BULGARIAN ACADEMY OF SCIENCES
INSTITUTE OF INFORMATION AND COMMUNICATION TECHNOLOGIES
Department: “Information Processes and Decision Support Systems”

Diellza Shani Berisha

Optimization Strategies and Algorithms for Solving Health System Problems During a Pandemic

Dissertation Thesis

for acquiring Educational and Scientific Degree “DOCTOR”

In Professional field 4.6. “Informatics and Computer Science”,
in Doctoral Program: Informatics

Scientific Supervisor: Prof. Dr. Vassil Georgiev Guliashki

Sofia, 2026

Table of Contents

Dedication

I dedicate this work wholeheartedly to my parents, to my unforgettable sister, who lives forever in my soul, and to my daughter, all of whom have been my light in the darkness, my strength in moments of exhaustion, and the reason I continued moving forward without giving up

Acknowledgments

With deep respect and sincere gratitude, I would like to thank all those who have supported me throughout this long and meaningful academic journey. This dissertation is not only the result of my scientific work, but also of the encouragement, sacrifice, understanding, and support I received during every stage of this process.

First and foremost, I express my deepest gratitude to my beloved parents for their unconditional love, sacrifice, patience, prayers, and continuous belief in me. Their support has been a constant source of strength and motivation, and this achievement belongs to them as well.

With profound love and lasting remembrance, I also thank my beloved sister, who is no longer physically present with me. Her memory, love, and spiritual presence have continued to give me strength throughout this journey. I dedicate this achievement also to her, with endless love and gratitude.

I sincerely thank my husband for his patience, understanding, and continuous support during my doctoral studies. His encouragement and stability helped me overcome the challenges of this demanding process.

My special gratitude goes to my daughter, Aja, who is my light, joy, and greatest motivation. Her presence has given me extraordinary strength to continue, and this achievement is also for her, as a reflection of a mother’s love, dedication, and perseverance.

With profound respect and sincere appreciation, I express my deepest gratitude to my scientific supervisor, Prof. Dr. Vassil Georgiev Guliashki, for his guidance, academic support, professional advice, patience, and continuous encouragement. His supervision has been essential to my scientific development and to the preparation of this dissertation.

I also sincerely thank Professor Galia for her kindness, support, valuable advice, and encouragement. I extend my gratitude to Rector Edmond Hajrizi for his support and contribution to an academic environment that promotes knowledge, research, and professional advancement.

Finally, I thank my friends, colleagues, and all those who supported and believed in me throughout this journey. This dissertation represents both an important academic achievement and a personal journey shaped by love, sacrifice, resilience, faith, and dedication.

Dissertation Structure

This dissertation is organized in a coherent sequence that guides the reader from the research problem to the proposed methodology, empirical preparation, computational evaluation, and final synthesis. The structure reflects the main objective of the study: to develop and evaluate optimization strategies and algorithms for supporting healthcare decision-making during pandemic conditions, with emphasis on resource-constrained healthcare systems and the Kosovo COVID-19 case study.

The introductory part presents the research background, motivation, problem statement, aim and objectives, scientific novelty, and main contributions. It positions pandemic healthcare management as an optimization and decision-support problem requiring forecasting, allocation, simulation, and operational evaluation.

Chapter 1 provides the literature review and theoretical foundation. It examines the multi-dimensional impact of pandemics and analyses existing approaches in epidemiological forecasting, healthcare operations optimization, supply-chain coordination, economic-health trade-offs, equity-aware optimization, and simulation-based decision support. The chapter identifies the gaps that motivate the proposed framework.

Chapter 2 presents the proposed integrated hybrid decision-support framework. It defines the problem scope, modelling assumptions, system architecture, epidemiological component, healthcare resource model, optimization strategy, simulation layer, and adaptive workflow. This chapter establishes the methodological basis for linking prediction, prescription, and operational validation.

Chapter 3 describes the empirical data collection, parameterization, and experimental design for the Kosovo COVID-19 case study. It presents the data sources, municipality inclusion logic, reliability controls, derived indicators, preprocessing, normalization, MATLAB input preparation, and scenario design required for the SEIR-kNN, MILP, and DES experiments.

Chapter 4 reports the computational results and analysis. It presents the SEIR-kNN forecasting and demand-estimation results, MILP-based hospital allocation outcomes, DES-based hospital-operations simulation results, and integrated evaluation of the SEIR-kNN-MILP-DES framework through indicators such as forecast accuracy, hospital and ICU demand, utilization, transfer feasibility, waiting time, throughput, unmet demand, and robustness.

The final part summarizes the thesis contributions, related publications, declaration of originality, bibliography, and appendix materials. Overall, the dissertation follows a complete doctoral pathway: it defines the problem, reviews existing knowledge, develops the methodology, prepares the case study, evaluates the results, and synthesizes the main contributions.

Keywords

Pandemic healthcare management; healthcare optimization; epidemiological forecasting; SEIR-kNN; mixed-integer linear programming; discrete event simulation; hospital capacity allocation; ICU demand management; robust decision support; Kosovo COVID-19 case study.

Roadmap of the dissertation

The thesis is structured as a coherent research roadmap that moves from the identification of the pandemic healthcare-management problem to the development, empirical parameterization, computational validation, and final synthesis of the proposed optimization framework.

The roadmap below summarizes the logical role of each part of the dissertation and shows how the individual chapters contribute to the integrated forecasting–optimization–simulation approach.

Introduction	Defines the research motivation, problem statement, aim, objectives, scientific novelty and contributions. It positions pandemic healthcare management as an optimization and decision-support problem requiring forecasting, allocation and operational evaluation.
Chapter 1	Builds the theoretical foundation through a critical review of pandemic impacts, epidemiological forecasting, healthcare operations optimization, logistics, equity-aware optimization and adaptive simulation-based decision support. The chapter identifies the gaps that justify the proposed framework.
Chapter 2	Develops the Integrated Hybrid Decision-Support Framework. It formalizes the decision problem and connects epidemiological modelling, healthcare-resource representation, robust multi-objective optimization, MILP-based allocation, DES validation and rolling-horizon adaptive logic.
Chapter 3	Prepares the Kosovo COVID-19 case study. It defines the empirical data sources, municipality inclusion logic, reliability controls, derived indicators, preprocessing, normalization, MATLAB input preparation and parameterization of the SEIR-kNN, MILP and DES experiments

Chapter 4	Presents the computational findings. It reports SEIR-kNN forecasting and demand estimation, MILP hospital allocation, peak-pressure scenarios, DES hospital-operation simulation, integrated KPI evaluation, sensitivity analysis, equity implications and practical decision-support interpretation.
Final part	Synthesizes the research outcomes through the thesis contributions, publications related to the dissertation, declaration of originality, bibliography and appendices containing reproducibility materials, supplementary data, model equations and MATLAB resources.

Methodologically, the dissertation follows the sequence: *pandemic healthcare problem definition → critical gap analysis → integrated hybrid framework design → Kosovo case-study data preparation → SEIR-kNN forecasting → MILP allocation → DES operational validation → integrated KPI, robustness and equity evaluation → final conclusions and contributions.*

.

List of Abbreviations

Abbreviation	Description
ABM	Agent-Based Modeling
AI	Artificial Intelligence
BMO	Barnacles Mating Optimizer
CoG	Centre of Government
COVID-19	Coronavirus Disease 2019
CPLEX	Optimization solver
CV	Coefficient of Variation
DES	Discrete Event Simulation
EMS	Emergency Medical Services
GDP	Gross Domestic Product
GPU	Graphics Processing Unit
GAMS	General Algebraic Modeling System
ICU	Intensive Care Unit
IHDSF	Integrated Hybrid Decision-Support Framework
ILO	International Labour Organization
KPI	Key Performance Indicator
kNN	k-Nearest Neighbour
LMIC	Low- and Middle-Income Country
LOS	Length of Stay
LSSVM	Least Squares Support Vector Machine
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MATLAB	Matrix Laboratory computing environment
MILP	Mixed-Integer Linear Programming
MIP	Mixed-Integer Programming
MPC	Model Predictive Control
MSSP	Multistage Stochastic Programming
NLP	Nonlinear Programming
NPI	Non-Pharmaceutical Intervention
ODE	Ordinary Differential Equation
OECD	Organisation for Economic Co-operation and Development
OR	Operations Research
PPE	Personal Protective Equipment
RMSE	Root Mean Square Error
SEIR	Susceptible-Exposed-Infectious-Recovered model
SEIR-H-ICU-D	Extended SEIR with hospitalized, ICU and deceased states
SEIR-kNN	SEIR with kNN parameter adaptation
SEIR-kNN-MILP-DES	Integrated forecasting-allocation-simulation framework
SIR	Susceptible-Infectious-Recovered model
SVEIR	Susceptible-Vaccinated-Exposed-Infectious-Recovered model
UNESCO	United Nations Educational, Scientific and Cultural Organization
WHO	World Health Organization

Introduction

The COVID-19 pandemic demonstrated that contemporary healthcare systems are exposed to complex, rapidly evolving, and highly uncertain crisis conditions. The sudden increase in infections, hospital admissions, intensive care demand, workforce pressure, and medical-supply shortages revealed that pandemic management cannot be treated only as a clinical or epidemiological issue. It is also an optimization and decision-support problem, because decision-makers must allocate scarce resources, coordinate hospitals, estimate future demand, and maintain service continuity while information is incomplete and conditions change over time.

This dissertation is focused on optimization strategies and algorithms for solving health-system problems during a pandemic. The research is positioned at the intersection of epidemiological modelling, operations research, healthcare resource allocation, machine-learning-based forecasting, and simulation-based evaluation. Its central idea is that pandemic response can be improved when forecasting, optimization, and operational simulation are not developed as separate tools, but integrated into one coherent decision-support framework.

The dissertation gives particular attention to resource-constrained healthcare systems, with Kosovo used as the empirical case study. Such systems often face limited hospital capacity, fragmented data reporting, constrained analytical infrastructure, and unequal regional access to healthcare services. These characteristics make the development of scalable, transparent, and practically applicable optimization methods especially important. The purpose of the introduction is to present the research motivation, define the problem addressed by the dissertation, formulate the aim and objectives, clarify the scientific novelty and contributions, and outline the structure of the dissertation.

Research Motivation and Problem Statement

The motivation for this research arises from the operational challenges experienced by healthcare systems during pandemic waves. When infection growth accelerates, hospitals must respond to uncertain and uneven demand. Intensive care units may become overloaded, patient waiting times may increase, staff may be reassigned under pressure, and available resources may not be distributed according to real-time needs. These challenges are amplified when hospital networks operate with limited coordination and when decision-makers do not have reliable predictive and optimization tools.

A major limitation identified in pandemic-response practice and literature is the fragmentation of analytical approaches. Epidemiological models are often used to estimate infection trajectories, but their outputs are not always connected directly to hospital-capacity planning. Optimization models may allocate resources, but they frequently treat demand as fixed or deterministic. Simulation models can reproduce operational bottlenecks, but they are often applied after decisions have already been made rather than as part of an adaptive planning process. As a result, decision support remains incomplete when forecasting, allocation, and operational evaluation are not integrated.

The problem becomes more difficult in resource-constrained environments. Data may be delayed, incomplete, or inconsistent; hospital capacities may differ substantially across regions; and the computational infrastructure required for sophisticated models may be limited. In such conditions, complex models that work in highly resourced settings may not be directly transferable. Therefore, the dissertation is motivated by the need for a framework that is scientifically rigorous but also feasible for practical implementation under imperfect data and limited resources.

The central research problem of the dissertation can be expressed as follows: how can epidemiological forecasting, healthcare resource optimization, and hospital-operation simulation be integrated into a scalable and uncertainty-aware decision-support framework for improving pandemic healthcare management in resource-constrained systems?

This problem involves several interrelated sub-problems: estimating future healthcare demand from epidemic dynamics; transforming predicted infections into hospital and ICU demand indicators; allocating patients or resources across facilities under capacity constraints; evaluating operational performance through waiting time, utilization, and unmet demand; and supporting adaptive decisions under uncertainty. Addressing these sub-problems requires a hybrid methodological approach rather than a single isolated model.

Aim and Objectives of the Dissertation

The main aim of this dissertation is to develop and evaluate an integrated hybrid decision-support framework that combines epidemiological forecasting, optimization-based resource allocation, and discrete event simulation for improving healthcare-system performance during pandemic conditions.

To achieve this aim, the dissertation defines the following objectives:

1. To analyse the state of the art in pandemic healthcare modelling, including epidemiological forecasting, healthcare operations optimization, hospital-capacity planning, logistics coordination, equity-aware allocation, and simulation-based decision support.
2. To identify the main methodological and practical gaps in existing approaches, with special attention to model fragmentation, data uncertainty, computational scalability, equity limitations, and applicability in resource-constrained healthcare systems.
3. To design an integrated hybrid decision-support framework that links SEIR-based epidemiological modelling, k-nearest-neighbour-supported demand estimation, mixed-integer linear programming, and discrete event simulation into a coherent workflow.
4. To formulate healthcare resource allocation as an optimization problem that considers demand, capacity, hospital utilization, transfer feasibility, and shortage minimization under pandemic pressure.

5. To construct and parameterize a Kosovo COVID-19 case study by preparing municipality-level epidemiological, demographic, and healthcare-capacity indicators suitable for computational experimentation.
6. To evaluate the proposed framework through experimental scenarios, including baseline demand, peak-pressure conditions, coordinated allocation, and optimized hospital-operation configurations.
7. To assess the decision value of the integrated framework using key performance indicators such as forecast accuracy, hospital and ICU demand, utilization, waiting time, unmet demand, and operational improvement.
8. To demonstrate how the proposed approach can support more evidence-based, adaptive, and operationally feasible pandemic decision-making in resource-constrained healthcare environments.

These objectives define a structured pathway from theoretical gap analysis to methodological development, empirical data preparation, computational experimentation, and evaluation of results. They also ensure that the dissertation remains focused on both scientific contribution and practical relevance.

Scientific Novelty and Contributions

The scientific novelty of the dissertation lies in the integration of forecasting, optimization, and simulation into a single decision-support framework for pandemic healthcare management. Instead of treating epidemic prediction, resource allocation, and operational evaluation as separate tasks, the proposed framework connects them sequentially and functionally. Forecasting outputs are transformed into demand indicators; demand indicators are used in optimization models; and optimization results are evaluated through simulation-based operational scenarios.

The dissertation contributes to the field of healthcare optimization and pandemic decision support in the following ways:

1. Integrated hybrid framework: The dissertation proposes a structured forecasting-optimization-simulation architecture for pandemic healthcare management. This framework supports the transition from epidemiological signals to operational decisions.
2. SEIR-kNN demand-estimation logic: The research combines interpretable epidemic modelling with similarity-based parameter adaptation in order to support municipality-level demand estimation under limited data conditions.
3. MILP-based hospital allocation model: The dissertation formulates coordinated hospital allocation as a mixed-integer linear programming problem, enabling capacity-aware distribution of demand and reduction of localized overload.

4. DES-based operational evaluation: Discrete event simulation is used to evaluate patient flow, waiting time, ICU utilization, and operational bottlenecks under alternative pandemic scenarios.
5. Resource-constrained case-study application: The framework is applied to Kosovo, offering a context-aware contribution for healthcare systems where data, capacity, and analytical infrastructure may be limited.
6. KPI-based integrated evaluation: The dissertation uses measurable performance indicators to compare baseline, peak-demand, and optimized scenarios, thereby supporting transparent assessment of the proposed decision-support logic.
7. Practical decision-support relevance: The proposed approach is designed not only as a theoretical model but also as a practical workflow that can support healthcare managers and policymakers during pandemic emergencies.

Through these contributions, the dissertation strengthens the connection between operations research, epidemiological modelling, healthcare management, and applied decision-support systems. Its novelty is not limited to the individual use of SEIR models, MILP, kNN, or DES, but is found in their structured integration for a specific and practically relevant healthcare problem.

Dissertation Organization

The dissertation is organized in a logical sequence that follows the development of the research from theoretical background to methodological formulation, empirical preparation, computational experimentation, and final evaluation.

1. The introductory part presents the motivation for the research, the problem statement, the aim and objectives, the scientific novelty, and the organization of the dissertation.
2. Chapter 1 presents the literature review and theoretical foundation of the research. It analyses optimization strategies and algorithms used for pandemic healthcare management, including epidemiological models, healthcare operations optimization, supply-chain coordination, economic-health trade-offs, equity-aware optimization, and simulation-based decision-support frameworks. The chapter identifies the main gaps that motivate the proposed research framework.
3. Chapter 2 develops the proposed integrated hybrid decision-support framework. It defines the formal problem scope, modelling assumptions, system architecture, epidemiological component, healthcare resource model, optimization strategy, simulation layer, and adaptive workflow. This chapter establishes the methodological foundation for linking forecasting, optimization, and simulation.

4. Chapter 3 presents the empirical data collection, parameterization, and experimental design. It describes the Kosovo COVID-19 case study, data sources, municipality inclusion logic, data preprocessing, indicator construction, MATLAB input preparation, and parameterization of the SEIR-kNN, MILP, and DES experiments. It also defines the experimental scenarios and reproducibility setup.
5. Chapter 4 reports and discusses the experimental findings. It presents the results of SEIR-kNN forecasting, MILP-based hospital allocation, DES-based hospital-operation simulation, integrated KPI evaluation, ablation comparison, sensitivity analysis, limitations, and practical implications for pandemic healthcare management.
6. The final part of the dissertation summarizes the main conclusions, thesis contributions, publications related to the dissertation, declaration of originality, bibliography, and appendices.

This structure follows a coherent doctoral research logic: the dissertation first explains why the problem is important, then reviews the existing knowledge, develops the proposed methodology, prepares the empirical case study, evaluates the results, and finally summarizes the scientific and practical contributions. In this way, the dissertation presents a complete and systematic research pathway for optimization strategies and algorithms applied to pandemic healthc

Chapter 1 – Literature Review on Optimization Strategies and Algorithms for Addressing Health System Challenges in Pandemics

1.1 Introduction

This chapter presents a systematic and critical review of optimization strategies, mathematical models, and decision-support frameworks developed to address healthcare system challenges during pandemics. The review focuses on the integration of epidemiological forecasting, healthcare capacity planning, supply chain coordination, economic considerations, governance mechanisms, and behavioral dynamics. Particular attention is given to the applicability, scalability, and operational feasibility of these approaches in resource-constrained healthcare systems.

The COVID-19 pandemic constituted an unprecedented global systemic shock. Since its emergence in late 2019, it has generated simultaneous pressures across healthcare, economic, and social systems. The World Health Organization (WHO) estimated approximately 14.9 million excess deaths during 2020–2021, while intensive care unit (ICU) overload was reported in a majority of affected countries [1], [2]. Even highly developed healthcare systems experienced severe shortages of ICU beds, ventilators, trained medical personnel, vaccines, and essential medical supplies [3], [4]. These shortages were further exacerbated by disruptions in global medical supply chains, infection-related workforce depletion, and rapidly fluctuating patient demand.

The crisis revealed that healthcare systems operate as components of tightly interconnected socio-economic networks. Public health measures implemented to contain viral transmission including lockdowns, mobility restrictions, and business closures triggered significant economic contraction. According to OECD assessments, global gross domestic product (GDP) declined by approximately 3.4% in 2020, marking the sharpest economic downturn in decades [5]. Simultaneously, labor market disruptions and prolonged school closures intensified social inequalities, disproportionately affecting vulnerable populations and low- and middle-income countries (LMICs) [6], [7]. These cascading and mutually reinforcing effects demonstrate that pandemic management cannot be treated as a purely epidemiological or operational problem;

rather, it represents a multi-layered dynamic system challenge requiring coordinated decision-making under uncertainty.

In response to these complexities, a substantial body of research has emerged at the intersection of operations research, epidemiology, health economics, and public policy. Mathematical and algorithmic approaches have been proposed to support critical decision-making tasks, including ICU capacity planning, vaccination and testing logistics, hospital network coordination, and adaptive intervention policies. These approaches encompass stochastic programming, mixed-integer linear and nonlinear optimization, dynamic programming, agent-based modeling, compartmental epidemiological models (SIR/SEIR), and hybrid metaheuristic frameworks [8], [9], [10], [11], [12]. More recent contributions have incorporated machine learning techniques for demand forecasting and uncertainty quantification, as well as equity-aware optimization formulations designed to address disparities in healthcare access [13], [14].

While the methodological diversity of this literature is considerable, its structural coherence remains limited. Many studies focus on isolated components of pandemic response such as ICU bed allocation, epidemiological forecasting, or supply chain resilience without integrating these elements into unified and adaptive decision-support architectures. Epidemiological models frequently operate independently of healthcare operations models; economic assessments are often detached from hospital capacity constraints; and behavioral dynamics are rarely embedded within formal optimization structures. This fragmentation constrains the ability of policymakers to evaluate trade-offs across health outcomes, economic stability, and social equity in a coherent and timely manner [15], [16].

A further limitation concerns operational feasibility. A significant proportion of advanced optimization and simulation-based models assume access to high-frequency, high-quality data streams and substantial computational resources. Such assumptions are rarely satisfied during large-scale health emergencies and are particularly unrealistic in resource-constrained environments characterized by fragmented reporting systems and limited analytical infrastructure [15], [17]. Consequently, the practical deployment of many theoretically sophisticated models remains restricted, especially in LMIC contexts.

From a contextual perspective, the limited attention devoted to resource-constrained healthcare systems represents a critical research gap. Regions such as Kosovo and the Western Balkans face structural constraints including limited surge capacity, constrained fiscal space, incomplete data availability, and institutional fragmentation. Models developed and validated in high-income settings therefore exhibit reduced transferability and scalability when applied to these environments [18], [19]. There is a pressing need for adaptive, computationally efficient, and context-aware optimization frameworks capable of operating under uncertainty and data scarcity while maintaining policy relevance.

This dissertation addresses these challenges by positioning pandemic response as an integrated, multi-objective, and dynamic optimization problem. It seeks to bridge the gap between methodological sophistication and operational applicability by developing a hybrid decision-support framework that combines epidemiological forecasting, healthcare resource optimization, supply chain coordination, and behavioral considerations within a unified architecture. The proposed framework emphasizes scalability, robustness under uncertainty, and explicit incorporation of equity and contextual constraints, with particular relevance for resource-limited healthcare systems. This methodological positioning is strengthened by robust optimization, multicriteria optimization, and rolling-horizon control theory [20], [21], [21], [22], [23], [24], [25], [26].

Accordingly, this chapter systematically classifies and critically evaluates existing optimization and coordination strategies, identifies methodological and contextual limitations, and delineates the research gaps that motivate the research goal and tasks presented in Section 1.5. Through this structured analysis, the chapter establishes the conceptual and methodological foundation for the integrated optimization framework developed in the subsequent chapters.

1.2. Systemic Multi-Dimensional Impact of Pandemics

The COVID-19 pandemic demonstrated that modern health crises are not isolated epidemiological events but complex systemic shocks affecting healthcare, economic, and social subsystems simultaneously. Between 2020 and 2021, the World Health Organization (WHO) estimated approximately 14.9 million excess deaths worldwide, while intensive care unit (ICU) overload was reported in a majority of affected countries [1], [2]. Beyond direct mortality, the pandemic triggered cascading disruptions across labor markets, education systems, and public finances, reinforcing pre-existing structural inequalities.

Unlike earlier outbreaks such as SARS (2003) or H1N1 (2009), COVID-19 unfolded within highly interconnected global networks characterized by intense mobility, tightly coupled supply chains, and synchronized policy responses. This structural interconnectedness amplified systemic stress and created feedback loops between health system strain, economic contraction, and social disruption [27], [28], [29]. The pandemic thus revealed the limitations of siloed policy and modelling approaches and underscored the necessity for integrated decision-support frameworks capable of addressing multi-layer interactions under uncertainty.

1.2.1. Health System Strain and ICU Overload

The most visible manifestation of systemic stress during COVID-19 was the rapid saturation of hospital and ICU capacity. Severe shortages of beds, ventilators, oxygen supplies, and trained personnel were observed across both high-income countries and low- and middle-income countries (LMICs) [3], [4]. In several regions, ICU occupancy exceeded baseline capacity during peak waves, necessitating emergency infrastructure expansion, inter-regional patient transfers, and temporary field hospitals.

Comparative analyses indicate that ICU overload was driven not only by infection incidence but also by pre-existing structural constraints, including limited surge capacity, inflexible staffing structures, fragmented hospital networks, and vulnerable medical supply chains [30], [8]. Systems with stronger coordination mechanisms and integrated data infrastructures generally demonstrated greater adaptive capacity.

In response, optimization and simulation models were developed to support ICU capacity planning, patient prioritization, and dynamic staff allocation. Multistage stochastic programming approaches have enabled proactive allocation under demand uncertainty [9], while hybrid epidemiological–optimization models have integrated infection dynamics with operational constraints [8]. Equity-aware formulations incorporating fairness metrics have further sought to mitigate disparities in critical care access [13].

However, systematic reviews highlight persistent barriers to real-time implementation. Many models require granular, high-frequency data and extensive computational resources, limiting their deployment during crisis conditions, particularly in LMIC contexts characterized by delayed reporting and fragmented information systems [15], [15]. Moreover, ICU optimization is frequently treated as an isolated operational problem, insufficiently integrated with upstream epidemiological forecasts and downstream supply chain constraints. These limitations expose a need for adaptive, computationally efficient frameworks capable of operating under data scarcity while maintaining policy relevance.

1.2.2. Economic Disruptions and Health–Economy Trade-offs

The public health measures required to contain viral transmission including lockdowns, mobility restrictions, and sectoral shutdowns generated profound macroeconomic effects. Global GDP contracted by approximately 3.4% in 2020, and more than 220 million full-time equivalent jobs were lost, disproportionately affecting informal workers, women, and young populations [5], [6], [6].

The magnitude of economic contraction varied substantially across countries, reflecting differences in fiscal capacity, economic structure, digital infrastructure, and labor market composition [27], [28]. LMICs faced particularly severe trade-offs due to constrained fiscal space and limited access to external financing, restricting their ability to simultaneously expand healthcare capacity and provide economic support [29], [31]. Additional macroeconomic studies confirm that pandemic policy responses produced heterogeneous sectoral, fiscal, and labor-market effects across countries and income groups [32], [33], [34], [15], [32], [35], [14], [36], [15], [17], [16], [18], [19], [15], [30], [8], [37].

Traditional macroeconomic models often treated health system strain as an exogenous shock, while epidemiological models largely abstracted from economic constraints. More recent hybrid economic–epidemiological frameworks have attempted to internalize these interactions by coupling infection dynamics with labor supply shocks and output losses [12], [38], [12]. Multi-objective optimization approaches further formalize trade-offs between infection control measures and economic sustainability [38], [39].

Despite their analytical advances, many integrated models remain computationally intensive and data-dependent, limiting their practical applicability in real-time crisis management, especially in resource-constrained environments. Simplified and scalable formulations that preserve essential health–economy interactions while remaining feasible under uncertainty remain underdeveloped.

1.2.3. Social Inequalities, Behavioral Heterogeneity, and Access Constraints

The pandemic amplified pre-existing social inequalities. Empirical studies demonstrate that infection risk, hospitalization rates, and mortality outcomes were strongly correlated with socio-economic status, housing density, occupational exposure, and access to healthcare services [40], [41]. Prolonged school closures affected approximately 1.6 billion students globally, with learning losses disproportionately concentrated among low-income and rural populations [7], [42], [43]. Health equity and behavioral studies further show that unequal exposure, trust, health literacy, and access barriers shaped infection and mortality outcomes [44], [45], [44], [45], [46], [47], [48], [49], [45], [46], [17], [30], [37], [50], [51], [52], [53], [46], [54], [55], [56], [27], [28], [29], [57], [58].

Disparities in testing access, treatment availability, and vaccination uptake further intensified inequities, particularly among marginalized communities [14], [59], [60]. These patterns were closely linked to structural determinants such as insurance coverage gaps, geographic accessibility, language barriers, and health literacy.

From a modelling perspective, social vulnerability and behavioral heterogeneity remain insufficiently integrated into formal optimization frameworks. Many models assume homogeneous compliance with non-pharmaceutical interventions and uniform access to healthcare services, oversimplifying real-world dynamics [44]. Agent-based and hybrid simulation studies demonstrate that incorporating heterogeneous behavioral responses significantly alters projected infection trajectories and resource demand patterns [61], [62].

Although equity-aware optimization models incorporating fairness metrics and vulnerability indices have been proposed [63], [64], equity considerations are often treated as secondary objectives rather than structurally embedded design principles. Furthermore, complex fairness-based formulations frequently depend on detailed socio-economic data that may be unavailable in LMIC contexts [18], [45]. This creates a tension between ethical rigor and operational feasibility, highlighting the need for parsimonious equity-aware frameworks adaptable to data-sparse environments.

1.2.4. Integrated System Perspective

The evidence across health, economic, and social domains demonstrates that the COVID-19 pandemic constituted a multi-layered systemic crisis characterized by reinforcing feedback loops. Health system overload triggered restrictive public health measures; these measures generated economic contraction and social disruption; and rising inequality and behavioral responses, in turn, influenced infection dynamics and healthcare demand.

Tables 1.1 and 1.1A and Figure 1.1. synthesize these interconnected effects. Rather than representing independent outcome categories, health strain, economic contraction, and social inequality form an integrated crisis loop. This systemic interdependence challenges traditional single-objective or siloed modelling approaches.

Table 1.1. Dimensions of Pandemic Impact (2020–2021)

Dimension	Key Indicators	Sources
Health Impact	Approximately 14.9 million deaths (direct and indirect); ICU overload reported in more than 70% of countries	[1]
Economic Impact	Global GDP contraction of −3.4% in 2020; more than 220 million full-time equivalent job losses	[3], [4]
Social Impact	Around 1.6 billion students affected by school closures; heightened inequality in LMICs	[65]

Table 1.1 synthesizes the core quantitative indicators discussed in the preceding subsections, demonstrating how the health crisis rapidly translated into economic contraction and widespread social disruption.

Table 1.A. highlights the pronounced asymmetries between high-income economies and low- and middle-income countries, underscoring the role of pre-existing structural capacities in shaping pandemic outcomes.

Table 1. A. Comparative Impact: High-income vs. LMIC Regions (2020–2021)

Dimension	High-income Countries	LMICs	Sources
Health (Mortality & ICU Capacity)	Lower mortality rates; ICU overload approximately 50–60%; relatively adequate ventilator and staffing capacity	Mortality rates two to three times higher; ICU overload exceeding 80%; severe shortages of oxygen and trained staff	[1], [3]
Economic (GDP & Employment)	GDP contraction around 2.5%; job losses affecting roughly 5% of the workforce; fiscal stimulus packages often exceeding 5% of GDP	GDP contraction frequently above 5%; informal job losses exceeding 15%; limited fiscal support (approximately 1–2% of GDP)	[3], [66]
Social (Education & Inequality)	Rapid transition to online education; relatively limited digital divide	Over 1.2 billion students affected; pronounced digital divide; restricted access to online education	[65], [3]

Figure 1.1. illustrates the interconnected nature of the COVID-19 crisis across health, economic, and social systems.

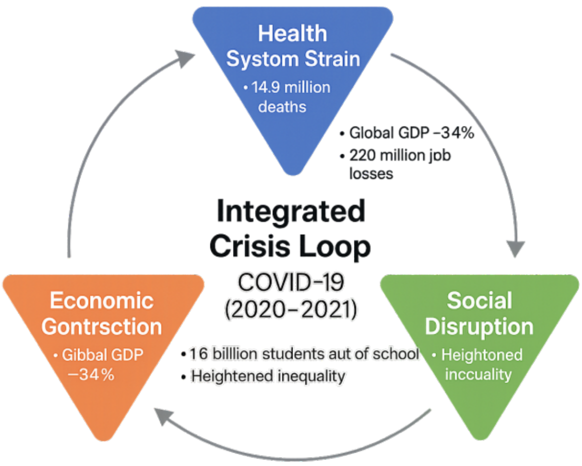

Figure 1.1. Integrated Crisis Loop (COVID-19 health, economic, social impact)

Evidence from international reports indicates that widespread ICU overload, reported in over 70% of countries, triggered stringent public health measures that contributed to a global GDP contraction of 3.4% and substantial employment losses, while prolonged school closures affected approximately 1.6 billion students worldwide [1], [3], [65]. These reinforcing feedback loops intensified inequality and, in turn, further strained healthcare systems.

From an optimization standpoint, pandemic response therefore requires:

- multi-objective formulations balancing health, economic, and equity outcomes,
- stochastic representations of demand and policy uncertainty,
- dynamic, adaptive decision mechanisms capable of operating in rolling horizons,
- explicit modeling of structural and contextual constraints in resource-limited systems.

The multi-dimensional nature of pandemic impact thus provides the foundational rationale for the integrated hybrid decision-support framework developed in this dissertation.

1.2.5 Critical Gap Analysis and Implications

The evidence reviewed in Sections 1.2.1–1.2.4 shows that COVID-19 generated tightly coupled health, economic, and social disruptions, forming reinforcing feedback loops that cannot be addressed with siloed analytical tools. Although the literature offers sophisticated models for forecasting, capacity planning, logistics, and policy analysis (see [45], [53], [55], [41], [42]), practical deployment is often constrained by data delays, limited computing resources, and weak cross-domain integration limitations that are amplified in resource-constrained healthcare systems.

These observations motivate the need for integrated, scalable, and uncertainty-aware decision-support architectures that (i) operate under imperfect data, (ii) remain computationally feasible in real time, and (iii) incorporate equity and contextual constraints as first-class design elements. The following sections therefore review optimization paradigms and hybrid modeling strategies with a focus on operational feasibility and integration potential.

1.3. Categories of Optimization and Coordination Strategies

The COVID-19 pandemic demonstrated that effective pandemic management requires coordinated interventions across multiple domains, integrating policy-level governance mechanisms with advanced, algorithm-driven technical solutions. Empirical evidence from the COVID-19 crisis indicates that fragmented approaches limited the ability of decision-makers to anticipate cross-domain feedback effects, resulting in delayed responses, inefficient resource allocation, and inconsistent policy outcomes. Traditional, siloed approaches where epidemiology, logistics, and socio-economic policies were addressed separately proved insufficient in mitigating cascading crises. Instead, research and practice emphasized the importance of holistic, data-driven strategies that combine multiple decision-making layers [65], [67], [14], [68]. Governance and public administration studies similarly emphasize that institutional coordination and information flows are decisive for crisis performance [31], [38], [39], [69].

1.3.1 Epidemiological and Predictive Modeling Paradigms

Accurate epidemiological forecasting constitutes the analytical foundation of pandemic decision-making. Forecasts of infection trajectories, hospitalization demand, and peak intensity directly influence healthcare capacity planning, resource allocation, and policy timing. Over the past decades, forecasting approaches have evolved from classical compartmental models to hybrid economic–epidemiological structures and, more recently, to machine learning–enhanced predictive systems.

- **Classical Compartmental Models**

Compartmental models such as SIR and SEIR remain the foundational framework for modeling infectious disease dynamics. Originally introduced by Kermack and McKendrick and later formalized and extended by Hethcote [11], these models partition populations into epidemiological states (e.g., Susceptible, Infectious, Recovered, Exposed) governed by differential equations. Extensions incorporating vaccination (SVEIR), waning immunity, and age stratification have improved realism while preserving interpretability. System dynamics and infectious disease modeling literature further emphasizes feedback structures, behavioral responses, and delayed information flows in epidemic processes [57], [70], [71].

The primary strength of compartmental models lies in their transparency and analytical tractability. Their relatively low data requirements and structural simplicity make them suitable for early-stage outbreak analysis and scenario exploration. However, they rely on homogeneity assumptions and fixed parameter structures that limit their ability to capture behavioral heterogeneity, mobility patterns, and rapidly changing intervention effects. Robust uncertainty treatment can complement such reduced-order models when parameter estimates are incomplete or delayed [20].

- **Hybrid Economic–Epidemiological Models**

To address the interaction between infection dynamics and economic activity, hybrid models have been developed that couple epidemiological compartments with macroeconomic or sector-level production structures. Goenka and Liu [12], for example, integrated a SARS-CoV-2 transmission model with a 63-sector economic framework, enabling sector-specific lockdown optimization. Similarly, Laurentz et al. [34] applied hybrid structures to balance ICU capacity constraints with economic output preservation.

Such models formalize health–economy trade-offs and allow policy optimization across competing objectives. By internalizing labor supply shocks, mobility restrictions, and healthcare capacity limits, they provide a more comprehensive representation of systemic interdependencies. However, these frameworks are typically computationally intensive and dependent on detailed sectoral and epidemiological data, limiting their transferability to data-constrained environments.

- **Machine Learning and Metaheuristic-Enhanced Models**

Recent developments incorporate artificial intelligence and metaheuristic optimization to improve predictive accuracy. Mustaffa et al. [72] combined the Barnacles Mating Optimizer (BMO) with Least Squares Support Vector Machines (LSSVM) to forecast daily COVID-19 cases, while Ahmed et al. [73] extended this framework by integrating vaccination dynamics (IBMO-LSSVM). Deep learning approaches have further enabled hotspot detection and real-time cluster identification [33].

Advanced hybrid models increasingly combine compartmental epidemiology, machine learning, and control theory. The SIDARE framework proposed by Kasis et al. [74] integrates agent-based urban mobility simulation with control-theoretic mechanisms, reducing forecasting errors relative to standard SEIR formulations. These approaches illustrate the shift toward adaptive and data-driven predictive systems capable of capturing nonlinear dynamics and policy feedback.

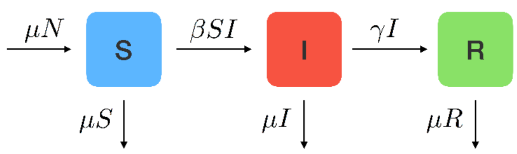

Figure 1.2.(a) Classical Compartmental Models (SIR/SEIR structures)

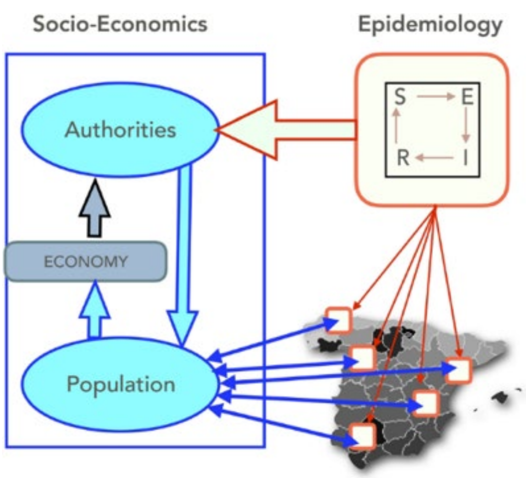

Figure 1.3.(b) Economic Epidemiological hybrid models

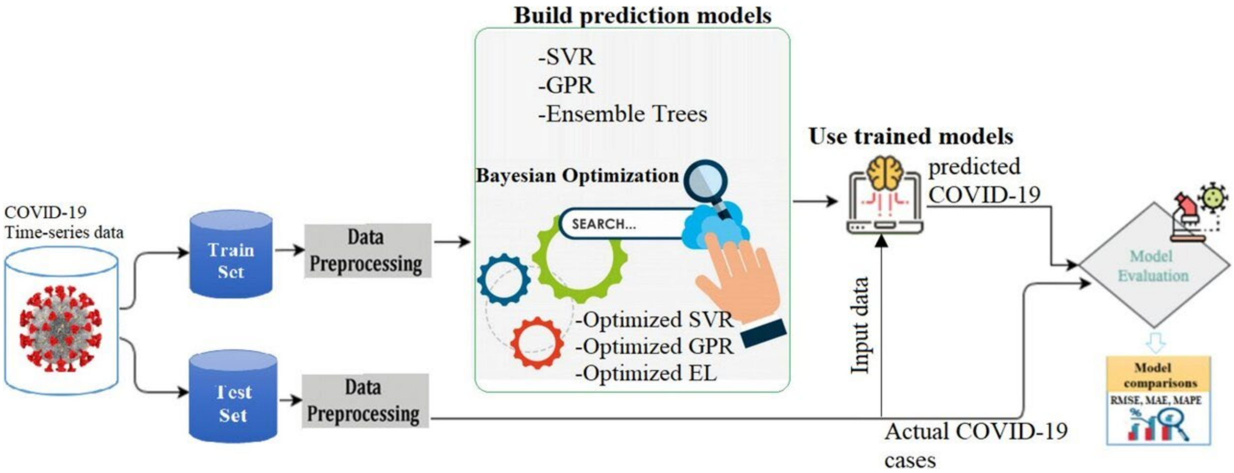

Figure 1.4.(c) AI-Machine enhanced Predictive Frameworks

Figure 1.5. Evolution of Forecasting Models for Pandemic Management

As shown in Figure 1.2(a–c), forecasting models evolved from classical compartmental structures to hybrid and AI-enhanced architectures. This progression reflects increasing structural integration and predictive sophistication.

➢ **From Forecasting Accuracy to Decision Relevance**

Although predictive performance has improved significantly, forecasting accuracy alone does not guarantee effective decision-making. A central limitation of many epidemiological and predictive models lies in their weak coupling with operational decision layers such as healthcare capacity planning, logistics coordination, and intervention timing [50], [51]. Forecast outputs are frequently treated as exogenous inputs to separate optimization models, without explicit feedback integration.

Data dependency presents an additional challenge. Machine learning–based and hybrid models typically require large-scale, high-frequency datasets for calibration. During crisis peaks, epidemiological data are often delayed, underreported, or inconsistent particularly in low- and middle-income contexts reducing robustness and increasing sensitivity to noise [52].

Furthermore, most predictive frameworks are evaluated primarily using statistical accuracy metrics (e.g., RMSE, MAE), while limited attention is devoted to how forecast uncertainty propagates into downstream optimization decisions. Policies that appear optimal under deterministic point forecasts may perform poorly under stochastic realization [53]. This highlights the importance of embedding uncertainty quantification directly within forecasting–optimization interfaces.

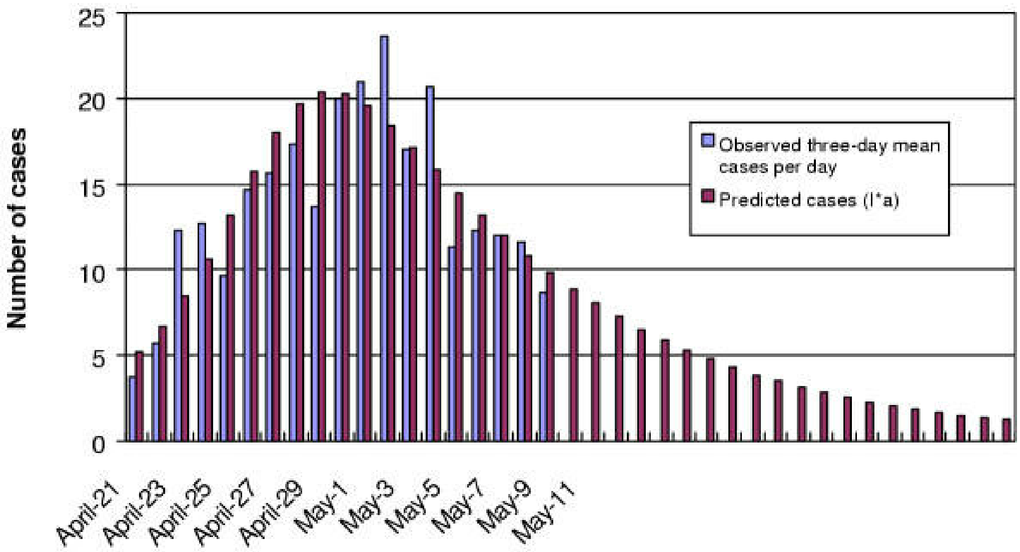

Figure 1.6.(a) Deterministic Compartmental Forecasting in year 2020 (SIR-Based Prediction)

Comparison between observed case counts and deterministic SIR-based model predictions over time. The Figure 1.3(a) illustrates the strengths and limitations of classical compartmental models in capturing epidemic peaks and decline phases.

While structurally interpretable and computationally efficient, such models rely on fixed transmission parameters and homogeneous mixing assumptions, which may limit predictive robustness under behavioral or policy shifts.

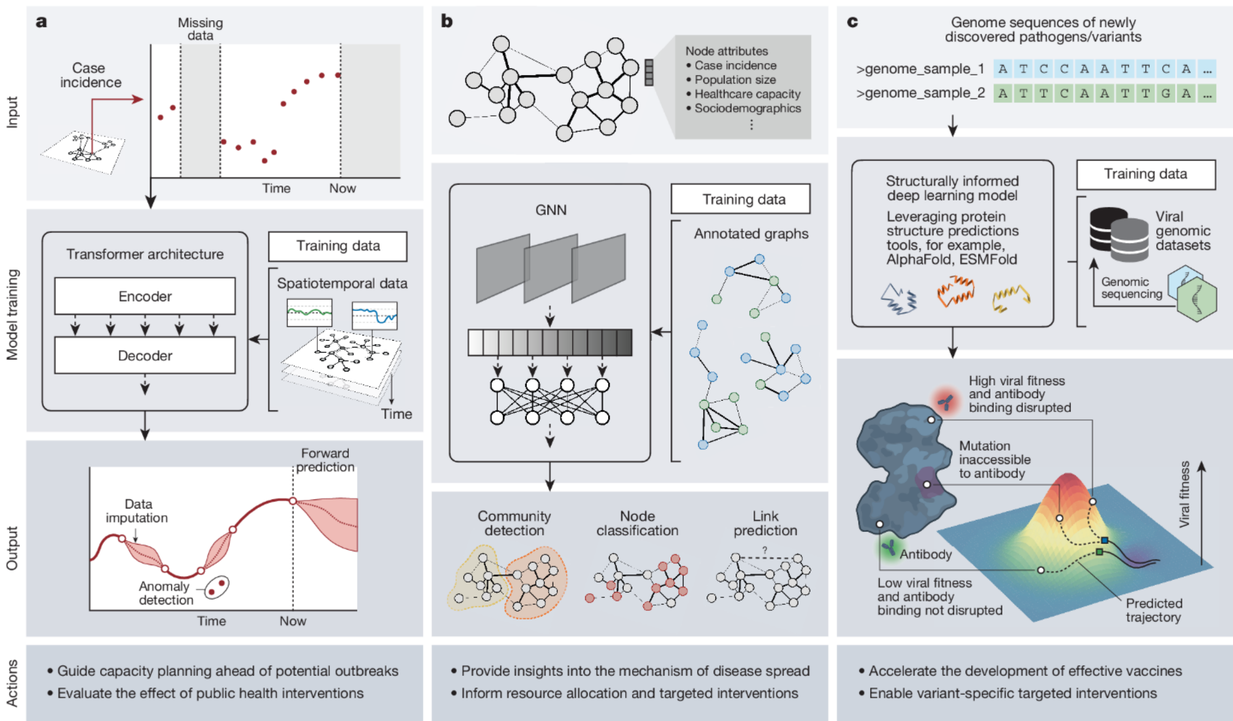

Figure 1.7.(b) AI-Enhanced Spatiotemporal and Network-Based Forecasting Architectures

Advanced predictive frameworks integrating transformer architectures, graph neural networks (GNNs), and structurally informed deep learning models. These approaches leverage spatiotemporal data, mobility networks, and genomic information to improve prediction accuracy, anomaly detection, and intervention targeting. Although highly adaptable and capable of modeling nonlinear dynamics, these architectures require extensive datasets and significant computational resources.

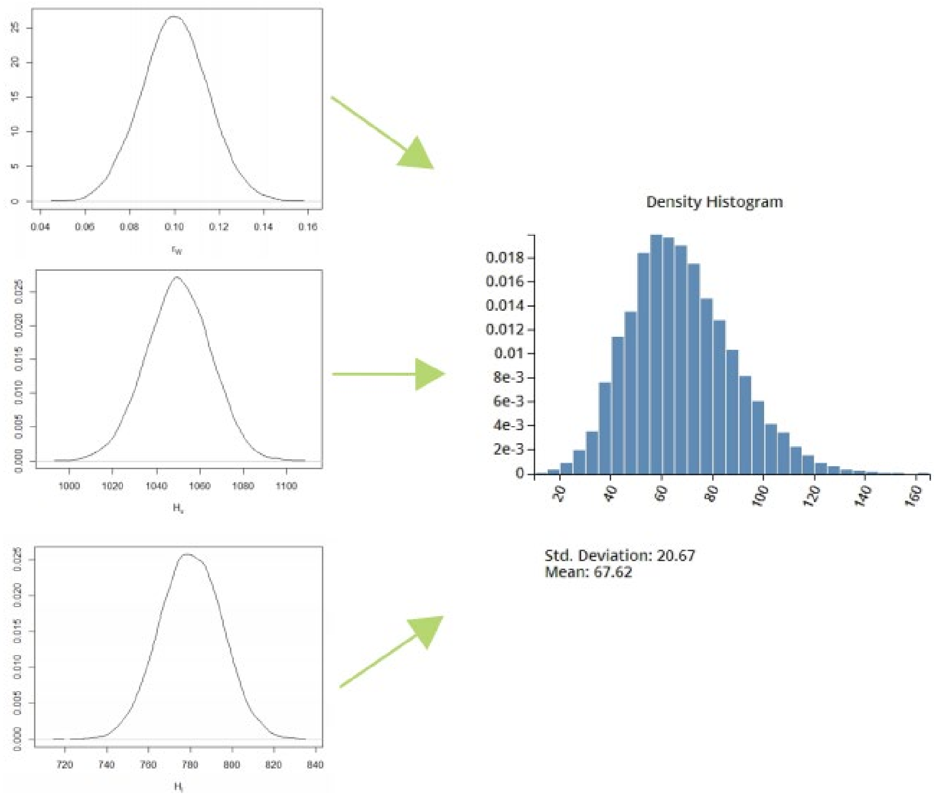

Figure 1.8.(c) Stochastic Parameter Estimation and Uncertainty Propagation in Forecasting
Figure 1.9. Comparison of Forecasting Model Categories

Probabilistic forecasting framework illustrating parameter uncertainty propagation into predictive distributions. Monte Carlo–based parameter sampling generates density histograms representing forecast variability (mean and standard deviation). This approach enhances robustness by quantifying uncertainty, enabling downstream integration with stochastic or robust optimization models.

Collectively, Figure 1.3(a–c) highlights the transition from deterministic curve fitting to data-driven architectures and, ultimately, to uncertainty-aware probabilistic forecasting suitable for integration with optimization-based decision support.

Classical SIR/SEIR models emphasize interpretability; hybrid models incorporate economic or policy dimensions; AI-driven frameworks enhance adaptability and predictive accuracy but increase data and computational demands.

➢ **Implications for Resource-Constrained Environments**

The transferability of advanced forecasting architectures to environments such as Kosovo remains constrained by limited surveillance coverage, fragmented data integration, and restricted computational capacity [18], [37]. Data-intensive AI-driven approaches may therefore exhibit reduced operational feasibility despite high theoretical accuracy.

In such contexts, reduced-order SEIR formulations augmented with Discrete Event Simulation (DES) and lightweight metaheuristics offer a pragmatic balance between interpretability, scalability, and robustness. Rather than maximizing predictive complexity, the focus shifts toward designing forecasting components that integrate seamlessly with optimization and simulation modules under uncertainty. This direction is consistent with adaptive and rolling-horizon decision logic, where forecasts are repeatedly updated as new evidence becomes available [25], [26].

In summary, the evolution of epidemiological and predictive models reflects a clear trajectory toward integration, hybridization, and AI augmentation. However, persistent challenges related to data dependency, uncertainty propagation, and weak operational coupling limit their standalone effectiveness. These observations reinforce the need to conceptualize forecasting not as an isolated analytical task, but as an embedded component of an integrated, adaptive decision-support framework an approach that underpins the methodological direction of this dissertation.

1.3.2 Healthcare Operations and Capacity Optimization

While epidemiological models estimate infection trajectories and demand surges, healthcare operations optimization translates projected demand into actionable resource allocation decisions. During pandemics, operational decision-making involves the allocation of intensive care unit (ICU) beds, ventilators, medical personnel, testing infrastructure, and inter-hospital transfers under severe uncertainty and time constraints [75], [76]. This operational perspective is supported by healthcare operations research and hospital capacity planning literature [75], [76].

➢ **ICU and Hospital Capacity Planning**

The rapid saturation of ICU capacity during COVID-19 prompted the development of optimization-based surge planning models. Multistage stochastic programming approaches have been applied to allocate ICU beds under uncertain demand scenarios [9], while scalable mixed-integer formulations have been proposed to determine surge expansion strategies and patient transfer coordination [37], [37].

These models typically aim to minimize unmet demand, mortality proxies, or system overload, subject to constraints on bed availability, staffing levels, equipment capacity, and regional coordination. In several studies, demand forecasts derived from SEIR models are integrated as input parameters to operational optimization layers [8].

Despite methodological sophistication, many ICU allocation frameworks assume centralized coordination and near-real-time data availability. In practice, hospital networks often operate with delayed reporting, decentralized governance structures, and heterogeneous infrastructure capacity, complicating real-time implementation [15], [17].

The flow diagram on Figure 1.4 illustrates the role of Centres of Government (CoGs) as central hubs connecting national leadership to health and economic ministries, which in turn interface with hospitals, supply networks, and the public through synchronized communication and data channels. Adapted from OECD governance reports [65].

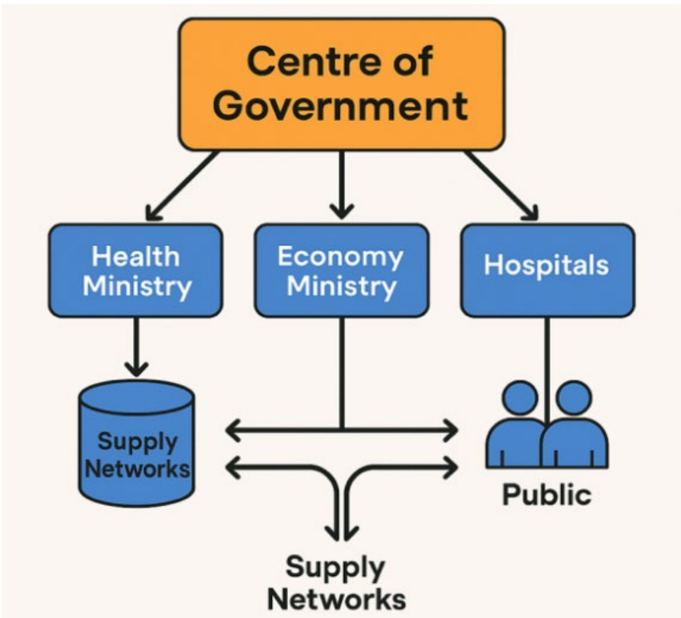

Figure 1.10. Centres of Government Coordination Flow

- **Dynamic Resource Allocation and Staff Scheduling**

Beyond bed allocation, pandemics require dynamic reassignment of healthcare personnel, reconfiguration of wards, and flexible scheduling under workforce shortages. Operations research models have addressed staff rostering and cross-training strategies to increase surge resilience [30], [30].

Discrete Event Simulation (DES) has been widely used to capture patient flows, waiting times, and bottlenecks in emergency departments and ICU units. When coupled with optimization modules, DES enables scenario-based evaluation of surge strategies and adaptive capacity planning [77]. Such hybrid simulation–optimization frameworks are particularly valuable in environments characterized by stochastic arrivals and resource variability.

However, DES-based systems can become computationally intensive when embedded within large-scale optimization loops, especially under rolling-horizon decision structures.

- **Stochastic and Robust Optimization Under Uncertainty**

Pandemic healthcare operations are inherently uncertain. Infection rates fluctuate, hospitalization probabilities vary, and supply chains experience disruption. To address this uncertainty, multistage stochastic programming models incorporate scenario trees representing possible demand realizations [9], [9].

Robust optimization techniques, grounded in the theoretical framework developed by Bertsimas and Tsitsiklis [78], provide an alternative approach by seeking solutions that remain feasible under worst-case parameter realizations. Robust surge planning models have been proposed to safeguard ICU capacity against extreme but plausible demand shocks [37].

While stochastic and robust approaches improve resilience, they introduce increased computational complexity. Large scenario trees and high-dimensional uncertainty sets may become intractable for real-time decision-making in resource-limited systems. Consequently, trade-offs arise between solution robustness and computational feasibility.

- **Network-Level Coordination and Inter-Hospital Transfers**

Healthcare systems function as interconnected networks rather than isolated facilities. Optimization models have therefore addressed patient transfer coordination across hospital networks to reduce localized overload [8]. Network flow formulations and decentralized coordination mechanisms allow redistribution of demand when regional capacity asymmetries arise.

However, empirical evidence suggests that transfer-based strategies are highly sensitive to travel time constraints, administrative fragmentation, and communication delays factors frequently underestimated in purely mathematical formulations.

- **Operational–Forecast Integration Challenges**

A recurring limitation across healthcare operations models concerns the interface between epidemiological forecasting and capacity optimization. Forecast outputs are often treated as deterministic inputs, neglecting uncertainty propagation and parameter instability. This decoupling may lead to allocation strategies that are optimal under expected demand but fragile under forecast error.

Furthermore, many operational optimization models prioritize efficiency objectives such as minimizing overflow or cost without explicitly incorporating equity considerations or contextual constraints. In resource-constrained environments, decisions must account for geographic disparities, informal care structures, and institutional limitations that standard formulations rarely capture.

- **Implications for Resource-Constrained Healthcare Systems**

The majority of healthcare operations optimization models were developed and validated in high-income settings with relatively stable data infrastructure and centralized coordination. Direct transfer to contexts such as Kosovo or the Western Balkans is therefore limited by:

- fragmented hospital networks,
- delayed or incomplete reporting systems,
- constrained fiscal space,
- limited computational infrastructure.

In such environments, scalable and parsimonious formulations are essential. Hybrid approaches combining reduced-order stochastic optimization, Discrete Event Simulation, and computationally efficient metaheuristics may provide a more viable balance between robustness and operational feasibility.

In summary, healthcare operations optimization has made substantial methodological progress in ICU allocation, surge capacity planning, and dynamic staff scheduling. Nevertheless, persistent challenges related to uncertainty propagation, computational scalability, and contextual adaptability limit real-world deployment. These constraints reinforce the need for integrated, adaptive, and equity-aware decision-support architectures capable of bridging forecasting and operational layers under uncertainty.

1.3.3 Supply Chain and Logistics Optimization

Pandemic response depends not only on hospital capacity but also on the resilience and coordination of healthcare supply chains. Shortages of personal protective equipment (PPE), ventilators, oxygen, pharmaceuticals, and vaccines during COVID-19 exposed structural fragilities in globally interconnected production and distribution networks. Unlike conventional supply chains optimized for efficiency and cost minimization, healthcare supply chains during pandemics must prioritize reliability, flexibility, and rapid surge response. Digital transformation and healthcare logistics reports further highlight visibility, traceability, and coordination as central supply-chain requirements during crisis conditions [79], [80].

- **Supply Chain Resilience and Viability**

Ivanov [16] introduced the concept of supply chain viability, emphasizing the ability of networks to survive and adapt under extreme disruption. Dynamic stockpiling strategies, multi-sourcing policies, and decentralized distribution structures were proposed to enhance resilience during demand surges [32].

Optimization models addressing pandemic logistics frequently incorporate network flow formulations, multi-echelon inventory control, and stochastic demand scenarios. These models seek to minimize shortage costs, transportation delays, or disruption impacts while satisfying service-level constraints. These formulations are grounded in supply chain theory, inventory resilience, and risk-management principles [81], [82], [83], [84].

However, pre-pandemic supply chain configurations were largely designed for lean efficiency, making them vulnerable to synchronized global shocks. The transition from cost-efficiency to resilience-oriented optimization represents a fundamental paradigm shift in healthcare logistics. The broader supply-chain risk literature similarly identifies redundancy, flexibility, and robust network design as essential resilience mechanisms [81], [84].

- **Medical Resource Allocation and Distribution**

During pandemic peaks, centralized allocation of scarce resources such as ventilators, oxygen tanks, antiviral drugs, and vaccines became critical. Optimization models have addressed:

- allocation of limited ventilators across regional hospitals,
- redistribution of stock between facilities,
- prioritization of high-risk populations,
- capacity-constrained distribution routing.

Multi-stage stochastic programming approaches have been proposed to account for uncertain demand and supply disruption [9], [9]. Robust optimization frameworks further safeguard allocation decisions against worst-case demand realizations [37].

Yet many allocation models assume reliable transportation infrastructure and real-time inventory tracking, assumptions frequently violated in LMIC contexts.

- **Vaccination and Testing Logistics**

Vaccination rollout represents a large-scale logistical optimization problem involving prioritization, cold-chain constraints, scheduling, and geographic coverage. Dynamic programming and network-based allocation models have been used to optimize vaccination strategies under limited supply [10], [10].

Network-driven prioritization approaches demonstrate that targeting high-centrality nodes in mobility networks can significantly reduce transmission rates [66]. These findings illustrate how epidemiological and logistical layers intersect within allocation decisions.

Testing logistics similarly requires balancing geographic accessibility, throughput capacity, and budget constraints. Equity-aware distribution models incorporate social vulnerability indices to improve fairness in access [13].

Despite methodological advances, large-scale vaccination optimization frameworks often rely on granular demographic and mobility data that may not be consistently available.

- **Forecast-Driven Logistics and ML Integration**

Recent research integrates demand forecasting with logistics optimization through machine learning–enhanced supply chain models. LSTM-based demand forecasting combined with stochastic inventory optimization has been proposed to improve responsiveness under volatile demand patterns [85].

Hybrid frameworks combining predictive analytics with metaheuristic optimization offer improved adaptability, particularly when classical deterministic models fail to capture nonlinear demand shifts. However, these approaches increase model complexity and computational requirements.

A persistent methodological challenge lies in synchronizing epidemiological forecasts, healthcare capacity constraints, and supply chain decisions within a unified optimization architecture.

- **Network Interdependencies and Cascading Disruptions**

Healthcare supply chains operate within broader economic production networks. Studies on production network propagation demonstrate how localized disruptions cascade across sectors [86]. Pandemic-induced workforce absenteeism, export restrictions, and transportation bottlenecks further amplify systemic vulnerability.

Most supply chain optimization models treat healthcare logistics as a partially isolated subsystem. However, the COVID-19 crisis demonstrated that global production, transportation infrastructure, and public policy measures jointly shape supply reliability. Incorporating these cross-sectoral dependencies remains analytically challenging.

- **Limitations in Resource-Constrained Environments**

In regions such as Kosovo and the Western Balkans, healthcare supply chains are characterized by:

- heavy reliance on imports,
- limited domestic production capacity,

- constrained cold-chain infrastructure,
- fragmented inventory reporting systems.

Data scarcity and logistical informality reduce the feasibility of highly granular, data-intensive optimization models. Consequently, simplified, scenario-based, and robust allocation frameworks may provide more practical support for decision-makers under uncertainty.

In summary, supply chain and logistics optimization has evolved toward resilience-oriented, stochastic, and hybrid modeling approaches. Nonetheless, persistent gaps remain in integrating logistics decisions with epidemiological forecasts, healthcare capacity constraints, and equity considerations within a unified and computationally feasible framework. Addressing these interdependencies constitutes a central challenge for integrated pandemic decision-support systems.

On Figure 1.5. is shown a schematic map of healthcare supply chains during the COVID-19 pandemic, highlighting critical nodes (manufacturers, transport hubs, hospitals) and key integration points for predictive analytics and network-based optimization. Adapted from network-based supply chain analyses and the co-authored survey [32], [87].

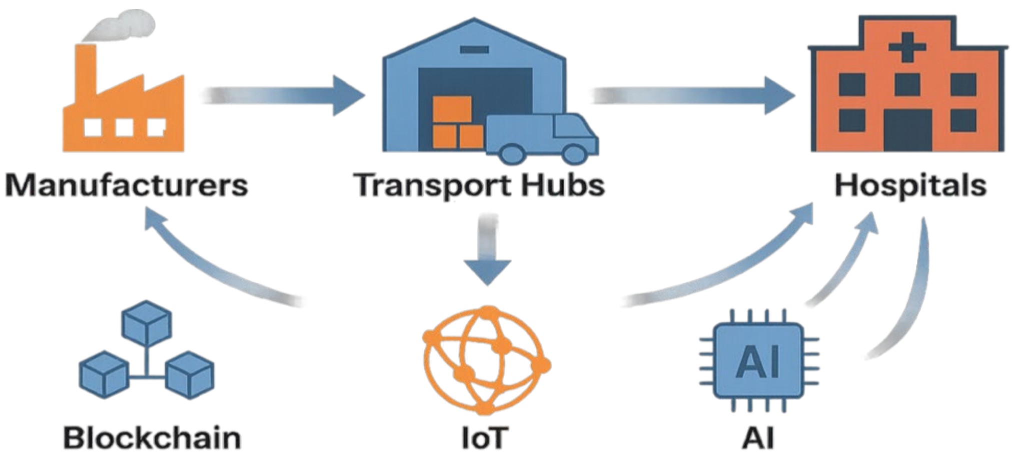

Figure 1.11. Healthcare Supply Chain Resilience Map

1.3.4 Multi-Objective and Economic–Health Optimization

Pandemic response inherently involves competing objectives. Policies designed to minimize infection rates and mortality may impose substantial economic costs, while rapid economic reopening can trigger renewed transmission waves and healthcare system overload. Consequently, pandemic management cannot be adequately represented as a single-objective optimization problem; rather, it requires formal multi-objective decision frameworks capable of balancing health protection, economic resilience, and social equity under uncertainty.

➢ **Health–Economy Trade-off Modeling**

Early epidemiological models largely optimized health outcomes in isolation. However, the economic consequences of lockdown measures, mobility restrictions, and sectoral shutdowns prompted the development of integrated economic–epidemiological frameworks. Hybrid models coupling SEIR dynamics with macroeconomic structures demonstrate how public health interventions simultaneously affect infection trajectories and economic output [12], [38], [12].

These models formalize trade-offs through welfare functions or composite objective functions combining mortality, output loss, and intervention costs. For example, optimal lockdown strategies are derived by minimizing weighted social welfare loss subject to epidemiological and healthcare capacity constraints [38], [39].

While analytically powerful, many such models rely on aggregated representative-agent assumptions and homogeneous sectoral responses, potentially overlooking heterogeneity in economic vulnerability and institutional capacity.

➢ **Formal Multi-Objective Optimization Approaches**

From an operations research perspective, pandemic decision-making can be structured as a multi-objective optimization problem, including the following objectives:

- Minimize mortality or ICU overload,
- Minimize economic contraction or fiscal expenditure,
- Minimize inequality or unmet demand,
- Maximize healthcare system resilience.

Multi-objective formulations generate Pareto-efficient solution sets rather than a single optimal solution. Decision-makers must then select preferred trade-off points based on policy priorities. This theoretical basis is well established in multicriteria and nonlinear multiobjective optimization [79], [72], [73].

Weighted-sum methods, ε-constraint approaches, and goal programming techniques have been applied to healthcare resource allocation and vaccination prioritization problems [36], [63], [63]. These approaches enable explicit incorporation of equity and capacity constraints within the objective structure. Evolutionary multi-objective methods also provide useful search mechanisms for Pareto-front exploration when exact scalarization is insufficient [88].

However, real-world implementation often reduces multi-objective problems to scalarized formulations using arbitrarily chosen weights. This may obscure the normative assumptions embedded within policy decisions.

➢ **Equity as an Objective vs. Constraint**

A critical methodological distinction concerns whether equity is treated as:

1. A secondary constraint, or
2. A co-equal optimization objective.

Fair resource allocation models incorporating Gini-based metrics, deprivation indices, or vulnerability weights illustrate the transition toward explicitly embedding fairness within objective functions [63], [64], [89]. Such formulations allow simultaneous evaluation of efficiency and equity outcomes.

Nevertheless, equity metrics often depend on detailed socio-economic data that may be unavailable or unreliable in resource-constrained environments. This creates tension between ethical rigor and operational feasibility.

- **Dynamic and Adaptive Multi-Objective Policies**

Pandemic dynamics evolve over time. Static optimization may yield policies optimal for a single period but suboptimal under changing infection rates, behavioral adaptation, and supply fluctuations.

Dynamic multi-objective frameworks incorporate rolling-horizon decision structures, allowing policy recalibration as new information becomes available. These approaches resemble model predictive control systems in engineering, where objectives are re-optimized periodically under updated state estimates.

However, multi-period stochastic multi-objective optimization significantly increases computational complexity. Scenario explosion, high-dimensional uncertainty, and nonlinear epidemiological interactions may render exact formulations intractable for real-time decision support.

- **Limitations of Existing Multi-Objective Pandemic Models**

Despite methodological advances, several limitations persist:

- Heavy reliance on deterministic forecasts,
- Limited propagation of epidemiological uncertainty into economic objectives,
- Simplified behavioral assumptions,
- Reduced applicability to decentralized or data-scarce healthcare systems.

Moreover, many economic–health optimization models are calibrated for high-income economies with stable fiscal buffers and advanced data infrastructure. Transferability to contexts such as Kosovo and the Western Balkans remains limited due to structural differences in healthcare capacity, labor markets, and governance mechanisms.

These limitations highlight the need for computationally efficient, uncertainty-aware multi-objective frameworks capable of integrating epidemiological forecasts, healthcare operations, and supply chain decisions within a unified architecture.

In summary, multi-objective and economic–health optimization represents a critical advancement in pandemic modeling by formalizing trade-offs across competing policy goals. However, persistent challenges related to uncertainty propagation, scalability, equity integration, and contextual adaptability limit practical deployment. These challenges motivate the development of hybrid, adaptive, and context-aware decision-support frameworks, as proposed in this dissertation.

1.3.5 Equity-Aware and Fairness-Based Optimization

Pandemics disproportionately affect vulnerable populations, exposing structural inequalities in healthcare access, socio-economic conditions, and institutional capacity. Traditional efficiency-oriented optimization models focused on minimizing total mortality or resource shortage may inadvertently exacerbate disparities if equity considerations are not explicitly incorporated.

Equity-aware optimization frameworks seek to balance efficiency and fairness through formal mathematical structures. Two principal approaches are observed in the literature:

1. **Equity as a Constraint**
   Resource allocation models impose proportionality or minimum coverage constraints across regions or demographic groups. This ensures baseline access but may limit overall system efficiency.
2. **Equity as an Objective Function**
   Multi-objective formulations integrate fairness metrics directly within the optimization objective. Measures such as Gini coefficients, deprivation indices, or weighted vulnerability scores are incorporated alongside efficiency criteria [63], [64], [89].

These formulations enable explicit trade-off analysis between aggregate system performance and distributive justice. In vaccination and ICU allocation models, embedding equity considerations has been shown to produce more socially acceptable and politically robust solutions without necessarily incurring significant efficiency losses.

However, equity modelling faces practical challenges:

- High data requirements for vulnerability assessment,
- Difficulty in quantifying social disadvantage,
- Increased computational complexity in multi-objective settings.

In low- and middle-income environments, incomplete socio-economic datasets and fragmented reporting systems may restrict the feasibility of highly granular fairness metrics. Consequently, parsimonious and robust equity-aware formulations capable of operating under data scarcity are essential for real-world deployment.

In summary, equity-aware optimization extends pandemic modelling beyond purely utilitarian efficiency objectives. Yet, its integration remains uneven, and operationally feasible fairness-based frameworks remain underdeveloped, particularly in resource-constrained contexts.

1.3.6 Simulation-Based and Adaptive Decision Frameworks

Pandemic response unfolds under rapidly evolving uncertainty. Static optimization models, even when multi-objective, may become suboptimal as infection dynamics, behavioral responses, and resource availability shift over time. This has motivated the development of simulation-based and adaptive decision-support frameworks.

- **Discrete Event Simulation (DES)**

Discrete Event Simulation is widely used to model patient flows, queue dynamics, ICU occupancy, and emergency department congestion. DES captures stochastic arrivals, service times, and bottleneck formation, providing a realistic representation of hospital operations [77]. When combined with optimization modules, DES enables scenario-based evaluation of surge strategies and resource reallocation policies. Simulation–optimization hybrids allow decision-makers to test alternative intervention strategies under varying demand realizations before implementation.

- **Rolling-Horizon and Adaptive Optimization**

Adaptive frameworks implement rolling-horizon decision processes, periodically updating forecasts and re-optimizing resource allocation as new data become available. This approach resembles model predictive control, where decisions are repeatedly recalibrated in response to state changes. This logic is closely related to approximate dynamic programming and model predictive control [25], [26].

Rolling-horizon optimization improves robustness under forecast error and demand volatility. However, repeated re-optimization increases computational burden and requires efficient solution algorithms to remain feasible in real time. However, the design of such adaptive systems requires explicit attention to computational tractability and the curse of dimensionality [25], [26].

- **Integration Challenges**

Despite their flexibility, simulation-based and adaptive frameworks face several limitations:

- High computational requirements in large-scale systems,
- Sensitivity to forecast uncertainty,
- Limited integration of equity and economic objectives within simulation loops.

Moreover, many simulation studies evaluate system performance retrospectively rather than embedding simulation directly into real-time decision pipelines.

In resource-constrained healthcare systems, lightweight simulation–optimization architectures combining reduced-order epidemiological models, DES modules, and computationally efficient metaheuristics may offer a practical balance between realism and feasibility.

In summary, simulation-based and adaptive decision frameworks enhance the temporal robustness of pandemic optimization. However, integration across forecasting, operations, logistics, and equity layers remains incomplete. Bridging these layers within a computationally scalable architecture constitutes a central challenge addressed in this dissertation.

1.4 Critical Review and Identified Gaps

This section synthesizes the evidence reviewed in Sections 1.1–1.3 and identifies the methodological, operational, and contextual gaps that limit the real-world effectiveness of current pandemic optimization and decision-support approaches. Across the literature, advances in forecasting, hospital operations, logistics, and policy analytics are substantial; however, their practical deployment remains constrained by fragmentation, data limitations, computational bottlenecks, insufficient equity integration, and weak real-time adaptivity constraints that are particularly acute in resource-constrained healthcare systems such as Kosovo.

1.4.1 Fragmented Modeling and Decision-Support Approaches

A dominant limitation is the persistence of siloed modeling. Most studies optimize isolated subsystems ICU capacity and transfers [9], forecasting and intervention evaluation [50], [51], logistics resilience [32], or vaccination prioritization [66] without an operationally coupled system architecture that supports coordinated decisions across domains. As a result, even strong domain-specific models frequently fail to support coherent policy choices when healthcare capacity, supply constraints, and behavioral compliance interact dynamically.

Although hybrid approaches exist [8], [87], integration is often conceptual rather than computationally operational: outputs from one model are treated as fixed inputs for another, without feedback loops that reflect how decisions reshape epidemic trajectories, logistics feasibility, and public compliance. This gap limits decision relevance precisely when cross-domain coordination is most critical.

To summarize the extent of cross-domain coverage across major modeling paradigms, Table 1.2 provides a structured comparison of representative approaches and highlights the persistent lack of integration across healthcare, logistics, economic, and behavioral dimensions.

Table 1.2. Comparison of Key Model Dimensions

Model Type	Healthcare Capacity	Supply Chain Integration	Economic Impacts	Behavioral Factors	Applicability in LMIC contexts (e.g., Kosovo)
MSSP / MILP	High (ICU-focused)	Minimal	No	No	Yes (simplified)
Hybrid LP + Stochastic	Moderate	Moderate	Partial	No	Limited (needs computing)
Graph-Theory Vaccination	None	No	No	No	Poor (needs full networks)
Dynamic Economic Models	Low	Low	High	No	Difficult (data scarcity)

1.4.2 Data Reliability, Timeliness, and Information-Flow Constraints

A second gap concerns data feasibility under crisis conditions. Many optimization and simulation frameworks implicitly assume timely, standardized, high-quality data streams for admissions, ICU occupancy, staffing, inventory, testing, and surveillance [90]. In practice, data are frequently delayed, incomplete, and inconsistently reported particularly in LMIC contexts undermining calibration, validation, and real-time updating [91]. Earlier operations research studies similarly emphasize that data quality and reporting structures strongly determine the usability of public-health decision-support systems [92], [93].

Crucially, data limitations do not merely reduce predictive accuracy; they propagate through decision chains. Forecast errors and delayed signals degrade downstream optimization outputs, increasing the risk of misallocation and localized shortages. This highlights the need for models explicitly designed for imperfect information, including mechanisms for uncertainty-aware forecasting, robust decision rules, and graceful degradation under missing data.

From a decision-analytic perspective, this motivates optimization frameworks that are explicitly designed for imperfect information, including robust and distributionally conservative decision rules that remain feasible under bounded data errors and delayed signals [20], [21], [21].

Figure 6 illustrates how reporting delays cascade across forecasting, optimization, and allocation layers, amplifying uncertainty and increasing the risk of healthcare system overload in resource-constrained environments. The figure highlights how even modest delays at the data collection stage can trigger compounding errors across downstream decision-support processes.

Figure 1.6 illustrates how reporting delays cascade across forecasting, optimization, and allocation layers, amplifying uncertainty and increasing the risk of healthcare system overload in resource-constrained environments. The figure highlights how even modest delays at the data collection stage can trigger compounding errors across downstream decision-support processes

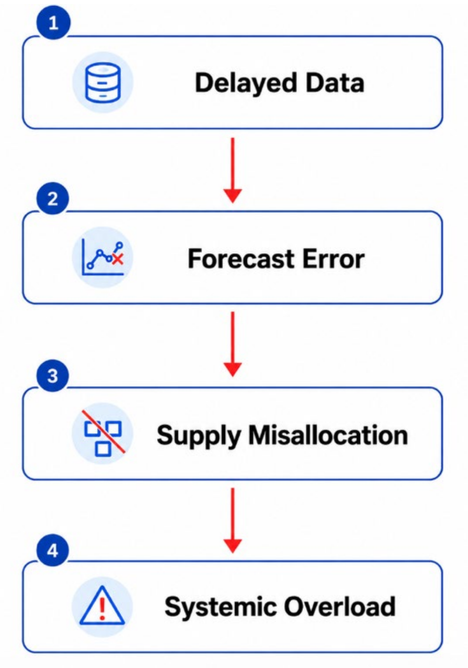

Figure 1.12. Impact of Data Delays on Decision Chains

Collectively, these findings indicate that data reliability and timeliness constraints represent a fundamental bottleneck for the practical deployment of advanced optimization and simulation models. Without explicit mechanisms to handle delayed, incomplete, and uncertain data, even methodologically sophisticated decision-support systems risk producing recommendations that are misaligned with real-world operational conditions, particularly in resource-constrained healthcare systems.

1.4.3 Scalability and Computational Complexity

Even when data are available, real-time feasibility is constrained by computational load. Multistage stochastic programming and large MILP formulations can become intractable as scenario trees grow, while AI/metaheuristic forecasting and simulation–optimization loops may require long runtimes that are incompatible with time-critical crisis operations [73], [36]. Robust formulations can reduce the need for large scenario trees by replacing extensive stochastic enumeration with tractable uncertainty sets, offering a practical compromise between resilience and computational feasibility in time-critical settings [20], [21], [21].

This creates a fundamental trade-off between model richness and operational usability. For real deployment especially in settings with limited analytical infrastructure scalable formulations, reduced-order representations, and computationally efficient solution methods become primary design requirements rather than secondary implementation details.

Table 1.3 compares typical runtime and infrastructure requirements across representative modeling approaches, highlighting the fundamental trade-off between modeling sophistication and operational feasibility in pandemic decision-support contexts.

Table 1.3. Computational Requirements of Models

Model Approach	Typical Runtime (National Dataset)	Specialized Software Needs	Hardware Requirements	Feasible for Kosovo?
IBMO-LSSVM	10–12 hours	MATLAB, custom scripts	High-performance servers (GPU)	No (needs simplification)
Metaheuristic Hybrid	6–8 hours	GAMS, VBA, Excel	Multi-core workstations	Partial (simplified only)
MSSP	2–3 hours	CPLEX / Gurobi	Moderate servers	Yes
DES + NLP (Proposed)	1–2 hours (scaled)	MATLAB, Simulink	Standard workstations	Yes (optimized approach)

These findings indicate that computational scalability constitutes a critical bottleneck for the practical deployment of advanced optimization and simulation models. Approaches that require long execution times, specialized software, or high-performance hardware may deliver strong analytical results, but they fail to meet the temporal and operational constraints of real-world crisis management. Consequently, scalability and computational efficiency emerge as essential design criteria for decision-support frameworks intended for deployment in resource-constrained healthcare systems.

1.4.4 Equity and Contextual Adaptation Limitations

While fairness-aware allocation models represent an important methodological advance [13], equity is still often treated as an auxiliary feature rather than a core design dimension. Formally, the efficiency–equity tension is a multicriteria optimization problem, where policy acceptability depends on transparent trade-offs rather than purely utilitarian objectives [22]. In practice, ad hoc scalarization weights may embed normative assumptions that remain implicit; multicriteria optimization provides a clearer structure for eliciting and documenting such trade-offs [22]. Broader health-equity literature also shows that distributive impacts must be treated as central design constraints rather than afterthoughts [94], [64], [95], [58], [23], [24].

Many equity approaches depend on detailed vulnerability and network data, which may be unavailable or unreliable in fragmented and resource-limited contexts. Similarly, technology-intensive logistics solutions (e.g., IoT/blockchain-based visibility) may not be scalable or cost-feasible in Kosovo and comparable systems [32], [6], [80].

As a result, there remains a gap between equity principles and deployable equity-aware decision support. A practical pathway is the development of parsimonious fairness mechanisms robust constraints, minimum-service guarantees, and scenario-based equity stress tests that remain implementable under data scarcity while preventing systematic disadvantage.
Table 1.4 evaluates the adaptability of major modeling strategies to LMIC contexts and highlights the persistent mismatch between infrastructure assumptions and practical deployment conditions in settings such as Kosovo.

Table 1.4. Adaptability of Models to Kosovo and LMICs

Strategy Type	Tested in LMICs?	Data Requirements	Infrastructure Needs	Practicality for Kosovo
Graph-Theory Vaccination	No	Full networks	High	Low
Supply Chain Optimization	Limited	Blockchain/IoT	Advanced	Low
MSSP with Simplified Equity Constraints	Yes	Aggregate data	Moderate	High

Collectively, these observations indicate that insufficient integration of equity and contextual constraints represents a major barrier to the practical deployment of pandemic response models in low- and middle-income settings. Models that fail to account for social vulnerability, institutional limitations, and infrastructure heterogeneity risk reinforcing existing inequalities rather than mitigating them. Addressing equity as a core design dimension rather than a secondary constraint is therefore essential for developing decision-support systems that are both effective and socially just in resource-constrained healthcare environments.

1.4.5 Insufficient Real-Time and Adaptive Decision Support

Despite progress in DES and ABM, relatively few studies achieve a fully integrated, real-time adaptive pipeline that links forecasting, operational optimization, and simulation-based evaluation in a rolling-horizon manner [8], [9]. Rolling-horizon decision architectures are closely aligned with the theory of Model Predictive Control (MPC), where policies are repeatedly re-optimized under updated state estimates and operational constraints [25]. For large-scale systems with uncertainty and dimensionality, approximate dynamic programming offers computationally viable alternatives for near-real-time decision making [26]. Forecasting models are commonly evaluated on statistical accuracy rather than on their contribution to downstream decision quality under uncertainty [53]. Meanwhile, operational optimization often assumes deterministic demand or weakly represents uncertainty propagation.

This leaves a major gap in end-to-end adaptive decision support: systems that can continuously ingest imperfect data, update forecasts, optimize allocations, and stress-test policies through simulation while remaining computationally feasible and context-aware.

Figure 1.7 presents a conceptual hybrid DES–MSSP–NLP framework that integrates epidemiological forecasts, hospital capacity data, and healthcare supply-chain information through a centralized optimization core, enabling real-time and context-aware decision support.

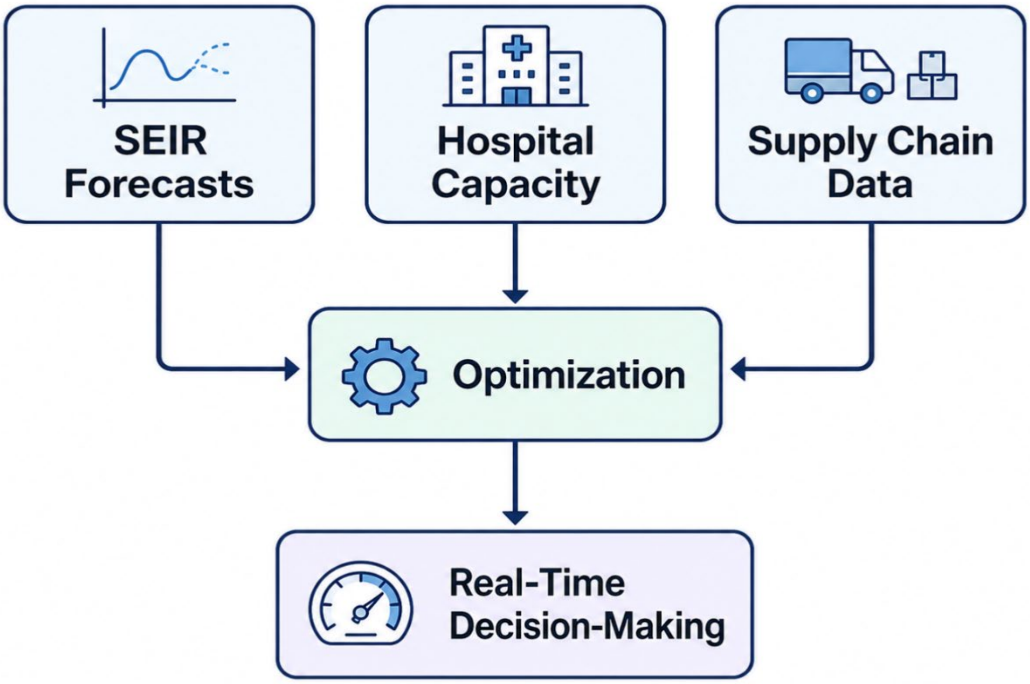

Figure 1.13. Proposed DES–MSSP–NLP Hybrid Framework

These limitations demonstrate that the lack of fully integrated, adaptive decision-support systems remains a critical gap in the existing literature. Without seamless coupling between forecasting, optimization, simulation, and operational healthcare data, decision-support tools struggle to deliver timely and actionable recommendations under rapidly evolving crisis conditions. This gap directly motivates the hybrid DES–MSSP–NLP framework proposed in this dissertation.

1.4.6 Synthesis of Identified Research Gaps

The critical review identifies five interrelated gaps that jointly limit real-world pandemic decision support:

1) **Fragmentation**: weak cross-domain coupling across forecasting, operations, logistics, and behavior.
2) **Data feasibility**: delayed, incomplete, and non-standardized data streams undermine real-time deployment.
3) **Computational scalability**: high complexity restricts timely re-optimization under uncertainty.
4) **Equity and contextual validity**: fairness mechanisms and infrastructure assumptions often do not transfer to LMIC settings.
5) **Limited adaptivity**: insufficient end-to-end rolling-horizon decision architectures linking forecasting, optimization, and simulation.

These gaps collectively motivate the dissertation's integrated hybrid approach, designed to balance analytical rigor with real-time feasibility, robustness to uncertainty, and contextual applicability in resource-constrained healthcare systems.

Methodologically, these gaps point toward hybrid architectures that combine uncertainty-aware optimization (robust or stochastic), multicriteria trade-off analysis, and rolling-horizon control principles to remain deployable under crisis constraints [20], [21], [21], [22], [23], [24], [25], [26].

1.5 Conclusions, Research Goal and Research Tasks

1.5.1 Conclusions

The literature reviewed in this chapter demonstrates substantial progress in pandemic response modeling, including advances in epidemiological forecasting, healthcare operations optimization, supply chain coordination, equity-aware allocation, and simulation-based evaluation. Nevertheless, the COVID-19 experience revealed that domain-specific methodological sophistication does not automatically translate into operational decision value under crisis conditions.

A recurring finding is the mismatch between model assumptions and real-world constraints. Many advanced frameworks require timely, standardized data streams and significant computational resources, while real decision environments particularly in resource-constrained health systems are characterized by delayed reporting, fragmented information flows, limited analytical capacity, and institutional constraints. These realities reduce the feasibility of complex models exactly when rapid adaptation is most needed.

The chapter further shows that pandemic management is inherently multi-objective and cross-domain: decisions about capacity expansion, logistics prioritization, and interventions must be evaluated jointly with equity considerations and behavioral responses. As a result, effective decision support requires integrated architectures that link forecasting, optimization, and simulation in a rolling-horizon manner, while explicitly handling uncertainty and maintaining computational feasibility.

These conclusions establish the motivation for the dissertation's research goal and tasks, which focus on developing an integrated, scalable, and adaptive hybrid decision-support framework tailored to resource-constrained healthcare systems, with particular emphasis on Kosovo.

1.5.2 Research Goal

Building upon the critical synthesis presented in Sections 1.1–1.5.1, the overarching goal of this dissertation is formulated as follows:

The goal of this dissertation is to develop an integrated, scalable, and uncertainty-aware hybrid decision-support framework that unifies epidemiological forecasting, healthcare resource optimization, logistics coordination, and equity considerations within a computationally efficient architecture suitable for real-time deployment in resource-constrained healthcare systems.

The proposed framework aims to bridge the gap between methodological sophistication and operational feasibility by:

- explicitly incorporating uncertainty in epidemiological and operational inputs;
- embedding equity and contextual constraints as structural model components rather than auxiliary criteria;
- ensuring computational scalability compatible with limited analytical infrastructure;
- enabling rolling-horizon, adaptive decision-making under imperfect and delayed data conditions.

The framework is specifically designed and validated in the context of Kosovo and comparable low- and middle-income healthcare systems, while maintaining generalizability to broader crisis-management environments.

1.5.3 Research Tasks

To achieve the stated research goal, the dissertation addresses the following interrelated research tasks:

Task 1 – Conceptual and System-Level Modeling

Develop a unified system architecture that formally links:

- extended SEIR-type epidemiological models,
- healthcare capacity and ICU resource allocation,
- supply chain and logistics constraints,
- behavioral and equity-aware decision variables.

This task establishes the structural integration necessary to overcome the fragmentation identified in Chapter 1.

Task 2 – Development of an Uncertainty-Aware Optimization Core

Design and implement a hybrid optimization module that:

- integrates multistage stochastic programming (MSSP) and nonlinear equity-aware formulations (NLP);
- incorporates robust or scenario-based uncertainty handling;
- balances efficiency and fairness objectives within a multi-criteria structure.

This task directly addresses the identified gaps related to uncertainty propagation and equity integration.

Task 3 – Integration with Discrete Event Simulation (DES)

Embed the optimization core within a Discrete Event Simulation (DES) environment to:

- model patient flows, hospital operations, and logistical processes;
- evaluate policy decisions under dynamic and uncertain conditions;
- enable scenario-based stress testing of healthcare system resilience.

This task ensures operational realism and supports adaptive, rolling-horizon decision-making.

Task 4 – Computational Scalability and Model Simplification

Develop computationally efficient solution strategies, including:

- reduced-order epidemiological representations,
- lightweight metaheuristics where exact optimization is infeasible,
- scalable scenario reduction techniques.

The objective is to ensure feasibility in data-scarce and infrastructure-limited environments such as Kosovo.

Task 5 – Contextual Validation and Policy Evaluation

Calibrate and validate the integrated framework using data representative of Kosovo and comparable healthcare systems, and:

- evaluate trade-offs between healthcare capacity, logistics feasibility, and equity outcomes;
- assess robustness under delayed and incomplete data scenarios;
- derive actionable policy recommendations for real-time crisis management.

The completion of these research tasks provides a structured pathway from theoretical gap identification to the development of a deployable, integrated hybrid decision-support system. The following chapter therefore presents the methodological foundations and mathematical formulations underpinning the proposed framework.

Chapter 2. Proposed Models and Methodology: Proposed Integrated Hybrid Decision-Support Framework

2.1 Introduction to the Proposed Research Framework

This chapter presents the methodological foundation of the dissertation and introduces an Integrated Hybrid Decision-Support Framework (IHDSF) for pandemic healthcare system optimization. The framework connects epidemiological forecasting, healthcare resource modelling, robust multi-objective optimization, simulation-based validation, and adaptive feedback control in a unified decision-support structure. Its purpose is to transform uncertain epidemiological information into operational decisions related to hospital capacity, ICU utilization, patient transfers, staff deployment, surge-capacity activation, and emergency resource allocation [96], [92], [20], [97], [98].

Unlike static planning approaches, the proposed methodology treats pandemic healthcare management as a dynamic decision process. Infection rates, hospitalization probabilities, ICU admission rates, recovery rates, staff availability, transfer feasibility, and reporting delays can change over short time intervals. Consequently, decisions that appear feasible under one information state may become infeasible or inefficient when new observations are incorporated. The IHDSF is therefore designed as a closed-loop framework in which forecasts, constraints, allocation policies, simulation outputs, and feedback are updated repeatedly [99], [3], [3], [18].

The framework is built around five methodological components. First, an epidemiological forecasting component estimates future infection, hospitalization, and ICU demand. Second, a healthcare resource modelling component represents effective capacity, including ward beds, ICU beds, ventilators, staff, and transfer possibilities. Third, a robust multi-objective optimization component generates allocation and coordination decisions under uncertainty. Fourth, a discrete-event simulation layer evaluates whether candidate decisions remain operationally feasible under stochastic patient-flow conditions. Fifth, an adaptive feedback layer enables rolling-horizon recalibration and re-optimization [100], [97], [98].

At a high level, the proposed framework can be summarized by the following closed-loop logic:

$$Data_t \rightarrow Forecast_t \rightarrow Capacity_t \rightarrow Optimize_t \rightarrow Simulate_t \rightarrow Feedback_t \rightarrow Data_{t+1} \tag{2.1}$$

The general state of the pandemic healthcare system at decision epoch t is represented by a state vector x_t, while operational actions are represented by a decision vector u_t. The system evolves according to

$$x_{t+1} = F(x_t, u_t, \theta_t), \quad t = 0,1, \dots, T_{max} - 1 \tag{2.2}$$

where θ_t denotes uncertain epidemiological and operational parameters. These parameters may include transmission rates, hospitalization probabilities, ICU admission probabilities, recovery rates, length-of-stay parameters, staff absenteeism, reporting delays, and operational disruptions. To account for uncertainty, the parameter vector is assumed to belong to an uncertainty set U_tThis allows the optimization model developed later in the chapter to search for decisions that remain feasible across a range of plausible scenarios, rather than only under a single deterministic forecast.

The detailed formulation of each component is developed in the subsequent sections. Section 2.2 formally defines the decision problem and scope. Section 2.3 specifies the modelling assumptions and constraints. Section 2.4 presents the architecture of the IHDSF. Sections 2.5, 2.6, 2.7, and 2.8 develop the epidemiological, resource, optimization, and simulation components. Section 2.9 defines the adaptive decision-support workflow, and Section 2.10 summarizes the chapter.

Figure 2.1. illustrates the interaction among the five components of the framework. Epidemiological forecasts produce demand projections, which are translated into capacity requirements by the resource modelling layer. The optimization layer generates allocation and coordination decisions under uncertainty. The simulation layer evaluates these decisions under stochastic operational conditions. Finally, the adaptive feedback layer incorporates newly observed data and simulation outputs into the next planning cycle.

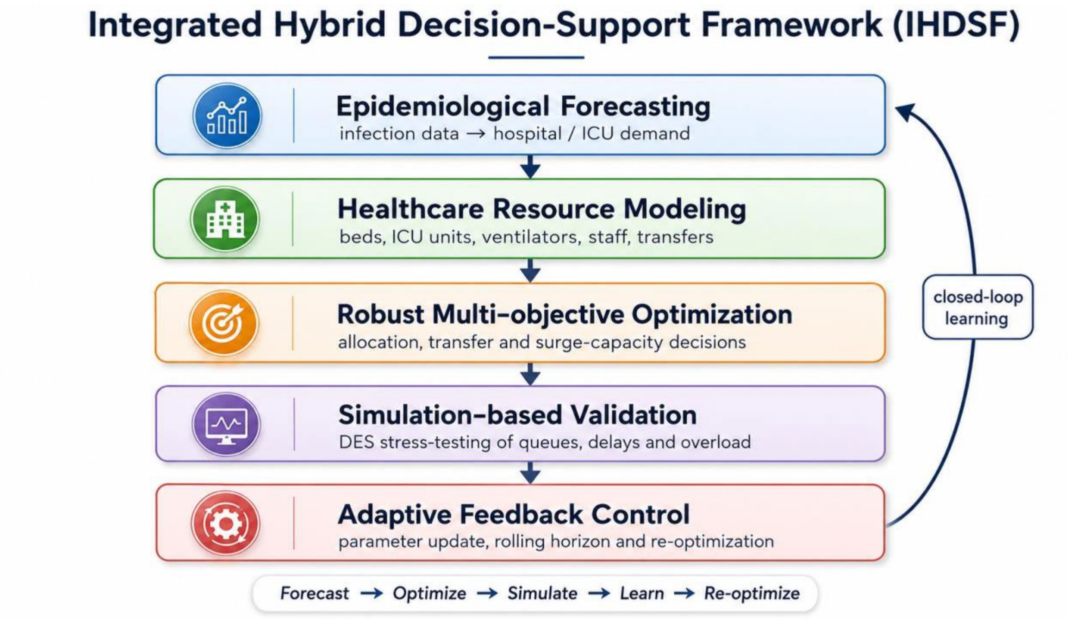

Figure 2.1. Conceptual architecture of the Integrated Hybrid Decision-Support Framework for pandemic healthcare system optimization

The contribution of the proposed framework is not the isolated use of SEIR modelling, optimization, or simulation. Its methodological value lies in connecting prediction, prescription, evaluation, and adaptation into a single decision-support process suitable for pandemic healthcare management under uncertainty. The robustness claim is interpreted conditionally: the framework improves feasibility for parameter realizations contained in the specified uncertainty set, but it does not eliminate all possible risks outside the modelled bounds [96], [96], [8].

2.2 Formal Problem Definition and Research Scope

A formal problem definition is necessary because pandemic healthcare response contains many interacting decisions, but not all of them are decision variables of the proposed model. The dissertation distinguishes between endogenous decisions, which the model can recommend, and exogenous conditions, which are treated as data, parameters, or scenarios. This distinction prevents the framework from becoming either too narrow to support regional coordination or too broad to remain computationally and institutionally realistic [101], [102], [65].

The central decision problem can be expressed in operational terms as follows: given uncertain trajectories of infection, hospitalization, and ICU demand, and given limited healthcare capacity distributed across a network of facilities, determine a sequence of allocation, transfer, staffing, and surge-capacity decisions that reduces avoidable shortage and overload while respecting clinical, logistical, and institutional constraints. This definition keeps the focus on healthcare system coordination rather than on population-level policy optimization [3], [103], [18].

The model is particularly relevant for regional systems in which hospitals differ in available beds, ICU units, staff, ventilators, and transfer accessibility. During pandemic surges, some facilities may face severe overload while others retain unused effective capacity. Without coordination, localized overload can trigger deterioration in waiting times, treatment quality, staff fatigue, and mortality-related risk. The IHDSF therefore treats the regional healthcare system as a network rather than as a set of isolated hospitals [92], [8], [20].

The purpose of this section is to define the pandemic healthcare optimization problem addressed in the dissertation. The proposed framework focuses on regional healthcare coordination under uncertain pandemic demand. It supports tactical and strategic decisions related to hospital and ICU capacity, patient transfers, staff and equipment redistribution, and surge-capacity activation. The model does not replace clinical judgment or national public health policymaking; rather, it supports system-level coordination when demand threatens to exceed available healthcare capacity.

2.2.1 System Definition and Decision Scope

The decision scope is tactical rather than purely strategic or bedside clinical. Tactical coordination includes decisions that can be revised over days or weeks, such as redistributing patients, activating emergency capacity, redeploying staff, or prioritizing scarce transfer resources. Strategic decisions, such as building new hospitals or redesigning national health policy, occur on longer time scales and are outside the optimization scope. Bedside clinical decisions, such as diagnosis, triage judgment, and treatment plans for individual patients, remain under the authority of clinical professionals [65], [104].

This scope also reflects the type of data that can realistically be available during a crisis. Facility-level occupancy, ICU load, admissions, staff availability, and transfer capacity are often reported at regular intervals, while individual-level clinical trajectories may be incomplete or protected by privacy requirements. The framework therefore uses aggregate operational variables in the optimization layer and reserves patient-level stochastic behaviour for the simulation layer, where detailed patient-flow uncertainty can be tested without overloading the prescriptive model [3], [103], [18].

The pandemic healthcare system is modelled over a discrete planning horizon

$$T = \{0,1,\dots,T_{\max}\} \tag{2.3}$$

where each period t represents a decision interval, such as one day or one week. Let I denote the set of hospitals, healthcare facilities, or modelled regions, and let $C = \{W, ICU\}$ denote the set of care levels, where W represents general ward care and ICU represents intensive care. If inter-facility transfers are allowed, feasible transfer routes are represented by $A \subseteq I \times I$.

The decision scope of the dissertation is the coordination of hospital resources during pandemic pressure. Included decisions are admissions or allocation of patients to facilities, inter-facility transfers, activation of surge capacity, redistribution of staff or critical equipment, and balancing of utilization across the healthcare network. Broader policies such as vaccination strategy, school closure, lockdown timing, travel restrictions, and macroeconomic interventions are treated as exogenous scenario inputs. These factors may influence epidemiological parameters, especially the transmission rate, but they are not optimized directly in Chapter 2.

2.2.2 Sets, States, Decisions and Uncertainty

The state vector combines epidemiological and operational information because neither dimension is sufficient by itself. Epidemiological information explains future demand pressure, while operational information explains whether the healthcare network can absorb that demand. For example, a forecasted increase in infections is operationally relevant only when it is translated into projected ward occupancy, ICU demand, staff requirements, and transfer needs.

The decision vector is also multi-dimensional because pandemic healthcare coordination rarely consists of one action. A policy may simultaneously admit patients locally, transfer selected patients to less-loaded facilities, activate temporary ICU capacity, redeploy staff, or request emergency equipment. Representing these actions in one vector allows the framework to evaluate trade-offs and interactions among them instead of optimizing each action separately.

Uncertainty is interpreted broadly. It includes not only epidemiological uncertainty, such as transmission and hospitalization rates, but also operational uncertainty, such as staff absenteeism, transfer delays, capacity reporting errors, or sudden changes in ICU availability. This broader uncertainty representation is essential because a decision that is robust to infection forecasts alone may still fail if staffing or transfer capacity changes unexpectedly.

At each decision period, the system is described by a state vector, a decision vector, and an uncertainty vector. A compact representation of the state is

$$x_t = (e_t, p_t, c_t, s_t) \tag{2.4}$$

where e_t represents epidemiological states, p_t represents patient census and flow states, c_t represents available capacity states, and s_t represents staffing and operational readiness states.

In concrete terms, x_t may include susceptible, exposed, infectious, hospitalized, ICU, recovered, and deceased populations; ward and ICU occupancy; effective capacity; staff availability; transfer feasibility; and emergency surge status.

Operational decisions are represented by

$$u_t = (a_t, y_t, z_t, r_t) \tag{2.5}$$

where a_t denotes admission or allocation decisions, $y_t$denotes inter-facility transfer decisions, z_t denotes surge-capacity activation decisions, and r_t denotes redistribution of critical resources such as staff, ventilators, or other equipment. This notation is intentionally general at this stage; the detailed decision variables are specified in the optimization component in Section 2.7.

The system evolves dynamically according to

$$x_{t+1} = F(x_t, u_t, \theta_t) \tag{2.6}$$

where F(x_t, u_t, θ_t) captures the coupled evolution of epidemiological states and healthcare system operations. The uncertainty vector θ_t may include transmission rates, hospitalization probabilities, ICU admission probabilities, recovery rates, mortality-related parameters, length-of-stay distributions, staff absenteeism, and reporting delays. Uncertainty is represented through

$$\theta_t \in \mathcal{U}_t \tag{2.7}$$

where $\mathcal{U}_t$ is the uncertainty set for period t. The uncertainty representation is necessary because pandemic decisions must remain feasible under multiple plausible epidemiological and operational scenarios, not only under one expected trajectory.

2.2.3 Demand, Capacity and Shortage Representation

The shortage variable is central to the methodology because it converts overload into an explicit mathematical object. In severe pandemic conditions, forcing all demand to be served may make the optimization model infeasible, while ignoring excess demand would hide the clinical consequences of overload. By introducing , the model remains solvable under extreme demand while still penalizing unmet need strongly in the objective function [92], [92], [105].$q_{i,c,t}$

The distinction between demand and treated census is equally important. Demand represents the number of patients requiring care, whereas treated census represents the number actually accommodated by the healthcare system. When demand exceeds effective capacity, the difference is not simply a technical imbalance; it represents delayed admission, denied admission, treatment under crisis standards, or other forms of unmet clinical need. For this reason, shortage is treated as a clinical and operational penalty rather than as a neutral slack variable.

Effective capacity is defined at the care-level because ward and ICU services are not interchangeable. A facility may have empty ward beds while lacking ICU beds, ventilators, or trained ICU staff. Conversely, ICU resources may be available but irrelevant for patients requiring only ward care. The separation of W and ICU prevents the model from overestimating the ability of the system to absorb demand through simple aggregation.

Healthcare demand for facility i, care level c, and period t is denoted by $D_{i,c,t}^{\text{dem}}$. Effective available capacity is denoted by $C_{i,c,t}^{\text{eff}}$. Effective capacity is not simply a nominal bed count; it incorporates the availability of beds, ICU resources, ventilators, staff, and other operational bottlenecks. Unmet demand is represented by the non-negative shortage variable $q_{i,c,t}$:

$$q_{i,c,t} \geq D_{i,c,t}^{\text{dem}} - C_{i,c,t}^{\text{eff}}, \quad q_{i,c,t} \geq 0 \tag{2.8}$$

This definition penalizes only positive shortage. It avoids the misleading interpretation that surplus capacity in one facility can automatically compensate for shortage in another facility unless explicit transfer and coordination constraints allow such redistribution. Shortage is therefore a clinically meaningful and operationally interpretable variable, not a signed imbalance measure.

When transfers are included, local shortage is evaluated after inbound and outbound transfers are considered. If $y_{ji,c,t}$ denotes the transfer of patients from facility i to facility j at care level c, the transfer-adjusted shortage expression is

$$q_{i,c,t} \geq D_{i,c,t}^{\text{dem}} + \sum_{j:(j,i)\in A} y_{ji,c,t} - \sum_{j:(i,j)\in A} y_{ij,c,t} - C_{i,c,t}^{\text{eff}} \tag{2.9}$$

This expression ensures that receiving additional patients increases local pressure, while transferring patients out reduces local demand pressure subject to feasibility constraints. Patient transfers are therefore not treated as free capacity; they are constrained by feasible routes, receiving capacity, emergency transport capacity, and clinical appropriateness.

2.2.4 Robust Multi-Objective Problem Formulation

The objective function is multi-objective because pandemic healthcare response involves competing priorities. Minimizing unmet demand is clinically urgent, but a policy that minimizes shortage by transferring many patients may exceed EMS capacity or create unnecessary logistical risk. A policy that minimizes transfers may be easier to implement but may leave localized hospitals overloaded. The multi-objective structure makes these trade-offs explicit and allows decision-makers to select priority weights or acceptable thresholds [22], [22].

The robust formulation is used because an expected-value solution can be fragile. If the model optimizes only for the average forecast, it may perform well under nominal conditions but fail when transmission, hospitalization, or staff absenteeism are worse than expected. The max-over-uncertainty term therefore protects decisions against adverse but plausible parameter realizations within the defined uncertainty set [20], [102], [102].

The formulation should not be interpreted as a claim of universal optimality. The solution is optimal relative to the specified model, data, constraints, uncertainty set, objective weights, and planning horizon. This careful interpretation is important for doctoral-level methodology because it avoids overstating what mathematical optimization can guarantee in a real public health emergency.

The high-level decision problem is formulated as a robust multi-objective optimization problem. The model seeks a feasible decision policy that minimizes unmet demand, mortality-related risk, logistical burden, and inequity in system load under uncertainty:

$$u^{\star} = \underset{u \in \mathcal{X}}{\operatorname{argmin}} \max_{\theta \in \mathcal{U}} J(u, \theta) \tag{2.10}$$

The system-level penalty can be decomposed into period-level penalties:

$$J(u, \theta) = \sum_{t \in T} J_t\,(u_t, \theta_t) \tag{2.11}$$

$$J_t(u_t, \theta_t) = w_1 Q_t(\theta_t) + w_2 M_t(\theta_t) + w_3 L_t(u_t) + w_4 E_t(u_t) \tag{2.12}$$

$$Q_t(\theta_t) = \sum_{i \in I} \sum_{c \in C} q_{i,c,t}\,(\theta_t) \tag{2.13}$$

In this formulation, Q_t denotes total unmet clinical demand, $M_t$denotes mortality-related system risk, $L_t$denotes logistical or transfer cost, and E_t denotes equity or load-balancing penalty. The weights w_1, w_2, w_3, and w_4 represent decision-maker priorities and determine the relative importance of the clinical, logistical, and equity-related objectives. In pandemic response, the shortage and mortality-related terms are usually assigned high priority because they are directly connected to system safety and clinical outcomes.

The feasible decision space X includes all policies satisfying physical capacity limits, staff feasibility, transfer-route feasibility, EMS capacity, surge-capacity bounds, non-negativity, integrality, and institutional constraints. The detailed forms of these constraints are developed in Sections 2.6 and 2.7.

2.2.5 Decision Levels and Stakeholder Context

The primary decision level addressed by the IHDSF is regional coordination. Hospital-level data provide local capacity and occupancy information, while regional authorities coordinate transfers, surge resources, and load balancing across the network. National-level policies define the broader scenario environment by affecting transmission dynamics, emergency resource availability, crisis standards, and institutional rules. These national policies are treated as exogenous scenario parameters rather than direct decision variables in the optimization model [65], [104].

The framework is intended to support, not replace, human decision-making. Hospitals and healthcare providers deliver care and report operational capacity. Public health authorities provide surveillance and epidemiological information. Regional coordinators use model recommendations to organize patient transfers and allocate scarce resources. Government authorities define the policy environment. Patients and communities generate healthcare demand and are affected by system performance. The final decision remains subject to clinical judgment, administrative feasibility, ethical standards, and legal authority.

Table 2.1. Core notation and problem elements

Element	Notation	Interpretation
Planning horizon	T = {0,1,…,T_max}	Discrete set of decision periods
Facilities or regions	I	Hospitals, healthcare facilities, or modelled regions
Care levels	C = {W, ICU}	Ward and intensive-care levels
Transfer arcs	A	Feasible inter-facility routes
System state	x_t	Epidemiological and operational status
Decision vector	u_t	Allocation, transfer, surge, and resource actions
Uncertainty vector	$\theta_t \in U_t$	Plausible epidemiological and operational realizations
Demand	$D_{i,c,t}^{\mathrm{dem}}$	Projected need for care
Effective capacity	$C_{i,c,t}^{eff}$	Operationally available care capacity
Shortage	$q_{i,c,t}$	Unmet demand when demand exceeds capacity
Objective penalty	J(u,θ)	Weighted clinical, logistical, and equity penalty
Feasible space	X	Physical, staffing, transfer, and institutional limits

2.3 Modeling Assumptions and Constraints

The proposed IHDSF is based on explicit modelling assumptions that define the scope and interpretation of the epidemiological, resource, optimization, simulation, and adaptive-control components. A pandemic healthcare system contains biological, behavioural, logistical, institutional, and political mechanisms that cannot all be represented at the same level of detail within a tractable decision-support model. The assumptions in this section specify the conditions under which the proposed models are valid and operationally meaningful [92], [103], [3].

The assumptions are grouped into three categories: epidemiological assumptions, data availability and quality assumptions, and computational and institutional constraints. The purpose of this section is not to develop the full SEIR model, the full resource model, or the full optimization model. Those components are developed in Sections 2.5, 2.6, and 2.7. Section 2.3 instead clarifies the modelling conditions that support the later formulations.

2.3.1 Epidemiological Assumptions

The piecewise-stability assumption is particularly important. Pandemic parameters can change because of new variants, behavioural adaptation, testing changes, vaccination effects, treatment improvements, or policy interventions. However, within a short planning interval it is reasonable to approximate parameters as stable enough to generate operational forecasts. Rolling-horizon recalibration then allows the model to update these parameters when new data become available [101], [11], [11].

The model also assumes that epidemiological projections are used as decision-support inputs, not as final policy conclusions. Infection and hospitalization forecasts provide demand signals, but resource allocation decisions require additional information about capacity, staff, transfers, and institutional feasibility. This separation of roles prevents the epidemiological model from being interpreted as a complete pandemic response model.

The use of aggregate compartments implies that individual heterogeneity is represented indirectly. Age, comorbidity, vaccination status, and clinical severity can influence hospitalization or ICU probabilities through parameter values or scenario definitions. If the empirical data support more detailed stratification, the framework can be extended to include patient classes; however, the core Chapter 2 formulation remains at the level required for system-wide coordination [101], [11], [106].

The epidemiological component follows an extended SEIR-type logic in which infection dynamics are linked to hospital and ICU demand. The model assumes that population-level compartmental dynamics provide a sufficiently informative approximation for system-level planning over short decision intervals. This aggregation is appropriate because the dissertation focuses on regional coordination and resource allocation, not individual patient diagnosis.

The main epidemiological assumptions are:

- A1. The population can be represented through aggregate epidemiological and clinical compartments.
- A2. Epidemiological parameters are treated as piecewise stable within each short planning interval and may be updated between intervals.
- A3. Vaccination, behavioural changes, and non-pharmaceutical interventions are not optimized directly in Chapter 2; their effects are reflected through changes in epidemiological parameters, especially the transmission rate.
- A4. A proportion of infectious individuals requires hospitalization, and a proportion of hospitalized patients requires intensive care.
- A5. Epidemiological projections serve as inputs to the healthcare resource allocation model.
- A6. The framework supports system-level decision-making and does not replace clinical diagnosis or individual patient-level judgment.

A compact compartmental accounting structure for region r may be written as

$$N_r(t) = S_r(t) + E_r(t) + I_r(t) + H_r(t) + \mathrm{ICU}_r(t) + R_r(t) + D_r(t) \tag{2.14}$$

where $S_r(t)$, $E_r(t)$, $I_r(t)$, $H_r(t)$, $\mathrm{ICU}_r(t)$, $R_r(t)$, and $D_r(t)$ denote susceptible, exposed, infectious, hospitalized, intensive-care, recovered, and deceased populations, respectively. The detailed epidemiological equations are developed in Section 2.5. ICU notation is used consistently as ICU(t). The mortality compartment is denoted by D(t) to avoid unnecessarily cumbersome notation.

Healthcare demand can be interpreted either as occupancy-based demand or admission-flow demand. In the occupancy-based interpretation,

$$D^{\mathrm{dem}}_{r,W,t} = H_{r,t}, \quad D^{\mathrm{dem}}_{r,\mathrm{ICU},t} = \mathrm{ICU}_{r,t} \tag{2.15}$$

This expression establishes the link between the epidemiological forecasting layer and the healthcare resource model. In the arrival-flow interpretation, demand is derived from transitions into ward and ICU compartments, which is useful for simulation of admissions, queues, waiting times, and length of stay.

2.3.2 Data Availability and Quality Assumptions

Data quality constraints are not merely technical inconveniences; they directly affect decision quality. Under-reporting of infections can delay recognition of future hospital pressure, while delayed ICU occupancy data can cause the optimization layer to underestimate immediate overload. Similarly, inconsistent definitions of available beds can create artificial capacity if beds are reported without the staff or equipment required to use them [103], [3].

The framework therefore assumes that data are harmonized before being used for forecasting, optimization, or simulation. Harmonization includes aligning time stamps, reconciling hospital-level and regional-level records, correcting obvious inconsistencies, imputing missing values where justified, and documenting the uncertainty introduced by data limitations. These steps are developed empirically in the data and parameterization chapter, but their methodological role is established here.

The observation-error representation also justifies the use of scenario and uncertainty analysis. When the current state is measured imperfectly, a single deterministic forecast may give a false sense of precision. By passing uncertainty forward into the robust optimization and simulation layers, the framework recognizes that data limitations are part of the decision environment rather than a separate preprocessing issue [20], [20], [102].

The framework assumes that epidemiological and operational data are available at a frequency compatible with the planning horizon. However, observed pandemic data are treated as imperfect measurements of the true system state. Infection counts may be affected by testing capacity, hospitalization records may lag behind actual admissions, capacity data may be revised retrospectively, and staff availability may change rapidly because of sickness, quarantine, fatigue, or redeployment.

A general observation model can be written as

$$\tilde{Y}_{k,t} = Y_{k,t-\ell_k} + \varepsilon_{k,t} \tag{2.16}$$

where $\tilde{Y}_{k,t}$ is the observed value of data element k at time t, $Y_{k,t-\ell_k}$ is the true delayed value, ℓ_k is the reporting delay, and $\varepsilon_{k,t}$ is an observation or revision error term. This representation clarifies why the framework cannot rely only on a single deterministic estimate of the current state. Instead, it uses uncertainty-aware forecasting, robust optimization, and simulation-based validation.

The operational data required by the IHDSF include epidemiological indicators, hospital admissions, ward and ICU occupancy, bed availability, ventilators, staff availability, EMS capacity, and transfer feasibility. These data support forecasting, capacity estimation, optimization, and simulation-based validation.

Table 2.2. Data categories and quality assumptions

Data category	Examples of variables	Main quality issue	Role in IHDSF
Epidemiological data	New infections, positivity rate, recoveries, deaths	Reporting delay and under-detection	Forecasting and parameter updating
Hospital demand data	Admissions, ward occupancy, ICU occupancy	Delayed or inconsistent reporting	Demand calibration and validation
Capacity data	Beds, ICU units, ventilators, oxygen capacity	Rapid changes during surge response	Effective capacity estimation
Workforce data	Available staff, absenteeism, redeployment	Sudden staff depletion	Staff-adjusted capacity estimation
Logistics data	Distances, EMS capacity, transfer times	Route and vehicle availability	Transfer feasibility and cost modelling

2.3.3 Computational and Institutional Constraints

Computational constraints require a balance between realism and tractability. Including every individual patient, staff member, ambulance, and clinical pathway inside the optimization model would make the problem difficult to solve repeatedly in a rolling-horizon setting. The proposed structure therefore uses aggregate decision variables in the optimization model and stochastic patient-flow detail in the DES layer. This division of labour preserves operational relevance while keeping the prescriptive model manageable [92], [105], [107].

Institutional constraints are equally important. Even when an allocation appears mathematically efficient, it may be infeasible if regional authorities cannot authorize transfers, if receiving hospitals cannot accept specific patient classes, or if crisis standards restrict certain decisions. These constraints are represented through feasible decision spaces, transfer arcs, capacity bounds, and human-in-the-loop review. The model therefore supports institutional decision-making rather than bypassing it [65], [104], [65].

The assumptions in this section define the boundary conditions of the IHDSF. Within these boundaries, the framework can generate transparent, reproducible, and uncertainty-aware recommendations. Outside these boundaries, for example when data are unavailable or institutional authority is absent, the model outputs should be interpreted as exploratory scenario analyses rather than direct operational recommendations.

The IHDSF is designed for tactical and strategic coordination rather than individual-level clinical decision-making. To preserve computational tractability, the optimization layer aggregates patients by facility, care level, and time period. This aggregation allows the model to include multiple hospitals, care levels, transfer arcs, time periods, and uncertainty scenarios without requiring an individual-level clinical simulator inside the optimization problem.

Core feasibility constraints include capacity, shortage, transfer, emergency transport, surge-capacity, non-negativity, and institutional-authority constraints. Patient census must satisfy

$$P_{i,c,t} \leq C_{i,c,t}^{\text{eff}} + q_{i,c,t}, \quad q_{i,c,t} \geq 0 \tag{2.17}$$

where $P_{i,c,t}$ is treated patient census, $C_{i,c,t}^{\text{eff}}$ is effective capacity, and $q_{i,c,t}$ is recorded unmet demand. Transfers must be feasible within the hospital network and limited by available transport capacity:

$$y_{ij,c,t} = 0 \text{ for } (i,j) \notin A, \quad \sum_i \sum_j \sum_c y_{ij,c,t} \leq \text{EMS}_t \tag{2.18}$$

Surge capacity is bounded by operational limits:

$$0 \leq z_{i,c,t} \leq S_{i,c,t}^{\text{surge,max}} \tag{2.19}$$

These constraints ensure that the model does not assume unlimited beds, staff, equipment, transfers, or institutional authority. They also define the operational realism of the optimization and simulation components developed later in the chapter. Institutional constraints are particularly important because cross-facility allocation and patient transfers usually require authorization by regional healthcare authorities and must remain compatible with clinical and ethical standards.

2.4 Architecture of the Integrated Hybrid Decision-Support Framework

The Integrated Hybrid Decision-Support Framework is designed as a closed-loop architecture that connects epidemiological forecasting, healthcare resource modelling, robust optimization, simulation-based evaluation, and adaptive feedback. Its purpose is to transform observed epidemiological and operational data into allocation, transfer, and surge-capacity decisions that remain feasible under uncertainty [103], [65], [92].

The framework consists of six interconnected modules: a Data and Monitoring Module, an Epidemiological Forecasting Module, a Healthcare Resource Modelling Module, an Optimization and Decision Module, a Simulation and Evaluation Module, and an Adaptive Feedback Module. These modules are not independent blocks; the output of each module becomes an input to subsequent modules, and feedback from simulation and observed data updates earlier assumptions, constraints, and parameters.

2.4.1 Overall System Architecture

The architecture separates the analytical tasks while preserving their interaction. Forecasting produces demand estimates, but those estimates are not used directly as decisions. Resource modelling converts demand into capacity pressure. Optimization transforms pressure and constraints into candidate actions. Simulation tests those actions under stochastic variability. Feedback updates the system for the next decision cycle. This modular structure makes the framework transparent and easier to validate [96], [96], [20].

The modular design also supports implementation flexibility. The forecasting module may use different estimation methods depending on available data; the optimization module may use exact or heuristic methods depending on problem size; and the simulation layer may use empirical or parametric distributions depending on hospital records. What remains fixed is the information flow among modules and the requirement that candidate policies be evaluated before implementation.

The architecture is therefore both methodological and practical. It provides a scientific structure for the dissertation and a potential operational structure for healthcare authorities. Each module can be calibrated, tested, and improved without requiring the entire framework to be redesigned. The Data and Monitoring Module collects epidemiological, hospital, staff, and logistics data. The Epidemiological Forecasting Module converts infection dynamics into projected ward and ICU demand. The Healthcare Resource Modelling Module estimates effective capacity and identifies bottlenecks. The Optimization and Decision Module generates robust multi-objective allocation and coordination decisions. The Simulation and Evaluation Module stress-tests candidate decisions under stochastic patient-flow conditions. Finally, the Adaptive Feedback Module updates parameters, constraints, and decisions in a rolling-horizon manner.

The modular flow can be expressed as

$$\mathrm{Input}_t = \left(\mathrm{Data}_t^{\mathrm{epi}}, \mathrm{Data}_t^{\mathrm{hosp}}, \mathrm{Data}_t^{\mathrm{staff}}, \mathrm{Data}_t^{\mathrm{log}}\right) \tag{2.20}$$

$$u_t^\star = G_{\mathrm{opt}}\left(\widehat{D}_{t:t+H_p}, C_t^{\mathrm{eff}}, \mathcal{U}_t, \mathcal{X}_t\right) \tag{2.21}$$

$$\hat{J}(u_t^\star) = G_{\mathrm{sim}}(u_t^\star, \Omega) \tag{2.22}$$

where $\widehat{D}_{t:t+H_p}$ is the demand forecast over the planning horizon, C_t^{eff} is the effective capacity estimate, $\mathcal{U}_t$ is the uncertainty set, $\mathcal{X}_t$ is the feasible decision space, Ω represents stochastic simulation conditions, and $\hat{J}u_t^\star$ is the estimated performance of the candidate policy.

Figure 2.2 illustrates the overall architecture and information flow of the Integrated Hybrid Decision-Support Framework.

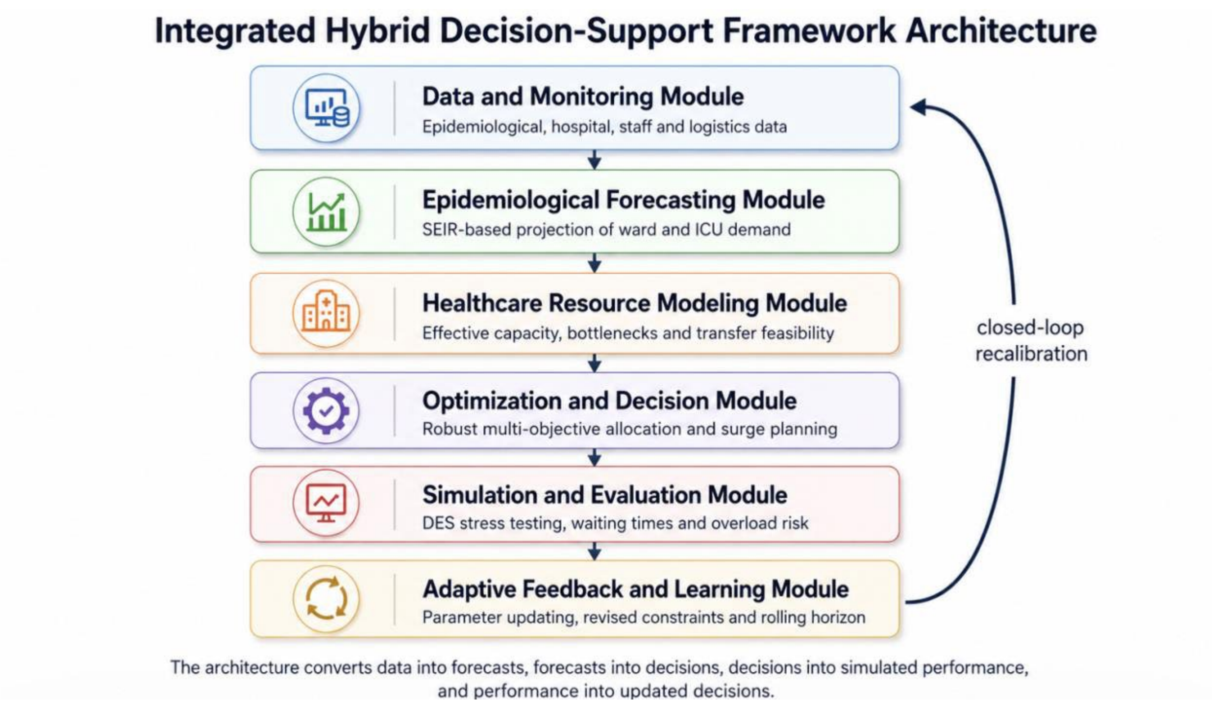

Figure 2.2. Overall architecture of the Integrated Hybrid Decision-Support Framework

Table 2.3 summarizes the principal modules of the Integrated Hybrid Decision-Support Framework and their roles in the decision-support process.

Table 2.3. Main modules of the Integrated Hybrid Decision-Support Framework

Module	Main input	Analytical role	Main output
Data and Monitoring	Epidemiological, hospital, staff, logistics data	State estimation and harmonization	Updated input vector
Epidemiological Forecasting	Surveillance data and parameters	SEIR-H-ICU-D forecasting	Ward and ICU demand forecast
Resource Modelling	Beds, ICU units, ventilators, staff, routes	Capacity and bottleneck modelling	Effective capacity and feasibility limits
Optimization and Decision	Demand, capacity, uncertainty sets	Robust multi-objective optimization	Allocation, transfer, surge decisions
Simulation and Evaluation	Candidate policy and stochastic inputs	DES stress testing	Waiting, overload, utilization indicators
Adaptive Feedback	Observed outcomes and simulation outputs	Rolling-horizon updating	Revised parameters and policies

2.4.2 Information Flow and Feedback Mechanisms

Feedback is not limited to correcting forecast errors. It also identifies structural weaknesses in proposed decisions. For example, simulation may reveal that an optimized policy creates excessive ICU queues, that a transfer plan overloads EMS capacity at peak times, or that a facility with nominal spare capacity cannot absorb patients because staff availability has declined. These feedback signals can be converted into additional constraints, penalty adjustments, or warning indicators [100], [96].

The closed-loop structure also supports learning over time. As the framework observes new outcomes, it can refine parameter estimates, uncertainty bounds, and operational thresholds. This does not imply automatic machine learning in every empirical implementation; rather, it means that the model is designed to incorporate updated evidence systematically. In this sense, the IHDSF is adaptive even when the updating method is a simple moving-window recalibration or scenario revision [20].

A key advantage of this architecture is interpretability. Each recommendation can be traced back to the demand forecast, capacity estimate, objective weights, constraints, uncertainty set, and simulation evaluation that produced it. Such traceability is important for accountability in public health decision-making, where recommendations must be explained to administrators, clinicians, and policy authorities.

The effectiveness of the framework depends on the ability of its modules to exchange information and adapt as system conditions change. The framework is sequential because data flow from forecasting to optimization and simulation, but it is also iterative because feedback from later modules updates earlier parameters and constraints. The conceptual information flow is:

$$\mathrm{Data}_t \rightarrow \mathrm{Forecast}_t \rightarrow \mathrm{Capacity}_t \rightarrow \mathrm{Optimize}_t \rightarrow \mathrm{Simulate}_t \rightarrow \mathrm{Feedback}_t \rightarrow \mathrm{Data}_{t+1} \tag{2.23}$$

Feedback operates at three levels. Epidemiological feedback updates transmission, hospitalization, ICU progression, recovery, and mortality parameters. Operational feedback updates occupancy, staff availability, transfer feasibility, and surge capacity. Simulation feedback identifies waiting-time violations, queue growth, unsafe overload probability, or unrealistic transfer volumes, and converts these indicators into revised constraints, penalties, or warnings for the next planning cycle.

An abstract representation of the recommended policy at time t is

$$\pi_t = \Phi\left(\mathrm{Data}_t, \widehat{D}_{t:t+H_p}, C_t^{\mathrm{eff}}, \mathcal{U}_t, \hat{J}_t\right) \tag{2.24}$$

where Φ denotes the integrated decision mapping that combines data, forecasts, capacity estimates, uncertainty bounds, and simulation-evaluated performance. This representation emphasizes that the IHDSF is not a collection of isolated models; it is a coordinated decision-support process linking prediction, prescription, evaluation, and adaptation.

2.5 Epidemiological Modeling Component

The epidemiological component is the predictive engine of the IHDSF. Its purpose is to translate disease transmission dynamics into operational demand signals for ward and ICU care. In the context of pandemic healthcare management, forecasting infections alone is not sufficient. Healthcare authorities require forecasts of hospital admissions, ICU demand, patient severity, recovery, and mortality-related pressure in order to plan resource allocation and capacity coordination [11], [108], [109].

The epidemiological model used in this dissertation must satisfy three requirements. It must be epidemiologically interpretable, computationally tractable for repeated recalibration, and directly connectable to healthcare resource requirements. For this reason, the dissertation adopts a SEIR-based compartmental structure extended with hospitalization, ICU, recovery, and mortality compartments.

2.5.1 Selection of a SEIR-Based Modeling Approach

The SEIR structure is also compatible with the time scale of healthcare planning. Hospital and ICU demand usually respond to infection dynamics with a delay, and this delay can be represented through exposed, infectious, hospitalized, and ICU compartments. The model therefore supports short- and medium-term forecasts that are useful for decisions such as surge activation, staff redeployment, and inter-facility transfer planning [108], [109].

Another advantage of the compartmental approach is parameter interpretability. Transmission, incubation, recovery, hospitalization, ICU progression, and mortality parameters have epidemiological meaning and can be estimated from surveillance or hospital data. This interpretability is valuable for decision-makers because it allows scenario assumptions to be discussed in clinically and operationally meaningful terms [102], [106].

The choice of SEIR does not imply that agent-based or network models are unsuitable in general. Such models may be appropriate when individual mobility, contact networks, or behavioural heterogeneity are the main research focus. In this dissertation, however, the main objective is to connect demand forecasting with resource allocation and simulation validation. For that purpose, an extended SEIR-H-ICU-D model provides an appropriate methodological balance.

Mathematical epidemiology commonly employs either compartmental models or agent-based models. Agent-based models can represent heterogeneous contact networks, individual behaviour, and mobility patterns, but they are computationally intensive and less straightforward to integrate repeatedly with optimization solvers in rolling-horizon settings. The IHDSF requires the forecasting layer to be recalibrated frequently and passed directly to the optimization engine. A compartmental model therefore offers a more appropriate balance between epidemiological realism and computational feasibility.

The SEIR foundation is suitable because it distinguishes between exposed individuals and infectious individuals. This distinction is important for pandemic healthcare management because hospital demand is delayed relative to infection exposure. The extended model links earlier transmission dynamics to later hospital and ICU demand, thereby making the forecast operationally useful for capacity planning and resource allocation.

2.5.2 Extended SEIR-H-ICU-D Structure

The additional H and ICU compartments are essential because they represent the direct operational burden placed on hospitals. A standard SEIR model may forecast infections accurately but still fail to inform resource decisions if it does not estimate how many patients require ward care or intensive care. The H compartment therefore represents patients requiring general hospital care, while the ICU compartment represents patients requiring critical care resources such as specialized staff, monitoring, ventilators, and higher care intensity.

The deceased compartment $D(t)$ is included as an absorbing state. It is not used to replace clinical outcome analysis, but it allows mortality-related system risk to be represented consistently in the forecasting and optimization layers. In the optimization model, mortality-related risk can be linked to shortage, ICU overload, or prolonged waiting times, depending on the empirical assumptions and available data.

The transition parameters ρ and κ are especially important for resource planning. The parameter ρ links infections to ward demand, while κ links hospitalized patients to ICU demand. These parameters may vary across waves, variants, age structures, or treatment practices. For this reason, they should be estimated or scenario-calibrated in the empirical chapter rather than treated as universal constants.

The model can be implemented in continuous time or discretized for operational decision periods. The optimization and simulation layers use discrete planning intervals, so epidemiological outputs are evaluated at decision epochs. This permits the model to connect continuous disease dynamics with daily or weekly operational decisions without requiring the optimization layer to solve continuous-time differential equations directly.

The standard SEIR model is insufficient for healthcare operations research because it does not distinguish mild infections from hospitalized or intensive-care patients. To bridge this gap, the dissertation adopts an extended SEIR-H-ICU-D structure. The regional population is divided into susceptible, exposed, infectious, hospitalized, ICU, recovered, and deceased compartments. ICU(t) is used consistently for the intensive-care compartment, and D(t) is used for the deceased compartment [11], [101], [11].

Let N denote the total population in the modelled region. The continuous-time dynamics of the extended model are represented by the following system of nonlinear ordinary differential equations:

$$\frac{dS}{dt} = -\beta(t)\frac{S(t)I(t)}{N} \tag{2.25}$$

$$\frac{dE}{dt} = \beta(t)\frac{S(t)I(t)}{N} - \sigma E(t) \tag{2.26}$$

$$\frac{dI}{dt} = \sigma E(t) - (\gamma_I + \rho)I(t) \tag{2.27}$$

$$\frac{dH}{dt} = \rho I(t) - (\gamma_H + \kappa)H(t) \tag{2.28}$$

$$\frac{d\mathrm{ICU}}{dt} = \kappa H(t) - (\gamma_{\mathrm{ICU}} + \mu)\mathrm{ICU}(t) \tag{2.29}$$

$$\frac{dR}{dt} = \gamma_I I(t) + \gamma_H H(t) + \gamma_{\mathrm{ICU}}\mathrm{ICU}(t) \tag{2.30}$$

$$\frac{dD}{dt} = \mu\,\mathrm{ICU}(t) \tag{2.31}$$

The model satisfies the population accounting identity

$$N = S(t) + E(t) + I(t) + H(t) + \mathrm{ICU}(t) + R(t) + D(t) \tag{2.32}$$

In this system, β(t) is the time-varying transmission rate, σ is the incubation rate, γ_I, γ_H, and γ_{ICU} are recovery rates, ρ is the transition rate from infection to hospitalization, κ is the transition rate from hospitalization to ICU, and μ is the ICU mortality rate. The hospitalization and ICU compartments create the direct bridge between epidemiological dynamics and healthcare operations.

For operational decision-making, model output is translated into demand signals. In an occupancy-based interpretation, ward and ICU demand are

$$D_{r,W,t}^{\mathrm{dem}} = H_{r,t}, \quad D_{r,\mathrm{ICU},t}^{\mathrm{dem}} = \mathrm{ICU}_{r,t} \tag{2.33}$$

In an admission-flow interpretation, demand may instead be derived from ρI(t) and κH(t), which represent new hospital admissions and new ICU admissions. The occupancy-based interpretation is appropriate when the optimization model allocates capacity to existing patient load, while the arrival-flow interpretation is appropriate when the simulation layer models admissions, queues, waiting times, and length of stay.

Figure 2.3 illustrates the transitions among the susceptible (S), exposed (E), infectious (I), hospitalized (H), intensive care unit (ICU), recovered (R), and deceased (D) compartments of the extended SEIR-H-ICU-D epidemiological model.

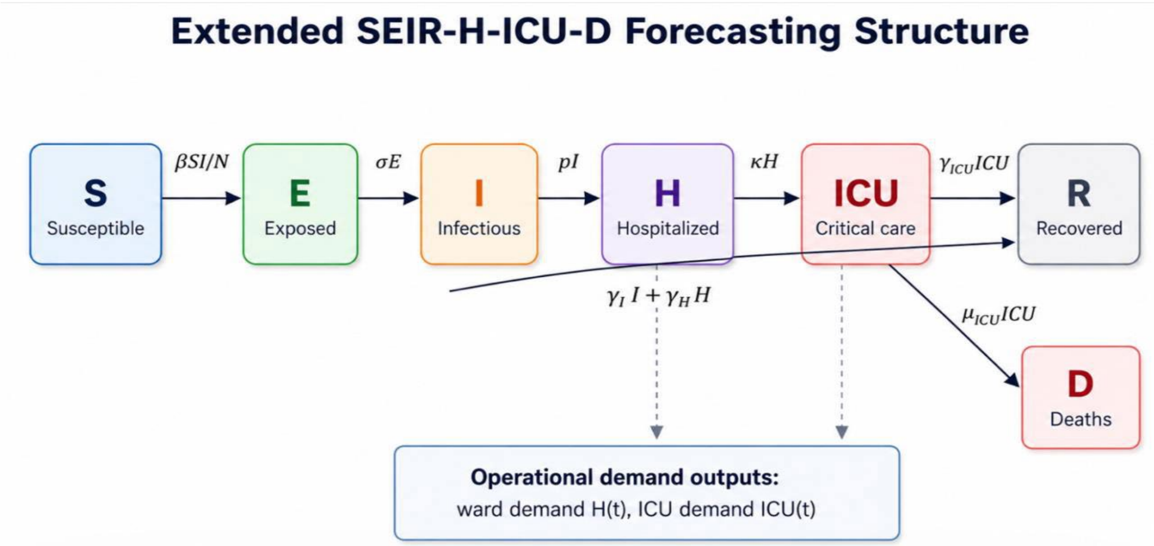

Figure 2.3. Extended SEIR-H-ICU-D state-transition structure used for healthcare demand forecasting

Table 2.4 defines the state variables and epidemiological parameters used in the extended SEIR-H-ICU-D model.

Table 2.4. State variables and epidemiological parameters

Notation	Definition	Operational interpretation
$S(t)$	Susceptible population	Individuals at risk of infection
$E(t)$	Exposed population	Infected but not yet infectious individuals
$I(t)$	Infectious population	Individuals capable of transmitting infection
$H(t)$	Hospitalized patients	Patients requiring ward-level care
$ICU(t)$	Intensive-care patients	Critically ill patients requiring ICU resources
$R(t)$	Recovered population	Individuals no longer requiring pandemic-related care
$D(t)$	Deceased population	Absorbing mortality state
$\beta(t)$	Effective transmission rate	Time-varying disease transmission intensity
σ	Incubation rate	Inverse of latent period
ρ	Hospitalization transition rate	Rate at which infectious individuals require hospital care
κ	ICU progression rate	Rate at which hospitalized patients deteriorate to ICU
$\gamma_i, \gamma_H, \gamma_{ICU}$	Recovery rates	Recovery from infectious, ward, and ICU states
μ	ICU mortality rate	Mortality rate among ICU patients

2.5.3 Handling Uncertainty in Epidemiological Parameters

The uncertainty set can be calibrated in several ways. If sufficient data are available, confidence intervals from parameter estimation can define lower and upper bounds. If data are limited, expert-defined scenarios may specify pessimistic, baseline, and optimistic parameter ranges. Historical variability across waves or regions can also inform admissible bounds. The important methodological requirement is that uncertainty bounds be documented and interpreted as part of the model assumptions.

There is a trade-off in setting the width of the uncertainty set. Narrow bounds may produce less conservative decisions but increase the risk that the real trajectory falls outside the protected region. Wide bounds may improve preparedness for adverse scenarios but can lead to overly conservative recommendations, excessive surge activation, or unnecessary transfers. Sensitivity analysis in the empirical chapters should therefore examine how decisions change when uncertainty bounds are varied.

The bounded uncertainty set also supports communication with decision-makers. Instead of presenting a single forecast as if it were certain, the framework can show how allocation decisions perform under a range of plausible epidemiological conditions. This helps decision-makers understand whether a recommendation is stable or highly sensitive to uncertain parameters.

The epidemiological parameters are not assumed to be perfectly known. During a pandemic, transmission rates, hospitalization probabilities, ICU progression rates, recovery rates, and mortality parameters may be uncertain because of reporting delays, asymptomatic infections, testing limitations, changes in clinical practice, and evolving interventions. At each decision epoch t, the critical parameter vector is defined as [3], [20], [102]

$$\theta_t^{\text{epi}} = \left(\beta_t, \rho_t, \kappa_t, \gamma_{I,t}, \gamma_{H,t}, \gamma_{\text{ICU},t}, \mu_t\right) \tag{2.34}$$

Uncertainty is represented by a bounded uncertainty set

$$\mathcal{U}_{\theta,t} = \{\theta_t\colon \theta_{m,t}^{L} \le \theta_{m,t} \le \theta_{m,t}^{U},\ m = 1, \dots, K\} \tag{2.35}$$

The bounds may be estimated from historical data, confidence intervals, surveillance uncertainty, scenario analysis, or expert assessment. The purpose of this representation is to prevent the optimization layer from relying on a single deterministic forecast. Instead, allocation decisions are evaluated over a range of plausible epidemiological scenarios. The robustness interpretation is conditional: the framework improves feasibility within the specified uncertainty set, but it does not guarantee protection against trajectories outside the modelled bounds.

2.6 Healthcare System and Resource Modeling

While the epidemiological component estimates the spatiotemporal demand generated by pandemic transmission, operational feasibility depends on the supply-side structure of the healthcare system. Pandemic healthcare capacity cannot be represented accurately as a static number of beds. A bed is operationally usable only when supported by appropriate staff, equipment, oxygen supply, clinical protocols, and transfer feasibility. Capacity must therefore be modelled as a dynamic and multidimensional construct [92], [105].

The healthcare system and resource modelling layer translates epidemiological demand into operational pressure. It defines the regional hospital network, quantifies effective ward and ICU capacity, models patient flows across care pathways, incorporates human-resource limitations, and defines utilization and load-balancing measures. These elements provide the feasibility structure for the optimization model in Section 2.7.

2.6.1 Hospital Network and Effective Capacity

The network representation is necessary because healthcare capacity is geographically and institutionally distributed. During a pandemic, overload is rarely uniform across all facilities. Some hospitals may face extreme ICU pressure while others retain reserve capacity. Modelling the system as G = (I,A) allows the framework to evaluate whether patient transfers or resource redistribution can reduce localized overload without creating new bottlenecks elsewhere.

Feasible transfer arcs are not purely geographical. A short physical distance does not guarantee transfer feasibility if ambulance capacity is unavailable, if clinical protocols prevent transfer, if the receiving hospital lacks the relevant care level, or if administrative authorization is absent. The arc set A therefore represents operational feasibility rather than only distance. Transfer costs can include distance, time, clinical risk, EMS burden, and coordination complexity [92], [110].d_{ij}

Effective capacity differs from nominal capacity in a way that is crucial for pandemic planning. Nominal capacity counts infrastructure, while effective capacity counts what can actually be used. For example, a hospital may report ICU beds but lack enough trained ICU nurses to operate all of them. Similarly, ventilator shortages can reduce ICU capacity even if beds and staff exist. The bottleneck formulation prevents these hidden constraints from being ignored.

The linear feasibility constraints are included because the optimization model is intended to be implementable. Minimum functions are intuitive for explanation, but mixed-integer optimization solvers typically require linear or linearized constraints. Representing bottlenecks through inequalities preserves the same operational logic while keeping the model compatible with standard optimization methods [92], [105].

The regional healthcare system is modelled as a directed network [92], [92]

$$G = (I, A)\mathcal{U}_{\theta,t} = \{\theta_t : \theta_{m,t}^L \leq \theta_{m,t} \leq \theta_{m,t}^U,\ m = 1, \ldots, K\} \tag{2.36}$$

where I denotes the set of healthcare facilities, hospitals, or modelled regions, and A denotes the set of feasible patient-transfer arcs. A directed arc (i,j) ∈ A indicates that a patient transfer from facility i to facility j is operationally feasible, subject to clinical appropriateness, transport availability, and receiving-capacity constraints. The notation avoids using N for hospitals because N denotes total population in the epidemiological model, and avoids using E for network edges because E(t) denotes the exposed compartment.

For each facility, time period, and care level, capacity is stratified into ward and intensive-care capacity. The notation in Equations (2.37) and (2.38) represents ward beds, ICU beds, ventilators, staff availability, and staffing ratios at each facility and period. Effective ward capacity is

$$C_{i,W,t}^{\text{eff}} = \min\{B_{i,W,t}, S_{i,W,t}^{\text{staff}}/r_W\} \tag{2.37}$$

and effective ICU capacity is

$$C_{i,\text{ICU},t}^{\text{eff}} = \min\{B_{i,\text{ICU},t}, V_{i,t}, S_{i,\text{ICU},t}^{\text{staff}}/r_{\text{ICU}}\} \tag{2.38}$$

These equations express capacity as a bottleneck-constrained quantity. A ward bed is not effective capacity without adequate staff, and an ICU bed is not effective ICU capacity without ventilators and trained ICU staff. For mixed-integer linear programming implementation, the minimum functions can be represented through linear feasibility constraints:

$$P_{i,W,t} \leq B_{i,W,t}, \quad r_W P_{i,W,t} \leq S_{i,W,t}^{\text{staff}} \tag{2.39}$$

$$P_{i,\text{ICU},t} \leq B_{i,\text{ICU},t}, \quad P_{i,\text{ICU},t} \leq V_{i,t}, \quad r_{\text{ICU}} P_{i,\text{ICU},t} \leq S_{i,\text{ICU},t}^{\text{staff}} \tag{2.40}$$

where $P_{i,c,t}$ denotes treated patient census in facility i, care level c, and time period t.

Table 2.5 summarizes the notation used to represent healthcare demand, capacity, staffing, transfers, and shortages.

Table 2.5. Healthcare resource modelling notation

Notation	Meaning	Unit / type
I	Healthcare facilities or modelled regions	Set
A	Feasible transfer arcs	Set
W	General ward care	Care level
ICU	Intensive care	Care level
$\boldsymbol{B_{i,W,t}}$	Ward beds	Beds
$\boldsymbol{B_{i,\mathrm{ICU},t}}$	ICU beds	Beds
$\boldsymbol{V_{i,t}}$	Ventilators	Units
$\boldsymbol{S_{i,c,t}^{\mathrm{staff}}}$	Available staff	Staff
$\boldsymbol{r_{\mathrm{W}}, r_{\mathrm{ICU}}}$	Staffing ratios	Ratio
$\boldsymbol{S_{i,c,t}^{\mathrm{seff}}}$	Effective capacity	Patients
$\boldsymbol{P_{i,c,t}}$	Treated patient census	Patients

2.6.2 Patient Flow and Transfer Constraints

Patient-flow conservation is the mechanism that prevents the model from creating or removing patients artificially. The census at the next time period must equal the current census plus admissions and inbound transfers, minus outbound transfers, discharges, deaths, and internal outflows, with internal inflows added where appropriate. This accounting structure is essential when the model is used over multiple periods because small inconsistencies can accumulate into unrealistic capacity states [110], [100].

Transfers are useful only when they are clinically and operationally feasible. Moving a patient from an overloaded hospital to a less-loaded one may reduce local pressure, but it consumes EMS resources and may expose the patient to transport risk. The model therefore treats transfers as decision variables with constraints and costs, not as automatic balancing mechanisms. This is important for ICU patients, for whom transfers may be clinically sensitive and resource-intensive.

Shortage is evaluated after transfers because transfers can change the distribution of demand across facilities. A hospital receiving inbound patients may become overloaded even if its original local demand was manageable. Conversely, a hospital transferring patients out may reduce its overload risk. The transfer-adjusted shortage formula therefore links network coordination directly to clinical unmet demand.

The EMS constraint provides a practical limit on regional coordination. Even if several hospitals have available beds, patients cannot be redistributed if ambulances, transfer teams, or transport routes are unavailable. During a pandemic surge, EMS capacity may itself become a bottleneck.

Including EMS_t prevents the optimization model from recommending transfer volumes that cannot be implemented.

Patients may be admitted to a ward, deteriorate and require ICU care, recover and be discharged, die, or be transferred to another facility. Patient movement is represented through flow conservation equations that track admissions, transfers, internal clinical transitions, recoveries, and deaths. Let $A_{i,c,t}$ denote new admissions, $y_{ij,c,t}$transfers from facility i to facility j, $\phi^{\text{in}}_{i,c,t}$internal inflows into care level c, $\phi^{\text{out}}_{i,c,t}$ internal outflows from care level c, $R_{i,c,t}$ recoveries or discharges, and $D_{i,c,t}$deaths.

$$P_{i,c,t+1} = \quad P_{i,c,t} + A_{i,c,t} + \sum_j y_{ji,c,t} - \sum_j y_{ij,c,t} + \phi^{\text{in}}_{i,c,t} - \phi^{\text{out}}_{i,c,t} - R_{i,c,t} - D_{i,c,t} \tag{2.41}$$

The distinction between inbound transfers $y_{ji,c,t}$ and outbound transfers $y_{ij,c,t}$ is essential for correct flow accounting. Every transfer leaving one facility enters another facility, provided the transfer arc is feasible. Capacity feasibility is represented as

$$P_{i,c,t} \leq C^{\text{eff}}_{i,c,t} \tag{2.42}$$

When demand exceeds available effective capacity, excess demand is recorded as shortage rather than forcing patient census above physical feasibility:

$$q_{i,c,t} \geq D^{\text{dem}}_{i,c,t} + \sum_j y_{ji,c,t} - \sum_j y_{ij,c,t} - C^{\text{eff}}_{i,c,t} \tag{2.43a}$$

$$q_{i,c,t} \geq 0 \tag{2.43b}$$

Transfers must also satisfy feasibility restrictions. A facility cannot transfer more patients out of a care level than the number currently assigned to that level, total transfers are limited by EMS capacity, and transfers are allowed only along feasible arcs:

$$\sum_j y_{ij,c,t} \leq P_{i,c,t} \tag{2.44}$$

$$\sum_i \sum_j \sum_c y_{ij,c,t} \leq \text{EMS}_t \tag{2.45}$$

$$y_{ij,c,t} = 0, \quad (i,j) \notin A \tag{2.46}$$

Figure 2.4 illustrates the patient-flow and healthcare-resource pathways from admission through ward and ICU care, including capacity-constrained inter-hospital transfers, recovery or discharge, and mortality.

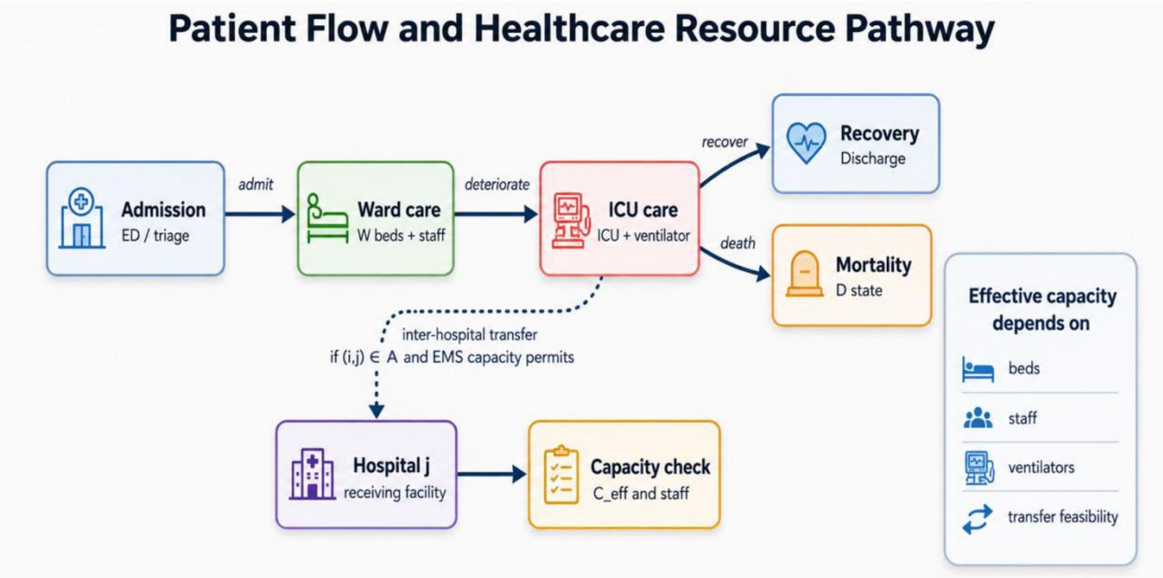

Figure 2.4. Patient-flow and healthcare-resource pathway under ward, ICU, transfer, recovery, and mortality dynamics

2.6.3 Workforce and Surge Capacity

The staff depletion parameter $\alpha_{i,t}$ can represent multiple operational pressures, including infection among healthcare workers, quarantine requirements, fatigue, burnout, or redeployment to other departments. Because these pressures can vary across facilities, $\alpha_{i,t}$ is indexed by facility and time. This allows the model to represent localized workforce crises rather than assuming uniform staff availability across the network.

Surge staffing is treated as limited because emergency staffing cannot be expanded indefinitely. Temporary redeployment may reduce capacity in other services, external support may be delayed, and staff may not be qualified for ICU care. The upper bound $S_{i,c,t}^{\text{surge,max}}$ therefore reflects operational realism. It also allows the optimization model to decide when surge staffing should be used and when other actions, such as transfers or load balancing, are preferable.

Workforce modelling also supports equity and resilience. A policy that keeps all hospitals close to maximum staff utilization may be mathematically feasible in the short term but unsafe over a longer horizon. By combining staffing constraints with utilization and load-balancing penalties, the framework discourages localized exhaustion and supports more sustainable regional coordination [92], [105].

During prolonged emergencies, human resources may become more restrictive than physical infrastructure. A hospital may have beds available but still be unable to admit additional patients if it lacks nurses, physicians, respiratory therapists, or ICU-trained staff. Workforce availability may decline because of infection, quarantine, fatigue, burnout, redeployment, or family-care responsibilities.

Let $S_{i,c}^{\text{base}}$ denote the baseline staff assigned to facility i and care level c. Let $\alpha_{i,t} \in [0,1]$ denote the staff depletion rate at time t. Let $S_{i,c,t}^{\text{surge}}$ denote temporary additional staff made available through emergency mobilization, redeployment, or external support. Available staff are modelled as

$$S_{i,c,t}^{\text{staff}} = S_{i,c}^{\text{base}}\left(1 - \alpha_{i,t}\right) + S_{i,c,t}^{\text{surge}} \tag{2.47}$$

Surge staffing is bounded by operational feasibility:

$$0 \leq S_{i,c,t}^{\text{surge}} \leq S_{i,c,t}^{\text{surge,max}} \tag{2.48}$$

This formulation distinguishes ward staff from ICU staff. ICU staffing constraints are usually more restrictive because intensive care requires specialized personnel and cannot be substituted easily by general ward staff.

2.6.4 Utilization and Load Balancing

Utilization is not included only as a descriptive indicator; it affects the quality and safety of the healthcare system. High utilization can reduce flexibility, increase waiting times, limit the ability to respond to sudden admissions, and intensify staff pressure. Monitoring utilization by facility and care level allows the framework to detect where overload risk is concentrated.

The load-balancing penalty is especially important in regional coordination. Without it, the optimizer might minimize total shortage while still leaving one hospital near collapse and another underused. Penalizing deviations from average utilization encourages a more even distribution of pressure, subject to transfer feasibility and clinical constraints. This does not mean that all hospitals must have identical utilization; rather, extreme imbalance is discouraged when feasible alternatives exist [92], [110].

The linearization of the absolute value requires two inequalities because the deviation can be positive or negative. Using only one inequality would capture only one direction of deviation and would not correctly represent $|\eta_{i,c,t} - \bar{\eta}_{c,t}|$. The auxiliary variable $e_{i,c,t}$ therefore ensures a correct linear representation of the load-balancing term.

The resource modelling layer also prepares the interface with simulation. The same variables used in optimization - treated census, capacity, transfers, utilization, and shortage - become inputs or validation targets for the DES layer. This shared variable structure is what allows the IHDSF to connect mathematical feasibility with stochastic operational performance.

To monitor facility pressure, the utilization ratio is defined as

$$\eta_{i,c,t} = \frac{P_{i,c,t}}{C_{i,c,t}^{\text{eff}}} \tag{2.49}$$

A utilization ratio close to 1 indicates full use of effective capacity, while a value above a predefined safety threshold indicates overload risk or operational stress. To prevent localized collapse, the framework includes a load-balancing penalty. Let $\bar{\eta}_{c,t}$ denote average utilization across the network for care level c:

$$\bar{\eta}_{c,t} = \frac{1}{|I|}\sum_{i} \eta_{i,c,t} \tag{2.50}$$

The network imbalance penalty is

$$E_t = \sum_{c} \sum_{i} \left|\eta_{i,c,t} - \bar{\eta}_{c,t}\right| \tag{2.51}$$

In mixed-integer linear programming, the absolute value is linearized through auxiliary variables $e_{i,c,t} \geq 0$. Both inequalities are required:

$$e_{i,c,t} \geq \eta_{i,c,t} - \bar{\eta}_{c,t} \tag{2.52a}$$

$$e_{i,c,t} \geq \bar{\eta}_{c,t} - \eta_{i,c,t} \tag{2.52b}$$

$$E_t = \sum_{c} \sum_{i} e_{i,c,t} \tag{2.53}$$

2.6.5 Summary and Link to Optimization

The healthcare resource modelling layer establishes the operational feasibility structure of the IHDSF. It links epidemiological demand projections with effective capacity, patient-flow conservation, transfer feasibility, staffing limitations, and load-balancing requirements. These elements provide the constraints and objective-function components used in the robust multi-objective optimization model developed in Section 2.7 [92], [105].

2.7 Optimization Strategies and Algorithmic Design

The epidemiological model forecasts demand, and the resource model defines operational feasibility. The central prescriptive task is to determine which allocation decisions should be implemented when the healthcare system faces uncertain demand, scarce resources, and conflicting clinical and logistical objectives. The optimization component is therefore the prescriptive engine of the IHDSF. It transforms demand forecasts and capacity constraints into allocation, transfer, surge-capacity, and staffing decisions [92], [8], [22].

Pandemic healthcare allocation is formulated as a robust multi-objective optimization problem because no single indicator captures the decision challenge. A strategy that minimizes transfers may leave some hospitals overloaded. A strategy that balances load aggressively may increase transport burden. A strategy that minimizes unmet demand may require emergency surge resources. The optimization model must therefore represent trade-offs among clinical protection, regional load balancing, logistical feasibility, and operational resilience.

2.7.1 Decision Variables and Feasible Space

The feasible decision space is not static. It may change as capacity, staff availability, transfer routes, or institutional rules change. For example, if a facility loses ICU staff because of absenteeism, the effective capacity and feasible allocation set must be updated. If EMS capacity declines, feasible transfer volumes must be reduced. The rolling-horizon workflow in Section 2.9 is designed to update as new operational information becomes available [92], [107].$X_t X_t$

The decision variables have different mathematical types. Transfers and admissions are usually integer because patients are discrete units. Surge activation may be binary when it represents whether an emergency unit is opened, or integer when it represents the number of temporary beds activated. Staff deployment may be continuous in aggregate models or integer if individual staff units are modelled. These choices should be aligned with the empirical data and solver implementation.

The model can also incorporate policy restrictions through X_t. For example, certain transfers may be prohibited for clinical reasons, certain facilities may be designated as COVID-19 referral centres, or crisis standards may impose priority rules. Encoding such restrictions in the feasible set ensures that recommendations remain consistent with real decision authority.

Let the decision vector at time period t be represented as

$$u_t = \left(a_{i,c,t}, y_{ij,c,t}, z_{i,c,t}, s_{i,c,t}^{\text{surge}}\right) \tag{2.54}$$

where $a_{i,c,t}$denotes admission or allocation decisions, $y_{ij,c,t}$denotes patient transfers from facility i to facility j, $z_{i,c,t}$ denotes activation of surge capacity, and $s_{i,c,t}^{\text{surge}}$ denotes emergency staff deployment. These variables are constrained by effective capacity, patient-flow, staffing, transfer, and institutional feasibility conditions.

The feasible decision space is denoted by X_t and includes all policies satisfying physical capacity, ICU availability, staff feasibility, transfer-route feasibility, EMS capacity, surge-capacity limits, and non-negativity or integrality requirements:

$$u_t \in \mathcal{X}_t, \quad t \in T \tag{2.55}$$

Table 2.6 defines the core decision variables used in the optimization layer.

Table 2.6. Core decision variables in the optimization layer

Variable	Meaning	Type
$\boldsymbol{a_{i,c,t}}$	Admission or allocation decision	Integer
$\boldsymbol{y_{ij,c,t}}$	Transfer from facility i to facility j	Integer
$\boldsymbol{z_{i,c,t}}$	Surge-capacity activation	Binary / integer
$\boldsymbol{s_{i,c,t}^{\text{surge}}}$	Emergency staff deployment	Continuous / integer
$\boldsymbol{q_{i,c,t}}$	Unmet demand	Continuous
$\boldsymbol{e_{i,c,t}}$	Load-balance deviation	Continuous

2.7.2 Multi-Objective Optimization Formulation

Objective normalization deserves explicit attention because the four components have different physical meanings and magnitudes. Unmet demand may be measured in patients, load imbalance in utilization deviations, transfers in distance-weighted patient movements, and surge use in activated emergency resources. Without normalization, the objective with the largest numerical scale could dominate the solution even if decision-makers did not intend that priority. Normalized objectives therefore make the weights interpretable as preference parameters rather than accidental unit conversions [22], [22].

Weight selection can be performed through expert consultation, policy priorities, sensitivity analysis, or scenario comparison. In emergency conditions, weights may be selected to give strong priority to ICU shortage and mortality-related risk, while still penalizing excessive transfers and inequitable load distribution. Because weights involve value judgments, they should be reported transparently in the empirical chapter. A useful practice is to compare several weight settings, such as clinically conservative, logistics-constrained, and equity-oriented configurations, and examine whether the recommended policies remain stable [22].

The epsilon-constraint formulation is a useful alternative when decision-makers are uncomfortable assigning direct numerical weights to ethical or clinical priorities. For example, the model may minimize unmet ICU demand while requiring transfer burden and surge use to remain below acceptable thresholds. This approach is often easier to explain to healthcare authorities because it frames non-clinical objectives as operational limits rather than as direct trade-offs against clinical shortage.

The unmet-demand objective is normally assigned the highest priority because shortage is directly connected to clinical risk and system failure. ICU shortage may receive a higher penalty than ward shortage because ICU patients are more clinically vulnerable and because delays in intensive care can have severe consequences. The penalty $p_c$therefore allows the model to differentiate the severity of shortages across care levels.

The load-balancing objective represents an equity and resilience consideration. Equity here does not mean identical treatment of all facilities regardless of context; it means avoiding preventable concentration of overload when the network has feasible alternatives. This objective is important in pandemic conditions because localized overload can create cascading consequences for neighbouring facilities and regional coordination.

The transfer-cost objective prevents the model from using transfers excessively. Transfers may reduce shortage but also consume EMS capacity, require coordination, and may create clinical risk for unstable patients. Including f_3 allows the model to balance the benefits of redistributing demand against the operational burden of doing so.

The surge-cost objective represents the fact that emergency capacity is not free. Temporary beds, redeployed staff, expanded ICU capacity, and emergency equipment involve financial, operational, and human-resource costs. Penalizing surge use encourages the model to activate emergency resources only when they are justified by demand pressure and clinical priorities.

Two common approaches to multi-objective optimization are supported. The weighted-sum formulation is transparent and easy to solve when priorities can be expressed as weights. The epsilon-constraint formulation is useful when one objective, such as minimizing clinical shortage, is primary and the others are constrained within acceptable limits. The empirical implementation may use either approach depending on decision-maker preferences.

The first objective minimizes unmet clinical demand. Let p_c denote the shortage penalty associated with care level c, where ICU shortage is assigned a higher penalty than ward shortage. The unmet-demand objective is

$$f_1 = \sum_{t\in T} \sum_{i\in I} \sum_{c\in C} p_c \, q_{i,c,t} \tag{2.56}$$

The second objective minimizes imbalance in utilization across the healthcare network. Using the auxiliary deviation variable $e_{i,c,t}$ from Section 2.6, the load-balancing objective is

$$f_2 = \sum_{t\in T} \sum_{c\in C} \sum_{i\in I} e_{i,c,t} \tag{2.57}$$

The third objective minimizes the logistical burden of patient transfers. Let d_{ij} denote distance, time, or generalized transfer cost between facilities i and j. The transfer-cost objective is

$$f_3 = \sum_{t \in T} \sum_{(i,j) \in A} \sum_{c \in C} d_{ij}\, y_{ij,c,t} \tag{2.58}$$

The fourth objective penalizes emergency surge-resource use. Let $h_{i,c}$ denote the marginal operational burden of surge activation for facility i and care level c. The surge-cost objective is

$$f_4 = \sum_{t \in T} \sum_{i \in I} \sum_{c \in C} h_{i,c}\, z_{i,c,t} \tag{2.59}$$

The scalarized multi-objective problem is written using u consistently as the decision variable:

$$\min_{u \in \mathcal{X}} Z = w_1 \tilde{f}_1 + w_2 \tilde{f}_2 + w_3 \tilde{f}_3 + w_4 \tilde{f}_4 \tag{2.60}$$

where $\tilde{f}_k$ denotes a normalized objective component and $w_k$the corresponding decision-maker-defined weight. Normalization is necessary because the objective components are measured in different units, including patients, utilization deviations, transfer distances, and surge-resource costs. An alternative epsilon-constraint formulation minimizes the clinically dominant objective while bounding other objectives:

$$\min_{u \in \mathcal{X}} f_1 \quad \text{subject to} \quad f_2 \leq \varepsilon_2,\ f_3 \leq \varepsilon_3,\ f_4 \leq \varepsilon_4 \tag{2.61}$$

Table 2.7 summarizes the objective functions of the optimization model and their operational interpretation.

Table 2.7. Objective functions and operational interpretation

Objective	Focus	Interpretation
$\boldsymbol{f_1}$	Unmet demand	Clinical shortage
$\boldsymbol{f_2}$	Load balancing	Regional equity and overload prevention
$\boldsymbol{f_3}$	Transfers	Logistical and EMS burden
$\boldsymbol{f_4}$	Surge use	Emergency resource burden

Figure 2.5 is illustrative rather than an empirical result. It shows the trade-off logic behind the multi-objective model: policies that reduce unmet demand may require more transfers and emergency capacity, while policies that reduce logistical burden may accept higher clinical shortage. The role of the decision-maker is to select a policy according to clinical priorities, institutional feasibility, and ethical considerations.

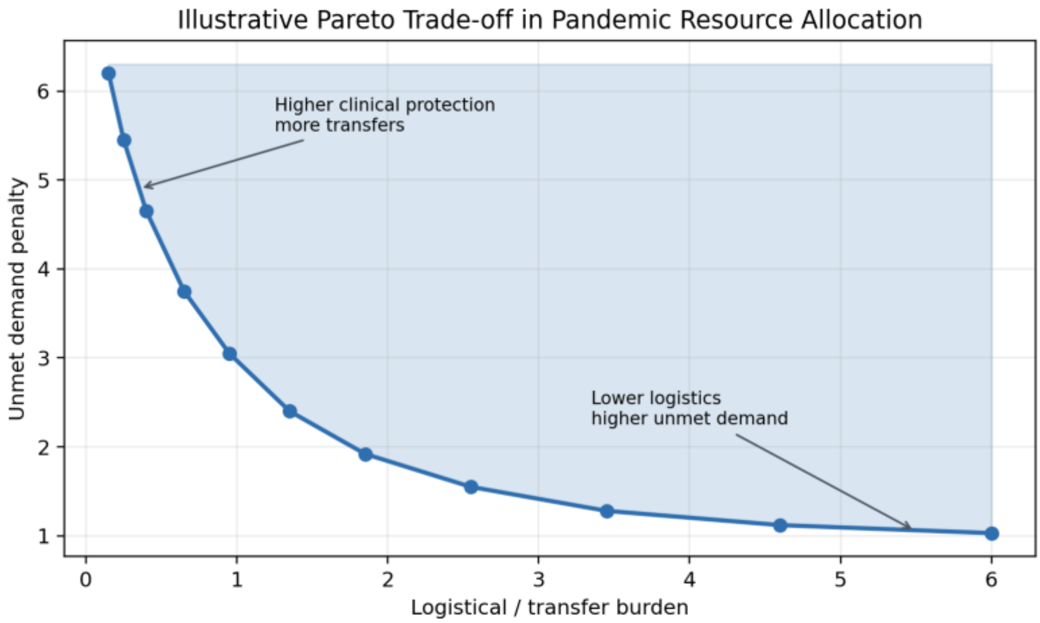

Figure 2.5. Illustrative Pareto trade-off between unmet demand and logistical transfer burden in pandemic resource allocation

2.7.3 Robust Counterpart Under Epidemiological Uncertainty

The robust counterpart is particularly useful in early outbreak phases, when parameter estimates are unstable and hospital demand may increase rapidly. Under such conditions, a deterministic solution based on a single forecast can produce fragile allocation plans. The robust formulation instead evaluates candidate decisions against adverse but plausible conditions, reducing the probability that a policy becomes infeasible shortly after implementation [20], [20], [102].

Robust optimization also supports transparency. Decision-makers can be informed that a policy was tested against a specified uncertainty set, such as higher hospitalization rates, increased ICU progression, or reduced staff availability. This makes the protection level explicit. If decision-makers want more conservative or less conservative recommendations, the uncertainty set or objective weights can be adjusted and the implications examined [20], [102].

The approach should not be confused with stochastic optimization. Stochastic optimization requires probability distributions over scenarios, while robust optimization protects against all realizations inside a specified uncertainty set. In practice, the IHDSF combines robust optimization with simulation: optimization protects feasibility under bounded uncertainty, and simulation evaluates stochastic patient-flow performance.

The deterministic formulation assumes that demand, capacity, and transition parameters are known. During pandemic response this assumption is too strong. Demand depends on uncertain epidemiological parameters such as transmission rates, hospitalization probabilities, ICU progression rates, recovery rates, and reporting delays. To account for this uncertainty, the IHDSF uses a robust counterpart of the multi-objective optimization model:

$$\min_{u\in\mathcal{X}}\max_{\theta\in\mathcal{U}_\theta}\left[w_1\tilde{f}_1(u,\theta)+w_2\tilde{f}_2(u,\theta)+w_3\tilde{f}_3(u)+w_4\tilde{f}_4(u)\right] \tag{2.62}$$

This formulation selects the feasible policy that performs best under the worst admissible realization of uncertainty. The robustness claim is conditional on the defined uncertainty set. The model improves feasibility for parameter values inside the defined uncertainty set, but it does not guarantee protection against all possible epidemic trajectories outside the modelled bounds. The width and calibration of this uncertainty set therefore influence the balance between preparedness and conservatism.

2.7.4 Core Constraints of the Optimization Model

The capacity constraint prevents the treated census from exceeding effective capacity. In real systems, crisis conditions may lead to care beyond normal capacity, but such overload should not be hidden inside the model. If emergency overflow is allowed, it should be represented explicitly through surge capacity, shortage, or safety-threshold penalties rather than by silently violating capacity limits.

The transfer constraints protect the model from recommending impossible patient movements. Outgoing transfers are limited by the patients currently assigned to a facility and care level. Total transfers are limited by EMS capacity. Transfer variables are set to zero when a route is infeasible. These constraints together ensure that regional load balancing remains operational rather than purely mathematical.

Non-negativity and integrality constraints preserve physical meaning. Negative patients, negative shortages, or fractional patient transfers are not operationally meaningful in most implementations. Some aggregate variables may be relaxed for computational purposes, but the final interpretation must remain consistent with discrete patient and resource units.

The optimization model is constrained by the healthcare resource model developed in Section 2.6. The essential feasibility conditions are summarized as follows:

$$P_{i,c,t} \le C^{\text{eff}}_{i,c,t} \tag{2.63}$$

$$q_{i,c,t} \ge D^{\text{dem}}_{i,c,t} + \sum_{j:(j,i)\in A} y_{ji,c,t} - \sum_{j:(i,j)\in A} y_{ij,c,t} - C^{\text{eff}}_{i,c,t} \tag{2.64}$$

$$\sum_{j:(i,j)\in A} y_{ij,c,t} \le P_{i,c,t} \tag{2.65}$$

$$\sum_{i}\sum_{j}\sum_{c} y_{ij,c,t} \leq \mathrm{EMS}_t \tag{2.66}$$

$$y_{ij,c,t} = 0, \quad \forall (i,j) \notin A \tag{2.67}$$

$$q_{i,c,t} \geq 0, \quad e_{i,c,t} \geq 0, \quad y_{ij,c,t} \in \mathbb{Z}_+ \tag{2.68}$$

These constraints ensure that the optimization model respects physical capacity, patient-flow feasibility, transportation limits, and implementation realism. Additional constraints may be added for clinical priority rules, regional policy restrictions, or scenario-specific emergency resource limits.

2.7.5 Hybrid Optimization Approach

The hybrid approach reflects the practical tension between mathematical rigor and real-time usefulness. Exact optimization is desirable because it provides strong guarantees for the specified formulation, but crisis decision-making may require solutions within limited time. When the problem becomes too large for exact methods, heuristic or simheuristic procedures can produce high-quality solutions that are operationally useful even without full optimality proofs [107], [111], [112], [113].

A simheuristic approach is especially relevant when stochastic performance must be considered. The optimization model can guide the search toward promising allocation policies, while simulation evaluates how those policies perform under random arrivals, length of stay, staff shocks, and transfer delays. Poorly performing policies can be rejected or penalized, and the search can be redirected toward more robust alternatives [111], [112], [113].

The selected algorithmic approach should be reported transparently in the empirical chapter. Important implementation details include solver type, optimality gap, time limit, number of scenarios, number of simulation replications, stopping criteria, and hardware. These details allow the computational results to be reproduced and assessed by reviewers.

The resulting model is a multi-period, multi-objective, network-based mixed-integer optimization problem. Large instances may become computationally demanding because they include multiple hospitals, care levels, transfer arcs, time periods, and uncertainty scenarios. Exact solvers can provide strong solutions for well-structured instances, but computational burden may become prohibitive in large-scale or frequently updated pandemic settings.

For this reason, the dissertation adopts a hybrid optimization approach. The short-term tactical horizon is solved using exact mixed-integer programming whenever tractable. For larger horizons or deep-uncertainty settings, the framework can use metaheuristic or simheuristic search to generate high-quality candidate policies within operationally acceptable time limits. Simulation-assisted search is particularly relevant when stochastic performance indicators must be evaluated repeatedly.

Algorithm 2.1. Hybrid optimization procedure

1. Initialize demand forecasts from the epidemiological model.
2. Update effective capacity, staffing, transfer feasibility, and surge limits.
3. Construct the multi-objective optimization formulation.
4. Define uncertainty bounds for epidemiological and operational parameters.
5. Solve the robust tactical model if the instance is tractable.
6. Activate metaheuristic or simheuristic search for large instances.
7. Generate candidate allocation, transfer, and surge-capacity policies.
8. Pass candidate policies to the simulation layer for stress testing.
9. Select a policy satisfying feasibility, robustness, and operational thresholds.
10. Implement the selected near-term policy within the rolling-horizon workflow.

The optimization layer defines allocation, transfer, surge-capacity, and staffing decisions within a feasible decision space determined by the healthcare resource model. Its objective balances unmet demand, load balancing, transfer burden, and surge-resource use, while the robust counterpart accounts for epidemiological and operational uncertainty. The resulting candidate policies are passed to the simulation-validation layer developed in Section 2.8.

2.8 Integration of Simulation and Optimization

The optimization model developed in Section 2.7 generates candidate resource allocation, transfer, and surge-capacity policies. However, a mathematically feasible policy may still perform poorly under stochastic patient-flow conditions. Pandemic healthcare systems are affected by random patient arrivals, uncertain clinical progression, variable length of stay, transfer delays, staff absenteeism, and ICU congestion. For this reason, the IHDSF includes a simulation-based validation layer [92], [8], [77].

Discrete Event Simulation (DES) is used to evaluate candidate policies under stochastic operational conditions. The DES layer represents patient arrivals, admissions, ward and ICU transitions, transfers, discharges, deaths, queues, and resource contention. Its purpose is not to replace the optimization model, but to test whether optimized policies remain operationally feasible when variability is introduced. The term digital-twin-like is used cautiously: the simulation is an operational representation of the healthcare network sufficient for stress testing, not a complete reproduction of every microscopic clinical detail [77], [100].

2.8.1 Discrete Event Simulation Layer

The DES layer is appropriate because healthcare operations are event-driven. Admission, bed assignment, deterioration, ICU transfer, recovery, discharge, death, and inter-facility transfer occur at specific times and change the system state. Modelling these events explicitly allows the framework to estimate waiting times, queues, and congestion patterns that are difficult to capture in a deterministic optimization model [77], [97].

For facility i, care level c, time period t, and simulation replication r, patient arrivals may be represented as

$$A_{i,c,t}^{(r)} \sim \text{Poisson}(\lambda_{i,c,t}) \tag{2.69}$$

where $\lambda_{i,c,t}$ is derived from the epidemiological demand forecast. Other arrival models, such as non-homogeneous Poisson processes, empirical sampling, or scenario-based profiles, may be used when supported by data. Length of stay may be modelled using positive right-skewed distributions such as lognormal, gamma, Weibull, or empirical distributions. For example,

$$\text{LOS}_k^{\text{ICU}} \sim \text{Lognormal}(\mu_{\text{ICU}}, \sigma_{\text{ICU}}^2), \quad \text{LOS}_k^{W} \sim \text{Lognormal}(\mu_W, \sigma_W^2) \tag{2.70}$$

The specific distribution should be selected during empirical calibration according to observed hospital data. The purpose of the DES layer is not to impose a universal distribution but to provide a flexible stochastic evaluation mechanism.

2.8.2 Patient Pathways and Performance Indicators

A patient may enter the system through emergency admission, be assigned to ward care, deteriorate to ICU, recover and be discharged, die, or be transferred to another facility. These pathways correspond to the epidemiological and resource components defined earlier, but DES adds stochastic timing and resource contention. At every event, resource feasibility is checked. If the required resource is unavailable, the patient enters a queue or shortage state, allowing the model to estimate waiting time and congestion risk.

For each candidate policy u, the DES layer estimates operational performance across R simulation replications:

$$\hat{J}(u) = \frac{1}{R}\sum_{r=1}^{R} J\,(u, \omega_r) \tag{2.71}$$

where ω_r denotes stochastic realization r and $J(u, \omega_r)$may include shortage, waiting time, transfer cost, ICU overload, utilization imbalance, and mortality-related risk. The main performance indicators are summarized in Table 2.8.

Table 2.8. Simulation-based performance indicators

Indicator	Symbol	Purpose
Waiting time	W_q	Delay and clinical risk
Queue length	L_q	Congestion measurement
Utilization	η	Resource pressure
Unmet demand	q	Clinical shortage
Transfers	$\mathrm{N}^{\mathrm{transfer}}$	EMS and coordination burden
Overload probability	$\mathrm{P}_{\mathrm{over}}$	Safety violation risk
Mortality-related risk	M	Clinical outcome proxy

These indicators are important because a policy may appear feasible in the optimization formulation but fail under stochastic operational stress. For example, an allocation plan may satisfy expected ICU capacity while still generating long ICU queues in the upper tail of the demand distribution [77], [100].

2.8.3 Coupling DES with the Optimization Module

The simulation-optimization coupling is bidirectional. The optimization model generates a candidate policy u*, while the DES layer evaluates this policy under stochastic patient-flow conditions. For each replication r, a stochastic scenario is sampled: [113], [112]

$$\omega_r = \left(A^{(r)}, \mathrm{LOS}^{(r)}, \xi_{\mathrm{staff}}^{(r)}, \xi_{\mathrm{transfer}}^{(r)}\right) \tag{2.72}$$

where $A^{(r)}$represents arrivals, $\mathrm{LOS}^{(r)}$ length-of-stay realizations, $\xi_{\mathrm{staff}}^{(r)}$ staff-availability shocks, and $\xi_{\mathrm{transfer}}^{(r)}$ transfer-delay or EMS-capacity shocks. The candidate policy is evaluated by

$$\mathrm{KPI}^{(r)} = \mathrm{DES}(u^{\star}, \omega_r) \tag{2.73}$$

and aggregated across replications as

$$\widehat{\mathrm{KPI}}(u^{\star}) = \frac{1}{R}\sum_{r=1}^{R} \mathrm{KPI}^{(r)} \tag{2.74}$$

The resulting indicators are compared with safety and operational thresholds. If the candidate policy violates these thresholds, feedback is returned to the optimization module through revised penalties, additional constraints, or updated parameters. This loop distinguishes policies that are optimal only under average conditions from policies that remain acceptable under stochastic stress.

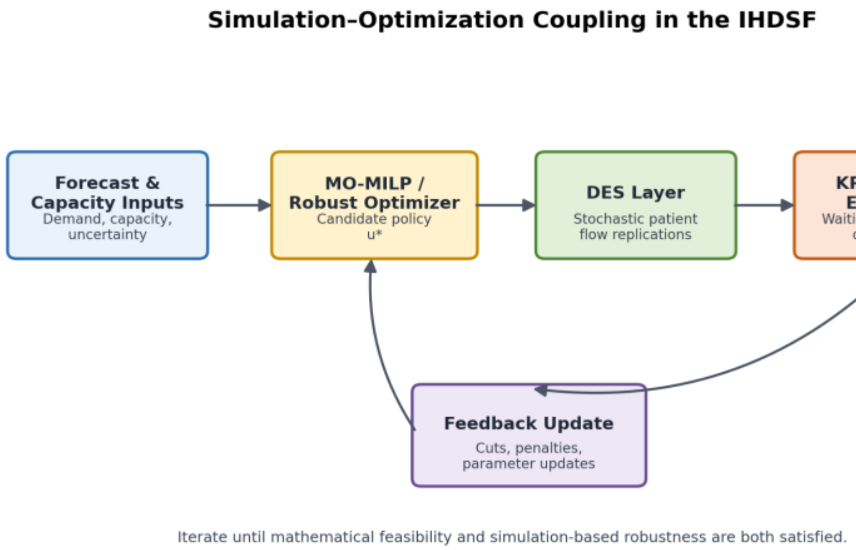

Figure 2.6. Simulation-optimization coupling in the IHDSF

2.8.4 Safety Thresholds, Feedback, and Validation

Safety thresholds translate simulation evidence into decision rules. A probabilistic ICU waiting-time threshold can be written as [77], [97]

$$\Pr\big(W_{q,\mathrm{ICU}}(u^\star,\omega) > W_{\mathrm{ICU}}^{\max}\big) \leq \alpha \tag{2.75}$$

where the right-hand threshold denotes the maximum acceptable ICU waiting time and α is the tolerated violation probability. Alternatively, a quantile-based safety threshold may be used:

$$Q_{0.95}\left(W_{q,\mathrm{ICU}}(u^\star,\omega)\right) \leq W_{\mathrm{ICU}}^{\max} \tag{2.76}$$

If a violation is detected, the objective function may be updated as

$$Z_{\mathrm{new}}(u) = Z(u) + \lambda_{\mathrm{sim}}\max\{0, \hat{v}(u) - v^{\max}\} \tag{2.77}$$

where the simulated violation measure is compared with the acceptable threshold and the simulation penalty controls the severity of the update. Alternatively, the feasible set may be refined by adding a simulation-informed constraint:

$$g_{\mathrm{sim}}(u) \leq b_{\mathrm{sim}} \tag{2.78}$$

The loop continues until the candidate policy satisfies both mathematical feasibility and simulation-based operational robustness. Verification and validation of the DES layer are necessary before using its outputs for decision support. Verification checks the internal implementation, including patient-flow conservation, non-negative queues, valid event ordering, capacity feasibility, and transfer accounting. Validation compares simulated outputs with empirical data or, where direct data are unavailable, expert review, scenario plausibility checks, and sensitivity analysis. Detailed simulation precision formulas and full verification tables can be placed in the appendices.

2.9 Decision-Support Workflow and Adaptive Logic

The IHDSF is designed as a rolling-horizon decision-support system. At each update time t, new epidemiological, hospital, staffing, and logistics data are collected and transformed into an updated system-state estimate. The model then recalibrates parameters, generates demand forecasts, solves the optimization problem, validates candidate decisions through simulation, and provides recommendations for human review. This cyclical structure is necessary because pandemic information becomes outdated quickly [114], [115].

2.9.1 Real-Time Decision-Support Process

The workflow is designed to be compatible with hospital information systems, bed-management platforms, public health surveillance data, staffing records, and regional coordination databases. Compatibility does not imply that the model is automatically integrated with electronic health records in every empirical context. The actual level of integration depends on the available data infrastructure and institutional permissions [115], [116].

Each update cycle produces a traceable recommendation. The decision-maker can see the input data, demand forecast, capacity estimate, uncertainty assumptions, optimized policy, simulation results, and safety warnings. This traceability is important for trust and accountability, especially when recommendations involve sensitive actions such as ICU transfers or emergency resource redistribution.

The real-time process should also include exception handling. If data are missing or inconsistent, the model should flag uncertainty rather than silently producing a precise recommendation. If optimization becomes infeasible, the system should report the binding constraints and shortage patterns. If simulation indicates unsafe performance, the recommendation should be withheld or marked for revision.

The term real-time is used in an operational sense. The framework updates whenever new data become available at the frequency supported by the healthcare system. In some contexts this may be hourly; in others it may be daily or weekly. The methodological logic remains the same: each update cycle transforms observed system data into validated and implementable recommendations.

The input vector at time t may be written as

$$Y_t = \left(Y_t^{\text{epi}}, Y_t^{\text{hosp}}, Y_t^{\text{staff}}, Y_t^{\text{log}}\right) \tag{2.79}$$

and the updated state estimate as

$$\hat{x}_t = \mathcal{H}(Y_t, \hat{x}_{t-1}) \tag{2.80}$$

where H denotes a data harmonization and state-estimation operator. This operator reconciles new observations with the previous system state, handles missing data, corrects reporting delays, and updates hospital occupancy and staff availability. The updated demand forecast over the planning horizon is

$$\widehat{D}_{t:t+H_p} = \mathcal{F}\left(\hat{x}_t, \hat{\theta}_t\right) \tag{2.81}$$

The optimization layer then produces a candidate policy

$$u_t^{\star} = \mathcal{O}\left(\widehat{D}_{t:t+H_p}, C_t^{\text{eff}}, \mathcal{U}_t, \mathcal{X}_t\right) \tag{2.82}$$

which is evaluated through simulation before implementation. The final recommendation is not an automatic command. It is a structured decision-support output reviewed by healthcare administrators, public health officials, clinical leaders, and regional coordinators.

2.9.2 Adaptive and Rolling-Horizon Mechanism

The distinction between planning horizon H_p and execution horizon $H_e$is central. The planning horizon allows the model to anticipate future demand peaks, while the execution horizon prevents the system from committing to a long plan based on uncertain forecasts. For example, a model may plan over 14 days but implement only the next 24 or 48 hours of transfers and staffing decisions.

Rolling-horizon control is appropriate for pandemic healthcare because it accepts that forecasts are provisional. Instead of requiring perfect long-term prediction, the framework repeatedly updates its forecast and decisions as new evidence arrives. This makes the methodology resilient to reporting corrections, sudden outbreaks, staff depletion, and changes in transfer feasibility.

The rolling-horizon mechanism also supports scenario testing. Decision-makers can compare how policies change under baseline, severe surge, staff-depletion, or EMS-restriction scenarios. If a recommendation remains stable across scenarios, it may be considered more robust. If it changes sharply, the model highlights where additional preparedness or contingency planning is needed.

The central challenge in pandemic healthcare logistics is uncertainty over longer horizons. A plan generated for several weeks may become obsolete after only a few days because of a localized outbreak, a sudden staff shortage, a reporting correction, or a change in public health restrictions. The IHDSF therefore adopts a rolling-horizon logic inspired by model predictive control and dynamic optimization [117], [98], [118].

At decision time t, the system optimizes over a planning horizon H_p, but implements only a shorter execution horizon H_e, where $H_e < H_p$. The optimization problem solved at time t can be written as

$$\min_{u_{t:t+H_p-1}} \max_{\theta_{t:t+H_p-1} \in \mathcal{U}_{t:t+H_p-1}} \sum_{\tau=t}^{t+H_p-1} J\left(x_\tau, u_\tau, \theta_\tau\right) \tag{2.83}$$

subject to the dynamic state transition and feasibility conditions

$$x_{\tau+1} = F(x_\tau, u_\tau, \theta_\tau), \quad u_\tau \in \mathcal{X}_\tau \tag{2.84}$$

Only the first part of the optimized policy is implemented:

$$\pi_t = \left(u_t^\star, u_{t+1}^\star, \dots, u_{t+H_e-1}^\star\right), \quad H_e < H_p \tag{2.85}$$

After the execution interval, the system observes new data, shifts the planning window forward, and repeats the forecasting-optimization-simulation cycle. This mechanism allows the framework to anticipate future demand while avoiding rigid commitment to outdated forecasts.

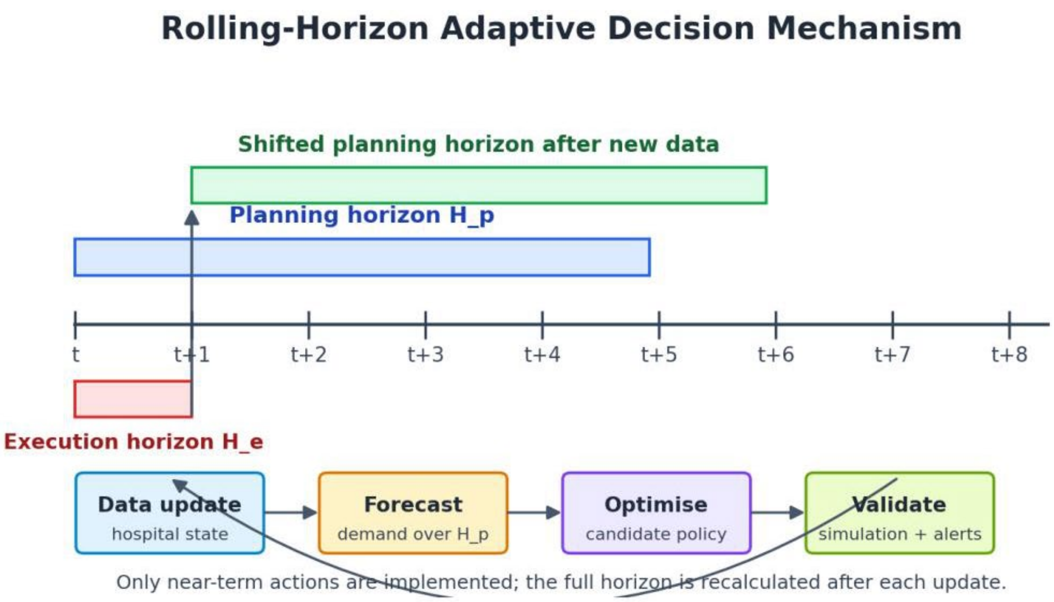

Figure 2.7. Rolling-horizon adaptive decision mechanism

2.9.3 Human-in-the-Loop Governance and Dashboard Outputs

Human review is particularly important because optimization objectives cannot fully encode all ethical and clinical considerations. For instance, a transfer may be mathematically efficient but inappropriate for a clinically unstable patient. A staffing recommendation may be feasible in aggregate but unacceptable because of fatigue or legal working-time limits. Human-in-the-loop governance ensures that model outputs are interpreted within the real institutional context [115], [117].

Dashboard outputs should be concise enough for crisis use. They should not overwhelm decision-makers with every model variable. Instead, they should emphasize actionable information: where demand is expected to exceed capacity, which facilities face ICU risk, which transfers are recommended, which resources are binding, and which safety thresholds are violated. Scenario comparisons can help decision-makers understand the consequences of alternative assumptions.

The governance layer also supports accountability. Recommendations should record the data version, parameter assumptions, objective weights, uncertainty set, solver status, and simulation validation results. This documentation allows later review of why a recommendation was made and whether the assumptions were appropriate.

The IHDSF supports human decision-making rather than automating crisis governance. Mathematical recommendations must remain subject to clinical ethics, legal authority, institutional rules, and real-time judgment by experienced healthcare administrators. The workflow therefore includes a human-in-the-loop stage between simulation validation and operational implementation.

Dashboard outputs may include projected ward and ICU demand, facility-level utilization, shortage risk, transfer recommendations, staff-depletion warnings, safety-threshold violations, and scenario comparisons. These outputs should be ranked by urgency and accompanied by uncertainty indicators. The model identifies feasible, high-performing, and risk-aware options; human authorities determine whether recommendations are clinically appropriate, ethically acceptable, and legally implementable.

2.9.4 Algorithmic Description of the Adaptive Workflow

The algorithm can be implemented with different levels of automation. In a research prototype, several steps may be performed manually, such as data preparation, parameter calibration, and review of simulation outputs. In an operational deployment, these steps could be integrated with dashboards and hospital information systems. The methodological contribution of Chapter 2 does not depend on full automation; it depends on the structured sequence of updating, forecasting, optimization, simulation validation, and human review.

The algorithm is also suitable for comparative evaluation. A static policy, a no-transfer policy, a greedy allocation policy, a deterministic optimization policy, and the full IHDSF policy can all be evaluated under the same scenarios. Such comparisons are important because they answer the practical question of whether the integrated framework performs better than simpler decision rules. This comparative logic is developed in the empirical and computational chapters.

Algorithm 2.2 summarizes the complete adaptive workflow.

Algorithm 2.2. Rolling-horizon adaptive decision-support procedure

1. Observe epidemiological, hospital, staffing, and logistics data at time t.
2. Harmonize data and estimate the current system state $\hat{x}_t$.
3. Update epidemiological and operational parameters $\hat{\theta}_t$.
4. Generate hospital and ICU demand forecasts over planning horizon H_p.
5. Update effective capacity, transfer feasibility, and staff availability.
6. Solve the robust multi-objective optimization problem over H_p.
7. Evaluate the candidate policy through DES replications.
8. Check safety thresholds, overload risk, and implementation feasibility.
9. If thresholds are violated, revise constraints or penalties and re-optimize.
10. If thresholds are satisfied, submit the validated recommendation to decision-makers.
11. Implement only the execution horizon H_e.
12. Observe the realized state, shift the horizon forward, and repeat.

Section 2.9 functions as the operational integration layer of Chapter 2. It connects epidemiological forecasting, resource modelling, robust optimization, and simulation validation into a coherent adaptive decision-support workflow [117], [98].

2.10 Chapter Summary

This chapter developed the proposed Integrated Hybrid Decision-Support Framework for pandemic healthcare system optimization. The framework integrates epidemiological forecasting, healthcare resource modelling, robust multi-objective optimization, discrete-event simulation, and adaptive rolling-horizon feedback into a unified methodological structure [96], [20], [8], [77], [117].

The chapter first formulated pandemic healthcare management as a dynamic decision-support problem under uncertainty. The system was represented through time-dependent epidemiological and operational states, decision variables, uncertainty sets, demand projections, effective capacity, shortage variables, patient transfers, and feasibility constraints. The model scope was defined around regional healthcare coordination, while broader public health policies were treated as exogenous scenario inputs.

The epidemiological component was developed using an extended SEIR-H-ICU-D structure that links infection dynamics with ward and ICU demand. This model provides the predictive layer of the framework by transforming infection trajectories into operational healthcare demand signals. Uncertainty in epidemiological parameters is represented through bounded uncertainty sets, allowing later optimization decisions to be evaluated across plausible scenarios.

The healthcare resource modelling layer represented the hospital network, ward and ICU capacity, staff availability, ventilators, transfer feasibility, patient flows, and load-balancing requirements. Effective capacity was defined as a bottleneck-constrained quantity rather than a nominal bed count, ensuring that the model reflects operational feasibility. Shortage, transfer, staffing, and utilization variables provide the direct bridge between demand forecasting and optimization.

The optimization component formulated pandemic healthcare management as a robust multi-objective decision problem. The objective balances unmet demand, mortality-related risk, transfer burden, utilization imbalance, and surge-resource use. The robust counterpart accounts for uncertainty in epidemiological and operational parameters, while the hybrid solution approach combines exact optimization with heuristic or simulation-assisted search when required.

The simulation layer introduced DES as a validation mechanism for candidate allocation policies. Simulation allows the framework to evaluate stochastic patient arrivals, length of stay, waiting times, queues, transfer delays, overload probabilities, and safety-threshold violations before operational implementation. Simulation feedback can then refine the optimization model through updated penalties, constraints, or parameters.

Finally, the adaptive workflow was formulated as a rolling-horizon mechanism. At each update cycle, new data are incorporated, parameters are recalibrated, demand forecasts are generated, candidate policies are optimized and simulation-tested, and only the near-term part of the policy is implemented. This makes the framework responsive to changing pandemic conditions while preserving forward-looking preparedness.

Overall, Chapter 2 provides the methodological basis for the empirical and computational analysis developed in the following chapter. The next chapter applies the proposed framework to data and scenarios, evaluates the performance of the optimization and simulation components, and discusses the resulting policy implications.

CHAPTER 3: EMPIRICAL DATA COLLECTION, PARAMETERIZATION AND EXPERIMENTAL DESIGN

3.1 Kosovo COVID-19 Case Study and Data Sources

Kosovo is used in this dissertation as the empirical healthcare case study for evaluating pandemic-related forecasting, optimization, and simulation strategies. The case study is relevant because Kosovo represents a resource-constrained healthcare environment in which pandemic pressure, hospital capacity, and regional service availability may differ substantially across municipalities. These conditions make the country a suitable setting for examining how data-driven decision-support models can support healthcare planning during pandemic emergencies.

The empirical focus of this chapter is the construction of a structured dataset that can support the later computational modules of the dissertation. The dataset combines epidemiological, demographic, and healthcare-capacity information at municipality level. The epidemiological component captures reported COVID-19 cases, while the demographic and capacity components provide population size and hospital-bed availability. These variables are required because the dissertation does not analyze infection pressure alone; it evaluates how infection pressure interacts with limited healthcare resources.

The case-study dataset includes the main Kosovo municipalities for which complete tabular information was available: Prishtina, Prizren, Peja, Gjakova, Mitrovica, Ferizaj, and Gjilan. Only municipalities with complete records were included in the experimental dataset. This criterion ensures that the same municipality-level observations can be used consistently across the SEIR-kNN forecasting module, the MILP allocation model, and the DES hospital-operations simulation.

The historical input period used for the municipality-level COVID-19 dataset is January-February 2021. This period provides the empirical basis for deriving infection burden, short-term growth, capacity-pressure indicators, and normalized input matrices. In the subsequent experimental design, these prepared inputs are used for forecasting preparation, hospital-demand estimation, patient-allocation optimization, and simulation-based operational validation. The computational results generated from these inputs are reported in Chapter 4

Table 3.1 presents the municipality-level epidemiological, demographic, and hospital-capacity data used to prepare the Kosovo case study.

Table 3.1. Kosovo municipality-level COVID-19 and hospital-capacity dataset used for experimental preparation

Municipality	Population	Hospital beds	January cases	February cases	Cumulative Jan-Feb cases
Prishtina	210,000	800	1,200	1,500	2,700
Prizren	178,000	600	800	900	1,700
Peja	130,000	400	500	600	1,100
Gjakova	94,000	300	400	450	850
Mitrovica	85,000	250	300	350	650
Ferizaj	110,000	350	500	550	1,050
Gjilan	100,000	300	450	500	950

The variables in Table 3.1 represent both sides of the pandemic healthcare-management problem. Reported COVID-19 cases represent epidemiological pressure, while hospital-bed availability represents the basic capacity of the healthcare system to absorb that pressure.

Population size is included because it allows comparison between municipalities with different demographic scales. The cumulative January-February case count provides a short-term input signal for later forecasting and demand-estimation procedures.

The selected municipalities represent major healthcare demand areas and regional hospital catchments in Kosovo. In later sections of this chapter, these municipality-level observations are transformed into derived epidemiological and healthcare-capacity indicators, normalized input matrices, and parameterized experimental inputs. These prepared data structures are then used by the SEIR-kNN, MILP, and DES modules.

The empirical base is documented through the final thesis bibliography, which consolidates the institutional and technical sources used for data and implementation. These include Kosovo public-health reporting, Ministry of Health reporting, WHO Kosovo COVID-19 material, COVID-19 dashboard or repository data, Johns Hopkins data where applicable, MATLAB documentation, and the SEIR-kNN, MILP, and DES studies underlying the computational modules. [119], [120], [121], [122], [123], [124], [125].

In summary, Section 3.1 defines the Kosovo COVID-19 healthcare case study and identifies the empirical variables required for the dissertation experimental design. The dataset provides the starting point for municipality inclusion, data-completeness verification, derived healthcare indicators, preprocessing, normalization, parameterization, and experimental workflow construction.

3.2 Municipality Inclusion, Data Completeness and Reliability

This section defines how the municipality-level observations described in Section 3.1 are selected, verified, and prepared for use in the dissertation experiments. The purpose of the section is not to repeat the complete dataset table, but to establish a transparent inclusion logic that makes the later forecasting, optimization, and simulation results reproducible. The same municipality-level records must be usable across the SEIR-kNN forecasting module, the MILP allocation model, and the DES hospital-operations simulation. For this reason, data completeness is treated as a prerequisite for experimental consistency rather than as a separate computational result.

The empirical unit of analysis is the municipality-level record. Each record represents one Kosovo municipality for which the required demographic, epidemiological, and healthcare-capacity variables are available in a consistent tabular form. The final experimental dataset contains the seven complete municipality records listed in Table 3.1: Prishtina, Prizren, Peja, Gjakova, Mitrovica, Ferizaj, and Gjilan. These observations form the common input base for the derived indicators, normalization procedures, parameterization steps, and scenario definitions developed in the remaining sections of this chapter.

The main inclusion principle is completeness of the mandatory variables. A municipality was retained in the experimental dataset only when the required values were available for population, hospital-bed capacity, January 2021 reported cases, and February 2021 reported cases. If any of these mandatory variables is absent, the record cannot be used consistently across the computational workflow without introducing imputation or unsupported assumptions. All municipalities included in the experiments satisfied completeness requirements. Only municipalities with complete records were included.

Table 3.2 reports the inclusion requirements and reliability controls used before the dataset is passed to the computational modules. It is intentionally organized by data element rather than by municipality.

The full row-level coverage audit is not repeated in the main text because it would duplicate the information already implied by the inclusion rule. Any detailed audit of incomplete or excluded records belongs in Appendix B.

Table 3.2. Inclusion requirements and reliability controls for the municipality-level dataset

Data element	Completeness requirement	Reliability control	Role in the experimental workflow
Municipality identifier	One unique municipality name per row	Harmonized spelling and no duplicated municipality entries	Links the same observation across derived indicators, normalized matrices, and later model inputs
Population	Available positive numerical value	Checked as the demographic denominator for rate-based indicators	Used for cases per 100,000 inhabitants, beds per 100,000 inhabitants, and normalization
Hospital beds	Available positive numerical value	Checked as the healthcare-capacity proxy for each included municipality	Used for capacity-pressure indicators, demand-capacity comparison, and MILP/DES parameterization
January 2021 cases	Available non-negative numerical value	Checked for period consistency and compatibility with the February value	Used as part of the historical input signal for cumulative cases, growth, and forecasting preparation
February 2021 cases	Available non-negative numerical value	Checked for the same reporting period and municipality definition	Used with January cases to construct cumulative Jan-Feb cases and short-term growth
Common input period	January-February 2021 available for every included municipality	Verified before computing derived indicators and normalized matrices	Ensures that all municipalities enter the experiments under the same temporal window

The completeness assessment is used as a data-quality filter, not as a mathematical model, optimization result, or methodological contribution. For this reason, no completeness-ratio equation is reported in the main text. A formula-based completeness ratio is retained in Appendix A for documentation, while the main chapter states the operational rule used for inclusion: complete records were retained and incomplete records were not used in the main experimental matrix.

3.2.1 Municipality Inclusion Logic

The municipality inclusion procedure follows a conservative design. The objective is to avoid unsupported numerical claims and to ensure that every row used in later computation has the same structure. A municipality was eligible for inclusion when it satisfied four conditions: the municipality identifier was unique, the population value was available, the hospital-bed value was available, and both January and February 2021 case values were available. These conditions are minimal but sufficient for the data-preparation tasks required in this dissertation.

The use of a conservative inclusion rule is appropriate because the later stages of analysis depend on the comparability of the observations. If a municipality had missing hospital capacity, the capacity-pressure indicator could not be computed. If January or February cases were missing, cumulative cases and short-term growth could not be computed reliably. If population were missing, rate-based comparisons across municipalities would not be meaningful. The inclusion rule therefore protects the consistency of the entire experimental chain.

The decision to use seven complete records is a data-availability decision for the main quantitative experiments. Additional municipalities are incorporated when the same mandatory variables are available in complete tabular form. In that case, the dataset size can be expanded and the same preprocessing, normalization, and parameterization workflow can be applied without methodological changes.

3.2.2 Data Completeness Requirements

Data completeness is evaluated at the level of the mandatory variables needed for the experimental workflow. The required variables correspond to the minimum information needed to connect the three analytical layers of the dissertation: epidemiological forecasting, capacity-aware optimization, and operational simulation. The population variable supports normalization and demographic comparison. Reported case counts provide the empirical signal for infection pressure. Hospital-bed capacity provides the baseline healthcare-resource constraint. Together, these variables make it possible to transform the Kosovo case study into a structured experimental dataset.

Completeness is checked before computing derived indicators. Derived measures such as cases per 100,000 inhabitants, beds per 100,000 inhabitants, capacity pressure, and growth rate are only meaningful when their underlying variables are available and internally consistent. Therefore, all derived indicators in the subsequent sections are based on complete municipality records only.

The completeness check also prevents a mismatch between textual interpretation and numerical evidence. A municipality may appear in a preliminary graphical reference, descriptive discussion, or broader case-study context, but it is not entered into a quantitative table or MATLAB input matrix unless the required tabular values are available. This distinction avoids overclaiming and ensures that every number reported in the experimental sections can be traced back to a complete input record.

3.2.3 Data Reliability and Internal Consistency Controls

The reliability controls applied in this section are internal consistency controls. They do not claim that epidemiological reporting is free from the limitations common during public-health emergencies. Instead, they ensure that the dataset used for modeling is coherent, reproducible, and suitable for the computational procedures developed in this dissertation.

Municipality names are treated as linkage keys. Each included municipality appears once and is written consistently across the dataset, derived-indicator table, normalized input matrix, and later model inputs. Numerical variables must have valid ranges: population and hospital-bed values must be positive, while reported case values must be non-negative. The epidemiological variables must also refer to the same input period across all included municipalities. The January-February 2021 window is therefore used consistently as the historical input period for the prepared dataset.

Finally, the same municipality set is maintained across the data-preparation stage and the parameterization stage. This ensures that SEIR-kNN inputs, demand-conversion parameters, MILP capacity constraints, and DES scenario assumptions are based on a common empirical foundation.

3.2.4 Treatment of Incomplete Records

Incomplete records are not used in the main quantitative matrix. This decision is made before model execution and is applied consistently across all subsequent computations. The purpose is to avoid mixing complete numerical observations with partially documented municipalities. Such mixing would be problematic because the forecasting, optimization, and simulation modules require structured inputs with the same variable definitions for each observation.

Only municipalities with complete records were included. The detailed explanation of any incomplete municipality, including the specific missing fields and the experimental decision, belongs in Appendix B rather than in the main chapter. This placement preserves transparency while avoiding unnecessary repetition in the dissertation body.

The methodology does not depend on the number seven as a fixed assumption. Additional complete municipality records can be added to the dataset and processed through the same workflow. The number seven reflects the complete records available for the current tabular dataset.

3.2.5 Implications for the Experimental Design

The inclusion and completeness rules defined in this section directly support the experimental design of the dissertation. They ensure that the same empirical observations can be passed from data preparation to parameterization and then to the computational models. This consistency is especially important because the dissertation evaluates an integrated forecasting-optimization-simulation workflow rather than a single isolated model.

For the SEIR-kNN module, the complete dataset provides the municipality-level feature vectors required for similarity-based parameter adaptation. For the MILP module, the same dataset provides the demographic and capacity basis for demand-capacity reasoning and hospital allocation scenarios. For the DES module, it provides the empirical context for patient-flow scenarios, ICU pressure, and operational stress assumptions.

The reliability framework also defines the boundary of interpretation. The results derived from the prepared dataset are interpreted as outputs of a structured Kosovo municipality-level case study, not as unrestricted claims for every possible hospital or patient-level situation in Kosovo. The dissertation therefore extracts the maximum consistent information from complete records while keeping incomplete or unsupported material outside the main experimental matrix.

3.3 Derived Epidemiological and Healthcare-Capacity Indicators

The municipality-level dataset defined in Sections 3.1 and 3.2 provides the raw empirical basis for the experimental design. However, raw population counts, reported case numbers, and hospital-bed values are not directly comparable across municipalities because the municipalities differ in demographic size and healthcare-capacity structure. For this reason, a set of derived epidemiological and healthcare-capacity indicators was constructed before the forecasting, optimization, and simulation modules were parameterized.

The purpose of this section is to transform the verified Kosovo municipality-level dataset into comparable analytical variables. The indicators are not intended to create a separate categorical clinical grouping model. Instead, they provide reproducible empirical signals that connect infection burden, hospital-capacity availability, and short-term case growth. These signals are later used to support SEIR-kNN feature construction, MILP demand-capacity preparation, and DES scenario design.

3.3.1 Indicator Construction and Analytical Role

Four indicators were retained for the main chapter: cumulative cases per 100,000 inhabitants, hospital beds per 100,000 inhabitants, capacity-pressure proxy, and January-February growth rate. These indicators were selected because they are directly derived from the complete dataset and because they support the dissertation workflow without introducing unsupported categorical thresholds.

The indicators were computed using standard epidemiological and healthcare-capacity transformations. To keep the notation concise, let C_m denote cumulative January-February cases, P_m denote population, B_m denote hospital beds, J_m denote January cases, and F_m denote February cases for municipality m. The retained indicators are defined as follows:

$$Cases\ per\ 100{,}000_m = \frac{C_m}{P_m} \times 100{,}000 \tag{3.1}$$

$$Beds\ per\ 100{,}000_m = \frac{B_m}{P_m} \times 100{,}000 \tag{3.2}$$

$$\text{Capacity} - \text{pressure proxy}_m = \frac{C_m}{B_m} \tag{3.3}$$

$$\text{Jan} - \text{Feb growth rate}_m = \frac{F_m - J_m}{J_m} \times 100 \tag{3.4}$$

Here, m denotes the municipality included in the complete experimental dataset. The first indicator standardizes infection burden by population size. The second indicator standardizes healthcare-capacity availability by population size. The third indicator relates the cumulative infection signal to available hospital beds and is therefore used as a simple capacity-pressure proxy. The fourth indicator captures the short-term percentage change in reported cases during the January-February 2021 historical input period.

No categorical threshold groups are assigned in this section. Fixed categories such as low, moderate, and high would require explicit external justification, for example through peer-reviewed literature, official guidance, quantile grouping, or clustering results. The capacity-pressure proxy is therefore used as a continuous empirical indicator, and the section reports measurable values only.

3.3.2 Municipality-Level Indicator Profile

Table 3.3 reports the derived epidemiological and healthcare-capacity indicators calculated for the municipalities included in the experimental dataset.

Table 3.3. Derived epidemiological and healthcare-capacity indicators for the Kosovo municipality dataset

Municipality	Cases per 100,000	Beds per 100,000	Capacity-pressure proxy	Jan-Feb growth rate
Prishtina	1285.71	380.95	3.38	25.00%
Prizren	955.06	337.08	2.83	12.50%
Peja	846.15	307.69	2.75	20.00%
Gjakova	904.26	319.15	2.83	12.50%
Mitrovica	764.71	294.12	2.60	16.67%
Ferizaj	954.55	318.18	3.00	10.00%
Gjilan	950.00	300.00	3.17	11.11%

The indicator profile shows that healthcare pressure is not explained by absolute case numbers alone. Prishtina and Gjilan exhibited the highest pressure indicators, suggesting stronger healthcare stress relative to available capacity. Ferizaj also shows relevant pressure when cumulative cases are interpreted against the available hospital-bed base. This supports the use of capacity-aware indicators instead of relying only on raw reported cases.

The comparison of cases and beds per 100,000 inhabitants further confirms the need for standardized indicators. A municipality may have fewer absolute cases than the largest municipality but still face relevant operational pressure if its hospital-capacity base is proportionally smaller. This distinction is important for later experimental modules, because forecasting, allocation, and simulation require capacity-aware inputs rather than national averages or raw counts alone.

Figure 3.1 visualizes the capacity-pressure proxy, while Figure 3.2 provides the standardized comparison between infection burden and healthcare-capacity availability.

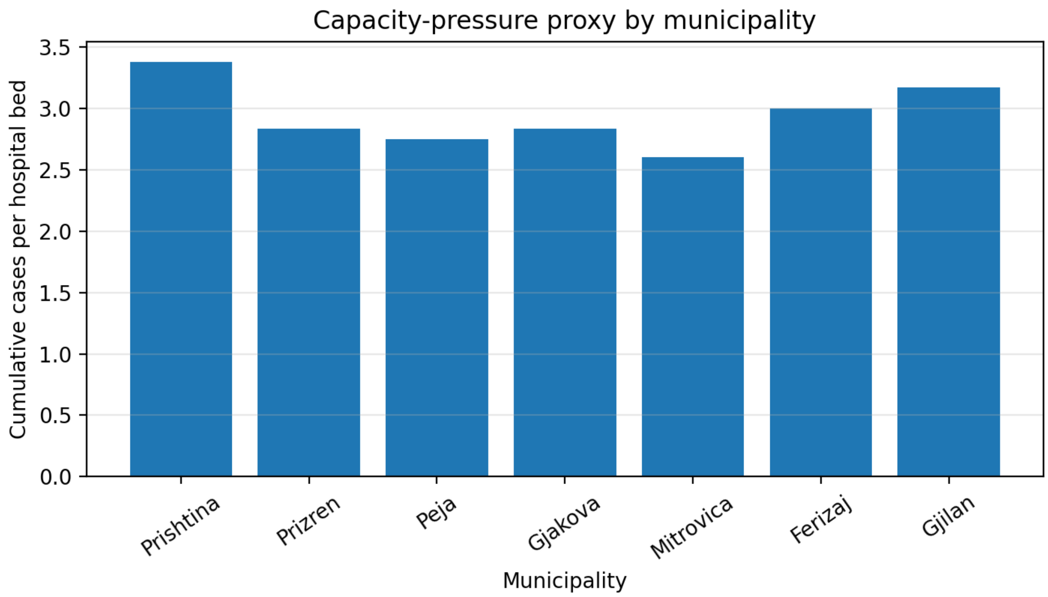

Figure 3.1. Capacity-pressure proxy by municipality

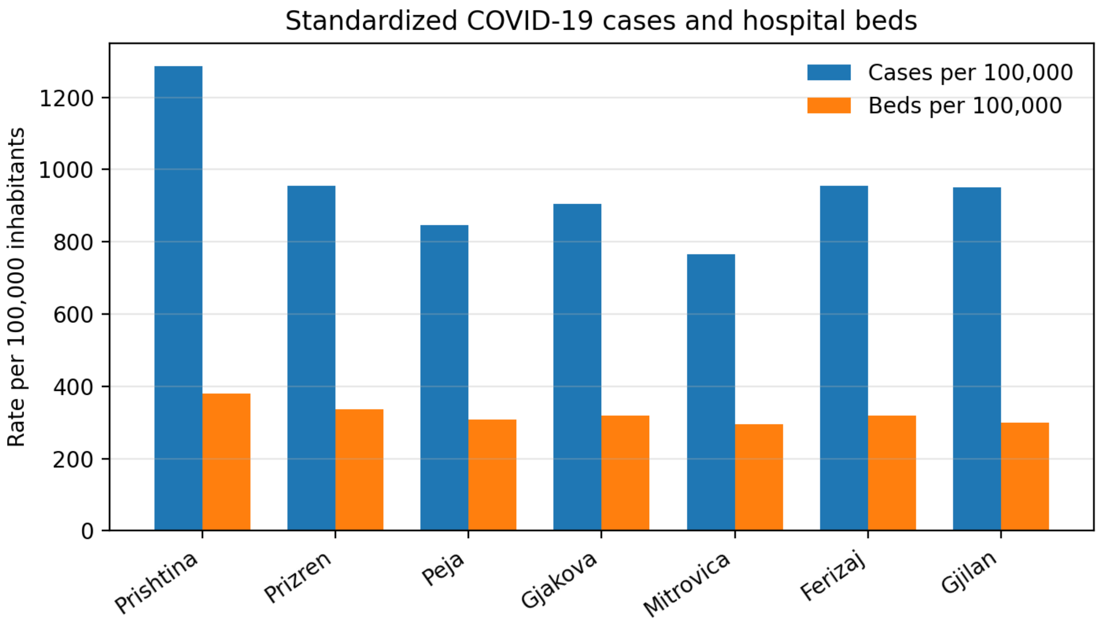

Figure 3.2. Standardized comparison of COVID-19 cases and hospital beds per 100,000 inhabitants

A separate hospital-bed-only chart is not included because hospital-bed values are already reported in Table 3.1 and would duplicate the same information. The section therefore keeps the two figures that add interpretive value: the pressure proxy and the standardized cases-versus-capacity comparison.

3.3.3 Role of the Indicators in the Experimental Design

The derived indicators provide the bridge between empirical data preparation and computational experimentation. In the SEIR-kNN component, the standardized epidemiological and healthcare-capacity variables support municipality-level feature construction and parameter adaptation. In the MILP component, the indicators help prepare demand and capacity inputs for allocation analysis. In the DES component, they support the design of operational scenarios in which patient-flow pressure is evaluated under limited healthcare resources.

The MATLAB implementation used to generate the indicator values and visualizations is not included in the main text. To keep the chapter focused on data preparation and experimental design, the full script and reproducibility materials are retained in Appendix B. This preserves transparency without interrupting the main academic narrative with implementation code.

In summary, Section 3.3 transforms the verified Kosovo municipality-level dataset into derived epidemiological and healthcare-capacity indicators suitable for parameterization and experimental design. The section uses the essential indicator table and the two most informative figures, applies continuous indicators, avoids unnecessary repetition, and prepares the dataset for the preprocessing, normalization, and MATLAB input-preparation steps presented in Section 3.4.

3.4 Data Preprocessing, Normalization and MATLAB Input Preparation

The preceding sections defined the Kosovo COVID-19 healthcare case study, the municipality-inclusion logic, and the derived epidemiological and healthcare-capacity indicators. The next step is to convert these empirical data into a structured computational input layer. This section therefore describes how the dataset was cleaned, normalized, summarized, and prepared as MATLAB input for the SEIR-kNN, MILP, and DES components of the dissertation.

This section is intentionally positioned in Chapter 3 because preprocessing and input-matrix preparation are part of the empirical data-design layer rather than the results chapter. The numerical outputs produced from these prepared inputs are reported later in Chapter 4. The role of Section 3.4 is to ensure that the computational experiments are based on the same verified dataset and that each module receives consistent inputs.

The preprocessing stage follows three principles. First, the same cleaned dataset is used across all computational modules. Second, numerical variables are standardized only where scale comparability is required, especially for the kNN component. Third, full MATLAB code, standard transformation formulas, detailed normalized matrices, and file-level exports are kept in Appendix B, while the main text reports only the methodological logic and the compact descriptive summary needed for doctoral-level transparency.

3.4.1 Data Cleaning and Feature Construction

Data cleaning was performed after the municipality-inclusion step described in Section 3.2. Because only complete municipality records are retained in the experimental dataset, no imputation is applied in the main analysis. This choice avoids introducing synthetic values into a small municipality-level dataset and ensures that forecasting, allocation, and simulation inputs remain traceable to observed data.

The cleaned dataset is organized at municipality level. Each row represents one municipality included in the complete experimental dataset, and each column represents either a raw variable or a derived feature required by the experimental workflow. The raw variables are population, hospital beds, January COVID-19 cases, February COVID-19 cases, and cumulative January-February cases. The constructed indicators are those defined in Section 3.3: cases per 100,000 inhabitants, beds per 100,000 inhabitants, capacity-pressure proxy, and January-February growth rate.

The feature-construction step has two roles. It improves comparability across municipalities with different demographic sizes, and it prepares capacity-aware signals for the later modules. The SEIR-kNN module uses the standardized feature profile for similarity-based parameter adaptation. The MILP module uses demand and capacity variables for allocation constraints. The DES module uses the prepared demand, capacity, and scenario inputs to represent patient-flow pressure under alternative operating conditions.

Table 3.4 summarizes the cleaning and feature-construction logic without reproducing the full MATLAB implementation. This avoids duplicating the same information in text, code, and tables.

Table 3.4. Data-cleaning and feature-construction procedures used before experimental modeling

Data element	Cleaning or construction action	Reason for inclusion	Role in later experiments
Municipality identifier	Names harmonized and kept in the same row order across all input structures	Prevents mismatch between municipal records and model matrices	Index for SEIR-kNN, MILP, and DES inputs
Population	Checked as a positive numerical value and retained without normalization for raw demographic interpretation	Defines population scale and supports per-capita indicators	SEIR-kNN scaling and demand normalization
Hospital beds	Checked as a positive capacity value and retained as the basic healthcare-capacity variable	Represents available local capacity before stress assumptions	MILP capacity vector and DES scenario capacity
January and February COVID-19 cases	Checked as non-negative numerical values and used to construct cumulative cases	Defines the historical input period used for experimental preparation	SEIR-kNN historical input and demand-estimation basis
Derived indicators	Cases per 100,000, beds per 100,000, capacity pressure, and growth rate generated from the raw variables	Transforms raw counts into comparable analytical variables	Parameterization, scenario design, and capacity-pressure interpretation
Incomplete records	Excluded from the main experimental matrix rather than imputed	Protects consistency and avoids unsupported synthetic observations	Ensures all included municipalities satisfy completeness requirements

The complete MATLAB script that performs these operations is provided in Appendix B.

3.4.2 Normalization and Descriptive Statistics

Normalization was required because the variables in the dataset are measured on different numerical scales. Population values are substantially larger than monthly case counts and hospital-bed counts, while derived variables such as capacity pressure and growth rate are relative measures. If all variables were used without standardization in a distance-based method, larger-scale variables would dominate similarity calculations and reduce the effect of healthcare-capacity and epidemiological-pressure indicators.

Standard normalization and descriptive measures were therefore used to prepare the municipality-level input matrix. The main text does not repeat the formulas for min-max normalization, z-score standardization, or coefficient of variation, because these are standard preprocessing expressions rather than methodological contributions. Their definitions and the exact MATLAB implementation are retained in Appendix B. In the main chapter, the focus is on why the transformations were needed and how the resulting inputs are used by the experimental modules.

Descriptive statistics were calculated as a quality-control step before the dataset was transferred to the computational models. The table is shortened to the three most useful quantities: mean, standard deviation, and coefficient of variation. This preserves the key information about scale and relative dispersion while avoiding an oversized table in the main text.

Table 3.5 summarizes the descriptive statistics used to verify the scale, dispersion, and preprocessing consistency of the experimental variables.

Table 3.5. Simplified descriptive statistics used for preprocessing control

Variable	Mean	Standard deviation	CV
Population	129,571.43	47,123.44	0.364
Hospital beds	428.57	199.70	0.466
January cases	592.86	308.80	0.521
February cases	692.86	395.21	0.570
Cumulative cases	1,285.71	703.39	0.547
Cases per 100,000	951.49	163.27	0.172
Beds per 100,000	322.45	29.41	0.091
Capacity pressure	2.94	0.26	0.090
Growth rate	15.40	5.47	0.355

Source: Prepared from the Kosovo municipality-level experimental dataset. Expanded descriptive statistics are retained in Appendix B for reproducibility.

The simplified statistics confirm that the dataset is heterogeneous. Case-count variables show stronger relative dispersion than the standardized capacity-pressure indicators, while beds per 100,000 inhabitants and capacity pressure are comparatively more compact. This supports the use of municipality-specific parameterization rather than a single national average. It also justifies normalization before kNN-based similarity comparison.

The normalized matrix is particularly important for the SEIR-kNN component. Without normalization, population would dominate the distance calculation and reduce the contribution of hospital beds and case counts. For this reason, the numerical normalized matrix is moved to Appendix B, while the main chapter keeps a single visual representation of the standardized input matrix.

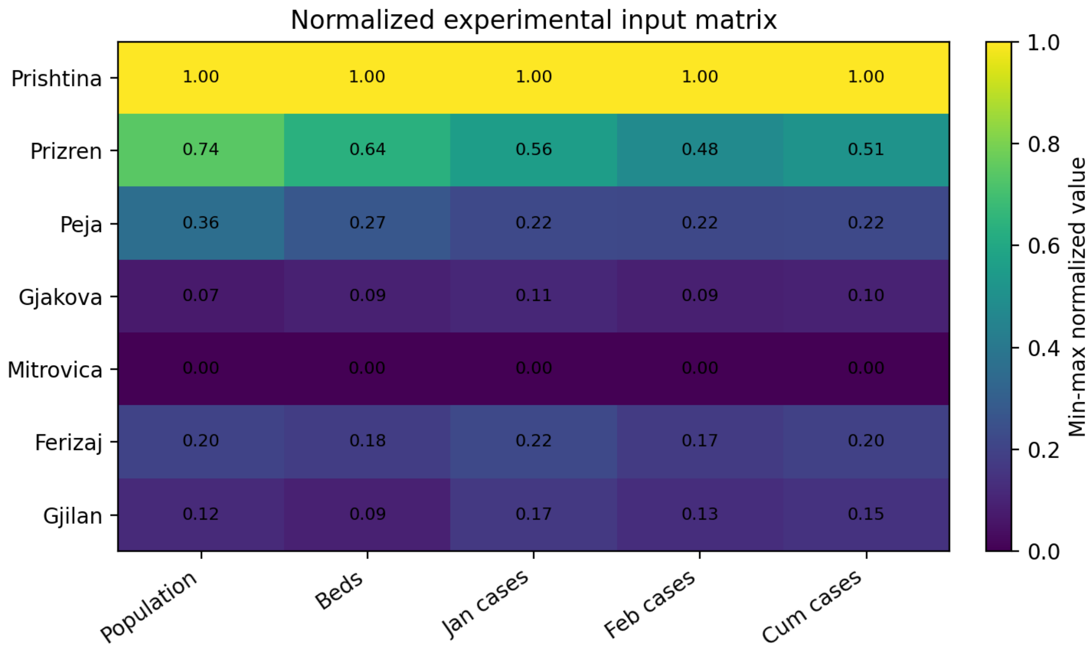

Figure 3.3. Normalized experimental input matrix used for SEIR-kNN parameter adaptation

3.4.3 MATLAB Input Matrix Preparation

After cleaning, feature construction, and normalization, the dataset was organized into MATLAB input structures required by the dissertation experiments. The SEIR-kNN module requires a normalized municipality-feature matrix. The MILP module requires demand and capacity inputs, together with allocation-related structures. The DES module requires operational parameters and scenario assumptions derived from the same preprocessed data layer.

The key requirement is consistency across modules. Forecasting, optimization, and simulation are not calibrated from disconnected datasets. All forecasting, optimization, and simulation experiments therefore use a common preprocessed dataset, ensuring reproducibility and consistency across analytical modules. This allows forecasts, demand signals, capacity constraints, and simulation scenarios to remain connected to the same empirical source.

Table 3.6. Analytical input structures generated from the preprocessed dataset

Input structure	Main contents	Used by module	Purpose in experimental design
Normalized municipality-feature matrix	Standardized population, beds, January cases, February cases, and cumulative cases	SEIR-kNN	Supports municipality similarity comparison and adaptive parameterization
Derived indicator matrix	Cases per 100,000, beds per 100,000, capacity pressure, and growth rate	SEIR-kNN, MILP, DES	Links epidemiological burden with healthcare-capacity pressure
Demand-preparation vector	Observed or forecast-derived demand signals by municipality or hospital node	MILP	Provides demand inputs for allocation and feasibility checks
Capacity-preparation vector	Available bed-capacity information and scenario-adjusted capacity assumptions	MILP, DES	Defines capacity constraints and operational limits
Scenario-parameter structure	Baseline, stress, and optimized configuration assumptions	DES and integrated evaluation	Supports simulation-based operational validation and robustness analysis

Detailed file names, MATLAB workspace objects, and output-export tables are not listed in the main text. They are retained in Appendix B as reproducibility resources. In the main chapter, the reproducibility statement is summarized as follows: experimental data, preprocessing outputs, and MATLAB workspace files were generated automatically and stored in reproducible formats.

Algorithm 3.1. Experimental preprocessing workflow

Step	Operation
1	Load and verify municipality-level COVID-19 and healthcare-capacity data.
2	Generate epidemiological and healthcare-capacity indicators.
3	Normalize selected variables and compute simplified descriptive statistics.
4	Export standardized MATLAB matrices and reproducibility materials.
5	Pass prepared outputs to the SEIR-kNN, MILP, and DES experimental modules.

MATLAB Script 3.1 implemented preprocessing, indicator generation, and export procedures. The script itself is not included in the main text because code listings are implementation details rather than dissertation results. Full code is provided in Appendix B. This keeps the chapter readable while preserving complete reproducibility.

In summary, Section 3.4 converts the verified Kosovo municipality-level dataset into a computational input layer for the dissertation experiments. It keeps data cleaning, normalization, descriptive control, and input-matrix preparation in Chapter 3, while transferring standard formulas, full code, detailed normalized matrices, and file-level exports to Appendix B. The resulting preprocessed dataset forms the common empirical basis for the parameterization of the SEIR-kNN, MILP, and DES modules in Section 3.5.

3.5 Parameterization of SEIR-kNN, MILP, and DES Experiments

After the Kosovo municipality-level dataset has been verified, transformed into derived epidemiological and healthcare-capacity indicators, and standardized into MATLAB-ready input structures, the next step is to define the experimental parameterization of the three computational modules used in the dissertation: SEIR-kNN forecasting, MILP-based hospital allocation, and DES-based hospital operations simulation. The purpose of this section is to specify the empirical inputs, parameter ranges, scenario assumptions, and cross-module linkages required for reproducible experimentation.

This section is written as part of Chapter 3, which functions as the data, parameterization, and experimental-design chapter. It is not a results section. Therefore, it does not report forecasting errors, allocation matrices, transfer-cost comparisons, waiting-time reductions, throughput improvements, p-values, confidence intervals, or integrated KPI findings. Those numerical outputs belong to Chapter 4, where the computational results are analyzed.

The mathematical foundations of the SEIR, MILP, and DES components are not repeated here because they belong to the methodology chapter. Similarly, MATLAB code, solver workflows, file names, event logs, and detailed reproducibility outputs are retained in Appendix B. The main text records only the parameter settings and experimental assumptions necessary for understanding how the later results are generated.

A common parameterization logic is maintained across all modules. The same preprocessed Kosovo dataset is used for forecasting, the forecast-derived demand signals are used for optimization, and the allocation and capacity assumptions are then used for simulation-based operational validation. This prevents the dissertation experiments from becoming disconnected model demonstrations and keeps the framework aligned with the integrated SEIR-kNN-MILP-DES decision-support design.

Table 3.7. Cross-module parameterization logic used in the experimental design

Computational module	Empirical input from Chapter 3	Main parameterization task	Output prepared for Chapter 4
SEIR-kNN forecasting	Complete municipality dataset, normalized features, and derived epidemiological/capacity indicators	Define epidemiological parameter ranges, kNN adaptation settings, prediction horizon, and demand-conversion assumptions	Forecast outputs, validation inputs, and hospital/ICU demand signals
MILP hospital allocation	Hospital demand signals, capacity vectors, hospital-node interpretation, and transfer-cost assumptions	Define demand-capacity matrices, transfer policies, unmet-demand variables, surge-capacity variables, and stress scenarios	Allocation scenarios, transfer-network inputs, feasibility diagnosis, and sensitivity settings
DES hospital operations simulation	Arrival assumptions, ICU and ward capacities, LOS assumptions, staffing assumptions, and scenario inputs	Define patient-flow parameters, scenario configurations, operational KPIs, and replication logic	Waiting-time, throughput, ICU-utilization, and unmet-demand evaluation design

Source: Author's experimental design based on the prepared Kosovo municipality-level dataset and the SEIR-kNN, MILP, and DES modules developed in the dissertation. Expanded implementation resources are retained in Appendix B.

3.5.1 SEIR-kNN Experimental Parameterization

The SEIR-kNN module provides the predictive layer of the dissertation. Its role is to transform the prepared Kosovo municipality-level dataset into short-term infection and demand signals that can later be used by the allocation and simulation modules. The forecasting experiment uses the SEIR formulation introduced in Chapter 2 with municipality-specific parameter adaptation; therefore, the SEIR equations are not repeated in this section.

The historical input period for the SEIR-kNN experiment is January-February 2021, while the prediction and validation period is March-April 2021. The municipality set is restricted to the seven complete records defined earlier in this chapter: Prishtina, Prizren, Peja, Gjakova, Mitrovica, Ferizaj, and Gjilan. The normalized feature matrix prepared in Section 3.4 is used for kNN-based similarity comparison and municipality-specific parameter adaptation.

The epidemiological parameter ranges follow the SEIR-kNN study used in the dissertation. The transmission parameter beta is searched in the unscaled interval 1.5-2.2 and converted to the effective daily transmission-rate scale used in the ODE simulation. The resulting effective municipality-level beta values are expected to fall approximately in the range 0.180-0.230 after scaling and correction. The exposed-to-infectious transition parameter sigma is evaluated in the interval 0.18-0.22, while the recovery/removal parameter gamma is evaluated in the interval

0.09-0.14. The number of nearest neighbors k is treated as an experimental hyperparameter and evaluated through sensitivity analysis rather than assumed arbitrarily.

Table 3.8 presents the parameter values and assumptions used in the SEIR-kNN experiment.

Table 3.8. SEIR-kNN experimental parameterization profile

Parameter group	Setting used in experimental design	Reason for inclusion	Result location
Municipality set	Seven complete municipalities: Prishtina, Prizren, Peja, Gjakova, Mitrovica, Ferizaj, and Gjilan	Ensures that forecasting uses consistent and complete records	Chapter 4, Section 4.2
Historical input period	January-February 2021	Provides the empirical basis for feature construction and parameter adaptation	Chapter 4 forecasting setup
Prediction/validation period	March-April 2021; 60-day forecasting horizon	Supports short-term pandemic demand estimation	Chapter 4 forecasting results
Feature basis	Normalized population, hospital beds, January cases, February cases, cumulative cases, and derived pressure indicators	Prevents scale dominance in distance-based kNN adaptation	Appendix B includes expanded matrices
Transmission parameter beta	Unscaled search interval: 1.5-2.2; effective daily rate after ODE scaling: approximately 0.180-0.230	Controls infection-spread intensity under municipality-specific adaptation	Numerical outputs in Chapter 4
Progression parameter sigma	0.18-0.22	Controls exposed-to-infectious transition	Numerical outputs in Chapter 4
Recovery/removal parameter gamma	0.09-0.14	Controls recovery/removal dynamics	Numerical outputs in Chapter 4
kNN hyperparameter k	Candidate values: k = 3, 5, 7, and 9	Tests the balance between local adaptation and smoothing	Sensitivity results in Chapter 4
Hospital-demand conversion	Baseline hospitalization signal: 15% of predicted cases; stress demand signal: 30% of predicted cases	Converts forecasted infections into operational hospital-demand inputs	Demand results in Chapter 4
ICU-demand assumption	Stress ICU signal defined as 12% of the stress hospital-demand signal	Prepares ICU demand for DES and stress testing	Operational results in Chapter 4

Source: Prepared from the SEIR-kNN experimental setup in the dissertation experimental setup. Full code, validation vectors, and exported MATLAB files are retained in Appendix B.

This parameterization makes the forecasting experiment reproducible without treating forecast outputs as findings in Chapter 3. The final forecast trajectories, error metrics, k-sensitivity results, and forecast-derived hospital and ICU demand outputs are reported later in Chapter 4.

3.5.2 MILP Experimental Parameterization

The MILP module provides the prescriptive allocation layer of the dissertation. Its purpose is to transform forecast-derived hospital demand into structured patient-allocation and hospital-coordination scenarios. The full mathematical formulation of the MILP model is not repeated in this section. Instead, this subsection defines the experimental inputs required to run the allocation scenarios: hospital-node interpretation, demand vectors, capacity vectors, transfer-cost inputs, policy variants, shortage variables, and sensitivity settings.

The MILP experiment uses a reduced five-node hospital network to represent the main operational hospital catchment areas relevant to the Kosovo case study. This reduced network is not intended to represent every healthcare facility in Kosovo. It is an experimental structure for testing allocation logic under controlled capacity and demand conditions, and it can be expanded when complete hospital-level demand, capacity, ICU, transfer-time, ambulance, and staffing data become available.

Table 3.9. MILP experimental parameterization profile

Parameter group	Setting used in experimental design	Experimental role	Result location
Hospital-network nodes	H1 = Prishtina/QKUK H2 = Prizren; H3 = Peja; H4 = Gjilan/Ferizaj regional node; H5 = Mitrovica/Gjakova regional node	Defines the reduced hospital-allocation network for the Kosovo case study	Chapter 4, Section 4.3
Capacity-sufficient demand vector	$D = [80, 20, 10, 50, 60]$	Tests whether coordination can remove local overload when total demand and total capacity are balanced	Chapter 4 allocation comparison
Capacity-sufficient capacity vector	$C = [50, 40, 30, 40, 60]$	Defines available receiving capacity in the reduced five-hospital scenario	Chapter 4 allocation comparison
Peak-demand stress vector	$D_{peak} = [450, 225, 150, 300, 375]$	Represents severe demand pressure where total demand exceeds total capacity	Chapter 4 peak-demand feasibility diagnosis

Peak-capacity vector	$C_{peak} = [300, 200, 150, 200, 250]$	Defines the stress-scenario capacity profile	Chapter 4 peak-demand feasibility diagnosis
Transfer-cost input	Combined transfer-cost matrix based on distance, traffic delay, and road-quality penalty	Allows comparison between local-priority and global-cost allocation logic	Expanded matrix in Appendix B; results in Chapter 4
Policy variants	No coordination; greedy nearest-available; MILP local-priority; MILP global-cost	Provides comparison between uncoordinated, heuristic, and optimized allocation strategies	Chapter 4 comparative results
Unmet-demand variable	Activated under capacity shortage or infeasible allocation conditions	Prevents the model from falsely claiming full coverage when total capacity is insufficient	Chapter 4 feasibility results
Surge-capacity variable	Activated when additional bed-equivalent capacity is required	Quantifies emergency expansion needs under peak demand	Chapter 4 stress analysis
Sensitivity settings	Demand multipliers and capacity-change scenarios, including moderate surge, severe surge, capacity loss, and recovery/surge-capacity addition	Tests robustness of unmet demand under stress conditions	Chapter 4 sensitivity analysis

Source: Prepared from the MILP allocation setup in the dissertation experimental setup. The full transfer-cost matrix, MATLAB solver implementation, solver diagnostics, and reproducibility files are retained in Appendix B.

The use of both capacity-sufficient and peak-demand parameter settings is essential. A capacity-sufficient scenario tests whether coordination can eliminate avoidable local overload when unused capacity exists elsewhere. A peak-demand stress scenario tests whether the model correctly identifies unavoidable shortages when total demand exceeds total capacity. This distinction prevents overclaiming and makes the later allocation results easier to interpret.

The MILP parameterization also separates experimental design from result interpretation. This section defines the demand vectors, capacity vectors, allocation variants, and sensitivity assumptions. The allocation matrix, transfer network, objective values, unmet demand, and comparison between allocation strategies are reported in Chapter 4.

3.5.3 DES Experimental Parameterization

The DES module provides the operational simulation layer of the dissertation. While SEIR-kNN produces demand signals and MILP prescribes allocation decisions, DES evaluates whether those decisions remain operationally feasible when patients arrive stochastically, ICU resources become congested, queues form, and staff capacity constrains throughput. In Chapter 3, the DES component is parameterized as an experimental design; the actual scenario outcomes are reported in Chapter 4.

The DES parameterization uses aggregate operational assumptions from the Kosovo hospital-operations simulation study. The parameters define patient arrivals, care pathways, ward and ICU lengths of stay, ICU capacity, staffing availability, simulation scenarios, and operational performance indicators. Patient-level event logs, Monte Carlo replication tables, confidence intervals, and statistical-test outputs are not included in the main text unless they are supported by final MATLAB reproducibility records. These materials are retained in Appendix B as reproducibility evidence.

Table 3.10. DES experimental parameterization profile

Parameter group	Setting used in experimental design	Experimental role	Result location
Arrival process	Average daily arrivals: 180 patients/day; stochastic arrivals represented through a random-arrival process	Defines patient inflow intensity under pandemic pressure	Chapter 4 DES results
Patient severity classes	Mild, moderate, and severe patient categories	Links triage outcomes to ward or ICU service requirements	Chapter 4 operational analysis
Base ICU capacity	15 ICU beds in the base configuration	Defines the main critical-care capacity constraint	Chapter 4 scenario comparison
Ward length of stay	3-5 days	Defines general-ward service duration assumptions	Appendix B for expanded DES inputs
ICU length of stay	7-10 days	Defines critical-care resource occupancy duration	Appendix B for expanded DES inputs
Staffing assumption	8 doctors and 12 nurses per shift	Represents staff capacity and service bottleneck assumptions	Chapter 4 operational interpretation
Scenario A: baseline	Moderate pressure under base capacity and staffing assumptions	Provides reference condition for operational comparison	Chapter 4 DES results
Scenario B: ICU overflow	Severe surge where ICU demand exceeds available capacity	Tests bottleneck behavior and queue formation under stress	Chapter 4 DES results
Scenario C: optimized configuration	10% ICU capacity increase combined with staff reallocation	Tests the intervention policy under improved operating conditions	Chapter 4 DES results

Operational KPIs	ICU utilization, average waiting time, daily throughput, unmet ICU demand	Defines the performance indicators used to evaluate hospital operations	Chapter 4 KPI dashboard and improvement table
Replication and statistical reporting	Monte Carlo replication and statistical testing retained only when MATLAB replication datasets and reproducibility records are available	Avoids presenting unsupported p-values and confidence intervals in the main text	Appendix B for reproducibility; Chapter 4 for supported results

Source: Prepared from the DES experimental setup in the dissertation experimental setup. Detailed event logic, entity tables, event logs, and scripts are retained in Appendix B.

This DES parameterization is sufficient for the Chapter 3 experimental design because it defines the hospital-flow assumptions before the computational results are reported. The section does not claim waiting-time reductions, throughput increases, ICU-utilization values, p-values, or confidence intervals. Those outputs are treated as results and are therefore reported in Chapter 4 after the scenario execution and validation logic are presented.

Across the three modules, the parameterization strategy preserves a consistent evidence chain. The preprocessed Kosovo dataset provides standardized inputs for SEIR-kNN forecasting. The forecast-derived demand signals define the demand side of the MILP allocation experiment. The resulting demand, allocation, and capacity assumptions define the DES patient-flow scenarios. This design allows Chapter 4 to evaluate results in a coherent sequence: forecasting performance, hospital allocation, operational simulation, and integrated assessment.

In summary, Section 3.5 specifies the experimental parameterization required to execute the SEIR-kNN, MILP, and DES modules. It defines the key input periods, parameter ranges, demand assumptions, capacity vectors, hospital-network scenarios, patient-flow parameters, and operational KPIs. By keeping equations, code, output files, and numerical results outside this section, the chapter remains aligned with its defined purpose: empirical data preparation, parameterization, and experimental design.

3.6 Experimental Design, Scenarios, and Reproducibility Setup

Sections 3.1-3.5 established the empirical and computational preparation layer of the dissertation. The Kosovo municipality-level dataset was defined, complete records were selected, epidemiological and healthcare-capacity indicators were derived, preprocessing and normalization were completed, and the SEIR-kNN, MILP, and DES modules were parameterized. The purpose of the present section is to connect these elements into one coherent experimental design.

This section does not report computational results. Instead, it defines how the prepared dataset, parameter settings, and scenario assumptions are passed through the integrated forecasting-optimization-simulation workflow. The numerical outputs generated from this workflow, including forecasting accuracy, allocation matrices, transfer-cost comparisons, waiting-time reductions, throughput changes, unmet-demand indicators, and integrated KPI results, are presented in Chapter 4.

The experimental design follows three principles. First, all computational modules use the same verified empirical data foundation. Second, each module has a clearly defined role within the dissertation workflow: SEIR-kNN produces demand-oriented forecasts, MILP converts demand and capacity information into allocation scenarios, and DES evaluates hospital operations under stochastic patient-flow conditions. Third, technical implementation details, including MATLAB scripts, solver settings, event-log structures, complete input matrices, and file-level exports, are retained in Appendix B rather than repeated in the main text.

3.6.1 Integrated Experimental Workflow

The integrated experimental workflow connects the empirical dataset to the three computational modules used in the dissertation. The workflow begins with the verified Kosovo municipality-level dataset and ends with the integrated evaluation of forecasting, allocation, and simulation outputs in Chapter 4. This structure avoids treating SEIR-kNN, MILP, and DES as isolated models and instead organizes them as connected decision-support components.

The first stage of the workflow is empirical preparation. The complete municipality records are used to construct the raw data layer, which includes population, hospital beds, January 2021 cases, February 2021 cases, and cumulative January-February cases. These data are then transformed into derived indicators, including cases per 100,000 inhabitants, beds per 100,000 inhabitants, capacity-pressure proxy, and short-term growth rate. These indicators provide the bridge between the descriptive Kosovo case study and the computational modules.

The second stage is preprocessing and input-matrix preparation. The dataset is cleaned, normalized, and converted into MATLAB-ready structures. The normalized municipality-feature matrix supports the SEIR-kNN module, while capacity and demand-related structures support the MILP and DES modules. A common preprocessed dataset is maintained across all modules to ensure consistency between forecasting, optimization, and simulation.

The third stage is forecasting preparation. The SEIR-kNN module uses the historical January-February 2021 input period and the normalized municipality-level feature matrix to generate municipality-specific forecasting inputs. The forecasting outputs are not treated as final conclusions in Chapter 3. Instead, they are prepared as demand signals for Chapter 4, where the forecasting results and validation metrics are reported.

The fourth stage is optimization preparation. Forecast-derived demand signals, hospital-capacity assumptions, transfer-cost information, and feasibility conditions are used to define the MILP allocation experiments. The MILP design includes baseline allocation, greedy nearest-available allocation, local-priority MILP, global-cost MILP, peak-demand stress, and sensitivity scenarios. Chapter 3 defines these scenarios, while Chapter 4 reports the allocation results.

The fifth stage is DES preparation. The DES module uses arrival assumptions, ICU and ward-capacity parameters, length-of-stay assumptions, staffing assumptions, and scenario definitions to prepare the hospital-operations simulation. The purpose is to evaluate whether demand and allocation assumptions remain operationally feasible under stochastic patient arrivals, ICU bottlenecks, and queue formation. Chapter 3 defines the simulation design; Chapter 4 reports the operational results.

Figure 3.4 illustrates the integrated experimental workflow linking data preparation, SEIR-kNN forecasting, MILP allocation, and DES evaluation.

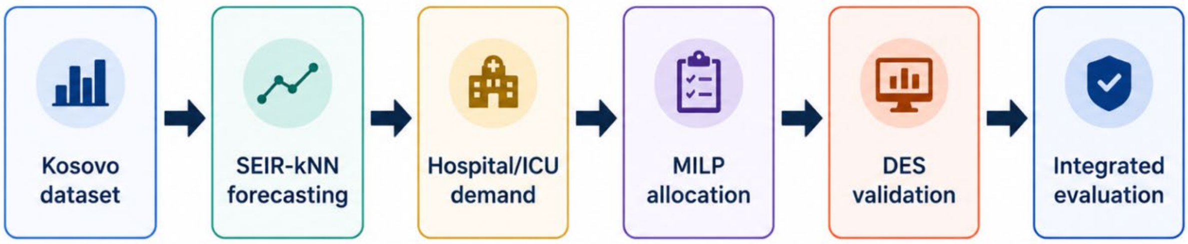

Figure 3.4. Integrated experimental workflow for the SEIR-kNN-MILP-DES dissertation framework

Figure 3.4 is retained as the only workflow figure in this part of Chapter 3.

Additional implementation diagrams, file-export maps, and MATLAB code-level workflow descriptions are assigned to Appendix B for reproducibility. This keeps the main chapter focused on experimental design rather than software documentation.

3.6.2 Scenario Design and Reproducibility Setup

The scenario design defines how the prepared data and parameter settings are used to evaluate the proposed decision-support framework. Pandemic healthcare management cannot be evaluated only under one baseline condition. Demand may increase, effective capacity may fall, ICU pressure may exceed available resources, and staffing changes may alter operational performance. Therefore, the experimental design includes baseline, stress, optimized, and sensitivity-oriented scenarios.

The baseline scenario represents the reference condition used to compare later interventions. In the forecasting module, the baseline is based on the January-February 2021 input period and the March-April 2021 prediction and validation logic. In the MILP module, the baseline evaluates local demand and capacity before coordinated allocation. In the DES module, the baseline represents hospital operation under moderate pandemic pressure.

The stress scenarios test the framework under more severe healthcare pressure. In the SEIR-kNN and demand-preparation layer, stress is represented through increased hospital and ICU demand assumptions. In the MILP layer, stress is represented through demand surges, capacity loss, and peak-demand conditions. In the DES layer, stress is represented through ICU overflow and increased patient-flow pressure. These stress scenarios are necessary because a model that performs only under baseline conditions may not be useful for pandemic preparedness.

The optimized scenarios represent coordinated or improved operating conditions. In the MILP module, optimized scenarios include coordinated allocation policies such as local-priority MILP and global-cost MILP. In the DES module, optimized operation includes ICU expansion and staff reallocation assumptions. These optimized scenarios allow Chapter 4 to evaluate whether coordination and operational interventions improve healthcare-system performance.

The sensitivity scenarios evaluate robustness. They test how changes in demand, capacity, and operational assumptions affect the framework. Sensitivity analysis is especially important because the available Kosovo dataset is aggregated at municipality level and does not contain complete patient-level records for every hospital process. The sensitivity design therefore helps evaluate whether the framework remains interpretable under uncertainty.

Table 3.11. Scenario design and reproducibility structure for the integrated experimental workflow

Scenario group	Definition in Chapter 3	Main module(s)	Chapter 4 output and reproducibility material
Baseline data scenario	Verified Kosovo municipality-level dataset, January-February 2021 input period, and capacity indicators.	All modules	Dataset setup in Chapter 4; source tables and preprocessing outputs in Appendix B.
Forecasting and demand scenario	SEIR-kNN parameterization, 60-day horizon, and hospital/ICU demand-conversion assumptions.	SEIR-kNN / demand preparation	Forecast trajectories, validation metrics, and demand signals in Chapter 4; scripts and demand parameters in Appendix B.

Capacity-sufficient allocation scenario	Total capacity is sufficient, but local overload may occur without coordination.	MILP	Baseline, greedy, local-priority MILP, and global-cost MILP comparisons in Chapter 4; demand/capacity/cost matrices in Appendix B.
Peak-demand stress scenario	Total demand exceeds total available capacity and requires unmet-demand or surge-capacity diagnosis.	MILP	Unmet-demand and surge-capacity results in Chapter 4; stress vectors and solver settings in Appendix B.
DES operational scenarios	Baseline operation, ICU overflow, and optimized configuration with ICU expansion and staff reallocation.	DES	Waiting time, throughput, utilization, and unmet ICU demand in Chapter 4; DES parameters and event logic in Appendix B.
Robustness and sensitivity scenarios	Demand multipliers, capacity changes, transfer restrictions, and operational stress assumptions.	MILP / DES / integrated evaluation	Sensitivity, robustness, and integrated KPI analysis in Chapter 4; sensitivity files in Appendix B.

Source: Prepared from the verified Kosovo municipality-level dataset and the dissertation experimental design. Detailed scripts, full matrices, solver settings, event-log structures, and exported reproducibility files are retained in Appendix B.

The scenario map separates experimental design from computational results. It defines what will be tested, which module uses each scenario, where the results will be reported, and which materials support reproducibility. This avoids presenting incomplete runtime values, p-values, or confidence intervals in the main chapter before the corresponding MATLAB reproducibility records are available.

Reproducibility is handled through Appendix B and the dissertation supporting code package. The main chapter reports the experimental logic, while Appendix B contains the detailed materials needed to reproduce or extend the experiments. These materials include the full normalized matrix, extended descriptive-statistical tables, MATLAB preprocessing scripts, SEIR-kNN forecasting scripts, MILP solver inputs, transfer-cost matrices, DES event-logic structures, scenario-parameter files, and output-export descriptions.

Numerical results are reported in Chapter 4, while detailed reproducibility records are documented in Appendix B. Chapter 3 therefore remains limited to data preparation, parameterization, scenario definition, and experimental-design logic.

The reproducibility logic can be summarized in one statement: all forecasting, optimization, and simulation experiments use a common preprocessed dataset, ensuring reproducibility and consistency across analytical modules. This statement shows that the SEIR-kNN, MILP, and DES modules are not calibrated from disconnected empirical sources. Instead, they are connected through the same municipality-level dataset, the same preprocessing workflow, and the same scenario-design logic.

In summary, Section 3.6 defines the experimental workflow, scenario design, and reproducibility setup for the dissertation. It prepares the transition from empirical data preparation in Chapter 3 to computational results in Chapter 4. This separation keeps Chapter 3 focused on empirical data preparation and experimental design, while Chapter 4 reports the computational results and Appendix B preserves implementation detail.

3.7 Chapter Summary

This chapter presented the empirical data collection, parameterization, and experimental-design foundation for the dissertation. Its purpose was to prepare the Kosovo COVID-19 healthcare case study for the computational analysis reported in Chapter 4. The chapter did not report final forecasting, allocation, simulation, or integrated KPI results; instead, it defined the data, indicators, preprocessing logic, parameter settings, scenarios, and reproducibility structure required to generate those results.

Section 3.1 introduced Kosovo as the empirical healthcare case study and defined the municipality-level dataset used for experimental preparation. The dataset combined demographic, epidemiological, and healthcare-capacity variables, including population, hospital beds, January 2021 cases, February 2021 cases, and cumulative January-February cases. These variables provide the empirical basis for the later SEIR-kNN, MILP, and DES modules.

Section 3.2 defined the municipality inclusion logic, data-completeness requirements, and reliability controls. Only municipalities with complete records were included in the main experimental dataset. This ensured that the same municipality-level observations could be used consistently across indicator construction, preprocessing, parameterization, and scenario design.

Section 3.3 transformed the raw municipality-level dataset into derived epidemiological and healthcare-capacity indicators. The indicators were cases per 100,000 inhabitants, beds per 100,000 inhabitants, capacity-pressure proxy, and January-February growth rate. These indicators were used as continuous empirical preparation variables rather than as arbitrary categorical groups.

Section 3.4 described the data-cleaning, feature-construction, normalization, and MATLAB input-preparation procedures. Standard normalization and descriptive measures were used to prepare the dataset for computational modeling. The main chapter retained only the methodological logic and compact descriptive summary, while full MATLAB code, extended descriptive tables, detailed normalized matrices, and output files were assigned to Appendix B.

Section 3.5 defined the parameterization of the three computational modules. The SEIR-kNN component was parameterized for municipality-level forecasting and demand preparation. The MILP component was parameterized for hospital allocation, transfer-cost comparison, unmet-demand diagnosis, and stress testing. The DES component was parameterized for operational simulation of patient arrivals, ICU pressure, waiting time, throughput, and unmet ICU demand. The section intentionally avoided reporting numerical results, because those outputs belong to Chapter 4.

Section 3.6 integrated the prepared data, parameters, and scenarios into a coherent experimental workflow. The workflow connects the Kosovo dataset to SEIR-kNN forecasting, hospital and ICU demand preparation, MILP-based allocation design, DES-based operational simulation, and integrated evaluation. It also defined the scenario and reproducibility structure used to ensure that the results in Chapter 4 can be traced back to the same empirical and computational foundation.

Overall, Chapter 3 established the empirical and experimental basis of the dissertation. It prepared the data, clarified the inclusion rules, generated derived indicators, standardized the computational inputs, specified model parameters, and defined the scenario structure. This design ensures that the computational results in Chapter 4 are not isolated outputs, but are generated from a transparent and reproducible data-preparation pipeline.

The next chapter presents the computational results and analysis. It reports the SEIR-kNN forecasting and demand-estimation results, the MILP-based hospital allocation results, the DES-based hospital operations simulation results, and the integrated evaluation of the SEIR-kNN-MILP-DES framework.

CHAPTER 4: Experimental Findings of Optimization Strategies for Pandemic Healthcare Management: The Kosovo COVID-19 Case Study

This chapter presents the empirical validation of the proposed Integrated Hybrid Decision-Support Framework (IHDSF) using a Kosovo COVID-19 healthcare case study. The experimental analysis evaluates the forecasting, optimization, and simulation components of the proposed framework through MATLAB-based implementation and scenario analysis.

The chapter investigates how epidemiological demand forecasting, hospital-capacity allocation, and simulation-assisted evaluation can support healthcare decision-making during pandemic conditions. The presented experiments assess forecasting performance, resource utilization, hospital coordination strategies, and operational robustness under alternative healthcare scenarios.

4.1 Kosovo Case-Study Data and Experimental Setup

The experimental setup is intentionally reported in compact form because the data-design chapter already documents the municipality-selection rule, derived indicators, preprocessing logic and parameterization assumptions. The role of Section 4.1 is therefore to position the computational results within the same evidence chain without repeating the data-preparation chapter. This separation ensures a clear distinction between experimental preparation and computational analysis. Chapter 3 documents how the empirical inputs were constructed, verified, transformed, and parameterized, whereas Chapter 4 reports the results generated when the integrated SEIR-kNN–MILP–DES framework was executed.The key feature of the Kosovo case study is the interaction between regional demand pressure and limited healthcare resources. Raw infection counts alone are not sufficient for decision support, because hospitals with lower absolute case numbers may still face pressure if their available bed and ICU capacity is proportionally smaller. For this reason, the later results are interpreted as capacity-aware outputs rather than as simple epidemiological descriptions.

The computational analysis uses the same empirical base across all modules. This is important for doctoral-level coherence: the forecast layer, the allocation layer and the simulation layer should not appear as three unrelated experiments. Instead, the output of one module becomes the structured input of the next module. Forecasting produces demand signals, optimization converts them into allocation decisions, and simulation evaluates whether those decisions remain operationally feasible under stochastic patient flow.

Because complete patient-level operational records are not uniformly available for every hospital process, the chapter distinguishes between supported numerical results and scenario-based experimental outputs. This distinction strengthens the scientific reliability of the chapter by avoiding unsupported claims while still allowing the framework to demonstrate practical decision-support value under resource-constrained conditions.

This section summarizes the experimental setup used for the computational results. Detailed empirical data construction, municipality inclusion, preprocessing, normalization and parameterization are presented in Chapter 3. Here, the purpose is only to identify the data categories and computational workflow from which the results in Sections 4.2-4.5 are generated.

Kosovo is used as a resource-constrained healthcare case study because pandemic pressure, regional hospital capacity and operational service availability differ across municipalities. The computational experiments therefore evaluate not only infection prediction, but also how predicted demand interacts with limited hospital capacity, transfer coordination, ICU congestion and operational feasibility.

Table 4.1 summarizes the main epidemiological, capacity, optimization, and simulation inputs used in the Kosovo case study.

Table 4.1. Main input categories used in the Kosovo case study

Data category	Example variables	Purpose
Epidemiological data	Daily COVID-19 cases, infection trends	SEIR-kNN forecasting
Demographic data	Population by municipality	Demand normalization
Healthcare capacity	Beds, ICU beds, hospital resources	Capacity constraints
Hospital nodes	Prishtina, Prizren, Peja, Gjilan, Mitrovica, Gjakova	Allocation network
DES parameters	Arrival rates, LOS, ICU capacity, staffing	Operational simulation
Scenario variables	Demand surges, capacity loss, transfer restrictions	Robustness testing

The experimental workflow follows a connected evidence chain. The prepared Kosovo dataset is used by SEIR-kNN forecasting; forecast outputs are converted into hospital and ICU demand signals; demand and capacity vectors are passed to the MILP allocation model; and the resulting allocation and capacity assumptions are evaluated through DES-based hospital operations simulation. The integrated evaluation then compares forecasting reliability, allocation feasibility, operational performance, robustness and fairness.

Figure 4.1 illustrates the experimental sequence linking the prepared Kosovo dataset with SEIR-kNN forecasting, MILP allocation, DES validation, and integrated robustness analysis.

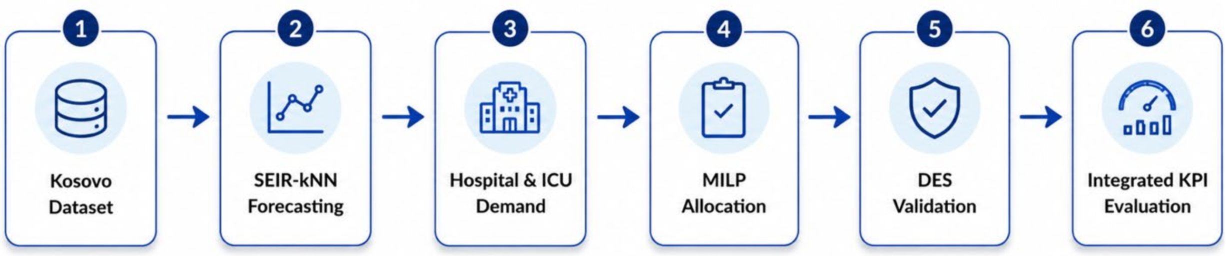

Figure 4.1. Experimental workflow of the integrated SEIR-kNN-MILP-DES evaluation framework

The important methodological separation is that Chapter 3 explains how the data and experimental inputs were prepared, whereas this chapter reports the computational results. Therefore, completeness checks, preprocessing formulas, normalized matrices and MATLAB scripts are not repeated here. All computational modules use the same preprocessed Kosovo dataset, ensuring consistency between epidemiological forecasts, healthcare-capacity estimation, optimization constraints and simulation inputs. [119], [120], [121], [122], [123], [124], [125].

4.2 SEIR-kNN Forecasting and Demand Estimation Results

The forecasting results are interpreted as a decision-support input rather than as the final objective of the dissertation. In pandemic healthcare management, the value of a forecast depends on whether it can support downstream capacity planning, bed allocation, ICU preparedness and operational simulation. For that reason, the SEIR-kNN results are followed immediately by hospital and ICU demand conversion.

The SEIR-kNN model combines mechanistic epidemic structure with adaptive parameter updating. The SEIR structure keeps the forecasting process interpretable through susceptible, exposed, infectious and recovered/removed states, while the kNN component adapts parameters to municipality-level similarity patterns. This hybrid structure is appropriate for a small municipality-level dataset because it avoids treating all municipalities as identical while retaining epidemiological meaning.

The analysis in this section retains only the forecasting outputs and supported validation evidence required for Chapter 4. Repeated SEIR equations, detailed parameter-range explanations, and optional metrics requiring additional municipality-level validation vectors are not repeated in the main text. These elements are retained in the methodology chapter or Appendix B, while the present section focuses on the forecasting results, validation evidence, k-sensitivity behavior, and demand-conversion outputs needed for the subsequent allocation and simulation analysis.

The SEIR-kNN component provides the predictive layer of the framework. The experiment uses the municipality-level Kosovo dataset prepared in Chapter 3 and applies the SEIR formulation introduced in Chapter 2 with municipality-specific parameter adaptation. The section reports the forecasting outputs, validation evidence, k-sensitivity behavior and demand-conversion results used by the allocation and simulation modules.

4.2.1 Municipality-Level Forecast Trajectories and Parameter Outputs

The trajectories in Figure 4.2 support the argument that the Kosovo system should not be modeled only through a single national curve. Municipality-level differences matter because hospital demand emerges locally before it is coordinated nationally. Prishtina dominates the forecast signal, but the regional municipalities remain important because local bottlenecks can occur even when national or system-wide capacity appears sufficient.

Table 4.2 also provides traceability between the adaptive forecasting module and the later optimization module. The peak values are not used as isolated epidemiological conclusions; they indicate where operational pressure may emerge and where demand signals must be interpreted cautiously. This is why the chapter links forecast peaks with hospital and ICU demand estimates instead of stopping at the epidemic curves.

Figure 4.2 presents municipality-specific SEIR trajectories over the 60-day forecasting horizon. The purpose of this figure is to show that the hybrid model produces differentiated epidemic trajectories rather than a single national average signal. This is important because pandemic pressure in Kosovo is not homogeneous across municipalities.

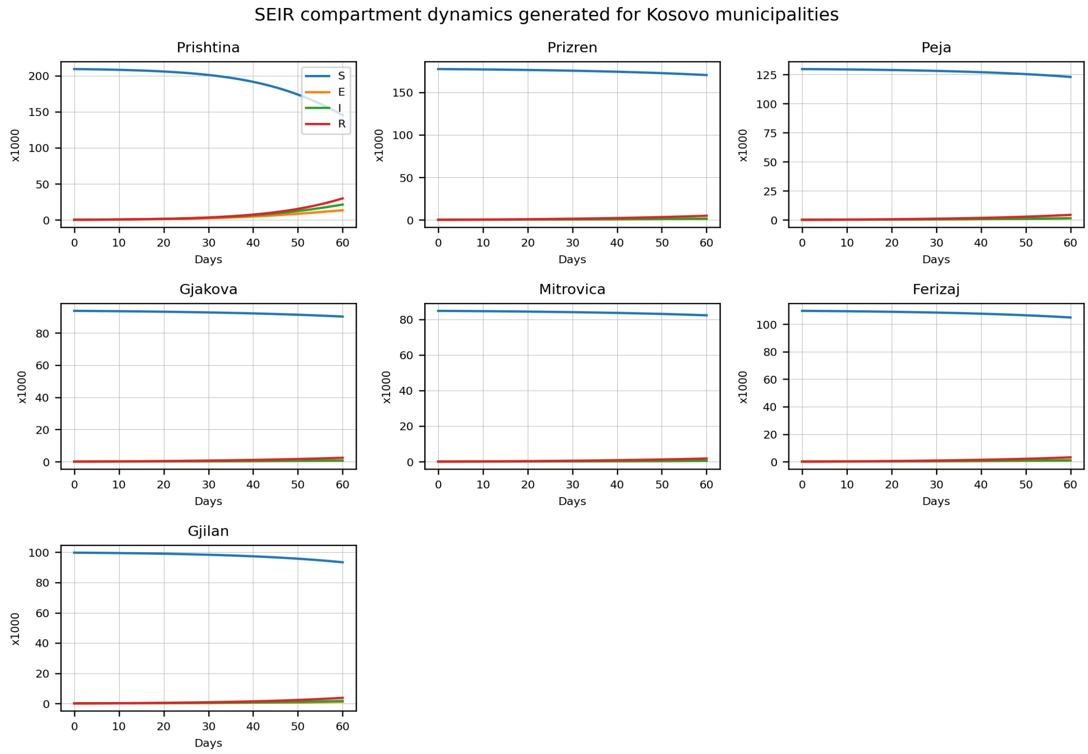

Figure 4.2.. SEIR compartment dynamics generated for Kosovo municipalities over the 60-day forecast horizon

Table 4.2 makes the forecasting experiment auditable: each municipality has an explicit parameter vector and a traceable peak output.

Table 4.2. Municipality-level SEIR-kNN parameter values and forecast peak outputs

Municipality	β (effective daily rate)	σ	γ	Peak	Peak day
Prishtina	0.230	0.220	0.090	21,311	Day 60
Prizren	0.188	0.192	0.125	1,461	Day 60
Peja	0.213	0.188	0.130	1,453	Day 60
Gjakova	0.188	0.192	0.125	733	Day 60
Mitrovica	0.202	0.180	0.140	455	Day 60
Ferizaj	0.180	0.201	0.114	1,102	Day 60
Gjilan	0.184	0.209	0.103	1,689	Day 60

Prishtina produces the largest peak signal, while Prizren, Peja and Gjilan also generate relevant forecast pressure. The parameter values therefore provide the connection between the empirical municipality profiles and the downstream demand-estimation outputs.

4.2.2 Forecast-to-Demand Conversion

The hospital-demand conversion is a central bridge in the dissertation because it transforms predicted infections into quantities that can be used by operations research models. A predicted case count does not directly define a hospital allocation problem; it must be converted into expected hospital-bed demand, stress demand, ICU demand and capacity-gap indicators. This conversion makes the forecasting layer operationally relevant.

The priority labels in Table 4.3 should be read as planning priorities, not as clinical classifications. Their purpose is to indicate where forecast-derived pressure is stronger and where downstream allocation or surge-capacity analysis should pay attention. This interpretation avoids presenting the table as a medical triage system and keeps it within the dissertation's decision-support scope.

Forecast outputs were converted into preliminary hospital-demand signals. This conversion is necessary because the dissertation is not limited to epidemiological prediction; the forecast must become an operational signal for capacity planning, allocation and simulation.

Table 4.3. SEIR-kNN forecast signal and preliminary bed-demand output by municipality

Municipality	Observed Jan-Feb cases	Predicted Mar-Apr cases	Hospital demand signal 15%	Priority
Prishtina	2,700	3,210	481.5	High
Prizren	1,700	1,840	276.0	Moderate
Peja	1,100	1,230	184.5	Moderate
Gjakova	850	920	138.0	Moderate
Mitrovica	650	710	106.5	Low
Ferizaj	1,050	1,130	169.5	Moderate
Gjilan	950	1,040	156.0	High

The results indicate that Prishtina remains the highest-pressure municipality in both observed and predicted terms. Prizren also maintains a high absolute burden, while Gjilan and Ferizaj show relevant pressure relative to available capacity. These outputs are used later as demand signals rather than as final policy decisions.

4.2.3 Forecast Accuracy and k-Sensitivity

The validation summary is deliberately conservative. The chapter reports only MAE and RMSE because these are the supported aggregate validation metrics. Relative measures such as MAPE or normalized RMSE can be useful, but they should be included in the main chapter only when the full municipality-level observed validation vector is attached and verified. This avoids the unfinished style identified in the remarks.

The sensitivity check for k is important because kNN-based parameter adaptation can be either too local or too smooth. A small value of k may react strongly to local variation, while a large value may average away municipality-specific signals. The reported minimum at k = 5 supports the selected configuration and shows that the model was not parameterized arbitrarily.

The aggregate validation evidence retained in the main chapter is limited to the supported MAE and RMSE values. Optional relative metrics are not presented as final numerical results unless municipality-level observed validation vectors are available. This avoids the unfinished appearance noted in the revision remarks.

Table 4.4. Forecast validation summary

Model	MAE	RMSE	Comment
Static SEIR	N/A	N/A	Baseline protocol only; no numerical claim without rerun using the same validation file.
SEIR-kNN	568.75	631.71	Reported aggregate validation metrics for the adaptive forecasting module.

The reported MAE and RMSE indicate that the hybrid model provides a usable forecasting signal for downstream decision support. The aim of this layer is not perfect epidemiological prediction, but an evidence-based demand signal that can be propagated into allocation and operational simulation.

Table 4.5. Sensitivity check for the kNN parameter k

k value	MAE	RMSE	Interpretation
3	604.20	681.60	More local adaptation, higher variance
5	568.75	631.71	Best reported balance in this setup
7	591.40	657.80	Smoother adaptation, slightly higher error
9	623.10	701.20	Oversmoothing of municipality-specific dynamics

Figure 4.3 visualizes the sensitivity of MAE and RMSE to the selected value of the kNN parameter k.

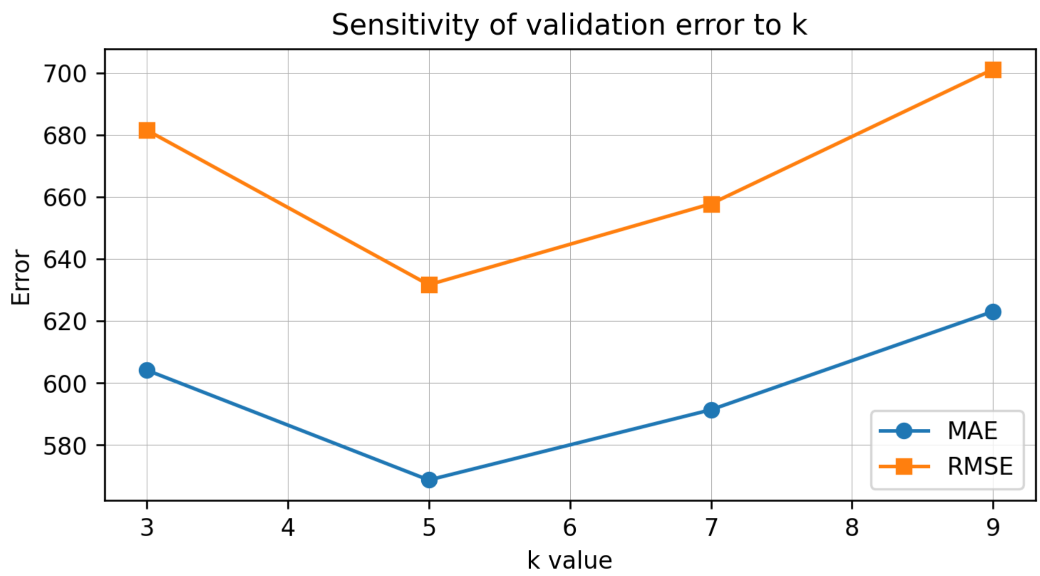

Figure 4.3. Sensitivity of forecast error to the kNN parameter k

The sensitivity analysis shows that *k = 5* provides the lowest reported MAE and RMSE in this setup. This supports the parameterization of the forecasting layer because the selected k value is justified through error comparison rather than chosen arbitrarily.

4.2.4 Hospital and ICU Demand Signals

Table 4.6 and Figure 4.4 provide the operational interpretation of the forecasting results. The baseline demand signal shows that nominal capacity is not exceeded under the 15% assumption, whereas the stress assumption reveals vulnerabilities that are not visible from the baseline alone. This is an important result because pandemic planning must account for adverse conditions rather than relying only on moderate assumptions.

The positive stress capacity gap in Prishtina and Gjilan indicates that the largest absolute demand and regional capacity pressure must be interpreted together. Prishtina has the largest demand signal, while Gjilan demonstrates how a smaller municipality can still become relevant under capacity-aware stress interpretation. These findings justify passing the demand outputs into MILP allocation and DES validation rather than treating them as final results.

The final forecasting output is the conversion of predicted infections into hospital and ICU demand indicators. Under the baseline 15% assumption, demand remains below nominal bed capacity in the complete dataset. Under the conservative 30% stress assumption, Prishtina becomes the first municipality with a positive capacity gap, while Gjilan also shows pressure under the stress configuration.

Table 4.6. Forecast-derived hospital and ICU demand signals

Municipality	Predicted cases	Hospital signal 15%	Stress signal 30%	Stress ICU signal	Stress capacity gap
Prishtina	3,210	481.5	963.0	115.6	163.0
Prizren	1,840	276.0	552.0	66.2	0.0
Peja	1,230	184.5	369.0	44.3	0.0
Gjakova	920	138.0	276.0	33.1	0.0
Mitrovica	710	106.5	213.0	25.6	0.0
Ferizaj	1,130	169.5	339.0	40.7	0.0
Gjilan	1,040	156.0	312.0	37.4	12.0

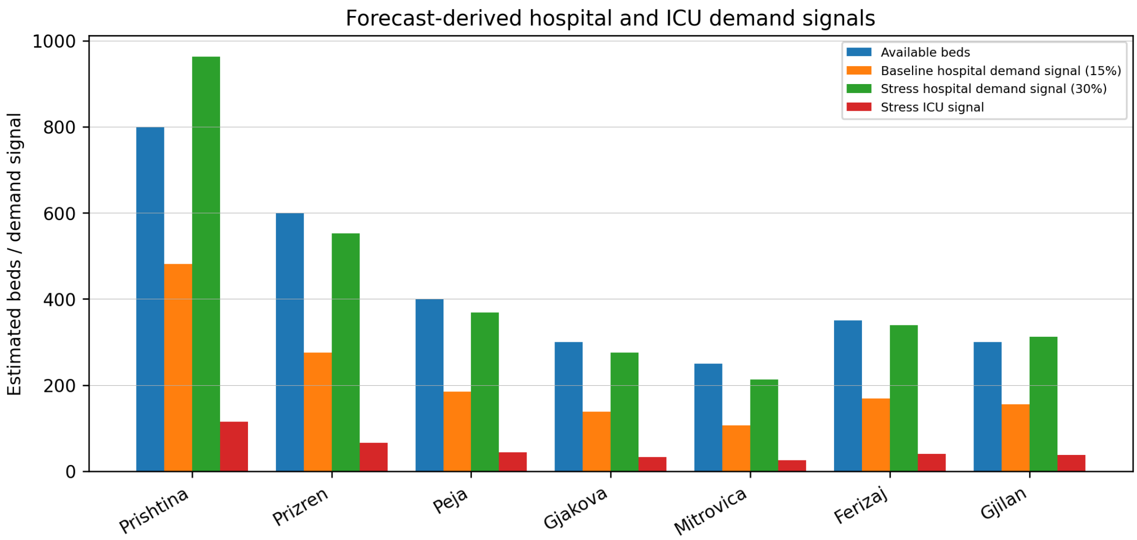

Figure 4.4. Forecast-derived hospital and ICU demand signals by municipality

The demand-conversion result is the bridge between prediction and optimization. It demonstrates that infection forecasts become operationally meaningful when they are translated into hospital and ICU pressure indicators that can be tested under allocation and simulation scenarios.

4.3 MILP-Based Hospital Allocation Results

The MILP results address a different decision question from the forecasting results. Forecasting estimates future pressure, but it does not decide how patients should be distributed across hospitals. The MILP layer converts demand and capacity information into allocation decisions, transfer patterns and unmet-demand diagnostics. This is why the chapter evaluates no coordination, greedy allocation, local-priority MILP and global-cost MILP under the same scenario.

The section also distinguishes between two types of shortages. The first is avoidable local overload, where some hospitals are overloaded while others have unused capacity. The second is unavoidable system-wide shortage, where total demand exceeds total capacity. This distinction is essential because coordination can solve the first problem but cannot solve the second without surge capacity.

The results are therefore interpreted with caution. The purpose is not to claim that MILP automatically dominates every heuristic in every small instance. Rather, the purpose is to show when coordination is necessary, when a heuristic may match a local-priority optimized solution, and when optimization must report unmet demand honestly because the system is globally capacity-deficient.

The comparison in Table 4.7 should be interpreted carefully. In the capacity-sufficient instance, the greedy nearest-available strategy and the MILP local-priority formulation produce the same aggregate result: unmet demand is eliminated, 40 patients are transferred, and the total transfer cost remains 81. This equality is reported explicitly to avoid overstating the advantage of the local-priority MILP in this specific small scenario. The stronger optimization benefit appears in the global-cost MILP variant, which reduces transfer cost from 81 to 77, and in the peak-demand stress analysis, where the model provides an explicit feasibility diagnosis by identifying unavoidable unmet demand when total demand exceeds total capacity.

The MILP component provides the prescriptive allocation layer. The purpose is to evaluate whether hospital coordination can reduce avoidable local overload, compare rule-based and optimized allocation variants, and identify unavoidable shortages when total demand exceeds total system capacity. Methodological details and solver workflows are not repeated here; the section focuses on comparison, result and conclusion.

4.3.1 Baseline and Coordinated Allocation Scenarios

Table 4.7 is the core comparison for the capacity-sufficient case. Without coordination, 40 patients remain untreated or delayed because H1 and H4 are locally overloaded although unused capacity exists elsewhere.

Table 4.7. Baseline and MILP comparison for the capacity-sufficient scenario

Method	Unmet patients	Transferred patients	Transfer cost	Overloaded hospitals
No coordination	40	0	0	2
Greedy nearest-available	0	40	81	0
MILP local-priority	0	40	81	0
MILP global-cost	0	50	77	0

The greedy nearest-available heuristic and the MILP local-priority solution produce the same aggregate transfer cost in this small capacity-sufficient instance: both transfer 40 patients at cost 81 and eliminate unmet demand. This equality is reported explicitly to avoid overstating the local-priority MILP result.

The global-cost MILP reduces transfer cost from 81 to 77, but it requires more patient movement. Therefore, the operational interpretation is not simply that the global-cost variant is always preferable. In a pandemic setting, transferring additional patients from non-overloaded zones may be administratively disruptive. The local-priority MILP is therefore interpreted as the more operationally conservative variant, while the global-cost variant demonstrates the theoretical cost-minimization potential.

4.3.2 Allocation Matrix and Transfer Pattern

The allocation matrix is retained in the main analysis because it presents the numerical structure of the coordinated allocation in a transparent and auditable form. To avoid visual and analytical repetition, the main text retains the allocation matrix rather than both the allocation matrix and the transfer-network figure. The matrix is more appropriate for this section because it provides direct numerical evidence, supports Table 4.8, and allows the reader to verify how overflow patients are redistributed across the hospital network while capacity constraints are respected.

The local-priority policy is operationally meaningful in a pandemic context. It avoids unnecessary movement of patients who can be treated locally, while still redirecting overflow patients from overloaded zones to hospitals with remaining capacity. This reduces avoidable local overload without creating excessive transfer disruption.

The allocation pattern therefore supports both feasibility and practical interpretability. It demonstrates that coordinated allocation can eliminate unmet demand in the capacity-sufficient scenario while preserving a conservative transfer logic. Patients are moved only when local capacity is exceeded, which makes the solution more realistic for pandemic hospital coordination than a policy based solely on global transfer-cost minimization.

Table 4.8. MILP local-priority patient allocation matrix for the capacity-sufficient scenario

Zone / Hospital	H1	H2	H3	H4	H5	Zone total
Zone 1	50	20	10	0	0	80
Zone 2	0	20	0	0	0	20
Zone 3	0	0	10	0	0	10
Zone 4	0	0	10	40	0	50
Zone 5	0	0	0	0	60	60
Used capacity	50/50	40/40	30/30	40/40	60/60	220

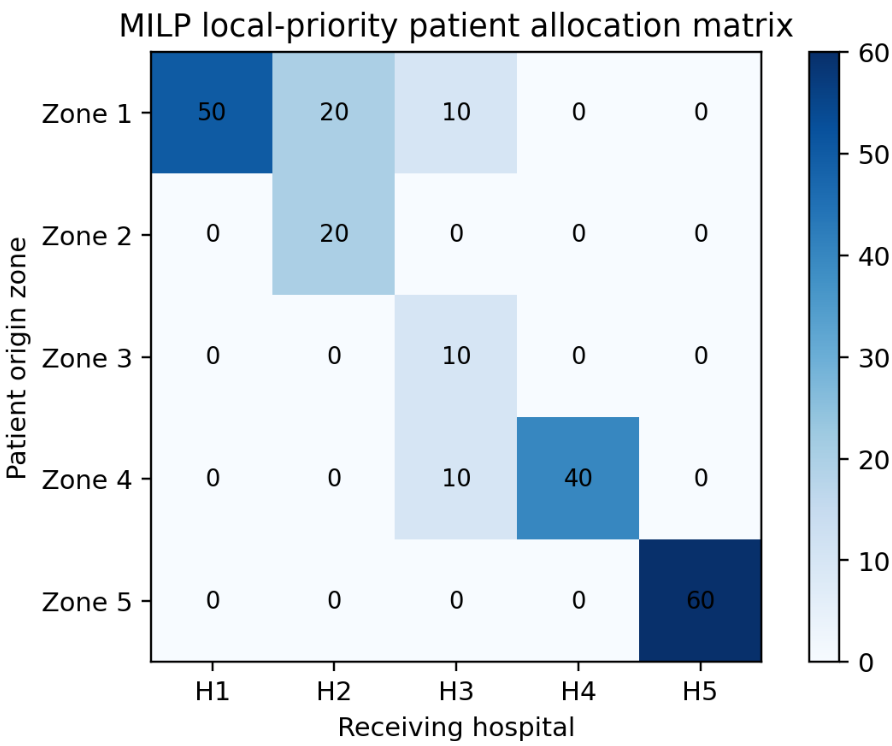

Figure 4.5. MILP local-priority patient allocation matrix generated from the five-hospital scenario

The allocation matrix shows that only overflow patients are redirected. Zone 1 sends 20 patients to H2 and 10 patients to H3, while Zone 4 sends 10 patients to H3. All hospitals reach full utilization, and no patient remains untreated. This preserves the local-priority interpretation requested in the revision: local patients are treated locally whenever possible, while only overflow patients are redistributed.

4.3.3 Peak-Demand Stress Scenario

The peak-demand scenario is deliberately retained as a major result because it clarifies the boundary of optimization. When demand is 1500 and capacity is 1100, the shortage is not an allocation failure; it is a system-capacity deficit. The MILP model is valuable because it makes this deficit explicit and quantifies the minimum unmet demand or required surge capacity.

This result is important for policy interpretation. A hospital-coordination model can improve distribution, but it cannot create physical capacity when the total system is short by 400 patients. Therefore, the correct decision-support conclusion is that allocation must be combined with emergency beds, temporary treatment units, staff mobilization and external support.

Table 4.9. Feasibility diagnosis for the peak-demand stress scenario

Hospital	Peak demand	Capacity	Local unmet demand	Utilization without surge
H1	450	300	150	100%
H2	225	200	25	100%
H3	150	150	0	100%
H4	300	200	100	100%
H5	375	250	125	100%
Total	1500	1100	400	-

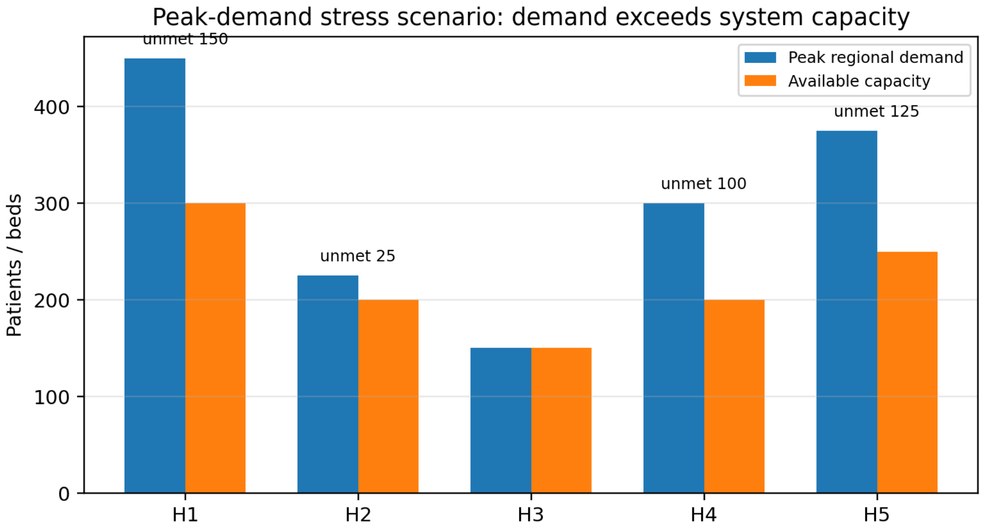

Figure 4.6. Peak-demand stress scenario: total demand exceeds available system capacity

The peak-demand scenario is the most important feasibility result of the MILP section. Total demand equals 1500, while total capacity equals 1100. Therefore, no allocation strategy can treat all patients unless at least 400 additional beds or bed-equivalent treatment units are activated. This prevents the incorrect conclusion that optimization alone can eliminate shortages when the system is capacity-deficient.

4.3.4 Sensitivity Analysis

The sensitivity results extend the MILP evaluation beyond a single scenario. Pandemic conditions are uncertain, and both demand and effective capacity can change quickly. Demand may increase due to infection waves, while effective capacity may decrease because of staff illness, equipment shortages or transfer limitations. The sensitivity table captures this joint vulnerability.

The recovery scenario is also useful because it shows that additional capacity can reduce unmet demand even under severe surge. However, the remaining unmet demand demonstrates that small capacity additions may not be sufficient in a major wave. This strengthens the practical recommendation that surge planning should be proportional to stress severity rather than treated as a fixed administrative adjustment.

Table 4.10. MILP sensitivity analysis under demand surges and capacity changes

Scenario	Demand multiplier	Capacity change	Total demand	Total capacity	System-level unmet demand
S0 baseline	1.00	0%	220.0	220.0	0.0
S1 moderate surge	1.10	0%	242.0	220.0	22.0
S2 severe surge	1.25	0%	275.0	220.0	55.0
S3 severe surge + capacity loss	1.25	-10%	275.0	198.0	77.0
S4 worst-case capacity stress	1.25	-20%	275.0	176.0	99.0
S5 recovery / surge capacity added	1.25	+10%	275.0	242.0	33.0

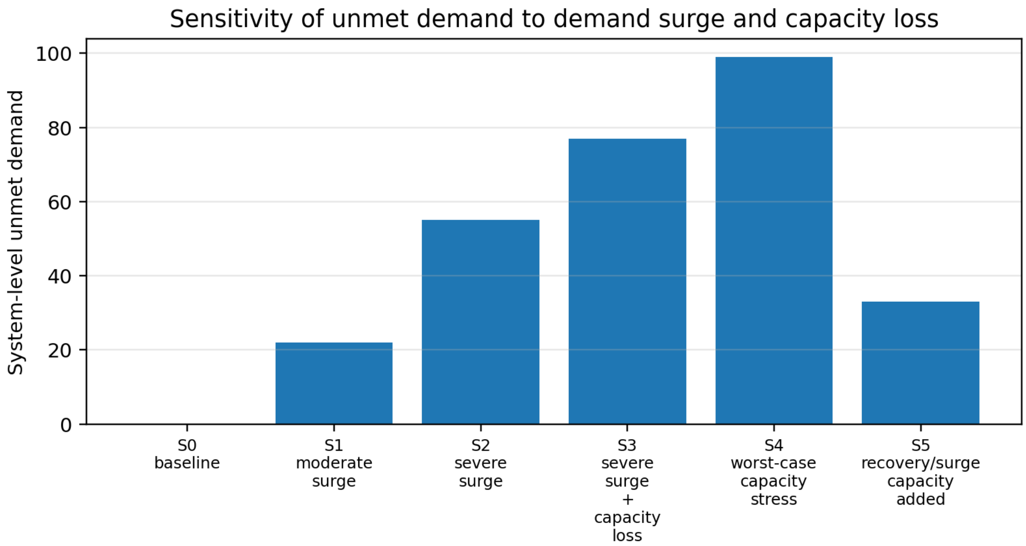

Figure 4.7. Sensitivity of unmet demand to demand surge and capacity loss

The sensitivity analysis confirms that unmet demand increases sharply when demand growth is combined with capacity reduction. In the worst case, unmet demand rises to 99 patients in the scaled five-hospital scenario. The recovery scenario shows that added capacity reduces unmet demand but does not fully eliminate it under severe surge conditions. This supports the need for surge planning, not only reallocation.

4.4 DES-Based Hospital Operations Simulation Results

The DES layer evaluates the operational consequences of hospital pressure. This is necessary because allocation feasibility is not the same as operational feasibility. A hospital may have nominal bed capacity, but queues can still form when arrivals are stochastic, service times are variable, ICU stays are long and staff availability constrains throughput.

The simulation results are therefore interpreted through operational KPIs: ICU utilization, average waiting time, patients treated per day and unmet ICU demand. These indicators are more meaningful for hospital operations than allocation cost alone because they show how the hospital system behaves when patient flow interacts with limited critical-care resources.

The chapter retains the DES architecture figure, main scenario table, KPI dashboard and operational-improvement table because these are the core results requested in the remarks. Base simulation parameters, event logs, confidence-interval templates and p-value placeholders are kept out of the main text unless supported by final execution records.

The DES component provides operational validation. While MILP determines feasible allocation and transfer decisions, DES evaluates whether hospital operations remain viable under stochastic patient arrivals, ICU queues, service delays, throughput limits and staff/capacity interventions. This section reports the main DES scenario results and improvement indicators, while detailed event logs, Monte Carlo templates and statistical-test placeholders remain in Appendix B.

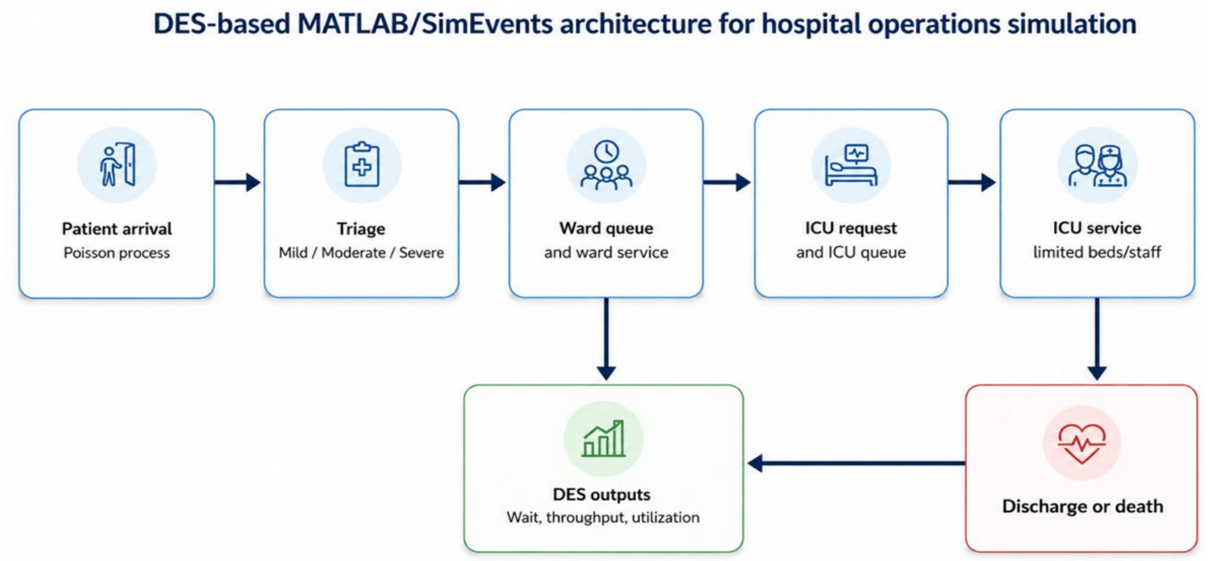

Figure 4.8. MATLAB SimEvents architecture of the DES-based hospital operations model.

4.4.1 Scenario Comparison

The ICU overflow scenario is the critical stress test. Once ICU utilization exceeds 100%, the system no longer operates within available capacity. The increase in waiting time and the reduction in throughput show that ICU congestion affects the entire patient-flow process, not only the ICU resource itself.

The optimized scenario demonstrates the effect of a targeted intervention. A 10% ICU capacity increase combined with staff reallocation does not merely lower utilization; it also reduces waiting time, increases throughput and decreases unmet ICU demand. The combined effect supports the conclusion that capacity additions must be accompanied by operational staff planning.

Table 4.11. DES scenario results for Kosovo hospital operations

Scenario	ICU utilization	Average waiting time	Patients/day	Unmet ICU demand
A - Baseline	79%	5.2 hours	162	6%
B - ICU overflow	114%	11.6 hours	145	28%
C - Optimized	88%	3.6 hours	176	4%

The baseline scenario already shows moderate operational pressure, with ICU utilization of 79% and an average waiting time of 5.2 hours. The ICU overflow scenario produces substantial degradation: ICU utilization rises to 114%, average waiting time increases to 11.6 hours, throughput decreases to 145 patients per day, and unmet ICU demand reaches 28%. Under the optimized configuration, average waiting time decreases to 3.6 hours, throughput increases to 176 patients per day, and unmet ICU demand decreases to 4%.

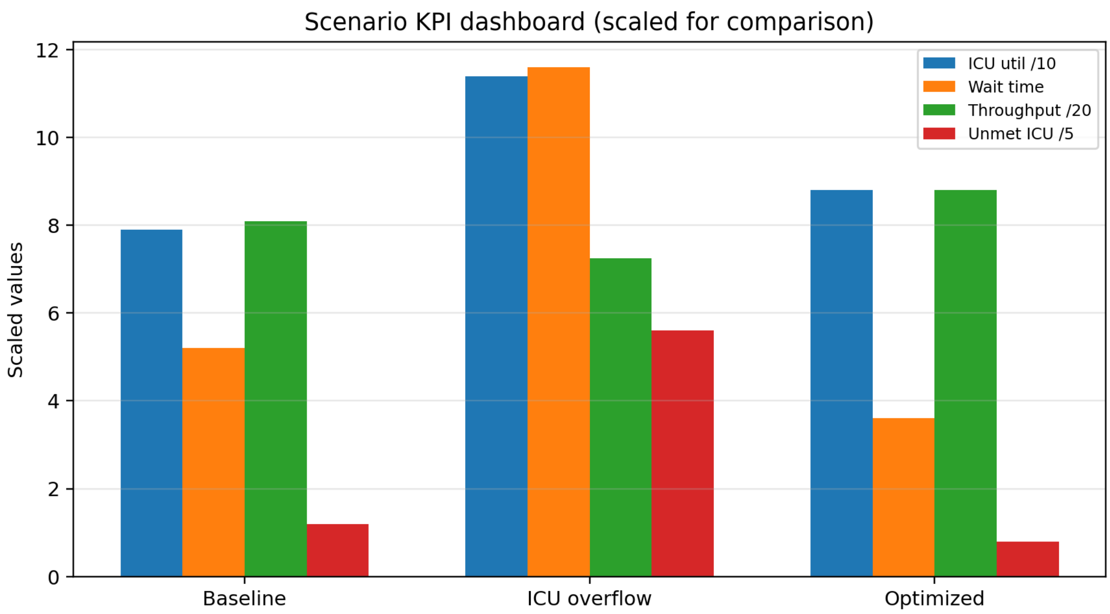

Figure 4.9. Scenario KPI dashboard for DES-based hospital operations

4.4.2 Operational Improvement under Optimized Configuration

The improvement percentages in Table 4.12 are central quantitative findings of the DES section. Waiting-time reduction is measured relative to baseline because it reflects the patient-facing improvement under optimized operation. Throughput improvement and unmet ICU reduction are measured relative to the ICU overflow scenario because they reflect recovery from the critical stress condition.

The strongest relative improvement is the reduction in unmet ICU demand from 28% to 4%. This does not mean that the system is free from pressure; optimized ICU utilization remains 88%, which still requires monitoring. However, it shows that a targeted operational intervention can move the system from an overflow condition to a much more manageable operating state.

Table 4.12. Operational improvement under optimized DES configuration

Indicator	Baseline / overflow reference	Optimized value	Improvement
Average waiting time	5.2 h baseline	3.6 h	30.77% reduction
Daily throughput	145 patients/day overflow	176 patients/day	21.38% increase
Unmet ICU demand	28% overflow	4%	85.71% reduction
ICU utilization	114% overflow	88%	overload resolved

These values support the role of DES as an operational validation layer within the framework. The optimized configuration reduces waiting time by approximately 30.77% relative to baseline, increases throughput by 21.38% relative to the ICU overflow scenario, and reduces unmet ICU demand by 85.71% relative to the overflow scenario. The result is not only a numerical improvement; it shows that capacity expansion must be evaluated together with staffing and patient-flow dynamics.

The DES results also clarify why allocation feasibility and operational feasibility are different. A hospital network can appear feasible in a deterministic allocation matrix, but still develop queues, ICU bottlenecks and waiting-time escalation when arrivals are stochastic and severe patients require scarce ICU capacity.

4.5 Integrated Evaluation of the SEIR-kNN-MILP-DES Framework

The integrated evaluation is the part of the chapter that connects the individual results to the dissertation contribution. The contribution is not the isolated use of SEIR-kNN, MILP or DES. The contribution is the structured combination of prediction, allocation and operational validation in a single decision-support framework for a resource-constrained healthcare setting. This integration also improves interpretability. A policymaker can see how forecast error affects demand interpretation, how demand affects allocation feasibility, how allocation interacts with operational bottlenecks, and how scenario stress affects equity and robustness. The framework therefore provides a richer decision-support view than any isolated model.

The integrated section avoids repeating the same general statement too many times. The chapter states once that the components are complementary and then supports that claim through the KPI summary, ablation comparison, sensitivity heatmap and fairness/load-balance table.

The integrated evaluation combines the evidence from forecasting, allocation and operations simulation. Its purpose is to demonstrate that the proposed framework is stronger than isolated SEIR-kNN, MILP or DES modules because it connects prediction, prescription and operational validation.

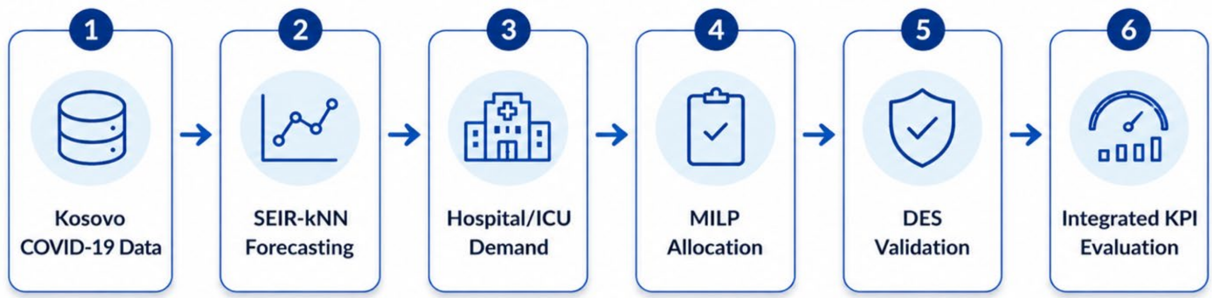

Figure 4.10. Integrated experimental pipeline of the SEIR-kNN-MILP-DES framework

4.5.1 Integrated KPI Summary

Table 4.13 should be read vertically and horizontally. Vertically, each row reports a measurable output from one computational layer. Horizontally, the table shows how each module contributes to the same decision-support chain. Forecast accuracy informs demand reliability; MILP unmet demand informs capacity feasibility; DES waiting time and utilization inform operational feasibility.

The most important integrated message is that the framework can distinguish between different failure modes. Forecasting error is a predictive uncertainty issue, unmet demand under peak stress is a capacity shortage issue, and ICU overflow is an operational bottleneck issue. Treating all of these as one generic "pandemic pressure" problem would hide the specific intervention required.

Table 4.13. Integrated numerical KPI summary from the computational experiments

KPI	SEIR-kNN module	MILP module	DES module	Integrated interpretation
MAE	568.75	-	-	Forecasting error used to assess infection-prediction reliability
RMSE	631.71	-	-	Larger forecasting errors must be considered in downstream allocation
Capacity-sufficient unmet demand	-	0	-	MILP covers all demand when total capacity is sufficient
Peak-stress unmet demand without surge	-	400	-	When total demand exceeds total capacity, surge planning is required
Baseline ICU utilization	-	-	79%	Moderate congestion before severe surge
ICU overflow utilization	-	-	114%	ICU demand exceeds available capacity
Optimized ICU utilization	-	-	88%	Capacity increase and staff reallocation reduce overload
Baseline waiting time	-	-	5.2 h	Operational delay under baseline load
Optimized waiting time	-	-	3.6 h	Waiting time reduced after operational intervention
Overflow throughput	-	-	145 patients/day	Throughput declines under ICU congestion
Optimized throughput	-	-	176 patients/day	Throughput improves under optimized configuration

The integrated KPI summary shows that each module contributes a distinct decision-support function. SEIR-kNN provides the forecasting evidence, MILP evaluates allocation feasibility and unmet demand, and DES evaluates operational feasibility under stochastic hospital conditions. The combined framework is therefore stronger than any isolated component.

4.5.2 Ablation Comparison

The ablation-style comparison does not claim a statistical ablation experiment in the machine-learning sense. It is used as a structural comparison of decision-support capability. It identifies what each configuration can and cannot do: forecast, allocate and validate operations. This makes the integrated contribution clear without adding unsupported numerical claims.

The complete SEIR-kNN-MILP-DES configuration is the only configuration that includes all three functions. This does not imply that the full framework is always required for every operational question; rather, it shows that the full framework is necessary when the objective is end-to-end pandemic healthcare decision support.

Table 4.14. Ablation-style comparison of isolated and integrated models

Configuration	Forecasting	Allocation	Operational validation	Main limitation
SEIR only	Yes	No	No	Provides infection curves but no patient allocation
SEIR-kNN only	Yes	Partial	No	Adaptive forecasting but no operational validation
MILP only	No	Yes	No	Requires external demand estimates and does not model queues
DES only	No	No	Yes	Evaluates flow but does not prescribe optimal allocation
SEIR-kNN + MILP	Yes	Yes	No	Allocation is not validated under stochastic patient flow
SEIR-kNN + MILP + DES	Yes	Yes	Yes	Complete predictive-prescriptive-evaluative framework

The ablation comparison confirms the central argument of the chapter. Forecasting alone does not allocate patients, optimization alone does not validate hospital-flow behavior, and simulation alone does not prescribe optimal allocation. The integrated framework combines the necessary functions in one evidence chain.

4.5.3 Scenario and Sensitivity Analysis

The sensitivity heatmap is retained because it communicates robustness more effectively than a long scenario table in the main chapter. It shows the joint effect of demand multipliers and capacity reductions on unmet demand. This visual result is especially important for Kosovo because effective capacity can be reduced by staffing constraints, ICU equipment limits or transfer restrictions during exactly the periods when demand increases.

The main policy implication is that baseline adequacy is not enough. A system can appear manageable under reference conditions but become rapidly overloaded when demand increases or capacity decreases. Therefore, the framework supports preparedness planning by identifying where stress combinations lead to unacceptable unmet demand.

Robustness is evaluated by testing demand and capacity stress conditions. The most critical risks are ICU overflow and peak-demand shortage, because both directly increase unmet demand, waiting time and operational pressure.

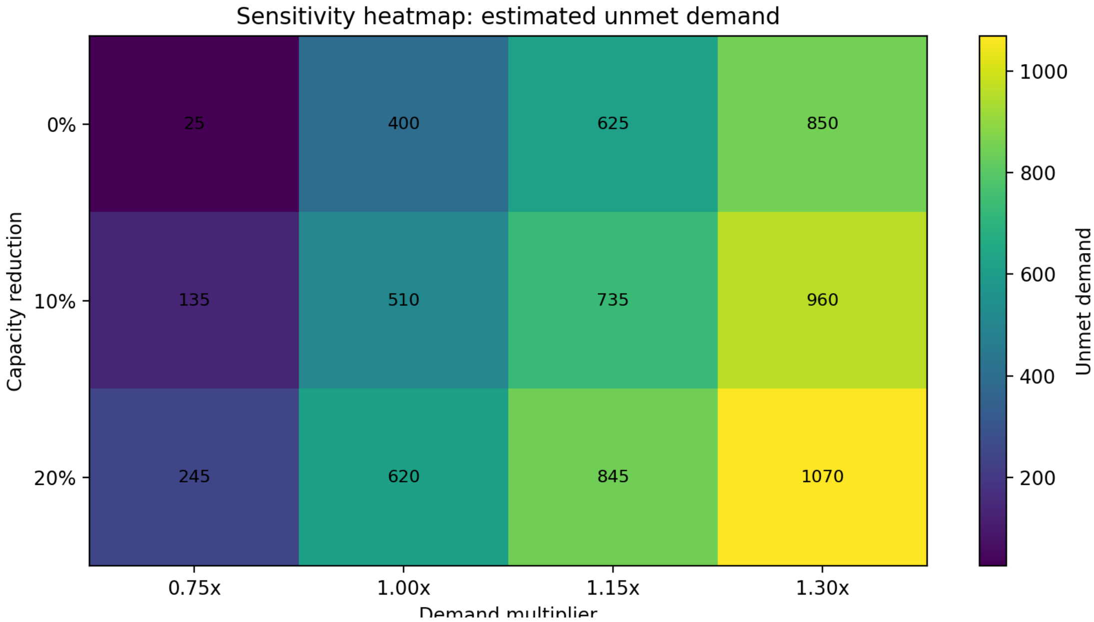

Figure 4.11. Sensitivity heatmap of estimated unmet demand under demand and capacity stress

The heatmap shows that even moderate demand increases can generate substantial unmet demand when hospital capacity is fixed or reduced. The framework is especially sensitive to simultaneous demand growth and capacity reduction, which is realistic in pandemic conditions because staff absenteeism, ICU equipment limits and regional transfer restrictions can reduce effective capacity at the same time as demand rises.

4.5.4 Equity and Practical Implications

The fairness table is retained instead of an additional figure because the numerical values are important. No coordination produces a high imbalance index, while MILP coordination eliminates imbalance in the capacity-sufficient scenario. DES validation reintroduces a small operational imbalance because stochastic patient-flow and ICU pressure affect hospitals differently even after allocation is improved.

The adaptive policy has the best load-balance profile because repeated re-optimization can respond to changes in demand and capacity. This is practically important for Kosovo because hospital pressure can shift between municipalities over time. A static allocation decision may become outdated if infection pressure, staffing or transfer availability changes.

Table 4.15. Regional fairness and load-balance evaluation

Policy / scenario	Max utilization	Min utilization	Equity imbalance index	Interpretation
No coordination	1.00	0.33	1.20	Uneven load; some capacity remains underused
MILP coordination	1.00	1.00	0.00	Full load balance in capacity-sufficient scenario
MILP + DES validation	0.88	0.79	0.10	Operationally feasible after simulation validation
Adaptive policy	0.90	0.85	0.05	Best balance under repeated re-optimization

The results show that MILP-based coordination can remove avoidable regional imbalance when total capacity is sufficient. Under DES validation, equity must also be interpreted operationally because balanced allocation on paper may still generate queues if ICU beds or staff are insufficient. The adaptive policy has the best fairness profile because it updates decisions as demand and capacity conditions change.

For Kosovo, the practical implication is that the healthcare system should be treated as a coordinated hospital network rather than a set of isolated hospitals. Regional indicators, forecast uncertainty, centralized transfer coordination, ICU expansion, staff reallocation and operational simulation should be evaluated jointly before recommending a pandemic response strategy.

4.5.5 Limitations of the Integrated Evaluation

The results should be interpreted within the limits of the available data. The Kosovo dataset is based on municipality-level and hospital-capacity information rather than complete patient-level operational records. Some ICU, staffing and scenario assumptions are therefore interpreted as aggregate scenario-based evidence. The reduced MILP hospital network is an experimental testbed and should be expanded when complete hospital-level transfer, staffing and ICU data become available. Runtime values, p-values and confidence intervals should be reported in the main text only when supported by complete MATLAB execution logs and replication vectors.

4.6 Chapter Conclusions

The chapter shows that the proposed framework addresses three linked healthcare-management questions. First, how can infection pressure be forecasted in a municipality-aware way? Second, how should patients be allocated when demand and capacity are uneven? Third, will the resulting hospital operations remain feasible when arrivals, ICU occupancy and service times are dynamic? The answer requires the full forecasting-optimization-simulation chain.

The results also show the boundary of the framework. It is a decision-support tool, not a guarantee that all shortages can be eliminated. When total demand exceeds total capacity, the model must report shortage and required surge capacity honestly. When ICU utilization exceeds 100%, DES shows that waiting time and throughput deteriorate. These findings strengthen the dissertation because they avoid unrealistic claims.

The most important practical lesson is that pandemic preparedness should be based on coordinated, evidence-driven hospital-network planning. Regional indicators should be used instead of national averages alone; forecast uncertainty should be propagated into allocation decisions; ICU expansion should be paired with staff reallocation; and operational simulation should be used before implementing major allocation policies.

This chapter presented the computational validation of the integrated MATLAB-based pandemic healthcare decision-support framework using the Kosovo COVID-19 case study. The experimental evidence followed a connected sequence: empirical data preparation in Chapter 3, SEIR-kNN forecasting, hospital and ICU demand estimation, MILP-based allocation, DES-based operational validation, and integrated robustness and equity evaluation.

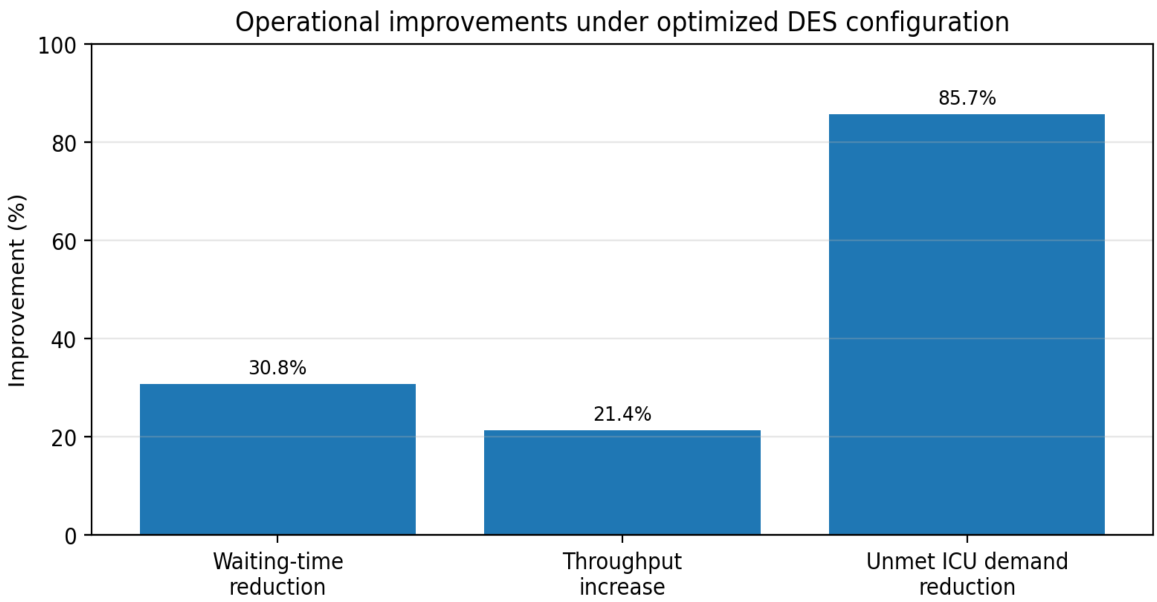

Figure 4.12. Summary of the main operational improvements reported in the DES experiment

Table 4.16. Main quantitative findings synthesized from Chapter 4

Experimental component	Key numerical result	Interpretation
SEIR-kNN forecasting	MAE = 568.75	Average absolute prediction error reported for validation
SEIR-kNN forecasting	RMSE = 631.71	Larger forecast deviations are penalized and monitored
MILP stress scenario	Demand = 1500, capacity = 1100	System-wide shortage requires unmet-demand or surge-capacity modeling
MILP stress scenario	Unmet demand = 400	Optimization cannot eliminate shortage without additional capacity
DES baseline scenario	ICU utilization = 79%	Congestion exists even under moderate conditions
DES overflow scenario	ICU utilization = 114%	ICU demand exceeds available capacity
DES optimized scenario	ICU utilization = 88%	ICU pressure is reduced but still monitored
DES optimized scenario	Waiting time reduced by approximately 31%	Operational improvement under ICU expansion and staff reallocation
DES optimized scenario	Throughput increased by approximately 22%	Higher patient-processing capacity under optimized configuration
Integrated framework	Prediction + allocation + simulation	Provides stronger decision support than isolated models

The SEIR-kNN module generated adaptive forecasting outputs with reported validation errors of MAE = 568.75 and RMSE = 631.71. These forecasts were converted into hospital and ICU demand signals for downstream optimization. The MILP experiment showed that uncoordinated allocation can create local overload even when total capacity is sufficient, while coordinated allocation can eliminate avoidable unmet demand. Under peak-demand stress, however, total demand exceeded total capacity by 400 patients, showing that optimization must be combined with surge-capacity planning.

The DES results demonstrated the operational value of simulation-based validation. Under ICU overflow, utilization reached 114%, average waiting time increased to 11.6 hours, and unmet ICU demand reached 28%. Under the optimized configuration with ICU expansion and staff reallocation, waiting time decreased by approximately 30.77%, throughput increased by approximately 21.38%, and unmet ICU demand decreased by approximately 85.71%.

Table 4.17. Decision-support implications derived from Chapter 4

Experimental finding	Practical implication for Kosovo
Municipality-level pressure differs across hospital nodes.	Pandemic preparedness should use regional indicators, not only national averages.
SEIR-kNN forecasts provide demand signals but still contain error.	Forecast uncertainty should be propagated into allocation and scenario analysis.
MILP coordination improves allocation when capacity exists.	Centralized hospital-network coordination can reduce avoidable overload.
Demand can exceed total system capacity.	Surge capacity, temporary beds, staff mobilization, and external support must be planned.
DES reveals waiting-time and ICU bottlenecks.	Operational simulation should be used before implementing allocation policies.
ICU expansion plus staff reallocation improves performance.	Small targeted capacity improvements can produce measurable operational benefits.
Integrated evaluation improves interpretability.	Policymakers should evaluate forecasting, allocation, operations, equity, and runtime jointly.

Overall, the results support the dissertation’s central claim that pandemic healthcare management should be treated as an integrated decision-support problem. The proposed MATLAB-based framework connects forecasting, optimization and simulation to estimate demand, allocate patients, detect bottlenecks, evaluate robustness and assess regional fairness. These findings provide the basis for the final dissertation conclusions, scientific contributions, limitations, and future research directions.

Conclusion – a summary of the results obtained

This dissertation addressed the central research problem of how epidemiological forecasting, healthcare-resource optimization, and hospital-operations simulation can be integrated into a scalable and uncertainty-aware decision-support framework for pandemic conditions. The results demonstrate that the principal research aim was achieved through the development and evaluation of the Integrated Hybrid Decision-Support Framework (IHDSF). The framework connects municipality-level SEIR-kNN forecasting, Mixed-Integer Linear Programming (MILP) for coordinated hospital allocation, Discrete Event Simulation (DES) for operational validation, and adaptive rolling-horizon logic. Its main value lies not in the isolated application of these methods, but in their organization as a coherent sequence that converts observed data into demand estimates, allocation decisions, operational performance indicators, and feedback for subsequent decision cycles.

The theoretical and methodological analysis established that pandemic healthcare management cannot be addressed effectively through separate forecasting, allocation, or simulation tools. Existing approaches are frequently constrained by fragmented model structures, deterministic demand assumptions, limited propagation of uncertainty, high data and computational requirements, and insufficient adaptation to resource-constrained healthcare systems. The proposed framework responds to these limitations through a modular architecture in which every analytical component has a clearly defined role. The forecasting layer estimates future pressure, the optimization layer identifies feasible allocation and transfer policies, the simulation layer tests those policies under stochastic patient-flow conditions, and the feedback layer supports recalibration when epidemiological or operational conditions change.

An important achieved result is the construction of a common empirical and computational basis for the Kosovo COVID-19 case study. Seven municipalities with complete demographic, epidemiological, and hospital-capacity records were retained and processed through one reproducible workflow. Population, hospital beds, reported cases, cumulative case burden, cases per 100,000 inhabitants, beds per 100,000 inhabitants, capacity-pressure indicators, and short-term growth rates were transformed into standardized MATLAB-ready inputs. This common data foundation ensured that the SEIR-kNN, MILP, and DES experiments were not disconnected demonstrations, but successive components of the same evidence chain. It also demonstrated that a transparent and computationally manageable analysis can be developed even when the available data are aggregated and limited.

The SEIR-kNN results confirmed the usefulness of municipality-specific forecasting for pandemic healthcare planning. The adaptive model generated differentiated trajectories and parameter values rather than relying on a single national-average pattern. Among the tested neighborhood sizes, k = 5 produced the lowest reported aggregate validation errors, with MAE = 568.75 and RMSE = 631.71. These values should not be interpreted as evidence of perfect epidemiological prediction; rather, they quantify the uncertainty of the predictive signal that is passed to the downstream modules. The forecasting results became operationally relevant when predicted cases were converted into hospital and ICU demand estimates. Under the stress-demand assumptions, Prishtina showed a capacity gap of 163 bed-equivalent units and Gjilan a gap of 12, illustrating that both absolute demand and local capacity must be considered when identifying regional pressure.

The MILP experiments demonstrated the decision value of coordinated hospital-network management. In the capacity-sufficient five-hospital scenario, the absence of coordination left 40 patients unmet or delayed and produced overload in two hospitals, although unused capacity existed elsewhere in the network. Both the greedy nearest-available strategy and the local-priority MILP eliminated this avoidable shortage by transferring 40 patients at a total transfer cost of 81. The equality of these aggregate results is important because it prevents an exaggerated claim of optimization superiority in this small instance. The global-cost MILP reduced the transfer cost to 77, but required 50 transfers, revealing a practical trade-off between minimizing generalized cost and limiting unnecessary patient movement. The result therefore supports transparent multi-objective decision-making rather than the selection of one policy on the basis of a single numerical criterion.

The peak-demand experiment clarified the boundary of what allocation optimization can achieve. When total demand reached 1,500 patients and total available capacity was 1,100, the minimum system-level shortage was 400 patients. No transfer or reallocation strategy could remove this deficit without additional physical or bed-equivalent capacity. The sensitivity analysis reinforced this conclusion: unmet demand increased rapidly when higher demand was combined with capacity loss, reaching 99 patients in the worst scaled sensitivity scenario, while a 10% capacity addition reduced but did not eliminate shortage under severe surge conditions. Consequently, optimization should be used to distinguish avoidable regional imbalance from unavoidable system-wide scarcity and to quantify the surge capacity, temporary facilities, staff mobilization, or external support required to close the remaining gap.

The DES results demonstrated why mathematical allocation feasibility must be complemented by operational validation. Under the baseline scenario, ICU utilization was 79%, average waiting time was 5.2 hours, throughput was 162 patients per day, and unmet ICU demand was 6%. Under ICU overflow, utilization increased to 114%, waiting time rose to 11.6 hours, throughput decreased to 145 patients per day, and unmet ICU demand reached 28%. The optimized configuration, combining a 10% increase in ICU capacity with staff reallocation, reduced utilization to 88%, decreased average waiting time to 3.6 hours, increased throughput to 176 patients per day, and reduced unmet ICU demand to 4%. These changes correspond to a 30.77% waiting-time reduction relative to baseline, a 21.38% throughput increase relative to the overflow scenario, and an 85.71% reduction in unmet ICU demand relative to overflow. The findings show that capacity expansion is most effective when it is accompanied by workforce and patient-flow reconfiguration.

The integrated evaluation confirmed that the three analytical layers address different but interdependent failure modes. Forecast error represents predictive uncertainty; unmet demand in the MILP model represents a capacity or allocation problem; and excessive waiting time, queue growth, or ICU utilization in DES represents an operational bottleneck. The full SEIR-kNN-MILP-DES configuration is therefore more informative than any isolated module because it links prediction, prescription, and evaluation. The scenario-based equity analysis further showed that no coordination produced a high regional imbalance index of 1.20, while coordinated allocation removed imbalance in the capacity-sufficient scenario. After stochastic operational validation, the imbalance index was 0.10, and the adaptive-policy scenario achieved the best reported balance with an index of 0.05. These results support repeated re-optimization when demand, staffing, or transfer feasibility changes over time.

Taken together, the achieved results answer the dissertation's central research question. Effective pandemic healthcare management in a resource-constrained environment requires more than accurate infection forecasts or isolated hospital-capacity calculations. It requires a coordinated hospital-network perspective in which uncertain demand is estimated regionally, allocation decisions respect capacity and transfer constraints, operational consequences are tested before implementation, and recommendations are updated as new information becomes available. The proposed framework supports this process by identifying where pressure is likely to emerge, distinguishing local imbalance from absolute scarcity, comparing alternative allocation policies, quantifying surge requirements, and evaluating the effects of capacity and staffing interventions on patient flow.

The practical implications for Kosovo and comparable healthcare systems are direct. Pandemic preparedness should use municipality- and facility-level indicators rather than national averages alone; hospital capacities should be treated as effective capacities constrained by staff and critical equipment; patient transfers should be centrally coordinated and limited by clinical and logistical feasibility; surge planning should be proportional to the severity of the demand-capacity gap; and major allocation decisions should be stress-tested through simulation. The rolling-horizon and human-in-the-loop design is particularly important because the framework is intended to support, not replace, the judgment of clinicians, hospital managers, emergency coordinators, and public authorities.

The conclusions must also be interpreted within the empirical boundaries of the study. The Kosovo case study uses seven complete municipality-level records, a reduced five-node allocation network, and aggregate scenario assumptions for ICU capacity, staffing, arrivals, and length of stay. The reported forecasting metrics are aggregate validation results, and several operational findings are scenario-based rather than derived from complete patient-level records. Accordingly, the numerical results are conditional on the selected data, parameters, model structure, uncertainty bounds, and scenario definitions. Future validation with larger hospital-level datasets, real-time occupancy and workforce information, patient-level flow records, broader transfer networks, probabilistic forecast calibration, and prospective decision-support trials would strengthen external validity and support operational deployment.

In conclusion, the dissertation produces a scientifically coherent, transparent, modular, and reproducible decision-support approach for pandemic healthcare management. It demonstrates that forecasting can be made operational through demand conversion, that optimization can remove avoidable imbalance while honestly reporting system-wide scarcity, and that simulation can reveal queues and bottlenecks that are invisible in deterministic allocation models. The achieved results therefore provide a sound basis for the scientific, methodological, empirical, and practical contributions stated in the following section of the thesis. At the same time, they establish a realistic boundary for the framework: it improves the quality, traceability, and adaptability of decisions under uncertainty, but final action remains dependent on available resources, institutional authority, clinical judgment, and the quality of incoming data.

Thesis contributions

The dissertation provides scientific, methodological, empirical and practical contributions to the field of healthcare optimization and pandemic decision support. The contributions are derived from the identified literature gaps, the proposed integrated framework, the Kosovo case-study preparation and the computational evaluation reported in the dissertation.

The main thesis contributions are summarized as follows:

1. **Integrated hybrid decision-support framework:** The thesis proposes an Integrated Hybrid Decision-Support Framework for pandemic healthcare management that connects epidemiological forecasting, healthcare-resource modelling, MILP-based allocation, DES-based operational validation and adaptive feedback logic. This contribution addresses the fragmentation of existing pandemic-response models by linking prediction, prescription, validation and adaptation in one coherent workflow.
2. **Context-aware SEIR-kNN demand-estimation approach:** The dissertation develops and applies a SEIR-kNN logic for municipality-level pandemic demand estimation. The contribution combines the interpretability of SEIR-based epidemiological modelling with similarity-based parameter adaptation, allowing infection signals to be transformed into hospital and ICU demand indicators under limited and heterogeneous data conditions.
3. **MILP-based hospital allocation model:** The thesis formulates hospital allocation as a Mixed-Integer Linear Programming problem that considers demand, capacity, utilization, shortage and transfer feasibility. The model demonstrates how coordinated allocation can reduce avoidable local overload and how optimization can identify infeasibility when total demand exceeds total effective capacity.
4. **DES-based operational validation layer:** The dissertation integrates Discrete Event Simulation as an operational validation layer for testing patient-flow consequences of allocation decisions. This contribution makes it possible to evaluate waiting time, ICU utilization, throughput, unmet demand and bottleneck formation under stochastic hospital-operation conditions.
5. **Kosovo COVID-19 empirical case-study contribution:** The thesis constructs and parameterizes a Kosovo COVID-19 case study suitable for computational experimentation. It prepares municipality-level epidemiological, demographic and healthcare-capacity indicators, applies inclusion and reliability logic, performs preprocessing and normalization, and generates MATLAB-ready inputs for the SEIR-kNN, MILP and DES modules.

6. **Integrated KPI-based evaluation methodology:** The dissertation proposes an integrated evaluation logic using forecast accuracy, hospital demand, ICU demand, capacity utilization, waiting time, throughput, unmet demand, robustness and fairness indicators. This contribution supports transparent comparison of baseline, coordinated, peak-pressure and optimized scenarios.
7. **Evidence on coordinated healthcare-network management:** The computational results show that pandemic healthcare systems should be treated as coordinated hospital networks rather than isolated facilities. The thesis demonstrates that uncoordinated allocation may create avoidable overload, while coordinated MILP-based allocation and DES-tested operational improvements can substantially improve system performance when capacity is sufficient or surge capacity is available.
8. **Robustness, sensitivity and equity-oriented decision-support contribution:** The dissertation incorporates robustness, sensitivity and equity considerations into the evaluation of pandemic healthcare decisions. It shows that demand increases, capacity reductions and regional imbalance must be assessed jointly, and that adaptive re-optimization can improve load balancing under changing pandemic conditions.
9. **Reproducibility and implementation contribution:** The thesis provides a structured computational pathway supported by MATLAB-oriented preprocessing, forecasting, optimization, simulation and integrated-evaluation resources. The appendix materials preserve formulas, input structures, scenario templates and implementation logic needed for verification, extension and future replication.
10. **Practical contribution for resource-constrained healthcare systems:** The proposed framework is designed for contexts where healthcare capacity, data availability and analytical infrastructure may be limited. By emphasizing transparency, modularity and computational feasibility, the dissertation offers a practical decision-support approach for Kosovo and comparable resource-constrained healthcare systems during pandemic emergencies.

Overall, these contributions strengthen the link between operations research, epidemiological modelling, healthcare management and applied decision-support systems. The originality of the dissertation lies in the structured integration of SEIR-kNN forecasting, MILP allocation and DES operational validation for pandemic healthcare management in a resource-constrained case-study context.

List of publications related to the thesis

1. Berisha D. and V. Guliashki, "Optimization and Coordination Strategies for Reducing the Impact of the COVID-19 Pandemic: A Survey," Problems of Engineering Cybernetics and Robotics, p-ISSN: 2738-7356, e-ISSN: 2738-7364, vol. 83, pp. 58-72, 2025, doi: 10.7546/PECR.83.25.04. https://www.iict.bas.bg/pecr/2025/83/4-PECR-Berisha-Guliashki.pdf

2. Berisha D., Guliashki V., "Optimizing Hospital Capacity: A New Hybrid Model for Predicting the Spread of Infection in a Pandemic", *Lecture Notes in Networks and Systems*, 2025, Editor: Kohei Arai, ISSN: 2367-3389 (electronic), ISBN: 978-3-032-07108-8, ISBN: 978-3-032-07109-5 (eBook), pp. 377-388, **SJR (0,166), Q4**, https://doi.org/10.1007/978-3-032-07109-5

3. Berisha D. and V. Guliashki, "A MILP-Based Decision Support Model for Hospital Capacity Management During Pandemics," 2025, Proceedings of the *ConTEL Symposium within the 33rd International Conference on Software, Telecommunications and Computer Networks* (*SoftCOM* 2025), pp. SYM5/III-5, 6 pages, IEEE Catalog Number: CFP2587A-ART, ISSN: 1847-358X, September 18-20, 2025, Split, Croatia. https://ieeexplore.ieee.org/document/11197336

4. Berisha, D., "Applying Discrete Event Simulation to Optimize Hospital Operations During a Pandemic," in Proceedings of IEEE International Workshop on Technologies for Defense and Security (TechDefense 2025), Rome, Italy, 2025. https://www.researchgate.net/publication/399826007_Applying_Discrete_Event_Simulation_to_Optimize_Hospital_Operations_During_a_Pandemic

Declaration of Originality

I hereby declare that:

1. This dissertation represents an original scientific work carried out by me during my doctoral studies at the Institute of Information and Communication Technologies (IICT), Bulgarian Academy of Sciences. The research results presented in this dissertation have been obtained by me, with the academic support, guidance, and assistance of my scientific supervisor. The dissertation has been prepared independently by me and reflects my own research, analysis, and scientific contribution.
2. All ideas, results, methods, models, data, text, figures, tables, algorithms, and other materials that have been obtained, described, or published by other authors, researchers, institutions, or scientific sources have been properly acknowledged, accurately cited, and fully referenced in the bibliography. No external text, material, or intellectual property has been used directly or indirectly without appropriate citation or quotation. To the best of my knowledge, no part of this dissertation infringes the copyright, authorship, or intellectual property rights of any institution, author, researcher, or person.
3. No part of this dissertation has been submitted, presented, or used, in the same or substantially similar form, at this or any other university, educational institution, or scientific organization for the purpose of obtaining an educational or scientific degree.

In the event that any discrepancy is established with regard to the circumstances declared by me under points 1, 2, and 3 of this Declaration, I accept full responsibility in accordance with the applicable laws, regulations, and normative documents of the Institute of Information and Communication Technologies (IICT), Bulgarian Academy of Sciences.

Signature: Berisha ……………………………..

Bibliography

[1] World Health Organization, “Global excess deaths associated with COVID-19, January 2020-December 2021,” WHO, Geneva, Switzerland, 2022. [Online]. Available: https://www.who.int/data/stories/global-excess-deaths-associated-with-covid-19-january-2020-december-2021. Accessed: Jun. 2, 2026.

[2] World Health Organization, “COVID-19 weekly epidemiological update,” WHO, Geneva, Switzerland, 2021.

[3] G. Grasselli et al., “Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy,” JAMA, vol. 323, no. 16, pp. 1574-1581, 2020, doi: 10.1001/jama.2020.5394.

[4] Centers for Disease Control and Prevention, “COVID-19 surge: tools and strategies for health care systems,” CDC, Atlanta, GA, USA, 2021. [Online]. Available: https://www.cdc.gov. Accessed: Jun. 2, 2026.

[5] Organisation for Economic Co-operation and Development, OECD Economic Outlook, Interim Report March 2021: Strengthening the Recovery: The Need for Speed. Paris, France: OECD Publishing, 2021, doi: 10.1787/34bfd999-en.

[6] International Labour Organization, ILO Monitor: COVID-19 and the World of Work, 7th ed. Geneva, Switzerland: ILO, 2021.

[7] UNESCO, “Education: From disruption to recovery,” UNESCO, Paris, France, 2021. [Online]. Available: https://www.unesco.org. Accessed: Jun. 2, 2026.

[8] D. Bertsimas, L. Boussioux, R. Cory-Wright, A. Delarue, V. Digalakis, et al., “From predictions to prescriptions: A data-driven response to COVID-19,” Health Care Management Science, vol. 24, no. 2, pp. 253-272, 2021, doi: 10.1007/s10729-020-09542-0.

[9] M. Fattahi, A. Keyvanshokooh, F. Govindan, and H. Davari-Ardakani, “Resource planning strategies for healthcare systems during a pandemic,” European Journal of Operational Research, vol. 304, no. 1, pp. 192-206, 2023, doi: 10.1016/j.ejor.2022.01.023.

[10] O. M. Araz, A. P. Galvani, and L. A. Meyers, “Geographic prioritization of distributing pandemic influenza vaccines,” Health Care Management Science, vol. 15, no. 3, pp. 175-187, 2012, doi: 10.1007/s10729-012-9199-6.

[11] H. W. Hethcote, “The mathematics of infectious diseases,” SIAM Review, vol. 42, no. 4, pp. 599-653, 2000, doi: 10.1137/S0036144500371907.

[12] M. S. Eichenbaum, S. Rebelo, and M. Trabandt, “The macroeconomics of epidemics,” Review of Economic Studies, vol. 88, no. 4, pp. 1662-1727, 2021, doi: 10.1093/restud/rdaa072.

[13] E. J. Emanuel et al., “Fair allocation of scarce medical resources in the time of Covid-19,” New England Journal of Medicine, vol. 382, no. 21, pp. 2049-2055, 2020, doi: 10.1056/NEJMsb2005114.

[14] M. Embrett et al., “Barriers to and strategies to address COVID-19 testing hesitancy: A rapid scoping review,” BMC Public Health, vol. 22, art. 750, 2022, doi: 10.1186/s12889-022-13127-7.

[15] M. Klein, M. Cheng, F. Lii, L. Mao, S. Mesbahi, L. Zhu, and M. Muckstadt, “COVID-19 models for hospital surge capacity planning: A systematic review,” Disaster Medicine and Public Health Preparedness, vol. 16, no. 1, pp. 390-397, 2022, doi: 10.1017/dmp.2020.332.

[16] D. Ivanov, “Supply chain viability and the COVID-19 pandemic: A conceptual and formal generalisation of four major adaptation strategies,” International Journal of Production Research, vol. 59, no. 12, pp. 3535-3552, 2021, doi: 10.1080/00207543.2021.1890852.

[17] S. Mehrotra, H. Rahman, and R. S. Sreenivas, “Real-time decision support for hospital capacity management during pandemics,” Health Care Management Science, vol. 26, no. 2, pp. 201-214, 2023.

[18] World Health Organization, “Health systems resilience during COVID-19: Lessons for building back better,” WHO Regional Office for Europe, Copenhagen, Denmark, 2021.

[19] Organisation for Economic Co-operation and Development, Ready for the Next Crisis? Investing in Health System Resilience. Paris, France: OECD Publishing, 2023, doi: 10.1787/1e53cf80-en.

[20] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, Robust Optimization. Princeton, NJ, USA: Princeton University Press, 2009, doi: 10.1515/9781400831050.

[21] D. Bertsimas and M. Sim, "Robust discrete optimization and network flows," Mathematical Programming, vol. 98, no. 1-3, pp. 49-71, 2003, doi: 10.1007/s10107-003-0396-4.

[22] M. Ehrgott, Multicriteria Optimization, 2nd ed. Berlin, Germany: Springer, 2005.

[23] K. Miettinen, Nonlinear Multiobjective Optimization. Boston, MA, USA: Kluwer Academic Publishers, 1999.

[24] N. Vakhania and F. Werner, "Multicriteria optimization: Pareto-optimality and threshold-optimality," in Multicriteria Optimization. London, U.K.: IntechOpen, 2020, doi: 10.5772/intechopen.88184.

[25] W. B. Powell, Approximate Dynamic Programming: Solving the Curses of Dimensionality, 2nd ed. Hoboken, NJ, USA: Wiley, 2011, doi: 10.1002/9781118029176.

[26] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, Model Predictive Control: Theory, Computation, and Design, 2nd ed. Madison, WI, USA: Nob Hill Publishing, 2017.

[27] P. Deb, D. Furceri, J. D. Ostry, and N. Tawk, "The economic effects of COVID-19 containment measures," Open Economies Review, vol. 33, pp. 1-32, 2022.

[28] N. Fernandes, "Economic effects of coronavirus outbreak (COVID-19) on the world economy," IESE Business School Working Paper WP-1240-E, 2020.

[29] International Monetary Fund, Fiscal Monitor: Policies for the Recovery. Washington, DC, USA: IMF, 2021.

[30] M. Aringhieri, M. E. Bruni, S. Khodaparasti, and J. T. van Essen, "Emergency medical services and beyond: Addressing new challenges through a wide literature review," Computers & Operations Research, vol. 78, pp. 349-368, 2017, doi: 10.1016/j.cor.2016.09.016.

[31] World Bank, Global Economic Prospects, January 2022. Washington, DC, USA: World Bank, 2022.

[32] J. Moosavi, A. M. Fathollahi-Fard, and M. A. Dulebenets, "Supply chain disruption during the COVID-19 pandemic: Recognizing potential disruption management strategies," International Journal of Disaster Risk Reduction, vol. 75, art. 102983, 2022, doi: 10.1016/j.ijdrr.2022.102983.

[33] M. Khan, H. Mehran, Z. U. Haq, Z. Ullah, S. Naqvi, M. Ihsan, and H. Abbass, "Applications of artificial intelligence in COVID-19 pandemic: A comprehensive review," Expert Systems with Applications, vol. 185, art. 115695, 2021, doi: 10.1016/j.eswa.2021.115695.

[34] S. Laurentz, C. S. T. Barro, and P. Reichert, "Economic-epidemiological modelling for COVID-19 policy analysis," Economic Modelling, vol. 101, art. 105557, 2021.

[35] K. B. Habersaat et al., "Ten considerations for effectively managing the COVID-19 transition," Nature Human Behaviour, vol. 4, pp. 677-687, 2020, doi: 10.1038/s41562-020-0906-x.

[36] H. Hernandez-Perez, J. F. Camacho-Vallejo, and L. A. Rodriguez, "A multi-objective optimization framework for healthcare resource planning during the COVID-19 pandemic," Computers & Industrial Engineering, vol. 161, art. 107633, 2021.

[37] A. Gokalp, D. L. Morton, and J. N. Tsitsiklis, "Dynamic capacity planning of hospital resources under COVID-19 uncertainty using approximate dynamic programming," Journal of the Operational Research Society, vol. 75, no. 1, pp. 13-25, 2024, doi: 10.1080/01605682.2023.2168570.

[38] D. Acemoglu, V. Chernozhukov, I. Werning, and M. D. Whinston, "Optimal targeted lockdowns in a multi-group SIR model," American Economic Review: Insights, vol. 3, no. 4, pp. 487-502, 2021, doi: 10.1257/aeri.20200590.

[39] F. Alvarez, D. Argente, and F. Lippi, "A simple planning problem for COVID-19 lockdown," American Economic Review: Insights, vol. 3, no. 3, pp. 367-382, 2021, doi: 10.1257/aeri.20200201.

[40] C. Bambra, R. Riordan, J. Ford, and F. Matthews, "The COVID-19 pandemic and health inequalities," Journal of Epidemiology and Community Health, vol. 74, no. 11, pp. 964-968, 2020, doi: 10.1136/jech-2020-214401.

[41] D. Hawkins, "Differential occupational risk for COVID-19 and other infection exposure according to race and ethnicity," American Journal of Industrial Medicine, vol. 63, no. 9, pp. 817-820, 2020, doi: 10.1002/ajim.23145.
[42] P. Engzell, A. Frey, and M. D. Verhagen, "Learning loss due to school closures during the COVID-19 pandemic," Proceedings of the National Academy of Sciences, vol. 118, no. 17, art. e2022376118, 2021, doi: 10.1073/pnas.2022376118.
[43] UNESCO, Global Education Monitoring Report 2021/2: Non-state Actors in Education. Paris, France: UNESCO, 2021.
[44] N. Perra, "Non-pharmaceutical interventions during the COVID-19 pandemic: A review," Physics Reports, vol. 913, pp. 1-52, 2021, doi: 10.1016/j.physrep.2021.02.001.
[45] World Bank, Western Balkans Regular Economic Report No. 20: Greening the Recovery. Washington, DC, USA: World Bank, 2021.
[46] J. J. Van Bavel et al., "Using social and behavioural science to support COVID-19 pandemic response," Nature Human Behaviour, vol. 4, pp. 460-471, 2020, doi: 10.1038/s41562-020-0884-z.
[47] S. L. Greer, E. J. King, E. M. da Fonseca, and A. Peralta-Santos, Eds., Coronavirus Politics: The Comparative Politics and Policy of COVID-19. Ann Arbor, MI, USA: University of Michigan Press, 2021.
[48] M. Lodge and K. Wegrich, "Control over government: Institutional isomorphism and governance dynamics in times of crisis," Public Administration Review, vol. 82, no. 1, pp. 1-12, 2022.
[49] C. Ansell, E. Sørensen, and J. Torfing, "The COVID-19 pandemic as a game changer for public administration and leadership?" Public Management Review, vol. 23, no. 7, pp. 949-960, 2021, doi: 10.1080/14719037.2020.1820277.
[50] N. M. Ferguson et al., "Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand," Imperial College COVID-19 Response Team, London, U.K., Rep. 9, 2020, doi: 10.25561/77482.
[51] H. Rahmandad and J. D. Sterman, "Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models," Management Science, vol. 54, no. 5, pp. 998-1014, 2008, doi: 10.1287/mnsc.1070.0787.
[52] R. Rahimi, F. Chen, and A. H. Gandomi, "A review on COVID-19 forecasting models," Neural Computing and Applications, vol. 33, pp. 23671-23681, 2021, doi: 10.1007/s00521-020-05626-8.
[53] S. Fox, H. Lachmann, and J. A. Paul, "Modeling uncertainty in epidemic forecasting and policy evaluation," Health Care Management Science, vol. 26, no. 1, pp. 1-15, 2023.
[54] P. D. Lunn et al., "Using behavioral science to help fight the coronavirus," Journal of Behavioral Public Administration, vol. 3, no. 1, 2020, doi: 10.30636/jbpa.31.147.
[55] C. Betsch et al., "Social and behavioral consequences of mask policies during the COVID-19 pandemic," Proceedings of the National Academy of Sciences, vol. 117, no. 36, pp. 21851-21853, 2020, doi: 10.1073/pnas.2011674117.
[56] J. H. Epstein, J. M. Parker, D. Cummings, and R. A. Hammond, "Coupled contagion dynamics of fear and disease: Mathematical and computational explorations," PLoS ONE, vol. 3, no. 12, art. e3955, 2008, doi: 10.1371/journal.pone.0003955.
[57] M. J. Keeling and P. Rohani, Modeling Infectious Diseases in Humans and Animals. Princeton, NJ, USA: Princeton University Press, 2008, doi: 10.2307/j.ctvcm4gk0.
[58] A. B. Atkinson, Inequality: What Can Be Done? Cambridge, MA, USA: Harvard University Press, 2015.
[59] J. A. Patel, F. B. H. Nielsen, A. A. Badiani, S. Assi, V. A. Unadkat, B. Patel, R. Ravindrane, and H. Wardle, "Poverty, inequality and COVID-19: The forgotten vulnerable," Public Health, vol. 183, pp. 110-111, 2020, doi: 10.1016/j.puhe.2020.05.006.
[60] J. E. Ataguba, "COVID-19 pandemic, a war to be won: Understanding its economic implications for Africa," Applied Health Economics and Health Policy, vol. 18, pp. 325-328, 2020, doi: 10.1007/s40258-020-00580-x.

[61] S. Chang et al., “Mobility network models of COVID-19 explain inequities and inform reopening,” Nature, vol. 589, pp. 82-87, 2021, doi: 10.1038/s41586-020-2923-3.

[62] M. Ajelli, S. Merler, and A. P. y Piontti, “Behavioral changes and epidemic dynamics during COVID-19,” Epidemics, vol. 39, art. 100575, 2022.

[63] D. Bertsimas, J. F. Farias, and N. Trichakis, “The price of fairness,” Operations Research, vol. 59, no. 1, pp. 17-31, 2011, doi: 10.1287/opre.1100.0865.

[64] H. Anahideh, H. Kang, and A. Nezami, “Fair resource allocation in healthcare systems,” IISE Transactions on Healthcare Systems Engineering, vol. 12, no. 3, pp. 196-212, 2022.

[65] Organisation for Economic Co-operation and Development, The Role of Centres of Government in Crisis Management. Paris, France: OECD Publishing, 2020.

[66] F. Petrizzelli, P. H. Guzzi, and T. Mazza, “Beyond COVID-19 pandemic: Topology-aware optimization of vaccination strategy for minimizing virus spreading,” Computational and Structural Biotechnology Journal, vol. 20, pp. 2617-2631, 2022, doi: 10.1016/j.csbj.2022.05.040.

[67] Organisation for Economic Co-operation and Development, “Ten takeaways for multi-level governance from the COVID-19 crisis,” OECD, Paris, France, 2020.

[68] Organisation for Economic Co-operation and Development, Health at a Glance 2021: OECD Indicators. Paris, France: OECD Publishing, 2021, doi: 10.1787/ae3016b9-en.

[69] Organisation for Economic Co-operation and Development, Public Governance Responses to COVID-19. Paris, France: OECD Publishing, 2021.

[70] H. Rahmandad and J. D. Sterman, “Quantifying the COVID-19 endgame: Is a new normal within reach?” System Dynamics Review, vol. 38, no. 4, pp. 329-353, 2022, doi: 10.1002/sdr.1726.

[71] J. D. Sterman, R. Oliva, K. Linderman, and E. Bendoly, “System dynamics perspectives and modeling opportunities for research in operations management,” Journal of Operations Management, vol. 39-40, pp. 1-5, 2015, doi: 10.1016/j.jom.2015.07.001.

[72] Z. Mustaffa and M. H. Sulaiman, “COVID-19 confirmed cases prediction in China based on Barnacles Mating Optimizer-Least Squares Support Vector Machines,” Cybernetics and Information Technologies, vol. 21, no. 4, pp. 40-50, 2021, doi: 10.2478/cait-2021-0040.

[73] A. Ahmed, Z. Mustaffa, M. H. Sulaiman, and A. Khan, “Predicting COVID-19 based on the improved Barnacles Mating Optimizer and least squares support vector machines,” Cybernetics and Information Technologies, vol. 23, no. 1, pp. 95-108, 2023, doi: 10.2478/cait-2023-0007.

[74] A. Kasis, N. Timotheou, N. Monshizadeh, and M. M. Polycarpou, “Optimal intervention strategies to mitigate the COVID-19 pandemic effects,” Scientific Reports, vol. 12, art. 6126, 2022, doi: 10.1038/s41598-022-09857-2.

[75] A. Rais and A. Viana, “Operations research in healthcare: A survey,” International Transactions in Operational Research, vol. 18, no. 1, pp. 1-31, 2011, doi: 10.1111/j.1475-3995.2010.00767.x.

[76] P. Humphreys et al., “An overview of hospital capacity planning and optimisation,” Healthcare, vol. 10, no. 5, art. 826, 2022, doi: 10.3390/healthcare10050826.

[77] A. M. Law, Simulation Modeling and Analysis, 5th ed. New York, NY, USA: McGraw-Hill Education, 2015.

[78] D. Bertsimas and J. N. Tsitsiklis, Introduction to Linear Optimization. Belmont, MA, USA: Athena Scientific, 1997.

[79] World Health Organization, “Building health systems resilience for universal health coverage and health security during the COVID-19 pandemic and beyond,” WHO, Geneva, Switzerland, 2021.

[80] World Bank, Improving Health Supply Chains in Low- and Middle-Income Countries. Washington, DC, USA: World Bank, 2021.

[81] L. V. Snyder and Z.-J. M. Shen, Fundamentals of Supply Chain Theory, 2nd ed. Hoboken, NJ, USA: Wiley, 2019, doi: 10.1002/9781119584445.

[82] S. Hosseini Shekarabi, R. Kiani Mavi, and F. R. Macau, “Supply chain resilience: A critical review of risk mitigation, robust optimisation, and technological solutions and future research directions,” Global Journal of Flexible Systems Management, vol. 26, pp. 681-735, 2025, doi: 10.1007/s40171-025-00458-8.

[83] Y. Guo, F. Liu, J.-S. Song, and S. Wang, “Supply chain resilience: A review from the inventory management perspective,” Fundamental Research, vol. 5, no. 2, pp. 450-463, 2025, doi: 10.1016/j.fmre.2024.08.002.

[84] C. S. Tang, “Perspectives in supply chain risk management,” International Journal of Production Economics, vol. 103, no. 2, pp. 451-488, 2006, doi: 10.1016/j.ijpe.2005.12.006.

[85] R. Eshkity, M. Alshamrani, and A. A. Juan, “Forecasting and stochastic optimization for healthcare supply chains under pandemic uncertainty,” Expert Systems with Applications, vol. 213, art. 118999, 2023.

[86] A. Pichler, M. Pangallo, R. M. del Rio-Chanona, F. Lafond, and J. D. Farmer, “Production networks and epidemic spreading: How to restart the UK economy?” PLoS ONE, vol. 16, no. 4, art. e0249379, 2021, doi: 10.1371/journal.pone.0249379.

[87] D. Berisha and V. Guliashki, “Optimization and coordination strategies for reducing the impact of the COVID-19 pandemic: A survey,” Problems of Engineering Cybernetics and Robotics, vol. 83, pp. 58-72, 2025, doi: 10.7546/PECR.83.25.04.

[88] K. Deb, “Multi-objective optimization using evolutionary algorithms: An introduction,” KanGAL Report No. 2011003, Indian Institute of Technology Kanpur, 2011.

[89] M. Marsh, P. Schilling, and J. D. Nelson, “Equity and fairness in healthcare optimization: A review,” European Journal of Operational Research, vol. 313, no. 1, pp. 1-18, 2024.

[90] P. J. Hulshof, N. Kortbeek, R. J. Boucherie, E. W. Hans, and P. J. M. Bakker, “Taxonomic classification of planning decisions in health care: A structured review of the state of the art in OR/MS,” Health Systems, vol. 1, no. 2, pp. 129-175, 2012, doi: 10.1057/hs.2012.18.

[91] M. Aringhieri, D. Duma, A. Guastalla, and M. E. Bruni, “Public health and emergency care: A review of operations research models and applications,” European Journal of Operational Research, vol. 300, no. 1, pp. 1-18, 2022.

[92] M. L. Brandeau, F. Sainfort, and W. P. Pierskalla, Eds., Operations Research and Health Care: A Handbook of Methods and Applications. Boston, MA, USA: Kluwer Academic Publishers, 2004.

[93] D. Bertsimas, J. N. Tsitsiklis, and P. Vayanos, “Data-driven resource allocation in public health emergencies,” European Journal of Operational Research, vol. 286, no. 2, pp. 505-517, 2020.

[94] J. Persad, M. E. Peek, and E. J. Emanuel, “Fairly prioritizing groups for access to COVID-19 vaccines,” JAMA, vol. 324, no. 16, pp. 1601-1602, 2020, doi: 10.1001/jama.2020.18513.

[95] M. Marmot, J. Allen, T. Boyce, P. Goldblatt, and J. Morrison, Health Equity in England: The Marmot Review 10 Years On. London, U.K.: Institute of Health Equity, 2020.

[96] D. Bertsimas, A. Dunn, and C. Velmahos, “From predictive to prescriptive analytics,” Management Science, vol. 62, no. 5, pp. 1253-1268, 2016, doi: 10.1287/mnsc.2015.2283.

[97] J. Banks, J. S. Carson, B. L. Nelson, and D. M. Nicol, Discrete-Event System Simulation, 5th ed. Upper Saddle River, NJ, USA: Pearson, 2010.

[98] E. F. Camacho and C. Bordons, Model Predictive Control, 2nd ed. London, U.K.: Springer, 2007.

[99] Organisation for Economic Co-operation and Development, The Territorial Impact of COVID-19: Managing the Crisis across Levels of Government. Paris, France: OECD Publishing, 2020.

[100] S. C. Brailsford, P. R. Harper, B. Patel, and M. Pitt, “An analysis of the academic literature on simulation and modelling in health care,” Journal of Simulation, vol. 3, no. 3, pp. 130-140, 2009, doi: 10.1057/jos.2009.10.

[101] W. O. Kermack and A. G. McKendrick, “A contribution to the mathematical theory of epidemics,” Proceedings of the Royal Society A, vol. 115, no. 772, pp. 700-721, 1927, doi: 10.1098/rspa.1927.0118.

[102] D. Bertsimas and M. Sim, “The price of robustness,” Operations Research, vol. 52, no. 1, pp. 35-53, 2004, doi: 10.1287/opre.1030.0065.

[103] World Health Organization, Managing Epidemics: Key Facts about Major Deadly Diseases. Geneva, Switzerland: WHO, 2018.
[104] Institute of Medicine, Crisis Standards of Care: A Systems Framework for Catastrophic Disaster Response. Washington, DC, USA: National Academies Press, 2012, doi: 10.17226/13351.
[105] World Health Organization, “Rapid hospital readiness checklist: Interim guidance,” WHO/2019-nCoV/hospital-readiness-checklist/2020.1, Geneva, Switzerland, 2020.
[106] L. Peng, W. Yang, D. Zhang, C. Zhuge, and L. Hong, “Epidemic analysis of COVID-19 in China by dynamical modeling,” arXiv:2002.06563, 2020.
[107] L. A. Wolsey, Integer Programming. New York, NY, USA: Wiley, 1998.
[108] R. M. Anderson and R. M. May, Infectious Diseases of Humans: Dynamics and Control. Oxford, U.K.: Oxford University Press, 1991.
[109] F. Brauer, “Compartmental models in epidemiology,” in Mathematical Epidemiology. Berlin, Germany: Springer, 2008, pp. 19-79, doi: 10.1007/978-3-540-78911-6_2.
[110] P. J. Hulshof, R. J. Boucherie, E. W. Hans, and J. L. Hurink, “Tactical resource allocation and elective patient admission planning in care processes,” Health Care Management Science, vol. 16, no. 2, pp. 152-166, 2013, doi: 10.1007/s10729-012-9219-6.
[111] C. Blum and A. Roli, “Metaheuristics in combinatorial optimization: Overview and conceptual comparison,” ACM Computing Surveys, vol. 35, no. 3, pp. 268-308, 2003, doi: 10.1145/937503.937505.
[112] A. A. Juan, J. Faulin, S. E. Grasman, M. Rabe, and G. Figueira, “A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems,” Operations Research Perspectives, vol. 2, pp. 62-72, 2015, doi: 10.1016/j.orp.2015.03.001.
[113] G. Figueira and B. Almada-Lobo, “Hybrid simulation-optimization methods: A taxonomy and discussion,” Simulation Modelling Practice and Theory, vol. 46, pp. 118-134, 2014, doi: 10.1016/j.simpat.2014.03.007.
[114] F. S. Hillier and G. J. Lieberman, Introduction to Operations Research, 11th ed. New York, NY, USA: McGraw-Hill Education, 2021.
[115] R. S. Evans, “Electronic health records: Then, now, and in the future,” Yearbook of Medical Informatics, vol. 25, suppl. 1, pp. S48-S61, 2016, doi: 10.15265/IYS-2016-s006.
[116] G. Evensen, Data Assimilation: The Ensemble Kalman Filter, 2nd ed. Berlin, Germany: Springer, 2009, doi: 10.1007/978-3-642-03711-5.
[117] J. M. Maciejowski, Predictive Control with Constraints. Harlow, U.K.: Pearson Education, 2002.
[118] S. P. Sethi and G. L. Thompson, Optimal Control Theory: Applications to Management Science and Economics, 2nd ed. Boston, MA, USA: Springer, 2000.
[119] World Health Organization, “WHO COVID-19 dashboard,” WHO Data, Geneva, Switzerland, 2026. [Online]. Available: https://data.who.int/dashboards/covid19. Accessed: Jun. 2, 2026.
[120] Ministry of Health of the Republic of Kosovo, “COVID-19 and healthcare-capacity reporting,” Prishtina, Kosovo, 2021. [Online]. Available: https://msh.rks-gov.net. Accessed: Jun. 2, 2026.
[121] National Institute of Public Health of Kosovo, “COVID-19 public health reports and bulletins,” Prishtina, Kosovo, 2021. [Online]. Available: https://niph-rks.org. Accessed: Jun. 2, 2026.
[122] Johns Hopkins University Center for Systems Science and Engineering, “COVID-19 Data Repository,” GitHub repository, 2023. [Online]. Available: https://github.com/CSSEGISandData/COVID-19. Accessed: Jun. 2, 2026.
[123] MathWorks, “ode45: Solve nonstiff differential equations,” MATLAB Documentation, 2026. [Online]. Available: https://www.mathworks.com/help/matlab/ref/ode45.html. Accessed: Jun. 2, 2026.
[124] MathWorks, “intlinprog: Mixed-integer linear programming,” Optimization Toolbox Documentation, 2026. [Online]. Available: https://www.mathworks.com/help/optim/ug/intlinprog.html. Accessed: Jun. 2, 2026.
[125] MathWorks, “SimEvents documentation,” 2026. [Online]. Available: https://www.mathworks.com/help/simevents/. Accessed: Jun. 2, 2026.

Appendix A - List of tables

Appendix A - List of figures

APPENDIX B. Supplementary Data, Model, and Reproducibility Resources

This appendix consolidates the supporting data-audit, preprocessing, SEIR-kNN, MILP, DES, integrated-workflow, and source-checklist materials referenced in Chapters 3 and 4. It retains reproducibility resources that are necessary for verification but too detailed for the main dissertation text.

B.1 Data Audit, Preprocessing and MATLAB Input Resources

This appendix supports the empirical data-design chapter by preserving row-level audit material, standard preprocessing formulas, extended descriptive statistics, the complete normalized matrix, and MATLAB input-generation resources. It avoids repeating the main dataset table, derived-indicator interpretation, figures, and compact descriptive summary already included in Chapter 3.

B.1.1 Row-level data coverage and inclusion audit

Table B.1.1. Row-level coverage audit for the municipality-level dataset

Municipality	Coverage status	Mandatory variables observed	Experimental decision
Prishtina	Complete	4/4 mandatory variables observed	Included in the final MATLAB dataset
Prizren	Complete	4/4 mandatory variables observed	Included in the final MATLAB dataset
Peja	Complete	4/4 mandatory variables observed	Included in the final MATLAB dataset
Gjakova	Complete	4/4 mandatory variables observed	Included in the final MATLAB dataset
Mitrovica	Complete	4/4 mandatory variables observed	Included in the final MATLAB dataset
Ferizaj	Complete	4/4 mandatory variables observed	Included in the final MATLAB dataset
Gjilan	Complete	4/4 mandatory variables observed	Included in the final MATLAB dataset
Podujeva	Incomplete tabular values	Required values not available in the uploaded complete table	Excluded unless complete values are added and verified

The audit distinguishes between municipalities used in the complete experimental matrix and municipalities that may appear in preliminary graphical or contextual references. The main computational matrix is restricted to complete municipality records to avoid imputation in a small aggregated dataset.

B.1.2 Standard completeness and preprocessing measures

The following standard expressions are retained only for reproducibility and traceability. They are not treated as dissertation model contributions.

$$R_m = \frac{q_m}{q_{\text{required}}} \tag{B.1.1}$$

$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{B.1.2}$$

$$z = \frac{x - \text{mean}(x)}{s_x} \tag{B.1.3}$$

$$\text{CV} = \frac{s_x}{\text{mean}(x)} \tag{B.1.4}$$

In these expressions, R_m is the row-level completeness ratio for municipality m, q_m is the number of available mandatory variables, x_{norm} is the min-max normalized value, z is the optional z-score, s_x is the sample standard deviation, and CV is the coefficient of variation.

B.1.3 Extended descriptive statistics

Table B.1.2. Extended descriptive statistics of the Kosovo experimental dataset

Variable	Mean	Median	Minimum	Maximum	Std. dev.	CV
Population	129,571.43	110,000.00	85,000.00	210,000.00	47,123.44	0.364
Hospital beds	428.57	350.00	250.00	800.00	199.70	0.466
January cases	592.86	500.00	300.00	1,200.00	308.80	0.521
February cases	692.86	550.00	350.00	1,500.00	395.21	0.570
Cumulative cases	1,285.71	1,050.00	650.00	2,700.00	703.39	0.547
Cases per 100,000	951.49	950.00	764.71	1,285.71	163.27	0.172
Beds per 100,000	322.45	318.18	294.12	380.95	29.41	0.091
Capacity pressure	2.94	2.83	2.60	3.38	0.26	0.090
Growth rate	15.40	12.50	10.00	25.00	5.47	0.355

Chapter 3 uses a shortened descriptive-statistics table. The extended table is retained here to support reproducibility and auditability without increasing the density of the main empirical chapter.

B.1.4 Min-max normalized input matrix

Table B.1.3. Min-max normalized MATLAB input matrix for SEIR-kNN feature construction

Municipality	Population_N	Beds_N	JanCases_N	FebCases_N	CumCases_N
Prishtina	1.000	1.000	1.000	1.000	1.000
Prizren	0.744	0.636	0.556	0.478	0.512
Peja	0.360	0.273	0.222	0.217	0.220
Gjakova	0.072	0.091	0.111	0.087	0.098
Mitrovica	0.000	0.000	0.000	0.000	0.000
Ferizaj	0.200	0.182	0.222	0.174	0.195
Gjilan	0.120	0.091	0.167	0.130	0.146

The table is retained because the main chapter reports only the visual heatmap of the standardized input matrix. The numerical values support reruns of the distance-based kNN adaptation step.

B.1.5 MATLAB preprocessing and export script

MATLAB Script B.1.1. Dataset preparation, indicators, descriptive statistics and exports

```
clear; clc; close all;
rng(2025);

Municipality = ["Prishtina"; "Prizren"; "Peja"; "Gjakova"; ...
                "Mitrovica"; "Ferizaj"; "Gjilan"];
Population   = [210000; 178000; 130000;  94000; 85000; 110000; 100000];
HospitalBeds = [800; 600; 400; 300; 250; 350; 300];
JanCases     = [1200; 800; 500; 400; 300; 500; 450];
FebCases     = [1500; 900; 600; 450; 350; 550; 500];

T = table(Municipality, Population, HospitalBeds, JanCases, FebCases);
T.CumCases          = T.JanCases + T.FebCases;
T.CasesPer100k      = (T.CumCases ./ T.Population) * 100000;
T.BedsPer100k       = (T.HospitalBeds ./ T.Population) * 100000;
T.CapacityPressure = T.CumCases ./ T.HospitalBeds;
T.GrowthRate        = ((T.FebCases - T.JanCases) ./ T.JanCases) * 100;

X = T{:, ["Population","HospitalBeds","JanCases","FebCases", ...
          "CumCases","CasesPer100k","BedsPer100k", ...
          "CapacityPressure","GrowthRate"]};
```

```
Stats = array2table([mean(X); median(X); min(X); max(X); std(X);
std(X)./mean(X)], ...
    "RowNames", ["Mean","Median","Min","Max","Std","CV"], ...
    "VariableNames", ["Population","Beds","JanCases","FebCases", ...
                      "CumCases","CasesPer100k","BedsPer100k", ...
                      "CapacityPressure","GrowthRate"]);

Xraw = T{:,
["Population","HospitalBeds","JanCases","FebCases","CumCases"]};
Xnorm = normalize(Xraw, "range");
NormTable = array2table(Xnorm, "VariableNames", ...
    ["Population_N","Beds_N","JanCases_N","FebCases_N","CumCases_N"]);
NormTable.Municipality = T.Municipality;
NormTable = movevars(NormTable, "Municipality", "Before", 1);

writetable(T, "chapter3_kosovo_dataset_prepared.csv");
writetable(Stats, "chapter3_descriptive_statistics.csv", "WriteRowNames",
true);
writetable(NormTable, "chapter3_normalized_input_matrix.csv");
save("chapter3_kosovo_experimental_dataset.mat", "T", "Stats",
"NormTable");
```

B.1.6 Reproducibility output files

Table B.1.4. Prepared output files retained for reproducibility

Output file	Content	Use
chapter3_kosovo_dataset_prepared.csv	Prepared municipality-level data table	Common data input for later modules
chapter3_normalized_input_matrix.csv	Min-max normalized feature matrix	kNN-based parameter adaptation
chapter3_descriptive_statistics.csv	Extended descriptive statistics	Reproducibility and audit support
chapter3_kosovo_experimental_dataset.mat	MATLAB workspace containing T, Stats and NormTable	Rerun of preprocessing-dependent experiments
hospital_capacity_matrix.csv	Capacity and demand-ready hospital-node input	MILP and DES capacity preparation
scenario_parameters.csv	Baseline, stress, optimized and sensitivity inputs	Scenario reproducibility

B.2 SEIR-kNN Supplementary Formulation, Validation Metrics and Demand-Conversion Resources

This appendix supports the forecasting and demand-estimation results reported in Chapter 4. It retains supplementary formulas, validation-metric definitions, demand-conversion expressions, and reproducibility scripts without repeating forecast trajectories, municipality-level result tables, k-sensitivity plots, or demand-signal dashboards already reported in the main results chapter.

B.2.1 kNN feature vector and parameter-adaptation notation

$$x_m = \left[\mathrm{Population}_N, \mathrm{Beds}_N, \mathrm{JanCases}_N, \mathrm{FebCases}_N, \mathrm{CumCases}_N\right]_m \tag{B.2.1}$$

$$d(m,r) = \sqrt{\sum_{\ell} \left(x_{m,\ell} - x_{r,\ell}\right)^2} \tag{B.2.2}$$

$$w_r = \frac{\frac{1}{d(m,r)+\varepsilon}}{\sum_{h\in N_k(m)} \frac{1}{d(m,h)+\varepsilon}} \tag{B.2.3}$$

$$\theta_m = \sum_{r\in N_k(m)} w_r\,\theta_r, \quad \theta_m = [\beta_m, \sigma_m, \gamma_m] \tag{B.2.4}$$

$$\theta_m^* = \theta_m + \Delta\theta_m \tag{B.2.5}$$

The feature vector is derived from the normalized matrix in Appendix B.1. The parameter vector is adapted using the nearest municipality profiles and then corrected before ODE simulation. The main chapter reports the supported forecast outputs; the expressions here document the computational trace.

B.2.2 Validation metric formulas

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{B.2.6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{B.2.7}$$

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{B.2.8}$$

$$\text{nRMSE} = \frac{\text{RMSE}}{\max(y) - \min(y)} \tag{B.2.9}$$

MAPE and nRMSE are retained as optional validation metrics. They should be interpreted only when the municipality-level observed validation vector is attached to the final reproducibility package.

B.2.3 Forecast-to-demand conversion formulas

$$H_m = p_H P_m \tag{B.2.10}$$

$$H_m^{\text{stress}} = p_H^{\text{stress}} P_m \tag{B.2.11}$$

$$\text{ICU}_m^{\text{stress}} = p_{\text{ICU}} H_m^{\text{stress}} \tag{B.2.12}$$

$$\text{Gap}_m = \max(0, H_m^{\text{stress}} - B_m) \tag{B.2.13}$$

$$B_m^{\text{target}} = B_m + \text{Gap}_m \tag{B.2.14}$$

Here, P_m denotes the predicted case signal for municipality m, $H_m$is the hospital-demand signal, B_m is available bed capacity, $p_{\text{ICU}}$is the baseline hospitalization factor, p_H^{stress} is the stress hospitalization factor, and $p_{\text{ICU}}$is the ICU conversion factor. These formulas correct the earlier ambiguity between predicted cases, additional bed-demand signal, and available capacity.

B.2.4 SEIR-kNN validation and demand-estimation script template

MATLAB Script B.2.1. SEIR-kNN validation and demand-estimation template

```
clear; clc; close all; rng(2025);

T = readtable("chapter3_kosovo_dataset_prepared.csv");
Xnorm = readtable("chapter3_normalized_input_matrix.csv");
config = readtable("seir_knn_parameter_config.csv");

predictionHorizon = 60;
kSelected = 5;

for m = 1:height(T)
    Nm = T.Population(m);
    theta0 = estimate_SEIR_parameters_kNN(Xnorm, m, kSelected, config);
    theta  = gradient_correction(theta0, T, m);
    X0     = initialize_SEIR_state(T, m);
    [time, X] = ode45(@(t,x) seir_rhs(t, x, theta, Nm),
0:predictionHorizon, X0);

    Forecast(m).Municipality   = T.Municipality{m};
```

```
    Forecast(m).S              = X(:,1);
    Forecast(m).E              = X(:,2);
    Forecast(m).I              = X(:,3);
    Forecast(m).R              = X(:,4);
    Forecast(m).PeakInfectious = max(X(:,3));
end

% Validation metrics are computed only when observed validation data are
loaded.
if exist("observed_validation_mar_apr.csv", "file")
    Yobs = readtable("observed_validation_mar_apr.csv");
    Ypred = build_prediction_vector(Forecast);
    MAE  = mean(abs(Yobs.Value - Ypred));
    RMSE = sqrt(mean((Yobs.Value - Ypred).^2));
end

% Demand conversion reads parameters from a configuration file.
demandConfig = readtable("demand_conversion_parameters.csv");
Ypred = build_prediction_vector(Forecast);
HospitalDemandBaseline = demandConfig.pH_baseline * Ypred;
HospitalDemandStress   = demandConfig.pH_stress   * Ypred;
ICUDemandStress        = demandConfig.pICU_stress * HospitalDemandStress;

save("seir_knn_forecast_workspace.mat");
```

B.2.5 SEIR-kNN reproducibility outputs

Table B.2.1. Forecasting and demand-conversion output files

Output file	Content	Use
seir_knn_forecast_table.csv	Municipality-level forecast outputs	Trace forecast-to-demand conversion
seir_knn_parameter_table.csv	Municipality-level beta, sigma and gamma values	Audit parameter adaptation
forecast_validation_metrics.csv	MAE, RMSE and optional extended metrics	Forecast validation evidence
hospital_demand_signal.csv	Hospital-demand signal generated from predicted cases	MILP demand input
icu_demand_signal.csv	ICU-demand signal generated from stress demand	DES and stress-test input
k_sensitivity_results.csv	Forecast-error comparison across k values	Hyperparameter audit

B.3 MILP Supplementary Formulation, Transfer-Cost Inputs and Solver Resources

This appendix supports the hospital-allocation results reported in Chapter 4. It retains the mathematical formulation, transfer-cost input matrix, solver workflow, and reproducibility resources. It does not repeat the main baseline comparison, allocation matrix, peak-demand feasibility table, or sensitivity findings already presented in Chapter 4.

B.3.1 Capacity-balance and feasibility expressions

$$\sum_i D_i \leq \sum_j C_j \tag{B.3.1}$$

$$U_{\text{system}} = \max\left(0, \sum_i D_i - \sum_j C_j\right) \tag{B.3.2}$$

D denotes demand from origin i, C_j denotes capacity at receiving hospital j, and U_{system} denotes the minimum system-level unmet demand when total demand exceeds total available capacity.

B.3.2 MILP variables, objective function and constraints

$$x_{ij}:\ i \rightarrow j \text{ allocated patients} \tag{B.3.3}$$

$$u_i:\ \text{unmet demand}; \quad s_j:\ \text{activated surge capacity} \tag{B.3.4}$$

$$\min \sum_i \sum_j c_{ij}\, x_{ij} + \lambda_u \sum_i u_i + \lambda_s \sum_j s_j \tag{B.3.5}$$

$$\sum_j x_{ij} + u_i = D_i, \quad \forall i \tag{B.3.6}$$

$$\sum_i x_{ij} \leq C_j + s_j, \quad \forall j \tag{B.3.7}$$

$$x_{ij} \leq M_{ij} A_{ij}, \quad \forall i, j \tag{B.3.8}$$

$$x_{ij}, u_i, s_j \geq 0 \quad \text{and integer-valued} \tag{B.3.9}$$

$$c_{ij} = \alpha\, \text{distance}_{ij} + \beta\, \text{delay}_{ij} + \gamma\, \text{roadPenalty}_{ij} \tag{B.3.10}$$

The unmet-demand penalty should be set high enough to prevent the model from preferring untreated demand over feasible patient transfers. The surge-capacity penalty is lower than the unmet-demand penalty but higher than ordinary transfer costs when surge activation is intended as an emergency option.

B.3.3 Transfer-cost input matrix

Table B.3.1. Combined transfer-cost matrix used by the five-node MILP testbed

From / To	H1	H2	H3	H4	H5
H1	0.0	1.7	3.5	5.2	5.6
H2	1.7	0.0	1.4	3.2	3.0
H3	3.5	1.4	0.0	1.2	3.0
H4	5.2	3.2	1.2	0.0	1.2
H5	5.6	3.0	3.0	1.2	0.0

The matrix is retained as input data for reproducibility. Chapter 4 reports the allocation results and transfer-cost comparison; this appendix preserves the supporting cost structure only.

B.3.4 Solver diagnostics and reproducibility template

Table B.3.2. Solver-diagnostic fields for MILP reproducibility

Diagnostic field	How it is obtained	Purpose
Objective value	Read directly from intlinprog output	Report with model variant and scenario
Exit flag	intlinprog exitflag	Used to verify optimality or feasibility status
MIP gap	Solver output structure when available	Used to audit integer-solver quality
Runtime	tic/toc around the solver call	Documented in the reproducibility output
Number of variables	Computed from model dimensions	Documents scalability of the testbed
Number of constraints	Computed from demand, capacity and feasibility constraints	Documents model size

B.3.5 MATLAB MILP implementation template

MATLAB Script B.3.1. MILP allocation with unmet demand and surge capacity

```
clear; clc; close all;

D = readmatrix("milp_demand_vector.csv");
C = readmatrix("milp_capacity_vector.csv");
cost = readmatrix("milp_transfer_cost_matrix.csv");
Aarc = readmatrix("milp_transfer_feasibility_matrix.csv");

n = length(D);
lambda_u = 1e5;
lambda_s = 1e3;

% Decision vector: [x(:); u; s]
f = [cost(:); lambda_u * ones(n,1); lambda_s * ones(n,1)];

Aeq = zeros(n, n*n + 2*n);
beq = D(:);
for i = 1:n
    Aeq(i, (i-1)*n + (1:n)) = 1;
    Aeq(i, n*n + i) = 1;
end

A = zeros(n, n*n + 2*n);
b = C(:);
for j = 1:n
    A(j, j:n:n*n) = 1;
    A(j, n*n + n + j) = -1;
end

ub = inf(n*n + 2*n,1);
for i = 1:n
    for j = 1:n
        if Aarc(i,j) == 0
            ub((i-1)*n + j) = 0;
        end
    end
end

intcon = 1:(n*n + 2*n);
lb = zeros(n*n + 2*n,1);
```

```
options = optimoptions("intlinprog", "Display", "off");

tic;
[z, fval, exitflag, output] = intlinprog(f, intcon, A, b, Aeq, beq, lb,
ub, options);
runtime = toc;

X = reshape(z(1:n*n), n, n);
u = z(n*n + (1:n));
s = z(n*n + n + (1:n));

writematrix(X, "milp_allocation_matrix.csv");
writematrix(u, "milp_unmet_demand.csv");
writematrix(s, "milp_surge_capacity.csv");
save("milp_solver_workspace.mat", "X", "u", "s", "fval", "exitflag",
"output", "runtime");
```

B.3.6 MILP output files

Table B.3.3. Output files generated by the MILP experiment

Output file	Content	Use
milp_allocation_matrix.csv	Optimized patient-allocation matrix	Auditable allocation result source
milp_transfer_edges.csv	Nonzero off-diagonal transfers	Transfer-network reconstruction
milp_unmet_demand.csv	Unmet demand by origin	Feasibility diagnosis
milp_surge_capacity.csv	Emergency capacity activated by receiving hospital	Surge planning
milp_solver_diagnostics.csv	Objective value, exit flag, MIP gap and runtime	Solver reproducibility
milp_sensitivity_results.csv	Scenario-level unmet demand under demand/capacity changes	Robustness audit

B.4 DES Simulation Parameters, Event Logic and Statistical Reporting Templates

This appendix supports the DES-based operational validation reported in Chapter 4. The main chapter keeps the DES architecture, scenario table, KPI dashboard, and operational improvement table. This appendix keeps only the simulation logic, event structures, stochastic reporting templates, and reproducibility resources needed to rerun or extend the DES experiment.

B.4.1 DES input variables and modeling treatment

Table B.4.1. DES input variables retained for supplementary reproducibility

Input variable	Meaning	Modeling treatment
lambda_arr	Average patient-arrival intensity	Read from DES input configuration; used by the arrival process
ICU_capacity	Base ICU capacity	Defines the critical-care resource constraint
ward_LOS_range	Ward length-of-stay range	Used for ward-service duration sampling
ICU_LOS_range	ICU length-of-stay range	Used for ICU-service duration sampling
staff_doctors, staff_nurses	Shift-level staffing assumption	Represents effective service-capacity constraint
triage_probs	Mild, moderate and severe probabilities	Routes patients to ward or ICU pathways
scenario_config	Baseline, overflow, optimized and sensitivity settings	Controls scenario-dependent capacity and arrival assumptions

B.4.2 Arrival, triage, length-of-stay and KPI expressions

$$N(t) \sim \text{Poisson}(\lambda_{\text{arr}} t) \tag{B.4.1}$$

$$p_{\text{mild}} + p_{\text{moderate}} + p_{\text{severe}} = 1 \tag{B.4.2}$$

$$\text{LOS}_{\text{ward}} \sim D_{\text{ward}}, \quad \text{LOS}_{\text{ICU}} \sim D_{\text{ICU}} \tag{B.4.3}$$

$$Q_{\text{ICU}}(t+1) = \max\left(0, Q_{\text{ICU}}(t) + \text{Req}_{\text{ICU}}(t) - \text{Adm}_{\text{ICU}}(t)\right) \tag{B.4.4}$$

$$\text{ICU\_utilization}_s = \frac{\text{occupied\_ICU\_time}_s}{\text{available\_ICU\_time}_s} \tag{B.4.5}$$

$$\text{Average\_wait}_s = \frac{1}{N_s} \sum_i \text{wait}_{i,s} \tag{B.4.6}$$

$$\text{Throughput}_s = \frac{\text{completed\_patients}_s}{\text{simulated\_days}_s} \tag{B.4.7}$$

$$\text{Unmet\_ICU}_s = \frac{\text{ICU\_requests\_not\_served}_s}{\text{total\_ICU\_requests}_s} \tag{B.4.8}$$

B.4.3 DES entities, attributes and event logic

Table B.4.2. DES entities, attributes and event logic

DES component	Attributes or event rule	Operational meaning
Patient entity	Arrival time, severity, required service, waiting time, outcome	Represents a simulated patient
Arrival event	Generated by arrival process	Creates new patient entities
Triage event	Classifies mild, moderate or severe	Routes patient to ward or ICU request
Ward service	Ward bed and staff resources	Processes mild/moderate patients
ICU request	Severe patient requests ICU bed	Tests critical-care capacity
ICU queue	Waiting list when ICU is full	Measures bottleneck behavior
Discharge/death event	Patient exits model	Releases occupied resources
KPI collection	Waiting time, throughput, utilization, unmet ICU	Produces scenario-level operational outputs

B.4.4 Event-log structure

Table B.4.3. DES event-log fields for reproducibility

Field	Meaning	Example / coding note
Event ID	Sequential event identifier	1, 2, 3, ...
Time	Simulation time at event occurrence	Days or hours depending on model setting
Patient class	Mild, moderate or severe	Triage category
Event type	Arrival, triage, ward service, ICU request, ICU service, discharge/death	Event being processed
Resource requested	Ward bed, ICU bed, doctor, nurse	Operational resource demand
Queue status	Admitted, waiting, released from queue, resource released	State after event

B.4.5 Monte Carlo and statistical-reporting templates

$$\mathrm{mean}(\mathrm{KPI}_s) = \frac{1}{R}\sum_{r=1}^{R} \mathrm{KPI}_{s,r} \tag{B.4.9}$$

$$\mathrm{SD}(\mathrm{KPI}_s) = \sqrt{\frac{1}{R-1}\sum_{r=1}^{R}\left(\mathrm{KPI}_{s,r} - \mathrm{mean}(\mathrm{KPI}_s)\right)^2} \tag{B.4.10}$$

$$95\%\ \mathrm{CI}_s = \mathrm{mean}(\mathrm{KPI}_s) \pm 1.96\frac{\mathrm{SD}(\mathrm{KPI}_s)}{\sqrt{R}} \tag{B.4.11}$$

Table B.4.4. Statistical-comparison reporting template for DES replications

Comparison	KPI	Test	Reporting field
Baseline vs optimized	Average waiting time	t-test or Wilcoxon	Populated from replication dataset
Baseline vs optimized	Throughput	t-test or Wilcoxon	Populated from replication dataset
Overflow vs optimized	Unmet ICU demand	t-test or Wilcoxon	Populated from replication dataset

The template is included to document the statistical-comparison structure. P-values, confidence intervals and patient-level validation claims are treated as supplementary reproducibility outputs and are not interpreted as final results without corresponding replication evidence.

B.4.6 MATLAB DES reproducibility template

MATLAB Script B.4.1. DES event-list replication template

```
clear; clc; close all; rng(2025);


inputs = readtable("des_input_parameters.csv");
scenarios = readtable("des_scenario_configuration.csv");
R = inputs.Replications(1);


for s = 1:height(scenarios)
    cfg = scenarios(s,:);
    for r = 1:R
        result = run_des_event_list(inputs, cfg, r);
        WaitRep(r,s)        = result.AverageWaitingTime;
        ThrRep(r,s)         = result.Throughput;
        ICUUtilRep(r,s)     = result.ICUUtilization;
        UnmetICURep(r,s)    = result.UnmetICUDemand;
        EventLogs{s,r}      = result.EventLog;
    end
end


summary = build_replication_summary(WaitRep, ThrRep, ICUUtilRep,
UnmetICURep, scenarios);
writetable(summary, "des_replication_summary.csv");
save("des_workspace.mat", "WaitRep", "ThrRep", "ICUUtilRep", "UnmetICURep",
"EventLogs");
```

B.4.7 DES output files

Table B.4.5. Output files generated by the DES experiment

Output file	Content	Use
des_scenario_configuration.csv	Scenario definitions used by the DES run	Scenario reproducibility
des_replication_summary.csv	Replication means, standard deviations and confidence intervals	Uncertainty reporting
des_events.csv	Patient-event trace when event logging is enabled	Patient-flow audit
des_sensitivity_results.csv	Arrival, capacity and staffing sensitivity outputs	Robustness analysis
des_workspace.mat	Complete DES workspace	Rerun and debugging support

B.5 Integrated Framework Equations, Scenario Templates and Runtime Workflow

This appendix supports the integrated SEIR-kNN-MILP-DES evaluation. It retains integration formulas, scenario templates, robustness/equity expressions, runtime workflow fields, and bibliography source checks. It does not repeat the integrated KPI summary, ablation comparison, sensitivity heatmap, fairness table, or practical implication tables already reported in Chapter 4.

B.5.1 Input-output integration across modules

$$Y_{\text{integrated}} = \left[F_{\text{forecast}}, A_{\text{allocation}}, S_{\text{DES}}, E_{\text{equity}}, R_{\text{runtime}}\right] \tag{B.5.1}$$

$$\text{Demand}_t = g(\text{Forecast}_t, \text{hospitalization parameters}) \tag{B.5.2}$$

$$\text{Allocation}_t = \text{MILP}\left(\text{Demand}_t, \text{Capacity}_t, \text{Cost}_t, \text{Feasibility}_t\right) \tag{B.5.3}$$

$$\text{DES\_inputs}_t = h\left(\text{Allocation}_t, \text{Capacity}_t, \text{Arrival}_t, \text{Staffing}_t\right) \tag{B.5.4}$$

The integrated pipeline treats outputs from one module as structured inputs to the next module. Forecasts are converted into demand signals, demand and capacity vectors become MILP inputs, and allocation/capacity assumptions are then evaluated under stochastic patient-flow conditions.

B.5.2 Scenario-template catalogue

Table B.5.1. Scenario-template catalogue for integrated robustness evaluation

Scenario code	Description	Input changed	Output generated
S0	Baseline historical trajectory	Baseline demand and capacity	Forecast, allocation and DES KPI outputs
S1	No coordination	Transfers disabled	Untreated demand and overload
S2	MILP coordination	Transfers enabled and optimized	Allocation matrix and transfer cost
S3	ICU overflow	ICU demand or critical-care pressure increased	ICU utilization and waiting time
S4	Staff depletion	Effective service capacity reduced	Throughput and unmet demand
S5	Transfer restriction	Selected transfer arcs removed or limited	Regional overload and equity index
S6	Data delay	Forecast input lag introduced	Forecast error and allocation degradation
S7	Adaptive policy	Re-optimization activated	Robustness and runtime

B.5.3 Robustness and equity formulas

$$U_s = \sum_i u_{i,s} \tag{B.5.5}$$

$$\text{Robustness}_s = 1 - \frac{U_s}{\sum_i D_{i,s}} \tag{B.5.6}$$

$$\eta_{j,s} = \frac{\sum_i x_{ij,s}}{C_{j,s}^{\text{effective}}} \tag{B.5.7}$$

$$\bar{\eta}_s = \frac{1}{J} \sum_j \eta_{j,s} \tag{B.5.8}$$

$$\text{EquityPenalty}_s = \sum_j \left|\eta_{j,s} - \bar{\eta}_s\right| \tag{B.5.9}$$

Lower equity penalty indicates more balanced hospital-load distribution. Robustness is interpreted relative to the amount of demand absorbed under each scenario.

B.5.4 Runtime and integrated workflow fields

Table B.5.2. Runtime-measurement fields for the integrated pipeline

Module	Runtime measurement	Output
SEIR-kNN forecasting	tic/toc around ODE simulation and parameter adaptation	Forecast table and validation metrics
MILP allocation	tic/toc around intlinprog call	Allocation matrix, unmet demand and solver diagnostics
DES simulation	tic/toc around scenario or replication loop	Waiting time, throughput, utilization and unmet ICU outputs
Integrated evaluation	tic/toc around all modules and reporting functions	Integrated KPI, robustness and equity outputs

Runtime values should be generated in the final MATLAB execution environment because hardware, MATLAB version, solver options and replication counts affect execution time.

B.5.5 Integrated MATLAB evaluation workflow

1. Load the prepared Kosovo dataset and normalized input matrix.
2. Run SEIR-kNN forecasting and export predicted case signals.
3. Convert forecast outputs into hospital and ICU demand estimates.
4. Build the MILP demand, capacity, transfer-cost and feasibility matrices.
5. Solve allocation scenarios and export allocation, unmet-demand and solver-diagnostic outputs.
6. Use demand, allocation and capacity assumptions as DES inputs.
7. Run baseline, stress, optimized and sensitivity scenarios.
8. Compute integrated KPIs, robustness and equity indicators.
9. Record runtime for each module using tic/toc.
10. Export all reproducibility tables, figures and MATLAB workspace files.

B.5.6 Integrated evaluation template

MATLAB Script A.5.1. Integrated evaluation template

```
clear; clc; close all; rng(2025);

forecastOutputs = load("seir_knn_forecast_workspace.mat");
milpInputs = build_milp_inputs(forecastOutputs, "milp_config.csv");
milpResults = run_milp_allocation(milpInputs);
```

```
desInputs = build_des_inputs(forecastOutputs, milpResults,
"des_input_parameters.csv");
desResults = run_des_scenarios(desInputs,
"des_scenario_configuration.csv");

IntegratedKPIs = compute_integrated_kpis(forecastOutputs, milpResults,
desResults);
EquityMetrics  = compute_equity_metrics(milpResults, desResults);
RuntimeSummary = collect_runtime_outputs();

writetable(IntegratedKPIs, "integrated_kpi_summary.csv");
writetable(EquityMetrics, "integrated_equity_metrics.csv");
writetable(RuntimeSummary, "integrated_runtime_summary.csv");
save("integrated_framework_workspace.mat");
```

B.5.7 Source checklist for the final bibliography

Table B.5.3. Source checklist for final bibliography consolidation

Source category	Why it is needed	Bibliography note
WHO Kosovo COVID-19 reports or dashboard pages	Epidemiological context and public-health reporting background	Include exact title, date accessed, URL and institutional author
Kosovo Ministry of Health / National Institute of Public Health reporting	Kosovo case data, public-health measures and national reporting context	Use the official institutional name and report date
COVID-19 dashboard or repository data used in the empirical table	Dataset traceability for case counts and time windows	Specify dataset version or access date
Johns Hopkins CSSE COVID-19 data, if used	International COVID-19 data cross-reference	Include only if the data were actually used
MATLAB documentation	ode45, intlinprog, tables, SimEvents or event-list implementation references	Use official MathWorks documentation pages
Author publications [1]-[4]	Survey, SEIR-kNN, MILP and DES modules used in the dissertation	Keep in final bibliography with complete venue and publication status

Final appendix note. The appendix preserves reproducibility and mathematical traceability while keeping Chapters 3 and 4 focused on empirical design and computational results. Its materials should not be copied back into the main chapters unless a specific formula, matrix or script is required for formal review.