POLITECNICO DI TORINO

ICT Engineering for Smart Societies

Traffic-Aware Network Sharing: A Data-Driven Analysis

Candidate
Alessandro Caramia

Supervisors
Prof. Minchela Meo
Prof. Daniela Renga
Prof. Loutfi Nuaymi

Thesis in collaboration with

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Academic Year 2025–2026

Acknowledgements

Vorrei anzitutto ringraziare le mie relatrici, Michela Meo e Daniela Renga, per il loro supporto, la loro disponibilità e l'aiuto offertomi nei momenti di difficoltà.

Un grandissimo grazie va ai miei genitori, a mia sorella, alle mie zie e, soprattutto, a mia nonna, per il suo immenso amore.

Ringrazio tutti i miei amici che mi hanno accompagnato durante questi anni di studio; in particolare Mattia e Simone, per le chiamate con cui ci siamo sostenuti a vicenda, e Roberto e Modesto, per la compagnia durante la stesura della tesi.

Un ringraziamento speciale va alla mia ragazza, Gaia Vitagliano, che mi ha supportato e sopportato con pazienza, comprensione e forza, soprattutto durante il periodo all'estero.

I would also like to thank Maoquan Ni for providing me with the material that allowed me to start this work, and for his presence and support throughout the development of this thesis. I would also like to thank Professor Loutfi Nuaymi, who supervised my period in France, as well as all the new friends and colleagues I met at IMT Atlantique.

A special thanks also goes to Manoj and Sam for all the conversations we shared, both constructive and light-hearted, which made this journey even more meaningful.

Infine, il ringraziamento più grande va a me stesso, per non essermi dato per vinto, per essermi messo in gioco e per aver portato a termine questo percorso.

Abstract

The continuous growth of mobile data demand has driven increasingly dense base station (BS) deployments, yet substantial redundant capacity remains during off-peak periods, leading to avoidable energy consumption. Network Sharing (NS) among operators offers a promising approach to improve resource utilization and reduce energy waste, provided that sharing decisions account for the traffic characteristics of the participating BSs. In this thesis, we investigate how these traffic characteristics shape NS performance in large metropolitan networks. Using real traffic measurements and deployment information, we characterize each BS through traffic profiles, volume indicators, and derived descriptors, and apply clustering techniques to identify groups of BSs exhibiting similar traffic behavior. These traffic-based descriptors are then coupled with NS performance metrics produced by a simulation framework to quantify both energy-saving potential and operational stability. Results show that BS pairs with well-aligned and temporally stable loads consistently enable higher energy reductions and fewer network reconfigurations, while mismatched or persistently high-load pairs provide limited gains and may induce frequent transitions.

To move beyond descriptive assessment, the thesis also introduces a predictive component aimed at supporting proactive NS operation. We formulate the anticipation of capacity-threshold exceedances as a classification task and evaluate lightweight machine-learning models using only information available up to the decision time (e.g., temporal encodings and recent load history). The prediction outcomes provide actionable decision variables for identifying risky conditions in advance and for reducing unnecessary switching. Overall, the study highlights the importance of traffic-aware design and shows how combining traffic-driven clustering with short-term prediction can guide more efficient and stable NS strategies.

Contents

List of Figures

List of Tables

Acronyms

AGCOM Autorità per le Garanzie nelle Comunicazioni.

ANFR Agence nationale des fréquences.

ARCEP Autorité de régulation des communications électroniques, des postes et de la distribution de la presse.

BS Base Station.

BSs Base Stations.

CAPEX Capital Expenditure.

DBI Davies–Bouldin index.

DSL Digital Subscriber Line.

DTW Dynamic Time Warping.

FANS Fixed Access Network Sharing.

FCNS Fair Cooperative Network Sharing.

HAC Hierarchical Agglomerative Clustering.

KPI Key Performance Indicator.

LLU Local Loop Unbundling.

MIMO Multiple-Input Multiple-Output.

MNOs Mobile Network Operators.

MOCN Multi-Operator Core Network.

MVNOs Mobile Virtual Network Operators.

NFV Network Function Virtualization.

NGA Next Generation Access.

NS Network Sharing.

OPEX Operational Expenditure.

POIs Points of Interest.

QoS Quality of Service.

RAN Radio Access Network.

RF Random Forest.

RRH Remote Radio Head.

SDN Software-Defined Networking.

VULA Virtual Unbundled Local Access.

Chapter 1

Introduction

Cell phone networks have become crucial to our daily lives. Every call, message, or video stream relies on a complex system of antennas, radio gear, and data centers. As more and more data is used, especially with the introduction of 5G and the trajectory towards 6G, companies that provide mobile service must find ways to offer better coverage and faster speeds without significantly increasing costs or using too much energy. Mobile data traffic also continues to grow steadily and is forecast to more than double by 2031, with 5G expected to carry the large majority of global mobile data traffic, further increasing the pressure on operators to expand capacity. [1] A major issue is that the usual way of building these networks creates duplicate infrastructure, since each provider installs their own antennas and equipment, even in areas where their networks overlap. This method is not only costly but also inefficient in terms of energy and environmental impact.

The advent of 5G increased network capacity through wider spectrum allocations, additional carriers, more advanced radio equipment (e.g., massive Multiple-Input Multiple-Output (MIMO)), and often denser site deployments, especially at higher frequencies. While these advances improve the energy efficiency per delivered bit, the larger number of sites and the increased signal-processing burden can raise the energy demand and the overall site footprint when infrastructure continues to be deployed in parallel by multiple operators. [2]

In this context, the Radio Access Network (RAN) accounts for the largest share of electricity consumption often exceeding 50%, making it the primary lever for energy-efficiency interventions. [3] Consequently, RAN energy demand directly impacts Operational Expenditure (OPEX), while the need to deploy additional sites and radio equipment increases Capital Expenditure (CAPEX). [4]

Reducing energy consumption and environmental impact is a key priority for both industry and society. Empirical evidence from European markets suggests that, when properly designed and supervised, Network Sharing (NS) can improve coverage and service quality while also lowering consumer prices and increasing operators' profitability, indicating that sustainability and competition need not be mutually exclusive. [5]

NS addresses this inefficiency by allowing multiple mobile network operators to share infrastructure, ranging from passive assets (e.g., sites and power supply) to active elements in the RAN. By reducing redundant deployments, NS enables traffic consolidation

across co-located networks and, during low-demand periods, the temporary deactivation of underutilized base stations. From an operator perspective, the adoption of sharing arrangements is tightly coupled with the availability of adequate economic incentives and business models, which can otherwise hinder the implementation of sharing in practice [6].

Despite its potential, a practical challenge remains: the benefits and operational risks of sharing depend strongly on the demand for traffic that changes over time, and existing studies often provide limited guidance on the traffic conditions that make sharing more effective and operationally stable. Building on recent work, NS can also be framed as a lever not only for operational savings but for broader sustainability goals, since reducing on-time and avoiding unnecessary deployments can contribute to lowering the overall footprint of the access network [4].

To address these challenges, this thesis develops a data-driven framework to analyze when network sharing is most beneficial and operationally sustainable under realistic traffic dynamics.

The main contributions of this thesis can be summarized as follows:

- First, real-world mobile traffic traces are processed to derive representative traffic descriptors for each Base Station (BS), combining both temporal usage patterns and average traffic volume. These descriptors are then used to cluster BSs into interpretable traffic-profile categories.

- Second, the relationship between combinations of co-located BS traffic profiles and a set of Key Performance Indicator (KPI) is systematically quantified. This makes it possible to identify which pairings are more likely to achieve high energy savings and stable operation under NS, thus providing practical insight for traffic-aware sharing planning.

- Third, a feasibility score is defined to provide a synthetic and flexible assessment of NS opportunities by jointly capturing the trade-off between energy-saving potential and operational costs, such as frequent state transitions. This score is especially useful because it allows different sharing configurations to be ranked according to operator priorities, offering a practical decision-support metric for selecting the most convenient NS strategies.

- Finally, to support operational decision-making under uncertain future demand, a prediction module is developed to anticipate traffic conditions that may compromise NS feasibility. By forecasting threshold exceedance events more accurately than traditional traffic prediction baselines, the proposed approach enables more reliable and proactive network-sharing decisions.

Overall, the thesis shows that network sharing should not be treated as a purely infrastructure-level optimization problem, but as a traffic-aware decision problem in which the temporal compatibility of demand profiles plays a central role in balancing energy savings, operational stability, and service reliability.

Chapter 2

Background and related work

This chapter reviews the main concepts and prior literature on NS in telecommunications, with a primary focus on mobile networks. We first introduce the commercial, industrial, and regulatory drivers behind NS and clarify the main sharing models adopted in practice and standardized in the mobile ecosystem. We then outline the evolution from early site-based cooperation to active RAN sharing, highlighting how densification and traffic variability make the effectiveness of NS strongly dependent on the temporal and spatial distribution of load. We also summarize related work on energy-efficient and traffic-aware NS, emphasizing the gaps that motivate our data-driven analysis and the methodology developed in the next chapters.

2.1 Commercial and Industrial Context

Mobile NS refers to operator-level agreements in which two or more Mobile Network Operators (MNOs) jointly use selected parts of the network infrastructure. Depending on the scope of cooperation, sharing may involve **passive** elements (e.g., sites, masts, shelters, power supply) and/or **active** elements (e.g., radio equipment, baseband processing) and, in some cases, spectrum resources. In practice, BS is shaped by both business incentives and technical feasibility: it is pursued to improve the economic sustainability of deployments while maintaining service differentiation and compliance with regulatory constraints.

2.1.1 Commercial Perspective

From a commercial standpoint, network sharing enables operators to reduce CAPEX and OPEX by avoiding duplication of infrastructure. Cost savings are particularly relevant in rural or low revenue areas, where standalone deployments are often uneconomical [5]. Sharing agreements facilitate the faster rollout of next-generation technologies such as 5G, allowing operators to maintain competitive pricing and improve service quality. Empirical evidence shows that network sharing can lead to lower retail prices, greater broadband penetration, and better network coverage, benefiting end users while supporting the market entry of smaller operators [5]. In some cases, sharing agreements have also

been used as a transitional strategy, allowing new entrants or Mobile Virtual Network Operators (MVNOs) to gain market presence while gradually building their own infrastructure. However, regulators monitor such agreements closely to ensure that efficiency gains do not reduce competitive incentives or lead to coordinated market behavior that might harm consumers.

2.1.2 Industrial Perspective

From an industrial perspective, NS promotes the efficient use of resources such as spectrum, radio sites, and energy. Operators may adopt passive sharing, which includes site, tower, and power infrastructure, or active sharing, which extends to RAN and, in some cases, spectrum resources [7]. The emergence of tower companies (TowerCos) and neutral-host models has increased the flexibility of site and infrastructure sharing, separating asset ownership from service provision and enabling multi-tenant access [7]. More recently, the adoption of virtualization technologies such as Network Function Virtualization (NFV) and Software-Defined Networking (SDN) has enabled dynamic and software-driven sharing arrangements, including network slicing in 5G, which allows multiple operators or tenants to share the same physical network while maintaining logical separation of services [8]. Beyond efficiency, sharing supports environmental sustainability, as fewer physical sites reduce energy consumption, CO_2 emissions, and the visual impact of mobile infrastructure [4].

2.1.3 Regulatory Perspective

NS agreements are typically assessed by national regulatory authorities under a combination of competition and public-interest criteria. Across jurisdictions, evaluations commonly consider: (i) the impact on retail competition and investment incentives, (ii) transparency and non-discrimination conditions (especially when wholesale access is involved), (iii) coverage obligations and service quality, and (iv) broader societal objectives, including environmental and visual-impact constraints.

For example, in the **United States**, sector oversight is led by the FCC (Federal Communications Commission), with NS arrangements generally scrutinized under spectrum licensing conditions and competition policy principles. In **France**, Autorité de régulation des communications électroniques, des postes et de la distribution de la presse (ARCEP) provides regulatory oversight for electronic communications markets, while spectrum planning and technical management are handled by Agence nationale des fréquences (ANFR); ARCEP has also issued guidelines discussing when sharing is acceptable and how it should be structured to preserve competitive dynamics. In addition, specific mechanisms have been used to improve coverage in underserved areas (zones blanches), including local roaming and infrastructure mutualization frameworks [9, 10]. In **Italy**, Autorità per le Garanzie nelle Comunicazioni (AGCOM) oversees electronic communications markets, with a mandate that includes ensuring fair competition and protecting users' rights; it can also issue opinions on sharing arrangements involving spectrum use and market structure, emphasizing efficiency gains while requiring safeguards for competition and transparency [11].

2.1.4 Illustrative Agreements

A few examples illustrate the diversity of NS arrangements and their motivations:

Free Mobile–Orange (France): legacy roaming sunset. Free Mobile has relied on Orange's 2G/3G networks for nationwide roaming since 2011, with the current agreement running until the end of 2025. A new contract submitted to ARCEP extends roaming on Orange's 2G/3G networks from 2026 to 2028, aligning with Orange's planned shutdown of 2G in 2026 and 3G in 2028, and Free Mobile's intention to decommission its own 3G network to repurpose spectrum for 4G/5G. The contract reflects the phase-out of legacy technologies, defined as a sunset mechanism for the older technologies, in favor of more efficient spectrum use for 4G/5G. [12]

Iliad–Wind Tre (Italy): joint venture and MOCN in selected areas. The agreement involved the implementation of a Multi-Operator Core Network (MOCN) in specific areas of Italy by Wind Tre and Iliad, allowing the two operators to share both radio access infrastructure and spectrum resources. AGCOM issued a favorable opinion, emphasizing that such cooperation could enhance the efficiency of 5G deployment and prevent unnecessary duplication of infrastructure, while underscoring the importance of regulatory oversight to preserve competition and market transparency. [13]

2.2 Evolution of Network Sharing in Mobile Networks

Infrastructure sharing in mobile networks has been studied for more than three decades and has progressively shifted from a purely cost-driven practice toward a key architectural and business component of modern cellular systems. A consistent driver across generations is the possibility to divide infrastructure costs among multiple MNOs, improving profitability and accelerating deployments, especially during expensive technology migrations [14].

The earliest forms of sharing were predominantly passive, limited to physical space and non-active assets such as towers, shelters, and cabinets [14]. In the same period, national roaming emerged as a practical mechanism to extend coverage, and early work already discussed how roaming could be part of cellular specifications to improve nationwide availability under certain frequency plans. Alongside these technical arrangements, sharing-related research and regulation were closely connected to market liberalization and to the introduction of new stakeholders such as MVNOs, often seen as a way to monetize spare network capacity.

With the growth of mobile broadband and the rising cost of 3G/4G rollouts, cooperation increasingly extended beyond sites to active elements in the RAN, enabling larger savings but also requiring stronger coordination and clearer isolation between operators. A key enabler of this transition was standardization, which progressively clarified architectures and responsibilities in multi-operator deployments; for example, 3GPP specifies reference configurations for network sharing (including MORAN/MOCN and related options) through an explicit architectural and functional description. In parallel,

trace-driven studies started to quantify how the benefits and trade-offs of sharing depend on realistic demand dynamics, showing that spatio-temporal traffic correlation between operators is a key factor when assessing sharing opportunities [15].

In 5G, infrastructure sharing is no longer only a cost-reduction tool: it is increasingly intertwined with the architectural vision of multi-tenancy and the decoupling of infrastructure from services. In particular, the evolution of 3GPP mechanisms from classical network sharing principles toward on-demand multi-tenant systems has been discussed as a pathway to support dynamically leased resources and service differentiation (e.g., via network slicing) on top of shared physical infrastructure [16]. This evolution broadens the sharing ecosystem beyond traditional cooperation, raising new challenges in resource management, and governance.

2.3 Models of Mobile Network Sharing

The main mobile network sharing configurations standardized within the 3GPP ecosystem and commonly adopted in practice include:

- **MORAN (Multi-Operator Radio Access Network)**: operators share RAN equipment (and often the site), while each operator retains dedicated spectrum resources. This reduces infrastructure duplication while preserving spectrum autonomy.
- **MOCN (Multi-Operator Core Network sharing)**: operators share RAN equipment and spectrum resources at the shared site, while maintaining independent core networks. Compared to MORAN, spectrum pooling can improve spectral utilization but typically requires tighter coordination.
- **National roaming**: a service-level form of sharing in which subscribers of one operator are allowed to use another operator's network in areas without coverage, without requiring a joint RAN deployment.
- **Virtualized / indirect sharing**: multi-tenant sharing enabled by virtualization and cloud-native network functions, providing logical separation between tenants and potentially more flexible resource allocation compared to traditional RAN-sharing models [17].

2.4 Fixed Access Network Sharing (FANS)

Fixed Access Network Sharing (FANS) has undergone substantial evolution since the mid-1990s, transitioning from copper-based Digital Subscriber Line (DSL) systems to advanced fiber optic infrastructures. This transformation has been primarily driven by the dual objective of promoting market competition while avoiding unnecessary duplication of physical assets. The development of FANS reflects both technical considerations, such as resource isolation, service differentiation, and network reliability, and economic

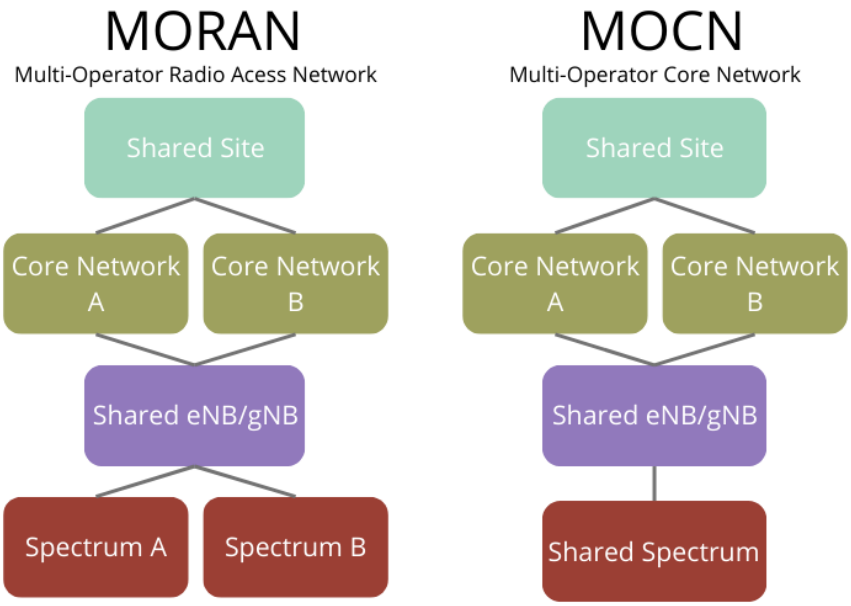


Figure 2.1: MORAN vs MOCN image

imperatives, including cost optimization and regulatory compliance. Unlike mobile network sharing, which focuses on the radio access segment, FANS is centered on the fixed 'last mile' connection between central offices and end users, allowing multiple operators to share physical and virtual access resources.

The earliest forms of FANS emerged during the liberalization of the telecommunications markets, mainly through Local Loop Unbundling (LLU) applied to DSL networks. LLU allowed alternative operators to lease sections of the incumbent's copper infrastructure connecting local exchanges to subscribers. Three principal models were implemented: line sharing, in which the incumbent retained control of the voice service while alternative operators exploited the higher frequency spectrum for broadband access; full unbundling, granting full control of the copper pair to the alternative operator; and bitstream access, where the incumbent provided high-speed connections aggregated at a limited number of handover points, allowing less operational control to the alternative operator.

Although these models supported competition and service innovation, they also faced limitations, including interference in shared cables, which required coordination and mitigation strategies between multiple operators.

The introduction of Next Generation Access (NGA) networks in the 2000s represented a major milestone, overcoming the inherent bandwidth and distance constraints of copper lines. Fiber-based enhancements enabled the delivery of high-speed broadband services, including . New wholesale access models evolved to balance flexibility and efficiency, such as Virtual Unbundled Local Access (VULA), which replicated much of the functional independence of LLU while reducing the costs associated with extensive backhaul infrastructure. This regulatory approach maintained the potential for service differentiation and innovation while ensuring economic sustainability.

Optical access networks, particularly Passive Optical Networks (PONs), have become the cornerstone of modern FANS due to their high capacity, low attenuation, and scalability. Nevertheless, the shared nature of PON infrastructures introduces challenges related to the allocation of optical resources and the prevention of interference. Time-division-based PON architectures allow dynamic bandwidth distribution among multiple operators, while wavelength-division-based systems offer dedicated optical channels to each operator, enhancing isolation at the expense of higher deployment costs.

The recent integration of SDN and NFV has further revolutionized the FANS paradigm. SDN introduces centralized programmability, enabling dynamic and policy-driven resource management, whereas NFV decouples network functions from proprietary hardware, allowing flexible and cost-effective deployment in virtualized environments. In this context, core access functions, such as those traditionally located at the optical line terminal, can now be virtualized within data centers, supporting multi-tenant architectures with logically isolated virtual networks. These advancements facilitate on-demand provisioning, significantly reduce capital expenditures through shared assets, and enable novel operational frameworks such as automated service orchestration and smart contract-based settlements, thus paving the way for a more open, efficient, and innovation-driven fixed access ecosystem. [8]

2.5 Related Work

NS has been widely investigated as a means to improve the energy efficiency and sustainability of mobile networks. An early study by Marsan *et al.* [18] provided one of the first quantitative assessments of the potential energy savings enabled by infrastructure sharing among multiple MNOs in European networks. Using analytical models, the authors showed that NS can significantly reduce energy consumption compared to independently operated networks, and highlighted that the achievable gains depend strongly on traffic underutilization during off-peak periods, which enables coordinated switch-off strategies while maintaining Quality of Service (QoS).

Building on this foundation, Renga *et al.* [19] investigated data-driven NS strategies in dense and energy-intensive 5G RANs. Using real-world mobile traffic traces, they showed that substantial energy savings can be achieved by dynamically consolidating traffic and deactivating underutilized Base Stations (BSs). Their results emphasized the importance of accounting for traffic variability and local conditions when designing effective NS policies.

More recently, [4] extended the evaluation of NS beyond operational energy by considering both energy consumption and carbon emissions, including embodied components associated with infrastructure life-cycle processes. This broader perspective further motivates NS as a lever for sustainability, while raising practical questions on how to design strategies that achieve large gains without introducing undesirable operational side effects.

While most studies focus on aggregate energy and sustainability benefits, the distribution of costs and benefits across participating operators has received comparatively less attention. Ni *et al.* [20] explicitly addressed this aspect by introducing a Fair Cooperative

Network Sharing (FCNS) framework, which offloads traffic between co-located BS pairs with the goal of balancing active time across operators and thus reducing asymmetries in operational burden. Complementary to fairness-aware control policies, Antonopoulos [21] proposed modeling multi-party sharing as a bankruptcy problem, providing resource-allocation rules and discussing applications across different 5G resource types.

Overall, prior work establishes that NS can deliver substantial efficiency and sustainability gains and has begun to address practical concerns such as operational side effects and fairness. However, limited attention has been devoted to understanding how the temporal structure of traffic demand can be used to guide NS decisions in a systematic and predictive way. In particular, the literature lacks a joint framework that links traffic-pattern characterization, compatibility between co-located BSs, synthetic decision metrics, and anticipation of critical traffic conditions. This thesis aims to fill this gap by proposing a data-driven approach that combines traffic clustering, pairwise analysis, feasibility scoring, and threshold-exceedance prediction for proactive network-sharing planning and operation.

Chapter 3

System Model

3.1 Scenario and entities

The scenario considered in this study focuses primarily on the Paris metropolitan area, which constitutes the main case study for the evaluation of traffic-aware NS strategies in a dense and high-demand urban environment. To assess the robustness and generality of the findings, we also replicate the same analysis pipeline on the Lyon metropolitan area as a secondary validation case. Paris and Lyon were selected because they exhibit different metropolitan scales and demand intensities, enabling the evaluation of NS under heterogeneous urban conditions while keeping the study tractable.

The analysis is based on real mobile traffic data provided by the NetMob dataset [22], released in the context of a public research challenge organized by a French mobile network operator. The objective of the NetMob initiative is to enable large-scale analysis of real cellular traffic and foster data-driven research on mobile network usage and performance. The dataset provides a high-resolution spatial representation of mobile traffic for 20 major French cities.

Traffic is aggregated over a regular spatial grid with a resolution of 100×100 m and a temporal resolution of 15 minutes. For each grid cell and time interval, the dataset reports traffic volumes associated with 68 popular mobile services, spanning social media platforms, video streaming, application downloads, web services, cloud computing, online gaming, and music streaming. Access to the dataset is granted upon request through the official NetMob platform, which also provides geo-referenced metadata in GeoJSON format, enabling the mapping between geographical coordinates and the grid indices used in the traffic data.

While the NetMob dataset offers a detailed spatial characterization of traffic demand, the primary objective of this work is to study NS mechanisms at the BS level. To this end, additional information describing the cellular infrastructure deployed in the considered areas is required. Specifically, this study relies on data provided by the ANFR [23], which maintains a comprehensive public database of all cellular base stations deployed in France. The ANFR database includes the geographical location of each antenna, the deployed radio access technologies, and the identity of the mobile network operator owning the infrastructure.

Column	Description
tile_id	Unique identifier of the spatial grid cell
time_point_k	Traffic volume at time index k
Number of columns	One column per discrete time slot
Spatial resolution	100×100 m grid cells
Temporal resolution	15 minutes

Table 3.1: Structure of the NetMob traffic data files as provided in raw format.

Column	Description
id	Unique identifier of the antenna site
operator	Mobile network operator owning the site
technology	Radio access technology (e.g., 2G, 3G, 4G, 5G)
coordinates	Geographic location of the antenna (latitude, longitude)
status	Operational status of the antenna (e.g., active)

Table 3.2: Relevant fields extracted from the ANFR base station dataset.

The joint use of the NetMob and ANFR datasets enables the association of spatially aggregated traffic demand to the underlying cellular infrastructure. This integration is a prerequisite for evaluating NS strategies under realistic traffic conditions derived from real-world measurements. In the considered scenario, the main entities of the system are the spatial grid cells, the base stations deployed within the studied cities, the mobile network operators owning such infrastructure, and discrete time intervals of 15 minutes over which traffic demand is observed.

Tables 3.1 and 3.2 summarize the structure of the two datasets employed in this work.

3.2 Traffic representation

This section describes how raw traffic measurements are transformed in order to enable traffic characterization, clustering, and the evaluation of NS strategies. The objective is to define a traffic abstraction that is sufficiently expressive to capture recurrent temporal patterns, while remaining suitable for network-level modeling and planning analysis.

3.2.1 Traffic aggregation across services

The NetMob dataset provides traffic volumes for 68 distinct mobile service categories. However, from the perspective of the radio access network, traffic generated by different applications contributes jointly to the overall load experienced by base stations, independently of the specific service generating it. Since this work focuses on NS mechanisms and their impact on energy consumption and base station operation, a service-agnostic traffic representation is adopted.

Accordingly, traffic volumes associated with all service categories are aggregated to obtain a single traffic signal per spatial grid cell and time interval. This aggregated traffic demand reflects the total load imposed on the network infrastructure and is therefore

more relevant for operator-level planning and performance evaluation than service-specific traffic breakdowns.

Let $x_i(k)$ denote the aggregated traffic demand associated with spatial grid cell i at discrete time index k. The variable $x_i(k)$ represents the total traffic volume observed within the corresponding grid cell during the k-th 15-minute interval. The resulting traffic representation consists of one discrete-time traffic signal per grid cell, capturing the temporal evolution of demand over the observation period.

This abstraction does not distinguish between individual users, sessions, or packet-level events, and instead models traffic demand at an aggregate level, which is appropriate for the network-level analysis conducted in this study.

3.2.2 Mapping of spatial indices

The NetMob dataset represents spatial information through a unique tile identifier associated with each grid cell. In order to enable spatial operations, visualization, and the association of traffic demand with cellular infrastructure, tile identifiers are mapped to two-dimensional row and column indices corresponding to the underlying grid structure.

For each city, the NetMob dataset specifies the number of grid columns and rows defining the spatial extent of the considered area. Given this information, the tile identifier is converted into row and column indices assuming a row-major ordering of the grid. Specifically, for a given tile identifier *tile_id* and a known number of columns N_{cols}, the corresponding indices are computed as

$$row\_index = \left\lfloor \frac{tile\_id}{N_{\text{cols}}} \right\rfloor, \qquad col\_index = tile\_id \bmod N_{\text{cols}},$$

This transformation enables the reconstruction of the two-dimensional spatial layout of traffic demand from the one-dimensional tile indexing used in the raw dataset.

A similar mapping is required for the cellular infrastructure data in order to align base station locations with the same grid representation. For each city, the minimum latitude and longitude defining the spatial bounds of the NetMob grid are extracted from the provided metadata. Base station geographic coordinates are then projected onto the grid by computing their relative position within the spatial extent of the city. Let $\phi_{\min}$ and $\lambda_{\min}$ denote the minimum latitude and longitude of the grid, and let N_{rows} and N_{cols} be the number of grid rows and columns, respectively. The relative grid coordinates associated with a base station located at latitude ϕ and longitude λ are computed as

$$row\_coord = \frac{\phi - \phi_{\min}}{\Delta_\phi}, \qquad col\_coord = \frac{\lambda - \lambda_{\min}}{\Delta_\lambda},$$

where Δ_ϕ and Δ_λ denote the latitude and longitude span of a single grid cell, respectively.

Unlike grid traffic indices, base station coordinates are retained as continuous values in order to enable distance-based association between grid cells and base stations.

This coordinate transformation ensures that both traffic demand and base station locations are expressed within a common spatial reference system, enabling consistent traffic assignment and infrastructure-aware analysis across different cities.

3.2.3 Infrastructure-aware filtering and operator selection

Since the objective of this work is to evaluate NS strategies at the radio access network level, traffic representation is constrained by the characteristics of the underlying cellular infrastructure. To ensure consistency and comparability across operators and locations, only base stations employing 4G technology and reported as operational are considered in the analysis. Furthermore, the study focuses on the four major mobile network operators active in France, which collectively account for the vast majority of nationwide traffic and infrastructure deployment.

3.2.4 Temporal windowing and traffic assignment to base stations

Traffic demand is observed over a temporal horizon of 28 consecutive days. In order to ensure comparable observation windows across operators and to avoid overlapping measurement periods, the dataset is partitioned into four non-overlapping weeks, and one week of traffic measurements is associated with each operator. This choice provides a balanced amount of data per operator while capturing the daily regularities of traffic demand.

For each operator, spatially aggregated grid traffic is mapped to the cellular infrastructure by assigning each grid cell to its nearest base station belonging to the same operator. This nearest-neighbor association is equivalent to constructing a Voronoi tessellation of the base station locations on an operator-wise basis, where each grid cell is assigned to the Voronoi region of the closest site. The traffic associated with a base station is then obtained by aggregating, at each time interval, the traffic of all grid cells assigned to it. This operator-specific mapping ensures that traffic demand is consistently associated with the infrastructure that would realistically serve it under normal operational conditions.

As a result, for each operator and for each of its base stations in the considered area, a one-week traffic time series is obtained, representing the evolution of aggregated demand over the selected observation window. These time series constitute the input for the traffic profiling and clustering analysis, and they are subsequently used to evaluate the performance of NS strategies.

3.3 Network Sharing model

The NS model and simulator adopted in this work build upon the framework originally developed by Maoquan Ni in [24]. While the underlying modeling assumptions and performance metrics follow the same baseline, we substantially revised the simulator implementation to enable *large-scale* evaluation across many operator pairs and base station combinations. In the original implementation, simulations were typically executed in small batches (e.g., 10×10 combinations) due to computational overhead. Here, we re-engineered the pipeline to support full-grid experiments by improving runtime efficiency and scalability.

These changes significantly reduce execution time and allow running extensive scenario sweeps while preserving the original model logic.

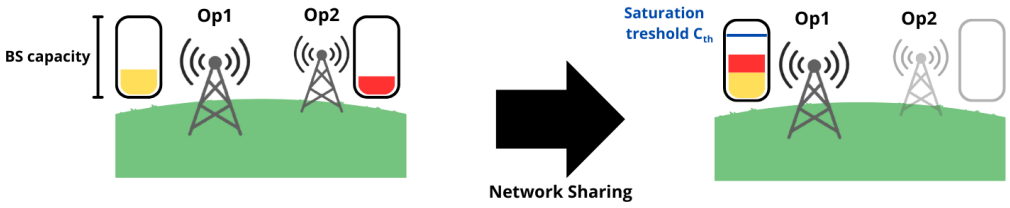


Figure 3.1: Schematic illustration of the pairwise network sharing model.

This section introduces the resulting NS model used throughout the thesis. The objective is to describe how traffic demand from multiple operators can be served through shared RAN infrastructure, and how sharing decisions affect base station operation, energy consumption, and switching behavior. The model targets planning-level analysis and therefore abstracts away radio-level details. Figure 3.1 provides a high-level overview of the considered NS mechanism. At each time interval, traffic can be offloaded between co-located base stations of two operators, subject to a saturation constraint, enabling traffic consolidation and potential switch-off opportunities.

3.3.1 Sharing scenario and assumptions

In this study, NS is evaluated on a *pairwise* basis: each experiment considers one pair of operators, and sharing is enabled only between the infrastructures owned by the two operators in the pair. Each operator owns and operates a set of base stations, as described in the previous sections.

NS is implemented through traffic offloading between co-located base stations belonging to different operators. Co-location is defined geometrically: two BSs of different operators are considered co-located if the Euclidean distance between their geographical coordinates is at most 50 m. Under this condition, we assume that the corresponding coverage areas are largely overlapping, since such short separations typically arise from deployments on the same site (e.g., rooftop or street-level locations) or in immediate proximity. Consequently, NS is considered feasible only among co-located base station pairs, while no sharing is allowed between non co-located sites.

The following assumptions are made throughout the analysis:

- Sharing decisions are taken at discrete time intervals, aligned with the traffic observation granularity.
- Radio-level effects such as interference, scheduling, and user association dynamics are not explicitly modeled.
- When NS is applied between a pair of co-located base stations, the base station that is switched off is assumed to consume negligible power.

These assumptions allow focusing on the impact of NS on base station utilization, energy consumption, and operational dynamics, while ensuring that sharing is only considered between infrastructures that realistically serve comparable geographical areas.

3.3.2 Base station capacity and saturation threshold

In this study, the absolute traffic volume served by each base station cannot be directly inferred from the NetMob dataset. Traffic values provided by NetMob are intentionally scaled by an unknown factor in order to preserve operator confidentiality and therefore do not represent the actual throughput handled by the network in physical units. Nevertheless, the dataset preserves relative traffic variations over time and across spatial locations, which are sufficient to compare load levels among base stations and to study NS dynamics.

Based on this observation, the base station capacity is defined in a relative and data-driven manner. For each base station b, a reference traffic time series is first obtained over the observation window assigned to the corresponding operator. To mitigate the impact of anomalous traffic spikes that could bias capacity estimation, the weekly traffic series is filtered using a three-sigma criterion[1], removing samples that deviate significantly from the typical traffic behavior.

Let $\widetilde{L}_b(t)$ denote the filtered traffic load of base station b at time interval t. The maximum observed traffic load over the filtered time series is then used to define the base station capacity. Specifically, the capacity C_b is set conservatively as

$$C_b = 0.9 \cdot \max_t \widetilde{\ell}_b(t), \tag{3.1}$$

where the factor 0.9 introduces a safety margin that accounts for unmodeled variability and avoids operating base stations at extreme utilization levels.

At each time interval t, let $\ell_b(t)$ denote the traffic load assigned to base station b after NS decisions. A saturation threshold $\theta \in (0,1]$ is introduced to prevent overload conditions and ensure stable operation. The admissible operating condition is given by

$$\ell_b(t) \leq \theta C_b. \tag{3.2}$$

This relative capacity model allows the evaluation of NS strategies based on comparative traffic loads, without requiring knowledge of absolute traffic volumes, while remaining consistent with the anonymized nature of the dataset.

3.3.3 Base station operational states

Each base station can operate in one of two states: active or switched off. A base station is considered active when it is serving traffic, either from its own operator or from shared traffic. Conversely, a base station can be switched off when its assigned traffic load is zero.

Let $s_b(t) \in \{0,1\}$ denote the operational state of base station b at time t, where $s_b(t) = 1$ indicates an active state and $s_b(t) = 0$ denotes a switched-off state. NS may enable traffic consolidation onto a reduced number of active base stations, thereby allowing other sites to be temporarily switched off.

[1]The three-sigma criterion removes samples lying more than three standard deviations away from the mean.

Parameter	Meaning	Value
N_{TX}	Number of transceivers	6
$P_{\max}$	Maximum RF output power	20 W
P_0	Power at zero load (per transceiver)	84 W
δ_p	Load-dependent slope	2.8

Table 3.3: RRH power model parameters adopted in this work.

The sequence of state transitions over time directly impacts both energy consumption and switching frequency, which are key performance metrics considered in this study.

3.3.4 Energy consumption model

The energy consumption of base stations is modeled using a load-dependent abstraction derived from established power models in the literature. In particular, the adopted model follows the approach proposed in [25], which describes a near-linear relationship between the Random Forest (RF) output power of a base station and its overall power consumption. This formulation is especially suitable for Remote Radio Head (RRH) architectures and has been widely used for energy-efficiency evaluations at the network level.

According to this model, the instantaneous power consumption $P_b^{\mathrm{in}}(t)$ of base station b at time interval t is expressed as

$$P_b^{\mathrm{in}}(t) = N_{\mathrm{TRX}} \left(P_0 + \delta_p \cdot P_b^{\mathrm{out}}(t) \right) \tag{3.3}$$

where N_{TRX} denotes the number of transceiver modules, P_0 represents the power consumption at zero RF output power, and δ_p is the slope of the load-dependent power component.

The RF output power $P_b^{\mathrm{out}}(t)$ is assumed to be proportional to the traffic load relative to the base station capacity. Specifically, it is modeled as

$$P_b^{\mathrm{out}}(t) = \frac{\ell_b(t)}{C_b} \cdot P_{\max} \tag{3.4}$$

where $\ell_b(t)$ is the traffic load assigned to base station b at time t, C_b is the reference capacity defined in 3.3.2 and $P_{\max}$ denotes the maximum RF output power.

Adopting the parameter values in Table 3.3 into eq.3.3 yields a linear relationship between base station power consumption and the normalized traffic load, enabling the evaluation of energy usage under varying traffic and NS conditions. While this model does not capture fine-grained radio effects or hardware-specific optimizations, it provides a realistic and analytically tractable approximation of base station energy behavior, and is therefore well suited for comparative evaluation of NS strategies and their associated energy savings.

3.4 Performance metrics

This section defines the performance metrics used to evaluate the impact of NS on energy efficiency and operational dynamics. The selected metrics are designed to capture both the potential energy benefits of consolidating traffic onto fewer active base stations and the operational cost associated with frequent state transitions. All metrics are computed on a pairwise basis, i.e., for each co-located base station pair belonging to a given operator pair.

3.4.1 Energy savings

Energy savings quantify the relative reduction in energy consumption achieved by NS with respect to a baseline scenario without sharing, in which both base stations of a co-located pair remain active. Let $E_{b_i}^{\mathrm{NS}}$ and $E_{b_j}^{\mathrm{NS}}$ denote the energy consumed by base stations b_i and b_j under NS, and let $E_{b_i}^{\mathrm{BL}}$ and $E_{b_j}^{\mathrm{BL}}$ denote the corresponding energy consumption in the baseline case.

The energy savings metric for a co-located base station pair (b_i, b_j) is defined as

$$S_E = 1 - \frac{E_{b_i}^{\mathrm{NS}} + E_{b_j}^{\mathrm{NS}}}{E_{b_i}^{\mathrm{BL}} + E_{b_j}^{\mathrm{BL}}}.$$

This metric represents the percentage reduction in energy consumption enabled by NS and constitutes the primary indicator of energy efficiency gains in this study.

3.4.2 Duty cycle

The duty cycle measures the fraction of time during which NS effectively enables traffic consolidation by switching off one base station in a co-located pair. Let K denote the total number of discrete time intervals in the observation window, and let $s_b(k) \in \{0,1\}$ denote the operational state of base station b at time interval k, where $s_b(k) = 1$ indicates that the base station is active and $s_b(k) = 0$ indicates that it is switched off.

The duty cycle is defined as

$$D = \frac{1}{K} \sum_{k=1}^{K} (1 - s_b(k)),$$

i.e., the fraction of time intervals in which the base station is switched off due to NS. This metric provides an intuitive measure of how frequently NS is applied in practice and allows energy savings to be directly related to the duration of switch-off periods.

3.4.3 Switching frequency

Switching frequency captures the operational dynamics induced by NS by quantifying how often a base station transitions between active and switched-off states. Frequent switching may introduce additional operational overhead and potential wear-and-tear, and therefore represents an important complement to purely energy-based metrics.

Let $s_b(k) \in \{0,1\}$ denote the operational state of the base station at time interval k. The total number of state transitions over the observation window is computed as

$$N_{\text{tr}} = \sum_{k=2}^{K} |s_b(k) - s_b(k-1)|,$$

which counts both ON→OFF and OFF→ON transitions. Since a complete switching cycle consists of two transitions, the number of full ON–OFF–ON switching cycles is given by $N_{\text{tr}}/2$.

To obtain a normalized and interpretable metric, switching frequency is expressed as the average number of switching cycles per day. In this study, each operator is associated with one week of traffic data, hence the daily switching frequency is defined as

$$F = \frac{N_{\text{tr}}}{2 \times 7}.$$

This metric provides a direct view of the average daily switching activity induced by NS and highlights potential trade-offs between energy savings and operational stability.

Chapter 4

Data-Driven Network Sharing Analysis

4.1 Overview of the proposed approach

This chapter introduces a data-driven methodology to characterize BS traffic dynamics and to support an assessment of NS feasibility. The key idea is to abstract raw traffic traces into a finite set of traffic profiles (or traffic types) and to evaluate whether these profiles can be used to anticipate NS-relevant operating conditions.

4.2 Weekly traffic profile definition and preprocessing

Let $x_i(t)$ denote the traffic demand associated with BS i at discrete time step t, where the sampling interval is fixed across the dataset. To capture both intra-day dynamics and weekday/weekend contrasts, we define for each BS a *weekly traffic profile* as the concatenation of traffic values over a full week:

$$\mathbf{x}_i = [x_i(1), x_i(2), \ldots, x_i(T)],$$

where T is the number of samples per week.

Basic preprocessing is applied to improve comparability across BSs and to reduce the impact of artifacts: (i) when clustering focuses on shape, profiles are normalized to remove scale effects; (ii) when using feature-space representations, features are standardized to have comparable scale. These steps ensure that clustering reflects genuine behavioral differences rather than magnitude dominance or measurement noise.

4.3 Traffic representations for profiling

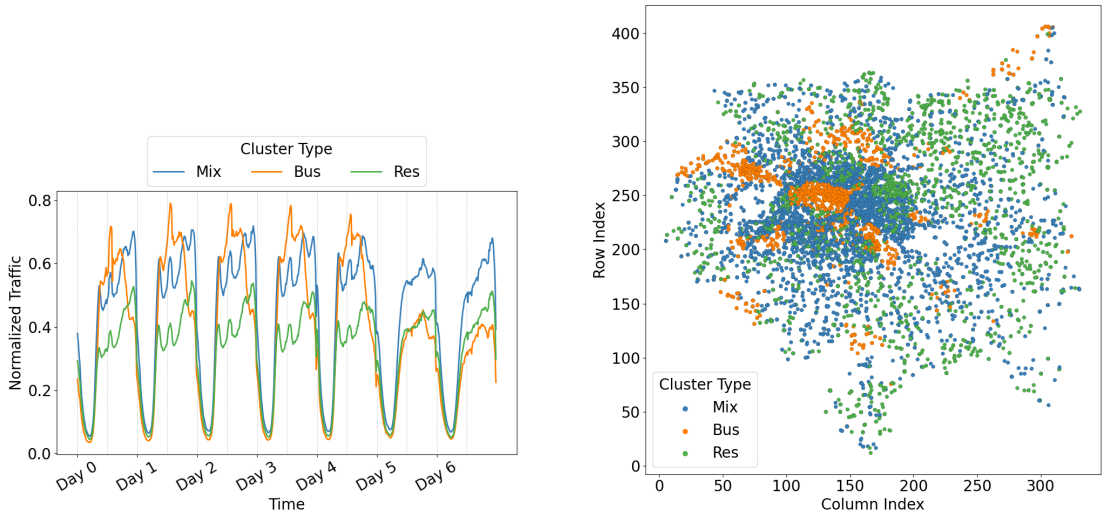
Traffic profiles can be represented in different ways depending on which aspect of demand is considered most relevant. Since our objective is to identify traffic types that are informative for NS feasibility, we compare multiple representations that emphasize complementary properties of the weekly profile.

4.3.1 Traffic-shape representation

The shape representation captures the temporal evolution of the weekly profile while reducing sensitivity to absolute magnitude. For each BS i, we define the normalized profile:

$$\tilde{x}_i(t) = \frac{x_i(t)}{\max_t x_i(t)},$$

Clustering in this space emphasizes similarities in daily/weekly cycles, peak timing, and relative night–day contrast.



(a) Mean normalized weekly profiles for the selected k-means clustering ($k = 3$).

(b) Spatial distribution of traffic-shape cluster entries.

Figure 4.1: Traffic types obtained from normalized weekly traffic-shape clustering

Figure 4.1 summarizes the outcome of the traffic-shape clustering. In Fig. 4.1a, the centroid (mean) normalized weekly profiles clearly differentiate three temporal behaviours. The **Business** (Bus) cluster is characterized by a pronounced daytime peak and reduced activity during evenings and night hours, the **Residential** (Res) cluster exhibits an evening-dominant profile, and the **Mixed-use** (Mix) cluster presents a comparatively flatter demand distribution over the day. Figure 4.1b maps the corresponding cluster assignments over the Paris area, highlighting spatial patterns that are consistent with the functional interpretation of the traffic types and their typical urban locations (e.g., central/business districts for Bus and residential belts for Res, with Mix distributed across heterogeneous areas). These patterns motivate the use of traffic types as compact descriptors of heterogeneous load dynamics that are relevant to downstream NS feasibility analysis.

Semantic validation To evaluate the semantic consistency of the identified clusters, we analyze the Points of Interest (POIs) within the coverage area of each BS, following

Table 4.1: Macro POI composition by traffic cluster.

Cluster	Entertainment	Office	Residential
Mixed-use (Mix)	29.24%	9.62%	61.14%
Business (Bus)	40.17%	14.89%	44.94%
Residential (Res)	22.17%	7.63%	70.19%

the methodology established by [26]. POI data is retrieved from OpenStreetMap [27] and categorized into three macro-classes: **Entertainment**, **Office**, and **Residential**. For each BS, the POI counts per category are normalized using Min–Max scaling, applied independently to each type. We then compute the average normalized composition for each cluster to derive the relative percentage contribution of each POI class.

As shown in Table 4.1, the results confirm a strong alignment between observed traffic patterns and land-use characteristics. The Business (Bus) cluster exhibits the highest concentration of Office POIs, while the Residential (Res) cluster is characterized by the largest share of Residential POIs. The Mixed-use (Mix) cluster occupies an intermediate position, mirroring its hybrid temporal behavior.

These distributions reflect the intrinsic spatial organization of the urban environment. While Residential POIs are spatially pervasive, appearing significantly across all clusters, their dominance is most pronounced in the Res cluster. In contrast, Office POIs are geographically concentrated, explaining their markedly higher proportion in the Bus cluster and limited presence elsewhere. The Entertainment category demonstrates a more balanced distribution, occurring more frequently in mixed or commercial zones than in purely residential areas. Ultimately, this POI composition validates the semantic interpretation of the clusters derived from the traffic data.

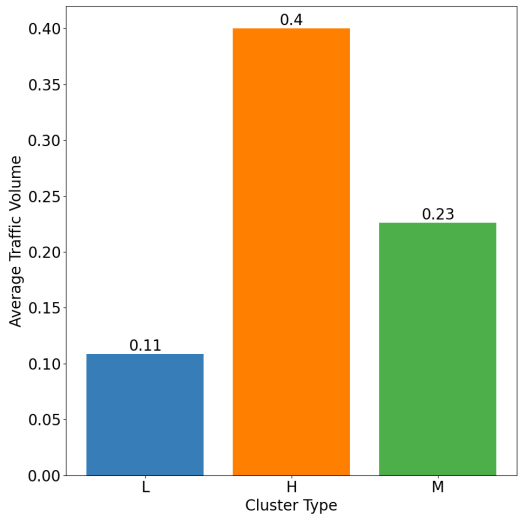
4.3.2 Average traffic volume representation

To isolate the effect of magnitude, we also consider a scalar representation based on the average weekly traffic volume:

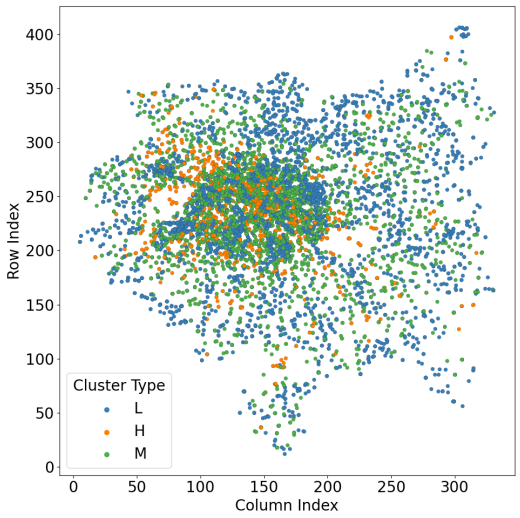
$$\bar{x}_i = \frac{1}{T} \sum_{t=1}^{T} x_i(t).$$

This representation groups BSs primarily by load intensity and is relevant because NS feasibility is directly constrained by available capacity.

Figure 4.2 summarizes the outcome of the average-volume clustering. Figure 4.2a reports the mean (normalized) traffic volume of the three clusters, which naturally separates the population into low-load (`L`), medium-load (`M`), and high-load (`H`) regimes. Figure 4.2b shows the corresponding spatial distribution across the metropolitan area, highlighting how these load regimes are geographically interleaved rather than strictly segregated. This volume-based partition is directly relevant to NS feasibility analysis, since higher average load implies reduced residual capacity and thus tighter headroom for offloading decisions.



(a) Mean (normalized) traffic volume per cluster.



(b) Spatial distribution of average-volume clusters.

Figure 4.2: Average-volume clustering results. Cluster labels (`L`, `M`, `H`).

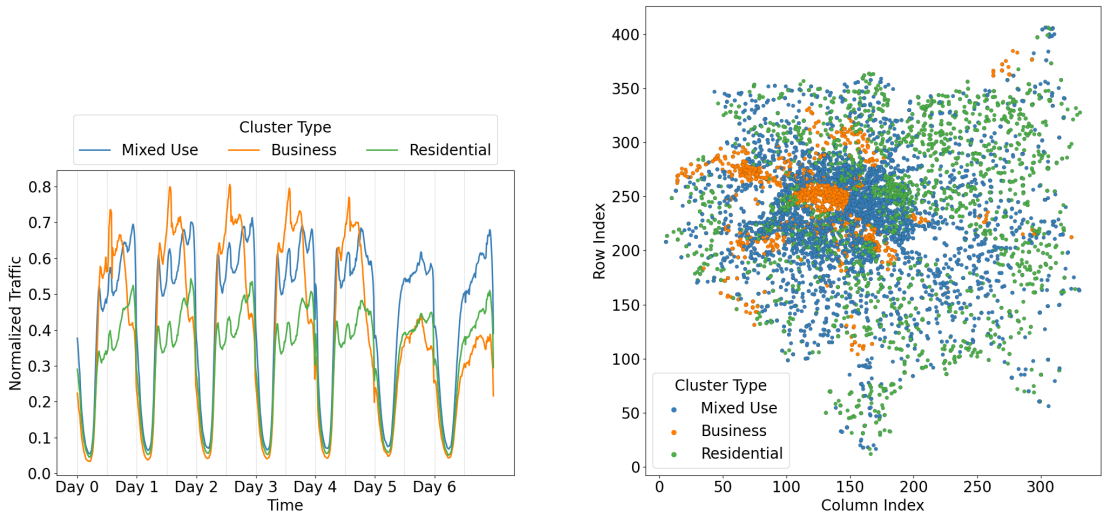
4.3.3 Engineered feature representation

As an alternative to clustering raw profiles, we map each weekly profile to a feature vector

$$\boldsymbol{\phi}_i = \phi(\mathbf{x}_i),$$

designed to summarize key behavioural properties such as variability, burstiness, peak/idle structure, weekday/weekend contrasts, and periodicity. Time-domain descriptors capture dispersion and burst structure, while frequency-domain descriptors capture the strength and location of periodic components. The complete list of engineered features and their definitions is reported in the appendix A, while this chapter focuses on how they are used for profiling and model selection.

Figure 4.3 reports the outcomes of the feature-based traffic-pattern clustering. The centroid profiles in Fig. 4.3a reveal three distinct temporal behaviours. The Business cluster is characterized by pronounced weekday daytime activity and a marked drop during night hours, the Residential cluster exhibits comparatively stronger evening demand and smoother weekday–weekend variations, and the Mixed-use cluster presents a more balanced load distribution across the week, lying between the previous two extremes. Figure 4.3b shows the corresponding spatial distribution of cluster assignments, highlighting coherent urban patterns: Business-type sites are concentrated in central and activity-dense areas, Residential-type sites are more prevalent in peripheral zones, while Mixed-use sites are distributed across heterogeneous neighbourhoods. Overall, these clusters provide an interpretable set of traffic types that compactly summarize heterogeneous demand dynamics and can be leveraged in the subsequent NS feasibility and prediction analyses.



(a) Mean normalized weekly traffic profiles for each cluster.

(b) Spatial distribution of feature-based cluster assignments.

Figure 4.3: Feature-based traffic-pattern clustering results.

Adequacy of the engineered representation. We validate the expressiveness of the engineered feature space through a supervised sanity check: NS-related KPIs are regressed from the pairwise descriptors using a Random Forest model and evaluated via cross-validation.

Since the KPIs are defined for co-located pairs, we construct pairwise descriptors as in (4.1), which capture both the overall load level (mean term) and the mismatch/compatibility (difference term).

$$\phi_{ij} \triangleq \left[\frac{\phi_i + \phi_j}{2}, \; \left| \phi_i - \phi_j \right| \right]. \tag{4.1}$$

This construction captures both the overall load level (mean terms) and the traffic mismatch/compatibility between the two sites, while keeping the representation compact (from 672 weekly samples to 18 BS-level descriptors, and their pairwise aggregation).

To interpret which traffic traits drive each KPI, we analyze Random Forest feature importances, computed as the normalized average reduction in regression impurity (variance/MSE) produced by splits on each feature across all trees, weighted by the number of samples reaching each split.

As shown in Fig. 4.4, the resulting R^2 values are consistently high across KPIs, indicating that the pairwise engineered representation retains the dominant traffic characteristics relevant to NS performance. Figure 4.5 further highlights a clear structure in the drivers of the different KPIs: **Energy Saved (%)** and **Duty Cycle** exhibit highly aligned importance rankings, dominated by descriptors capturing peak-to-idle structure, load concentration and burstiness (e.g., Absolute Max Traffic_diff, Area Under Curve_mean, Mean Traffic Load_mean, Burstiness Index_mean). In contrast, **Switching Frequency** shows

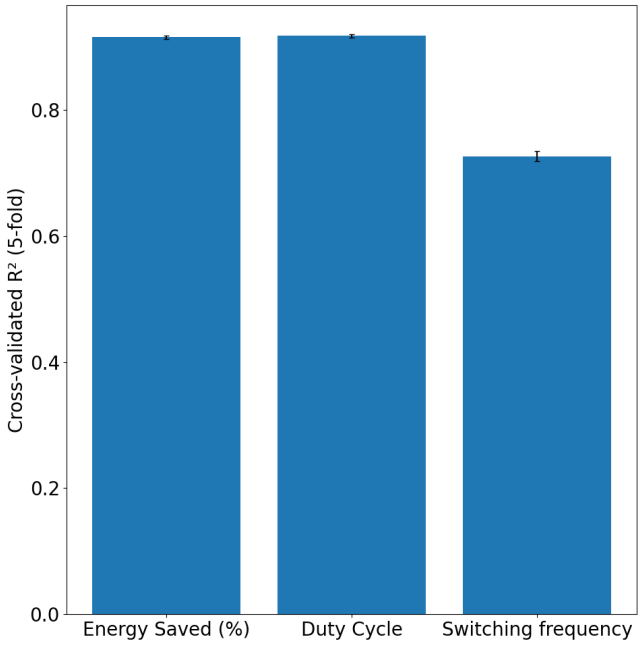


Figure 4.4: Cross-validated R^2 performance of Random Forest regressors predicting NS KPIs

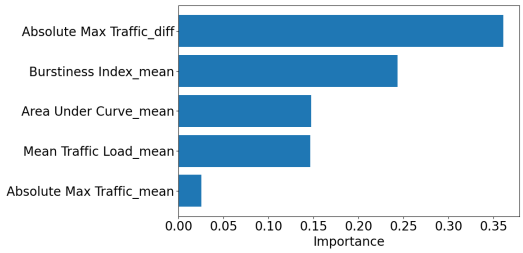
a distinct signature, with higher relevance of variability and time/frequency-structure indicators (e.g., Variance of Weekday Load_mean and Spectral Centroid_mean), consistent with the fact that frequent transitions are mainly induced by irregular fluctuations rather than by average load alone. Overall, the concentration of importance on a small subset of descriptors suggests that the regression task could be further simplified by retaining only the Top-k features, reducing dimensionality and complexity with limited impact on performance.

4.4 Clustering methods considered

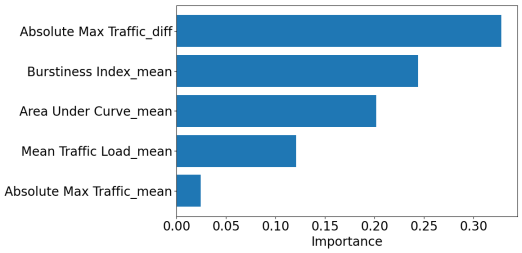
We apply multiple unsupervised clustering techniques to the representations described in Section 4.3. Let $\mathbf{z}_i$ denote the representation of BS i (normalized shape vector, scalar volume, or feature vector). The objective is to partition the set $\{\mathbf{z}_i\}_{i=1}^{N}$ into k groups such that samples in the same group exhibit similar traffic behaviour.

4.4.1 k-means clustering

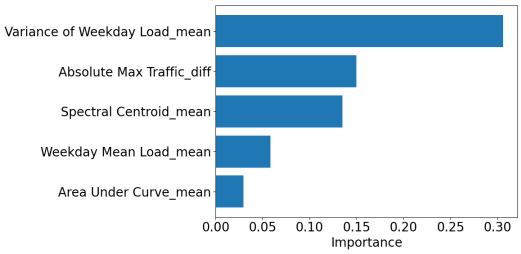
When $\mathbf{z}_i$ lies in an Euclidean space (e.g., engineered feature vectors or normalized profile vectors compared with Euclidean distance), we consider k-means clustering. Given a target number of clusters k, k-means seeks a partition $\{C_1, \ldots, C_k\}$ and centroids



(a) Energy Saved (%)



(b) Duty Cycle



(c) Switching frequency

Figure 4.5: Top-ranked feature importances.

$\{\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k\}$ that minimize the within-cluster sum of squares:

$$\min_{\{C_m\},\{\boldsymbol{\mu}_m\}} \sum_{m=1}^{k} \sum_{i \in C_m} \|\mathbf{z}_i - \boldsymbol{\mu}_m\|^2.$$

The algorithm alternates between (i) assigning each sample to the nearest centroid and (ii) updating centroids as the mean of assigned samples, until convergence. k-means is computationally efficient and produces compact clusters, but it implicitly favours roughly spherical clusters and can be sensitive to initialization; for this reason, multiple random initializations are typically used.

4.4.2 Hierarchical agglomerative clustering

To capture potentially non-spherical structures and to explore multiple resolutions, we consider Hierarchical Agglomerative Clustering (HAC). HAC starts with N singleton clusters and iteratively merges the two closest clusters until only one cluster remains. The notion of distance between clusters is defined by a linkage criterion. Given two clusters A and B, common linkages include:

- **Single linkage:** $d(A, B) = \min_{i \in A,\, j \in B} d(\mathbf{z}_i, \mathbf{z}_j)$,
- **Complete linkage:** $d(A, B) = \max_{i \in A,\, j \in B} d(\mathbf{z}_i, \mathbf{z}_j)$,
- **Average linkage:** $d(A, B) = \frac{1}{|A||B|} \sum_{i \in A} \sum_{j \in B} d(\mathbf{z}_i, \mathbf{z}_j)$,

- **Ward linkage:** merges clusters that produce the smallest increase in within-cluster variance (typically used with Euclidean distance).

The result is a dendrogram that can be cut at a chosen level to obtain k clusters. HAC can work with generic distance measures (not only Euclidean), which is useful when using time-series distances such as Dynamic Time Warping (DTW).

4.4.3 Other clustering variants

Depending on the representation, additional clustering approaches can be considered (e.g., spectral clustering in feature space). In all cases, the goal remains to obtain traffic types that are (i) well-separated according to internal validity criteria, (ii) not trivially imbalanced, and (iii) meaningful for downstream NS-related inference.

4.5 Clustering model selection and validation

Selecting an appropriate clustering configuration involves choosing both (i) the clustering method and (ii) the number of clusters k (when applicable). Since no ground-truth labels are available, we rely on internal validation metrics and distribution checks to compare candidate partitions. Let $\mathcal{C} = \{C_1, \ldots, C_k\}$ denote a partition of the N samples.

4.5.1 Silhouette score

The silhouette score measures how well each sample fits within its assigned cluster compared to the closest alternative cluster. For a sample i, let $a(i)$ be the average distance from i to all other points in its own cluster, and let $b(i)$ be the minimum (over clusters $C \neq C(i)$) of the average distance from i to points in cluster C. The silhouette of i is:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i),\, b(i)\}} \in [-1,1].$$

Values close to 1 indicate that i is well matched to its cluster and far from neighbouring clusters, values near 0 indicate overlap, and negative values suggest possible misassignment. The overall silhouette score is the mean of $s(i)$ over all samples; higher values indicate more compact and better separated clusters.

4.5.2 Davies–Bouldin index

The Davies–Bouldin index (DBI) quantifies average similarity between each cluster and its most similar (closest) cluster, combining within-cluster scatter and between-cluster separation. Let $\boldsymbol{\mu}_m$ be the centroid of cluster C_m and define the within-cluster scatter:

$$S_m = \frac{1}{|C_m|} \sum_{i \in C_m} \|\mathbf{z}_i - \boldsymbol{\mu}_m\|.$$

For two clusters m and n, define:

$$R_{mn} = \frac{S_m + S_n}{\|\boldsymbol{\mu}_m - \boldsymbol{\mu}_n\|}.$$

Then the DBI is:

$$\mathrm{DBI} = \frac{1}{k} \sum_{m=1}^{k} \max_{n \neq m} R_{mn}.$$

Lower DBI indicates clusters that are compact (small S_m) and well separated (large centroid distances), hence better partitions.

4.5.3 Cluster-size balance and skewness

Internal validity metrics may favor partitions that are compact/separated but operationally unhelpful if they produce highly imbalanced cluster sizes (e.g., one dominant cluster and many tiny clusters). To discourage such degenerate solutions, we assess the distribution of cluster sizes $\{n_m = |C_m|\}_{m=1}^{k}$.

A simple indicator is the skewness of the cluster-size distribution:

$$\mathrm{Skew}(n) = \frac{\frac{1}{k}\sum_{m=1}^{k}(n_m - \bar{n})^3}{\left(\frac{1}{k}\sum_{m=1}^{k}(n_m - \bar{n})^2\right)^{3/2}}, \qquad \bar{n} = \frac{1}{k}\sum_{m=1}^{k} n_m.$$

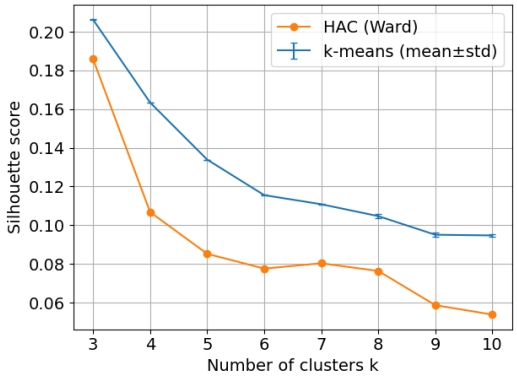
High positive skewness indicates that a small number of clusters contain most samples, which may reduce interpretability and limit the usefulness of profile combinations. In practice, skewness is used as a complementary selection signal to ensure that the chosen clustering yields a meaningful variety of traffic types.

4.5.4 Selection rationale

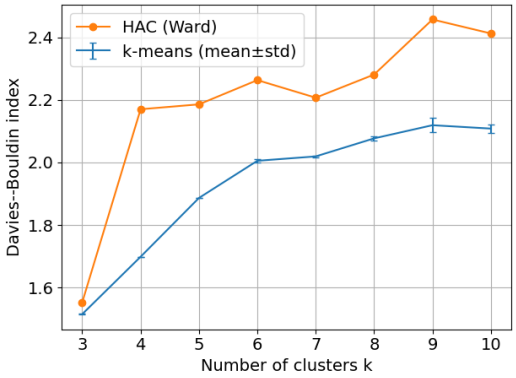
The final clustering configuration is selected by jointly considering (i) compactness and separation, quantified by the silhouette score (higher is better) and the DBI (lower is better), and (ii) the balance of the resulting partition, assessed through the skewness of the cluster-size distribution (values closer to zero indicate more balanced cluster sizes). Since no ground-truth labels are available, these internal criteria provide a principled basis to compare candidate methods and values of k. In the following, we evaluate k-means and HAC on the traffic-shape representation defined in Section 4.3.1, varying the number of clusters in the range $k \in \{3, \ldots, 10\}$.

Figure 4.6 highlights a consistent advantage of k-means over HAC on the traffic-shape space. Across all tested values of k, k-means achieves higher silhouette scores (Fig. 4.6a) and lower DBI (Fig. 4.6b), indicating clusters that are simultaneously more compact and better separated. Regarding the choice of k, both separation-based criteria point to $k = 3$, which yields the highest silhouette and the lowest DBI among the tested configurations. As expected, increasing k tends to improve size balance (Fig. 4.6c), with skewness moving closer to zero at $k = 4$, but this comes at the cost of a noticeable reduction in silhouette and an increase in DBI.

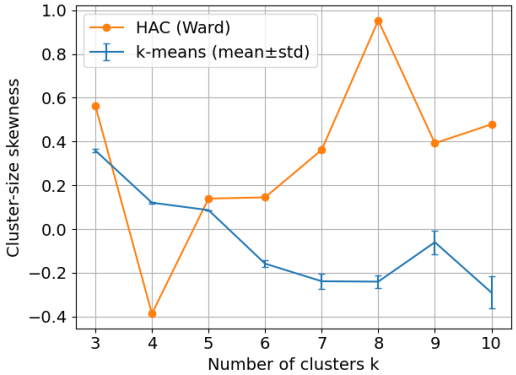
Overall, we prioritize cluster separation and stability, since the resulting traffic types are used downstream to characterize heterogeneous demand dynamics and to support NS-related analyses. Therefore, we adopt k-means with $k = 3$ as the final clustering configuration for the traffic-shape representation in the remainder of this work.



(a) Silhouette vs. k.



(b) DBI vs. k.



(c) Size skewness vs. k.

Figure 4.6: Comparison of k-means and hierarchical clustering (HAC) on normalized weekly traffic-shape profiles for $k \in \{3, \ldots, 10\}$.

4.6 Joint distribution of traffic-shape and volume classes

The traffic-shape clustering (Bus/Res/Mix) and the average-volume clustering (L/M/H) provide two complementary views of BS demand. We first inspect how the two classifications co-occur across the population of sites.

Figure 4.7 reports the contingency matrix of BS counts for the combination of traffic-shape and average-volume labels. The distribution is clearly non-uniform, indicating that temporal behaviour and load intensity are not independent. In particular, low-volume class (L) contains the largest number of sites across all traffic-shape types, and it is especially prevalent for Residential and Mixed-use patterns. Conversely, high-volume (H) sites are comparatively fewer, while medium-volume (M) sites are concentrated mainly in the Mixed-use and Residential groups. Overall, this cross-tabulation provides a compact summary of how traffic types and load regimes co-exist in the dataset and motivates the pair-level stratification introduced next.

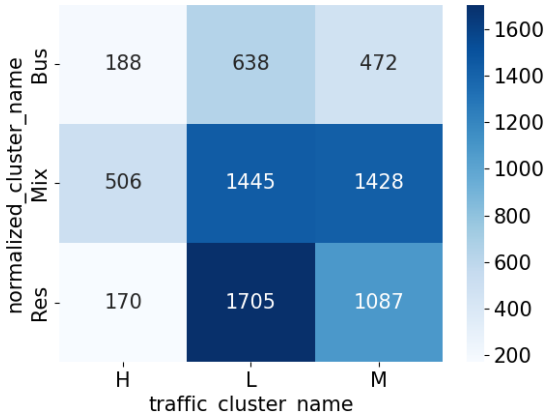


Figure 4.7: Contingency matrix of BS assignments across traffic-shape classes (rows) and average-volume classes (columns).

4.7 Joint label combinations for co-located pairs

We now lift BS-level labels to the level of co-located pairs used in the NS evaluation. Let (i, j) denote a co-located pair. Each element of the pair is associated with a pattern label and a volume label, thus inducing a *pattern-pair* class and a *volume-pair* class.

For traffic-shape types, we define the pattern-pair label as

$$c_{ij}^{\text{pat}} = (c_i^{\text{pat}}, c_j^{\text{pat}}),$$

which yields combinations such as Bus–Bus, Res–Res, or Mix–Res. Similarly, for average volume we define

$$c_{ij}^{\text{vol}} = (c_i^{\text{vol}}, c_j^{\text{vol}}),$$

which yields combinations such as H–H, M–L, or L–L.

Finally, we define a *joint pair label* as

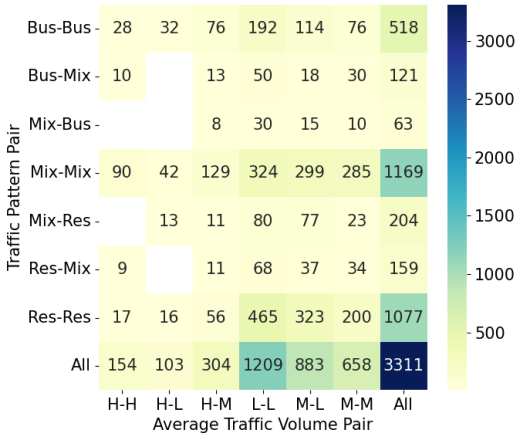
$$c_{ij}^{\text{joint}} = (c_{ij}^{\text{pat}}, c_{ij}^{\text{vol}}),$$

which enables a fine-grained stratification of pairs simultaneously by temporal behaviour and load regime. This representation will be used to understand how NS-related KPIs vary across different pair classes.

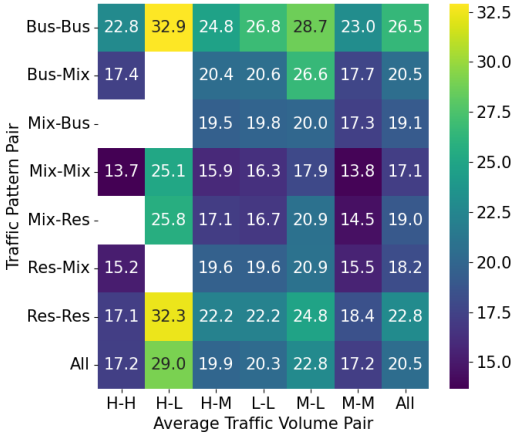
4.8 KPI stratification by joint pair labels

To assess whether the proposed profiling is informative for NS feasibility and performance, we analyze how NS-related KPIs vary across pair classes. For each joint class $(c_{ij}^{\text{pat}}, c_{ij}^{\text{vol}})$ we aggregate KPI values over all pairs belonging to that class, obtaining a matrix indexed by pattern-pair labels (rows) and volume-pair labels (columns). This visualization summarizes how KPIs change across temporal compatibility regimes and typical load-intensity regimes.

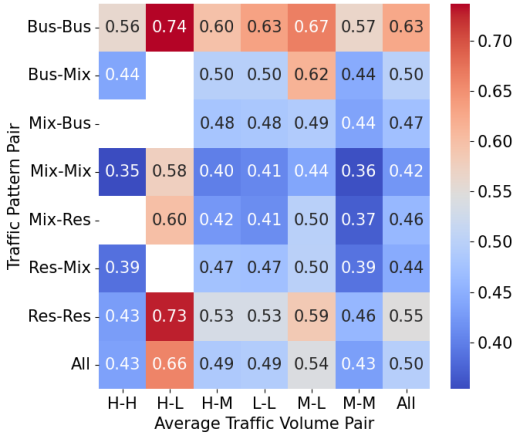
Some joint combinations occur rarely, leading to unreliable averages. To avoid drawing conclusions from insufficient support, we filter out joint classes whose number of samples falls below a minimum threshold. Filtered or non-observed combinations are shown as blank (white) cells in Figure 4.8a.



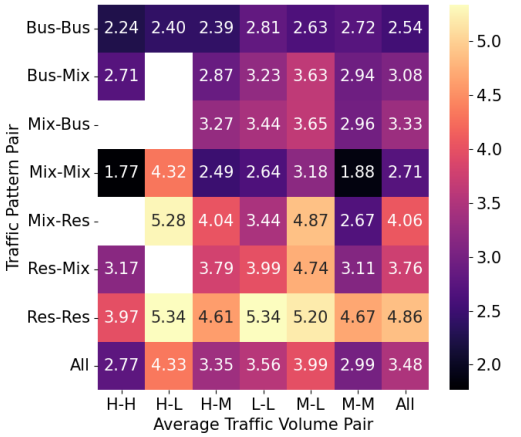
(a) Number of co-located pairs per class.



(b) Average energy-saving indicator for each class.



(c) Average duty-cycle indicator for each class.



(d) Average frequency of state transitions for each class.

Figure 4.8: KPI stratification by joint pair labels.

Turning to the KPI heatmaps, the energy-saving indicator (Fig. 4.8b) reveals a dominant volume-driven ordering. The most distinctive contrast is associated with load complementarity, with H–L volume pairs achieving the largest average savings across most pattern combinations. This is consistent with the intuition that sharing opportunities increase when one sector can absorb traffic while the other can be switched or offloaded. When this favorable volume pairing coincides with regular temporal behaviour (notably

Bus–Bus pattern pairs), the corresponding cross-combination reaches the best performance in terms of energy saving. Conversely, homogeneous high- and medium-load volume pairs (H–H and M–M) yield very limited potential, indicating that sustained load on both sectors constrains the ability to consolidate traffic without violating capacity margins. Compared to these constrained regimes, L–L pairs exhibit noticeably higher savings, suggesting that low-demand co-locations provide more headroom for switching/offloading actions.

The duty-cycle heatmap (Fig. 4.8c) provides an operational complement to the energy-saving indicator. Low duty-cycle values identify classes where switching actions are rarely feasible or triggered, which typically aligns with reduced NS potential in Fig. 4.8b. In particular, H–H (and often M–M) combinations tend to exhibit low duty cycles, reflecting the limited time windows in which traffic can be safely consolidated. By contrast, classes involving at least one low-load sector (e.g., H–L or L–L) generally allow higher duty cycles, i.e., longer or more frequent activation of NS actions, which is coherent with their higher energy-saving potential. Overall, the duty-cycle map helps distinguish whether high savings come from few strong opportunities or from sustained feasibility over time.

Finally, the transition-frequency heatmap (Fig. 4.8d) reveals meaningful behavioural differences associated with traffic-shape pairs. Bus–Bus pairs consistently exhibit among the lowest transition values across volume combinations, indicating that their regular and predictable temporal profiles enable stable operation with fewer switching events. Low transition counts are also observed in the H–H volume column, but for a different reason: when both sectors operate near sustained high load, the margin for NS is small, which naturally suppresses switching activity. At the opposite end, Res–Res pairs show systematically higher transition frequencies across most volume combinations, reflecting sharper evening peaks and steeper temporal gradients typical of residential demand. Such pronounced variations create more switching opportunities, but may also imply increased actuation activity (and potentially higher wear) due to more frequent state changes.

4.9 Feasibility score: balancing energy gains and switching stress

Another practical outcome of the KPI-based analysis is the ability to rank NS candidates and identify which pairings are truly attractive in operational terms, beyond the sole question of technical feasibility. Indeed, NS can unlock substantial energy savings by consolidating traffic on fewer active BSs and switching redundant resources off, but it can also induce frequent sleep/active transitions that increase operational complexity and may accelerate hardware wear due to repeated power-state changes. Since MNOs operate heterogeneous infrastructures (different vendors, generations, cooling conditions, and maintenance policies), the acceptable trade-off between energy gains and switching stress is inherently operator-dependent. To capture this variability, we introduce a composite **feasibility score** that merges energy saving and switching frequency into a single tunable indicator, allowing each MNO to weight the two objectives according to its own operational priorities.

4.9.1 Score definition

For each co-located pair, we consider two KPIs: (i) the percentage energy saved, denoted by ES (higher is better), and (ii) the frequency of state transitions, denoted by SF (lower is better). Because the two metrics have different scales, both are mapped to [0,1] through min-max normalization over the set of evaluated pair instances:

$$\widetilde{ES} = \frac{ES - ES_{\min}}{ES_{\max} - ES_{\min}}, \qquad \widetilde{SF} = \frac{SF - SF_{\min}}{SF_{\max} - SF_{\min}}. \tag{4.2}$$

Since fewer transitions are preferable, we convert switching into a stability term:

$$\widetilde{ST} = 1 - \widetilde{SF}. \tag{4.3}$$

The feasibility score is then defined as a weighted combination:

$$\mathrm{FS} = w_{\mathrm{ES}} \cdot \widetilde{ES} + w_{\mathrm{SF}} \cdot \widetilde{ST}, \qquad w_{\mathrm{ES}} + w_{\mathrm{SF}} = 1. \tag{4.4}$$

Large w_{ES} encodes an energy-oriented policy, whereas large w_{SF} encodes a conservative policy that prioritizes stability (i.e., fewer transitions). In our visualization we report two illustrative configurations: F1 uses $(w_{\mathrm{ES}}, w_{\mathrm{SF}}) = (0.5,0.5)$, while F2 uses (0.3,0.7), giving higher importance to switching stress.

4.9.2 Percentile-based admissibility and decision frontiers

To translate FS into an operational selection rule, we adopt a percentile-based cutoff. Given all scores observed in the dataset, we define an admissibility threshold as the ρ-th percentile of the score distribution:

$$\tau = Q_\rho(\mathrm{FS}), \qquad \rho \in (0,1), \tag{4.5}$$

and we retain only pair instances such that $\mathrm{FS} \geq \tau$. In practice, ρ controls how selective the policy is: for example, $\rho = 0.3$ discards the bottom 30% of pair instances under the chosen weights (retaining the top 70%). Because FS is linear in $(\widetilde{ES}, \widetilde{SF})$, the condition $\mathrm{FS} = \tau$ corresponds to a straight *iso-score* line in the (SF, ES) plane: increasing switching (moving right) must be compensated by higher energy saving (moving up). The plotted lines F1 and F2 are exactly these iso-score frontiers evaluated at the same percentile level $\rho = 0.3$ under two different weight settings; points above a frontier satisfy $\mathrm{FS} \geq \tau$ and are considered admissible under that policy.

4.9.3 Empirical trade-offs across pairing classes

Figure 4.9 reports the joint distribution of the two KPIs, together with class centroids (markers) and the two example decision frontiers. The scatter plots can be read as a trade-off map: the most desirable region is the upper-left (high ES, low SF), while the right-hand side highlights cases where high switching must be justified by sufficiently higher energy gains.

When pairs are grouped by traffic-shape combinations (e.g., Bus–Bus, Res–Res, Mix–Mix), the plot exhibits a clearer structure. Business–Business pairs tend to concentrate

in a favorable region, combining relatively high energy savings with limited switching, thus remaining admissible even under the more conservative frontier F2. Conversely, Mix–Mix pairs cluster at lower savings, meaning that—even if technically feasible—they often provide limited benefit and are more likely to be filtered out by a score-based policy. Residential-involving combinations span a wider region: they can reach high energy savings, but often at the cost of higher switching, making them particularly sensitive to the chosen weight on stability (w_{SF}).

When grouping pairs by average-volume combinations (e.g., L–L, M–L, H–M), the dispersion within each class is larger, showing that volume alone is not sufficient to uniquely characterize operational convenience. Still, centroids reveal systematic differences: some mixed-volume classes achieve reasonable savings with moderate switching, while other classes populate regions of low savings or high switching. Here, the score acts as a practical filter that rejects "not worth it" instances (low savings and/or excessive switching) even when they belong to the same coarse volume category.

Overall, the feasibility score provides a compact and actionable knob to convert heterogeneous KPI outcomes into an operator-specific admissible set of NS opportunities. By tuning $(w_{\mathrm{ES}}, w_{\mathrm{SF}})$ and the selectivity parameter ρ, an MNO can operationalize its own balance between sustainability benefits and infrastructure/operational constraints, obtaining a customized definition of which pairing conditions are worth activating at scale.

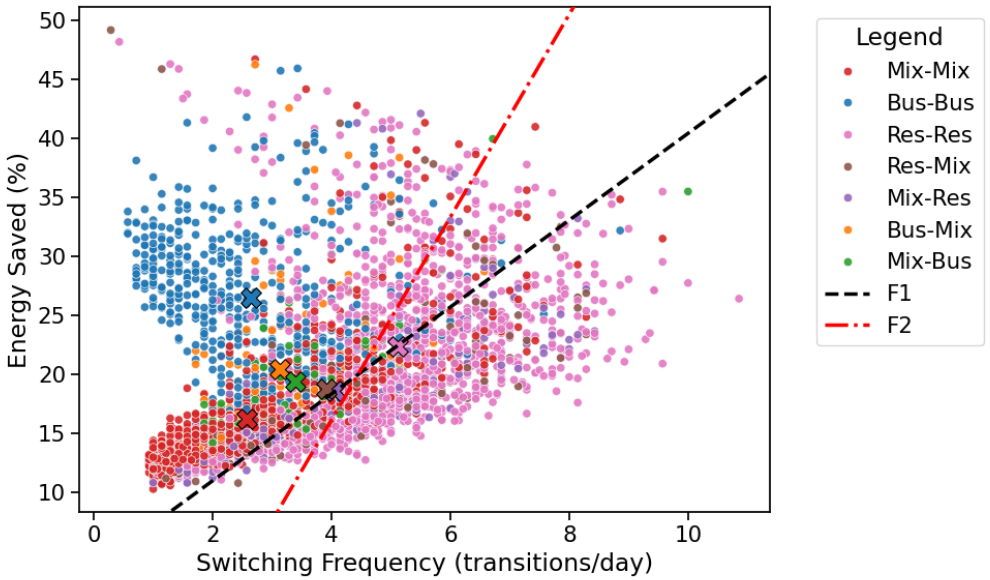
4.10 Explanatory power of traffic profiling

While the heatmaps in Section 4.8 provide a qualitative view of how KPIs vary across pair classes, we also quantify the explanatory power of the proposed profiling by measuring how much of the KPI variability can be explained by cluster-derived labels. To this end, we compute cross-validated coefficients of determination (R^2) for a set of NS-related KPIs using different labeling schemes as predictors. Higher R^2 indicates that the corresponding scheme better captures systematic variations of the KPI.

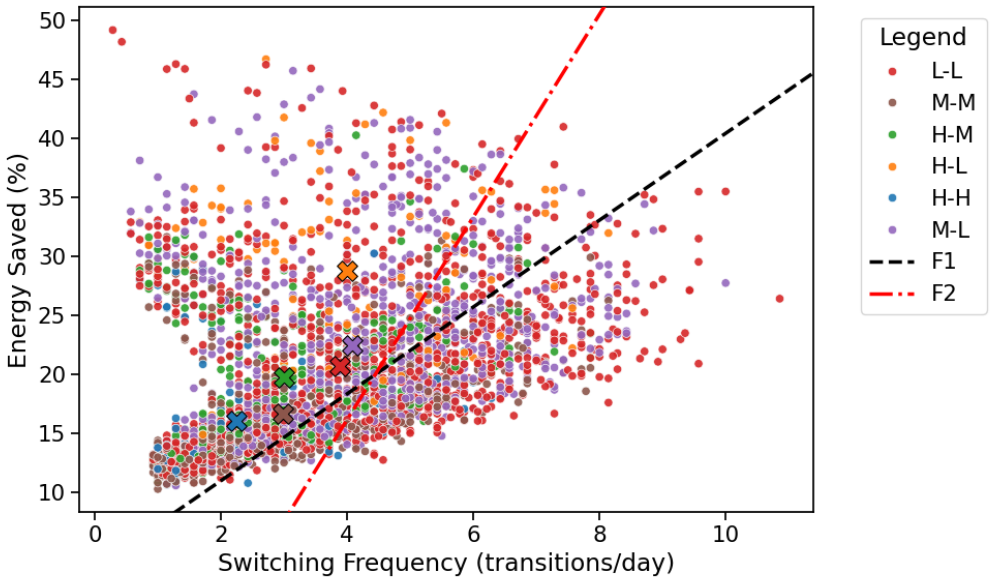
We consider five alternatives:

- **Pattern only**: traffic-shape label.
- **Volume only**: average-volume label.
- **Pattern+Volume**: combined descriptor.
- **Features only**: label derived from clustering in engineered-feature space.
- **Pattern+Volume+Features**: combined descriptor including all label sources.

Figure 4.10 compares the cross-validated R^2 obtained by linear models using different label sets as predictors. A clear outcome is that **Pattern+Volume** consistently improves over using **Pattern only** or **Volume only** across all KPIs, showing that the two descriptors capture complementary sources of KPI variability: temporal behaviour and average load regime. Moreover, the **Pattern+Volume+Features** scheme yields



(a) Pattern-based pairing classes.



(b) Volume-based pairing classes.

Figure 4.9: Trade-off between energy saving (ES) and switching frequency (SF) across NS pair instances.

the highest R^2 overall, indicating that engineered-feature clustering provides additional information beyond the coarse shape/volume taxonomy. Conversely, **Volume only** exhibits the weakest explanatory power, suggesting that average traffic intensity alone is

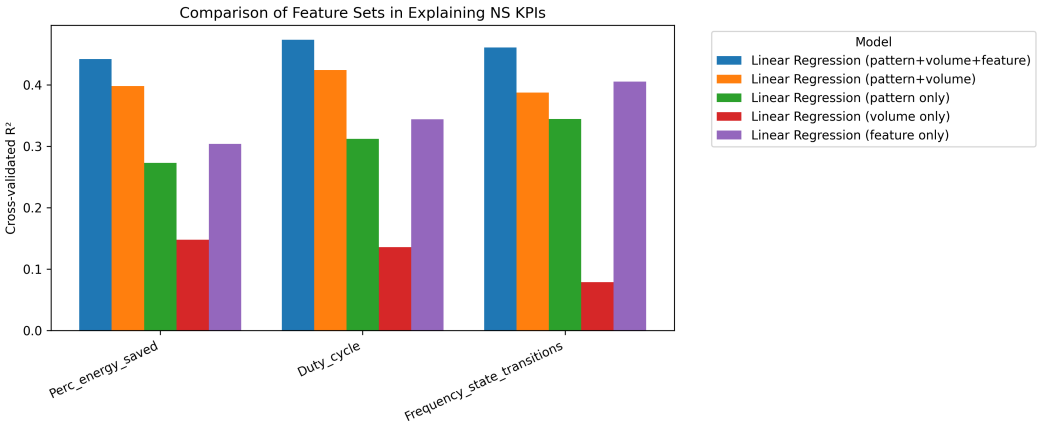


Figure 4.10: Cross-validated R^2 comparison across labeling schemes for three NS-related KPIs.

insufficient to explain KPI variations without accounting for temporal compatibility effects. Overall, these results support the use of joint labels for KPI stratification and motivate including engineered features when the goal is maximum explainability.

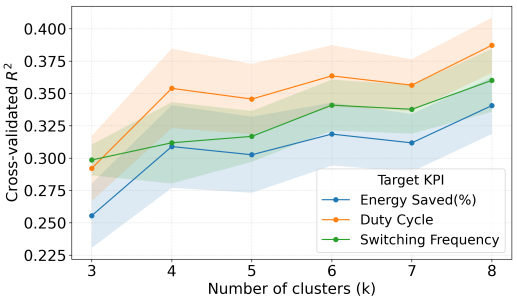
To assess the impact of representational granularity, we evaluate cross-validated R^2 as a function of the number of clusters k. We report the mean performance across folds and the associated variability for each KPI, considering three alternative profiling pipelines: traffic-shape clustering, volume-based clustering, and feature-based clustering.

Figure 4.11 shows that the relationship between k and explanatory power depends on both the target KPI and the profiling pipeline. For most KPIs, traffic-shape clustering (Fig. 4.11a) improves as k increases, suggesting that finer-grained temporal taxonomies capture additional variability relevant to NS potential. In contrast, for average-volume clustering (Fig. 4.11b), gains are mainly observed for **Energy Saved** and **Duty Cycle**, while **Switching Frequency** exhibits limited R^2 and minimal sensitivity to k, indicating that load intensity alone does not explain the temporal dynamics that drive switching activity. Finally, feature-based profiling (Fig. 4.11c) generally achieves higher R^2 and remains more stable across k, consistent with the richer information encoded in engineered descriptors. Overall, these trends support selecting a small k for interpretability, while showing that larger k can yield measurable gains when the objective is to maximize KPI explainability.

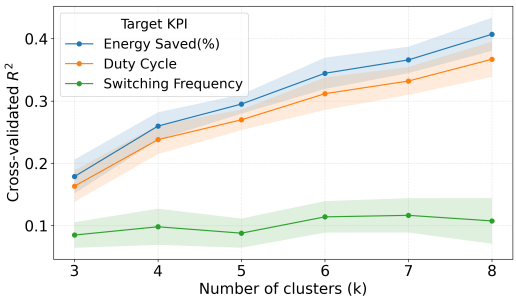
4.11 Cross-city validation: Paris vs. Lyon

To assess whether the trends observed in the primary study area are city-specific or general, we replicate the same KPI stratification on a second metropolitan region (Lyon) and compare the resulting distributions with Paris. The objective is to verify that the relationship between traffic-pattern combinations and NS-related KPIs remains stable under a different urban scale and demand intensity.

Figure 4.12 compares the KPI distributions across traffic-pattern pairs for Paris and



(a) Traffic-shape clustering.



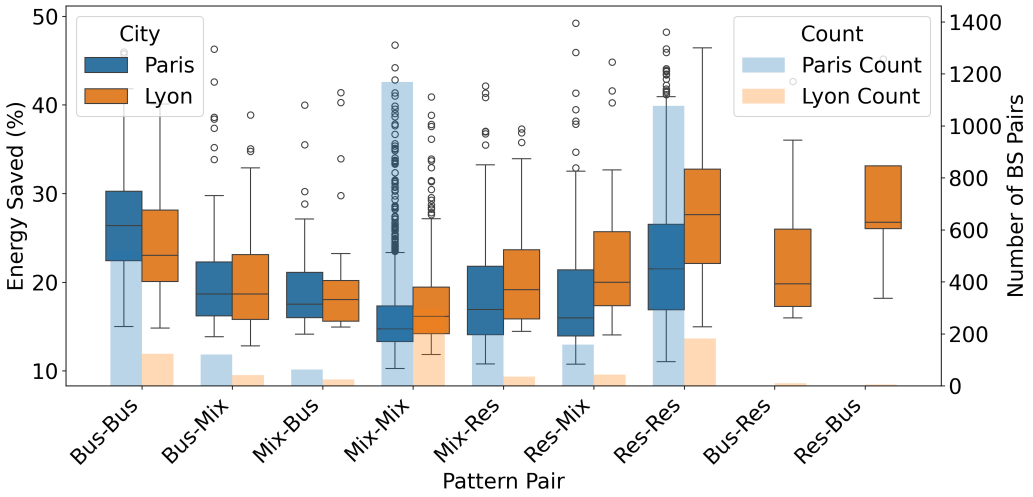
(b) Average-volume clustering.

(c) Engineered-feature clustering.

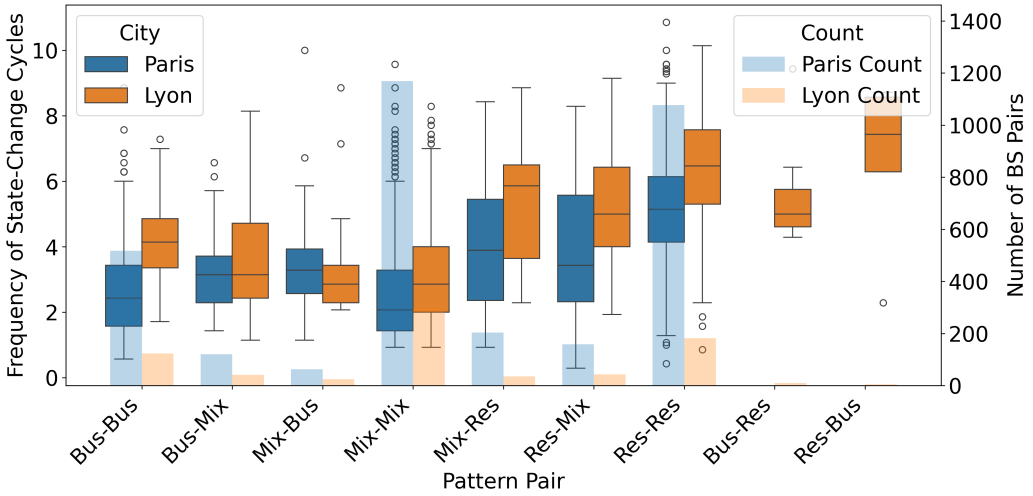
Figure 4.11: Cross-validated R^2 versus the number of clusters k (mean across folds with variability bands)

Lyon. Overall, the results confirm that traffic-pattern combinations drive NS performance in a consistent way across cities. Pattern pairs associated with more temporally regular and stable loads (e.g., Bus–Bus and Mix–Mix) systematically exhibit higher energy-saving potential and lower state-change cycle frequencies in both Paris and Lyon (Figs. 4.12a–4.12b). Conversely, combinations involving Residential patterns tend to show higher transition frequencies, with the highest switching activity typically observed for asymmetric pairs (e.g., Res–Bus, Res–Mix), reflecting stronger temporal imbalance between the two co-located sectors.

While the absolute KPI values differ slightly across cities, these differences are largely explained by the different number of available pairs per class (secondary-axis bars) and by city-level differences in demand intensity. Importantly, the relative ranking of pattern pairs is largely preserved across Paris and Lyon, indicating that the link between traffic behaviour and NS-related KPIs is not specific to a single metropolitan area. This cross-city validation strengthens the interpretation of the clustering-derived labels as transferable descriptors of NS-relevant operating regimes.



(a) Energy saved (%) across pattern-pair classes. Boxplots show KPI distributions for Paris and Lyon; shaded bars on the secondary axis report the number of BS pairs in each class.



(b) Frequency of state-change cycles across pattern-pair classes. Boxplots show KPI distributions for Paris and Lyon; shaded bars on the secondary axis report the number of BS pairs in each class.

Figure 4.12: Cross-city comparison of KPI distributions by traffic-pattern pair. Paris and Lyon exhibit consistent ordering across pattern combinations, supporting the generality of the observed trends.

4.12 Key Findings

This chapter showed that traffic profiling provides an effective lens to interpret and assess network sharing opportunities in a data-driven way. A first key result is that **traffic**

shape and **average traffic volume** play complementary roles. Traffic shape captures the temporal organization of demand and is especially informative for understanding the stability of network-sharing operation, while average volume captures the load level and therefore the residual capacity available for traffic consolidation. Considering only one of these dimensions would provide an incomplete view of NS feasibility.

A second important outcome is that their **joint use has stronger explanatory power** than either descriptor alone. The combined pattern–volume representation more effectively explains the variability of the considered NS-related KPIs, confirming that both the temporal structure of demand and its magnitude are needed to characterize the performance of co-located BS pairs. This supports the use of joint traffic labels as compact yet informative descriptors for NS-oriented analysis and planning.

Finally, the results highlight a fundamental **trade-off between energy savings and operational stability**. Some pairings enable large energy gains but at the cost of more frequent switching activity, whereas others provide more stable operation with lower savings. The proposed feasibility score makes this trade-off explicit and configurable, offering a practical way to rank NS opportunities according to operator priorities. Overall, these findings show that network sharing should be evaluated not only in terms of potential energy reduction, but also in terms of the traffic conditions that determine how reliably and sustainably such gains can be achieved.

Chapter 5

Prediction of Threshold Exceedance

5.1 Motivation: anticipating NS feasibility

The data-driven analysis in Chapter 4 shows that NS opportunities (e.g., traffic offloading and temporary switch-off of redundant base stations) are fundamentally constrained by feasibility conditions. In particular, a switch-off or offloading action is only safe when the neighboring BS has enough residual capacity to absorb additional load without entering congestion regimes. This feasibility condition is not static: utilization fluctuates over time due to strong diurnal and weekly patterns, as well as short-term bursts driven by local events and service-specific demand.

In an operational setting, NS decisions must therefore be made ahead of time, based on the expected future load. Relying solely on instantaneous utilization can be misleading: a base station that is currently underloaded may approach saturation shortly after, making an offloading decision unsafe. Conversely, conservative decisions that avoid switching whenever uncertainty is high may miss substantial energy-saving opportunities. This motivates the need for predictive mechanisms capable of anticipating high-load regimes and near-saturation events that reduce the available headroom for sharing.

To bridge the gap between retrospective NS evaluation and actionable control, this chapter introduces a prediction layer that forecasts capacity-threshold exceedance events at the level of candidate sharing pairs. These events provide a compact, decision-oriented representation of feasibility: they indicate whether the host BS, if required to serve the combined demand of the pair, is expected to exceed a given safety threshold within the next horizon, and therefore whether it can reliably accommodate offloaded traffic from the other BS.

In the following, we formalize exceedance events and compare two strategies to anticipate them: an indirect workflow based on forecasting utilization and thresholding the forecast, and a direct workflow where exceedances are predicted as a classification task. We then evaluate these strategies under time-consistent validation and quantify their accuracy using event-level metrics aligned with NS decision requirements.

5.2 Problem definition

NS feasibility is constrained by the ability of an active BS to absorb additional load without exceeding its capacity. In our framework, feasibility is evaluated for a candidate sharing pair (i, j), where BS i is the potential **host** and BS j is the **donor** whose traffic may be offloaded.

Let $\ell_i(t) \geq 0$ denote the traffic load served by BS i at time t, and let $C_i > 0$ be the capacity of BS i (in the same units as ℓ_i). For a decision taken at time t, we consider a prediction horizon $h \geq 1$ and the future combined load that BS i would need to serve if it also absorbed the traffic of BS j:

$$L_{ij}(t+k) = \ell_i(t+k) + \ell_j(t+k), \qquad k = 1, \ldots, h.$$

We define a horizon-level exceedance event for threshold $\tau \in (0,1)$ as

$$y^{(\tau)}_{ij,h}(t) = \mathbb{I}\Big(\max_{k=1,\ldots,h} L_{ij}(t+k) \;\geq\; \tau\, C_i \Big). \tag{5.1}$$

In this work we focus on two thresholds: a **caution** threshold $\tau_{\mathrm{c}} = 0.8$ and a **near-saturation (QoS)** threshold $\tau_{\mathrm{q}} = 0.95$. The threshold τ_{c} identifies high-load regimes where the margin for offloading becomes limited, while τ_{q} captures near-saturation regimes where additional offload is likely infeasible or unsafe. An exceedance therefore flags that offloading j onto i is expected to be unsafe within the next h slots.

Traffic is sampled every $\Delta = 15$ minutes. We consider the set of horizons

$$h \in \{1,4,8,16\},$$

corresponding to 15 minutes, 1 hour, 2 hours, and 4 hours ahead, respectively. Longer horizons provide earlier planning capability for NS decisions but increase uncertainty, since the combined load $L_{ij}(t+k)$ must be anticipated further into the future.

5.3 Prediction approaches compared

We compare two workflows to predict the future feasibility-related exceedance event $y^{(\tau)}_{ij,h}(t)$ defined in Section 5.2.

5.3.1 Indirect (forecast-then-threshold)

The indirect workflow first forecasts the future traffic loads of the pair and then derives the exceedance decision by thresholding the implied combined load. Given forecasts $\hat{\ell}_i(t+k)$ and $\hat{\ell}_j(t+k)$ for $k = 1, \ldots, h$, we form

$$\hat{L}_{ij}(t+k) = \hat{\ell}_i(t+k) + \hat{\ell}_j(t+k),$$

and compute the predicted exceedance as

$$\hat{y}^{(\tau)}_{ij,h}(t) = \mathbb{I}\Big(\max_{k=1,\ldots,h} \hat{L}_{ij}(t+k) \;\geq\; \tau\, C_i \Big). \tag{5.2}$$

Feature group	#	Purpose (high-level)
Temporal features	6	Encode daily/weekly periodicity (cyclic + ordinal time signals).
Recent load history	16	Provide short-term context for ℓ_i and ℓ_j (level and local evolution).
Pair features	9	Describe combined load, headroom proxies, and short-term statistics.
Total	**31**	

Table 5.1: Summary of feature groups used for pair-level prediction.

This approach can reuse standard time-series forecasting models (statistical or machine-learning-based). However, it optimizes a regression objective on future traffic values, whereas NS control depends on a horizon-level binary event. Consequently, small forecasting errors near the threshold can translate into decision flips that directly affect feasibility assessment.

5.3.2 Direct (event classification)

In the direct workflow, a classifier predicts $y^{(\tau)}_{ij,h}(t)$ directly from information available up to time t. The model outputs either a binary label or a score interpreted as exceedance likelihood. By training directly on the horizon-level exceedance event, the learning objective matches the downstream decision variable used by the NS logic and focuses capacity on the near-threshold region that determines feasibility.

5.4 Input features

For each decision time t and candidate sharing pair (i, j), we build a feature vector $\mathbf{x}_{ij}(t) \in \mathbb{R}^{31}$ using only information observable up to time t. Features are designed to (i) encode the strong daily/weekly periodicities of mobile traffic and (ii) summarize the most recent pair dynamics (level, trend, volatility), which are key drivers of near-future threshold exceedances.

5.4.1 Temporal features

To represent periodic behavior without artificial discontinuities at cycle boundaries, we include cyclic encodings of hour-of-day and day-of-week: `hour_sin`, `hour_cos`, `day_sin`, `day_cos`. This mapping places time on the unit circle, so that late-night and early-morning slots (e.g., 23:45 and 00:00) are close in feature space. We also retain integer representations `hour_of_day` and `day_of_week`, which are useful for tree-based models and provide a direct ordinal signal.

5.4.2 Recent load history

We include the most recent traffic samples for both BS i (host) and BS j (donor) using the current slot and the previous 7 lags:

$$\texttt{i\_load\_lag\_k} = \ell_i(t-k), \qquad \texttt{j\_load\_lag\_k} = \ell_j(t-k), \qquad k = 0,\ldots,7.$$

With a 15-minute sampling period, this corresponds to a 2-hour history window. These lag features provide local temporal context (current level and short-term evolution) that is typically most informative at the horizons considered in this work.

5.4.3 Pair features

Because feasibility depends on the combined demand under offloading, we include current loads and their sum:

$$\texttt{load\_i\_t} = \ell_i(t), \quad \texttt{load\_j\_t} = \ell_j(t), \quad \texttt{sum\_now} = \texttt{load\_i\_t} + \texttt{load\_j\_t}.$$

We further include a conservative proxy of instantaneous headroom under a fixed caution threshold (0.8):

$$\texttt{margin\_now} = 0.8\,C_i - \texttt{sum\_now}, \qquad \texttt{margin\_ratio} = \frac{\texttt{margin\_now}}{0.8\,C_i},$$

where C_i is the capacity of BS i in the same units as ℓ. Finally, to capture short-term variability beyond individual lags, we compute summary statistics of the combined load $\ell_i + \ell_j$ over the previous 8 slots (from $t-1$ to $t-8$): `sum_mean_past`, `sum_std_past`, `sum_max_past`, `sum_min_past`.

The same feature set is used for both indirect and direct workflows, and for all horizons h, so that performance differences can be attributed to the prediction formulation and model choice rather than to different input information.

5.5 Predictive models

To instantiate both strategies, we consider three model families: Random Forest (RF), LightGBM, and linear SVM-based baselines. For the **indirect** strategy, we employ regression models to forecast traffic and then derive exceedance decisions (RF regressor, LightGBM regressor, and a linear regressor baseline such as Ridge/LinearSVR). For the **direct** strategy, we employ classifiers trained to predict $y^{(\tau)}_{ij,h}(t)$ directly (RF classifier, LightGBM classifier, and Linear SVM).

5.5.1 Random Forest (RF)

Random Forest is an ensemble method based on bagging of decision trees. Each tree is trained on a bootstrap sample of the training data, and at each split only a random subset of features is considered, which reduces correlation between trees and improves generalization. For classification, the forest prediction is obtained by majority vote across trees; for regression, predictions are averaged. RF can capture nonlinear interactions and is robust to heterogeneous feature scales.

Fold f	Training window (days)	Test window (day)	Slots (train / test)
1	days 1–3	day 4	288 / 96
2	days 1–4	day 5	384 / 96
3	days 1–5	day 6	480 / 96
4	days 1–6	day 7	576 / 96

Table 5.2: Rolling/blocked train–test folds used for time-consistent evaluation (one week, 15-min slots).

5.5.2 LightGBM

LightGBM belongs to the family of gradient-boosted decision trees, where trees are added sequentially to correct the errors of the current ensemble. Formally, at iteration m the model is updated as $F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \eta f_m(\mathbf{x})$, where η is the learning rate and f_m is the newly added tree. LightGBM supports both regression and classification objectives (e.g., squared error and logistic loss) and is efficient on large datasets due to histogram-based split finding and optimized tree growth.

5.5.3 Linear Support Vector Machine (Linear SVM)

A Linear SVM learns a maximum-margin separating hyperplane $f(\mathbf{x}) = \mathbf{w}^\top \mathbf{x} + b$ for binary classification. The parameters are obtained by minimizing a regularized hinge-loss objective, controlled by a penalty parameter C that trades margin width against training errors. The model outputs a signed decision score; the predicted label is obtained from its sign. Since the method is sensitive to feature scaling, input features are standardized before training.

5.5.4 Train-Test Split

Because cellular traffic exhibits strong temporal correlation, we adopt a time-consistent blocked evaluation (rolling-origin train–test splits) to prevent information leakage. Let T be the total number of time samples in the week ($T = 7 \times 96 = 672$ slots). We use a test block length of one day (96 slots) and a growing training set. For fold f, the training interval contains only past samples and the test interval is the immediately following block:

$$\text{train}_f = [0,\, t_f) \qquad \text{and} \qquad \text{test}_f = [t_f,\, t_f + 96),$$

with $t_f \in \{288,384,480,576\}$, i.e., the first test day starts after three full training days.

Since the horizon-level label $y^{(\tau)}_{ij,h}(t)$ depends on future samples up to $t + h$, ground-truth labels are only defined for time indices $t \leq T - h - 1$. Accordingly, within each fold we evaluate only the time samples for which ground truth exists (i.e., the last h slots of the week are excluded from evaluation).

5.5.5 Evaluation metrics

Both the **direct** and **indirect** workflows ultimately output a binary decision on feasibility, i.e., a prediction of the horizon-level exceedance event $y_{ij,h}^{(\tau)}(t) \in \{0,1\}$ (Section 5.2), for $\tau \in \{\tau_c, \tau_q\}$. Accordingly, we evaluate all approaches primarily in terms of classification performance on these exceedance indicators. Let TP, FP, TN and FN denote true positives, false positives, true negatives and false negatives, respectively.

Precision,accuracy, recall and F1-score. Precision measures the fraction of predicted exceedances that are correct:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}.$$

Accuracy measures the fraction of correct predictions:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}.$$

Recall (sensitivity) measures the fraction of true exceedances that are detected:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}.$$

Their harmonic mean defines the F1-score:

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}.$$

These metrics are informative under class imbalance, which may arise depending on the chosen threshold τ and horizon h.

5.6 Feasibility Event statistics

We analyze the empirical properties of the feasibility-related exceedance events $y_{ij,h}^{(\tau)}(t)$ introduced in Section 5.2, focusing on two thresholds defined in 5.2: a *caution* threshold $\tau_c = 0.8$ and a *near-saturation (QoS)* threshold $\tau_q = 0.95$. In the following, we denote the corresponding events as $y_{ij,h}^{(\tau_c)}(t)$ and $y_{ij,h}^{(\tau_q)}(t)$.

5.6.1 Horizon-driven imbalance of exceedance events

Table 5.3 reports the class balance of exceedance events as a function of the prediction across the horizons. Two clear trends emerge.

First, the positive rate (fraction of exceedances) increases monotonically with h for both thresholds. This is expected from the horizon-level definition (5.1) as h grows, there are more opportunities for the combined load L_{ij} to cross the threshold, hence exceedances become more frequent.

Second, for any fixed horizon, exceedances are more frequent at the lower threshold τ_c than at τ_q, since $\tau_c C_i$ is easier to exceed than $\tau_q C_i$.

Overall, these statistics indicate that the classification task becomes progressively more imbalanced (more positives) as the horizon increases, which must be considered when interpreting precision/recall trade-offs and comparing performance across horizons.

h	Label	Rows	Pos	Neg	Pos rate
1	$y_{ij,h}^{(\tau_q)}$	362,340	174,491	187,849	0.4816
4	$y_{ij,h}^{(\tau_q)}$	360,720	194,352	166,368	0.5388
8	$y_{ij,h}^{(\tau_q)}$	358,560	210,476	148,084	0.5870
16	$y_{ij,h}^{(\tau_q)}$	354,240	237,546	116,694	0.6706
1	$y_{ij,h}^{(\tau_c)}$	362,340	199,493	162,847	0.5506
4	$y_{ij,h}^{(\tau_c)}$	360,720	217,724	142,996	0.6036
8	$y_{ij,h}^{(\tau_c)}$	358,560	233,062	125,498	0.6500
16	$y_{ij,h}^{(\tau_c)}$	354,240	259,783	94,457	0.7334

Table 5.3: Horizon-level feasibility exceedance events statistics.

5.6.2 Cross-horizon agreement and divergence

Since horizon-level labels are defined through a maximum over a future window, labels at different horizons are not independent. Longer horizons include shorter ones as a subset of the future window, which induces correlation between $y_{ij,h_1}^{(\tau)}(t)$ and $y_{ij,h_2}^{(\tau)}(t)$ for $h_2 > h_1$. To quantify this effect, Table 5.4 reports two agreement measures computed on aligned samples (i.e., on the intersection of time indices where both horizons have defined ground truth):

(i) the overall agreement accuracy (fraction of time–pair samples where the two labels match), and *(ii)* the positive-class Jaccard index, defined as

$$J^{+} = \frac{|\{y_{h_1}^{(\tau)} = 1\} \cap \{y_{h_2}^{(\tau)} = 1\}|}{|\{y_{h_1}^{(\tau)} = 1\} \cup \{y_{h_2}^{(\tau)} = 1\}|},$$

which focuses specifically on the overlap of exceedance events.

Agreement is high for nearby horizons. For instance, for both thresholds the pair $(h_1, h_2) = (4,8)$ yields agreement accuracies above 0.95 and positive Jaccard indices above 0.91. As the horizon gap increases, agreement decreases: comparing $h_1 = 1$ against $h_2 = 16$ yields an agreement of about 0.81 for both thresholds, and positive Jaccard indices around 0.72–0.75. This divergence reflects that longer windows are more likely to include at least one exceedance even when the short-horizon window does not.

These results provide useful context for interpreting predictive performance across horizons. Because labels remain strongly correlated for adjacent horizons, performance trends may appear "flat" across h in some settings, while larger horizon differences correspond to genuinely different decision problems.

h_1	h_2	Label	Aligned samples	Agreement acc.	J^+
1	4	$y_{ij,h}^{(\tau_q)}$	360,720	0.9440	0.8961
1	8	$y_{ij,h}^{(\tau_q)}$	358,560	0.8960	0.8229
1	16	$y_{ij,h}^{(\tau_q)}$	354,240	0.8128	0.7208
4	8	$y_{ij,h}^{(\tau_q)}$	358,560	0.9521	0.9185
4	16	$y_{ij,h}^{(\tau_q)}$	354,240	0.8683	0.8036
8	16	$y_{ij,h}^{(\tau_q)}$	354,240	0.9163	0.8752
1	4	$y_{ij,h}^{(\tau_c)}$	360,720	0.9480	0.9139
1	8	$y_{ij,h}^{(\tau_c)}$	358,560	0.9013	0.8481
1	16	$y_{ij,h}^{(\tau_c)}$	354,240	0.8171	0.7506
4	8	$y_{ij,h}^{(\tau_c)}$	358,560	0.9534	0.9282
4	16	$y_{ij,h}^{(\tau_c)}$	354,240	0.8687	0.8210
8	16	$y_{ij,h}^{(\tau_c)}$	354,240	0.9156	0.8849

Table 5.4: Cross-horizon agreement for feasibility exceedance events at thresholds $\tau_q = 0.95$ (QoS) and $\tau_c = 0.8$ (caution).

5.7 Model comparison

5.7.1 Comparative accuracy of direct and indirect strategies

We use the shorthand $y^{(c)}(t) := y_{ij,h}^{(\tau_c)}(t)$ and $y^{(q)}(t) := y_{ij,h}^{(\tau_q)}(t)$. Figures 5.1–5.2 compare direct (solid lines) and indirect (dashed lines) predictions across horizons and model families for the caution event $y^{(c)}$ ($\tau_c = 0.8$) and the QoS event $y^{(q)}$ ($\tau_q = 0.95$). Across both events, the direct formulation paired with tree-based learners (RF/LightGBM) achieves the best overall performance, with consistently higher accuracy and F1-score than the corresponding indirect workflow for all horizons.

- **Caution exceedance $y^{(c)}$ (Fig. 5.1).** Direct RF and direct LightGBM provide the highest accuracy and F1-score across all horizons, with a mild degradation as h increases. The indirect approach remains comparatively stable with h, but at a systematically lower accuracy/F1 level. This behavior is consistent with the indirect pipeline accumulating regression errors across the horizon and then applying a hard thresholding step, which can produce decision flips even when the traffic forecast error is small.

- **QoS exceedance $y^{(q)}$ (Fig. 5.2).** The same ranking holds, and the performance gap between direct and indirect is even more evident in precision. Direct RF/LightGBM preserve high recall while maintaining substantially higher precision than the indirect variants, indicating fewer false alarms for a similar ability to detect violations.

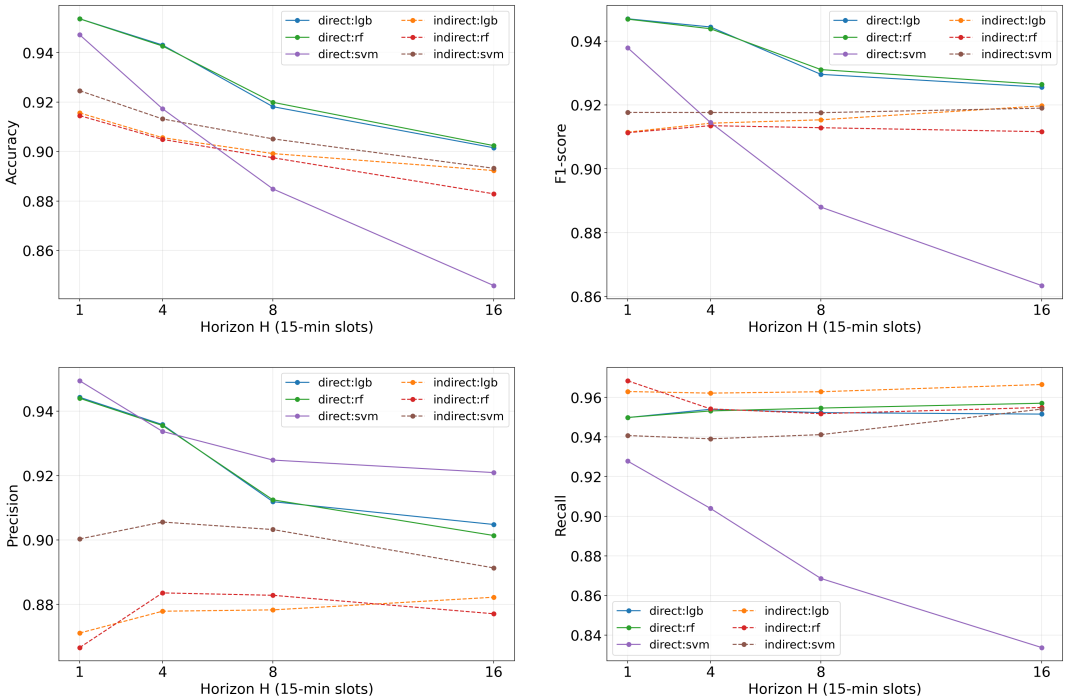


Figure 5.1: Direct vs indirect performance for the caution exceedance $y^{(c)}$ across horizons.

A consistent observation across both events is the weak sensitivity of the indirect F1-score to the horizon. This can occur because the horizon-level exceedance labels are strongly correlated across nearby horizons and are often dominated by samples that are far from the threshold (i.e., clearly safe or clearly unsafe). In such cases, changes in traffic-forecast RMSE with h do not necessarily translate into many additional threshold decision flips, yielding nearly flat F1 trends even when the underlying regression task becomes harder.

Finally, the linear SVM baseline underperforms for the direct task, and the gap grows with h. This is coherent with the feature distribution (highly non-Gaussian and heavy-tailed load descriptors): a linear separator struggles to represent the nonlinear boundary induced by the max-over-horizon operator and the capacity constraint, whereas tree ensembles can model feature interactions and non-monotone effects more effectively.

5.7.2 False alarms vs missed violations: impact on NS decisions

For NS control, the two error types have asymmetric operational impact: false negatives (FN) correspond to missed exceedances and may lead to unsafe sharing decisions, while false positives (FP) correspond to overly conservative decisions that reduce potential energy savings. The precision–recall decomposition in Figs. 5.1–5.2 provides a direct proxy for this trade-off.

- **Indirect strategy is more conservative in recall but noisier in precision.**

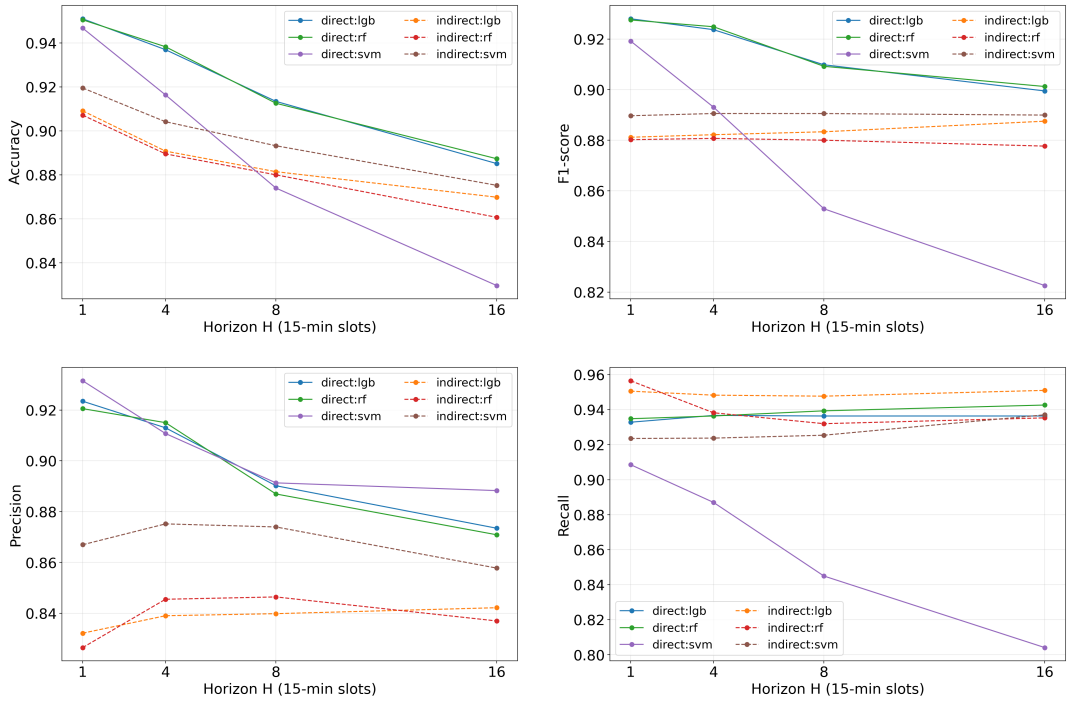


Figure 5.2: Direct vs indirect performance for the QoS exceedance $y^{(q)}$ across horizons.

For both $y^{(c)}$ and $y^{(q)}$, indirect variants tend to sustain high recall, indicating relatively few missed violations. However, their precision is consistently lower, meaning a larger fraction of predicted exceedances are false alarms. In NS terms, this corresponds to an approach that avoids risk at the price of reduced sharing opportunities.

- **Direct RF/LightGBM improves precision without sacrificing recall.** Direct tree-based models achieve recall comparable to the indirect method while improving precision markedly, especially for $y^{(q)}$. This indicates fewer unnecessary blocking decisions for the same level of QoS protection, which is desirable when NS aims to reduce energy consumption without increasing violation risk.

- **Horizon effects and operational tuning.** As h increases, the exceedance event becomes more likely, and the decision boundary effectively shifts toward a broader set of risky operating points. In this regime, maintaining high recall is critical because FN errors become more costly. The direct tree-based models remain robust under this shift, whereas the linear baseline exhibits a stronger recall degradation for larger h, implying an increasing FN risk.

From an engineering standpoint, these results suggest the following guideline for NS feasibility prediction: direct tree-based classifiers are preferable when the control objective depends primarily on the quality of the feasibility event (high precision with high recall), while the indirect approach remains attractive when traffic forecasts are required

downstream for other tasks, provided that the thresholding stage is calibrated to limit false alarms (e.g., via asymmetric costs or decision thresholds).

5.8 Deployment considerations

5.8.1 Computational cost and scalability

The two prediction workflows differ substantially in computational structure and, consequently, in scalability.

In the *indirect* workflow, the learning task is a multi-step traffic forecasting problem. For each decision time t, the predictor produces a vector of future loads over the horizon, $\{\hat{\ell}_i(t+k), \hat{\ell}_j(t+k)\}_{k=1}^{h}$, for all BSs involved in the subset. This implies (i) a larger output space than a binary event, and (ii) a heavier inference pipeline, since predictions must be generated for all stations and all horizon steps before the exceedance indicator can be derived. Moreover, training and evaluation are carried out on time-series structures (lookback windows, context stitching for rolling folds), which increases both CPU time and memory footprint as the number of BSs, the lookback length, and h grow.

In the *direct* workflow, the model predicts the horizon-level event $y_{ij,h}^{(\tau)}(t)$ directly from engineered features. At inference time, each pair–time sample requires only a single forward pass of a classifier (binary output), without producing intermediate trajectories. This reduces the dimensionality of the prediction target and makes the runtime scale approximately linearly with the number of evaluated pair–time instances. In practice, the direct approach is therefore better suited to large-scale evaluation across many pairs, especially when h is large or when forecasting all traffic trajectories is not required by downstream components.

Finally, model family impacts scalability. Tree ensembles (RF and LightGBM) dominate runtime during training but remain efficient at inference, while linear baselines are typically fastest but may lack expressive power for nonlinear feasibility boundaries. LightGBM is particularly attractive in deployment scenarios due to its optimized training and inference routines and its ability to handle large datasets with limited memory overhead.

5.8.2 Risk-aware operating point selection

Feasibility prediction is ultimately used to support NS control decisions, where different error types have asymmetric operational consequences. A missed exceedance (false negative) may result in attempting sharing in conditions that lead to capacity violations, whereas a false alarm (false positive) prevents sharing even when it would have been feasible, reducing potential energy savings.

An operational deployment should therefore select an operating point that matches the desired risk posture:

- **QoS-critical operation (risk-averse).** When the priority is to avoid overload or QoS degradation, the predictor should be tuned to maximize recall for $y^{(q)}$ (i.e., minimize false negatives), even if precision decreases. In practice, this corresponds to conservative decisions that may reduce NS opportunities but keep violation probability low.

- **Energy-driven operation (opportunity-seeking).** When energy saving is prioritized and occasional conservatism is undesirable, higher precision is preferred to reduce false alarms and enable more sharing actions. This shifts the operating point toward fewer predicted exceedances, at the cost of a potentially higher missed-violation risk.

Although the evaluation in this chapter reports fixed-threshold decisions from each model, the same framework can be extended to explicit risk control. For models that output calibrated scores (e.g., probabilistic classifiers), a decision threshold can be selected to satisfy a target violation risk (e.g., $\mathbb{P}[\text{FN}] \leq \epsilon$) or to minimize an asymmetric cost function where false negatives carry a higher penalty than false positives. This risk-aware tuning is consistent with the NS objective of enabling energy reduction while respecting safety constraints and can be adapted to operator-specific policies and traffic regimes.

Chapter 6

Conclusions

Network sharing (NS) is increasingly regarded as a key instrument to reduce the energy footprint and operational costs of mobile networks, especially in dense 5G deployments where multiple operators often install redundant infrastructure at the same locations. In principle, NS enables traffic consolidation: when demand is low, part of the infrastructure can be switched off and the remaining base stations can serve the aggregated load. In practice, however, NS is constrained by a fundamental feasibility condition: an offloading or switch-off action is safe only if the assisting base station has enough residual capacity to absorb additional traffic without entering congestion. Since traffic is highly time-varying and exhibits both regular seasonality and unpredictable bursts, feasibility cannot be assessed reliably from instantaneous utilization alone. This makes NS not only an optimization problem, but also a traffic-aware decision problem.

Within this broad context, this thesis focused on a specific and actionable question: which traffic conditions make network sharing more feasible and more beneficial than others? Rather than treating traffic variability as noise, the goal was to understand how systematic differences in traffic behavior across base stations translate into different NS opportunities. In particular, the thesis investigated whether certain traffic types systematically yield higher headroom, more stable consolidation decisions, and better overall performance, and how these types can be identified from data in a way that supports practical NS planning.

To answer this question, the thesis carried out a data-driven analysis based on real mobile traffic measurements and evaluated NS performance on co-located BS pairs under feasibility constraints. By profiling traffic and grouping BSs according to their behavior, the analysis made it possible to pinpoint which combinations of co-located traffic profiles lead to better NS outcomes in the general application of the model. The results show that feasibility and performance are strongly tied to traffic compatibility. In particular, the most favorable cases are observed when the two base stations exhibit a homogeneous traffic-shape combination, i.e., similar temporal patterns, and at the same time a complementarity in traffic volume, i.e., an imbalance in average load that creates exploitable headroom on the less loaded infrastructure. This combination is not only qualitatively preferable, but also quantitatively impactful: the achievable energy savings vary markedly across traffic pair types, reaching around 33% in favorable combinations and dropping to

about 14% in unfavorable ones. This gap highlights that whether sharing is worth it is largely determined by the underlying traffic pairing, and motivates traffic-aware selection as a necessary step for effective NS.

A second key outcome of the analysis is that energy gains cannot be assessed in isolation, because they are coupled with operational dynamics. Consolidation decisions can increase the frequency of base-station state transitions, especially when traffic is bursty or when two co-located cells have synchronized peaks that repeatedly violate feasibility. Higher switching frequency is operationally undesirable because frequent on/off cycling can increase control overhead and may accelerate wear and degradation mechanisms in the infrastructure. Therefore, the thesis supports the need for stability-aware control, where part of the energy gain is traded for smoother operation through safeguards such as hysteresis, minimum on/off times, or window-based decisions.

Building on these findings, the thesis moved from retrospective characterization to improving the applicability of NS in operational settings. Since feasibility depends on whether the assisting base station will remain below a critical utilization threshold in the near future, the work formulated the anticipation of unsafe regimes as a threshold-surpassing prediction problem. Instead of forecasting traffic only as a continuous time series, the objective was to predict whether (and when) utilization would exceed a critical level that makes offloading unsafe. This predictive layer improves decision-making by reducing unsafe activations ahead of upcoming peaks and by limiting overly conservative behavior when low-load conditions are expected to persist. In this sense, the thesis improves the practical usefulness of well-known traffic prediction approaches by tailoring them to the control objective that matters for NS: correctly identifying near-saturation events, with improved accuracy and precision, so that sharing actions are both safer and more effective. Overall, the thesis shows that the effectiveness of network sharing depends critically on which traffic is being shared. By identifying the traffic combinations that systematically maximize feasibility and quantifying their impact on achievable savings and by adding a predictive mechanism to anticipate threshold surpassing, the work contributes toward making NS decisions more actionable, safer, and operationally stable in real networks.

6.1 Future work

Future work can broaden and strengthen the proposed framework in several directions. Extending the control scope beyond co-located pairs to multi-cell and multi-operator clusters is a natural next step: larger sharing groups may unlock additional savings, but require scalable coordination strategies. Another direction is the development of multi-objective policies that explicitly trade energy savings with operational stability, QoS risk, and fairness across operators, potentially through constrained optimization or Pareto-based formulations. On the prediction side, uncertainty-aware forecasting and probabilistic models for threshold exceedance would enable risk-limiting decisions with explicit confidence levels, improving robustness under bursts and atypical events.

Appendix A

Features

1. **Coefficient of Variation (CV):**

$$\mathrm{CV} = \frac{\sigma_x}{\mu_x}$$

   Measures relative variability of traffic, independent of its absolute magnitude.

2. **Burstiness Index:**

$$B = \frac{\max(x(t)) - \mu_x}{\mu_x}$$

   Quantifies how strongly the traffic peaks exceed the average load.

3. **Area Under Curve (AUC):**

$$\mathrm{AUC} = \int_0^T x(t)\,dt \approx \sum_{t=1}^{T-1} \frac{x(t) + x(t+1)}{2}\,\Delta t$$

   Represents the total traffic volume over the considered period.

4. **Number of Peaks:**

$$\mathrm{NumPeaks} = \#\{\, t \mid x(t) > h_{\mathrm{high}} \;\&\; x(t-1) < x(t) > x(t+1) \,\}$$

   Counts the number of local maxima above the high threshold h_{high}.

5. **Average Peak Duration:**

$$\mathrm{AvgPeakDur} = \frac{1}{N_{\mathrm{high}}} \sum_{i=1}^{N_{\mathrm{high}}} d_i$$

   where d_i is the duration of the i-th high-load interval ($x(t) > h_{\mathrm{high}}$).

6. **Average Low Period Duration:**

$$\mathrm{AvgLowDur} = \frac{1}{N_{\mathrm{low}}} \sum_{i=1}^{N_{\mathrm{low}}} \ell_i$$

   where ℓ_i is the duration of the i-th low-load interval ($x(t) < h_{\mathrm{low}}$).

7. **Mean Ramp-Up Rate:**

$$R_i = \frac{1}{4} \sum_{j=p_i-4}^{p_i-1} \Delta x(j), \qquad \text{MeanRampUp} = \frac{1}{N_{\text{peaks}}} \sum_{i=1}^{N_{\text{peaks}}} R_i$$

where p_i is the position of the i-th peak.

8. **Weekday Mean Load:**

$$\mu_{\text{weekday}} = \frac{1}{480} \sum_{t=1}^{480} x(t)$$

9. **Weekend Mean Load:**

$$\mu_{\text{weekend}} = \frac{1}{192} \sum_{t=481}^{672} x(t)$$

10. **Weekend/Weekday Ratio:**

$$R_{W/W} = \frac{\mu_{\text{weekend}}}{\mu_{\text{weekday}}}$$

11. **Absolute Max Traffic:**

$$x_{\max} = \max_t |x(t)|$$

12. **Variance of Peak Duration:**

$$\text{VarPeakDur} = \text{Var}(d_1, d_2, \ldots, d_{N_{\text{high}}})$$

13. **Variance of Low Period Duration:**

$$\text{VarLowDur} = \text{Var}(\ell_1, \ell_2, \ldots, \ell_{N_{\text{low}}})$$

14. **Variance of Weekday Load:**

$$\text{VarWeekday} = \text{Var}(x(1), x(2), \ldots, x(480))$$

15. **Variance of Weekend Load:**

$$\text{VarWeekend} = \text{Var}(x(481), x(482), \ldots, x(672))$$

16. **Spectral Entropy:**

$$P[k] = \frac{A[k]}{\sum_m A[m]}, \qquad H = -\sum_k P[k] \log_2 P[k]$$

where $A[k]$ is the amplitude spectrum. Lower entropy indicates stronger periodicity.

17. **Spectral Centroid:**

$$C = \frac{\sum_k f_k A[k]}{\sum_k A[k]}$$

Describes the **center of mass** of the spectrum; higher C implies faster variations.

18. **Phase of Dominant Frequency:**

$$\phi_{\text{dom}} = \arg\max_k A[k]$$

Represents the phase shift of the dominant periodic component, i.e., the timing of the main daily or weekly cycle.

Bibliography

[1] Ericsson. Ericsson mobility report november 2025, November 2025.

[2] Nicola Piovesan, David López-Pérez, Antonio De Domenico, Xinli Geng, Harvey Bao, and Mérouane Debbah. Machine learning and analytical power consumption models for 5G base stations. *IEEE Communications Magazine*, 60(10):56–62, 2022.

[3] Next G Alliance. Green G: The path toward sustainable 6G. Technical report, Alliance for Telecommunications Industry Solutions (ATIS), 2022.

[4] Daniela Renga, Michela Meo, and Loutfi Nuaymi. Network sharing: A pathway to sustainability and carbon footprint mitigation in radio access networks. In *Proceedings of the 36th International Teletraffic Congress (ITC 36)*, 2025.

[5] Pantelis Koutroumpis, Pau Castells, and Kalvin Bahia. To share or not to share? the impact of mobile network sharing for consumers and operators. *Information Economics and Policy*, 65:101061, 2023.

[6] Nina Slamnik-Kriještorac, Haris Kremo, Marco Ruffini, and Johann M. Marquez-Barja. Sharing distributed and heterogeneous resources toward end-to-end 5G networks: A comprehensive survey and a taxonomy. *IEEE Communications Surveys & Tutorials*, 22(3):1592–1628, 2020.

[7] GSMA. RAN sharing: Overview of regulatory and competition aspects. Technical report, GSMA, 2019.

[8] Nima Afraz, Frank Slyne, Harleen Gill, and Marco Ruffini. Evolution of access network sharing and its role in 5G networks. *Applied Sciences*, 9(21):4566, 2019.

[9] Autorité de régulation des communications électroniques et des postes (ARCEP). Lignes directrices: Partage de réseaux mobiles. Technical report, ARCEP, 2016.

[10] Autorité de régulation des communications électroniques et des postes (ARCEP). Décision n. 2016-0076 du 2 février 2016 approuvant un projet de contrat de partage des infrastructures mobiles en zones blanches. Journal officiel de la République française (JORF) n. 0073, 26 mars 2016, texte n. 97, 2016.

[11] Centro di Ateneo per i Diritti Umani "Antonio Papisca", Università degli Studi di Padova. The italian communications regulatory authority: Functions and composition in 2025, 2025.

[12] ARCEP. Mobile network sharing. ARCEP, 2025.

[13] AGCOM. Agcom opinion on iliad and wind tre MOCN joint venture. AGCOM, 2025.

[14] Lorela Cano, Antonio Capone, and Brunilde Sansò. On the evolution of infrastructure sharing in mobile networks: A survey. *ITU Journal on Future and Evolving Technologies*, 1(1), December 2020.

[15] Paolo Di Francesco, Francesco Malandrino, and Luiz A. DaSilva. Mobile network sharing between operators: A demand trace-driven study. In *Proceedings of the 2014 ACM SIGCOMM Workshop on Capacity Sharing Workshop (CSWS '14)*, pages 39–44, 2014.

[16] Konstantinos Samdanis, Xavier Costa-Perez, and Vincenzo Sciancalepore. From network sharing to multi-tenancy: The 5G network slice broker. *IEEE Communications Magazine*, 2016.

[17] 3GPP. Network sharing evolution, 2024.

[18] Marco Ajmone Marsan and Michela Meo. Network sharing and its energy benefits: A study of european mobile network operators. Technical report, Politecnico di Torino, 2010.

[19] Daniela Renga, Maoquan Ni, Marco Giuseppe Ajmone Marsan, and Michela Meo. Sharing RANs for energy efficiency. In *Proceedings of the 25th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024.

[20] Maoquan Ni, Daniela Renga, Marco Ajmone Marsan, and Michela Meo. On fairness in network sharing. In *Proceedings of the IEEE 31st International Symposium on Local and Metropolitan Area Networks (LANMAN)*, pages 1–6. IEEE, 2025.

[21] Angelos Antonopoulos. Bankruptcy problem in network sharing: Fundamentals, applications and challenges. *IEEE Wireless Communications*, 27(4):81–87, 2020.

[22] Orlando E. Martínez-Durive, Sachit Mishra, Cezary Ziemlicki, Stefania Rubrichi, Zbigniew Smoreda, and Marco Fiore. The NetMob23 dataset: A high-resolution multi-region service-level mobile data traffic cartography. arXiv preprint, 2023.

[23] ANFR. Observatoire des sites 4G en service en france, 2025.

[24] Maoquan Ni. Sustainability in mobile networks: Analyzing the feasibility and benefits of infrastructure sharing. Master's degree thesis, Politecnico di Torino, March 2024.

[25] W. M. Wajda, G. Auer, P. Skillermark, and Y. Jading. INFOSO-ICT-247733 EARTH Deliverable D2.3: Energy Efficiency Analysis of the Reference Systems, Areas of Improvements and Target Breakdown. Project Deliverable D2.3, EARTH Project (FP7), European Commission, 2012. [Online]. Accessed: 2026-02-08.

[26] Han Wang, Junjie Xu, Xing Zhang, Pan Hui, and Dawn Song. Understanding mobile traffic patterns of large scale cellular towers in urban environment. In *Proceedings of the 2015 Internet Measurement Conference (IMC)*, pages 225–238, Tokyo, Japan, 2015. ACM.

[27] OpenStreetMap contributors. Openstreetmap. Online, 2026. [Online]. Accessed: 2026-02-08.