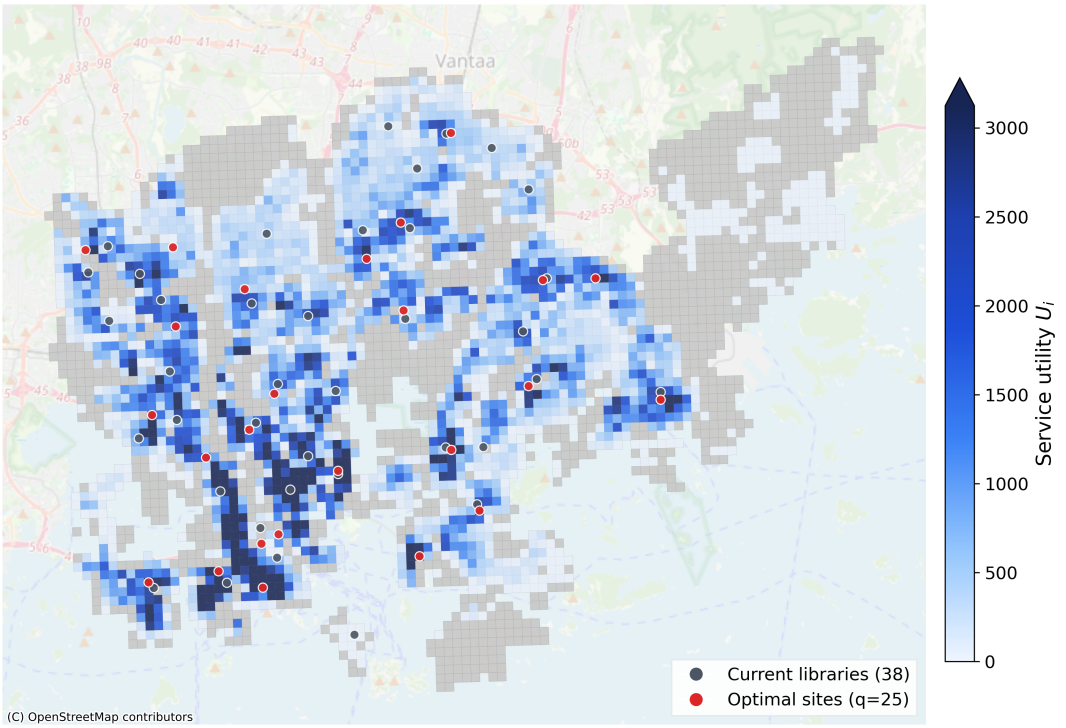


Figure 26: Optimal ground-up network matching the baseline utility. The selected sites are shown as red circles and the existing 38 libraries as gray circles for reference. Cells are colored by total utility on the baseline color scale.

5.2.6 Equivalent network size ($q = 38$)

As mentioned, the current baseline size of the service network is 38 libraries. Here, we compare the optimal service network computed ground-up to the baseline library network. Figure 27 illustrates the optimal network and the baseline service locations. The figure shows that the eastern part of Helsinki receives more libraries in the optimal scenario than in the existing service network. Furthermore, 75.8% of the population is above the threshold of comfortable access, an increase of 15.0 percentage points or 96 166 people. The total utility increases by 12.3%, most of the increase being captured by the Adults and Older adults profiles. Figure 28 shows the corresponding per-capita utility distribution.

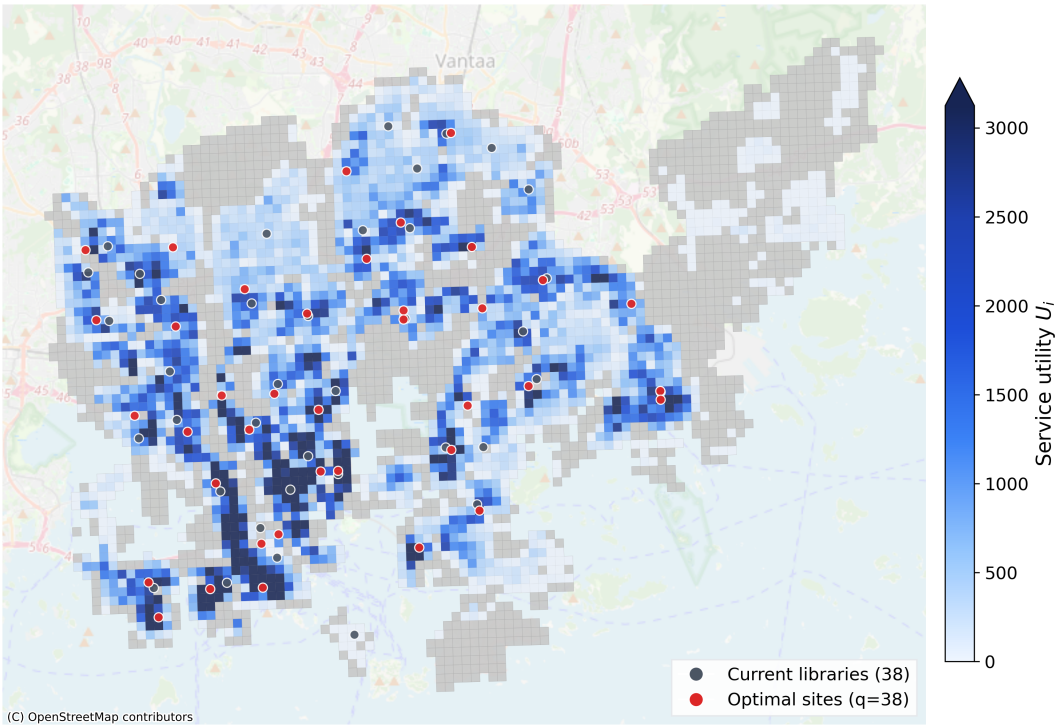


Figure 27: Ground-up optimal network with the same number of sites as the baseline ($q = 38$). The selected sites are shown as red circles and the existing 38 libraries as gray circles for reference. Cells are colored by total utility on the baseline color scale.

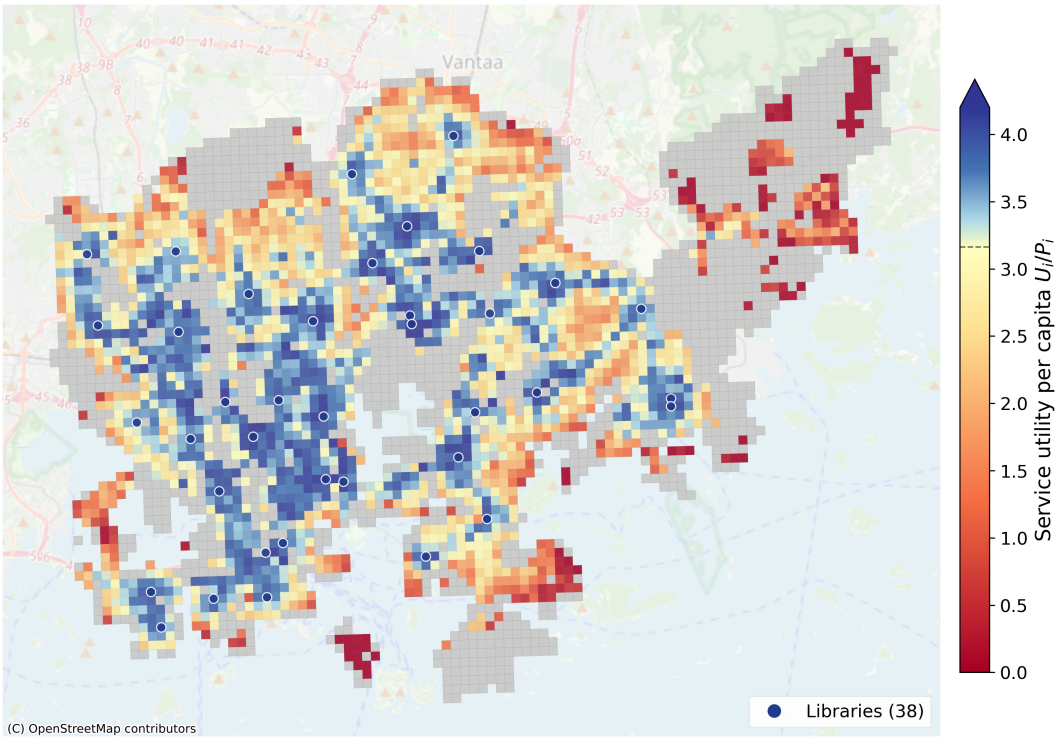


Figure 28: Per-capita utility U_i/P_i for the ground-up optimal network at $q = 38$ (all profiles). The network is optimized with $\tau_{\max} = 60$ min and the 2036 population prediction. The color gradient is centered on the comfortable-access threshold.

	Young 20–29	Adults 30–49	Middle-aged 50–64	Older 65+	**Adults 20+ (combined)**	Δ vs. baseline
U_r	431,973	982,531	406,955	327,237	**2,148,696**	+235,128
(% of U_r^{base})	(105.9 %)	(118.8 %)	(101.5 %)	(117.6 %)	**(112.3 %)**	(+12.3 %)
$\bar{u}_r^{\text{pc}}$	3.83	3.96	2.97	2.32	**3.36**	+0.37
$\tilde{u}_r^{\text{pc}}$	3.89	3.60	3.00	2.16	**3.09**	+0.31
$u_r^{\text{threshold}}$	3.20	4.00	2.40	2.40	**3.16**	0.00
cells above threshold	80.2 %	40.1 %	92.2 %	40.0 %	**46.6 %**	+12.2 pp
P_r	112,765	248,170	137,080	141,199	**639,213**	0
well served	106,458	159,308	135,273	83,567	**484,605**	+96,166
(% of P_r)	(94.4 %)	(64.2 %)	(98.7 %)	(59.2 %)	**(75.8 %)**	(+15.0 pp)
not well served	6,307	88,862	1,806	57,632	**154,608**	-96,166
(% of P_r)	(5.6 %)	(35.8 %)	(1.3 %)	(40.8 %)	**(24.2 %)**	(-15.0 pp)

Table 17: Utility summary, scenario: ground-up optimal network with the baseline-equivalent number of sites ($q = 38$). Rows: total utility, mean and median per-capita utility, the comfortable-access threshold, the share of populated cells above it, the population, and the residents above and below their own profile's threshold. The combined well-served count is the sum of the per-profile counts. The last column shows the change compared to the 38-library baseline.

The optimal network raises the well-served population share more than the well-served cell share, which shows that the objective favors densely populated areas. The network that maximizes total utility does not necessarily represent the most equitable placement.

6 Discussion and concluding remarks

The goal of this thesis was to construct a need-based model of service accessibility and an optimization framework to evaluate and optimize service networks. We applied the framework in the City of Helsinki for public libraries and determined how the network could be improved using a population prediction for the year 2036. We answered three research questions: 1. How can service accessibility be modeled so that it includes distance decay, need-based demographic differences, and the diminishing returns of multiple nearby services? 2. How accessible is the current service network of libraries in Helsinki? 3. How much room is there to improve the current service network of 38 libraries?

The baseline calculation of the current service network shows that the accessibility of the network is high. Within 25 minutes, 99.0% of the adult population (20+) in Helsinki can reach their nearest library, and within 45 minutes the entire adult population can. The accessibility threshold of comfortable access was set to $a = 0.5$. This is the level that reaches 80% of the maximum achievable benefit per cell. For an older adult, $a = 0.5$ corresponds to a single library 8 minutes away or two libraries 12 minutes away. With this definition, 60.8% of the population in 2036 will be above the comfortable-access threshold. It is up to future research and model users to define travel time limits that fit the modeled context best.

It is clear that the network has been built cumulatively. Each library location has been decided based on the assumptions and the network that was active at the time. One of the optimization approaches we selected was to build the network assuming that no libraries exist in Helsinki. This way we could calculate scenarios where the network is optimized as a whole, and not iteratively. From the optimization, we found that the same level of utility calculated for the baseline of 38 libraries could be achieved with 25 locations. The interpretation of this result should be that it is possible to evaluate service networks based on optimization and that the results can highlight inefficiencies in the evaluated set of services. However, this does not mean that there are too many libraries in Helsinki, but the allocation of the libraries should be examined critically.

We found that 50% of baseline utility could be achieved with 5 libraries, 70% with 10, and 100% with 25. With every additional service location, the utility gained is diminishing. For example, with optimal placement, increasing the library count from 25 to 38, an increase of 52%, raised the total utility from 101.2% to 112.3% of the baseline utility. It seems that it is possible to serve the population in 2036 better without increasing the number of libraries. Furthermore, assessing the locations in the optimal grid of 38 libraries shows that the network is more balanced than the baseline, with eastern parts of Helsinki receiving more libraries. This can be interpreted to mean that for the year 2036, there will be a need to shift some of the library services from the western parts of Helsinki to the eastern parts. Also, in some locations the model seems to place two libraries near each other. The locations are in high-density areas, where higher demand for library resources could be expected. In addition to creating service networks from the ground up, we determined optimal locations for a single additional service as well as for a single service removal. The optimal location for

adding a service to the current network is Punavuori, Helsinki. The optimal location for removing a service is Suomenlinna, Helsinki.

6.1 Limitations and assumptions

In this thesis, for practical reasons we used objective measures of the urban environment. In reality, subjective measurements such as surveyed happiness and satisfaction also play a significant role in assessing the livability and utility of an area. However, most factors affecting a resident's satisfaction with, for example, the library network, can be determined through objective measures such as travel time. It is left for future research to determine the weight and effect of subjective factors on the models.

The model that was used in this thesis does not consider the value of multipurpose trips. Here a multipurpose trip is defined as a typical behavioral pattern where an individual bundles several destinations into a single journey. For example, visiting a library might happen while traveling from work to home, or as a continuation of a grocery shopping trip. The model we developed can be interpreted to assume single-purpose trips. In reality a library located in a service hub would likely increase the total utility of a service network more than a library farther away from other everyday services. As a result, the model might undervalue libraries that are near frequently visited locations.

As with all models, assumptions need to be made to reflect the real world as closely as possible. These assumptions usually reflect the values and viewpoints of the organization or working unit that sets them, which creates bias in the model. In this thesis, we made multiple assumptions about population profile preferences, and it is important to acknowledge that they affect the results significantly.

The results depend on the selected values of n_{rs}, h_{rs} and a^*. We did not perform a sensitivity analysis of these values, so the robustness of individual results, such as the eastern reallocation in the 38-library scenario, was not tested. A sensitivity analysis of the values used in this thesis is recommended, especially if the model is applied in decision-making.

In the thesis, we used two functions, based on literature and discussions with City of Helsinki representatives, to measure the behavioral and preferential patterns of the residents. First, we implemented a Gaussian-form function to model how willingness to travel changes with travel time. Second, we used a concave function to model the diminishing returns of utility as the number of libraries near a cell increases. In practice, this steered the model away from clustering services to the most densely populated areas. For future research, it is advisable to evaluate different functions and measure how they affect model behavior. These could include per-cell caps for utility, power functions or just optimizing for the nearest facility.

It should also be noted that the equity constraint used in the optimization model is a soft constraint with a cell-specific penalty. This means that coverage does not strictly dominate the utility score, and with a small number of services the model can leave a sparsely populated cell uncovered. If a strict coverage guarantee is required, the constraint should be changed into a hard constraint.

The optimization runs used a relative MIP gap of 1%, which is of the same order of magnitude as the smallest reported effects, such as the 1.5% gain from adding one library. The reported utility values are not affected by this, because each solution is a feasible network evaluated with the exact utility function. The gap means that slightly better networks may exist. Results such as reaching the baseline utility with 25 libraries remain valid, but the selected sites are not guaranteed to be the exact optimal locations.

6.2 Data considerations

We used a travel time matrix from 2023, developed by the University of Helsinki [16]. The matrix provided an excellent base for travel time estimation, but many cells did not contain travel times to all other cells. We filtered out cells with fewer than 100 valid travel times and were left with 3 783 cells that contain enough data to represent public transport travel realistically. However, the missing travel times could still restrict the optimal solutions that the optimizer can find.

The population predictions from the City of Helsinki are for the year 2036, while the travel time matrix is from 2023. This means that the model has an assumption that travel times do not change during this period. Furthermore, the predictions contained information that was not used in this thesis. For example, population marked with cell identification number 9999999 was dropped, since that identifier does not refer to an actual cell but represents residents with no permanent residence.

Through the travel time matrix, it would have been possible to consider other transportation modes, such as car, walking and bicycle. We used public transport, because it offers a balance between walking and car travel. It also represents the most accessible option, since it does not depend on owning e.g. a car or a bicycle.

For the optimization model, we did not constrain the solution space, for example based on building permits or existing infrastructural limitations. Constraining the solution space is advisable, since it would decrease computation time significantly, which would allow for new objectives such as resource allocation, service models, capacity constraints or resource efficiency evaluation.

A further suggested improvement to the optimization model is to include service location attractiveness. In reality, all libraries have a relative attractiveness that is determined by the facility characteristics such as size, collection and additional services. This should be assessed in future research to weight each library by the volume of services it provides. In addition, considering the capacity of each library could increase the accuracy of the model. We did not consider capacity as a variable in this thesis.

6.3 For city planners

This thesis presents a concrete and repeatable approach for evaluating and optimizing service networks in the urban area. As an evaluation tool, the utility model developed can be used to compare areas and quantify the differences in utility levels. Furthermore, the optimization model can generate scenarios for different numbers of services or

determine how much improvement could be expected with different service networks. The model uses population predictions instead of the current population, which enables proactive decision-making. Also, the model can be extended beyond libraries to other public services. These could include services such as health stations, schools, daycare, or sports facilities.

6.4 Concluding remarks

The contribution of this thesis is in the suggested approach of combining travel-time accessibility, distance decay, diminishing returns, need-based weighting of a heterogeneous population, and equity into a single generalizable model. The approach highlights that it is possible, methodologically and computationally, to create mathematical models that give additional information for public service placement decisions.

As demonstrated for Helsinki, the model gives a starting point for evaluating complex service network problems in a computationally efficient setting. It helps city planners make informed and transparent decisions, and its explicit control of efficiency and equity supports value-based choices. With population predictions as input, the model provides a forward-looking tool to allocate public services to locations that maximize service accessibility in growing cities.

References

[1] Timur Abbiasov, Cate Heine, Sadegh Sabouri, Arianna Salazar-Miranda, Paolo Santi, Edward Glaeser, and Carlo Ratti. The 15-minute city quantified using human mobility data. *Nature Human Behaviour*, 8(3):445–455, 2024. doi:10.1038/s41562-023-01770-y.

[2] Shabbir Ahmed and Alper Atamtürk. Maximizing a class of submodular utility functions. *Mathematical Programming*, 128(1–2):149–169, 2011. doi:10.1007/s10107-009-0298-1.

[3] Benjamin Büttner, Cecília Silva, Louis Merlin, and Karst Geurs. Just around the corner: Accessibility by proximity in the 15-minute city. *Journal of Urban Mobility*, 6:100095, 2024. doi:10.1016/j.urbmob.2024.100095.

[4] Xiang Chen and Pengfei Jia. A comparative analysis of accessibility measures by the two-step floating catchment area (2SFCA) method. *International Journal of Geographical Information Science*, 33(9):1739–1758, 2019. doi:10.1080/13658816.2019.1591415.

[5] Richard Church and Charles ReVelle. The maximal covering location problem. *Papers of the Regional Science Association*, 32(1):101–118, 1974. doi:10.1111/j.1435-5597.1974.tb00902.x.

[6] City of Helsinki. Helsinki map service, 2026. Accessed: 11.5.2026. URL: https://kartta.hel.fi/?setlanguage=fi.

[7] City of Helsinki. Population prediction data, 2026. Closed source data. Accessed: 11.5.2026.

[8] City of Helsinki, Urban Research and Statistics. Population of the Helsinki Metropolitan Area by sub-area by age 1974- (Table 004r). Helsinki Region Open Statistical Databases, 2024. Accessed: 11.5.2026. URL: https://stat.hel.fi/pxweb/fi/Aluesarjat/.

[9] Stefano Coniglio, Fabio Furini, and Ivana Ljubić. Submodular maximization of concave utility functions composed with a set-union operator with applications to maximal covering location problems. *Mathematical Programming*, 196(1–2):9–56, 2022. doi:10.1007/s10107-022-01884-7.

[10] Mark Dooris and Zoe Heritage. Healthy Cities: Facilitating the active participation and empowerment of local people. *Journal of Urban Health*, 90(1):74–91, 2013. doi:10.1007/s11524-011-9623-0.

[11] Tammy Drezner, Zvi Drezner, and Dawit Zerom. Facility dependent distance decay in competitive location. *Networks and Spatial Economics*, 20(4):915–934, 2020. doi:10.1007/s11067-020-09507-4.

[12] Ahmed Atef Elsawy, Hany M. Ayad, and Dina Saadallah. Assessing livability of residential streets – case study: El-Attarin, Alexandria, Egypt. *Alexandria Engineering Journal*, 58(2):745–755, 2019. doi:10.1016/j.aej.2019.06.005.

[13] Dick Ettema and Marinel Schekkerman. How do spatial characteristics influence well-being and mental health? Comparing the effect of objective and subjective characteristics at different spatial scales. *Travel Behaviour and Society*, 5:56–67, 2016. doi:10.1016/j.tbs.2015.11.001.

[14] John Eyles, Stephen Birch, Susan Chambers, Jeremiah Hurley, and Brian Hutchison. A needs-based methodology for allocating health care resources in Ontario, Canada: development and an application. *Social Science & Medicine*, 33(4):489–500, 1991. doi:10.1016/0277-9536(91)90331-6.

[15] Bilal Farhan and Alan T. Murray. Distance decay and coverage in facility location planning. *The Annals of Regional Science*, 40(2):279–295, 2006. doi:10.1007/s00168-005-0041-7.

[16] Christoph Fink, Elias Willberg, and Tuuli Toivonen. Helsinki region travel time matrix 2018–2023, May 2024. Data set. doi:10.5281/zenodo.11220980.

[17] Natalia Gonzalez Bohorquez, Lisa Stafford, Steven M. McPhail, Shayma Mohammed Selim, Sanjeewa Kularatna, and Christina Malatzky. Disability, equity, and measurements of livability: A scoping review. *Disability and Health Journal*, 17(1):101521, 2024. doi:10.1016/j.dhjo.2023.101521.

[18] S. Louis Hakimi. Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, 12(3):450–459, 1964. doi:10.1287/opre.12.3.450.

[19] S. Louis Hakimi. Optimum distribution of switching centers in a communication network and some related graph theoretic problems. *Operations Research*, 13(3):462–475, 1965. doi:10.1287/opre.13.3.462.

[20] Kamyar Hasanzadeh, Marketta Kyttä, and Greg Brown. Beyond housing preferences: Urban structure and actualisation of residential area preferences. *Urban Science*, 3(1):21, 2019. doi:10.3390/urbansci3010021.

[21] Tapani Heino and Mikko Niemi. Personal communication, 3 2026. Tapani Heino (Johtava asiantuntija, Kaupunginkanslia, Strategiaosasto, Data- ja analytiikkatiimi) and Mikko Niemi (Datatieteilijä, Kaupunginkanslia, Strategiaosasto, Data- ja analytiikkatiimi), City of Helsinki.

[22] Carl Higgs, Hannah Badland, Koen Simons, Luke D. Knibbs, and Billie Giles-Corti. The Urban Liveability Index: developing a policy-relevant urban liveability composite measure and evaluating associations with transport mode choice. *International Journal of Health Geographics*, 18(1):14, 2019. doi:10.1186/s12942-019-0178-8.

[23] Shuju Hu, Wei Song, Chenggu Li, and Jia Lu. A multi-mode Gaussian-based two-step floating catchment area method for measuring accessibility of urban parks. *Cities*, 105:102815, 2020. doi:10.1016/j.cities.2020.102815.

[24] Sylvia J. T. Jansen, Henny C. C. H. Coolen, and Roland W. Goetgeluk, editors. *The Measurement and Analysis of Housing Preference and Choice*. Springer, Dordrecht, 2011. doi:10.1007/978-90-481-8894-9.

[25] Tom Kauko. An analysis of housing location attributes in the inner city of Budapest, Hungary, using expert judgements. *International Journal of Strategic Property Management*, 11(4):209–225, 2007. URL: https://www.proquest.com/scholarly-journals/analysis-housing-location-attributes-inner-city/docview/205498126/se-2.

[26] George Kephart and Yukiko Asada. Need-based resource allocation: different need indicators, different results? *BMC Health Services Research*, 9:122, 2009. doi:10.1186/1472-6963-9-122.

[27] Kevin M. Leyden, Michael J. Hogan, Lisa D'Arcy, Brendan Bunting, and Sanne Bierema. Walkable neighborhoods: Linkages between place, health, and happiness in younger and older adults. *Journal of the American Planning Association*, 90(1):101–114, 2024. doi:10.1080/01944363.2022.2123382.

[28] Jianxiao Liu, Han Bi, and Meilian Wang. Using multi-source data to assess livability in Hong Kong at the community-based level: A combined subjective-objective approach. *Geography and Sustainability*, 1(4):284–294, 2020. doi:10.1016/j.geosus.2020.12.001.

[29] Liang Liu, Hongbing Lyu, Yifan Zhao, and Dingxi Zhou. An improved two-step floating catchment area (2SFCA) method for measuring spatial accessibility to elderly care facilities in Xi'an, China. *International Journal of Environmental Research and Public Health*, 19(18):11465, 2022. doi:10.3390/ijerph191811465.

[30] Wei Luo and Yi Qi. An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health & Place*, 15(4):1100–1107, 2009. doi:10.1016/j.healthplace.2009.06.002.

[31] Wei Luo and Fahui Wang. Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design*, 30(6):865–884, 2003. doi:10.1068/b29120.

[32] Michael T. Marsh and David A. Schilling. Equity measurement in facility location analysis: A review and framework. *European Journal of Operational Research*, 74(1):1–17, 1994. doi:10.1016/0377-2217(94)90200-3.

[33] André Ngamini Ngui and Philippe Apparicio. Optimizing the two-step floating catchment area method for measuring spatial accessibility to medical clinics in Montreal. *BMC Health Services Research*, 11(1):166, 2011. doi:10.1186/1472-6963-11-166.

[34] Sławomir Palicki. Housing preferences in various stages of human life cycle. *Real Estate Management and Valuation*, 28(1):91–99, 2020. doi:10.2478/remav-2020-0008.

[35] Tytti P. Pasanen, Timo Lanki, Taina Siponen, Anu W. Turunen, Pekka Tiittanen, Vuokko Heikinheimo, Maija Tiitu, Jaana I. Halonen, and Arto Viinikka. What makes a liveable neighborhood? Role of socio-demographic, dwelling, and environmental factors and participation in Finnish urban and suburban areas. *Journal of Urban Health*, 101(6):1207–1220, 2024. doi:10.1007/s11524-024-00927-y.

[36] Arpan Paul and Joy Sen. A critical review of liveability approaches and their dimensions. *Geoforum*, 117:90–92, 2020. doi:10.1016/j.geoforum.2020.09.008.

[37] John Radke and Lan Mu. Spatial decompositions, modeling and mapping service regions to predict access to social programs. *Geographic Information Sciences*, 6(2):105–112, 2000. doi:10.1080/10824000009480538.

[38] Charles S. ReVelle, Horst A. Eiselt, and Mark S. Daskin. A bibliography for some fundamental problem categories in discrete location science. *European Journal of Operational Research*, 184(3):817–848, 2008. doi:10.1016/j.ejor.2006.12.044.

[39] Charles S. ReVelle and Ralph W. Swain. Central facilities location. *Geographical Analysis*, 2(1):30–42, 1970. doi:10.1111/j.1538-4632.1970.tb00142.x.

[40] Tobias Riepl and Simon Grabow. Assessing livability through human need theory: Insights from Vienna. *Cities*, 167:106336, 2025. doi:10.1016/j.cities.2025.106336.

[41] Sachin Kumar Sahu and Vandana Agrawal. Location satisfaction in residential areas. *International Review for Spatial Planning and Sustainable Development*, 13(4):1–19, 2025. doi:10.14246/irspsd.13.4_1.

[42] Marta Sapena, Michael Wurm, Hannes Taubenböck, Devis Tuia, and Luis A. Ruiz. Estimating quality of life dimensions from urban spatial pattern metrics. *Computers, Environment and Urban Systems*, 85:101549, 2021. doi:10.1016/j.compenvurbsys.2020.101549.

[43] Yue Shao and Wei Luo. Supply-demand adjusted two-steps floating catchment area (SDA-2SFCA) model for measuring spatial access to health care. *Social Science & Medicine*, 296:114727, 2022. doi:10.1016/j.socscimed.2022.114727.

[44] Xun Shi, Jennifer Alford-Teaster, Tracy Onega, and Dongmei Wang. Spatial access and local demand for major cancer care facilities in the United States. *Annals of the Association of American Geographers*, 102(5):1125–1134, 2012. doi:10.1080/00045608.2012.657498.

[45] Statistics Finland. Grid net for statistics 250 m x 250 m. Geographic data service, Statistics Finland, 2022. Published 18.10.2022; next update scheduled for March 2027. Dataset available in GeoPackage format from the Statistics Finland geographic-data service. Accessed: 25.5.2026. URL: https://stat.fi/en/services/statistical-data-services/geographic-data/statistical-areas/grid-net-for-statistics-250-m.

[46] Statistics Finland. Kotitalouksien ja henkilöiden jakautuminen tulokymmenyksen ja kotitalouden elinvaiheen mukaan, 1989–2024, 2025. StatFin Database, table statfin_tjt_pxt_12ew. Accessed: 20.4.2026. URL: https://statfin.stat.fi/PxWeb/pxweb/fi/StatFin/StatFin__tjt/12ew.px/.

[47] Constantine Toregas, Ralph Swain, Charles ReVelle, and Lawrence Bergman. The location of emergency service facilities. *Operations Research*, 19(6):1363–1373, 1971. doi:10.1287/opre.19.6.1363.

[48] United Nations Department of Economic and Social Affairs, Population Division. World urbanization prospects 2025: Summary of results. Technical Report UN DESA/POP/2025/TR/NO. 12, United Nations Department of Economic and Social Affairs, Population Division, 2025. UN DESA Population Division website. Accessed: 12.5.2026. URL: https://population.un.org/wup/assets/Publications/undesa_pd_2025_wup2025_summary_of_results_final.pdf.

[49] Antti Vasanen. Beyond stated and revealed preferences: the relationship between residential preferences and housing choices in the urban region of Turku, Finland. *Journal of Housing and the Built Environment*, 27(3):301–315, 2012. doi:10.1007/s10901-012-9267-8.

[50] Hao Wang, Qiping Shen, Bo-sin Tang, Chen Lu, Yi Peng, and LiYaning Tang. A framework of decision-making factors and supporting information for facilitating sustainable site planning in urban renewal projects. *Cities*, 40:44–55, 2014. doi:10.1016/j.cities.2014.04.005.