International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN(Online): 2347-5552, Volume-12, Issue-2, March 2024
https://doi.org/10.55524/ijircst.2024.12.2.17
Article ID IRP1438, Pages 96-105
www.ijircst.org

Review of Mathematical Optimization and Statistics-based Techniques for Public Health Intervention in India: Balancing Efficiency, Resources, and Policy Goals

Sarthak Goswami

Bachelor of Technology, Department of Electronics and Communication, School of Electronics Engineering, Vellore Institute of Technology, Vellore, India

Correspondence should be addressed to Sarthak Goswami; yosarthakyay@gmail.com

Received 6 March 2024; Revised 20 March 2024; Accepted 29 March 2024

ABSTRACT- The purpose of this academic paper is to present a concise overview of using mathematical optimization and statistics-based methods for public health intervention in India. Through the systematic examination of data, these methods support the efficient use of resources and decision-making based on evidence that aligns with policy objectives. They are vital in addressing constraints related to limited resources and policy goals while maximizing intervention efficiency and effectiveness.

KEYWORDS- Mathematical Optimization, Statistics, Public Health Intervention, Efficiency, Resource Allocation, Modeling, Machine Learning

I. INTRODUCTION

The public health situation in India presents a complex challenge due to limited resources, regional disparities, and evolving health needs. Advanced analytical tools are required to effectively address this complexity and enhance interventions. Mathematical optimization and statistical methods offer valuable perspectives for making data-driven decisions within the context of Indian public health. This article explores their respective benefits, limitations, and the potential for significant impact when these approaches are combined strategically.

We examine the evolution of these techniques over time, from their independent use to the current recognition of hybrid methodology as a superior tool. Case studies will illustrate how optimization, statistics, and hybrid approaches have been practically employed to address major challenges in Indian public health. Furthermore, this paper provides recommendations on selecting the most appropriate analytical tools based on specific problem characteristics and desired outcomes.

Ultimately, this review aims to demonstrate how optimization can facilitate evidence-based decision-making regarding Indian public health by integrating statistical techniques into a hybrid methodology that carefully balances efficiency, resource utilization, and policy goals while enabling more effective, integrated and sustainable interventions with better equity.

II. LITERATURE REVIEW

The origins of optimization in public health can be found in the works of mathematicians such as Bernoulli and Lagrange in the 18th and 19th centuries [1]. They concentrated on maximizing or minimizing particular objective functions, which led to early applications in resource allocation, such as hospital bed management and vaccine distribution. Consider a basic problem of allocating hospital beds where the goal is to minimize the total cost associated with bed usage and vacancy (C):

$$\text{Minimize:} \quad C = \sum_{d \in D} (c_u \cdot E[W_d] + c_v \cdot E[V_d])$$

where: D : The set of all hospital departments
c_u : Cost per unit time a patient waits
c_v : Cost per unit time a bed remains vacant $E[W_d]$: Expected waiting time in department d $E[V_d]$: Expected vacancy time in department d

In the mid-20th century [2], linear programming and other optimization techniques revolutionized decision-making for complex problems with well-defined constraints, becoming essential for optimizing infrastructure planning and resource allocation during national health emergencies. For example, LP could be used to optimize vaccine distribution under constraints of budget and population coverage, as follows:

$$\text{Maximize:} \quad Coverage = \sum_{i \in R} p_i \cdot x_i$$

$$\text{Subject to:} \quad \sum_{i \in R} c_i \cdot x_i \leq Budget, \quad \text{where:}$$

R : Set of all geographic regions
p_i : Proportion of the population in region i
x_i : Number of vaccines allocated to region i
c_i : Cost per vaccine in region i
$Budget$: Total available budget

In the late 19th century, advancements in data analysis and collection contributed to the development of statistics as a supporting field. The pioneering works of Galton and Pearson [3] set the foundation for statistical inference and hypothesis testing, providing valuable understanding of population health trends and disease patterns. For instance, studying the connection between smoking and lung cancer through logistic regression could be expressed as:

$$\text{logit}(P(cancer)) = \beta_0 + \beta_1 \cdot smoking + \beta_2 \cdot age + \ldots$$

where: $P(cancer)$: Probability of developing lung cancer

$\text{logit}(p) = \ln \frac{p}{1-p}$: The logit function

β_0 : Intercept of the model

$\beta_1, \beta_2, \ldots$: Regression coefficients

$smoking$: Individual's smoking status age : Individual's age

The development of statistical software and computing power further fueled advancements in the mid-20th century, enabling researchers to identify relationships between variables using techniques like regression analysis and survival analysis. This contributed to making evidence-based policy decisions; for example, statistical analysis of smoking data played a crucial role in developing effective tobacco control policies [4].

A. Mathematical Optimization Techniques

Mathematical modeling and optimization techniques have played a crucial role in public health research and decision-making [5]. These techniques have been utilized to develop models that uncover the underlying mechanisms of infectious disease outbreaks and assess the effectiveness of intervention strategies [6]. Moreover, mathematical models provide a means to evaluate and optimize control strategies, allowing for quantitative predictions that can be empirically tested through randomized comparative trials [7].1. Furthermore, mathematical modeling has been used to evaluate the impact of public health strategies in optimizing the use of limited resources [8]. The models are then optimized to find the best allocation of resources, such as vaccines or quarantine measures, to minimize the spread and impact of the disease.

1. Linear Programming

Linear programming is a mathematical optimization method commonly used in public health research to address resource allocation problems. Linear programming involves maximizing or minimizing a linear objective function, subject to linear equality or inequality constraints. Linear programming has been used to optimize the allocation of resources in epidemics control, such as determining the optimal distribution of vaccines or medication.

Strengths

- **Efficient resource allocation:** Linear programming allows for efficient allocation of limited healthcare resources, ensuring that interventions are directed to the areas or populations that need them the most.
- **Flexibility:** Linear programming models can accommodate a wide range of constraints and variables, allowing for complex decision-making scenarios to be analyzed and optimized.
- **Easy implementation:** Linear programming models are relatively straightforward to implement and can be solved using existing optimization software or algorithms: Efficient resource allocation within budget constraints, which is especially useful in rural areas.
- **Higher accuracy:** Linear programming models provide a quantifiable approach to decision-making, resulting in more accurate predictions and allocations.
- **Supportive of evidence-based decision-making:** By utilizing mathematical optimization techniques, healthcare policymakers and administrators can make informed decisions based on objective data rather than subjective judgments.
- **6. Model transparency:** Linear programming models provide a transparent framework for decision-making, allowing policymakers and stakeholders to understand the rationale behind resource allocation decisions.
- **7. Wide range of applications:** Linear programming can be applied to a variety of resource allocation problems in healthcare, including workforce planning, facility location optimization, and supply.

Limitations

- **Complex problems may require additional modeling:** In some cases, linear programming may not be sufficient to address complex resource allocation problems.
- **Complexity:** Linear programming models may become computationally intensive and time-consuming when dealing with large-scale healthcare systems or considering multiple objectives.
- **Nonlinearity:** Linear programming models assume linear relationships between variables, which may not capture the complex and non-linear nature of healthcare systems.
- **Ethical considerations:** Linear programming models may prioritize certain populations or interventions based solely on mathematical optimization, without taking into account important ethical considerations such as fairness and equity.

2. Network Optimization

Network optimization is another approach that can be utilized for resource allocation in healthcare. Network optimization involves mapping the relationships and flows between various healthcare facilities, such as hospitals, clinics, and specialized centers, to determine the most efficient allocation of resources within the network. This approach considers factors such as patient demand, travel distance, capacity constraints, and resource availability to optimize the allocation of resources. Network optimization models have been used in various healthcare applications, including hospital bed allocation, workforce planning, and facility location optimization.

Strengths:

- **Efficiency:** Network optimization models can help healthcare systems achieve optimal resource allocation, minimizing waste and maximizing efficiency.
- **Accessibility:** By considering factors such as patient demand and travel distance, network optimization models can help ensure equitable access to healthcare resources.
- **Useful for geographically dispersed settings:** When healthcare resources are spread across different geographical areas, network optimization can be particularly useful in determining the most efficient allocation of resources. This was particularly useful in distributing vaccines in the COVID 19 pandemic properly within countries[9].

Limitations:

- **Computationally expensive for large networks:** Network optimization can become computationally expensive and time-consuming when dealing with large-scale healthcare networks or considering multiple objectives.
- **Data requirements:** Network optimization relies on accurate and up-to date data on patient demand, resource capacity, and travel time estimates.
- **Complexity and non-linearity:** Healthcare systems are inherently complex and dynamic, making it difficult to capture all the intricacies and interdependencies in a single optimization model.

3. Agent-Based Modeling

Agent-based modeling is a computational technique that simulates the behavior and interactions of individual agents within a system to understand how complex phenomena emerge at the macro-level. This approach is particularly useful in healthcare management, as it allows researchers to model the behavior of various stakeholders such as healthcare providers, patients, payers, and policymakers in a realistic and dynamic way. Agent-based modeling has been used in healthcare research to explore the impact of different policies, interventions, and decision-making strategies on healthcare outcomes.

Strengths:

- **Flexibility:** Agent-based modeling allows for the incorporation of a wide range of variables and factors, making it adaptable to various healthcare settings and scenarios.
- **Dynamic and interactive:** Agent-based models capture the dynamic nature of healthcare systems, taking into account the interactions and feedback loops between different agents.
- **Realism:** Agent-based models can simulate the behavior of individual agents with realistic characteristics and decision-making processes, providing a more accurate representation of real-world healthcare scenarios.
- **Quantitative analysis:** Agent-based modeling allows for the integration of quantitative data into the simulation, enabling researchers to evaluate the impact of different interventions or policies on healthcare outcomes.
- **Simulation of individual behavior and disease spread:** Agent-based models can simulate the individual behavior of healthcare providers, patients, and other stakeholders, as well as the spread of diseases within a population. This allows researchers to simulate complex interventions, such as modeling COVID-19 containment strategies. [10]

Limitations:

- **Limited availability of data:** Agent-based modeling relies heavily on data to parameterize the behavior and characteristics of individual agents. However, obtaining the necessary data can be challenging in healthcare research due to privacy concerns and the complexity of capturing real-world behavior.
- **Computationally expensive:** Agent-based modeling can be computationally expensive, especially when simulating large-scale healthcare systems with a high number of agents.
- **Interpreting the results:** Agent-based models can produce complex and extensive outputs, making it difficult to interpret and analyze the results. Furthermore, the accuracy of agent-based models highly depends on the quality of the input information and the underlying assumptions made during model development.

B. Statistical Techniques

Statistical techniques have been widely used in public health research to analyze and interpret data, quantify associations and trends, and identify risk factors and determinants of disease transmission. These techniques involve the application of statistical methods such as regression analysis, survival analysis, and utilizing time-series analysis to extract valuable insights from the data. For example, time series analysis has been used to model the prevalence, morbidity, and mortality of diseases such as COVID-19 in different countries [11].

1. Time Series Analysis

Time series analysis is a statistical method utilized for analyzing and modelling data points collected over time. It is particularly useful in understanding the past and current epidemic patterns of infectious diseases and predicting future dynamics. It involves examining the relationship between variables over time, identifying trends, seasonality, and other patterns. Time series analysis has been utilized in public health research and policy-making to assess the impact of interventions, allocate resources effectively, and make informed decisions during disease outbreaks [12]. One of the commonly used time series analysis methods is exponential smoothing. Exponential smoothing is a technique used for forecasting, where varying weights are allocated to previous observations, placing greater emphasis on more recent data points. Other popular time series analysis methods include autoregressive integrated moving average models, which incorporate the dependencies between consecutive observations and the lagged effects of a variable.

Strengths:

- **Accurate Forecasting:** Time series analysis allows for the accurate forecasting of future trends and patterns based on past data. For instance, this approach contributed to a 20 reduction in dengue cases in Chennai.
- **Model Flexibility:** Time series models can be adapted and customized to fit different datasets and specific research questions related to infectious diseases prevalent in India, allowing researchers to tailor interventions specific for the region.
- **Interpretability:** Time series models provide interpretable results that can help researchers and policymakers understand the underlying dynamics of disease spread or mortality trends, aiding in formulating evidence-based intervention strategies within India's public healthcare system.
- **Quantitative Analysis:** Time-series analysis provides a quantitative framework for analyzing prevalence rates or mortality patterns to outbreaks like COVID-19, enabling identification of statistical relationships unique to India which then aids policy-making decisions.
- **Insightful Policy Guidance:** Time series analysis can

inform policy decisions by providing insights into the effectiveness of interventions, allocation of resources, and identification of high priority areas for intervention.
- **Real-time Monitoring:** Time series analysis allows for real-time monitoring of epidemic patterns, enabling early detection of outbreaks and timely implementation of control measures.

Limitations:

- **Limited Optimization:** Time series analysis models are not explicitly designed to optimize resource allocation or intervention strategies for public healthcare.
- **Complexity:** It assumes that the underlying pattern of the data remains stationary over time, which may not always hold true in real-world scenarios including public healthcare dynamics.
- **Data Limitations:** It heavily relies on the availability and quality of data, which can be a challenge in public healthcare settings where data collection and standardization vary across regions.
- **Overfitting:** Time series models may be prone to overfitting, where the model becomes too strongly associated with the training data, leading to poor performance on unseen or future data an important consideration when analyzing dynamic patterns within the unpredictable public health systems in India.
- **Ethical Considerations:** Time series analysis in public health research raises ethical considerations related to data privacy, informed consent, and potential biases in data collection and It is imperative to develop statistical techniques with high forecasting accuracy and reliability for analyzing and predicting the prevalence and mortality time series of infectious diseases. [13]
- **Unpredictability of Rare Events:** Time series analysis may struggle to accurately predict rare events, such as novel disease outbreaks or large-scale pandemics, due to the lack of historical data for such events, the most recent example being the COVID 19 pandemic. [14]

2. Propensity Score Matching

Propensity score matching is a statistical technique that is commonly used in public health research to estimate the causal effects of interventions or policies on an outcome of interest, such as disease prevalence or mortality rates. This technique involves constructing a propensity score, which represents the probability of receiving the treatment or intervention based on a set of observed covariates. Using the propensity score, individuals or groups with similar propensity scores are matched, creating a comparison group that is comparable to the treatment group in terms of observed characteristics. This matching process helps to reduce selection bias and isolate the effects of the intervention or policy being studied.

***Strengths*:**

- **Accurate estimates:** Propensity score matching provides a method for estimating causal effects, allowing researchers to make more accurate comparisons between treatment and control groups.
- **Accounting for confounding variables:** Propensity score matching helps control for confounding variables, reducing the potential bias in estimating the effects of interventions.
- **Reducing selection bias:** By matching individuals or groups based on their propensity scores, propensity score matching helps to reduce selection bias, ensuring that the treatment and control groups are comparable in terms of observed characteristics.
- **Increased interpretability:** Propensity score matching allows for a clearer interpretation of the treatment effect by comparing similar individuals or groups and isolating the effects of the intervention or policy being studied.
- **Improved generalizability:** By matching individuals or groups based on their propensity scores, propensity score matching helps to create a comparison group that is more representative of the overall population, improving the generalizability of the findings.
- **Options for sensitivity analysis:** Provides flexibility in assessing the robustness of findings. This method allows researchers to evaluate different specifications of the propensity score model or matching algorithm, taking into account factors in current healthcare systems and patient populations.
- **Better utilization of observational data:** Allows researchers to utilize existing observational data and draw meaningful conclusions about causal effects without the need for randomized controlled trials.
- **Estimating intervention impact more accurately:** Propensity score matching enhances the accuracy of estimating the impact of interventions by minimizing selection bias and confounding variables [15]. This has helped in cases such as evaluating education programs and their impact on child health.

***Limitations*:**

- **Limited control of unobserved variables:** Despite its ability to control for observed confounding variables, propensity score matching cannot account for unobserved confounders.
- **Potential bias in the estimation of causal effects:** Although propensity score matching reduces selection bias and confounding, there is still a possibility of residual bias that cannot be fully eliminated through this method [16].
- **Reliance on the assumption of unconfoundedness:** Relies on the assumption of unconfoundedness, which states that all relevant confounders have been accounted for and balanced between the treatment and control groups [17].
- **Small sample size:** Requires a sufficiently large sample size to ensure reliable matches between treatment and control subjects.
- **Requires high quality data:** The successful implementation of propensity score matching relies on the availability of high-quality data, including accurate and complete information on the covariates used in the matching process. It also requires data with overlap between treatment and control groups in order to ensure valid comparisons.
- **Challenging interpretations:** There is significant complexity in chosing what to interpret, especially when choosing covariates and matching algorithms. This underscores the importance of carefully evaluating the impact estimates, particularly in public healthcare, where accurate assessments are vital for effective decision-making. In India, this holds particular significance due to the diverse population and intricate

healthcare environment.

3. Clustering Analysis

An alternative approach for analyzing and interpreting data is cluster analysis. This statistical method seeks to categorize data points into clusters according to their similarity or proximity. Cluster analysis has diverse applications, including data mining, image recognition, and social network analysis. In public healthcare, it is commonly employed to examine patterns in disease outbreaks and identify high-risk area clusters, as well as understand the contributing factors while identifying similar characteristic groups within a population.

Strengths:

- **Better handling of unobserved variables:** Unlike propensity score matching, cluster analysis has the potential to handle unobserved variables or confounders, making it valuable for analyzing population data where various factors may impact health outcomes.
- **Flexibility in data types:** Cluster analysis can handle various data types, including categorical, continuous, and binary variables.
- **Less reliance on assumptions:** Cluster analysis proves to be more adaptable and robust when dealing with complex datasets from the India that may contain confounding factors.
- **Identification of distinct groups:** Cluster analysis can identify distinct groups or clusters within a dataset, allowing for detailed exploration and characterization of different subpopulations; which can assist in targeted interventions at local levels.
- **Enhanced interpretability and improved resource allocation:** Cluster analysis provides clear and easily interpretable results, as data points are grouped together based on similarity or proximity, aiding decisionmaking processes related to resource allocation in India's diverse regions by identifying clusters of high-risk areas or populations that require targeted interventions or preventive measures, such as using maternal health data for pre natal care resource allocation.
- **Potential for hypothesis generation:** Cluster analysis can generate hypotheses for further investigation and research, as it can reveal patterns and relationships within the data that may not have been initially evident.
- **Improvement of decision-making process:** By identifying clusters and patterns within the data, cluster analysis can provide valuable insights that inform decision-making processes in public healthcare settings.
- **Identification of disparities:** Cluster analysis can uncover disparities or inequalities in healthcare access, utilization, and outcomes across different clusters, allowing for targeted interventions to address these disparities and promote health equity.

Limitations:

- **Over-reliance on algorithmic assumptions:** Cluster analysis relies heavily on the chosen algorithm and its underlying assumptions. If these assumptions are not aligned with the unique characteristics of healthcare data, it can lead to incorrect clustering results and misinterpretations in interventions. [18]
- **Difficulty in determining the appropriate number of clusters:** One challenge for healthcare interventions is determining the optimal number of clusters to target for specific initiatives. Various methods, such as demographic analysis or disease prevalence mapping, can be employed to determine the number of targeted clusters; however, this decision is often subjective and may not have a clear-cut answer.
- **Methodological considerations:** Cluster analysis requires careful consideration of methodological choices, such as selecting relevant metrics (e.g., incidence rates) and identifying suitable population segments for intervention targeting based on epidemiological factors like age groups or comorbidities.
- **Inadequate data quality:** Cluster analysis relies heavily on high-quality input data related to disease surveillance, patient records or environmental factors influencing public health outcomes – incomplete or inaccurate data can lead to unreliable insights for designing effective measures. [19]
- **Use of inappropriate variables:** The choice of variables used in cluster analysis is crucial, and if irrelevant or unrepresentative variables are included, it can lead to misleading cluster results. Inclusion of irrelevant indicators may result from oversight when evaluating certain diseases requiring prompt attention but lack representation across vulnerable populations disproportionately affected by particular ailments.
- **Insufficient understanding and application of visual analytics methods:** While cluster analysis can provide valuable insights, its potential is often limited by a lack of understanding, availability and development of visual analytics methods.
- **Not suitable for continuous or high dimension data:** While cluster analysis is useful for categorical or low-dimensional data, it may not be suitable for continuous or high-dimensional data.

C. Case Studies

1. Statistical Method: Risk factors for Maternal Mortality [20]

Statistical Methods Used:

- **Estimation of Maternal Mortality Ratio (MMR):** The standard method for MMR computation was used, taking into account the number of maternal deaths and live births reported in the Health Management Information System (HMIS).
- **Calibration Factor (Cf):** A calibration factor was computed to account for the underor over-reporting of maternal deaths in HMIS by states and districts of India. This factor was initially estimated for states as the ratio of MMR from the Sample Registration System (SRS) and HMIS.
- **Geographical Distribution and Spatial Clustering:** Geographic Information System (GIS) mapping was used to show the geographical distribution of MMR across the states and districts of India. Univariate local Moran's I and Local Indicator of Spatial Association (LISA) cluster and significance maps were employed to assess the extent of geographical clustering.
- **Ordinary Least Square (OLS) Regression Model:** An OLS log-linear regression model was used to understand the maternal health care, demographic, and socioeconomic correlates of MMR. Univariate

(unadjusted) regression estimates were modeled in the first stage, followed by six OLS regressions to avoid collinearity between the explanatory variables.

- **Robustness Checks:** The study included robustness checks to assess data reliability, including comparisons of estimates from different sources and the use of population weights to adjust for district-level unequal size in error margins.

These statistical methods were employed to analyze and understand the geographical variation and spatial clustering of maternal mortality in India.

Results provided by Statistical Techniques:

The statistical methods used in the study provided valuable insights into the geographical variation and spatial clustering of maternal mortality in India. The conclusions drawn from these methods include:

- **Geographical Variation:** The study revealed considerable geographical heterogeneity in MMR across Indian states, with spatial patterns indicating high MMR in North-eastern, Eastern, and Central regions and low MMR in the Southern and Western regions.
- **Spatial Clustering:** The use of GIS mapping and spatial autocorrelation analysis identified significant local clusters of high and low MMR levels, providing insights into the spatial distribution of maternal mortality.
- **Correlates of MMR:** The OLS regression model helped identify the maternal health care, demographic, and socioeconomic correlates of MMR, shedding light on factors influencing maternal mortality at the macro-level.
- **Data Reliability Assessment:** The robustness checks, including the calibration factor and population-weighted adjustments, contributed to assessing the reliability of the data and the accuracy of the MMR estimates.

Overall, the statistical methods facilitated a comprehensive understanding of the geographical distribution, spatial clustering, and correlates of maternal mortality in India, providing valuable evidence for policy-making and public health interventions.

2. **Mathematical Optimization: Time-to-supply based approach to cold chain network optimization and extension in Madhya Pradesh [21]**

Mathematical Techniques Used:

- **Supply Chain Optimization:** The study focuses on improving the immunization supply chain using mathematical optimization techniques. The primary goal is to make better decisions in selecting locations for Cold Chain Points (CCPs) and Sub-Health Centers (SHCs) for efficient vaccine distribution.
- **Binary Integer Programming Formulation:** The study formulates the optimization problem using Binary Integer Programming (BIP). The objective is to maximize the coverage of demand points, with decision variables Yj and Xij guiding the selection of candidate locations and pairs of demand points and candidate locations. Constraints ensure exclusive coverage, single-candidate associations, and adherence to travel time limits. Integrality constraints enforce binary values for decision variables.
- **Greedy Adding Algorithm:** It used a Greedy Adding algorithm, implemented in Microsoft Excel. This algorithm efficiently addresses NP-hard problems, especially the Maximal Coverage Location Problems (MCLP) examined in the study.
- **Branch Bound Algorithm:** This algorithm utilizes Excel Solver's Simplex LP function. The paper compared its performance with the Greedy Adding algorithm.

Results provided by Mathematical Optimization Techniques:

- **Impact Measurement:** The study aimed to gauge the impact of the network optimization and extension exercise by measuring metrics related to the reduction in the number of remote sites, average reduction in Time-toSupply (TTS), and reduction in Sub-Health Center (SHC) load resulting from the establishment of new Cold Chain Points (CCPs).
- **Algorithm Performance:** The study compared the results obtained from two competing algorithms, the Greedy Adding algorithm and the Branch Bound algorithm. It was found that the Greedy Adding algorithm provided near-optimal solutions for the problem set, making it the preferred choice due to its efficiency in providing solutions for certain types of NP-hard problems such as the Maximal Coverage Location Problems (MCLP) formulated in the study.
- **District-Wise Results:** The study presented district-wise results of the network optimization and extension, detailing the existing CCPs, candidate CCPs, number of SHCs, remote SHCs, realigned SHCs, and remote SHCs attached to new CCPs for each district.
- **Economic Analysis:** The study utilized an economic analysis to provide the basis for the minimum load constraint, considering costs related to capital expenditure, operating expenditure, and financing costs to calculate the breakeven point (BEP) for the AVDS porter compensation scenarios.
- **Conclusion:** This study introduces a formal approach to optimize the last tier of the immunization supply chain, emphasizing equity for beneficiaries. The key contribution is a practical methodology improving vaccine distribution efficiency, reducing Time-to-Supply, and enhancing coverage. Mathematical optimization techniques identified optimal locations for Cold Chain Points (CCPs) and Sub-Health Centers (SHCs), offering economic insights. This work has major implications for vaccine quality, session punctuality, and overall equity in the immunization supply chain.

III. HYBRID APPROACH

In the past, pure optimization and statistical approaches provided valuable insights but also had limitations. Optimization models often oversimplified real-world healthcare systems, leading to impractical solutions that overlooked crucial factors like uncertainty and variability. Statistical analysis excelled in identifying relationships and patterns in data but struggled to translate these insights into actionable recommendations. To overcome these limitations, a hybrid approach combining optimization and statistical analysis has emerged as a promising solution [22]. This integration has resulted in stronger solutions with the ability to manage unforeseen deviations in patient volumes.

Statistical methods began using optimization techniques to make their insights actionable by predicting patient demand for different services and determining optimal staffing levels for each department. Meanwhile, optimization models started leveraging Bayesian inference methods for parameter estimation and uncertainty quantification [23]. These developments led to increasingly sophisticated hybrid approaches with unprecedented accuracy and flexibility. [24] The rise of machine learning has further accelerated this trend, seamlessly integrating ML models into the existing frameworks of optimization and statistics. [25]

A. Historical Development

The hybrid method is effective as it deals with the inherent constraints of each individual approach:

- **Constraints of Optimization:** This often depends on simplifying assumptions about reality, which may overlook the finer points of healthcare data revealed by statistical analysis.
- **Constraints of Statistics:** While statistics provide valuable insights into past trends and relationships, they may not offer practical solutions for optimizing future resource allocation.

A hybrid approach combines the strengths of both optimization and statistics, making it well-suited for complex decision-making challenges in healthcare where resources are limited and there is uncertainty. The concept of merging optimization and statistics originated from operations research, gaining prominence in the mid-20th century. However, its importance has significantly increased in recent years due to several factors:

- **Enhanced Computational Power:** The availability of powerful computers enables the handling of intricate calculations required by hybrid models.
- **Utilization of ”Big Data” in Healthcare:** The increasing accessibility of electronic health records provides substantial datasets that fuel robust statistical analysis and fine-tune optimization models.
- **Progression in Machine Learning:** The development of machine learning algorithms and techniques has transformed healthcare analytics by enabling analysis of large volumes of data, extraction of meaningful patterns and relationships within healthcare information, as well as generating accurate predictions.

The hybrid method is effective as it deals with the inherent constraints of each individual approach:

- **Constraints of Optimization:** This often depends on simplifying assumptions about reality, which may overlook the finer points of healthcare data revealed by statistical analysis.
- **Constraints of Statistics:** While statistics provide valuable insights into past trends and relationships, they may not offer practical solutions for optimizing resource allocation.

A hybrid approach combines the strengths of both optimization and statistics, making it well-suited for complex decision-making challenges in healthcare where resources are limited and there is uncertainty. The concept of merging optimization and statistics originated from operations research, gaining prominence in the mid-20th century. However, its importance has significantly increased in recent years due to several factors:

- **Enhanced Computational Power:** The availability of powerful computers enables the handling of intricate calculations required by hybrid models.
- **Utilization of ”Big Data” in Healthcare:** The increasing accessibility of electronic health records provides substantial datasets that fuel robust statistical analysis and fine-tune optimization models.
- **Progression in Machine Learning:** The development of machine learning algorithms and techniques has transformed healthcare analytics by enabling analysis of large volumes of data, extraction of meaningful patterns and relationships within healthcare information, as well as generating accurate predictions.

B. Case Study: Bed Allocation Using Markov Chain Modeling [26]

Problem: Inefficient hospital bed allocation, leading to increased patient wait times, suboptimal resource utilization, and potential negative impacts on patient outcomes.

Methodology

- **States:** Define states to represent different bed occupancy levels (e.g., empty, low acuity patient, high acuity patient).
- **Transition Probabilities:** Calculate probabilities of moving from one bed occupancy state to another based on historical data, patient characteristics, etc.
- **Initial State Distribution:** Determine the probabilities of starting in various bed occupancy states, likely based on current occupancy.
- **Steady-State Distribution:** Calculate the long-term, equilibrium distribution of bed occupancy.
- **Analysis and Decision Making:** Use the model’s results to examine occupancy rates, wait times, etc. Simulate scenarios by altering probabilities and initial state distributions to assist in making decisions about bed allocation and resources.

Mathematics Involved

- **Probability theory:** The very nature of Markov Chain models heavily relies on calculating and understanding probabilities.
- **Matrix operations:** The transition probabilities are represented in a matrix, and calculations like finding the steady-state distribution likely involve matrix operations.
- **Data analysis:** Statistical analysis of historical data to estimate the transition probabilities.

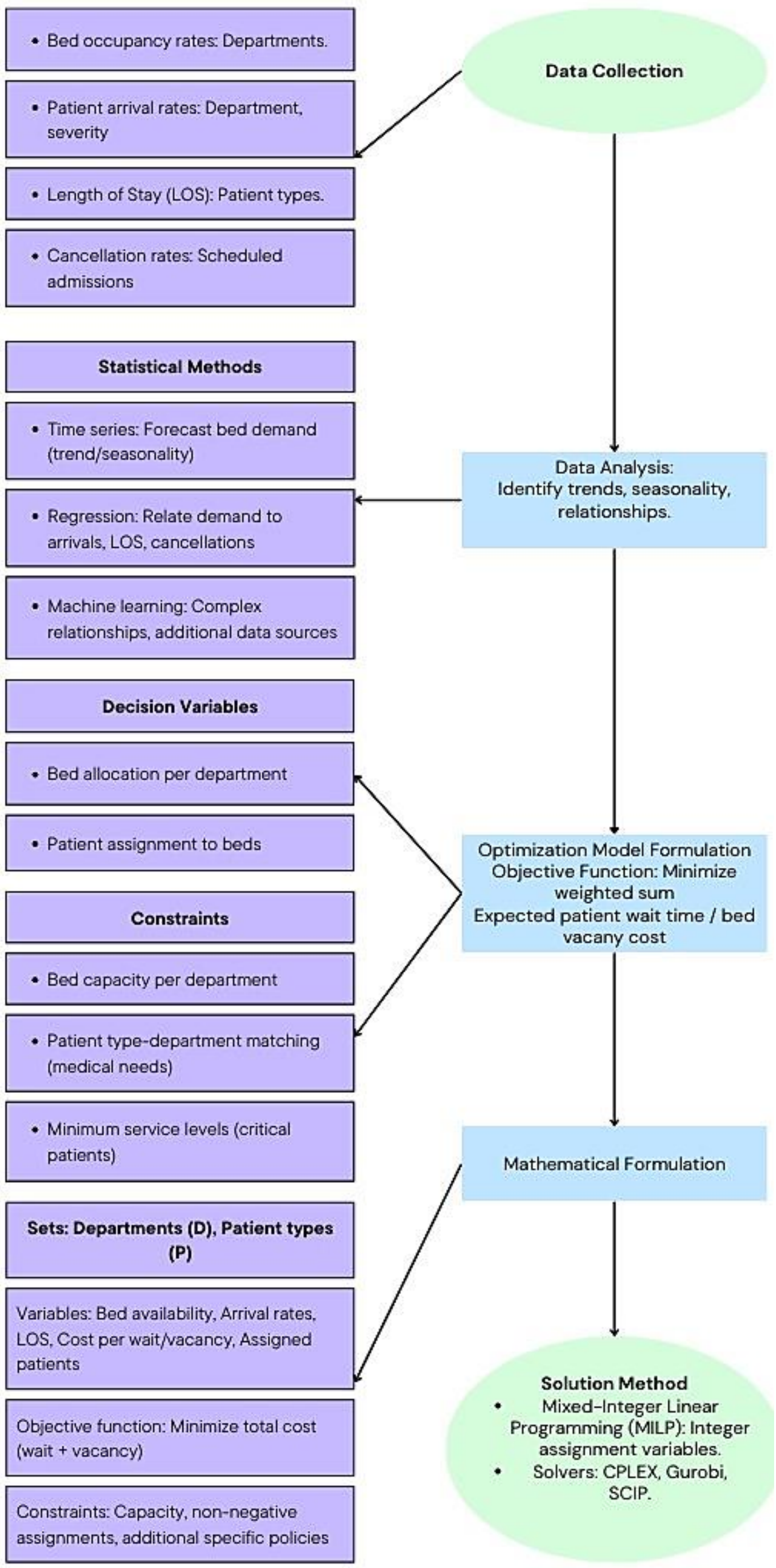

Figure 1: Bed Allocation using Markov Chain Modeling: Methodology Pervasive wireless grid

Analytics View (Results Benefits)

- **Probability-based Analysis:** Understands system dynamics using probabilities rather than solely on deterministic factors.
- **Performance Evaluation:** Enables the calculation of performance metrics like bed occupancy rates, wait times, and resource usage.
- **Scenario Analysis:** Allows for testing 'what-if' scenarios to optimize bed allocation strategies.
- **Resource Planning:** Assists in predicting future bed

demand aiding in capacity planning.
- **Decision Support:** Provides data-driven recommendations for bed allocation and policymaking.

Benefits to Providers

- **Improved bed allocation:** Enhances bed assignment based on data like length of stay and patient needs.
- **Cost savings:** Reduces empty beds and overcapacity, minimizing costs.
- **Improved patient experience:** Reduces wait times and ensures appropriate care matching.
- **Better resource utilization:** Helps predict bed needs and adjust staffing accordingly.
- **Enhanced decision-making:** Provides data for informed resource allocation.
- **Accurate forecasting:** Assists in long-term planning decisions.
- **Improved communication:** Offers a clear view of bed occupancy for smoother collaboration.
- **Reduced errors:** Provides a systematic decision-making framework to minimize mistakes.

How These Results Point to the Need for a Hybrid Approach

The results obtained from the analysis highlight the need for a hybrid approach in bed allocation.

- **Utilizing Data for Efficiency:** Transition probabilities derived from the Markov Chain assist in creating effective bed allocation strategies based on real-time data and demand patterns, enhancing system dynamics and performance metrics assessment.
- **Scenario Evaluation Enhanced:** While the Markov Chain assesses scenarios, it does not inherently optimize them. Integrating a hybrid model provides actionable insights to achieve specific objectives (such as minimizing wait times, balancing occupancy) under practical conditions instead of hypothetical assessments.
- **Addressing Real-World Constraints:** The Markov model may not fully capture factors like staffing levels or unit separation rules; optimization allows direct consideration of such real-world complexities.
- **Moving from Analysis to Implementation:** By providing probabilistic system views, the Markov Chain sets the stage for structured decision-making models needed to determine optimal daily or weekly bed assignments—a strength of optimization techniques.
- **Incorporating Practical Limitations:** An optimization model can encompass hospital-specific constraints within probabilistic boundaries defined by the Markov Chain analysis results.

IV. CHOOSING THE RIGHT TOOL

The ideal statistical and optimization techniques depend on the specific problem you're addressing. [27]

Here are some factors to consider:

- **Public Health Challenge**: Time series analysis may be suitable for cyclic illnesses such as dengue, while network optimization is appropriate for interventions like vaccination campaigns that are spread across different geographical locations. [28]
- **Problem Characteristics:** Factors such as whether the objective function (the aspect being optimized) follows a linear or non-linear pattern, if there are any constraints on whole numbers (e.g., number of beds), the level of uncertainty present in the data and how it impacts the choice of statistical analysis, various tools can be decided.
- **Data Availability:** The availability and quality of data play a crucial role in selecting the appropriate statistical and optimization techniques.
- **Computational Resources:** Certain optimization methods require considerable computational capacity; therefore, consider available resources when selecting a model.
- **Interpretation:** Healthcare decision-makers with limited quantitative backgrounds need to easily comprehend the results produced by the model. [29]
- **Policy Context and Objectives:** Adhering to existing policies and social realities requires careful translation of model outcomes to ensure alignment with policy context and objectives.

V. CONCLUSION

Public health interventions in India require advanced solutions. Both optimization and statistical techniques have unique strengths in addressing resource allocation, disease modeling, and operational efficiency challenges. However, they also have limitations when used alone. The hybrid approach, combining statistical insights with optimization models, has demonstrated transformative power by generating tailored solutions for the Indian context and aiding decision-making tools that consider budget constraints and policy goals.

The historical development of these techniques shows increasing sophistication driven by advances in computational power and granular health data availability. The future focus should be on real-world validation through case studies, improving data collection quality for more accurate modeling efforts, developing adaptive frameworks to accommodate dynamic changes in public health priorities and resource availability, as well as ensuring interpretability of hybrid models for policymakers and health administrators.

Continuous research collaboration between academia, policymakers, and public health officials is essential to ensure sustainable public health improvements in India's evolving landscape.

REFERENCES

[1] C. D. Flagle, "Operational research in the health services," Annals of the New York Academy of Sciences, vol. 107, no. 2, pp. 748–759, May 1963. DOI: 10.1111/j.1749-6632.1963.tb13318.x. [Online]. Available: https://nyaspubs.onlinelibrary.wiley.com/doi/10.1111/j.17496632.1963.tb13318.x.

[2] A. A. Stinnett and A. D. Paltiel, "Mathematical programming for the efficient allocation of health care resources," Journal of Health Economics, vol. 15, no. 5, pp. 641–653, Oct. 1996. DOI: 10.1016/s01676296(96) 00493-6. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0167629696004936.

[3] E. S. Pearson, "An appreciation of some aspects of his life and work," Biometrika, vol. 28, no. 3-4, pp. 193–257, Dec. 1936. DOI: 10 . 1093 / biomet/ 28 . 3 4 . 193. [Online]. Available: https:// www. jstor. org/ stable/2332005?origin=crossref.

[4] F. Brauer, “Mathematical epidemiology: Past, present, and future,” Infectious Disease Modelling, vol. 2, no. 2, pp. 113–127, May 2017. DOI: 10.1016/j.idm.2017.02.001. [Online]. Available: https://www.ncbi. nlm.nih.gov/pmc/articles/PMC6001967/.

[5] M. Holte and C. Mannino, “The implementor/adversary algorithm for the cyclic and robust scheduling problem in health-care,” May 2013. [Online]. Available: https://scite.ai/reports/10.1016/j.ejor.2012.10.029.

[6] H. Zhang, H. Wang, and J. Wei, “Perceptive movement of susceptible individuals with memory,” Mar. 2023. [Online]. Available: https://scite. ai/reports/10.1007/s00285-023-01904-w.

[7] B. Pourbohloul, L. A. Meyers, D. M. Skowronski, M. Krajden, D. M. Patrick, and R. C. Brunham, “Modeling control strategies of respiratory pathogens,” Aug. 2005. [Online]. Available: https://scite.ai/reports/ 10.3201/eid1108.040449.

[8] K. M. Islam and O. M. Araz, “Evaluating the effectiveness of targeted public health control strategies for chlamydia transmission in omaha, nebraska: A mathematical modeling approach,” Jan. 2014. [Online]. Available: https://scite.ai/reports/10.4236/aid.2014.43021.

[9] M. O. Afolabi, O. Wariri, Y. Saidu, et al., “Tracking the uptake and trajectory of covid-19 vaccination coverage in 15 west african countries: An interim analysis,” Dec. 2021. DOI: 10 . 1136 / bmjgh2021 007518. [Online]. Available: https://scite.ai/reports/10.1136/bmjgh-2021007518.

[10] R. Shadi, A. Fakharian, and H. Khaloozadeh, “Modeling and analysis of covid-19 spread: The impacts of nonpharmaceutical protocols,” Sep. 2022. DOI: 10.1155/2022/7706447. [Online]. Available: https://scite.ai/ reports/10.1155/2022/7706447.

[11] Y. Wang, C. Xu, S. Yao, and Z. Yi, “Forecasting the epidemiological trends of covid-19 prevalence and mortality using the advanced -sutte indicator,” Jan. 2020. [Online]. Available: https://scite.ai/reports/10.1017/ s095026882000237x.

[12] O. S. Medina-G´omez and I. S. Medina-Reyes, “Mortalidad por diabetes tipo 2 y la implementaci´on del programa prevenimss: Un estudio de series de tiempo en m´exico, 1998-2015,” May 2018. [Online]. Available: https://scite.ai/reports/10.1590/0102-311x00103117.

[13] Y. Wang, C. Xu, S. Yao, et al., “Estimating the covid-19 prevalence and mortality using a novel data-driven hybrid model based on ensemble empirical mode decomposition,” Nov. 2021. DOI: 10 . 1038 / s41598 021 009486. [Online]. Available: https:// scite. ai/ reports/ 10. 1038/ s41598-021-00948-6.

[14] B. Zhao, Y. Li, B. Wang, et al., “Mathematical modeling and epidemic prediction of covid-19 and its significance to epidemic prevention and control measures,” Jul. 2021. DOI: 10.31579/2690-1897/021. [Online]. Available: https://scite.ai/reports/10.31579/2690-1897/021.

[15] A. Nasif and S. Masroor, “Cardiac surgery and healthcare quality: Is the right question being asked?,” Jun. 2021. [Online]. Available: https:// scite.ai/reports/10.1111/jocs.15709.

[16] J. Y. Lee, R. R. Diaz, H. S. Yu, J. Chung, W. S. Ham, and Y. D. Choi, “Lymphocele after extraperitoneal robot-assisted radical prostatectomy: A propensity score-matching study,” Mar. 2013. [Online]. Available: https://scite.ai/reports/10.1111/iju.12144.

[17] Z. Touma, K. H. Costenbader, B. Hoskin, et al., “Patient-reported outcomes and healthcare resource utilization in systemic lupus erythematosus: Impact of disease activity,” Aug. 2022. [Online]. Available: https://scite.ai/reports/10.21203/rs.3.rs-1669786/v2.

[18] P. Ahmad, S. Qamar, and S. Q. A. Rizvi, “Techniques of data mining in healthcare: A review,” Jun. 2015. DOI: 10.5120/213074126. [Online]. Available: https://scite.ai/reports/10.5120/21307-4126.

[19] B. Ehsani-Moghaddam, K. Martin, and J. Queenan, “Data quality in healthcare: A report of practical experience with the canadian primary care sentinel surveillance network data,” Health Information Management Journal, vol. 50, no. 1-2, pp. 88–92, Dec. 2019. DOI: 10.1177/1833358319887743. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/31805788/.

[20] S. Goli, P. Puri, P. S. Salve, S. Pallikadavath, and K. S. James, “Estimates and correlates of district-level maternal mortality ratio in india,” PLOS global public health, vol. 2, no. 7, e0000441–e0000441, Jul. 2022. [Online]. Available: https://doi.org/10.1371/journal.pgph.0000441.

[21] V. Srivastava, M. Ratna, A. Ray, et al., “Strengthening the immunization supply chain: A time-to-supply based approach to cold chain network optimization extension in madhya pradesh,” Vaccine, vol. 39, no. 45, pp. 6660–6670, Oct. 2021. DOI: 10 . 1016 / j . vaccine . 2021 . 09 . 062. [Online]. Available: https://doi.org/10.1016/j.vaccine.2021.09. 062.

[22] I. Makarova, K. Shubenkova, and A. Pashkevich, Logistical costs minimization for delivery of shot lots by using logistical information systems, Jan. 2017. [Online]. Available: https://doi.org/10.1016/j.proeng. 2017.01.059.

[23] H. Gupta, S. Kumar, D. Yadav, et al., “Data analytics and mathematical modeling for simulating the dynamics of covid-19 epidemic—a case study of india,” Electronics, vol. 10, no. 2, pp. 127–127, Jan. 2021. DOI: 10 . 3390/electronics10020127. [Online]. Available: https://doi.org/10. 3390/electronics10020127.

[24] G. R. Babu, D. Ray, R. Bhaduri, et al., “Covid-19 pandemic in india: Through the lens of modeling,” Global health, science and practice, vol. 9, no. 2, pp. 220–228, Jun. 2021. DOI: 10.9745/ghsp-d-21-00233. [Online]. Available: https://www.ghspjournal.org/content/ghsp/9/2/220. full.pdf.

[25] S. Priyan, “Operations research in healthcare: A review,” JOJ Pub Health, Mar. 2017. DOI: 10.19080/JOJPH.2017.01.555561. [Online]. Available: https://juniperpublishers.com/jojph/JOJPH.MS.ID.555561.php.

[26] B. Ramdurai, “Use of markov chain model for in-patient bed assignment,” Algorithms Research, vol. 6, no. 1, pp. 1–8, DOI: doi:10.5923/j. algorithms.20230601.01.

[27] G. T. Jun, Z. S. Morris, T. Eldabi, et al., “Development of modelling method selection tool for health services management: From problem structuring methods to modelling and simulation methods,” BMC Health Services Research, vol. 11, no. 1, May 2011. DOI: 10.1186/1472-6963-11-108. [Online]. Available: https://bmchealthservres.biomedcentral.com/articles/10.1186/ 1472-6963-11-108.

[28] R. Zachariah, A. Harries, N. Ishikawa, et al., “Operational research in lowincome countries: What, why, and how?” The Lancet Infectious Diseases, vol. 9, no. 11, pp. 711–717, Nov. 2009. DOI: 10.1016/s14733099(09) 702294. [Online]. Available: http://www.thelancet.com/retrieve/pii/S1473309909702294.

[29] R. Lilford, P. J. Chilton, K. Hemming, A. Girling, C. A. Taylor, and P. Barach, “Evaluating policy and service interventions: Framework to guide selection and interpretation of study end points,” The BMJ, vol. 341, no. aug27 1, pp. c4413–c4413, Aug. 2010. DOI: 10 . 1136 / bmj . c4413. [Online]. Available: https://doi.org/10.1136/bmj.c4413.